NONLINEAR SIGNAL PROCESSING FOR EFFICIENT PHYSICAL LAYER DESIGN IN 5G BEYOND

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MURAT BABEK SALMAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JULY 2023

Approval of the thesis:

NONLINEAR SIGNAL PROCESSING FOR EFFICIENT PHYSICAL LAYER DESIGN IN 5G BEYOND

submitted by **MURAT BABEK SALMAN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İlkay Ulusoy
Head of Department, **Electrical and Electronics Engineering** _____

Assist. Prof. Dr. Gökhan Muzaffer Güvensen
Supervisor, **Electrical and Electronics Engineering, METU** _____

Prof. Dr. Tolga Çiloğlu
Co-supervisor, **Electrical and Electronics Engineering, METU** _____

Examining Committee Members:

Prof. Dr. Orhan Arıkan
Electrical and Electronics Engineering, Bilkent Univ. _____

Assist. Prof. Dr. Gökhan Muzaffer Güvensen
Electrical and Electronics Engineering, METU _____

Prof. Dr. Ali Özgür Yılmaz
Electrical and Electronics Engineering, METU _____

Assist. Prof. Dr. Serkan Sarıtaş
Electrical and Electronics Engineering, METU _____

Assist. Prof. Dr. Özlem Tuğfe Demir
Electrical and Electronics Engineering, TOBB ETÜ _____

Date:12.07.2023

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Murat Babek Salman

Signature :

ABSTRACT

NONLINEAR SIGNAL PROCESSING FOR EFFICIENT PHYSICAL LAYER DESIGN IN 5G BEYOND

Salman, Murat Babek
Ph.D., Department of Electrical and Electronics Engineering
Supervisor: Assist. Prof. Dr. Gökhan Muzaffer Güvensen
Co-Supervisor: Prof. Dr. Tolga Çiloğlu

July 2023, 164 pages

Nonlinear power amplifier (PA) distortion is a significant hardware impairment that imposes limitations on communication systems, hindering their performance. This thesis presents a comprehensive analysis framework for hardware impairments and develops nonlinear signal processing techniques to mitigate the detrimental effects of nonlinear distortion on communication system performance. The evaluation is based on several criteria, including bit-error-rate (BER), achievable information rate (AIR), and out-of-band radiation.

The thesis's contribution can be divided into two independent parts. The first part focuses on uplink communication, where the user terminal (UT) is equipped with a nonlinear PA. For uplink communication, the investigation encompasses the effects of frequency-selective communication channels and symbol-rate signal processing. A critical phenomenon named as 'distortion amplification', which significantly degrades detection performance, is discovered. To overcome distortion amplification, the use of a bank of nonlinear processing units for different fractional delays is proposed for both single-input-single-output (SISO) and multiple-input-multiple-output (MIMO)

systems, leading to substantial improvements in detection performance.

The second part of thesis examines downlink communication scenarios, where base stations (BS) are equipped with nonlinear PAs. Firstly, a comprehensive analysis of received distortion power is conducted to investigate the effects of multi-user diversity and frequency selectivity. Then, based on the obtained distortion power, a power allocation scheme is proposed to enhance spectral efficiency for linear precoders. Lastly, a low-complexity digital pre-distortion (DPD) design is developed for two-stage hybrid beamforming systems with a fully connected analog beamforming structure. The proposed DPD structure achieves sufficient linearization performance with reduced complexity compared to state-of-the-art methods.

Keywords: Power amplifiers, Nonlinear Distortion, Post-distortion, Digital pre-distortion, Distortion amplification, Distortion Analysis

ÖZ

5G VE ÖTESİ SİSTEMLERDE FİZİKSEL KATMAN İÇİN ETKİN DOĞRUSAL OLMAYAN SİNYAL İŞLEME

Salman, Murat Babek
Doktora, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Dr. Öğr. Üyesi. Gökhan Muzaffer Güvensen
Ortak Tez Yöneticisi: Prof. Dr. Tolga Çiloğlu

Temmuz 2023 , 164 sayfa

Nonlineer güç amplifikatörü (PA) distorsiyonu, iletişim sistemlerinin yeterli performansı elde etmesini önemli ölçüde sınırlayan önemli bir donanım bozulmasıdır. Bu tez, donanım bozulmaları için kapsamlı bir analiz çerçevesi sunmakta ve iletişim sistemi performansını olumsuz yönde etkileyen nonlineer distorsiyonun etkilerini azaltmak için nonlineer sinyal işleme teknikleri geliştirmektedir. Değerlendirme, hata oranı (BER), elde edilebilir bilgi oranı (AIR) ve bant dışı radyasyon gibi çeşitli kriterlere dayanmaktadır.

Tezin katkısı iki bağımsız bölüme ayrılabilir. İlk bölüm, kullanıcı terminalinin (UT) nonlineer PA ile donatıldığı yukarı bağlantı iletişimine odaklanmaktadır. Yukarı bağlantı iletişimi için, frekans seçici iletişim kanallarının ve sembol örnekleme işlem şemalarının etkileri incelenmektedir. Sembol tespit performansını önemli ölçüde bozan 'distorsiyon amplifikasyonu' olarak adlandırılan önemli bir olgu keşfedilmektedir. Distorsiyon amplifikasyonunu aşmak için, tek girişli-tek çıkışlı (SISO) ve çoklu girişli-çoklu çıkışlı (MIMO) sistemler için nonlineer işlem birimi bankasının kulla-

nımı önerilmekte ve tespit performansında önemli iyileştirmeler sağlanmaktadır.

Tezin ikinci bölümü, baz istasyonlarının (BS) nonlineer PA'larla donatıldığı aşağı bağlantı iletişimi senaryolarını incelemektedir. İlk olarak, çok kullanıcılı çeşitlilik ve frekans seçiciliğinin etkilerini araştırmak için kapsamlı bir alındı distorsiyon gücü analizi gerçekleştirilir. Ayrıca, elde edilen distorsiyon gücü ifadesine dayanarak bir güç tahsis şeması önerilmektedir. Son olarak, tam bağlantılı analog hüzme oluşturma yapısına sahip iki aşamalı hibrit MIMO sistemleri için düşük karmaşıklıkta bir sayısal ön bozma (DPD) tasarımı geliştirilmiştir. Önerilen DPD yapısı, mevcut yöntemlere göre daha düşük karmaşıklıkla yeterli doğrusallaştırma başarımı sağlamaktadır.

Anahtar Kelimeler: Güç yükselticiler, Doğrusal olmayan bozulma, sonradan bozma, Sayısal ön bozma, Bozulma güçlendirme, Bozulma analizi

To my family

ACKNOWLEDGMENTS

First of all, I would like to thank and express my gratitude to my supervisor, Assist. Prof. Dr. Gökhan Muzaffer Güvensen. He has not only been a great advisor and mentor, he is always there for me whenever I needed. I feel very lucky to work with a supervisor who has such a friendly personality. He also spared his own time whenever I struggled, and I have learned so much during technical discussions.

I feel extremely privileged to have Prof. Dr. Tolga Çiloğlu as a co-supervisor. His calmness and way of approaching problems inspired me, whenever, I was rushing for something, he has slow me down to deeply investigate the problem.

I would like to thank my thesis monitoring committee members Prof. Dr. Orhan Arıkan and Prof. Dr Ali Özgür Yılmaz for their valuable comments and suggestions. I would also like to thank my PhD defense jury members, Assist. Prof. Dr. Serkan Sarıtaş and Assist. Prof. Dr. Özlem Tuğfe Demir for their comments and suggestions on the thesis.

I would like to thank the Scientific and Technological Research Council of Turkey (TÜBİTAK) for the financial support TÜBİTAK 2214-A scholarship they gave me during my visit to Sweden.

I would like to acknowledge ASELSAN Inc. for providing both technical knowledge and financial support. Especially, I would like to thank Soner Yeşil for introducing the concept of full duplex communications. The core contribution of this thesis came to my mind during one of our weekly meetings. Also, I would like to thank Dr. Ali Bulut Üçüncü for re-directing me to the DPD design problem, which I had lost hope in this topic previously. Thanks to the research project under his supervision, one of the important parts of the thesis is completed.

I would like to present my gratitude to Prof. Dr. Emil Björnson from KTH Royal Institute of Technology, Sweden for giving me the opportunity to visit his research

group. I would like to also thank all CoS division for the friendly environment during my visit.

I would like to also thank my roommates from Room-214, Semih Yağlı, Mustafa Çökmezoğlu, and Alperen Gövdeli, for the nicest/cleanest room in METU. Alperen Gövdeli deserves special credit because even if he drives me crazy sometimes, life is beautiful to him. I would also like to thank my dearest friends Nehir Kavruk, Alaaddin Sak, Ziynet Kavruk, Mehmet Kavruk, Oğuz Çakmak for their enjoyable friendship. I also would like to thank my officemates at METU, Anıl Kurt, Uğur Berkay Saraç, Safa Çelik, Utku Çelebi, and Esen Özbay. Lastly, I would like to thank Hasan Aykut Şatana and Altuğ Kaya for their during our full duplex project.

I would like to thank my family for the support and encouragement they provided me for my entire life. Most importantly, I would like to thank my beautiful wife, Elif, for always supporting and bearing with me. I owe all my accomplishments to her.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

AIR	Achievable information rate
AWGN	Additive white Gaussian noise
BER	Bit-error-rate
BS	Base station
CDF	Cumulative distribution function
CSI	Channel state information
DA-SSD	Distortion aware symbol by symbol detector
DPD	Digital pre-distortion
DFT	Discrete Fourier Transform
FDE	Frequency domain equalization
FEC	Forward error correction
FT	Fast time
GMI	Generalized mutual information
GPR	Gaussian process regression
IDFT	Inverse Discrete Fourier Transform
IB-DFE	Iterative block decision feedback equalizer
ISI	Inter-symbol interference
IUI	Inter-user interference
LDPC	Low-density parity-check code
MIMO	Multiple-input-multiple-output
MM	Modified metric
MMSE	Minimum mean squared error
nIC	Nonlinear interference cancellation
NN	Neural network

NP	Nonlinear post-distortion
NPL	Nonlinear parameter learning
OFDM	Orthogonal Frequency-Division Multiplexing
OOB	Out-of-band
OBO	Output backoff
PA	Power amplifier
PU-IB-DFE	Per user IB-DFE
QAM	Quadrature amplitude modulation
SC	Single carrier
SE	Spectral efficiency
SDNR	Signal-to-distortion-plus-noise ratio
SNR	Signal-to-noise ratio
SISO	Single-input-single-output
ST	Slow time
UE	User equipment
UT	User terminal
VS	Volterra series

CHAPTER 1

INTRODUCTION

1.1 Literature Review and Motivation

The advent of communication technologies has opened up new frontiers for communication systems, with the potential to revolutionize the way we interact with each other and the world around us. Likewise, the emergence of new tools for human interaction has led to an increase in the throughput demand placed on wireless networks [4]. Meeting the growing demand for higher throughput per user, as well as the densification of communication networks, presents significant challenges in modern communication systems. To satisfy the required data rates, communication system design has evolved over time and continues to do so to meet future needs [5]. Early communication systems were primarily based on wired telephone networks, which transmitted voice signals [6]. The introduction of the first generation (1G) wireless networks in the 1980s enabled cellular communication, which relied solely on analog communication [7]. The subsequent evolution of wireless networks resulted in the emergence of fully digital communication schemes in second generation (2G) networks [7]. The primary objective of both 1G and 2G cellular systems was to transmit audio signals. However, the integration of the internet into everyday life rendered these networks inadequate, prompting the introduction of third generation (3G) systems. 3G networks represented a significant milestone in the evolution of communication systems by enabling wireless data transmission [7]. Achievable data rates have been improved further by forth generation (4G) systems by exploiting spatial diversity with the help of antenna arrays [8]. Currently, studies mainly focus on developing systems in the context of fifth-generation (5G) systems to fully benefit from smartphones and other devices connected to the internet [9].

In order to provide sufficient user experience, more sophisticated communication system designs were developed for each successive generation of wireless networks, achieving higher spectral efficiencies. For instance, multiple-input multiple-output (MIMO) systems have emerged as the most intriguing communication technology in recent times, attracting significant attention from researchers and industry experts. The growing interest in MIMO systems stems from their ability to enhance data rates, increase spectral efficiency, improve link reliability, and mitigate the effects of multipath fading. Various studies have demonstrated the substantial performance gains achieved through MIMO techniques in different wireless communication scenarios. In [10] a comprehensive overview of MIMO systems, including their principles, applications, and performance metrics, is provided.

The advancements of new generations in the field of wireless communication have brought about a host of fresh challenges. These include inter-cell interference, the need to utilize higher frequencies, growing demand for energy efficiency, and hardware imperfections [10–13]. Already, significant studies have been carried out, and countless complicated signal processing techniques have been developed to multiply the achievable spectral efficiency by mitigating these issues [14]. Nevertheless, there exist unresolved and even unidentified problems that set limitations impeding the attainment of optimal performance. One of the contributions of this thesis is to identify a problem that is neither addressed nor observed in the current literature.

The desire to optimize spectral efficiency, energy efficiency, and cost efficiency at once brings significant restrictions on hardware quality. Consequently, hardware impairments set some important obstacles in achieving the desired performance of communication systems since low-cost components are preferred to reduce the overall system cost while preserving the power efficiency [15]. Power amplifiers (PAs) are stated to be one of the key components due to their dynamic behavior across front-end modules. PAs exhibit nonlinear characteristics, because of inherent heat dissipation, near the saturation region, where they operate at maximum efficiency. This also means that a greater portion of the input power is converted into useful output power, resulting in reduced power consumption and longer battery life in portable devices. The improved power efficiency is particularly valuable in wireless communication systems where power consumption is a critical concern. Therefore, either

user equipment (UE) or base stations, which are equipped with low-cost nonlinear PA, transmit nonlinearly distorted signals since low power efficiency brings a serious burden to portable wireless equipment [16]. Furthermore, the goal of achieving high data rates with limited spectral resources makes higher-order quadrature amplitude modulation (QAM) constellations attractive for future-generation wireless networks [17–19]. However, PA nonlinearities on the transmitted signal degrade the received signal quality for single carrier (SC) transmission with higher modulation orders or orthogonal frequency division multiplexing (OFDM) modulation, where a high peak-to-average power ratio is a major issue.

Impacts of hardware impairments are even more evident for massive MIMO systems, where the base station is equipped with many antennas. Massive MIMO systems have attracted significant attention in the current literature thanks to their promise of multiplying the spectral efficiency greatly [20]. Multi-user (MU) massive MIMO systems enable to serve more than one user simultaneously by using spatial diversity so that different users can share the same time/frequency resources [21]. Besides, the spatial diversity of the users influences the radiation of the distortion signal, and spatial radiation of the distortion was studied in [22, 23]. These studies numerically demonstrate that the nonlinear distortion tends to decorrelate as the number of users increases. The majority of the current literature uses uncorrelated distortion assumption and analyzes/designs massive MIMO systems without taking distortion correlation into account. However, recent studies and this thesis have shown that neither the distortion correlation nor its effects can be ignored.

Hardware imperfections also have very dominant effects on full-duplex communication [24]. Full duplex systems aim to double the spectral efficiency by enabling simultaneous transmission and reception at the same time and frequency resources. However, strong self-interference (SI) poses a significant challenge to achieving the ultimate goal [25]. Furthermore, the SI signal is $\sim 100\ dB$ stronger than the desired signal-of-interest (SoI), which makes SI cancellation inevitable [26]. In addition, distortion due to PA nonlinearities leads to an additional challenge since the nonlinear distortion signal is much stronger than the SoI. Therefore, accurate estimation of the nonlinear distortion has vital importance [27]. Consequently, sophisticated signal processing tools are needed to estimate and cancel the nonlinear distortion part of the

SI signal.

Two major signal processing approaches are considered for mitigating nonlinear distortion due to PAs. The first approach aims to linearize the PA output signal. Among the PA linearization methods, the digital pre-distortion (DPD) method has attracted most of the attention in the research community due to its low complexity and reconfigurability compared to its analog counterparts [28–30]. DPD algorithms apply a nonlinear pre-processing, and it nonlinearly modifies the input signal such that the joint effects of DPD unit and PA yields the desired linear signal. Implementation of DPD not only mitigates the but also reduces the power leakage to adjacent channels, which is also referred out-of-band (OOB) radiation due to spectral re-growth. Although the DPD can linearize a single PA effectively, its implementation in massive MIMO systems requires a different perspective since per-antenna DPD is costly for fully digital beamforming systems and not practical for analog-digital hybrid beamforming structures.

However, the success of DPD algorithms is not always guaranteed, especially for PAs operated in the saturation region, in which the nonlinear function is not invertible. In addition, DPD methods may require the implementation of complex signal-processing tools such as deep neural networks [29, 30]. However, implementing such complex methods can be cumbersome for the user equipment [16]. Consequently, the second class of signal processing scheme, namely nonlinear post-distortion or nonlinear equalization, is developed to mitigate the effects of nonlinear distortion at the receiver side to improve the data detection accuracy [31, 32]. The nonlinear post-distortion framework treats PA characteristics as a nonlinear channel. By elevating this viewpoint, nonlinear equalizers are proposed in the aforementioned studies.

1.2 Summary of Contributions

This thesis presents a comprehensive signal processing framework aimed at addressing the challenges posed by PA nonlinearities in communication systems. The study encompasses a broad perspective, examining the effects of nonlinear distortion and proposing effective mitigation strategies for uplink communication scenarios where

non-ideal PAs are integrated into the user equipment (UE). Additionally, it investigates the impact of non-ideal PAs deployed in the base station (BS) on downlink communication.

One of the most important shortcomings of the current literature is the treatment of nonlinear distortion in the downsampled signal domain for the SC modulation scheme, which is employed in uplink communication. All the studies related to the effects of PA nonlinearities on the performance of SC-modulated systems ignore the implications of pulse shaping filter and aliasing due to downsampling. However, in Chapter 3, it is shown that PA amplifier nonlinearity distorts the pulse shaping filter such that memory effects, namely nonlinear inter-symbol interference (ISI), are observed at the received signal.

In Chapter 4, nonlinear post-distortion methods that take nonlinear ISI into account is developed to mitigate the nonlinear distortion at the receiver side. The proposed nonlinear post-distortion methods are employed in the detection of the nonlinearly distorted symbols for uplink communication, where the communication channels exhibit frequency selectivity. The previous studies either handle the linear communication channel and UE PA nonlinearity jointly or assume perfect decoupling of the linear channel and PA non-linearity. By using perfect decoupling, they propose to use zero-forcing or minimum mean square error (MMSE) filtering to mitigate the effects of the frequency selectivity. Chapter 4 contains the first work that shows the mutual effects of the received distortion signal in down-sampled signal domain and the linear channel. It is the first work that reveals a significant phenomenon called **distortion amplification**, which is the result of the mismatch between the effective channels of the desired and distortion signals in downsampled signal domain. In Chapter 4, the effects of distortion amplification are shown via numerical simulations for a single-input single-output (SISO) system. In addition, a distortion-aware receiver structure based on **fractional sampling** is developed to mitigate the effects of distortion amplification by elevating the diversity in the temporal domain.

The complexity of the post-distortion based on the fractional-sampling framework introduced in Chapter 4, requires complex nonlinear signal processing for each fraction. However, it might be cumbersome for multi-user MIMO systems since complexity

also increases significantly with the number of users. In addition, the implementation of fractional sampling is not straightforward for reduced complexity operations. In Chapter 5, we extend the post-distortion approach to a multi-user MIMO system with a hybrid beamforming hardware structure. Then, we develop a sophisticated distortion-aware receiver based on nonlinear distortion cancellation by using the nonlinear distortion estimation method introduced in Chapter 5. By using the distortion cancellation approach, the required complexity is significantly decreased. In addition, in Chapter 5, an extensive signal-to-distortion-plus-noise (SDNR) analysis is carried out, and the effects of the distortion amplification are analytically investigated.

In addition to the investigation of PA nonlinearities on the user side, how the nonlinear distortion is radiated from the multi-antenna base stations equipped with nonlinear PAs is also considered in this thesis. Although it is known that nonlinear distortion disperses in different directions in multi-user scenarios, the quantitative analysis of the received distortion power has not been extensively explored in the literature. All existing studies that derive closed-form SDNR expressions are based on the assumption of uncorrelated distortion across the antenna array. However, in Chapter 6, the validity of this assumption for massive MIMO systems has been investigated, revealing its limited applicability, particularly as the number of antennas increases. Furthermore, within the scope of Chapter 6, a method for finding optimal power allocation coefficients considering distortion correlation across the antenna array has been developed.

Despite the reduction obtained through the spatial diversity techniques, the adverse effects stemming from nonlinear distortion cannot be disregarded and necessitate effective mitigation strategies. Designing DPD for fully connected hybrid beamforming systems poses significant challenges. The inherent complexity of these systems, which integrate both analog and digital components, exacerbates the difficulty of DPD design. The presence of multiple antennas and radio frequency (RF) chains in hybrid beamforming architectures necessitates careful consideration of the nonlinear distortion effects that arise from power amplifiers. In such systems, accurately modeling and compensating for nonlinearities becomes increasingly intricate due to the interactions between the analog and digital domains. Moreover, the coupling between the different RF chains introduces additional complexities in capturing and mitigat-

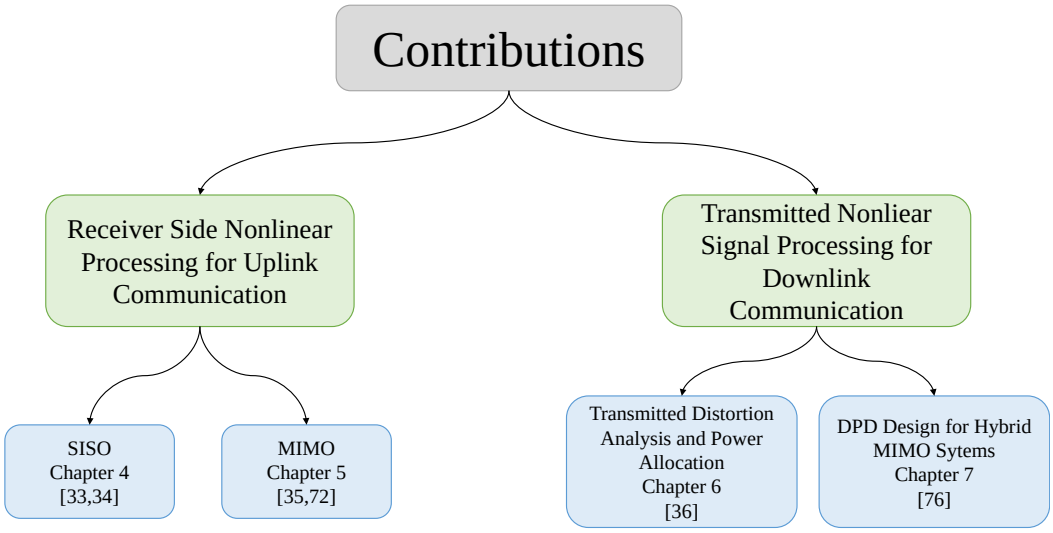


Figure 1.1: Contributions of the thesis.

ing nonlinear distortions effectively. Consequently, the design of DPD algorithms for fully connected hybrid beamforming systems demands sophisticated techniques and careful optimization to achieve satisfactory performance. In Chapter 7, we propose a correlation-based reduced complexity DPD design for hybrid beamforming systems. In the proposed method, the complexity is reduced by avoiding redundant signal processing operations by eliminating DPD implementations in the correlated antenna branches.

The organization of the thesis is summarized in Figure 1.1 and listed below:

In Chapter 2, preliminaries for PA amplifier nonlinearities, which will be used throughout the thesis, are introduced.

In Chapter 3, nonlinear ISI analysis is introduced, and employed post-distortion and nonlinear modeling algorithms are introduced.

In Chapter 4, a receiver structure for single-carrier SISO uplink transmission with frequency domain equalization (FDE) that is exposed to PA nonlinearities is proposed, and several nonlinear distortion modeling techniques are proposed to mitigate the nonlinear distortion [33], [34].

In Chapter 5, the receiver structure proposed in Chapter 4 is extended to MIMO channels, where a hybrid beamforming structure is employed [35].

In Chapter 6, downlink communication is considered for both single user [36] and multi-user massive MIMO systems to examine the received distortion characteristics. A closed-form SDNR expression is derived, and by using the derived expression, a power allocation scheme is proposed.

In Chapter 7, a reduced complexity DPD method is proposed for fully-connected hybrid massive MIMO systems.

Chapter 8, contains the conclusion that can be drawn from the thesis.

CHAPTER 2

PRELIMINARIES ON POWER AMPLIFIER NONLINEARITIES

In this chapter, preliminary information on PA nonlinearities is presented. Firstly, various nonlinear PA modeling tools, namely the Saleh and generalized memory polynomial models will be introduced. In addition, the effects of PA nonlinearities on the output signal for both time and frequency domains will be demonstrated. Furthermore, essential metrics to evaluate the impact of PA nonlinearities will be introduced.

2.1 Power amplifier characterization

In order to investigate the effects of PA nonlinearities, one should have either valid nonlinear PA models or a hardware setup with a PA to obtain meaningful observation. Hence, we use both options in this thesis to provide comprehensive numerical and empirical results.

2.1.1 Saleh Model

The Saleh model [37] is one of the widely employed models in the literature, representing both amplitude-amplitude (AM-AM) and amplitude-phase (PM-AM) distortions. For the input signal at time n x_n, the amplitude and the phase of the PA output signal, namely $|y_n|$ and ϕ_n^y, respectively, can be expressed as

$$|y_n| = \frac{g_0|x_n|}{1 + (|x_n|/A_o)^2}, \tag{2.1}$$

$$\phi_n^y = \frac{\alpha|x_n|^2}{1 + \beta|x_n|^2}, \tag{2.2}$$

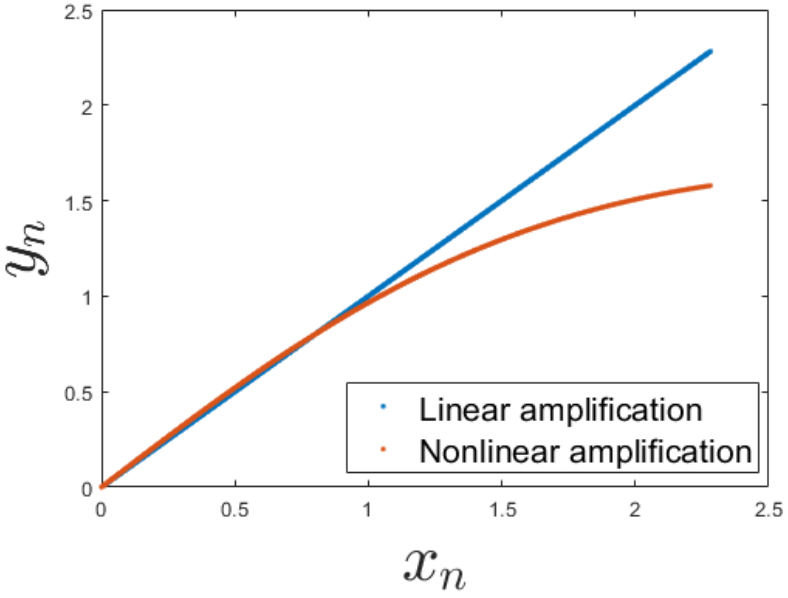


Figure 2.1: Input/output relation for Saleh Model PA and the linear PA.

where g_0 is the gain of the PA, A_o determines the saturation point of the PA, α and β are the parameters related to the phase distortion. Then the overall PA output signal can be expressed as

$$y_n = |y_n| e^{j\phi_n^x} e^{j\phi_n^y} \tag{2.3}$$

where ϕ_n^x is the phase of the input signal. In addition, from (2.1) and (2.2), it can be observed that the PA output only depends on the current input signal, which means that the Saleh model does not reflect the memory effects. In order to investigate the nonlinearity effects, the input-output relationship of the Saleh model is studied as demonstrated in Figure 2.1. As seen from Figure 2.1, the PA output signal tends to saturate as the amplitude of the input signal increases; consequently, the output of the PA cannot be expressed by a linear function of the input signal. Therefore, one should include additional nonlinear terms to the linear signal term to reflect such kind of saturation behavior.

In addition to the distortion in the time domain, nonlinearity causes unwanted signal radiation on the adjacent band channel, which is called out-of-band (OOB) radiation. The OOB radiation can be observed by investigating the power spectral density, which

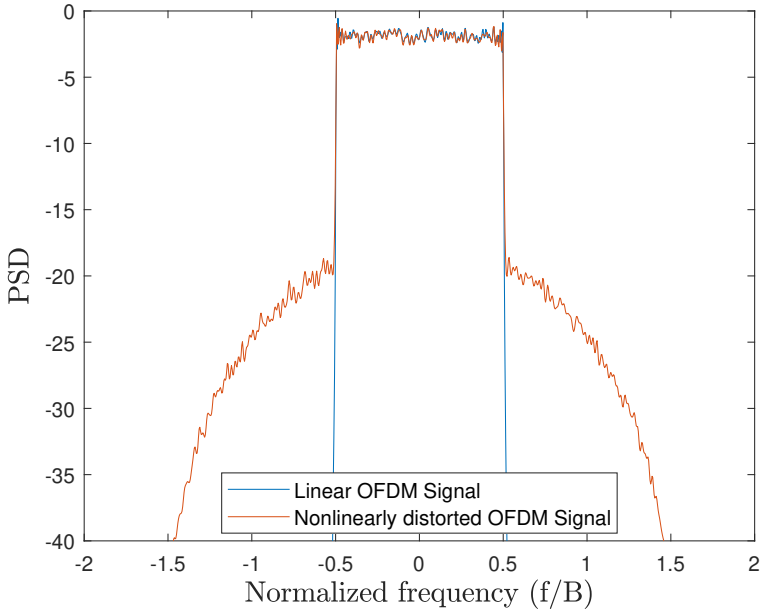


Figure 2.2: PSD for both linear signal and nonlinearly distorted signal.

is shown in Figure 2.2. It should be noted that for Figure 2.2, the input signal is generated as an OFDM waveform with upsampling factor 4. As can be observed from the figure, the linear OFDM signal does not have nonzero PSD for normalized frequencies $-0.5{-}0.5$; on the other hand, nonlinear distorted OFDM signal covers a broader range of frequency band due to spectral regrowth, and it is called as out-of-band radiation. OOB radiation is an undesired signal property of a nonlinearly distorted signal since it degrades the quality of the received signal for other users, which are being served at the adjacent band.

2.1.2 Generalized Memory Polynomial Model

Although the Saleh model provides insights into the nonlinear characteristics of PAs, it cannot capture the hardware's inherent memory effects. Hence, in order to have a comprehensive investigation, in this thesis, a more general model, which is widely employed in the literature, called the "Generalized Memory Polynomial (GMP)" model, is also considered. The nonlinear function that represents PA can be expressed

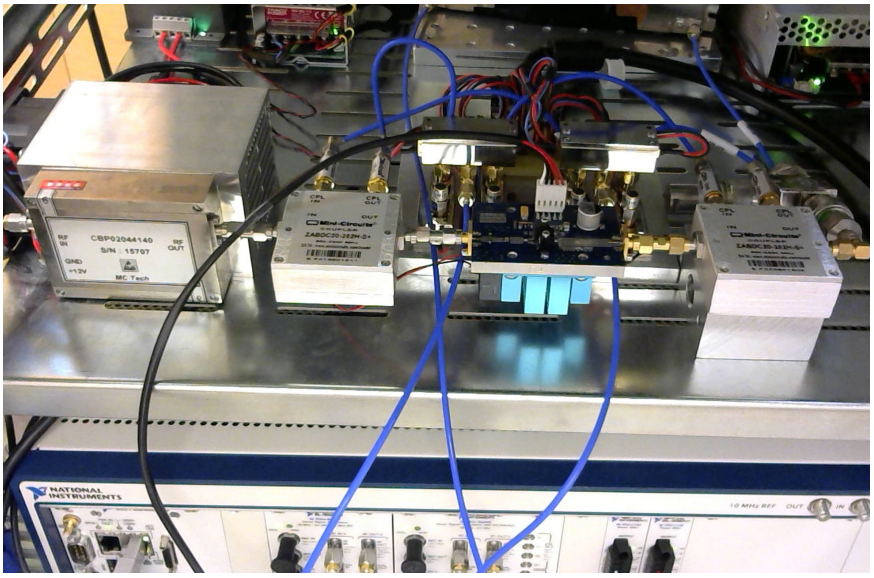


Figure 2.3: Hardware setup used to obtain GMP coefficients [2].

as

$$y_n = \sum_{k=0}^{K_{\mathrm{PA}}-1} \sum_{l=0}^{L_{\mathrm{PA}}-1} \sum_{w=-W_{\mathrm{PA}}+1}^{W_{\mathrm{PA}}-1} \omega_{k,l,w} x_{n-l} |x_{n-l+w}|^k, \tag{2.4}$$

where $\omega_{k,l,w}$'s are the model coefficients, K_{PA} represents the nonlinearity order, and L_{PA} and W_{PA} represent the memory effects. It should be noted that if one selects $W_{\mathrm{PA}} = 1$, then the resultant expression becomes the standard "Memory Polynomial (MP)" model. Throughout the thesis, both GMP and MP models are used to model PA nonlinearity. For this purpose, we employed both coefficients supplied by the previous studies or extracted from a real hardware setup. For instance, in Chapter 4, we employed MP coefficients provided in [1]. To extract the GMP coefficients from the physical hardware, we employ the experimental test setup presented in [2], as depicted in Figure 2.3.

The block diagram of the hardware setup can be seen in 2.4, where it can be seen that measurements are captured from a GaN PA. In order to extract the GMP coefficients, the training sequence of length N_{tr} is formed using MATLAB software. The generated training sequence can be expressed as

$$\mathbf{x}^{\mathrm{tr}} = \left[x_0^{\mathrm{tr}}, x_1^{\mathrm{tr}}, \ldots, x_{N_{\mathrm{tr}}}^{\mathrm{tr}}\right]. \tag{2.5}$$

Then the sequence $\mathbf{x}^{\mathrm{tr}}$ is sent to the system shown in Figure 2.4, and passed through

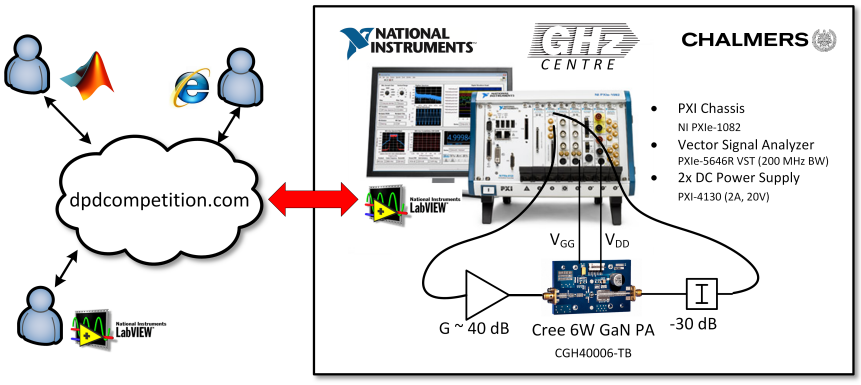


Figure 2.4: Block diagram of the hardware setup used to obtain GMP coefficients [2].

the PA, which can be represented by an arbitrary nonlinear function $\psi(\cdot)$, where each element of PA output can be represented by $y_n^{\mathrm{tr}} = \psi(\{x_n^{\mathrm{tr}}\})$. Then the PA output sequence

$$\mathbf{y}^{\mathrm{tr}} = \left[y_0^{\mathrm{tr}}, y_1^{\mathrm{tr}}, \ldots, y_{N_{\mathrm{tr}}}^{\mathrm{tr}}\right], \tag{2.6}$$

is captured by the receiver chain. Then GMP model coefficients can be obtained by using the least squares (LS) method as described in [38]

$$\boldsymbol{\omega} = \left(\mathbf{X}_{\mathrm{tr}}^{\mathrm{H}}\mathbf{X}_{\mathrm{tr}}\right)^{-1}\mathbf{X}_{\mathrm{tr}}^{\mathrm{H}}\mathbf{y}^{\mathrm{tr}} \tag{2.7}$$

where $\mathbf{X}_{\mathrm{tr}} \in \mathbb{C}^{J\times N_{\mathrm{tr}}}$ is the data matrix for the training sequence whose elements are constructed by using the basis functions in (2.4) as $x_{n-l}|x_{n-l+w}|^k$, $\boldsymbol{\omega}$ is the vector containing weight coefficients $\omega_{k,l,w}$, J is the number of coefficients, and N_{tr} is the length of the training sequence.

In order to investigate the nonlinearity characteristics of the real PA provided in [2], in Figure 2.5, input-output amplitude characteristics are demonstrated. Unlike the input-output relationship in Figure 2.1, which has one-to-one mapping, the output of the PA in [2] can have different values for the same input amplitude. The memory effects account for this behavior since, depending on the neighboring samples, the PA output can have different values. Therefore, it is important to consider the memory effects while handling the PA nonlinearities, and GMP representation can successfully cover memory effects.

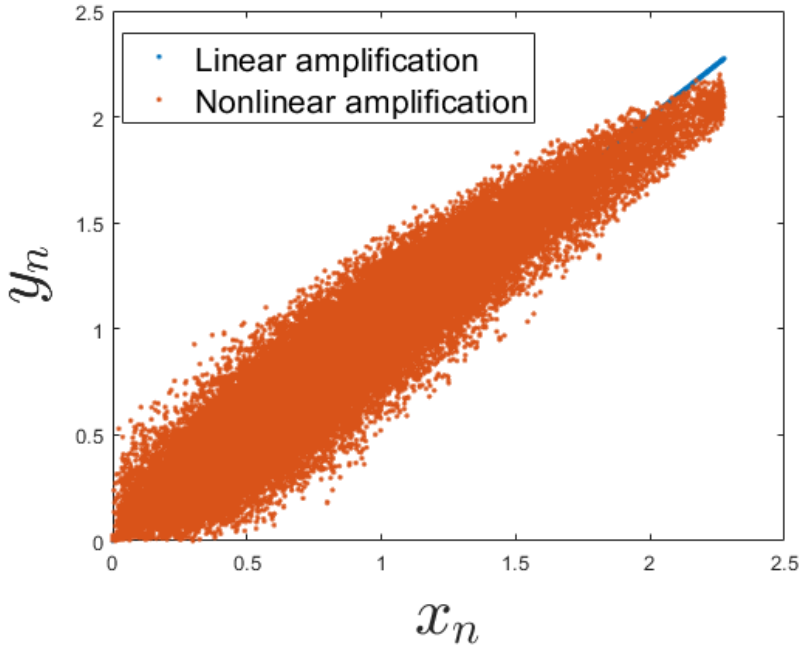


Figure 2.5: Input/output relation for PA in [2] and the linear PA.

2.2 System Parameters Related to Hardware Nonlinearities

2.2.1 Power backoff

Power backoff is an important system parameter that determines the operating region of PAs. In some practical systems, in order to avoid nonlinear distortion, the output power of the amplifier is reduced. The power reduction in the output power is called *output backoff* (OBO); however, increasing OBO is not desired since it reduces the efficiency of the power amplifier. Mathematically, OBO is defined as

$$\text{OBO} \triangleq \frac{P_{\text{max}}}{P_{\text{av}}}, \tag{2.8}$$

where P_{max} and P_{av} are the maximum output power that PA can supply, and average transmit power at the current operating point, respectively.

2.2.2 Adjacent Band Channel Leakage Ratio (ACLR)

Adjacent Band Channel Leakage Ratio (ACLR) is one of the most important metrics that measures how much power is transmitted at the adjacent band, which can also be

considered as the leakage power to the neigbouring channel. ACLR is defined by the ratio between the leakage power and the in-band power as [39]

$$\mathrm{ACLR} = \frac{\max\left\{\int_{-3B/2}^{-B/2} S_y(f)\mathrm{d}y, \int_{B/2}^{3B/2} S_y(f)\mathrm{d}y\right\}}{\int_{-B/2}^{B/2} S_y(f)\mathrm{d}y} \tag{2.9}$$

The maximum allowable ACLR is determined by the standards as ~ -44 dB [13]. In addition, ACLR can also be considered to be an quantitative indicator of the significance of the OOB radiation since it involves the effects of the OOB power over the whole adjacent band.

CHAPTER 3

EQUIVALENT NONLINEAR ISI CHANNEL ANALYSIS

In this section, the nonlinear inter-symbol-interference channel for single-carrier (SC) modulated signals, which occurs as a result of the downsampling of the received nonlinearly distorted signal due to violation of the Nyquist-1 criterion, is investigated.

3.1 Signal Model

The objective of this section is to analyze the effects of downsampling on the nonlinearly distorted received signal; therefore, a simple line-of-sight transmission is investigated for the equivalent nonlinear ISI channel analysis. Firstly data symbols a_k from a QAM alphabet are generated, and a pulse shaping filter is applied to those symbols as

$$x_n = \sum_{k=-\infty}^{k=\infty} a_k p_{n-\mu k}, \tag{3.1}$$

where x_n is the desired transmitted signal, n is the discrete time index, μ is the up-sampling factor, and p_n is the pulse shaping filter, which is selected to satisfy Nyquit-1 criterion to avoid linear ISI. In order to transmit signal x_n, it is up-converted to the carrier frequency, which is then passed through the nonlinear PA represented as $\tilde{x}_n = \psi(x_n)$. It should be noted that here perfect hardware besides non-ideal PA is assumed. For the sake of generality, neither of the models presented in Chapter 2 is considered, an arbitrary nonlinear function $\psi(\cdot)$ is investigated. In the next section, the nonlinear ISI channel as a result of the downsampling operation is studied.

3.2 Discrete-Time Equivalent Model for Nonlinear ISI Channel

In order to gain insight into the characteristics of nonlinear distortion, the distortion term in the received signal is analyzed. For analytical convenience, we considered a memoryless nonlinear model, expressed by arbitrary nonlinear basis functions, $\psi_q(\cdot)$, where q is the nonlinearity order. Then, we can express the transmitted signal after nonlinear amplification as

$$\tilde{x}_n = \sum_{q=0}^{Q-1} \varpi_q \psi_q \left(\sum_{k=0}^{N-1} a_k p_{n-\mu k} \right), \tag{3.2}$$

where $\{\varpi_q\}$ are the model coefficients, $\psi_0(x) = x$ corresponds to linear term, and N is the length of the data sequence. After matched filtering (MF), the received signal can be written as

$$y_n = \varpi_0 \sum_{k=0}^{N-1} a_k \hat{p}_{n-\mu k} + \sum_{l=-\infty}^{\infty} \sum_{q=1}^{Q-1} \varpi_q \psi_q \left(\sum_{k=0}^{N-1} a_k p_{l-\mu k} \right) p^*_{n-l}, \tag{3.3}$$

where $\hat{p} = p_n \circledast p^*_{-n}$ is ideal Nyquist-1 raised cosine impulse response. After sampling, $n = \mu m$, received signal at symbol rate becomes

$$y_m = \varpi_0 a_m + \sum_{q=1}^{Q-1} \sum_{l=-\infty}^{\infty} \varpi_q \psi_q \left(\sum_{k=0}^{N-1} a_k p_{l-\mu k} \right) p^*_{\mu m-l}. \tag{3.4}$$

From (3.4), it can be observed that the term in $\psi_q(\cdot)$ is a unique function for a specific sequence, $\{a_n\}$. Consequently, $\sum_{l=-\infty}^{\infty} \psi_q \left(\sum_{k=0}^{N-1} a_k p_{l-\mu k} \right) p^*_{\mu m-l}$ can be expressed by a nonlinear mapping which depends on symbol index m and symbol sequence $\{a_n\}$. Therefore, (3.4) can also be represented by employing a basis change as

$$y_m = \varpi_0 a_m + \sum_{q=1}^{Q-1} \varpi_q \psi_q^{(\{a_n\},m)} \left(\sum_{l=-\infty}^{\infty} \sum_{k=0}^{N-1} a_k p_{l-\mu k} (\tilde{p}^{(q)}_{\mu m-l})^* \right), \tag{3.5}$$

where $\psi_q^{(\{a_n\},m)}(\cdot)$ represents the unique mapping, which depends on the sequence $\{a_n\}$ and sampling instant m. $\tilde{p}_l^{(q)}$ is the equivalent matched filter impulse response distorted by the nonlinearity, $\psi_q(\cdot)$, which is also defined by another sequence dependent nonlinear mapping[1]. By changing the summation order in nonlinearity, the expression is further simplified to

[1] In this work, we are not interested in finding the mapping, but existence of such mapping is sufficient for rest of the analysis.

$$y_m = \varpi_0 a_m + \sum_{q=1}^{Q-1} \varpi_q \psi_q^{(\{a_n\},m)} \left(\sum_{k=0}^{N-1} a_k \bar{p}_{m-k}^{(q)} \right). \tag{3.6}$$

If the impulse response, $\bar{p}_m^{(q)} = p_l \circledast (\tilde{p}_{-l}^{(q)})^*|_{l=\mu m}$, in (3.6) were ideal raised cosine then $\psi_q^{(\{a_n\},m)}(\cdot)$ would depend only on the single symbol a_m, $\psi_q^{(\{a_n\},m)} = \psi_q^{(a_m)}$. However, due to nonlinear distortion, there is a mismatch in matched filtering, which creates ISI over sequence. Nevertheless, it can be observed that $\bar{p}_l$ is a decaying function, hence, it is assumed that nonlinearity depends on the sequence of a reduced dimension, $\mathbf{a}_m = [a_{m-L_q+1}, ..., a_m, ..., a_{m+L_q-1}]^T$. Consequently, received signal at symbol rate is expressed as

$$y_m = \varpi_0 a_m + \sum_{q=1}^{Q-1} \varpi_q \psi_q^{(\mathbf{a}_m)} \left(\sum_{k=-L_q+1}^{L_q-1} a_{m-k} \bar{p}_k^{(q)} \right). \tag{3.7}$$

It can be inferred from (3.7) that even for systems subjected the memoryless nonlinearity, distortion signal is a nonlinear function of transmitted sequence. Therefore, sequence detection should be performed in order to decode the signal [40]. Consequently, by limiting the memory of the sequence, efficient algorithms can be employed to detect the transmitted sequence. In the literature, several methods employing memoryless detection schemes were proposed [41, 42], which can be considered as the special cases of (3.7), where $\mathbf{a}_m = a_m$,

$$y_m = \varpi_0 a_m + \sum_{q=1}^{Q-1} \varpi_q \psi_q^{(a_m)} \left(\bar{p}_0^{(q)} \, a_m \right). \tag{3.8}$$

Eventually, received signal becomes a symbol dependent function and it can be represented by a symbol dependent coefficient, $\varpi^{(a_m)}$ as, $y_m = \varpi^{(a_m)} a_m$, where

$$\varpi^{(a_m)} = \varpi_0 + \left(\sum_{q=1}^{Q-1} \varpi_q \psi_q^{(a_m)} \left(\bar{p}_0^{(q)} \, a_m \right) \right) \frac{1}{a_m}. \tag{3.9}$$

The complexity of the optimal detector for (3.7) is tremendous; hence, instead of directly implementing the optimal detector, a nonlinear predictor will be designed to estimate a_m. This is more suitable, especially for the presence of the dispersive channel, which increases the computational complexity even further.

CHAPTER 4

AN EFFICIENT QAM DETECTOR VIA NONLINEAR POST-DISTORTION BASED ON FDE BANK UNDER PA IMPAIRMENTS

4.1 Introduction

In this chapter, we propose a receiver structure with frequency domain equalization (FDE) for single-carrier uplink transmission that is exposed to PA nonlinearities. A two-stage approach is adopted, in which linear communication channel is equalized at the first stage, and it is followed by post-distortion, where nonlinear distortion is reduced. In the literature, nonlinear processing techniques are proposed, which perform memoryless compensation of nonlinear distortion together with FDE. However, in Chapter 3, we show that even if memoryless nonlinearity exists, the received signal is impaired by nonlinear inter-symbol interference. Therefore, we propose a class of symbol rate post-distortion techniques, which use neighboring received symbols to suppress the nonlinear interference. Three different post-distortion methods, Gaussian process regression (GPR), neural network (NN) and Volterra series (VS) based post-distorters, are considered. Also, a combiner, which intelligently combines the outputs of fractional delayed bank of FDEs after post-distortion, is proposed to overcome performance degradation of FDE for frequency selective channels under nonlinear distortion. Performances of the proposed techniques are compared with state-of-the-art approaches in terms of bit error rate (BER) and achievable information rate (AIR) metrics. Simulation results demonstrate that post-distortion methods together with bank of FDE outperform state-of-the-art techniques.

4.2 Related Literature

In the literature, there are many methods proposed for the linearization of PAs at the transmitter. Digital predistortion (DPD), which aims to linearize the transmitted signal subjected to nonlinear distortion, is the most popular technique among others [29, 30, 38]. DPD is a suitable solution for downlink transmission since base stations (BS) have quite sufficient computational power to implement complex DPD algorithms. However, implementing DPD at UEs is quite costly; thus, techniques avoiding the utilization of DPD are more suitable for uplink transmission [16, 43]. Post-distortion techniques, which operate on nonlinearly distorted received signals, are developed in order to implement complex compensation algorithms at BS [31, 32, 40, 44, 45]. In [32] and [40], VS based post-distorters are employed as a nonlinear equalizer to detect the transmitted symbols. Besides, in [31], GPR is also adopted as nonlinear equalizer for nonlinear channels to improve the detection performance. However, nonlinear equalizers proposed in [31, 32, 40, 45], treat the nonlinearity together with the wireless channel, which increases parameter optimization complexity. In addition, decision region optimization-based detectors [41, 42] are recently proposed to decode nonlinearly distorted symbols. These techniques exploit constellation-point-dependent distortion, where nonlinear distortion becomes more distinctive as symbol power increases.

Recently, methods that decouple linear and nonlinear channels are proposed in order to reduce computational complexity [44, 46]. These studies consider the system as a memoryless nonlinearity followed by linear wireless channel. By exploiting Hammerstein channel assumption, first, linear wireless channel is equalized by using hybrid decision feedback FDE (HDFDE) then the equalized signal is processed by the memoryless post-distorter, which is the estimated inverse of the memoryless nonlinearity. Complex valued B-spline neural network is utilized in order to extract the inverse nonlinear model [44, 46, 47]. However, pulse shaping, which creates inherent nonlinear memory, is not considered in these studies. The significance of the memory due to pulse shaping is shown numerically in [41, 42] for memoryless PAs. These studies demonstrate that nonlinear ISI yields I/Q correlated distortion and developed modified decision metrics (MM) that outperform conventional detectors by taking

the distortion term into account. However, these detectors do not consider memory in detection, yet they only consider nonlinear ISI as additional distortion. In addition, the effects of nonlinear distortion on the FDE performance are not considered in these studies since the signal model is formed in the symbol-sampled domain, where perfect decoupling of linear and nonlinear systems is possible.

In [48], an iterative block decision feedback equalizer (IB-DFE), called *robust receiver*, is proposed. It considers nonlinearity as the source of an additional noise component. In this study, the feedforward equalization matrix is modified by taking the distortion power into account. Due to the lack of nonlinear receiver processing, the performance of this method is expected to be limited especially for higher-order constellations. In [49], an iterative approach is proposed for multiple-input and multiple-output (MIMO) SC-FDE systems employing quadrature phase shift keying (QPSK) modulation. This approach estimates the nonlinear distortion via an amplifier model by using the tentative decisions at each iteration and cancels the distortion in the frequency domain in order to improve the performance of the *robust receiver*. This study also considers a symbol-sampled channel domain, where both the desired and the distortion signals are exposed to the same effective channel in the frequency domain; hence, the effects of sampling are not considered. Also, the performance of the proposed algorithm depends on the tentative decisions at the first iteration since the algorithm converges in a few iterations. However, for higher-order modulations, tentative decisions can be erroneous, which may degrade the overall performance. Therefore, a nonlinear compensation, before calculating the tentative decisions, may improve the performance of the iterative approach.

In addition to state-of-the-art decision metrics, [41] and [42] develop a framework to evaluate the performance of a system impaired by transceiver non-idealities. Capacity expression based on generalized mutual information (GMI) metric [50] is adopted to obtain mismatched decoding capacity [51, 52]. In this thesis, mismatched decoding capacity is utilized in order to find a lower bound on AIR.

In this chapter, a receiver structure for uplink transmission is proposed, and extensive analysis is carried out for the received signals impaired by transceiver nonlinearities. Two-stage equalization is considered, where the first stage equalizes the linear chan-

nel, and the second stage compensates for the nonlinear distortion.

4.2.1 Contributions

In this chapter, a receiver structure for uplink transmission is proposed, and extensive analysis is carried out for the received signals impaired by transceiver nonlinearities. Two-stage equalization is considered, where the first stage equalizes the linear channel, and the second stage compensates for the nonlinear distortion. Main contributions can be summarized as follows:

- Firstly, a nonlinear distortion analysis is conducted to investigate the effects of frequency-selective channels on the received nonlinear distortion signal. The analysis reveals that the effective channels of the desired signal and the distortion terms are different, and conventional feedforward FDE causes *distortion amplification*. To the authors' knowledge, this phenomenon is not discussed in the literature.

- It is observed that the performance of the receiver, operating at the symbol rate, is significantly affected by the sampling instant due to the difference between the channels of desired and distortion signals. A receiver structure, which utilizes fractionally spaced samples, is proposed to improve the performance by exploiting channel diversity. In this structure, soft symbol estimates, produced by nonlinear post-distortion filters at each fraction, are combined to detect the transmitted symbols by taking the distortion for each effective channel into account.

- An equivalent nonlinear ISI channel model analysis is carried out, which shows that even for memoryless nonlinearity, the received signal suffers from nonlinear distortion depending on the neighboring symbols due to pulse shaping. Performance improvement is achieved compared to state-of-the-art techniques [41, 42, 44] by including memory with cross-terms in the nonlinear processing as presented in ISI channel analysis, even for the presence of memoryless nonlinearity.

- Three nonlinear post-distortion techniques, based on NN, GPR and VS structures following the linear FDE, are proposed. These methods are independent of instantaneous channel state information (CSI). In addition, the best linear unbiased estimator (BLUE) [53] is adopted to further decrease the complexity of GPR hyper-parameter optimization.

4.3 System Model

4.3.1 SC-FDE Based Transmitter Model

A conventional SC-FDE-based transmission scheme shown in Figure 4.1 is considered. At each data block, N_D number of P-QAM symbols, $[a_0, a_1, ..., a_{N_D-1}]^T_{N_D \times 1}$ with unit energy, $\mathbb{E}\{|a_k|^2\} = 1$, are transmitted. A cyclic prefix (CP) and cyclic suffix (CS) are added to prevent inter-block interference and create a circulant channel matrix. Then, the symbol vector for the transmitted block becomes

$$\mathbf{a} = [a_{N_D-N_{CP}}, \ldots, a_{N_D-1}, a_0, a_1, \ldots, a_{N_D-1}, a_0, \ldots, a_{N_{CS}-1}]^T, \tag{4.1}$$

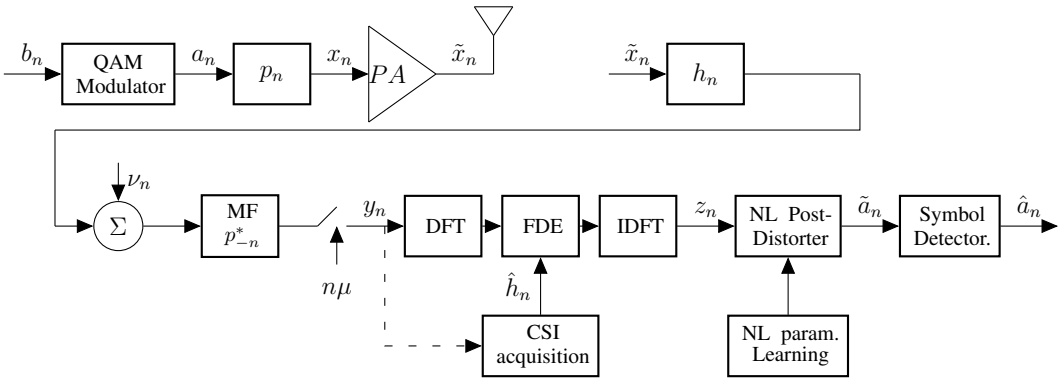


Figure 4.1: Transciever structure for SC-FDE transmission system.

where N_{CP} and N_{CS} are the lengths of CP and CS respectively. In discrete time, signal to be transmitted can be expressed as

$$x_n = \sum_{k=0}^{N-1} a_k p_{n-\mu k}, \tag{4.2}$$

where p_n is the upsampled pulse shaping filter response with upsampling factor μ and $N = N_D + N_{CP} + N_{CS}$. Then, the transmitted signal, $x_n = |x_n|e^{j\phi_{i,n}}$, is fed to PA, which nonlinearly distorts the signal. In this study, Saleh model [37], which is commonly employed in simulations, and a realistic PA model [1], which is extracted from an actual hardware are considered. For Saleh model, output of the nonlinear PA, $\tilde{x}_n = |\tilde{x}_n|e^{j\phi_{i,n}}e^{j\theta_n}$, can be expressed as

$$|\tilde{x}_n| = \frac{g_0|x_n|}{1 + (|x_n|/A_{sat})^2}, \quad \theta_n = \frac{\alpha|x_n|^2}{1 + \beta|x_n|^2}, \tag{4.3}$$

where g_0, A_{sat}, α and β are the model parameters. In addition to Saleh model, we considered a model that is extracted based on the measurements on the GaN power amplifier [1]. This model is based on memory polynomial [38] and the output of the PA is represented as

$$\tilde{x}_n = \sum_{k=0}^{K_b-1} \sum_{l=-P_b+1}^{P_b-1} \sum_{m=-P_c+1}^{P_c-1} c_{k,l,m} x_{n-l} |x_{n-l-m}|^{2k}, \tag{4.4}$$

where $c_{k,l,m}$'s are the model coefficients. In general, PA output can be represented by a generic nonlinear function as $\tilde{x}_n = \Psi(\{x_n\})$. Details of PA models will be presented in Section 4.8. Then, the discrete-time baseband received signal after matched filtering (MF) becomes

$$y_n = \left(\sum_{l=0}^{L-1} h_l \tilde{x}_{n-l} + \nu_n \right) \circledast p^*_{-n}, \tag{4.5}$$

where h_l's $l = 0, 1, ..., L-1$ are complex channel coefficients, $\circledast$ denotes convolution sum, and ν_n is zero-mean complex additive white Gaussian noise (AWGN) with variance $\mathbb{E}\{|\nu_n|^2\} = N_0$.

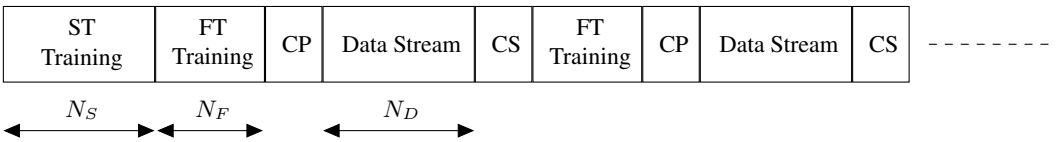


Figure 4.2: Frame structure for the transmition scheme.

4.3.2 Frame Structure

The frame structure, which is employed in this study, is shown in Figure 4.2. At the beginning of the transmission, a training sequence is sent to perform parameter learning for the nonlinear post-distortion. In the proposed system, nonlinearly distorted transmitted signals are captured by the observation chain of UE and sent to BS, and nonlinear parameter learning (NPL) is performed under no channel impairments with high signal-to-noise ratio (SNR)[1]. Since PA characteristics show a slow change compared to wireless channel, NPL, which can be referred as slow time (ST) training, is not performed very often. Thus, NPL does not bring any overhead to the system. ST sequence of length, N_S, is only transmitted when the nonlinear characteristics of the PA are changed. After NPL stage, data transmission starts. A block fading system is assumed and the CSI is not available. Since channel estimation is beyond the scope of this study, a sufficiently long training sequence having N_F symbols is transmitted at the beginning of each data block. Least squares (LS) method is employed to obtain CSI in fast time (FT) training stage.

4.4 Distortion Analysis and Pre-processing : Bank of FDE

In this section, an analysis on the effects of nonlinear distortion on FDE performance under frequency selective channels is presented. Also, based on this analysis, a receiver is proposed to decrease the effects of nonlinear distortion by exploiting channel diversity.

4.4.1 Distortion Analysis for Frequency Selective Channels

To analyze the effects of frequency selective channels on a nonlinearly distorted signal, signal under consideration should be decomposed into linear and nonlinear parts so that each term can be analyzed individually. Decomposition can be performed by

[1] Alternatively, NPL can be performed at the BS after CSI acquisition and equalization stages under channel impairments and at relatively low received SNR levels by repeating the same ST training sequence in order to increase effective SNR for NPL.

using Bussgang decomposition as

$$\tilde{x}_n = \alpha_x x_n + \gamma_n, \qquad \alpha_x = \frac{\mathbb{E}[x_n^* \tilde{x}_n]}{\mathbb{E}[x_n^* x_n]}, \tag{4.6}$$

where x_n and $\tilde{x}_n$ are defined in (4.2) and (4.3), respectively. Bussgang coefficient, α_x, can be found by Wiener filtering given in (4.6) and γ_n is the remaining distortion term. Received signal after matched filtering can be expressed in frequency domain, $Y(e^{j\omega})$, as

$$Y(e^{j\omega}) = \alpha_x X(e^{j\omega}) H(e^{j\omega}) P^*(e^{j\omega}) + \Gamma(e^{j\omega}) H(e^{j\omega}) P^*(e^{j\omega}), \tag{4.7}$$

where $X(e^{j\omega}) = \sum_{m=0}^{N_D-1} a_m P(e^{j\omega}) e^{-jm\omega\mu}$ is the desired signal spectrum, which is equivalent to $X(e^{j\omega}) = P(e^{j\omega}) A(e^{j\mu\omega})$ and $A(e^{j\omega}) = \sum_{m=0}^{N_D-1} a_m e^{-jm\omega}$. In (4.7), $H(e^{j\omega}) = \sum_{k=0}^{L-1} h_k e^{-jk\omega}$ and $P(e^{j\omega}) = \sum_{m=-\infty}^{\infty} p_m e^{-jm\omega}$ are Fourier transforms of h_n and p_n. Fourier transform of the distortion, $\Gamma(e^{j\omega}) \triangleq \sum_{n=0}^{\mu N_D-1} \gamma_n e^{-jn\omega}$, is related to power spectral density (PSD) of the distortion term as

$$S_\Gamma(e^{j\omega}) = \lim_{N_D \to \infty} \mathbb{E}\left[\frac{1}{\mu N_D} |\Gamma(e^{j\omega})|^2\right] \tag{4.8}$$

which is given in [54] as

$$S_\Gamma(e^{j\omega}) = \sum_{s=1}^{\infty} p_{\Gamma,s} \underbrace{S_X(e^{-j\omega}) \circledast ... \circledast S_X(e^{-j\omega})}_{s \ terms} \underbrace{S_X(e^{j\omega}) \circledast ... \circledast S_X(e^{j\omega})}_{s+1 \ terms}, \tag{4.9}$$

where $S_X(e^{j\omega})$ is PSD of x_n, and $p_{\Gamma,s}$ can be considered as the power of the s^{th} order nonlinearity. Furthermore, it can be inferred from (4.9) that spectral regrowth occurs due to nonlinearity. Consequently, (4.7) is simplified to

$$Y(e^{j\omega}) = \alpha_x A(e^{j\mu\omega}) |P(e^{j\omega})|^2 H(e^{j\omega}) + \Gamma(e^{j\omega}) H(e^{j\omega}) P^*(e^{j\omega}). \tag{4.10}$$

After decimation, sampling rate is reduced to symbol rate, and downsampled signal in frequency domain, namely $Y^d(e^{j\omega})$, can be expressed as

$$\begin{aligned} Y^d(e^{j\omega}) = & \underbrace{\alpha_x A(e^{j\omega}) \frac{1}{\mu} \sum_{i=0}^{\mu-1} H(e^{j(\frac{\omega-2\pi i}{\mu})}) |P(e^{j(\frac{\omega-2\pi i}{\mu})})|^2}_{\Phi(e^{j\omega}):\ Linear\ Term} + \\ & \underbrace{\frac{1}{\mu} \sum_{i=0}^{\mu-1} \Gamma(e^{j(\frac{\omega-2\pi i}{\mu})}) H(e^{j(\frac{\omega-2\pi i}{\mu})}) P^*(e^{j(\frac{\omega-2\pi i}{\mu})})}_{\Psi(e^{j\omega}):\ Distortion\ Term}, \end{aligned} \tag{4.11}$$

where $\frac{\alpha_x}{\mu}\sum_{i=0}^{\mu-1} H(e^{j(\frac{\omega-2\pi i}{\mu})})|P(e^{j(\frac{\omega-2\pi i}{\mu})})|^2$ can be interpreted as the effective channel frequency response, $H_{eff}(e^{j\omega})$. In addition, PSD of downsampled distortion signal can be expressed as

$$S_\Psi(e^{j\omega}) = \frac{1}{\mu}\sum_{i=0}^{\mu-1} S_\Gamma\left(e^{j(\frac{\omega-2\pi i}{\mu})}\right)\left|H\left(e^{j(\frac{\omega-2\pi i}{\mu})}\right)\right|^2\left|P\left(e^{j(\frac{\omega-2\pi i}{\mu})}\right)\right|^2. \tag{4.12}$$

From (4.11), it can be observed that the channels experienced by data symbols, a_m's, and the distortion signal are different, in general. Therefore, a particular frequency, ω_f, which yields fading, $H_{eff}(e^{j\omega_f}) \approx 0$, for symbol spectrum, may not cause fading for distortion. Consequently, distortion power is amplified during FDE operation since large gain is applied to corresponding frequency component of distortion to equalize deep fades of $H_{eff}(e^{j\omega})$. On the other hand, if communication channel is sparse such that non-zero taps only exist at symbol times then $H(e^{jw})$ becomes periodic with $\frac{2\pi}{\mu}$. Therefore, for such a channel, $H(e^{j(\frac{\omega-2\pi i}{\mu})})$ term can be moved outside the summations. Consequently, both linear and distortion terms experience similar channels, which does not yield any distortion amplification. In Figure 4.3, PSD's of both distortion and linear signal terms are shown. It is observed from Figure 4.3(a) that for a symbol rate channel both linear and distortion terms experience the same equivalent channel. However, for a more realistic upsampled channel model, it can be observed from Figure 4.3(b) that linear and distortion signals are subject to different equivalent channels where linear signal is having a deep fade but distortion term is not. Consequently, distortion amplification occurs after FDE operation and amplified distortion is spread over all symbols via IDFT operation. Besides, from Figure 4.3(c), it can be seen that if the sampling instant is shifted by $T/4$ then the resulting effective channel may not have deep fades for the same wireless medium. Hence, we propose a receiver structure, which is shown in Figure 4.4, that exploits channel diversity obtained from different sampling instances.

4.4.2 Channel Acquisition and FDE Operation

Sampled signal is divided into μ_B branches such that the signal on the i^{th} branch is given as

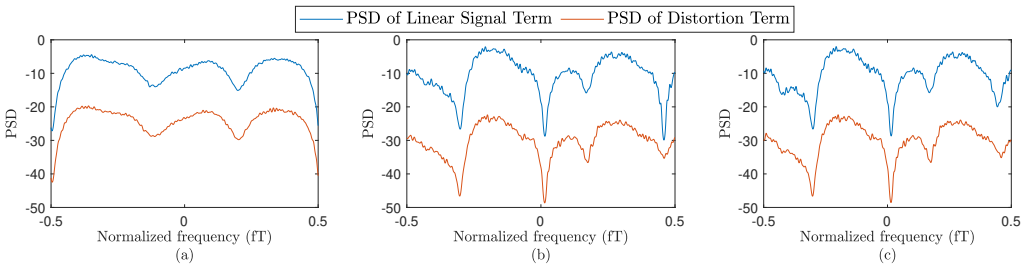


Figure 4.3: PSD of signals for (a) symbol sampled channel and a random channel realization from the given Rayleigh model in *Numerical Evaluations* with fractional delay of (b) 0 and (c) $T/4$.

$$y_n^{(i)} = y_{n\mu+(i-1)\frac{\mu}{\mu_B}}, \quad i = 1, ..., \mu_B \tag{4.13}$$

where it can be assumed that both cyclic prefix and suffix parts of the data block are discarded. For each branch standard FDE procedure is followed. Firstly, by using fast time training sequence, LS estimate of channel, $\hat{\mathbf{h}}^{(i)} = [\hat{h}^{(i)}_{-L_b+1}, ..., \hat{h}^{(i)}_0, ..., \hat{h}^{(i)}_{L_f-1}]^T$, for the i^{th} branch is obtained as

$$\hat{\mathbf{h}}^{(i)} = ([\mathbf{A}]^H \mathbf{A})^{-1} [\mathbf{A}]^H \mathbf{y}^{(i)}_{\text{fast}}, \tag{4.14}$$

where $\mathbf{A} \in \mathbb{C}^{N_F \times (L_b+L_f-1)}$ is the data matrix for FT sequence and the received sequence for FT training stage of the i^{th} branch is

$$\mathbf{y}^{(i)}_{\text{fast}} = [y^{(i)}_0, y^{(i)}_1, ..., y^{(i)}_{N_F-1}]^T_{N_F \times 1}. \tag{4.15}$$

Note that estimated channel in (4.14) has also anticausal parts to recover synchronization errors. In this study, conventional channel estimation is modified in order to decrease the effects of nonlinear distortion on estimation accuracy. For this purpose, elements of the data matrix are modified as, $[\mathbf{A}]_{k,l} = \bar{a}_{(k-l)+L_b-1}$, where $\bar{a}_n = \left(\Psi(x_m) \circledast p^*_{-m}\right)|_{m=\mu n}$ is the nonlinearly distorted symbol after MF and downsampling corresponding to the symbol a_n of the FT sequence. These symbols can be obtained during ST training stage since the FT sequence is not changed during transmission. If a_n's are employed as the elements of $\mathbf{A}$ then the nonlinear distortion would behave as additional noise; hence, the effective SNR for CSI acquisition decreases.

FDE is performed for all branches on the received information signals. DFT of the received signal is computed as $\mathbf{y}^{(i)}_f = \mathbf{Q}^H \mathbf{y}^{(i)}$ where $\mathbf{y}^{(i)} = [y^{(i)}_0, y^{(i)}_1, ..., y^{(i)}_{N_D-1}]^T$

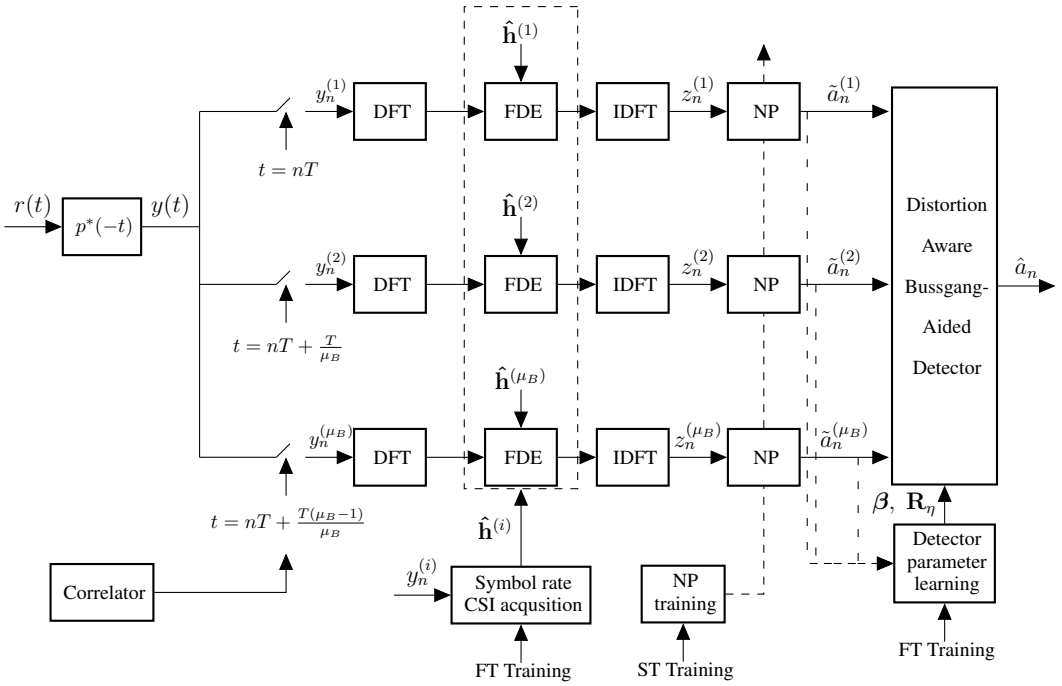


Figure 4.4: Block diagram for proposed bank of FDE's based detector.

and $\mathbf{Q}$ is the $N_D \times N_D$ normalized DFT matrix with the $(m,n)^{th}$ element, $q_n^m = \frac{1}{\sqrt{N_D}} e^{j2\pi mn/N_D}$. Then, the received signal in the frequency domain can be expressed as

$$\mathbf{y}_f^{(i)} = \text{diag}\{\lambda_k^{(i)}\}_{k=0}^{N_D-1} \mathbf{a}_f + \mathbf{n}_f^{(i)}, \tag{4.16}$$

where $\mathbf{a}_f \triangleq \mathbf{Q}^H \mathbf{a}$, $\mathbf{n}_f^{(i)} \triangleq \mathbf{Q}^H \boldsymbol{\nu}^{(i)}$, where $[\boldsymbol{\nu}^{(i)}]_n \triangleq (\nu_n \circledast p_{-n}^*)|_{n \to n\mu + (i-1)\frac{\mu}{\mu_B}}$, and the estimated equivalent channel frequency response can be expressed as $\lambda_k^{(i)} = \frac{1}{\sqrt{N_D}} \sum_{l=-L_b+1}^{L_f-1} \hat{h}_l^{(i)} e^{-j2\pi kl/N_D}$ for $k = 0, \ldots, N_D - 1$. To perform FDE, minimum mean squared error (MMSE) estimation type filtering is applied in frequency domain as

$$\mathbf{z}_f^{(i)} = \text{diag}\left\{ \frac{[\lambda_k^{(i)}]^*}{|\lambda_k^{(i)}|^2 + \delta} \right\}_{k=0}^{N_D-1} \mathbf{y}_f^{(i)}, \tag{4.17}$$

where $\delta = \frac{N_o}{E_s}$ is the regularization parameter, E_s is the average signal power and N_o is the noise variance. After performing FDE, time domain signal, $\{z_n^{(i)}\}$ for the i^{th} branch, is obtained via IDFT, $\mathbf{z}^{(i)} = \mathbf{Q}\mathbf{z}_f^{(i)}$. Note that, in this study, only linear FDE is considered; however, the performance of the receiver can be improved further by utilizing iterative FDE techniques [48, 55]. By using decision feedback, as in IB-DFE [48], remaining interference effects can be reduced so that superior detection

performance can be achieved. Following the linear equalization, post-distortion is performed to suppress the nonlinear distortion.

4.5 Symbol Rate Nonlinear post-distortion based on FDE Bank

In this section, a symbol rate nonlinear post-distortion (NP) approach for higher-order constellations is considered. In addition, a combiner, which employs FDE bank, will be developed to exploit diversity to minimize distortion amplification effects.

4.5.1 Post Processing: Symbol rate nonlinear post-distortion

Different symbol rate NP techniques based on GPR, augmented real-valued time-delay neural network (ARVTDNN) and VS are considered. The mentioned nonlinear post-distortion units are trained in ST training mode, whose details will be presented in the subsequent sections. The aim of the NP is to nonlinearly modify the incoming signal so that desired output is produced. However, one should note that we are not interested in finding the inverse model of the PA, $\Psi^{-1}(\cdot)$ but performing a symbol rate nonlinear processing, which eliminates the nonlinear ISI. A detailed analysis on equivalent nonlinear ISI channel is provided in Chapter 3. In addition, the analysis shows that combination of neighboring symbols are affected by the nonlinearity. Hence methods, which utilizes cross-terms, should be considered for NP to reduce the received nonlinear distortion.

4.5.2 Proposed Method for Slow-Time Training Data Set Formation

In order to train the post-distortion methods that are proposed in the thesis, a proper training set should be constructed. Since the objective is to minimize the detection error, it is desired to avoid imperfections other than the hardware impairments such as AWGN and channel frequency selectivity. Otherwise, the training quality may degrade significantly. It should be noted that the main responsibility of the nonlinear post-distortion unit is to recover the transmitted (*desired*) QAM symbols from the nonlinearly distorted received signals. In addition, it is assumed that the effects of

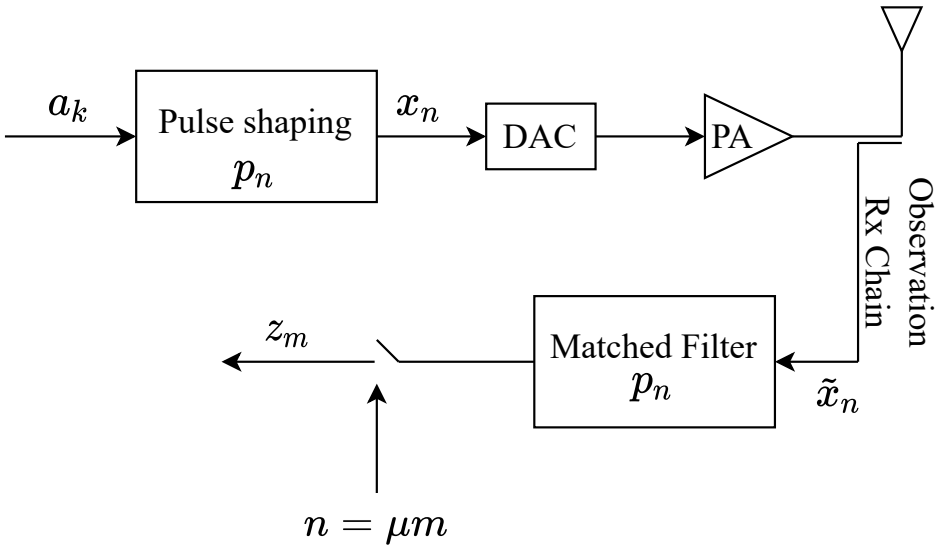


Figure 4.5: Block diagram of the hardware setup used to obtain training set at the UE.

the frequency-selective communication channel are mitigated prior to nonlinear processing via FDE. Therefore, nonlinear post-distortion is used to eliminate the solely nonlinear distortion in the down-sampled signal domain. Hence, in this study, two methods are proposed to acquire training samples such that they only contain nonlinear distortion after downsampling.

4.5.2.1 *First Alternative*: Acquisition of training samples via observation receiver at the UE

In order to capture the PA output, an observation receiver can be equipped at the UE. The block diagram for the proposed data acquisition structure is shown in Figure 4.5. In the proposed approach, the training QAM symbols are generated and pass through the PA, which nonlinearly distorts the signal $\tilde{x}_n$. Then the nonlinearly distorted signal $\tilde{x}_n$ is captured by employing an additional Rx chain. The captured signal is converted to digital signals and matched filtering is applied to that signal. Finally, matched filter output is sampled with sampling rate μ to obtain training symbols for nonlinear post-distortion z_n. It should be noted that since the observation receiver directly measures the PA output, SNR for parameter learning is high and the frequency response of the observation channel is flat.

4.5.2.2 *Second Alternative*: Acquisition of training samples at the BS based on channel integration

An alternative procedure for the construction of the training set under channel impairments with lower SNR is also proposed. For this case, ST training sequence is not captured by the UE, but received ST sequence, passing through a wireless channel, at BS is used for learning as depicted in Figure 4.1. For this case, in order to obtain channel diversity against noise and distortion amplification after FDE stage, ST period can be divided into K segments and the same training sequence is sent at each segment. Then, we can denote, n^{th} symbol of i^{th} training sequence as $a_n^{(i)}$:

$$a_n^{(i)} = a_n,\ i = 1, \ldots, K. \tag{4.18}$$

Each training segment is sent consecutively and they are exposed to different communication channels. After channel acquisition and linear equalization (FDE) stages, the received symbol for each ST training segment can be written in time domain as

$$z_n^{(i)} = \Omega(a_n) + \eta_n^{(i)}, \tag{4.19}$$

where $\Omega(\cdot)$ is the nonlinear function representing PA, $\eta_n^{(i)}$ is the additional distortion for the i^{th} segment due to mismatch between linear wireless channel and channel experienced by the nonlinear distortion term, and AWGN. After receiving all data segments, they can be integrated as

$$z_n = \sum_{i=1}^{K} w_i z_n^{(i)} \tag{4.20}$$

where w_i is the weight of each segment. One can simply use w_i as, $w_i = 1/K$. Thanks to this integration, received SNR is enhanced at least by the factor of K. The number of segment can be chosen to be high enpugh to have sufficient SNR for NPL; however, it should not be long enough to create latency problems. Then, $\{a_n, z_n\}$ sequence can be utilized to train NPL.

4.5.3 GPR based NP

GPR is developed to predict the output of a nonlinear system by minimizing the mean squared error in function space [56]. It can be considered as the Wiener solution of the nonlinear identification problem since system output is assumed to have Gaussian distribution. The nonlinear model based on GPR for real, $a_n^I = \mathrm{Re}\{a_n\}$, and imaginary, $a_n^Q = \mathrm{Im}\{a_n\}$, parts of the desired NP output is described as [56, 57]

$$a_n^I = \Omega_I(\bar{\mathbf{z}}_n) + \nu_n^I \quad \text{and} \quad a_n^Q = \Omega_Q(\bar{\mathbf{z}}_n) + \nu_n^Q, \tag{4.21}$$

where ν_n^I, ν_n^Q are the I/Q parts of the modeling error, which are zero-mean white Gaussian processes (GP) independent of other sequences with variance σ_ν^2, and the latent functions $\Omega_I(\cdot)$ and $\Omega_Q(\cdot)$ are the functions to be identified. Input vector, $\bar{\mathbf{z}}_n$, has the augmented form as

$$\begin{aligned}\bar{\mathbf{z}}_n = [&\mathrm{Re}\{z_{n+M-1}\}, ..., \mathrm{Re}\{z_n\}, ..., \mathrm{Re}\{z_{n-M+1}\},\\ &\mathrm{Im}\{z_{n+M-1}\}, ..., \mathrm{Im}\{z_n\}, ..., \mathrm{Im}\{z_{n-M+1}\}]^T,\end{aligned} \tag{4.22}$$

where M is the memory depth. Training set for GPR based NP is defined as, $\{a_n, \bar{\mathbf{z}}_n\}_{n=0}^{N_s-1}$ where a_n is the desired output. In this section, we focus on modelling the in-phase component; however, one can use the same procedure for the imaginary part. In order to express distribution of GP model in (4.21), required functions, vectors and matrices are defined as follows:

- Kernel function, which is the cross correlation between different samples of the latent process, is defined as $k(\bar{\mathbf{z}}_p, \bar{\mathbf{z}}_q) \triangleq \mathbb{E}\{\Omega_I(\bar{\mathbf{z}}_p)\Omega_I(\bar{\mathbf{z}}_q)\}$.
- Kernel matrix $\mathbf{K}$ is defined as the correlation matrix with entries, $[\mathbf{K}]_{p,q} = k(\bar{\mathbf{z}}_p, \bar{\mathbf{z}}_q)$.
- Kernel steering vector, namely $\mathbf{k}_* \triangleq \mathbb{E}\{\mathbf{a}_{\text{slow}}\Omega_I(\bar{\mathbf{z}}_*)\}$, is defined as the cross-correlation between the training and the test signals, and

  $$\mathbf{a}_{\text{slow}} \triangleq [a_0^I, a_1^I, ..., a_{N_s-1}^I]^T \tag{4.23}$$

  is the training sequence vector. It is used to predict the test symbol, a_*^I by using z_*, where $*$ denotes time indices in data sequence, p^{th} element of $\mathbf{k}_*$ can be calculated as $[\mathbf{k}_*]_p = k(\bar{\mathbf{z}}_p, \bar{\mathbf{z}}_*)$ due to the independence of $\Omega_I(\bar{\mathbf{z}}_p)$ and ν_p^I.

Based on defined kernel functions, the predictive distribution[2] can be written as [56, 57]

$$p(a_*^I|\{a_n, \bar{\mathbf{z}}_n\}_{n=0}^{N_s-1}, \bar{\mathbf{z}}_*) \sim \mathcal{N}\left(\mu_*, \sigma_*^2\right), \tag{4.24}$$

where μ_* and σ_*^2 are mean and variance of the distribution, and they can be obtained as

$$\mu_* = \mathbf{k}_*^T \left(\mathbf{K} + \sigma_\nu^2 \mathbf{I}\right)^{-1} \mathbf{a}_{\text{slow}}, \quad \sigma_*^2 = k(\bar{\mathbf{z}}_*, \bar{\mathbf{z}}_*) - \mathbf{k}_*^T (\mathbf{K} + \sigma_\nu^2 \mathbf{I})^{-1} \mathbf{k}_* + \sigma_\nu^2. \tag{4.25}$$

Then, symbol estimates at the output of NP becomes the mean of the predictive distribution:

$$\tilde{a}_*^I = \mathbb{E}\{a_*^I|\{a_n, \bar{\mathbf{z}}_n\}_{n=0}^{N_s-1}, \bar{\mathbf{z}}_*\} = \mathbf{k}_*^T \left(\mathbf{K} + \sigma_\nu^2 \mathbf{I}\right)^{-1} \mathbf{a}_{\text{slow}}. \tag{4.26}$$

In this study, kernel function is chosen as the exponential function that is widely employed in literature, $k(\bar{\mathbf{z}}_p, \bar{\mathbf{z}}_q) = \sigma_f^2 \exp\left(-\bar{\mathbf{z}}_p^T \mathbf{Q}^{-1} \bar{\mathbf{z}}_q\right)$, where $\mathbf{Q} = \text{diag}\{c_i^2\}_{i=1}^{4M-2}$ and c_i's are the length scale parameters and σ_f^2 is the signal variance. Maximum likelihood estimation is used to find the hyper-parameters, $\beta = [\sigma_f, \sigma_\nu, c_1, ..., c_{4M-2}]^T$ [31, 57].

Hyper-parameter optimization is a time consuming process, and training time exponentially increases with the length of the training sequence. Hence, higher-order P-QAM processing requires a long training sequence to train GPR parameters. In order to reduce the training time, we propose BLUE type post processing such that training sequence is divided into non-overlapping segments to train different GPR units and intelligently combine the output of each unit.

Estimated symbol for i^{th} segment with training set $\{a_n^{(i)}, \bar{\mathbf{z}}_n^{(i)}\}_{n=0}^{N_S^{(i)}-1}$ is obtained as $\tilde{a}_*^{I,(i)} = [\mathbf{k}_*^{(i)}]^T \left(\mathbf{K}^{(i)} + \sigma_\nu^2 \mathbf{I}\right)^{-1} \mathbf{a}_{\text{slow}}^{(i)}$, where $\mathbf{k}_*^{(i)}$ is the cross-correlation vector, $\mathbf{K}^{(i)}$ is the kernel matrix, and $\mathbf{a}_{\text{slow}}^{(i)}$ is the training symbols for the i^{th} segment, $N_S^{(i)}$ is the training length for i^{th} segment, and S is the number of segments: $\sum_{i=1}^{S} N_S^{(i)} = N_S$. For simplicity, we dropped I/Q indices, and variance of each symbol estimate, $(\sigma_*^i)^2$, can be obtained as

$$(\sigma_*^i)^2 = k(\bar{\mathbf{z}}_*^{(i)}, \bar{\mathbf{z}}_*^{(i)}) - [\mathbf{k}_*^{(i)}]^T \left(\mathbf{K}^{(i)} + \sigma_\nu^2 \mathbf{I}\right)^{-1} [\mathbf{k}_*^{(i)}] + \sigma_\nu^2, \tag{4.27}$$

where $\bar{\mathbf{z}}_*^{(i)}$ is test data for i^{th} segment. Estimates are fused based on BLUE combiner

[2] GPR prediction is linear Wiener filtering in function space and it predicts the best function in MMSE sense by using correlation information via Kernel functions. Hence, (4.25) still yields valid linear MMSE (LMMSE) estimate in function space for non-Gaussian distortion.

as in [53]

$$\tilde{a}_* = \sum_{i=1}^{S} w_*^i \tilde{a}_*^{(i)}, \tag{4.28}$$

where $\mathbf{w}_* = [w_*^1, ..., w_*^S]^T$ is the weight vector obtained as, $\mathbf{w}_* = \frac{\mathbf{D}_*^{-1}\mathbf{1}}{\mathbf{1}^T\mathbf{D}_*^{-1}\mathbf{1}}$, where $\mathbf{D}_* \triangleq \text{diag}\{(\sigma_*^1)^2, ..., (\sigma_*^S)^2\}$ is the diagonal variance matrix.

4.5.4 Neural Network (NN) based NP

In this chapter, ARVTDNN structure shown in Figure 4.6, which is employed to design digital predistortion unit in [29] via behavioral modelling, is used as the nonlinear post-distorter. However, the functionality of the NN in this study is different compared to that of DPD. In [29], NN is used for nonlinear system identification via nonlinear regression so that the inverse of the nonlinear function is obtained. However, in this study, we are employing NN to estimate transmitted symbols chosen from a discrete alphabet.

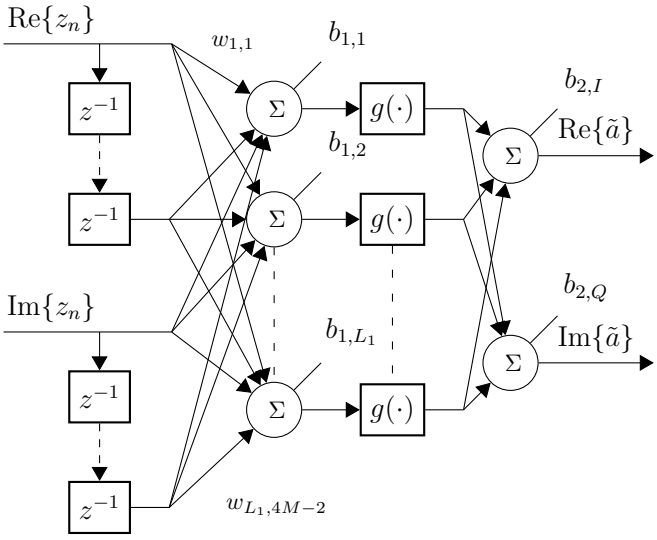


Figure 4.6: ARVTDNN architecture.

The same input signal, $\bar{\mathbf{z}}_n$, which is defined for GPR estimation, is also used for NN post-distortion. Then, symbol estimate at the output of NN can be expressed as $\tilde{a}_n = \Omega_I(\bar{\mathbf{z}}_n) + j\Omega_Q(\bar{\mathbf{z}}_n)$, where $\Omega_I(\bar{\mathbf{z}}_n)$ and $\Omega_Q(\bar{\mathbf{z}}_n)$ are the estimates of real and

imaginary parts respectively. The cost function is defined as

$$J = \frac{1}{2N_S} \sum_{n=0}^{N_S-1} (\mathrm{Re}\{a_n\} - \Omega_I(\bar{\mathbf{z}}_n))^2 + (\mathrm{Im}\{a_n\} - \Omega_Q(\bar{\mathbf{z}}_n))^2. \tag{4.29}$$

In the proposed network structure, there is a single hidden layer together with one input and output layers. Therefore, network output can be expressed as

$$\Omega_I(\bar{\mathbf{z}}_n) = \mathbf{w}_I^T g(\mathbf{W_1}\bar{\mathbf{z}}_n + \mathbf{b}_1) + b_{2,I}, \tag{4.30}$$

where $\mathbf{w}_I \in \mathbb{C}^{L_1 \times 1}$ is the weight vector of the output layer, similarly output for the quadrature part is obtained by the weight vector $\mathbf{w}_Q$ and $b_{2,Q}$. $\mathbf{W_1} \in \mathbb{C}^{L_1 \times 4M-2}$ is the weight matrix for the input layer with elements $[\mathbf{W_1}]_{k,l} = w_{k,l}$ which connects l^{th} neuron of the input layer to k^{th} neuron at the hidden layer, $\mathbf{b}_1$ and $\mathbf{b}_2$ are the bias vectors, L_1 is the number of neurons in hidden layer and $g(\cdot)$ is the activation function. It can be noted from (4.30) that nonlinearity $g(\cdot)$ is applied to the combination of the neighboring received signals as motivated by the nonlinear ISI analysis in Chapter 3. In this study, hyperbolic tangent sigmoid transfer function, $g(x) = \frac{2}{1+e^{-2x}} - 1$, is employed, which introduces the nonlinearity required for nonlinear compensation. At each epoch, cost function is evaluated and weights of NN coefficients are updated via back-propagation by using the Levenberg-Marquardt algorithm [58].

Furthermore, as part of the research effort, the application of a nonlinear post-distortion technique is augmented with the utilization of a developed NN model for the purpose of modeling the nonlinear self-interference in duplex communication systems in [34]. The proposed neural network structure is specifically designed to estimate the slowly changing nonlinear characteristics exhibited by power amplifiers. By accurately estimating the nonlinear distortion leakage, it becomes possible to effectively cancel out the adverse effects introduced to the received signal. The provision of precise distortion estimation greatly enhances the likelihood of successful decoding of the intended signal-of-interest. This integrated approach contributes to the advancement of techniques employed in duplex communication systems, enabling improved signal recovery and overall system performance.

4.5.5 Volterra series based NP

In this chapter, Volterra series (VS) is also adopted for NP. Post-distortion output can be given by VS basis functions of third order [38] as

$$\tilde{a}_n = \sum_{p=-M+1}^{M-1} \phi_p z_{n-p} + \sum_{p=-M+1}^{M-1} \sum_{q=-M+1}^{M-1} \sum_{r=-M+1}^{M-1} \psi_{p,q,r}\ z_{n-p} z_{n-q}^{*} z_{n-r}, \tag{4.31}$$

where ϕ_p and $\psi_{p,q,r}$ are the VS coefficients. To represent identification problem in a more convenient way, post-distortion output is represented in matrix-vector form as, $\tilde{\mathbf{a}} = \mathbf{Z}\boldsymbol{\phi}$, where $[\tilde{\mathbf{a}}]_n = \tilde{a}_n$, $\boldsymbol{\phi} \in \mathbb{C}^{((2M-1)^3+(2M-1))\times 1}$ is the coefficient vector, whose elements are the Volterra series coefficients, $\mathbf{Z} \in \mathbb{C}^{N_S \times ((2M-1)^3+(2M-1))}$ is the data matrix formed by nonlinearly distorted training symbols after standard FDE operation, whose elements are basis functions for VS, $z_{n-p} z_{n-q}^{*} z_{n-r}$. Then, LS method is performed to estimate VS series coefficients as

$$\boldsymbol{\phi} = \left(\mathbf{Z}^H \mathbf{Z}\right)^{-1} \mathbf{Z}^H \mathbf{a}, \tag{4.32}$$

where $[\mathbf{a}]_n = a_n$ is the vector of ST training symbols.

4.5.6 An Adaptive QAM Detection based on nonlinear post-distortion bank

In this section, a detector structure, which is novel in the aspect of taking the distortion amplification into account, is proposed. In order to reduce the effects of distortion amplification, the use of channel diversity is proposed in this study. Therefore, we propose that fusion of the estimates for different sampling instances improves the overall detection performance by taking the distortion powers for different effective channels into account.

In Figure 4.4, NP produces soft symbol estimates, $\tilde{a}_n^{(i)} = \Omega\left(\left\{z_n^{(i)}\right\}\right)$ as the output for each branch, where $\Omega(\cdot)$ denotes the output of nonlinear regression function. Then, soft symbol estimates, $\tilde{\mathbf{a}}_n = [\tilde{a}_n^{(1)}, \tilde{a}_n^{(2)}, ..., \tilde{a}_n^{(\mu_B)}]^T$, are fed into the distortion-aware symbol by symbol detector (DA-SSD), which is based on the generalized Bussgang decomposition for non-Gaussian signals [59]

$$\tilde{\mathbf{a}}_n = \boldsymbol{\beta} a_n + \boldsymbol{\eta}_n, \tag{4.33}$$

where $\boldsymbol{\beta}$ is the Bussgang coefficient vector computed during fast time training period as

$$\boldsymbol{\beta} = \frac{\mathbb{E}\{\tilde{\mathbf{a}}_n a_n^*\}}{\mathbb{E}\{|a_n|^2\}} \approx \left(\sum_{n=0}^{N_F-1} \tilde{\mathbf{a}}_n a_n^*\right)\left(\sum_{n=0}^{N_F-1} |a_n|^2\right)^{-1}, \tag{4.34}$$

and $\mathbf{R}_\eta$ is the autocorrelation matrix of distortion vector, $\boldsymbol{\eta}_n$, which is calculated as

$$\mathbf{R}_\eta = \mathbb{E}\{(\tilde{\mathbf{a}}_n - \boldsymbol{\beta} a_n)(\tilde{\mathbf{a}}_n - \boldsymbol{\beta} a_n)^H\} \approx \frac{1}{N_F}\sum_{n=0}^{N_F-1} (\tilde{\mathbf{a}}_n - \boldsymbol{\beta} a_n)(\tilde{\mathbf{a}}_n - \boldsymbol{\beta} a_n)^H. \tag{4.35}$$

Although the data symbols, a_n, are not Gaussian, it is shown in [59] and [60] that Bussgang decomposition can be employed for modulated signals. This decomposition can be considered as the LMMSE estimate of non-Gaussian signal, $\tilde{\mathbf{a}}_n$, by using desired symbol, a_n. The aim is to decompose the received signal into the desired and the uncorrelated distortion terms.

Soft symbol estimates are calculated by combining branch outputs to minimize the overall distortion power. For this purpose, minimum variance distortionless response (MVDR) method is adopted to find the combiner vector, $\mathbf{w}_c \in \mathbb{C}^{\mu_B \times 1}$,

$$\min \mathbf{w}_c^H \mathbf{R}_\eta \mathbf{w}_c \qquad \text{s.t.} \qquad \mathbf{w}_c^H \boldsymbol{\beta} = 1, \tag{4.36}$$

which gives the combiner vector as

$$\mathbf{w}_c = \frac{\mathbf{R}_\eta^{-1}\boldsymbol{\beta}}{\boldsymbol{\beta}^H \mathbf{R}_\eta^{-1}\boldsymbol{\beta}}. \tag{4.37}$$

Then the soft symbol estimate, $\hat{a}_n$, and the estimate of the distortion variance, $\hat{\sigma}_a^2$, which will be utilized while calculating log-likelihood ratio (LLR) values, become

$$\hat{a}_n = \mathbf{w}_c^H \tilde{\mathbf{a}}_n, \qquad \hat{\sigma}_a^2 = \mathbf{w}_c^H \mathbf{R}_\eta \mathbf{w}_c = \frac{1}{\boldsymbol{\beta}^H \mathbf{R}_\eta^{-1}\boldsymbol{\beta}}. \tag{4.38}$$

By using $\hat{a}_n$ and $\hat{\sigma}_a$, LLR for the k^{th} bit of symbol a_n, $L(b_n^k)$, can be calculated as

$$L(b_n^k) = \log_2\left(\frac{\sum_{a_n \in A_a^{0,k}} \exp\left(-\frac{1}{\hat{\sigma}_a^2}\left[(\mathrm{Re}\{\hat{a}_n\} - \mathrm{Re}\{a_n\})^2 + (\mathrm{Im}\{\hat{a}_n\} - \mathrm{Im}\{a_n\})^2\right]\right)}{\sum_{a_n \in A_a^{1,k}} \exp\left(-\frac{1}{\hat{\sigma}_a^2}\left[(\mathrm{Re}\{\hat{a}_n\} - \mathrm{Re}\{a_n\})^2 + (\mathrm{Im}\{\hat{a}_n\} - \mathrm{Im}\{a_n\})^2\right]\right)}\right), \tag{4.39}$$

where A_a is QAM symbol alphabet, $A_a^{0/1,k}$ is the ideal constellation point with bit $0/1$ at given bit location, k. It should be noted that, LLR values are calculated assuming

a Gaussian distribution, which is a reasonable assumption for two reasons. First, if there is no distortion amplification, nonlinear distortion is compensated by NP. Consequently, thermal noise, which has Gaussian distribution, becomes the dominant term. Second, if distortion amplification occurs, distortion is multiplexed by IDFT; hence, resultant distortion is approximately Gaussian.

4.6 Achievable Information Rate (AIR) based on mismatched decoding capacity

In order to evaluate AIR performance of the proposed receiver, a lower bound on the constrained capacity in terms of GMI metric is considered. In [51,52,61], mismatched decoding capacity is given as

$$C_P = \log_2 P - E_{a,\tilde{\mathbf{a}}}\left[\log_2\left(\frac{\sum_{a'\in A_a}\tilde{p}(\tilde{\mathbf{a}}|a')}{\tilde{p}(\tilde{\mathbf{a}}|a)}\right)\right], \tag{4.40}$$

where P is the modulation order. $\tilde{p}(\tilde{\mathbf{a}}|a')$ is the *mismatched* probability density function (PDF), which is used since exact knowledge on PDF is not available due to nonlinearity. Hence, an approximate exponential PDF, $\tilde{p}(\tilde{\mathbf{a}}|a)$, is assumed based on the decompostion given in (4.33) as

$$\tilde{p}(\tilde{\mathbf{a}}|a) = \frac{1}{(\pi)^{\mu}|\mathbf{R}_{\eta}|}\exp\left[-(\tilde{\mathbf{a}}-\boldsymbol{\beta}a)^{H}\mathbf{R}_{\eta}^{-1}(\tilde{\mathbf{a}}-\boldsymbol{\beta}a)\right], \tag{4.41}$$

where Gaussian distribution is assumed for the unknown PDF. Consequently, a *lower* bound on the AIR is obtained due to mismatch between the actual and assumed PDF's. AIR, which is obtained by using the exact PDF, is shown to be greater than the mismatched decoding capacity given in (4.40) [51]. *Outage probability* is another performance metric that can be used to measure the performances of the detectors in fading channels. Outage probability, P_{out} can be defined as the probability of the instantaneous AIR being less then a threshold $C_P^T : P_{out} = \Pr\{C_P < C_P^T\}$. It can also be inferred as the packet error rate since a significant amount of block errors occur if instantaneous AIR of the system falls below the defined threshold.

Table 4.1: Complexities of Baseline Schemes

	Channel Est.	FDE	NP	LLR Calc.
IB-DFE [49]	$\mathcal{O}(LN_F/N_D)$	$\mathcal{O}(N_{it}log_2(N_D))$	$\mathcal{O}(N_{it}((2M-1)L_1))$	$\mathcal{O}(N_{it}log_2(P)2P)$
MM [41]	$\mathcal{O}(LN_F/N_D)$	$\mathcal{O}(log_2(N_D))$	-	$\mathcal{O}(log_2(P)6P)$
DA-SSD (NN)			$\mathcal{O}(\mu_B((2M-1)L_1))$	
DA-SSD (GPR)	$\mathcal{O}(\mu_B LN_F/N_D)$	$\mathcal{O}(\mu_B log_2(N_D))$	$\mathcal{O}(\mu_B N_S)$	$\mathcal{O}(log_2(P)2P)$
DA-SSD (VS)			$\mathcal{O}(\mu_B(2M-1)^3)$	

4.7 Complexity Analysis

In this section, computational complexity of the proposed detector is evaluated and compared with the baseline techniques in terms of number of FLOP's per symbol in Table 4.1. Technique proposed in [49] has the same complexity as conventional IB-DFE with additional complexity of applying the transmitter chain. In [49], the same PA for transmitter is used at the receiver; however, in practice a PA model should be used to perform nonlinear distortion cancellation. For comparison, we included complexity of ARVTDNN structure in Table 4.1. In addition, FDE and NP operations are repeated for N_{it} iterations. Major complexity of MM receiver comes from LLR calculation especially for higher order modulations since matrix vector multiplications are performed for each symbol. Despite the lack of NP unit, complexity of [41] is comparable to NP based methods. Complexity of the proposed architecture increases linearly with the number of FDE branches. Each branch performs the conventional feedforward equalization together with NP for each symbol. Complexities of the proposed DA-SSD and IB-DFE in [49] schemes are dominated by the implementation of NP; however, implementation of NP is necessary to improve the performance significantly. Also note that the receiver in [49], does not require the FDE bank structure since it is developed for symbol sampled domain, where distortion amplification is not the issue. Therefore, the proposed DA+SSD can be employed together with the IB-DFE structure to improve the performance of the systems for realistic fractionally delayed channels.

4.8 Numerical Evaluations

Numerical results are presented to evaluate the performance of the proposed receiver structure. The proposed NP and DA-SSD techniques are compared with the state-of-the-art methods. Two different cases are considered for performance evaluations.

In the first case, the Saleh model is used as the memoryless nonlinear PA. Both AWGN channel and a clustered mmWave channel model containing 3 rayed 6 clusters with a line-of-sight component with 10 dB Rician factor [3] are considered. In simulations, the channel length is chosen to cover 16 symbols with unity gain. For this model, input signal power is scaled in such a way that the normalized output power, which is the ratio of average output power to the maximum output power of the amplifier, refers to output backoff (OBO). In the second scenario, a model extracted from an actual GaN PA, [1], and another PA, available at [2], are utilized. To test the performance of the proposed architecture, the transmitted signal passes through a dispersive channel generated according to either a clustered mmWave channel model or a Rayleigh distributed COST-207 channel model [62]. Also, the constellation order is selected as 256 QAM.

In data transmission, the block length is selected to be $N_D = 8100$, and the FT training sequence length is $N_F = 3000$ symbols. A *root raised cosine* filter with 0.3 roll-off factor and $\mu = 8$ is used as p_n. For nonlinear parameter learning, an ST training sequence of length $N_S = 16384$ is used. To utilize the proposed BLUE combiner for GPR estimates, the training sequence is divided into 8 segments. ARVTDNN structure with 30 neurons is employed. Unless otherwise stated, memory depth for GPR, NN, and VS-based NPs is chosen as ± 2 samples.

PA Model description: In simulations, model parameters for the Saleh model are chosen as in [41], $g_0 = 2$, $A_{sat} = 1$, $\alpha = 2$ and $\beta = 1$. Memory polynomial coefficients for the GaN PA, given in [1], are derived from the measurements of a PA by using LS fitting. The PA is operated at $2\,GHz$ with $200\,MHz$ sampling rate and $40\,MHz$ signal bandwidth. Model parameters are $K_b = 5$, $P_b = 4$, and $P_c = 1$. In addition, GaN PA [2] is also modeled by using ARVTDNN structure to speed up the simulations. In this model, upsampled signal samples at the output of the actual PA are utilized to

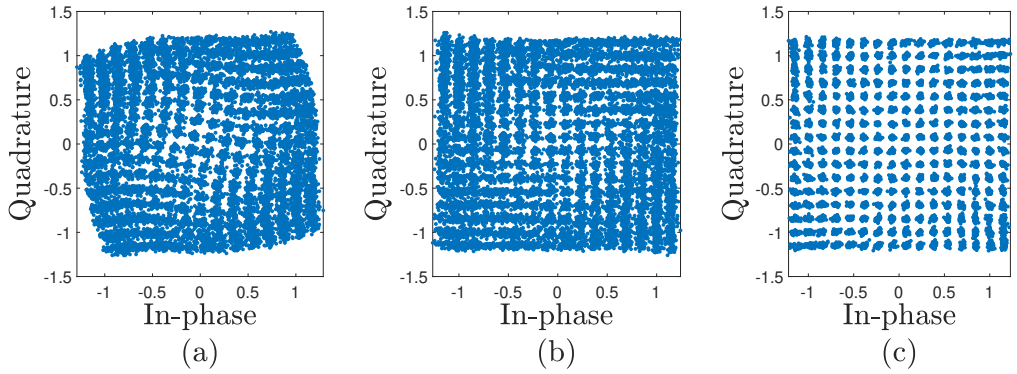


Figure 4.7: Scatterplots for different receivers (a) conventional (b) MM (c) NP with NN.

extract the model. Memory length of the model is chosen as, ± 5, and the number of neurons is 50. Normalized modelling error is $\sim -50\ dB$ in the MSE sense.

Nonlinear Post-distortion Evaluation: Before investigating the performances of the receiver structures in detail, received symbols are shown in Figure 4.7 for Saleh model with $5\ dB$ OBO. In Figure 4.7(a), the warping effect due to nonlinear behavior can be observed since the conventional receiver does not attempt to compensate for the nonlinearity. Received symbols after memoryless correction is shown in Figure 4.7(b). However, there is a strong nonlinear ISI remaining. NP methods compensate for nonlinear ISI together with the warping effect as can be seen in Figure 4.7(c).

4.8.1 Performance evaluations with Saleh model PA

4.8.1.1 AWGN channel

Firstly, AIR analysis and BER performance evaluations are carried out for different OBO values for 1024 QAM. For this scenario, the bank of FDE is not utilized since the AWGN channel with $E_s/N_o = 50$ dB is considered. To observe the effects of nonlinear ISI, memoryless NP with memory depth $M = 1$ is also considered. In Figure 4.8(a), AIRs obtained by GMI analysis are shown for different receiver structures. It can be observed that even for lower OBOs, higher AIR, ~ 9.5 bps/Hz, is

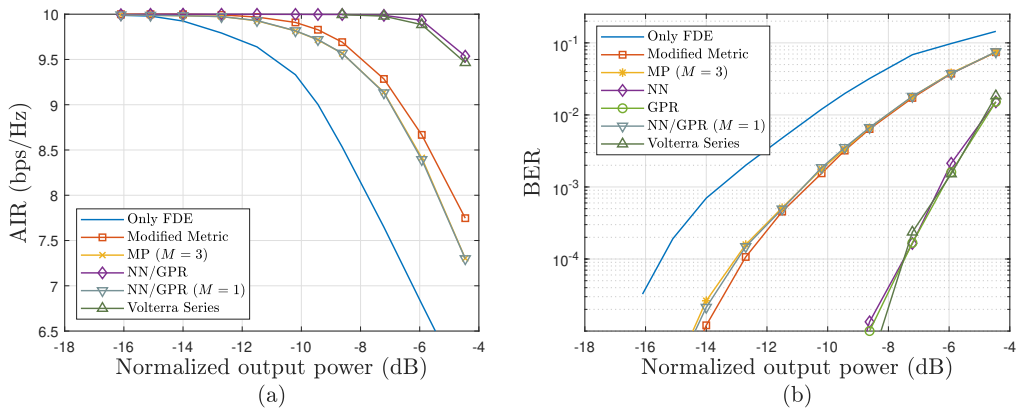


Figure 4.8: (a) AIR and (b) BER of the receivers for Saleh Model for different OBOs for fixed $E_s/N_o = 50\ dB$.

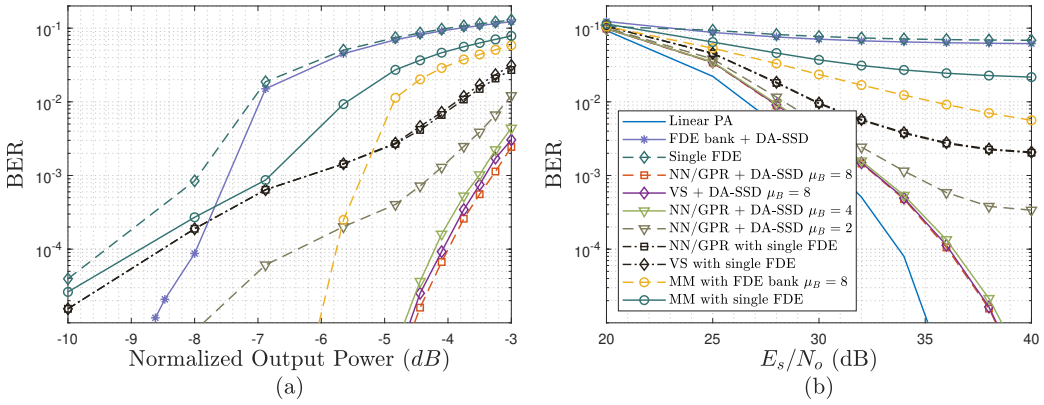


Figure 4.9: (a) BER vs. OBO (b) BER vs. E_s/N_o curves for Saleh Model and mmWave channel with code rate $r = 8/9$.

achieved by the NPs compared to the conventional and MM-based receivers. Lastly, the need for memory in detection is obvious since memoryless NP yields the same performance as MM, which is in compliance with the nonlinear ISI analysis given in Chapter 3. From Figure 4.8(b), it can be seen that the proposed NN, GPR, and VS-based NPs significantly outperform the other methods in terms of uncoded BER performances. Besides, the performance of the NP based on memory polynomial (MP) is investigated. However, MP does not provide any improvement since NP requires cross-terms as in NN, GPR, and VS, which is also in compliance with the nonlinear ISI analysis in Chapter 3.

4.8.1.2 Clustered mmWave channel

In this section, clustered mmWave channel is considered for 256 QAM. For this scenario, a low-density parity check (LDPC) encoder is used with code rate, $r = 8/9$. In Figure 4.9(a), BER vs. OBO performances of receivers are shown for fixed $E_s/N_o = 40\ dB$. In addition, the performance of the proposed DA-SSD is also evaluated by using the different number of FDE branches. Besides, for the MM detector, an FDE bank is also utilized and the branch, which yields the minimum distortion variance during FT training, is chosen to decode the sequence. It can be observed that at lower OBO values, nonlinear distortion is dominant; however, as OBO increases post-distortion algorithms compensate for nonlinearity. However, even for lower OBOs, the proposed DA-SSD improves the performance of the receivers. As OBO increases, the need for NP decreases since forward error correction (FEC) is sufficient to decode the desired symbols. However, the need of employing FDE bank is permanent since distortion amplification still occurs, which cannot be handled by FEC. For instance, the FDE bank+DA-SSD receiver outperforms the receiver that employs NP with a single FDE branch for OBO greater than $\sim 8\ dB$. Besides, DA-SSD with 2 branches provides improvement compared to NP with a single branch. As can be noted from Figure 4.9(a), the performances of the receivers with 4 and 8 branches are almost identical.

The robustness of the receivers to AWGN is also investigated for the same scenario with $5\ dB$ OBO in Figure 4.9(b). The figure demonstrates that the proposed NP structure is robust to additive Gaussian noise at the receiver side. There is $\sim 4\ dB$ difference in required E_s/N_o to achieve BER of 10^{-5} between the systems with ideal linear PA and the proposed NP+DA-SSD architecture. In addition, methods that do not employ DA+SSD or NP suffer from significant error floor.

In addition, a lower code rate, $r = 2/3$, is also employed for 3 dB OBO. In Figure 4.10, it can be observed that the effects of distortion amplification dominate the system performance since significant error floor cannot be eliminated even with higher FEC capabilities for systems with a single FDE branch. Also, the BER performance of the MM detector is improved by using a lower code rate. However, its performance improvement is limited since nonlinear distortion is also effective for such low OBO

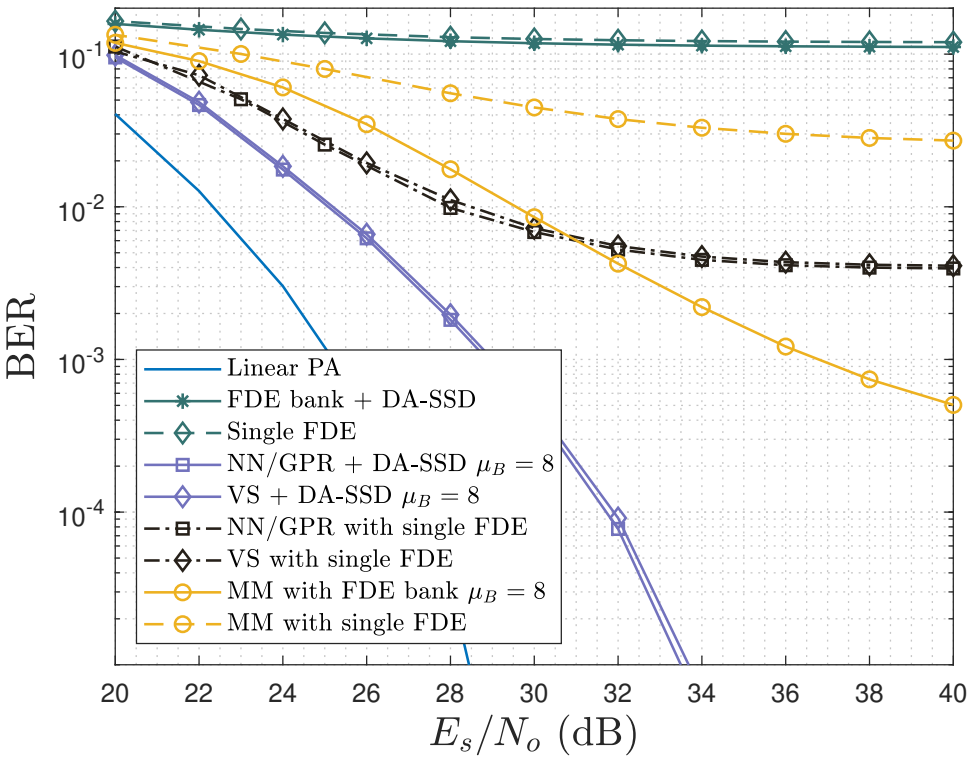


Figure 4.10: BER vs. E_s/N_o curves for Saleh Model with 3 dB OBO and mmWave channel with code rate $r = 2/3$.

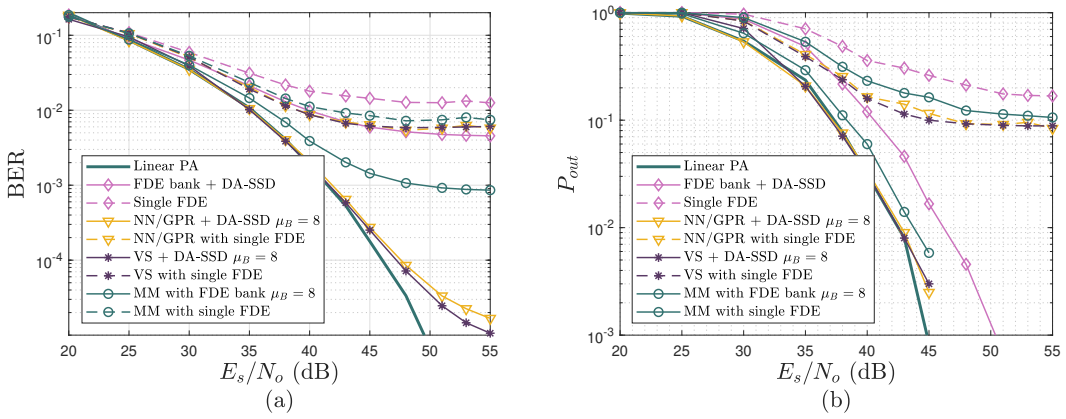


Figure 4.11: (a) Uncoded BER vs. E_s/N_o (b) P_{out} vs. E_s/N_o curves for GaN Model [1] .

levels and NP with memory is needed to mitigate the nonlinear ISI. Otherwise, there is significant error floor since residual nonlinear distortion significantly degrades the detection performance.

4.8.2 Performance evaluations for GaN PA model [1] for Rayleigh Channel

In this section, the performances of receivers are evaluated for the PA model given in [1]. In Figure 4.11(a), uncoded BER performances of the receivers are evaluated. Detectors that rely on a single FDE suffer from significant error floor. However, employing a fractionally delayed FDE bank with DA-SSD decreases the error floor levels substantially. In addition, the performance of memoryless MM-based detector is far from those of NP-based detectors.

To test the block error performances of the detectors, P_{out} performances are presented in Figure 4.11(b). The threshold for the AIR is chosen as $C_P^T = 7.5\ bps/Hz$. It is apparent that receivers with a single FDE branch experience outages very often; hence, it can be concluded that observed performance degradation stems from fluctuations in the AIR of the system due to channel variations. When the AIR falls below the threshold, block errors occur, and overall BER increases. However, NP+DA-SSD reduces P_{out} by giving more weight to the branch having less distortion.

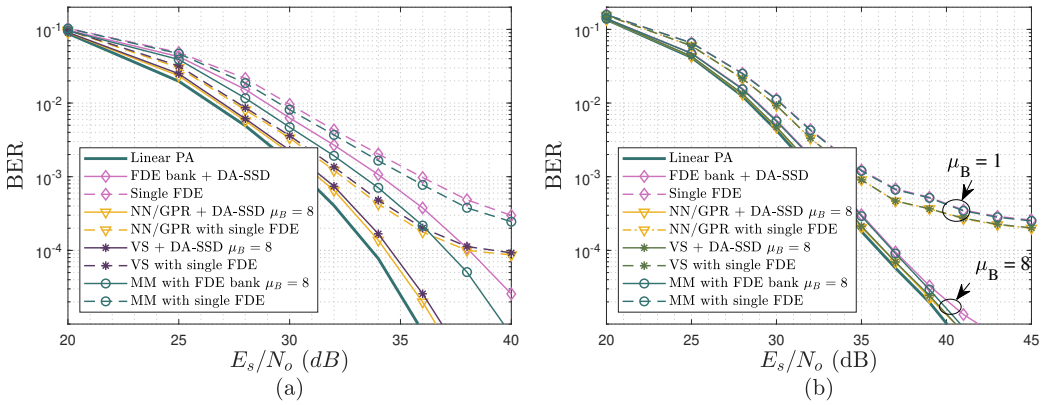


Figure 4.12: BER vs. E_s/N_o for GaN PA [2] for (a) mmWave with code rate $r = 9/10$ (b) Rayleigh channels with code rate $r = 2/3$.

4.8.3 PA Model based on real measurements [2]

In this section, BER performances of the receivers are compared for an actual hardware [2].

4.8.3.1 Clustered mmWave Channel

BER vs. E_s/N_o results are presented in Figure 4.12(a), where LDPC encoder is used with rate $r = 9/10$. It can be observed that performances of NP+DA-SSD methods are very close to that of linear PA. NP methods with single branch suffer from an error floor around BER of 10^{-4}. However, employing DA-SSD without any NP or MM with FDE bank eliminates the floor but required E_s/N_o increases.

4.8.3.2 COST-207 Rayleigh Channel

In Figure 4.12(b), BER results are demonstrated for Rayleigh channel with code rate $r = 2/3$. It is observed that BER performances of the all techniques with DA-SSD detector is close to that of the ideal system. Even without any NP method, DA+SSD performs close to ideal since FEC can handle errors due to PA nonlinearities. However, receivers with single branch suffer from error floor since distortion amplification is dominant and FEC cannot handle outage due to distortion amplification.

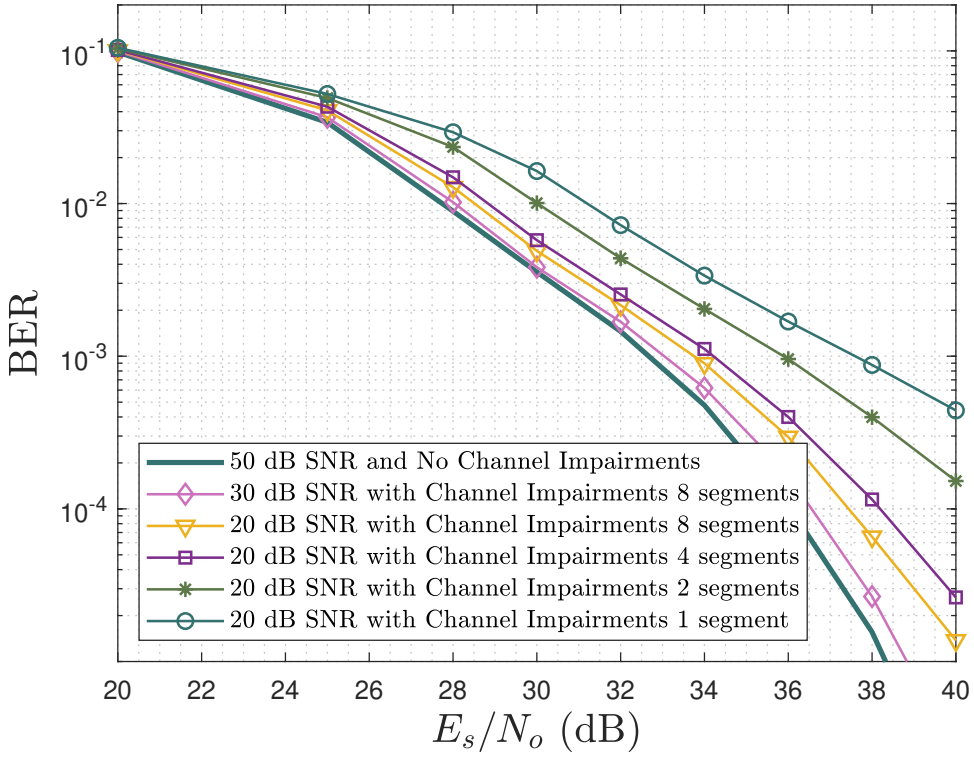


Figure 4.13: Performance of the NN based Nonlinear post-distortion for NPL with different SNR with channel impairments.

Consequently, it can be concluded that FEC can handle the errors stemming from the nonlinear distortion if there is no distortion amplification. However, distortion amplification would have dramatic effects depending on the channel characteristics as can be observed from Rician and Rayleigh fading case, where the system severely suffers from outage as in Figure 4.11(b). In the presence of distortion amplification, FEC cannot handle the burst errors. As a result, the need of NP decreases with the help of FEC; however, performance improvement is permanent in the case of channel-dependent distortion amplification thanks to DA-SSD.

4.8.4 Nonlinear parameter learning under channel impairments

The simulation experiments were conducted to evaluate the performance of the system under various SNR levels and with different numbers of segments, as depicted in Figure 4.13. The figure presents the BER curves for the Saleh Model with a 5 dB

OBO and a mm-Wave channel employing a code rate of 8/9. Specifically, we investigated the impact of practical NPL on the performance of the proposed NN+DA-SSD, considering a parameter $\mu_B = 8$.

In our simulation setup, each training segment for the spread spectrum detector consists of 2048 symbols. We observed that the proposed learning procedure mitigates the performance degradation caused by channel impairments. Moreover, the observed degradation was found to be negligible for practical SNR values, indicating the feasibility and effectiveness of our proposed learning approach.

These findings provide valuable insights into the performance characteristics of the NN+DA-SSD system, particularly considering different SNR levels, the number of segments, and the presence of NPL in the mm-Wave channel. The results highlight the robustness of the proposed learning methodology, which demonstrates promising performance even in the presence of realistic channel impairments.

4.8.5 The effects of the length of the FT sequence

Lastly, the effects of the length of the FT training sequence and channel estimation accuracy are studied. For this purpose, BER vs. training length (for given SNR values required to achieve BER of 10^{-5} are shown in Figure 4.14. It can be seen that the channel estimation performance of the proposed approach is very close to that of linear PA, which also requires $N_F = 1500$ training symbols.

4.9 Conclusions

In this chapter, we addressed the effects of PA nonlinearities impaired by memory. It is shown that even if a memoryless nonlinearity exists, nonlinear ISI occurs since the Nyquist-1 criterion is violated. Therefore, we developed a post-distortion scheme to reduce nonlinear distortion power by taking the nonlinear ISI into account. The proposed scheme divides the problem into two independent parts. In the first part, a standard FDE procedure is employed, where the linear wireless channel is equalized so that the dimension of the problem is reduced. Then, it is followed by nonlinear

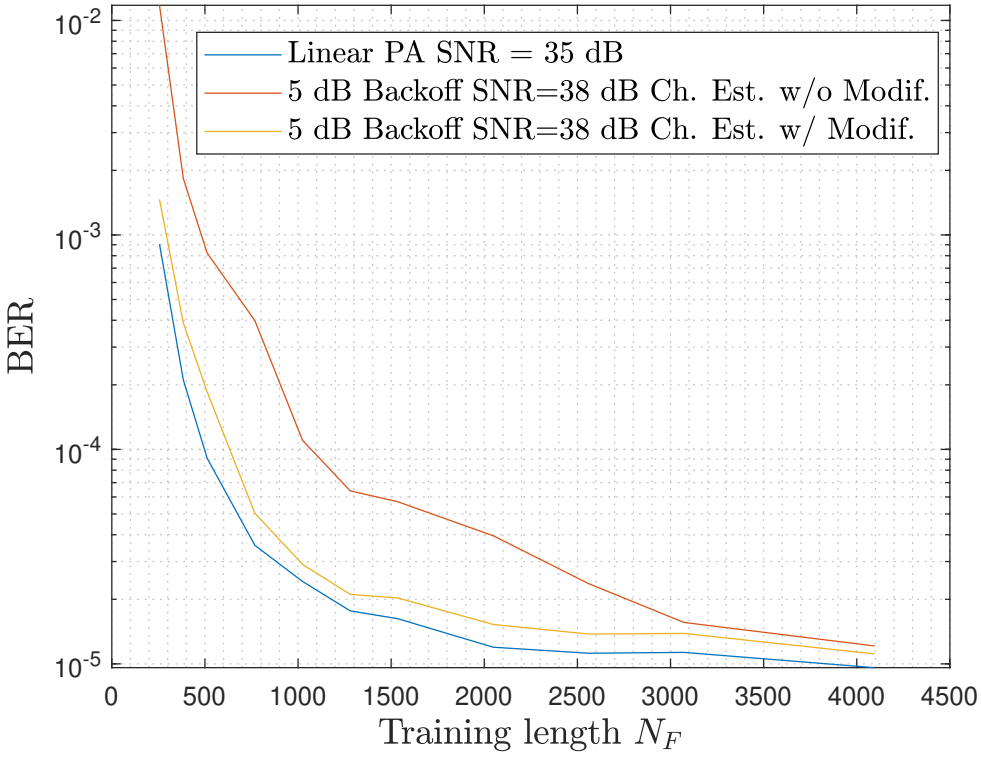


Figure 4.14: Performance evaluations for different training length BER vs. N_F.

processing, which suppresses the nonlinear ISI. In addition, the effects of dispersive communication medium are considered. For this case, it is observed that FDE performance is significantly affected by the nonlinear PA due to distortion amplification. To overcome this problem, we propose to utilize samples obtained by a fractional delayed sampling so that channel diversity is obtained. By taking the fading effects of each timing offset into account, DA-SSD based on FDE bank is proposed. Performance improvement provided by the proposed structure is shown via numerical simulations.

CHAPTER 5

NONLINEAR DETECTION FOR MULTIUSER HYBRID MIMO SYSTEMS AND PERFORMANCE ANALYSIS

5.1 Literature Review

In the literature, post-distortion or nonlinear distortion cancellation techniques at the receiver side have been deployed to mitigate the nonlinear distortion [33,40,44,45,48, 63–68]. In [40], Volterra series (VS) based post-distortion is employed as a nonlinear equalizer to detect the transmitted symbols from the nonlinearly modified received signals. Machine learning algorithms have demonstrated significant efficacy in performing nonlinear signal processing, as evidenced by various studies [45, 67, 68]. For instance, a method for digital signal recovery from the nonlinearly distorted received signal via a deep NN (DNN) structure is proposed in [67] for satellite channels. In addition, the neural network (NN) structure is further employed for joint equalization of both the linear wireless channel and the nonlinearity as proposed in the studies [45, 68]. In these studies, nonlinear equalization is performed to mitigate the combined effects of frequency selective linear wireless channel and PA nonlinearity. However, joint linear and nonlinear equalization brings an additional computational burden to the system due to the requirement of frequent nonlinear parameter learning during communication.

Methods that decouple linear and nonlinear channels are proposed in order to reduce the computational complexity [33, 44, 48, 63–66]. In mobile environments, the linear communication channel exhibits significantly faster variations compared to the nonlinear behavior of the hardware. Thus, it is practical to continuously estimate the linear time-varying channel to account for the mobility-induced changes rather than fre-

quently performing nonlinear parameter learning (NPL). Consequently, the required processing time for NPL is reduced, and the algorithms, which decouple linear and nonlinear parameter learning, may improve the system performance in terms of age of information metric [69, 70]. Furthermore, two-stage detection schemes has gained popularity for detection purposes in NOMA schemes to improve detection accuracy and covarage [71]. In [63] and [64], a receiver, which employs nonlinear distortion cancellation based on tentative decisions following FDE, is proposed for OFDM modulation. Also, as the interest in SC-FDE schemes increases, receiver design under hardware impairments has attracted significant attention [33, 44, 48, 65, 66]. In these studies, the linear wireless channel is equalized at the first stage, and the nonlinear processing follows it to reduce the effects of nonlinear distortion. In [48], a *robust receiver* is introduced for analog-to-digital converter (ADC) nonlinearities. In [65], a nonlinear distortion cancellation scheme based on tentative decisions is adopted to SC modulation with iterative block decision feedback equalization (IB-DFE) receiver for MIMO systems. Likewise, [44] presents a method based on a nonlinear hybrid decision feedback equalizer for single-input single-output (SISO) systems. This method equalizes the linear channel through a decision feedback equalizer and mitigates nonlinearity through the inverse of the power amplifier (PA) nonlinearity. Based on [44], the authors extend their work to fully digital MIMO scheme to perform multi-user detection in [66].

In the aforementioned studies [44, 63–66], the effect of fractionally delayed channel response on nonlinear distortion after downsampling is ignored by simply assuming a symbol-rate channel model. However, this assumption is shown to be inaccurate, and leads to a remarkable performance degradation depending on the mismatch between the channels of linear and nonlinear terms of the transmitted signal after downsampling, and this mismatch causes *distortion amplification* after performing minimum mean square error (MMSE) filtering for realistic upsampled channel models [33]. In order to overcome the *distortion amplification* problem and exploit channel diversity in SISO systems, a non-iterative receiver structure, which is based on a bank of fractionally delayed linear FDEs, is proposed, and a nonlinear post-distortion algorithm is developed to mitigate the nonlinear distortion for single user channel [33]. Moreover, in [72], nonlinear post-distortion is deployed for hybrid MIMO systems, where users

of different groups are differentiated in the spatial domain by using a sophisticated analog beamformer.

Implementing the bank of FDEs in [33] is not straightforward for MIMO architecture, which employs MMSE multi-user detection as in [48, 65, 66] since there are many active parallel operating radio frequency (RF) chains. Therefore, in this study, we focus on converting the multi-user MIMO detection into the multiple single-user detection problems by exploiting spatial correlation of the user channels. In [72], it is achieved by using generalized eigen-beamformer (GEB) [73] to form orthogonal user beams so that inter-user-interference (IUI) is suppressed almost perfectly. However, in practice, to reduce the hardware complexity, non-orthogonal beams are created, and IUI becomes a significant problem [74]. In this study, in order to keep the hardware complexity low, a simple phase-only analog beamforming scheme is utilized. Thus, residual IUI affects the MIMO system performance greatly, especially for higher-order constellations, even if the users are not closely spaced in the angular domain. Thus, the primary motivation in this study is to propose a reduced complexity and efficient receiver architecture for nonlinear multi-user uplink SC-FDE based MIMO channels, when practical BDMA-based analog beamforming is employed.

5.1.1 Contributions

Based on the aforementioned discussion, the main contributions lie with the proposed *multi-user (MU) receiver structure* and *analytical performance evaluation* for hybrid MIMO-based uplink SC BDMA under PA impairments:

- In this chapter, a novel receiver structure, that utilizes a bank of NP units, which are fed by FDE outputs for different sampling instances, is proposed for fully connected hybrid massive MIMO systems that GEB as the analog beamformer. Two different NP methods, which are based on Volterra series expansion and augmented real-valued time-delay neural network (ARVTDNN), are considered. Employed NP methods also include memory so that nonlinear inter-symbol-interference (ISI) is suppressed. Besides, a decision metric, which combines outputs of different post-distortion units for each user, is proposed to

exploit channel diversity so that distortion amplification is compensated.

- Novel aspects of the proposed MU receiver structure can be explained in terms of distortion modeling, nonlinear interference cancellation (nIC), distortion-aware MU detection, and computational complexity. First, it is shown that a simple single-layer feedforward NN structure can accurately mimic the nonlinear distortion of each user channel, which contains memory emerging from matched filtering and decimation operations in addition to PA nonlinearities. As a result, the NN model is extracted by a practical slow-time learning stage, and it is seen to bring significant nIC capability for improved multi-user detection.

- Second, a novel and special implementation type of nIC, based on the idea of decoupling the *slow* and *fast* varying components of the nonlinear channel, has been proposed. The nonlinear distortion is mitigated for each user with the help of extracted *slow-time* NN-based distortion model and *fast-time* fractionally sampled BDMA channel. Then, beamspace channel equalization is performed at each fractional delay by a per user IB-DFE (PU-IB-DFE) algorithm, which avoids performing joint MMSE detection. Finally, the distortion amplification effect on IUI is also mitigated at the detector stage via proper filtering of fractionally delayed IB-DFE outputs based on the residual interference statistics.

- Aside from the proposed structure, this work includes an analytical performance evaluation, which discloses an important *distortion amplification phenomenon* for nonlinear IUI affecting the performance of MU detection. Moreover, it has been analytically demonstrated that the proposed MU receiver architecture improves the BER performance considerably under PA impairments.

To sum up, in the considered hybrid MIMO framework, both numerical evaluations and analytical methods have been employed to demonstrate the significant reduction in distortion amplification effect on multi-user interference at the detector stage. This reduction is achieved through a proper combination of fractionally delayed beamspace equalizer outputs, which considerably minimizes nonlinear distortion power and enhances multi-user detection capability. These results are compared

to state-of-the-art schemes, and the proposed method has been shown to outperform them.

5.2 A Nonlinear Detector for Uplink SC-FDE mm-Wave Hybrid Massive MIMO under Hardware Impairments

5.2.1 Introduction

In this chapter, we propose a detector structure for fully connected hybrid massive MIMO systems with transceiver impairments at user equipments. In literature, memoryless compensation of nonlinear effects at the receiver side is recently proposed for systems employing frequency domain equalization. However, in these studies, the system model is restricted to be symbol sampled, where pulse shaping and fractional delays in channel impulse response (CIR) are not taken into consideration. However, memory effects of nonlinear distortion become more apparent when the system is considered in the upsampled signal domain. In this chapter, we propose a receiver architecture, where the received signal of each user is nonlinearly modified by the post-distortion unit with memory to recover the transmitted symbols. In addition, based on the proposed receiver structure, a novel detection criterion is introduced in order to compensate the distortion amplification due to fractional delays in the CIR. Lastly, performance improvement brought by the proposed detector is verified by numerical simulations.

5.2.2 System Model

5.2.3 mm-Wave Uplink Channel Model

In this section, uplink transmission, where user terminals are equipped with nonlinear PA's, is considered. In the considered system, U single antenna users are communicating with a BS, which is equipped with K_t antennas and D RF chains, $D < K_t$. In the considered hybrid beamforming structure, user grouping is employed so that spatially closer users are jointly processed, where number of RF chains, which are

assigned to a specific group, is significantly less than the total number of RF chains. Consequently, employing user grouping significantly reduces complexity of the receiver structure. U users are grouped into G groups each having U_g users based on users' spatial statistics. Received signal, $\mathbf{y}_n = [y_n^{(1)} y_n^{(2)}, ..., y_n^{(K_t)}]^{\mathrm{T}}$ at the BS can be expressed as,

$$\mathbf{y}_n = \sum_{l=0}^{L-1} \mathbf{H}_l \tilde{\mathbf{x}}_{n-l} + \boldsymbol{\nu}_n. \tag{5.1}$$

where $\tilde{\mathbf{x}}_n = [(\tilde{\mathbf{x}}_n^{(1)})^{\mathrm{T}}, (\tilde{\mathbf{x}}_n^{(2)})^{\mathrm{T}}, ..., (\tilde{\mathbf{x}}_n^{(G)})^{\mathrm{T}}]^{\mathrm{T}}$ is the transmitted signal vector of the users of G groups, $\tilde{\mathbf{x}}_n^{(g)} = [\tilde{x}_n^{(g_1)}, \tilde{x}_n^{(g_2)}, ..., \tilde{x}_n^{(g_{U_g})}]^{\mathrm{T}}$, $\mathbf{H}_l \in \mathbb{C}^{K_t \times U}$ is the l^{th} multipath component (MPC) of the CIR and $\boldsymbol{\nu}_n$ is the additive zero-mean circularly symmteric white complex Gaussian noise. Each MPC of the channel is assumed to be composed of uncorrelated rays which are locally scattered according to wide sense stationary uncorrelated scattering model [75], where statistics of channel vectors of different users can be given by the channel covariance matrix (CCM),

$$\mathbb{E}\{\mathbf{h}_l^{(g_u)} [\mathbf{h}_{l'}^{(g'_{u'})}]^H\} = \mathbf{R}_l^{(g)} \delta_{gg'} \delta_{uu'} \delta_{ll'}, \tag{5.2}$$

where $\mathbf{R}_l^{(g)}$ is the CCM of l^{th} MPC for the users in group g and $\mathbf{h}_l^{(g_u)}$ is the channel vector for u^{th} user of the g^{th} group and it corresponds to a vector in $\mathbf{H}_l$. In this study, block transmission is considered. In each block, N_D M-QAM symbols, $\{a_k\}_{k=0}^{N_D-1}$ are transmitted by employing SC modulation. Transmitted signal for g_u^{th} user is formed as,

$$x_n^{(g_u)} = \sum_{k=0}^{N_D-1} a_k^{(g_u)} p_{n-\mu k}, \tag{5.3}$$

where p_n is the upsampled pulse shaping filter with upsampling factor μ and $a_k^{(g_u)}$'s are i.i.d. QAM symbols with $\mathbb{E}\{a_k^{(g_u)} (a_{k'}^{(g'_{u'})})^*\} = E_s^{(g_u)} \delta_{kk'} \delta_{gg'} \delta_{uu'}$. In order to realize FDE, cyclic prefix and suffix are appended to each block. Then, $x_n^{(g_u)}$ is fed into the PA, which outputs nonlinearly distorted signal, $\tilde{x}_n^{(g_u)} = \psi_n(\{x_n{}^{(g_u)}\}_{n=0}^{N-1})$, where $\psi_n(\cdot)$ represents nonlinearity. In this study, two different nonlinear PA models, Saleh model [37] and generalized memory polynomials (GMP) [1], are considered in order to define nonlinearity.

5.2.3.1 Hybrid Beamforming Architecture

Hybrid beamforming architecture is employed at BS in order to use a relatively small number of RF chains compared to the number of antennas. Received signal processing is performed in two stages. At the first stage, analog beamforming (BF) performs spatial processing such that spatially uncorrelated users are decoupled. Then, digital beamforming is applied on reduced dimensional signal to carry out ZF-FDE for multi-user detection.

Statistical Beamforming: In the considered system, analog BF is designed based on slowly varying spatial characteristics of users, which is in compliance with JSDM framework. Different beamformers are used for different groups, which projects K_t dimensional input into D_g dimension, where D_g is the number of RF chains dedicated to group g. BF output, $\mathbf{y}_n^{(g)} \in \mathbb{C}^{D_g \times 1}$ for each group can be written as,

$$\mathbf{y}_n^{(g)} = [\mathbf{S}^{(g)}]^H \mathbf{y}_n = \sum_{l=0}^{L-1} \mathbf{H}_{eff,l}^{(g)} \tilde{\mathbf{x}}_{n-l}^{(g)} + \boldsymbol{\xi}_n^{(g)}, \tag{5.4}$$

where $\mathbf{H}_{eff,l}^{(g)} \triangleq [\mathbf{S}^{(g)}]^H \mathbf{H}_l^{(g)}$ is the effective channel matrix in reduced dimension, $\mathbf{H}_l^{(g)} = [\mathbf{h}_l^{(g_1)}, \mathbf{h}_l^{(g_2)}, ..., \mathbf{h}_l^{(g_{U_g})}]$ is the channel matrix of g^{th} group, and

$$\boldsymbol{\xi}_n^{(g)} = \sum_{\substack{g'=1 \\ g' \neq g}}^{G} \sum_{l=0}^{L-1} [\mathbf{S}^{(g)}]^H \mathbf{H}_l^{(g')} \tilde{\mathbf{x}}_{n-l}^{(g')} + [\mathbf{S}^{(g)}]^H \boldsymbol{\nu}_n \tag{5.5}$$

is the interference-plus-noise term.

In digital domain, $\mathbf{H}_{eff,l}^{(g)}$ will be utilized while designing digital beamformer. In this study, angle only generalized eigenspace beamforming (AO-GEB) [76] is employed since MPC's are equalized in frequency domain, where it is not necessary to distinguish different MPC's. Consequently, analog BF matrix for group g is obtained as the D_g dominant eigenvectors as the generalized eigenvalue problem

$$\mathbf{S}^{(g)} = \text{eigs}(\mathbf{R}_{\text{sum}}^{(g)}, \mathbf{R}_{\mathbf{y}}, D_g), \tag{5.6}$$

where $\mathbf{R}_{\text{sum}}^{(g)} \triangleq \sum_{l=0}^{L-1} \mathbf{R}_l^{(g)}$ and $\mathbf{R}_{\mathbf{y}} \triangleq \mathbb{E}\{\mathbf{y}_n \mathbf{y}_n^H\} = \sum_{g=1}^{G} \left(\sum_{u=1}^{U_g} E_s^{(g_u)} \right) \sum_{l=0}^{L-1} \mathbf{R}_l^{(g)} + N_o \mathbf{I}$. This can be considered as the generalization of Capon beamformer to general rank signal models [76]. Note that, fully connected analog beamforming architecture is utilized, where each RF chain is the combination of all antenna outputs.

5.2.4 Synchronization, Channel Estimation and Digital Beamforming

5.2.4.1 Frame Structure and Decoder Architecture

In this study, the same frame structure, proposed in [33], is also employed, where two-stage training is performed. Transmission starts with sending a long training sequence of length N_S in order to perform nonlinear parameter learning for NP. Training data is obtained by capturing PA outputs for the training sequence and it is transmitted to BS. This stage is referred as *slow time* training since it is only sent when nonlinear characteristics of the PA change, which is a slow process compared to signaling interval. Data transmission starts after completing the slow-time training. A training sequence of length N_F, which is also called *fast time* training, is transmitted at the beginning of each frame so that instantaneous channel state information (CSI) is acquired by using the least squares (LS) estimator.

In addition, it is shown in [33] that equivalent channels in the symbol sampled domain for the desired signal and distortion terms are not the same in the case of fractionally delayed CIRs. For some channels, the linear signal term may experience deep fading; however, the equivalent channel frequency response for the distortion term may not have deep nulls. In order to equalize the desired term experiencing deep fading, a significant gain is applied to both desired signal and distortion terms, which causes significant distortion amplification. Consequently, overall received distortion power increases tremendously. To compensate effects of distortion amplification, fractionally spaced samples are used to achieve channel diversity. In this study, bank of FDE structure, which is developed in [33] for single-input-single-output (SISO) channels, is employed in mm-Wave MIMO channels. In this structure, signals are obtained with different sampling offsets for branches and FDE is performed for each branch individually. For each FDE output, NP is performed to compensate for nonlinear effects at the receiver.

5.2.4.2 MIMO Channel Estimation

Group processing is performed to eliminate intra-group interference via ZF-FDE for each sampling branch, $i = 0, \ldots, \mu - 1$. Consider received samples for i^{th} sampling branch after analog BF for group, g, in vector form, $\mathbf{y}^{(g),i}$

$$\mathbf{y}^{(g),i} \triangleq \text{vec}\left\{ \left[\mathbf{y}_0^{(g),i}, \mathbf{y}_1^{(g),i}, \ldots, \mathbf{y}_{N_F-1}^{(g),i} \right]_{D_g \times N_F} \right\} \tag{5.7}$$

where $\mathbf{y}_n^{(g),i} = \left(\mathbf{y}_n^{(g)} \circledast p_{-n}^* \right)_{n \to n\mu+i}$ is the sampled signal for i^{th} branch, $\circledast$ denotes convolution. $\mathbf{y}^{(g),i} \in \mathbb{C}^{D_g N_F \times 1}$ is the signal vector that contains all training signals on all RF chains. Similarly, effective channel for u^{th} user of group g can be defined in vector form as

$$\mathbf{h}^{(g_u),i} \triangleq \text{vec}\left\{ \left[\mathbf{h}_{-L_b+1}^{(g_u),i}, ..., \mathbf{h}_0^{(g_u),i}, ..., \mathbf{h}_{L_f-1}^{(g_u),i} \right] \right\}, \tag{5.8}$$

note that anti-causal components are also estimated in order to take synchronization errors into consideration. By using channel vectors of users in the group g, the overall channel vector for that group can be constructed as

$$\mathbf{h}^{(g),i} \triangleq \text{vec}\left\{ [\mathbf{h}^{(g_1),i}, \mathbf{h}^{(g_2),i}, ..., \mathbf{h}^{(g_{U_g}),i}]. \right\}. \tag{5.9}$$

Then, by using channel representation given in (5.9), received signal, $\mathbf{y}^{(g),i}$, can be expressed as

$$\mathbf{y}^{(g),i} = \left[[\underbrace{\mathbf{A}^{(g_1)} \mathbf{A}^{(g_2)} \cdots \mathbf{A}^{(g_{U_g})}}_{\triangleq \mathbf{A}^{(g)}}] \otimes I_{D_g} \right] \mathbf{h}^{(g),i} + \boldsymbol{\xi}^{(g),i}, \tag{5.10}$$

where $\mathbf{A}^{(g_u)} \in \mathbb{C}^{N_F \times (L_b + L_f - 1)}$ is the convolution matrix for the fast time training sequence of user g_u, whose elements are $[\mathbf{A}^{(g_u)}]_{k,l} = a_{(k-l)+L_b-1}^{(g_u)}$, and $\boldsymbol{\xi}^{(g),i} \triangleq \text{vec}\{[\boldsymbol{\xi}_0^{(g),i} \boldsymbol{\xi}_1^{(g),i} \cdots \boldsymbol{\xi}_{N_F-1}^{(g),i}\}]$, $\boldsymbol{\xi}_n^{(g),i} = \left(\boldsymbol{\xi}_n^{(g)} \circledast p_{-n}^* \right)_{n \to n\mu+i}$, and $\otimes$ is the Kronecker product operator. Least squares method can be used to obtained channel estimate, $\hat{\mathbf{h}}^{(g),i}$, as

$$\hat{\mathbf{h}}^{(g),i} = \left[\left((\mathbf{A}^{(g)})^H \mathbf{A}^{(g)} \right)^{-1} (\mathbf{A}^{(g)})^H \otimes I_{D_g} \right] \mathbf{y}^{(g),i} \tag{5.11}$$

By using estimated channel coefficients obtained in vector form, effective channel matrix of the l^{th} MPC of g^{th} group is formed as, $\hat{\mathbf{H}}_{eff,l}^{(g),i} = \left[\hat{\mathbf{h}}_l^{(g_1),i}, \hat{\mathbf{h}}_l^{(g_2),i}, ..., \hat{\mathbf{h}}_l^{(g_{U_g}),i} \right]$.

5.2.4.3 Digital Pre-processing: Linear Beamspace FDE

In order to detect symbols of each user, regularized ZF-FDE is employed for each branch, i, where digital beamforming matrix is defined in frequency domain as,

$$\mathbf{W}_k^{(g),i} = \left([\mathbf{\Omega}_k^{(g),i}]^H \mathbf{\Omega}_k^{(g),i} + \mathbf{N_o} \right)^{-1} [\mathbf{\Omega}_k^{(g),i}]^H, \tag{5.12}$$

where $\mathbf{\Omega}_k^{(g),i} = \frac{1}{\sqrt{N_D}} \sum_{l=-L_b+1}^{L_f-1} \hat{\mathbf{H}}_{eff,l}^{(g),i} e^{-j2\pi kl/N_D}$, is the estimated frequency response, $\mathbf{N_o} = \mathrm{diag}\{\frac{N_o}{E_s^{(gu)}}\}_{u=1}^{U_g}$, and N_o is the noise variance. Intra-group users are discriminated in the frequency domain via FDE; hence, the received signal vector of group, g is transformed into the frequency domain as,

$$\mathbf{y}_{f,k}^{(g),i} = \frac{1}{\sqrt{N_D}} \sum_{n=0}^{N_D-1} \mathbf{y}_n^{(g),i} e^{-j2\pi kn/N_D}, \tag{5.13}$$

hence, equalized signal in frequency domain is obtained as, $\mathbf{z}_{f,k}^{(g),i} = \mathbf{W}_k^{(g),i} \mathbf{y}_{f,k}^{(g),i}$, which is transformed into time domain,

$$\mathbf{z}_n^{(g),i} = \frac{1}{\sqrt{N_D}} \sum_{k=0}^{N_D-1} \mathbf{z}_{f,k}^{(g),i} e^{j2\pi kn/N_D}, \tag{5.14}$$

where $\mathbf{z}_n^{(g),i} = [z_n^{(g_1),i}, z_n^{(g_2),i}, ..., z_n^{(g_{U_g}),i}]^{\mathrm{T}}$ is the equalized symbol vector of users of the group g, which is only corrupted by nonlinear PA. It is assumed that all linear channel effects are compensated via ZF-FDE operation and the remaining nonlinear distortion will be compensated by the following NP unit.

5.2.5 Distortion Aware Symbol Detection via Nonlinear Post-Distortion

In this section, a distortion-aware detector is developed, which performs detection by using the NP outputs of each branch. The proposed detector combines NP outputs of each branch intelligently by using distortion covariance information acquired during fast-time training. Two approaches, based on Volterra series and neural network, are considered to design NP units.

5.2.5.1 Digital Post-Processing: Nonlinear Post-distortion

The objective of the nonlinear post-distortion is to modify ZF-FDE output nonlinearly so that desired symbol is obtained. For notational simplicity, group and bank indices,

(g) and i superscripts are ignored.

5.2.5.2 Volterra Series based NP

Post-distortion output can be given by Volterra series basis functions of third order as

$$\tilde{a}_n = \sum_{p=-L_v+1}^{L_v-1} \phi_p z_{n-p} + \sum_{p=-L_v+1}^{L_v-1} \sum_{q=-L_v+1}^{L_v-1} \sum_{r=-L_v+1}^{L_v-1} \psi_{p,q,r}\ z_{n-p} z_{n-q}^{*} z_{n-r}, \tag{5.15}$$

where ϕ_p and $\psi_{p,q,r}$ are the Volterra series coefficients. To represent the identification problem in a more convenient way, the post-distortion output is represented in matrix-vector form as, $\tilde{\mathbf{a}} = \mathbf{Z}\boldsymbol{\phi}$, where $[\tilde{\mathbf{a}}]_n = \tilde{a}_n$, $\boldsymbol{\phi} \in \mathbb{C}^{((2L_v-1)^3+(2L_v-1))\times 1}$ is the coefficient vector, $\mathbf{Z} \in \mathbb{C}^{N_S\times((2L_v-1)^3+(2L_v-1))}$ is the data matrix formed by nonlinearly distorted training symbols after standard FDE operation, whose elements are, basis functions for Volterra series, $z_{n-p}z_{n-q}^{*}z_{n-r}$. Then, the LS method is performed to estimate Volterra series coefficients as

$$\boldsymbol{\phi} = \left(\mathbf{Z}^H\mathbf{Z}\right)^{-1}\mathbf{Z}^H\mathbf{a}, \tag{5.16}$$

where $[\mathbf{a}]_n = a_n$ is the vector of training symbols.

5.2.5.3 Neural Network based NP

In addition to Volterra series based post-distortion, ARVTDNN structure, which is originally proposed for DPD design in [29], is used as the NP. In this study, ARVTDNN structure, shown in Figure 4.6, is used to predict transmitted symbols by using non-linearly distorted received signals, z_n. Predicted symbol, $\tilde{a}_n$, can be expressed as, $\tilde{a}_n = \Omega_I(\bar{\mathbf{z}}_n) + j\Omega_Q(\bar{\mathbf{z}}_n)$, where $\Omega_I(\bar{\mathbf{z}}_n)$ and $\Omega_Q(\bar{\mathbf{z}}_n)$ are the estimates of real and imaginary parts. $\bar{\mathbf{z}}$ is the input vector, which is defined as

$$\begin{aligned}\bar{\mathbf{z}}_n = [&Re\{z_{n+P-1}\}, ..., Re\{z_n\}, ..., Re\{z_{n-P+1}\},\\ &Im\{z_{n+P-1}\}, ..., Im\{z_n\}, ..., Im\{z_{n-P+1}\}]^{\mathrm{T}}.\end{aligned} \tag{5.17}$$

where P is the memory depth. Nonlinear function, $\Omega_I(\bar{\mathbf{z}}_n)$, which is used to predict the in-phase part can be expressed as

$$\Omega_I(\bar{\mathbf{z}}_n) = \mathbf{w}_I^{\mathrm{T}} g(\mathbf{W_1}\bar{\mathbf{z}}_n + \mathbf{b}_1) + b_{2,I}, \tag{5.18}$$

where $\mathbf{W_1}$ is the input layer weight matrix with elements $[\mathbf{W_1}]_{k,l} = w_{k,l}$. $\mathbf{b}_1$ and $\mathbf{b}_2 = [b_{2,I}, b_{2,Q}]^{\mathrm{T}}$ are the bias vectors and $g(\cdot)$ is the activation function. $\mathbf{w}_I$ is the coefficient vector, which gives in-phase part of the output signal, similarly $\mathbf{w}_Q$ and $b_{2,Q}$ are used to obtain quadrature part of the output. Hyperbolic tangent sigmoid transfer function, $g(x) = \frac{2}{1+e^{-2x}} - 1$, is employed, which introduces the nonlinearity required for nonlinear compensation. At each epoch, the cost function, which is given as

$$J = \frac{1}{2N_S} \sum_{n=0}^{N_S-1} (\mathrm{Re}\{a_n\} - \Omega_I(\bar{\mathbf{z}}_n))^2 + (\mathrm{Im}\{a_n\} - \Omega_Q(\bar{\mathbf{z}}_n))^2, \tag{5.19}$$

is evaluated and weights are updated via back-propagation by using the Levenberg–Marquardt algorithm [58].

5.2.5.4 Adaptive QAM Detection based on NP Bank

In this section, proposed detector structure based on FDE bank is introduced. In [33], it is shown that nonlinear distortion amplification occurs since equivalent channels for linear and nonlinear components of the received signal are subject to different effective channels. To prevent such undesired signal amplification, distortion-aware symbol by symbol detector (DA-SSD) is proposed in [33] to exploit channel diversity. In this study, DA-SSD is extended to MIMO systems to achieve similar diversity gain where ZF-FDE causes distortion amplification. Consider multi-dimensional Bussgang decomposition for NP output vector for g_u^{th} user, $\tilde{\mathbf{a}}_n^{(g_u)} \triangleq [\tilde{a}_n^{(g_u),0}, \tilde{a}_n^{(g_u),1}, ..., \tilde{a}_n^{(g_u),\mu-1}]^{\mathrm{T}}$, which is combination of the received signals at different sampling branches,

$$\tilde{\mathbf{a}}_n^{(g_u)} = \boldsymbol{\beta}^{(g_u)} a_n^{(g_u)} + \boldsymbol{\eta}^{(g_u)}, \tag{5.20}$$

where $\boldsymbol{\beta}^{(g_u)}$ is the vector Bussgang coefficient computed at fast time training period as,

$$\begin{aligned}\boldsymbol{\beta}^{(g_u)} &= \frac{\mathbb{E}\{\tilde{\mathbf{a}}_n^{(g_u)}(a_n^{(g_u)})^*\}}{\mathbb{E}\{|a_n^{(g_u)}|^2\}} \\ &\approx \left(\sum_{n=0}^{N_F-1} \tilde{\mathbf{a}}_n^{(g_u)}(a_n^{(g_u)})^*\right)\left(\sum_{n=0}^{N_F-1} |a_n^{(g_u)}|^2\right)^{-1},\end{aligned} \tag{5.21}$$

and $\mathbf{R}_\eta^{(g_u)}$ is the correlation matrix of distortion vector, $\boldsymbol{\eta}^{(g_u)}$,

$$\begin{aligned}\mathbf{R}_\eta^{(g_u)} &= \mathbb{E}\{(\tilde{\mathbf{a}}_n^{(g_u)} - \boldsymbol{\beta}^{(g_u)}a_n^{(g_u)})(\tilde{\mathbf{a}}_n^{(g_u)} - \boldsymbol{\beta}^{(g_u)}a_n^{(g_u)})^H\} \\ &\approx \frac{1}{N_F-1}\sum_{n=0}^{N_F-1}(\tilde{\mathbf{a}}_n^{(g_u)} - \boldsymbol{\beta}^{(g_u)}a_n^{(g_u)})(\tilde{\mathbf{a}}_n^{(g_u)} - \boldsymbol{\beta}^{(g_u)}a_n^{(g_u)})^H.\end{aligned} \tag{5.22}$$

Assuming Gaussian distribution, PDF of the NP output vector, $\tilde{\mathbf{a}}_n^{(g_u)}$, can be expressed as,

$$\tilde{p}(\tilde{\mathbf{a}}_n^{(g_u)}|a_n) = \frac{\exp\left\{-[\boldsymbol{\Delta}_n^{(g_u)}(a_n)]^H(\mathbf{R}_\eta^{(g_u)})^{-1}\boldsymbol{\Delta}_n^{(g_u)}(a_n)\right\}}{(\pi)^\mu|\mathbf{R}_\eta^{(g_u)}|}, \tag{5.23}$$

where, $\boldsymbol{\Delta}_n^{(g_u)}(a_n) \triangleq \tilde{\mathbf{a}}_n^{(g_u)} - \boldsymbol{\beta}^{(g_u)}a_n$. By using the PDF in (5.23), DA-SSD is expressed as,

$$\hat{a}_n^{(g_u)} = \arg\min_{a_n}\ (\Delta_n^{(g_u)}(a_n))^H(\mathbf{R}_\eta^{(g_u)})^{-1}(\Delta_n^{(g_u)}(a_n)). \tag{5.24}$$

Note that due to its memoryless structure, decision metric introduced in (5.24) has the same computational complexity as the conventional QAM detector.

5.3 Numerical Results

Numerical results are presented to assess the performance of the proposed detector. In addition, systems with linear PA and nonlinear PA with DPD unit are also considered so as to obtain a benchmark for the performance limits. In the considered scenario, summarized in Table 5.1, there are 8 users, which are grouped into 4 groups each having 2 users and each group contains 2 MPC's with equal power, $\sum_{l=0}^{L-1}\mathrm{tr}\left(\mathbf{R}_l^{(g)}\right) = 1$ and 256 QAM is selected as constellation order. Each MPC consists of a cluster of rays, which spans a symbol duration. In this study, MPC cluster centers are separated by 16 symbols. There are $K_t = 120$ antenna in the BS with uniform linear array geometry and they are connected to $D = 12$ RF chains. Slow time training sequence

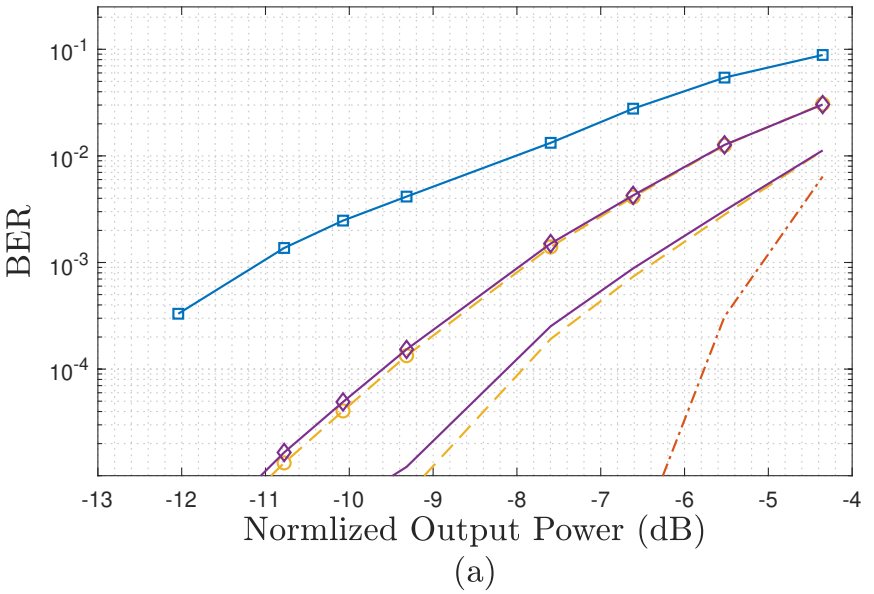
Table 5.1: Scenario

Group	1^{st} MPC Angular Sector	2^{nd} MPC Angular Sector
1	$[-31°, -29°]$	$[-21°, -19°]$
2	$[-11°, 9°]$	$[-1°, 1°]$
3	$[9°, 11°]$	$[19°, 21°]$
4	$[29°, 31°]$	$[39°, 41°]$

of length $N_S = 8192$ is used for nonlinear parameter learning, a sufficiently long fast time training $N_F = 3000$ symbols are used to acquire CSI, and transmission block length is set to $N_D = 8192$. In this study, a lower bound on achievable information rate (AIR) is obtained via GMI metric assuming a Gaussian distribution [33].

Firstly, performance of the proposed detectors are demonstrated for different output backoff levels for Saleh model for noise free case. It can be observed that both BER, in Figure 5.1 (a), and AIR, in Figure 5.1 (b), performances of conventional detectors are significantly degraded, however, NP algorithms provide significant performance improvement. Besides, it can be seen that as output power increases, which makes PA's behave more nonlinear, performances of all methods degrade. It can be observed that proposed DA-SSD outperforms the NP with single FDE branch by exploiting channel diversity. Furthermore; AIR performance of the proposed DA-SSD is very close to that of systems with DPD unlike the BER performance. This stems from the fact that occurrence of bulk errors, which can be avoided by using channel coding, increases BER when post-distortion algorithms are employed; however, such bulk errors are not observed for system with DPD since PA is already linearized. In this study, DPD is designed by using GMP basis functions as proposed in [38].

Furthermore, robustness against AWGN is also studied for Saleh model, where output backoff is chosen as, 9.3 dB. For this purpose, WGN is added to the received signal according to the average SNR definition given in [33]. It can be observed that post-distortion algorithms are not significantly affected by the presence of AWGN. It can be seen from BER results, shown in Figure 5.2, that DA-SSD performs similar to linear PA until it reaches its error floor. It can also be observed that use of FDE bank



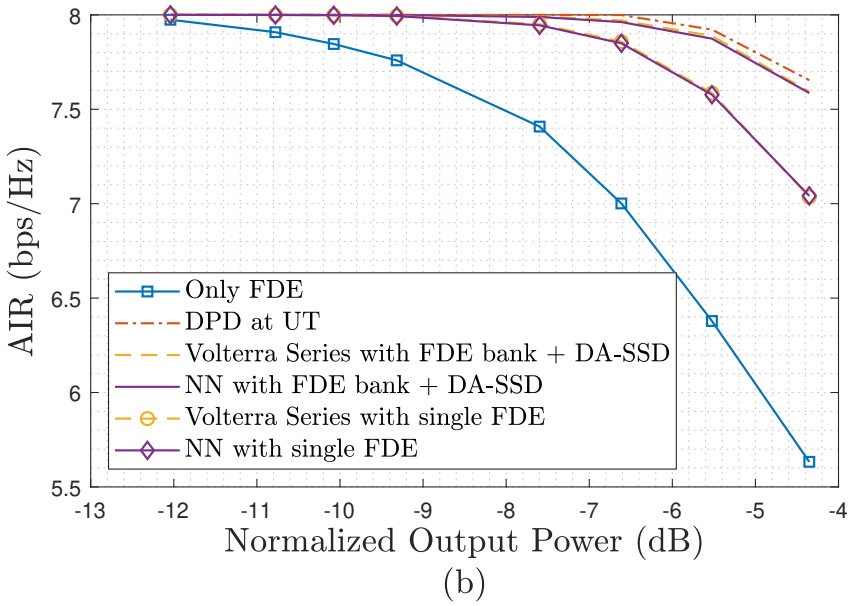
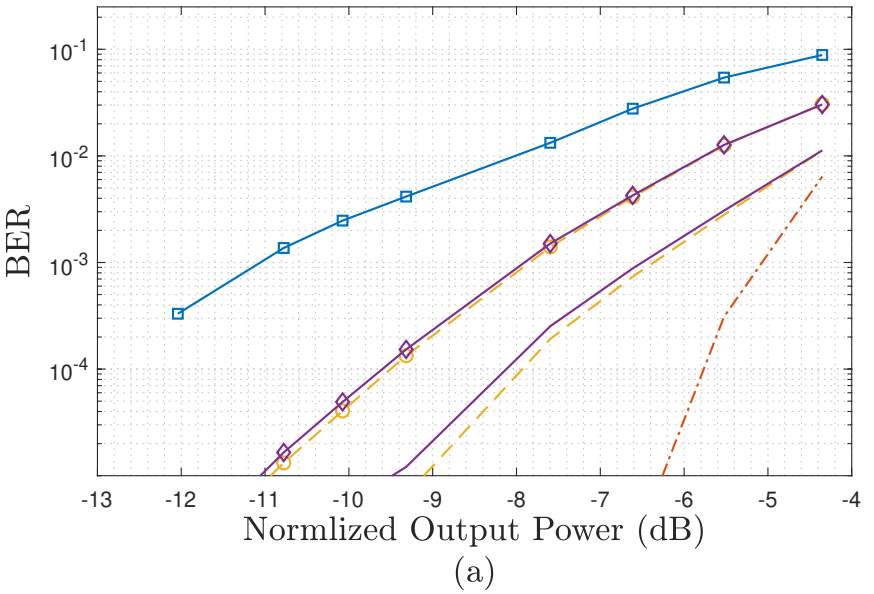
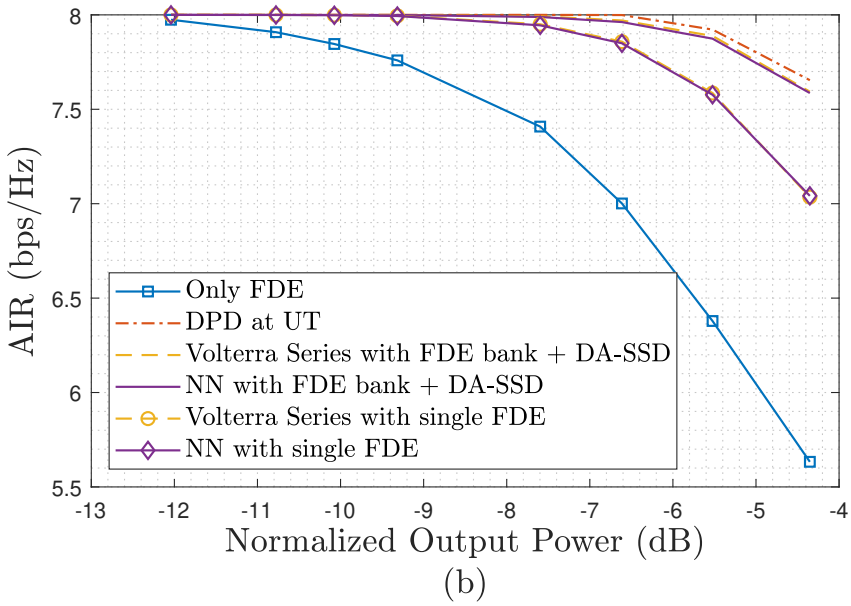


Figure 5.1: BER (a) and AIR (b) vs. output backoff curves for Saleh Model.



(a)



(b)

Figure 5.2: BER vs.SNR curves for Saleh Model with 9.3 dB output backoff.

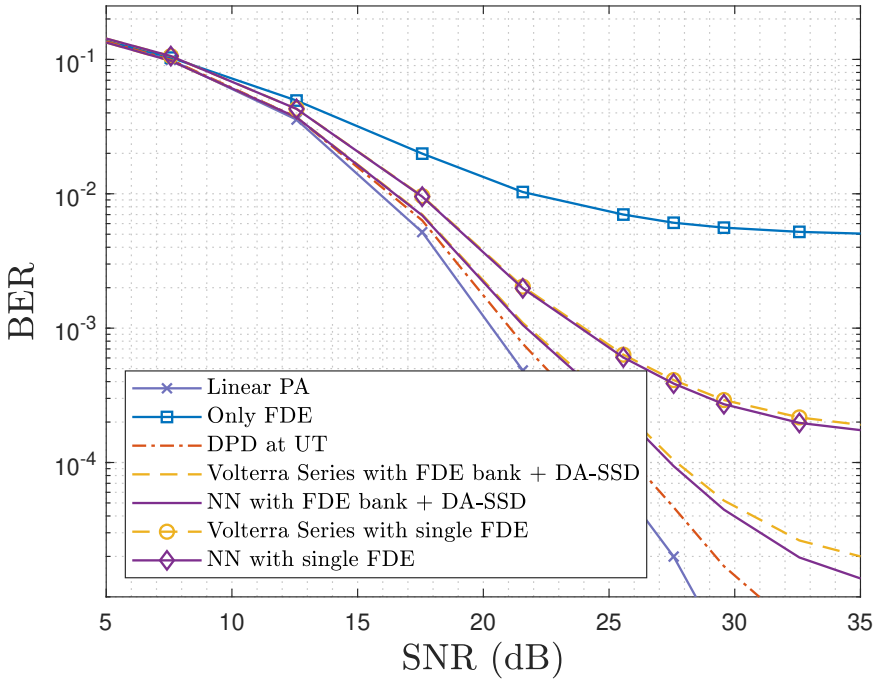


Figure 5.3: BER vs. SNR curves for GaN PA model.

reduces error floor significantly compared to systems with single FDE branch.

Lastly, performances of receiver structures are evaluated for a GMP model, that is extracted from a GaN amplifier [1]. From Figure 5.3, it can be observed that similar conclusions can be drawn for a more realistic PA model. Performance of the proposed DA-SSD approaches to performance of system employing DPD. However, systems, which employ NP with single ZF-FDE branch, suffer from significant performance degradation since relatively high error floor is present.

5.3.1 Conclusions

In this chapter, a novel detector structure is proposed to compensate nonlinear PA effects under dispersive communication channel via bank of ZF-FDE followed by NP. It can be concluded that conventional detectors, which only employs ZF-FDE, perform poorly compared to methods that employ NP and DPD. On the other hand, systems performing NP suffer from performance degradation compared to system with DPD prior to nonlinear PA that is as expected since signal at the output of the PA is almost

linear for systems employing DPD. Hence, nonlinear effects, which cannot be handled after FDE operation, are significantly reduced. As a result, it can be concluded that post-distortion algorithms provides a trade-off between conventional detectors and DPD solution, which requires complex signal processing tasks at UE's. As a future work, application of the proposed method to systems, which contain receiver side nonlinearities such as low resolution analog-to-digital converters (ADC)'s, will be investigated.

5.4 Low Complexity Nonlinear Detection for Multiuser Hybrid MIMO Systems and Performance Analysis

5.4.1 Introduction

This section proposes an analytical framework and a system of distortion-aware receiver structures for single-carrier uplink hybrid beamforming-based multiple-input-multiple-output (MIMO) channels. Proposed single-user processing units cooperate at the base station (BS) in an iterative manner for multi-user symbol detection under hardware impairments. Firstly, in the proposed system, an iterative nonlinear interference cancellation (nIC) based on feedforward neural network (NN) modeling is developed to eliminate the nonlinear distortion effects of power amplifiers (PAs). Furthermore, depending on the fading pattern of the instantaneous channel, it is analytically shown that a large gain may be applied to the distortion signal during the equalization stage, which yields distortion amplification causing a significant decrease in the signal-to-distortion-plus-noise ratio (SDNR). A reduced complexity distortion aware detector, based on a bank of fractionally delayed iterative block decision feedback equalization (IB-DFE) receivers, is introduced to adjust the combiner weights for fractionally delayed IB-DFE outputs so that distortion power is minimized before detection. Consequently, multi-user detection capability is enhanced considerably. Secondly, an asymptotic SDNR analysis is performed. The theoretical analysis, verified via numerical evaluations, clearly shows that distortion amplification limits the achievable SDNR at the BS, and the proposed reduced complexity multi-user receiver architecture improves the BER performance by effectively mitigating this effect.

5.4.2 System Model

5.4.2.1 Uplink SC-FDE Transmission

This chapter considers uplink SC-FDE transmission for frequency selective MIMO systems, where U single-antenna users communicate with a BS, which is equipped with U RF chains and K antennas ($K > U$), and user terminals are equipped with nonlinear PAs. In the proposed system, simultaneously served users are chosen in a manner that enables their spatial differentiation within beamspace. As a consequence, a single radio frequency (RF) chain is allocated per user to detect their respective symbols. Received signal at the BS front-end before beamforming (in full dimension) $\mathbf{y}_{\text{full},n} \in \mathbb{C}^{K\times 1}$ at time index n can be expressed as

$$\mathbf{y}_{\text{full},n} = \sum_{u=1}^{U}\sum_{l=0}^{L-1} \mathbf{h}_{\text{full},l}^{(u)} \tilde{x}_{n-l}^{(u)} + \boldsymbol{\nu}_{\text{full},n}, \tag{5.25}$$

where $\tilde{x}_n^{(u)}$ is the nonlinearly distorted transmitted signal of user u, $\mathbf{h}_{\text{full},l}^{(u)} \in \mathbb{C}^{K\times 1}$ is the l^{th} multipath component (MPC) of the channel response for u^{th} user, and $\boldsymbol{\nu}_{\text{full},n} \in \mathbb{C}^{K\times 1}$ is the additive zero-mean circularly symmetric spatially and temporally complex white Gaussian (CSCG) noise. In this work, a spatially correlated Rician fading channel model for each MPC, which consists of both a non-zero mean and a Rayleigh component, is adopted [77]. Channel vector, $\mathbf{h}_{\text{full},l}^{(u)}$ for $l = 0, \ldots, L-1$, of each user is modeled by spatially correlated multivariate complex Gaussian as

$$\mathbf{h}_{\text{full},l}^{(u)} = \bar{\mathbf{h}}_{\text{full},l}^{(u)} + \tilde{\mathbf{h}}_{\text{full},l}^{(u)}, \tag{5.26}$$

where $\bar{\mathbf{h}}_{\text{full},l}^{(u)}$ is the mean component, $\tilde{\mathbf{h}}_{\text{full},l}^{(u)} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_l^{(u)}\right)$ is the Rayleigh component, and $\mathbf{R}_l^{(u)}$ is the channel covariance matrix of $\tilde{\mathbf{h}}_{\text{full},l}^{(u)}$. According to the uncorrelated scattering assumption, the Rayleigh distributed term of the multipath component is formed by the composition of many rays stemming from far-field scatterers [78] and [79]. Hence, the resultant channel covariance matrix (CCM) of the multipath component $\tilde{\mathbf{h}}_{\text{full},l}^{(u)}$ can be expressed as the superposition of the CCMs of the individual rays, where CCM of each ray $\mathbf{R}_{l,p}^{(u)}$ for $p = 1, \ldots, P$ can be written as [75]

$$\mathbf{R}_{l,p}^{(u)} = \frac{\gamma_l^{(u)}}{P} \mathbf{q}(\theta_p)[\mathbf{q}(\theta_p)]^{\mathrm{H}}, \tag{5.27}$$

where $\mathbf{q}(\theta)$ is the array response vector, $\gamma_l^{(u)}$ is the power delay profile for Rayleigh components, θ_p is the angle of arrival (AoA) of the ray p. Uniform power angular spectrum is assumed, and a uniform linear array with half wavelength antenna separation is considered, where $[\mathbf{q}(\theta)]_k \triangleq \frac{1}{\sqrt{K}} e^{jk\pi \sin(\theta)}$ so that the power of each ray becomes $\mathrm{Tr}(\mathbf{R}_{l,p}^{(u)}) = \frac{\gamma_l^{(u)}}{P}$. Then the overall CCM can be obtained as the sum of the CCMs of individual rays with the AoA values spanning over the set $[\theta_c^{u,l} - \Delta^{u,l}\ \theta_c^{u,l} - \Delta^{u,l}]$. According to the diffuse scattering model, the superposition of CCMs of many rays can be represented by the integral, which is given as

$$\mathbf{R}_l^{(u)} = \lim_{P\to\infty} \sum_{p=1}^{P} \frac{\gamma_l^{(u)}}{P} \mathbf{q}(\theta_p)[\mathbf{q}(\theta_p)]^{\mathrm{H}} = \frac{\gamma_l^{(u)}}{2\Delta^{u,l}} \int_{\theta_c^{u,l}-\Delta^{u,l}}^{\theta_c^{u,l}+\Delta^{u,l}} \mathbf{q}(\theta)[\mathbf{q}(\theta)]^{\mathrm{H}} d\theta, \qquad (5.28)$$

where $2\Delta^{u,l}$ is the angular spread, $\theta_c^{u,l}$ is the mean AoA, and $\theta_p = \theta_c^{u,l} - \Delta^{u,l} + \frac{p-1}{P} 2\Delta^{u,l}$. In (5.26), $\bar{\mathbf{h}}_{\text{full},l}^{(u)} = \kappa_l^{(u)} \mathbf{q}(\theta_c^{u,l})$, where $\kappa_l^{(u)}$ is the non-random complex constant, and then Rician factor can be given as $\left|\kappa_l^{(u)}\right|^2 / \gamma_l^{(u)}$.

In this section, block transmission is considered. In each block, N_D M-QAM symbols $\{a_m^{(u)}\}_{m=0}^{N_D-1}$ are transmitted by employing SC modulation after appending proper cyclic prefix (CP) and cyclic suffix (CS) as

$$x_n^{(u)} = \sum_{m=-N_{CP}}^{N_D+N_{CS}-1} a_m^{(u)} p_{n-\mu m}, \qquad (5.29)$$

where $x_n^{(u)}$ is the transmitted signal of the u^{th} user, N_{CP} and N_{CS} are the lengths of CP and CS, respectively, p_n is the upsampled pulse shaping filter with upsampling factor μ, and $a_m^{(u)}$'s are i.i.d. QAM symbols with $\mathbb{E}\{a_m^{(u)}(a_{m'}^{(u')})^*\} = \delta_{tt'}\delta_{uu'}$. Then, the transmitted signal $x_n^{(u)} = |x_n^{(u)}| e^{j\phi_n^{(u)}}$ is fed to the PA, which nonlinearly modifies the signal as $\tilde{x}_n^{(u)} = \Psi^{(u)}\left(\left\{x_n^{(u)}\right\}\right)$, where $\Psi^{(u)}(\cdot)$ represents the nonlinearity of the PA of the user u.

At the BS, a receiver structure based on hybrid beamforming, which is summarized in Figure 5.5, is employed. In the considered system, a simple phase-only analog beamformer $\mathbf{b}^{(u)}$ of u^{th} user, formed by the steering vector $\mathbf{q}\left(\theta_c^{u,0}\right)$, which focuses on the statistically strongest MPC, is employed to reduce the hardware complexity.

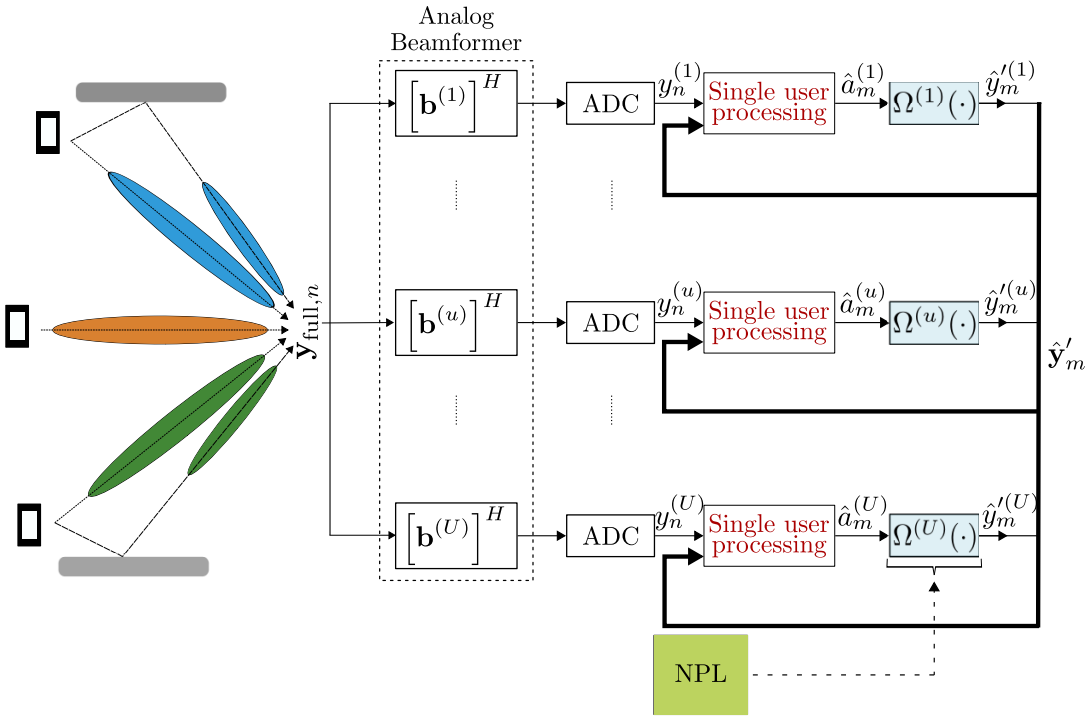


Figure 5.4: Block diagram for hybrid beamforming based system.

Then, beamformer output is obtained for each user as

$$\begin{aligned} y_n^{(u)} &= \left[\mathbf{b}^{(u)}\right]^{\mathrm{H}} \mathbf{y}_{\text{full},n} \\ &= \sum_{l=0}^{L-1} h_l^{(u,u)} \tilde{x}_{n-l}^{(u)} + \sum_{\substack{u'=1 \\ u' \neq u}}^{U} \sum_{l=0}^{L-1} h_l^{(u',u)} \tilde{x}_{n-l}^{(u')} + \nu_n^{(u)}, \end{aligned} \tag{5.30}$$

where $h_l^{(u,u)} \triangleq [\mathbf{b}^{(u)}]^{\mathrm{H}} \mathbf{h}_{\text{full},l}^{(u)}$ is the effective channel of the intended user u, $h_l^{(u',u)} \triangleq [\mathbf{b}^{(u)}]^{\mathrm{H}} \mathbf{h}_{\text{full},l}^{(u')}$ is the leakage channel of user u' to the intended user, and the effective noise $\nu_n^{(u)} \triangleq [\mathbf{b}^{(u)}]^{\mathrm{H}} \boldsymbol{\nu}_{\text{full},n}$.

Although the users are separated in beamspace via analog beamforming, there still exists a considerable amount of IUI due to the suboptimality of the analog stage. The remaining IUI and inter symbol interference (ISI) are mitigated in the digital domain as shown in Figure 5.4. Hence, we propose a reduced complexity receiver structure, which estimates and eliminates the interference of other users on the intended user beam. After eliminating the interference due to other users up to some extent via analog beamforming, the per-user equalization procedure is carried out in the digital domain. However, due to the suboptimality of the beamformer, a significant amount

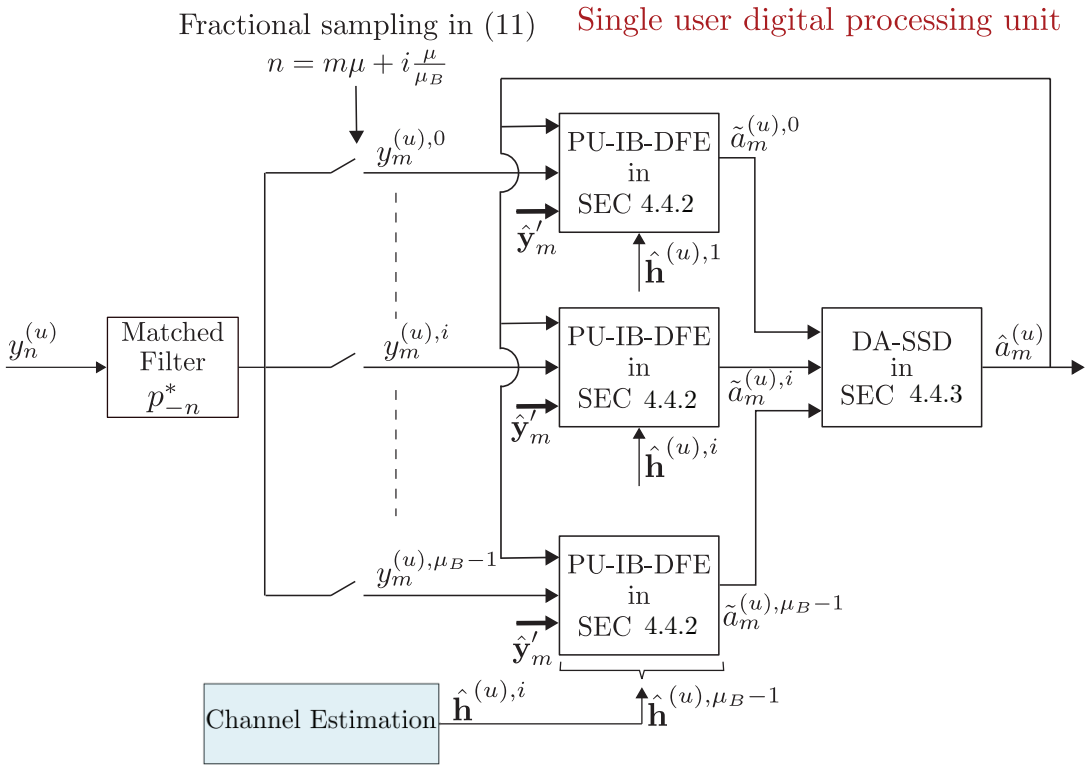


Figure 5.5: Receiver structure for a single user.

of interference remains to be canceled. Therefore, in this chapter, we propose a reduced complexity receiver structure, which estimates and eliminates the interference of other users on the intended user beam.

5.4.2.2 System Description

The proposed system contains different subsystems, which collaborate to decode the information symbols, which are shown in Figure 5.4.

- The first stage is the analog beamforming in Figure 5.4, which is used to separate user signals in beamspace up to some extent. However, due to the suboptimality of the beamformer, a significant amount of interference remains to be canceled out via the proposed digital processing.
- The second stage is the single-user processing unit, shown in Figure 5.5, where symbol detection via per-user IB-DFE is performed by using tentative symbol

estimates of both intended and interfering users. This block utilizes the channel estimates of intended and leakage channels. These estimates are acquired via small-time scale operations at each coherence block, which enable accurate estimation and cancellation of the received interference signal. In addition to symbol estimates, nonlinear distortion estimates, which are obtained through the nonlinear model $\Omega^{(u)}(\cdot)$ as shown in Figure 5.4, are also exploited in the processing stage to reduce the nonlinear distortion effects.

- The last sub-block is the nonlinear distortion estimation stage, where a nonlinear PA model is applied to the tentative decisions to eliminate the nonlinear distortion. Note that nonlinear characteristics of hardware are stationary during communication; hence, "Nonlinear Parameter Learning (NPL)" for nonlinear function ($\Omega^{(u)}(\cdot)$) estimation is performed infrequently on a *large-time scale* through slow time (ST) training by using N_S symbols.

5.4.3 Parameter Learning and Channel Estimation

5.4.3.1 Nonlinear Parameter Learning (NPL)

In this section, a NN-based nonlinear (NL) model is introduced. This model will be used to estimate and cancel out the nonlinear distortion from the received signal through IB-DFE with nonlinear interference cancellation (nIC) block in Figure 5.4.

To construct the model structure, we consider a PA model, expressed by a set of arbitrary basis functions $\psi_q(\cdot)$, where q is the nonlinearity order. NPL is examined under ideal conditions (noise-free, flat fading channel) for each user, and we dropped the user index (u) in the following discussion for the sake of simplicity. However, the NL model can be extracted at the BS per user basis under channel impairments by using the block integration capability over the ST training sequence to enhance the received training signal quality. Then, we can approximate the transmitted signal after nonlinear amplification $\tilde{x}_n = \Psi(\{x_n\})$, as

$$\tilde{x}_n = \sum_{q=0}^{Q-1} \varpi_q \psi_q \left(\sum_{t=0}^{N-1} a_t p_{n-\mu t} \right), \tag{5.31}$$

where $\{\varpi_q\}$ are the basis coefficients, and $\psi_0(x) = x$ corresponds to linear term. After applying matched filtering (MF) in the NPL stage, the received signal is downsampled to the symbol rate of $n = \mu m$. Following these steps, the received signal is rescaled to a magnitude of $\varpi_0 = 1$ at the symbol rate, which can be expressed after performing basic mathematical manipulations, as detailed in [33] as

$$y'_m = a_m + \sum_{q=1}^{Q-1} \varpi_q \psi_q^{(\mathbf{a}_m)} \left(\sum_{t=-L_q+1}^{L_q-1} a_{m-t} \bar{p}_t^{(q)} \right), \tag{5.32}$$

where $\psi_q^{(\mathbf{a}_m)}(\cdot)$ is the sequence dependent nonlinearity, and $\bar{p}_t^{(q)}$ is the downsampled nonlinearly distorted filter response. It is important to note that $\bar{p}_t^{(q)}$ is not a Nyquist-1 pulse; hence, $\sum_{t=-L_q+1}^{L_q-1} a_{m-t}\bar{p}_t^{(q)}$ term yields ISI, where L_q is the truncated memory. Besides, the term, which is exposed to ISI, is nonlinearly distorted at the symbol rate. Consequently, we should construct a model that should, firstly, model the ISI-related term, $a_{\text{ISI},m}^{(q)} \triangleq \sum_{t=-L_q+1}^{L_q-1} a_{m-t}\bar{p}_t^{(q)}$. Here, it is crucial to note that the required NL model for distortion characterization contains *memory*, which emerges from matched filtering and decimation operations. This term can equivalently be expressed as

$$\mathbf{a}_{\text{ISI},m} = \mathbf{W}_{\text{ISI}} \mathbf{a}_m, \tag{5.33}$$

where $\mathbf{a}_{\text{ISI},m} = \left[a_{\text{ISI},m}^{(0)}, a_{\text{ISI},m}^{(1)}, \ldots, a_{\text{ISI},m}^{(Q-1)} \right]^{\text{T}}$ is the vector, which contains ISI terms for different nonlinearity orders, $\mathbf{W}_{\text{ISI}} \in \mathbb{C}^{Q \times 2L_q-1}$ is the coefficient matrix with elements $[\mathbf{W}_{\text{ISI}}]_{q,l} = \bar{p}_{L_q-1-l}^{(q)}$, where $\bar{p}_n^{(0)} = \delta_n$, and $\mathbf{a}_m \in \mathbb{C}^{2L_q-1 \times 1}$ is the symbol vector with elements $[\mathbf{a}_m]_k = a_{m+k-L_q+1}$. Then, nonlinearity is applied to the vector

$$\mathbf{a}_{\psi,m} = \mathbf{\Psi}\left(\mathbf{a}_{\text{ISI},m}\right), \tag{5.34}$$

where $\mathbf{\Psi}(\cdot) = \left[\psi_0^{(\mathbf{a}_m)}(\cdot), \psi_1^{(\mathbf{a}_m)}(\cdot), \ldots, \psi_{Q-1}^{(\mathbf{a}_m)}(\cdot) \right]^{\text{T}}$ is the vector function containing the nonlinear basis functions, and $[\mathbf{a}_{\psi,m}]_q = \psi_q^{(\mathbf{a}_m)}(a_{\text{ISI},m}^{(q)})$. Based on (5.32) and (5.34), the model output is obtained as the combination of the elements of nonlinear basis functions outputs

$$y'_m = \boldsymbol{\varpi}^{\text{T}} \mathbf{a}_{\psi,m}, \tag{5.35}$$

where $\boldsymbol{\varpi} \in \mathbb{C}^{Q \times 1}$ is the vector, which contains basis coefficients, $[\boldsymbol{\varpi}]_q = \varpi_q$. As a result, by combining the expressions (5.33)-(5.35), model output can be expressed as

$$y'_m = \boldsymbol{\varpi}^{\text{T}} \mathbf{\Psi}\left(\mathbf{W}_{\text{ISI}} \mathbf{a}_m\right). \tag{5.36}$$

From (5.36), one can observe that the model has a similar structure to a feedforward neural network in [29]. Therefore, we adopt the augmented real-valued time-delay neural network (ARVTDNN) structure in this study. The output of the model can be expressed as

$$y'_m = \Omega(\mathbf{a}_m) = \Omega_{\mathrm{I}}(\mathbf{a}_m) + j\Omega_{\mathrm{Q}}(\mathbf{a}_m), \tag{5.37}$$

where $\Omega_{\mathrm{I}}(\cdot)$ and $\Omega_{\mathrm{Q}}(\cdot)$ are model functions of real and imaginary parts, respectively. The real part of the network output can be expressed as

$$\Omega_{\mathrm{I}}(\mathbf{a}_m) = \mathbf{w}_{\mathrm{I}}^{\mathrm{T}} g\left(\mathbf{W}_1 \begin{bmatrix} \mathrm{Re}\{\mathbf{a}_m\} \\ \mathrm{Im}\{\mathbf{a}_m\} \end{bmatrix} + \mathbf{b}_1\right) + b_{2,\mathrm{I}}, \tag{5.38}$$

where $\mathbf{w}_{\mathrm{I}} \in \mathbb{C}^{L_1 \times 1}$ is the weight vector and $b_{2,\mathrm{I}}$ is the bias term of the output layer. Similarly, output for the quadrature part is obtained by the weight vector $\mathbf{w}_{\mathrm{Q}}$ and the bias term $b_{2,\mathrm{Q}}$. In (5.38), $\mathbf{W}_1 \in \mathbb{C}^{L_1 \times 4L_q - 2}$ is the weight matrix for the input layer, $\mathbf{b}_1$ and $\mathbf{b}_2$ are the bias vectors, L_1 is the number of neurons in the hidden layer, and $g(x) = \frac{2}{1+e^{-2x}} - 1$ is the activation function, which is shown to provide minimum modeling error compared to the other activation functions in [29].

Via this analysis, we observed that a simple single-layer ARVTDNN is sufficient to capture PA nonlinearities; hence, using more complicated structures, such as convolutional NN (CNN) and DNN, can be avoided to keep the computational complexity lower.

5.4.3.2 Neural Network Training

In order to train the neural network in a supervised manner, the first step is to generate the training data. For this purpose, a symbol sequence a_m of length N_{S} is generated from a QAM alphabet. Next, the training waveform x_n is obtained after upsampling and pulse shaping stages, as expressed in (5.29). The waveform x_n is then passed through the power amplifier (PA) of the corresponding UE, and the upsampled training sequence is acquired, resulting in labeled data.

As the next step, the upsampled training data $\tilde{x}_n$ is captured. This can be done either in the transmitter side by using an observation chain or it can be directly received at the BS. After matched filtering and downsampling, the labeled symbols $y'_m =$

$\left(\tilde{x}_n * p^*_{-n}\right)|_{n=\mu m}$ are obtained. These labeled symbols serve as the training pairs $\{a_m, y'_m\}_{m=0}^{N_\mathrm{S}-1}$ that are used to train the NN coefficients.

The weights of NN coefficients are updated via back-propagation by using the Levenberg-Marquardt algorithm [58]. NN training aims to find a model that minimizes the mean square of the estimation error. Hence, the loss function to be minimized can be expressed as

$$J = \frac{1}{2N_\mathrm{S}} \sum_{m=0}^{N_\mathrm{S}-1} (\mathrm{Re}\{y'_m\} - \Omega_I(\mathbf{a}_m))^2 + (\mathrm{Im}\{y'_m\} - \Omega_Q(\mathbf{a}_m))^2. \tag{5.39}$$

Equivalently, a simple gradient descent algorithm can be performed to update NN coefficients. In order to train the NN coefficients $\mathbf{W}_1$, $\mathbf{w}_\mathrm{I}$ and $\mathbf{w}_\mathrm{Q}$, an iterative coefficient update procedure is employed. Then the update equations of $\mathbf{w}_\mathrm{I}$ and $\mathbf{w}_\mathrm{Q}$ for the i^{th} iteration can be expressed as,

$$\mathbf{w}_\mathrm{I}[i] = \mathbf{w}_\mathrm{I}[i-1] + \lambda \mathbf{a}'_i e_{\mathrm{I,i}}, \tag{5.40a}$$

$$\mathbf{w}_\mathrm{Q}[i] = \mathbf{w}_\mathrm{Q}[i-1] + \lambda \mathbf{a}'_i e_{\mathrm{Q,i}} \tag{5.40b}$$

where the error signals $e_{\mathrm{I,i}} \triangleq \mathrm{Re}\{y'_i\} - \Omega_I(\mathbf{a}_i)$, $e_{\mathrm{Q,i}} \triangleq \mathrm{Im}\{y'_i\} - \Omega_Q(\mathbf{a}_i)$ and the input vector $\mathbf{a}'_i \triangleq g\left(\mathbf{W}_1 \left[\mathrm{Re}\{\mathbf{a}_i^\mathrm{T}\}\ \mathrm{Im}\{\mathbf{a}_i^\mathrm{T}\}\right]^\mathrm{T}\right)$. Weight vector for the p^{th} neuron can be represented as the p^{th} row of $\mathbf{W}_1$, namely $\mathbf{w}_1^p = [\mathbf{W}_1]_{p,:}$, and it is updated as

$$\mathbf{w}_1^p[i] = \mathbf{w}_1^p[i-1] + \lambda \left(\bar{\mathbf{a}}_i \left[\mathbf{w}_\mathrm{I}[i]\right]_p e_{\mathrm{I,i}} + \bar{\mathbf{a}}_i \left[\mathbf{w}_\mathrm{Q}[i]\right]_p e_{\mathrm{Q,i}}\right) \tag{5.41}$$

where $\bar{\mathbf{a}}_i \triangleq \left(1 - g\left((\mathbf{a}'_i)^\mathrm{T} \mathbf{w}_1^p[i-1]\right)^2\right) \mathbf{a}'_i$, and λ is the learning rate.

5.4.3.3 Nonlinear Interference Prediction

In the considered system, the NN model is used to predict the nonlinearly modified signal by using the tentative decisions obtained after receiver processing $\hat{a}_m$ so that nonlinear distortion can be canceled. A nonlinear model at symbol rate is extracted via NN for each user. The model output is estimated by using the tentative decisions $\hat{a}_m^{(u)}$ as $\hat{y}_m'^{(u)} = \Omega^{(u)}\left(\hat{\mathbf{a}}_m^{(u)}\right)$. By using the model output, distortion signal for u^{th} user, namely $\hat{d}_m^{(u)}$, can be simply estimated as

$$\hat{d}_m^{(u)} = \hat{y}_m'^{(u)} - \hat{a}_m^{(u)}. \tag{5.42}$$

5.4.3.4 Channel Estimation with Fractional Sampling

In this study, we consider FDE to compensate for the effects of frequency selectivity of channels. It is shown numerically in [33] that distortion amplification occurs due to a mismatch between the channels of linear and nonlinear signal components; hence, a receiver structure based on a fractionally spaced FDE bank is proposed to exploit channel diversity. The same approach is adopted, and the effect of shifting the sampling instant fractionally is shown analytically in Section 5.4.5, and a nonlinear FDE method is proposed, which takes both distortion amplification and multi-user scenario into account.

Channel acquisition at BS is performed for each user u and fractional sampling instant $i = 0, \ldots, \mu_B - 1$ by taking the interference of other users on the intended user beam into account, where μ_B is the number of sampling instants as shown in Figure 5.5. The received signal for the i^{th} branch after analog beamforming in (5.30) and MF is expressed as

$$\mathbf{y}^{(u),i} \triangleq \left[y_0^{(u),i}, y_1^{(u),i}, \ldots, y_{N_T-1}^{(u),i} \right]^{\mathrm{T}}, \tag{5.43}$$

where $y_m^{(u),i} = \left(y_n^{(u)} * p_{-n}^{*} \right)_{n=m\mu+i\frac{\mu}{\mu_B}}$ is the downsampled signal for the i^{th} branch, and $\mathbf{y}^{(u),i} \in \mathbb{C}^{N_T \times 1}$ is the signal vector that contains all training symbols on the u^{th} RF chain and N_T is the number of training symbols. By using this training sequence, the fractionally delayed channel, which represents the interference channel of user u' onto u^{th} user's beam, $\mathbf{h}^{(u',u),i} = \left[h_{-L_b+1}^{(u',u),i}, \ldots, h_0^{(u',u),i}, \ldots, h_{L_f-1}^{(u',u),i} \right]^{\mathrm{T}}$ can be estimated via different estimation techniques [72].

5.4.4 Distortion-Aware Iterative Detector with nIC

In this section, the proposed nIC-based receiver, which is the single-user digital processing unit detailed in Figure 5.6, is introduced. The advantages of the proposed algorithm over existing methods are two-fold in terms of performance and complexity. Firstly, existing MIMO receivers proposed in [48] and [65] do not take distortion amplification into account. Therefore, they suffer from performance degradation since they do not aim to mitigate amplified nonlinear distortion. In addition, these methods perform joint MMSE detection, and they require matrix inversion.

The proposed PU-IB-DFE algorithm reduces the problem dimension to a single user by performing beamspace operations as detailed in the "Complexity Analysis" Section 5.4.6. Therefore, the proposed algorithm reduces the equalization complexity to single scalar multiplication, which is lower than that of existing methods [48, 65]. Secondly, the proposed IB-DFE receiver estimates nonlinear distortion independent of the sampling branch; hence, a single NN implementation is sufficient. Hence, it has significantly lower computational complexity than the SISO receiver, which requires multiple complex NN implementations [33]. In addition, since the proposed algorithm proposes an iterative decision-directed detection, its performance is also superior than that of the non-iterative method in [33] as shown in the "Performance Evaluations" Section 5.5.

In the considered system, a nonlinear FDE technique, called IB-DFE with nIC, is used to suppress ISI and IUI iteratively.

5.4.4.1 Description of the proposed algorithm

The proposed algorithm is developed based on the received signal expression in (5.30), which contains two main signal terms.

- The first signal component is related to the intended signal $\sum_{l=0}^{L-1} h_l^{(u,u)} \tilde{x}_{n-l}^{(u)}$. However, this signal is nonlinearly distorted, as explained in Section 5.4.3.1. Hence, the first step is to cancel the nonlinear distortion term in symbol rate, which is denoted as $d_m^{(u)}$. The distortion estimate is obtained using the tentative decisions as in (5.42). Furthermore, the distortion estimate is convolved with the channel estimate $h_m^{(u',u),i}$ to obtain the received distortion signal. The frequency domain equivalent of this step, which corresponds to the cancellation of the "Self nonlinear distortion" term, is given in (5.44).

- The second signal component is related to the IUI component in (5.30), which is given as $\sum_{\substack{u'=1 \\ u' \neq u}}^{U} \sum_{l=0}^{L-1} h_l^{(u',u)} \tilde{x}_{n-l}^{(u')}$. Note that this signal component is a function of both leakage channel $h_l^{(u',u)}$ and the nonlinearly distorted interfering users' signals $\tilde{x}_{n-l}^{(u')}$. The leakage signal affects the received signal quality, which in turn, degrades the detection performance. Hence, cancellation of the "Nonlin-

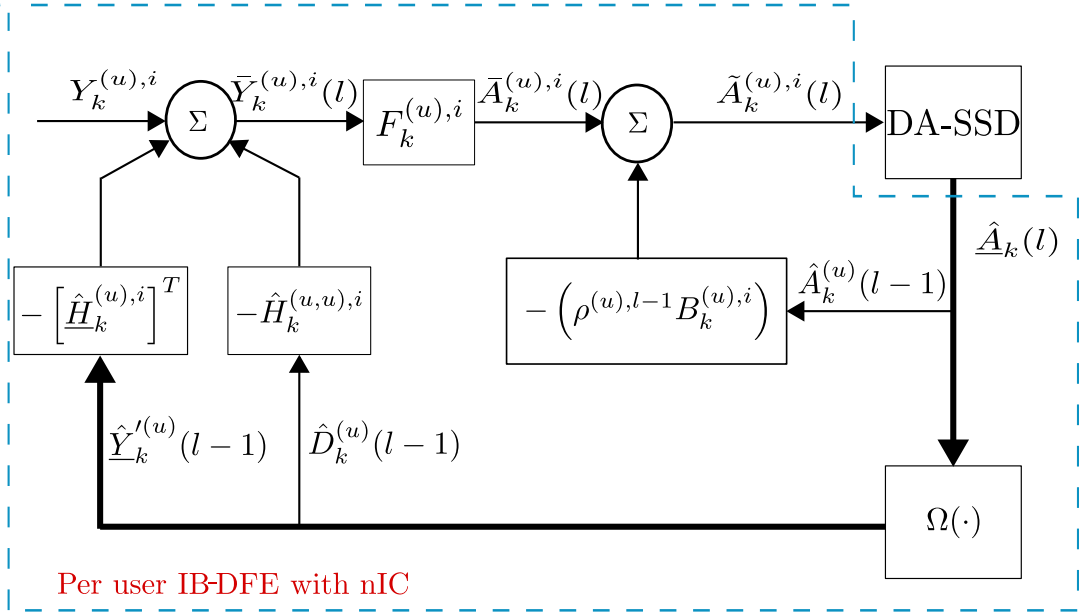


Figure 5.6: PU-IB-DFE with nIC block. For simplicity, all variables are represented in DFT domain.

ear IUI" term in (5.44) is performed in *downsampled signal domain* to increase the received signal quality.

- After removing nonlinear distortion and the IUI, conventional IB-DFE operations are applied on the received signal per user basis to obtain soft symbol estimates.

- By using the PU-IB-DFE algorithm, soft symbol decisions are obtained for each sampling branch, where the resultant distortion power varies over sampling branch output depending on channel frequency response. Consequently, in the last stage, soft symbol estimates are combined in a way that the ultimate distortion power is minimized.

5.4.4.2 IB-DFE with nIC

Consider the frequency domain representation of the received signal after MF for the i^{th} sampling branch for user u, $y_m^{(u),i}$ in (5.43), which is denoted by $Y_k^{(u),i}$. One can obtain the signal $\bar{Y}_k^{(u),i}(l)$ from $Y_k^{(u),i}$ after nIC for all users and linear IUI cancellation

for the interfering users at the l^{th} iteration as

$$\bar{Y}_k^{(u),i}(l) = Y_k^{(u),i} - \underbrace{\sum_{\substack{u'=1 \\ u' \neq u}} \hat{H}_k^{(u',u),i} \hat{Y}_k'^{(u')}(l-1)}_{\text{Nonlinear IUI}} - \underbrace{\hat{H}_k^{(u,u),i} \hat{D}_k^{(u)}(l-1)}_{\text{Self nonlinear distortion}}, \tag{5.44}$$

where the frequency domain signals are defined as $Y_k^{(u),i} = \mathcal{F}\left\{y_m^{(u),i}\right\}$ is the received signal, $\hat{H}_k^{(u',u),i} = \mathcal{F}\left\{\hat{h}_m^{(u',u),i}\right\}$ is the estimated effective channel frequency response of user u', seen by u^{th} RF chain, $\hat{D}_k^{(u)}(l) = \mathcal{F}\left\{\hat{d}_m^{(u)}(l)\right\}$ is the estimated distortion signal of user u at l^{th} iteration obtained from (5.42), and $\hat{Y}_k'^{(u)}(l) = \mathcal{F}\left\{\hat{y}_m'^{(u)}(l)\right\}$ is the estimate of nonlinearly distorted signal in (5.42).

After interference cancellation, IB-DFE type equalization [48, 65] for the u^{th} user on (5.44) is performed, and we obtain the soft symbol estimate in the frequency domain for the i^{th} sampling branch at the l^{th} iteration as

$$\tilde{A}_k^{(u),i}(l) = F_k^{(u),i} \bar{Y}_k^{(u),i}(l) - \rho^{(u),l-1} B_k^{(u),i} \hat{A}_k^{(u)}(l-1), \tag{5.45}$$

where $F_k^{(u),i}$ is the feedforward MMSE filter given by

$$F_k^{(u),i} = \left[\hat{H}_k^{(u,u),i}\right]^* \left(\left|\hat{H}_k^{(u,u),i}\right|^2 + \epsilon\right)^{-1}, \tag{5.46}$$

and $\epsilon = \frac{1}{SDNR}$ is the regularization parameter. In this study, the regularization parameter is selected as a function of SDNR as proposed in the *robust receiver* in [48]. In (5.45), reliability metric $\rho^{(u),l}$ is found as

$$\rho^{(u),l} = \mathbb{E}\left\{\hat{a}_m^{(u)}(l)\left(a_m^{(u)}\right)^*\right\}\left(\mathbb{E}\left\{\left|a_m^{(u)}\right|^2\right\}\right)^{-1}, \tag{5.47}$$

where $\hat{a}_m^{(u)}(l)$ is the hard decision produced by the detector, $\hat{A}_k^{(u)}(l) = \mathcal{F}\left\{\hat{a}_m^{(u)}(l)\right\}$ at the l^{th} iteration, and $B_k^{(u),i}$ is the feedback filter, which is used to eliminate the remaining *linear* ISI of the intended user, which can be given as

$$B_k^{(u),i} = F_k^{(u),i} \hat{H}_k^{(u,u),i} - 1. \tag{5.48}$$

In the next section, a distortion-aware detector, which combines soft symbol estimates obtained for each fractional delay, namely $\tilde{a}_m^{(u),i}(l) = \mathcal{F}^{-1}\left\{\tilde{A}_k^{(u),i}(l)\right\}$, will be introduced.

5.4.4.3 Distortion-Aware Adaptive QAM Detection with IB-DFE

In (5.45), soft symbol estimates are obtained for each sampling branch, and it is shown in Section 5.4.5 that SDNR is different on the fractional sampling channels due to the distortion amplification phenomenon. From (5.46), it can be observed that the applied gain to equalize the frequency selective channel is roughly proportional to $1/\hat{H}_k^{(u,u),i}$; therefore, if the magnitude of the channel frequency response $\hat{H}_k^{(u,u),i}$ is small then the applied gain can be significantly large. Consequently, to equalize the linear signal, the distortion term is multiplied with a large gain. This may yield amplification of the distortion signal since its channel characteristics differ as can be seen from the equivalent channel representation analysis in Appendix A, specifically from (A.11). Furthermore, this effect is more evident in practical spatially correlated mm-wave channels since all the antenna elements have a similar frequency response due to channel correlation across antenna elements. On the other hand, for a rich scattering environment, the effects of distortion amplification might be mild since each antenna observes an uncorrelated channel, which results in channel hardening. Therefore, the fading pattern of the effective channel might be smoother; however, the spatially correlated channel model, which is employed in this study, is considered to be more practical for next-generation wireless networks [75]. Hence, a distortion-aware symbol-by-symbol detector (DA-SSD), which is introduced in [33], can be adapted to multi-user detection problems as a feasible solution for the considered system. DA-SSD is used to estimate the transmitted symbols for each user independently, and those symbol estimates are exploited to cancel IUI and nonlinear distortion throughout the iterations (5.44) and (5.45).

The vector input of the DA-SSD for user u can be formed by the soft symbol estimates as $\tilde{\mathbf{a}}_m^{(u)} \triangleq \left[\tilde{a}_m^{(u),0}, \tilde{a}_m^{(u),1}, ..., \tilde{a}_m^{(u),\mu_B-1}\right]^{\mathrm{T}}$, and it can be expressed by employing generalized Bussgang decomposition for non-Gaussian signals, which is a useful tool for quantifying the distortion power of the nonlinearly distorted signal [59]. Thanks to the quantification of the distortion power, a combiner, which minimizes the overall distortion power can be calculated. Via the generalized Bussgang decomposition, the soft symbol estimate vector $\tilde{\mathbf{a}}_m^{(u)}$ can be represented by the desired useful term $a_m^{(u)}$

and the uncorrelated distortion-plus-noise (DN) vector $\boldsymbol{\eta}_m^{(u)}$ as

$$\tilde{\mathbf{a}}_m^{(u)} = \boldsymbol{\beta}^{(u)} a_m^{(u)} + \boldsymbol{\eta}_m^{(u)}, \tag{5.49}$$

where $\boldsymbol{\beta}^{(u)}$ is the Bussgang coefficient vector, which is found in the training phase as

$$\hat{\boldsymbol{\beta}}^{(u)} \approx \left(\sum_{m=0}^{N_T-1} \tilde{\mathbf{a}}_m^{(u)} \left(a_m^{(u)}\right)^* \right) \left(\sum_{m=0}^{N_T-1} |a_m^{(u)}|^2 \right)^{-1}. \tag{5.50}$$

The correlation matrix for the DN vector can be obtained by using training symbols as

$$\begin{aligned} \hat{\mathbf{R}}_\eta^{(u)} &= \mathbb{E}\{(\tilde{\mathbf{a}}_n^{(u)} - \boldsymbol{\beta}^{(u)} a_n^{(u)})(\tilde{\mathbf{a}}_n^{(u)} - \boldsymbol{\beta}^{(u)} a_n^{(u)})^{\mathrm{H}}\} \\ &\approx \frac{1}{N_T} \sum_{m=0}^{N_T-1} (\tilde{\mathbf{a}}_m^{(u)} - \hat{\boldsymbol{\beta}}^{(u)} a_m^{(u)})(\tilde{\mathbf{a}}_m^{(u)} - \hat{\boldsymbol{\beta}}^{(u)} a_m^{(u)})^{\mathrm{H}}. \end{aligned} \tag{5.51}$$

Note that the correlation matrix $\hat{\mathbf{R}}_\eta^{(u)}$ can also be calculated analytically as expressed in (A.13).

Symbol estimates from different branch outputs can be combined by using minimum variance distortionless response (MVDR) filter. Combiner vector for u^{th} user $\mathbf{w}_c^{(u)} \in \mathbb{C}^{\mu_B \times 1}$ can be expressed as

$$\mathbf{w}_c^{(u)} = \left(\left(\hat{\mathbf{R}}_\eta^{(u)}\right)^{-1} \hat{\boldsymbol{\beta}}^{(u)} \right) \left(\left(\hat{\boldsymbol{\beta}}^{(u)}\right)^{\mathrm{H}} \left(\hat{\mathbf{R}}_\eta^{(u)}\right)^{-1} \hat{\boldsymbol{\beta}}^{(u)} \right)^{-1}. \tag{5.52}$$

The soft symbol estimate at l^{th} iteration is obtained as

$$\tilde{a}_m^{(u)}(l) = \left(\mathbf{w}_c^{(u)}\right)^{\mathrm{H}} \tilde{\mathbf{a}}_m^{(u)}(l). \tag{5.53}$$

Then, hard detector outputs $\hat{a}_m^{(u)}(l)$ are obtained from $\tilde{a}_m^{(u)}(l)$ based on minimum distance criterion. Also, after the last IB-DFE iteration in (5.45), low density parity check (LDPC) decoding can be employed as forward error correction (FEC).

5.4.5 SDNR Analysis for the Proposed Detector

In this section, SDNR analysis of the proposed detector in Section 5.4.4 will be performed to present the performance improvement brought by the use of fractionally spaced FDE bank structure in hybrid MIMO systems. In the analysis, based on IB-DFE with nIC and DA-SSD, residual distortion power is derived and used to calculate the received SDNR.

We start the analysis by expressing the soft symbol estimates after IB-DFE with nIC in (5.45) and MVDR combining in (5.53), which are described in Section 5.4.4.2 and 5.4.4.3 [1], as

$$\tilde{A}_k^{(u)}(l) = \sum_{i=0}^{\mu_B-1} w_c^{(u),i}\left[F_k^{(u),i}\bar{Y}_k^{(u),i}(l) - \rho^{(u),l-1}B_k^{(u),i}\hat{A}_k^{(u)}(l-1)\right], \tag{5.54}$$

where $w_c^{(u),i} = \left[\mathbf{w}_c^{(u)}\right]_i$, and $\bar{Y}_k^{(u),i}$ can be expressed as

$$\begin{aligned}\bar{Y}_k^{(u),i}(l) = Y_k^{(u),i} &- H_k^{(u,u),i}\hat{D}_k^{(u)}(l-1) \\ &- \sum_{\substack{u'=1\\u'\neq u}}^{U}\left[H_k^{(u',u),i}\left(\hat{A}_k^{(u')}(l-1) + \hat{D}_k^{(u')}(l-1)\right)\right],\end{aligned} \tag{5.55}$$

where, $\hat{D}_k^{(u)}(l)$ is the estimate of $D_k^{(u)}$ in (A.9) based on tentative decisions at l^{th} iteration via (5.42). Soft symbol estimate expression in (5.54) can be further written as in (5.56) by using the equivalent channel representations derived in Appendix A, including the noise term, so that each term can be investigated individually

$$\begin{aligned}\tilde{A}_k^{(u)}(l) = &\underbrace{P_k^{(u,u)}A_k^{(u)} - Q_k^{(u)}\hat{A}_k^{(u)}(l-1)}_{(\Lambda_k^1)} + \underbrace{P_k^{(u,u)}\left(D_k^{(u)} - \hat{D}_k^{(u)}(l-1)\right)}_{(\Lambda_k^2)} \\ &+ \underbrace{\sum_{\substack{u'=1\\u'\neq u}}^{U} P_k^{(u',u)}\left(A_k^{(u')} - \hat{A}_k^{(u')}(l-1)\right)}_{(\Lambda_k^3)} + \underbrace{\sum_{\substack{u'=1\\u'\neq u}}^{U} P_k^{(u',u)}\left(D_k^{(u')} - \hat{D}_k^{(u')}(l-1)\right)}_{(\Lambda_k^4)} \\ &+ \underbrace{\sum_{u'=1}^{U}\sum_{i=0}^{\mu_B-1} w_c^{(u),i}F_k^{(u),i}\tilde{D}_k^{(u',u),i} + \sum_{i=0}^{\mu_B-1} w_c^{(u),i}F_k^{(u),i}\mathcal{V}_k^{(u),i}}_{(\Lambda_k^5)},\end{aligned} \tag{5.56}$$

where the scalars are defined as $P_k^{(u',u)} \triangleq \left[\sum_{i=0}^{\mu_B-1} w_c^{(u),i}F_k^{(u),i}H_k^{(u',u),i}\right]$, $Q_k^{(u)} \triangleq \left[\sum_{i=0}^{\mu_B-1} w_c^{(u),i}B_k^{(u),i}\rho^{(u),l-1}\right]$, and $\tilde{D}_k^{(u',u),i}$ is mismatched distortion defined in (A.11).

The MVDR filter in (5.52) can be analytically calculated by employing the correlation matrix $\mathbf{R}_\eta^{(u)}$ of the remaining distortion, and $\mathbf{R}_\eta^{(u)}$ can be computed as

$$\mathbf{R}_\eta^{(u)} = \sum_{u'=1}^{U}\mathbf{R}_d^{(u',u)}, \tag{5.57}$$

[1] In the analysis, both nonlinear modeling and channel estimation errors are not considered assuming that both are estimated accurately via training.

where $(i,r)^{th}$ entity of $\mathbf{R}_d^{(u',u)}$ is the correlation between distortion in i^{th} and r^{th} branches, expressed by (A.13) in Appendix A.

By using the symbol estimate expression in (5.56), we can calculate the distortion-plus-noise power, namely $S_{\text{DN}}^{(u)}(l)$, of the received signal at l^{th} iteration as

$$S_{\text{DN}}^{(u)}(l) = \frac{1}{N_D} \sum_{k=0}^{N_D-1} \mathbb{E}\left\{ \left| A_k^{(u)} - \tilde{A}_k^{(u)}(l) \right|^2 \right\}. \tag{5.58}$$

By assuming the orthogonality of the symbols for different users, and the distortion signals[2], which is based on generalized Bussgang theorem, we can express $S_{\text{DN}}^{(u)}(l)$ as the summation of the following power terms as

$$S_{\text{DN}}^{(u)}(l) = \frac{1}{N_D} \sum_{k=0}^{N_D-1} \left[S_{\text{PQ},k}^{(u)}(l) + S_{d,k}^{(u)}(l) + J_k^{(u)}(l) + J_{d,k}^{(u)}(l) \right] + \tilde{S}_d^{(u)}. \tag{5.59}$$

Under these assumptions, different terms in (5.59) can be calculated as follows. Remaining signal power due to imperfect decision feedback in IB-DFE operations $S_{\text{PQ},k}^{(u)}(l) \triangleq \mathbb{E}\left\{ |A_k^{(u)} - \Lambda_k^1|^2 \right\}$ can be calculated as

$$S_{\text{PQ},k}^{(u)}(l) = \left| P_k^{(u,u)} - 1 \right|^2 + \left| Q_k^{(u)} \right|^2 - 2\,\text{Re}\left\{ \rho^{(u),l-1} \left(P_k^{(u,u)} - 1 \right)^* Q_k^{(u)} \right\}, \tag{5.60}$$

where $\mathbb{E}\left\{ |A_k^{(u)}|^2 \right\} = \mathbb{E}\left\{ |\hat{A}_k^{(u)}(l)|^2 \right\} = 1$. The remaining distortion power $S_{d,k}^{(u)} \triangleq \mathbb{E}\left\{ |\Lambda_k^2|^2 \right\}$ due to imperfect symbol estimation at the previous iteration can be computed as

$$S_{d,k}^{(u)}(l) = 2\left(1 - \text{Re}\left\{ \rho_d^{(u),l-1} \right\}\right) \left| P_k^{(u,u)} \right|^2 \mathbb{E}\left\{ \left| D_k^{(u)} \right|^2 \right\}, \tag{5.61}$$

where $\rho_d^{(u),l}$ is the reliability factor for the distortion estimation, which is defined as

$$\rho_d^{(u),l} \triangleq \mathbb{E}\left\{ \hat{d_n}^{(u)}(l) \left({d_n}^{(u)} \right)^* \right\} \left(\mathbb{E}\left\{ \left| {d_n}^{(u)} \right|^2 \right\} \right)^{-1}. \tag{5.62}$$

Similar to the power term of the intended user, we can calculate the remaining power of interfering users after IUI cancellation and nIC. Residual power due to non-ideal symbol detection of interfering users $J_k^{(u)}(l) \triangleq \mathbb{E}\left\{ |\Lambda_k^3|^2 \right\}$ can be obtained as

[2] Note that distortion estimation error term Λ_k^4 and mismatch distortion related term Λ_k^5 are assumed to be orthogonal since estimation error stems from incorrect tentative decisions; however, mismatch distortion is due to the channel mismatch, which is uncorrelated to tentative decisions.

$$J_k^{(u)}(l) = \sum_{\substack{u'=1 \\ u' \neq u}}^{U} 2\left(1 - \mathrm{Re}\left\{\rho^{(u'),l-1}\right\}\right)\left|P_k^{(u',u)}\right|^2. \tag{5.63}$$

Residual power due to non-ideal nIC of interfering users $J_{d,k}^{(u)}(l) \triangleq \mathbb{E}\left\{\left|\Lambda_k^4\right|^2\right\}$ is obtained as

$$J_{d,k}^{(u',u)}(l) = \sum_{\substack{u'=1 \\ u' \neq u}}^{U} 2\left(1 - \mathrm{Re}\left\{\rho_d^{(u'),l-1}\right\}\right)\left|P_k^{(u',u)}\right|^2 \mathbb{E}\left\{\left|D_k^{(u')}\right|^2\right\}. \tag{5.64}$$

Finally, the residual power of the combination of mismatch distortion and the thermal noise $\tilde{S}_d^{(u)} \triangleq \frac{1}{N_D}\sum_{k=0}^{N_D-1}\mathbb{E}\left\{\left|\Lambda_k^5\right|^2\right\}$ can be computed by using the correlation matrix of the residual distortion component in (5.57) as $\tilde{S}_d^{(u)} = \left[\mathbf{w}_c^{(u)}\right]^{\mathrm{H}}\mathbf{R}_\eta^{(u)}\mathbf{w}_c^{(u)}$. SDNR expression can be written by combining all terms as $SDNR^{(u)}(l) = 1/S_{\mathrm{DN}}^{(u)}(l)$, and one can obtain the uncoded BER expression by taking the following expectation for M-QAM constellation with respect to channel realizations as

$$P_b \approx \frac{1}{U}\sum_{u=1}^{U}\frac{4}{\log_2(M)}\mathbb{E}_{\left\{h_m^{(u',u)}\right\}}\left\{Q\left(\sqrt{3\frac{SDNR^{(u)}}{M-1}}\right)\right\}. \tag{5.65}$$

In (5.65), the Gaussian assumption is made for the distortion term after nIC and IB-DFE. However, the remaining distortion does not exactly have a Gaussian probability density function (*pdf*). It is stated in [80] that the *pdf* has a relatively smaller slope on the tail due to nonlinear distortion. In the IB-DFE receiver, additional non-Gaussian interference comes from imperfect decision feedback. Therefore, relatively high tails departing from Gaussian distribution become more problematic since they yield bit errors for uncoded systems. We employed FEC to reduce the effects of tails in the *pdf* of the residual distortion term after MVDR combining $[\mathbf{w}_c^{(u)}]^{\mathrm{H}}\boldsymbol{\eta}_n^{(u)}$. Therefore, LDPC coding with a high code rate ($r = 8/9$) is utilized to compensate for errors due to high tails.

5.4.6 Complexity Analysis

In Table 5.2, we compared the proposed receiver's computational complexity with baseline techniques based on the number of floating-point operations (FLOPs) per symbol for one iteration. For the comparison, we considered the complexities of

both the conventional Robust MMSE and the Robust MMSE with nIC techniques. When comparing the complexities of the considered techniques, it is noted that the proposed PU-IB-DFE with nIC and Robust MMSE with nIC [65] are both heavily influenced by the nonlinear distortion estimation. Specifically, in order to create the input of an activation function of the NN-based nonlinear model, $4L_q$ real multiplications are needed to perform real/imaginary processing. Consequently, $4L_qL_1$ real multiplications are required to generate L_1 neuron inputs. However, one should note that employing the FDE bank does not increase NN-based nIC complexity since the same NL model output is used for each sampling branch, unlike in [33], where NN-based predictors are employed at each branch. Robust MMSE [48] may seem to have lower computational complexity due to lack of nIC; however, it suffers from significant performance degradation due to remaining nonlinear distortion, which will be presented in numerical results. Besides, it should be noted that nonlinear distortion estimation for each symbol should not be repeated at every iteration when the tentative decisions for a specific symbol estimate remain the same. The total number of computations might be decreased by avoiding unnecessary nonlinear processing, which in turn, reduces the circuit power consumption. This is a common and necessary operation for the algorithms that employ nIC, such as Robust MMSE with nIC [65] and the proposed PU-IB-DFE + DA-SSD. Hence, the energy efficiency of the system can be improved further. Furthermore, to estimate the received distortion in a multi-user MIMO scenario, the estimated distortion signal in the frequency domain must be multiplied by each user's intended/leakage channel. This operation requires $\mathcal{O}(U^2)$ real multiplications for Robust-MMSE with nIC [65] and $\mathcal{O}(\mu_B U^2)$ for the proposed algorithm. On the other hand, the proposed IB-DFE algorithm for multi-user MIMO scenarios has lower complexity required for frequency domain equalization than the MMSE-type receivers, which require matrix inversion for each FFT bin. In Robust MMSE algorithms, the equalization matrix is calculated after performing matrix inversion, requiring $\mathcal{O}\left(U^3\right)$ multiplications, followed by matrix multiplication, requiring $\mathcal{O}\left(U^3\right)$ multiplications. However, the proposed algorithm only requires scalar multiplications for each sampling branch, resulting in $\mathcal{O}(\mu_B U)$ multiplications due to the reduced effective dimension enabled by analog beamforming. It is important to acknowledge that the proposed algorithm necessitates computation for each sampling beach, resulting in a complexity increase by μ_B. Nonetheless, it should be noted

Table 5.2: Complexities of Baseline Schemes

	Distortion estimation via NN	Effective NL interference estimation	Distortion cancellation	Matrix inversion	Matrix multiplication for equalization	MVDR combiner
Robust MMSE Robust [47]	-	-	-	$\mathcal{O}\left(U^3\right)$	$\mathcal{O}\left(U^3\right)+\mathcal{O}\left(U^2\right)$	-
Robust MMSE with nIC [64]	$\mathcal{O}\left(U4L_qL_1\right)$	$\mathcal{O}\left(U^2\right)$	$\mathcal{O}\left(U\right)$	$\mathcal{O}\left(U^3\right)$	$\mathcal{O}\left(U^3\right)+\mathcal{O}\left(U^2\right)$	-
PU-IB-DFE with nIC +DA-SSD	$\mathcal{O}\left(U4L_qL_1\right)$	$\mathcal{O}\left(\mu_B U^2\right)$	$\mathcal{O}\left(\mu_B U\right)$	-	$\mathcal{O}\left(\mu_B U\right)$	$\mathcal{O}\left(\mu_B U\right)$

that the PU-IB-DFE architecture that is being suggested reduces the problem's dimensionality into scalar operations. Therefore, the additional complexity that arises from using multiple branches is not comparable with that of MMSE-type equalization. Consequently, it can be concluded that the proposed PU-IB-DFE with the nIC scheme provides superior performance with relatively low computational complexity.

5.5 Performance Evaluations

In this section, the performance of the proposed IB-DFE with nIC method is evaluated and compared with conventional state-of-the-art schemes via computer simulations.

5.5.1 Simulation Settings

In the considered scenario, a BS with $K = 64$ antennas is serving $U = 8$ mobile users, which randomly reside in 8 angular sectors distributed over $[-45°, 45°]$, and a single RF chain is dedicated to a single user. A spatially correlated channel model similar to that in [75] and [3], containing 5 MPCs per-user at fractional delays, which are distributed randomly over 16 symbol length, is considered. Each MPC is formed in a clustered manner, where each cluster contains additional μ channel taps at fractional delays around its center. Relative powers of each cluster are 0, -3, -5, -10 and -20 dB, Rician factor of MPCs are set to 10 dB, and angular spread of each MPC

cluster is 2°. The first three strong MPCs originate from the same angular sectors as the intended user, and the remaining MPCs are received over $[-45^{\circ}, 45^{\circ}]$.

In data transmission, the block length is selected to be $N_D = 8100$. A *root raised cosine* filter with 0.3 roll-off factor and $\mu = 8$ is used for p_n. In the simulations, all users are assumed to be synchronized; however, even for asynchronous cases, channel estimation, and equalization will handle the synchronization. For NPL, a training sequence of length 16384 symbols is sent before data transmission and is not repeated during the transmission phase. The number of neurons in the NN structure is 30, and memory length is chosen as ± 3. In this study, the Saleh model, [37], is employed in simulations, output of the PA, $\tilde{x}_n = |\tilde{x}_n| e^{j\phi_n} e^{j\theta_n}$, can be expressed as $|\tilde{x}_n| = \frac{g_0 |x_n|}{1+(|x_n|/A_{sat})^2}$, $\theta_n = \frac{\alpha |x_n|^2}{1+\beta |x_n|^2}$, where $g_0 = 2$, $A_{sat} = 1$, $\alpha_0 = 2$ and $\beta_0 = 1$ are the model parameters [33].

5.5.2 Simulation Results

5.5.2.1 SDNR Analysis Results

In this section, the performance of PU-IB-DFE is evaluated based on the analytical SDNR expression derived in Section 5.4.5. Firstly, we verify the analytical SDNR expression obtained in Section 5.4.5 by comparing it with the residual SDNR values (after PU-IB-DFE with nIC) based on link-level simulations via Monte Carlo (MC) trials. In Figure 5.7, CDF of the received SDNR of users for different channel realizations are presented for both one, $\mu_B = 1$, (PU-IB-DFE + DA-SSD(1)) and four, $\mu_B = 4$, (PU-IB-DFE + DA-SSD(4)) sampling branches, where output backoff (OBO) of PA model is set to 4 dB. OBO is defined as OBO $\triangleq \frac{P_{\text{max}}}{P_{\text{av}}}$, where P_{max} and P_{av} are maximum output power that PA can supply, and average transmit power at the current operating point, respectively. The results demonstrate a close match between the analytical SDNR curves and the empirical outcomes. Additionally, it is evident that the adoption of fractional sampling can lead to a significant improvement in the SDNR of the system. Our investigation reveals that for specific channel instances, the PU-IB-DFE technique with $\mu_B = 1$ experiences a severe deterioration in the received SDNR, which results in a substantial number of bit errors beyond the FEC's error

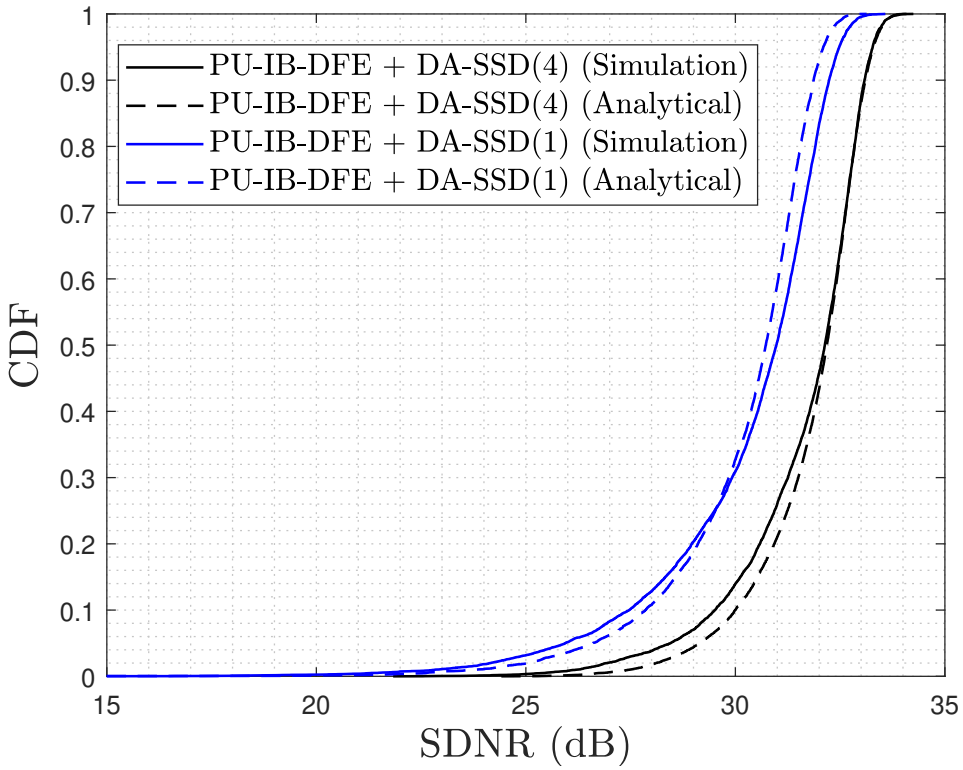


Figure 5.7: CDF of the received SDNR for the proposed PU-IB-DFE method with and without employing FDE bank.

correction capability. However, using different sampling instances avoids such a noticeable decrease in received SDNR; hence, overall system performance is enhanced remarkably. For instance, the probability of SNDR value being less than 25 dB, which can be considered as an outage threshold for 256-QAM transmission, is close to 3% for PU-IB-DFE + DA-SSD(1). This presents the frequency of the outage event, which indicates the significance of the distortion amplification. On the other hand, the CDF tail of PU-IB-DFE + DA-SSD(4) method starts at relatively higher SNDR values; hence, an outage event does not frequently occur for this method, which results in improved BER performance. However, the probability of SDNR being less than the threshold is less than 0.5%, which indicates performance improvement achieved by the detector.

5.5.2.2 Monte Carlo Simulations Results

In this section, BER performances of the considered methods for 256 QAM are evaluated and compared via MC simulations over 5000 channel realizations, where the instantaneous channel is estimated via the least squares method. In MC simulations, LDPC is used for channel coding with a code rate $r = 8/9$. In the simulations, Robust MMSE [48] and Robust MMSE with nIC [65] receivers are also included. In addition, the approximate BER expression in (5.65) is verified via different setups.

In Figure 5.8 (a), BER results for MC trials are shown for 4 dB OBO. It can be seen that PU-IB-DFE + DA-SSD(1) without fractional sampling suffers from severe error floor due to distortion amplification phenomenon; however, the proposed method with DA-SSD(4) decreases the floor to a lower level. The proposed PU-IB-DFE + DA-SSD(4) improves the performance by using channel diversity via MVDR combining. The proposed MVDR combiner assigns larger combining weights to the sampling branches, which does not experience distortion, by exploiting the distortion correlation matrix $\mathbf{R}_{\eta}$. Furthermore, it can also be observed that the BER performance of the proposed per-user IB-DFE algorithm is very close to the state-of-the-art methods, namely Robust MMSE [48] and Robust MMSE with nIC [65]. MMSE-type filtering requires matrix inversion for each frequency bin, which is avoided in the proposed receiver thanks to per-user processing enabled by spatial pre-processing. In Figure 5.8, implementation of the single user post-distortion method in [33] is also considered. It can be seen that the method in [33] suffers from significant performance loss due to lack of IUI cancellation, and the proposed nIC method is necessary to perform enhanced multi-user detection.

In Figure 5.8 (b), uncoded BER curves of PU-IB-DFE without any nIC and DA-SSD for $\mu_B = 1, 4$ for the same scenario are demonstrated by using analytical SDNR expression in Section 5.4.5. In addition, an uncoded analytical BER curve for the case of linear PA is also obtained for the sake of completeness. It can be observed that the BER of the coded case has a similar tendency and match with the analytical uncoded BER curves. The reason is that burst errors due to distortion amplification, which cannot be avoided with the use of such a high code rate ($r = 8/9$), are the primary source of overall bit errors. FEC is only effective in compensating the outliers due to

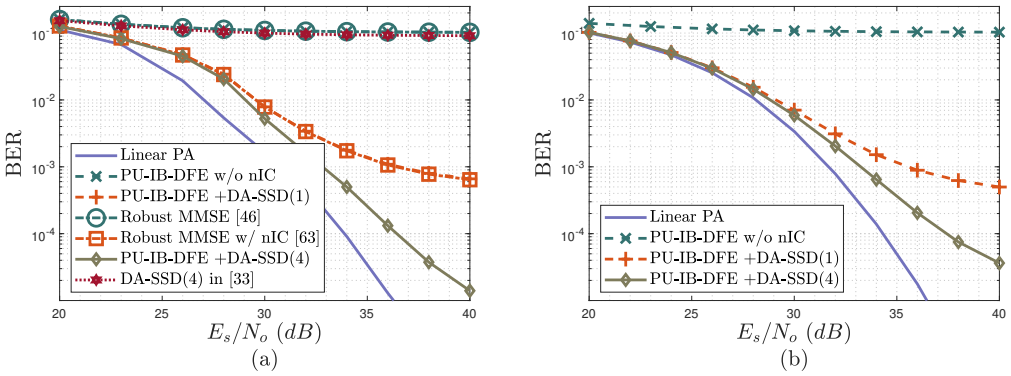


Figure 5.8: BER vs. E_s/N_o curves for 256 QAM for 4 dB OBO (a) numerical results with code rate $r = 8/9$ (b) analysis results.

residual nonlinear distortion after nIC.

BER results for different PA operating points, which is indicated by output backoff (OBO), and different constellation orders, namely 265 and 1024 QAM constellations, are presented for 50 dB E_s/N_o in Figure 5.9. Lower OBO values correspond to the scenarios where PA is operated closer to saturation; hence, nonlinearity is more severe. Figure 5.9 (a), provides the numerical BER performances for PU-IB-DFE with DA-SSD(4), DA-SSD(1), and w/o nIC methods. It is observed that the proposed PU-IB-DFE + DA-SSD(4) exhibits robustness to power amplifier nonlinearities. Even for lower OBO levels, it can achieve superior BER performances compared to the methods which do not employ fractional sampling. In addition, for the 1024 QAM constellation, the proposed method provides better BER performance than PU-IB-DFE + DA-SSD(1) with 256 QAM. For both constellations, the proposed method can eliminate the error floor. However, DA-SSD(1) suffers from error floor even for a relatively linear PA operating region. Analytical BER results are shown in Figure 5.9 (b), where it can be observed that the analytical BER expression in (5.65) provides a good approximation for the practical system. The performed analysis is able to capture outage events stemming from SDNR drops due to distortion amplification.

BER performance of the proposed algorithm is also evaluated in a different scenario with a memory polynomial-based PA model, which is extracted from a PA and used in [3] as shown in Figure 5.10. A more dense communication system, where a BS with

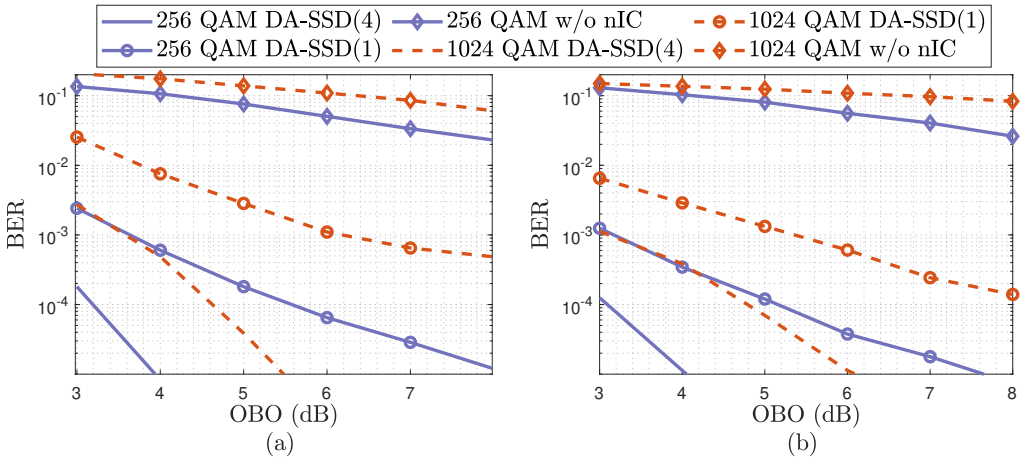


Figure 5.9: BER vs. OBO curves for 256 and 1024 QAM constellations (a) numerical BER results with code rate $r = 8/9$ (b) analytical BER results.

$K = 100$ antennas serves $U = 16$ users, is considered. In addition, for this case, alternative nonlinear basis functions (NBF), namely VS and memory polynomial (MP) based modeling, are considered for nonlinear modeling. Consistent with previous evaluations, the proposed PU-IB-DFE algorithm outperforms the state-of-the-art algorithms, as seen from Figure 5.10 (a). Even though the MP-based modeling has the same structure as the PA model, it suffers from performance degradation since it is shown in (5.32) that nonlinearity after downsampling contains a memory effect prior to the NBF. However, in the MP-based model, memory effects are considered after the nonlinearity. Hence, MP is not a proper choice for symbol-sampled processing. VS expansion-based method provides BER performance close to NN modeling since it contains cross-memory terms for NBF. However, it suffers from high computational complexity. Lastly, the BER curves based on SDNR analysis are shown in Figure 5.10 (b). Even though the analysis assumes a memoryless nonlinearity, it can accurately estimate the distortion amplification effects since the NN modeling compensates for memory effects. Therefore, analytical BER is consistent with the numerical results.

5.6 Conclusions

In this chapter, we developed a multi-user nIC method in the downsampled signal domain for multi-user hybrid MIMO beamforming framework. In the considered

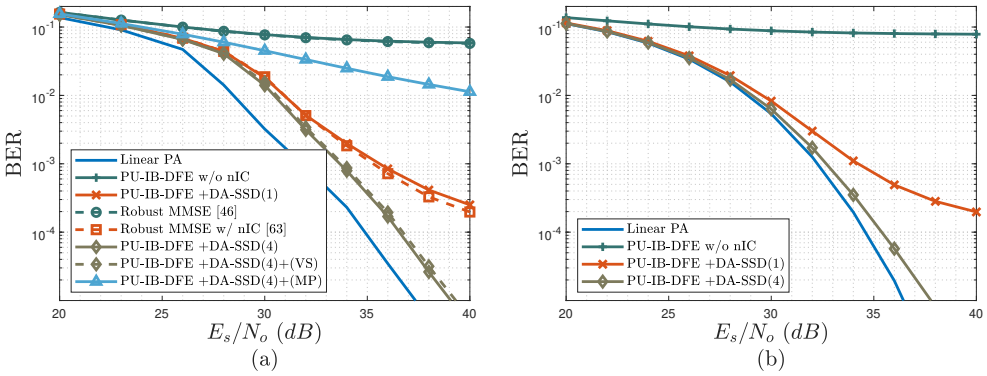


Figure 5.10: BER vs. E_s/N_o curves for 256 QAM for polynomial PA model in [3] (a) numerical results with code rate $r = 8/9$ (b) analysis results.

approach, the nonlinearly distorted IUI and nonlinear distortion component of the intended users are removed from the received signal so that a reduced complexity per-user IB-DFE receiver can be implemented. Then, fractionally spaced IB-DFE receiver outputs are combined intelligently via the distortion-aware detector to reduce the distortion amplification effects. Also, a theoretical framework is constructed to analyze the system performance under PA nonlinearities. It is observed that the proposed approach achieves superior performance compared to state-of-the-art methods, and the theoretical framework is verified by comparing the analytical results with those of MC simulations.

CHAPTER 6

ANALYTICAL NONLINEAR DISTORTION CHARACTERIZATION AND POWER ALLOCATION FOR MULTI-USER FREQUENCY-SELECTIVE MASSIVE MIMO CHANNELS

6.1 Introduction

In this section of the thesis, our attention is directed towards the downlink communication aspect, where the occurrence of nonlinear distortion is observed at the PAs of the BSs. We delve into analyzing the characteristics of received distortion in fully digital massive MIMO systems, examining the influence of different system parameters on the received distortion. It is demonstrated that the spatial correlation of distortion has significant effects, particularly in the context of massive MIMO systems. The analysis results highlight the importance of spatial correlation in distortion modeling.

Building upon the insights gained from the analysis, a power allocation scheme is proposed, taking the spatial correlation of distortion into account. Through extensive evaluations, we show that the proposed power allocation approach surpasses the performance of methods that overlook the spatial correlation of distortion. This finding underscores the significance of incorporating spatial correlation in the power allocation process to optimize system performance.

6.2 Literature Review

The Bussgang decomposition technique has emerged as a prominent tool extensively employed for analyzing the impact of hardware nonlinearities on communication systems [59]. By decoupling the nonlinearly distorted signal into two uncorrelated

components, namely the useful component and the uncorrelated distortion component, the Bussgang decomposition enables a comprehensive investigation of these effects [81]. Various studies in the literature have utilized the Bussgang decomposition to examine the consequences of hardware distortion, treating it as an additional additive noise [82,83]. For MIMO systems, researchers have performed the Bussgang decomposition individually for each antenna, modeling the distortion as spatially and temporally uncorrelated noise [82–85]. Nevertheless, although the Bussgang decomposition offers valuable insights, it falls short in adequately representing the spectral characteristics of the nonlinear distortion. Subsequently, in order to attain a modeling approach that better captures the power spectral density (PSD) of the distortion term, researchers have devised alternative methodologies [22,86,87]. Notably, several studies [48,86,87] have employed polynomial expansion techniques to model the spectral characteristics of nonlinear distortion on an individual basis for each antenna. These methodologies enable the estimation of distortion power at various frequency components. This modeling approach has been found to yield accurate results, particularly in scenarios where the number of users is relatively large. In such cases, the inputs to the antennas tend to exhibit decorrelation as the number of users increases, thereby enhancing the efficacy of the modeling approach [88]. Consequently, these studies suggest that as the number of antennas increases, the detrimental effects of nonlinear distortion diminish.

However, the spatially uncorrelated distortion noise assumption is not accurate for massive MIMO systems, where the number of antennas is much larger than the number of users. For instance, in [89] and [90], it is shown for single-user cases that the distortion signal is also beamformed towards the intended user. This means that the distortion is spatially correlated, and it is coherently combined at the receiver side. In [89], the spatial correlation of the distortion signal is exploited to devise a precoding vector that minimizes the received distortion power. Furthermore, this study optimizes the transmit power to mitigate the adverse effects of nonlinear distortion. In the works of [22, 23], the spatial distribution of radiated distortion power is examined for both multi-user and frequency-selective channel scenarios, employing the Ito-Hermite polynomial expansion representation to model the power spectral density (PSD) of the nonlinear distortion signal. Numerical simulations demonstrate that

the distortion is not only beamformed towards the intended user direction but also in other directions. Additionally, it is observed that, as the number of users increases, the number of directions towards which the distortion is beamformed grows exponentially. Consequently, this leads to a reduction in received distortion power at the intended user terminal. Moreover, when the number of users is sufficiently large, the distortion exhibits isotropic radiation, thereby validating the assumption of spatially uncorrelated distortion for such cases [82–85]. The framework introduced in [23] is adopted in [83] to obtain a semi-analytical semi-analytical BER expression. In [83], the autocorrelation matrix of the distortion across the antenna array is calculated for the given channel state information so that spatial correlation is taken into account in BER analysis. In [91], spatial correlation of the distortion is considered for correlated MIMO channels, and the signal-to-distortion-plus-interference-plus-noise ratio is derived for given CSI. However, this study does not present the effects of spatial diversity and frequency selectivity.

In the existing literature, several studies have been conducted with the objective of optimizing the transmit power of BSs. One fundamental approach is to employ power backoff techniques to operate the PAs within their linear region, as discussed in [92]. Furthermore, in the work presented by authors in [89], the transmit power is adaptively adjusted based on the instantaneous channel state information (CSI). Similarly, a power allocation framework for distributed MIMO systems considering hardware impairments is developed in [93]. This framework, similar to the approach in [89], tackles the power allocation problem by utilizing the instantaneous CSI. However, this study primarily focuses on employing maximum ratio transmission precoding, which has limited capability in suppressing inter-user interference, for digital preprocessing.

Moreover, joint precoding and power allocation frameworks have been proposed in [94] and [95]. Nonetheless, these investigations concentrate on MIMO systems operating under high-load conditions. As a consequence, these studies make the assumption of spatially uncorrelated distortions in their optimization problems. It is worth noting that this assumption may not hold in scenarios where there is a substantial correlation between distortions introduced by different PAs.

Thus, while previous research has made valuable contributions to power optimization in BSs, there is still a need for comprehensive approaches that address the challenges posed by hardware impairments in MIMO systems, considering both spatial correlation among distortions and inter-user interference suppression capabilities of digital pre-processing techniques.

6.2.1 Contributions

This chapter presents a comprehensive analysis that derives a fully closed-form expression for the received distortion power, which, to the best of our knowledge, has not been previously conducted in the literature. The derived expression captures the effects of spatial diversity and frequency selectivity on the received distortion. Notably, it is a function of slowly varying parameters, primarily the user locations. It is crucial to consider spatial correlation when dealing with massive MIMO systems, where the number of antennas is extremely large. Neglecting distortion correlation in such systems leads to significant mis-evaluation.

By utilizing the developed analysis framework, a power allocation problem is formulated while considering distortion correlation. The proposed power allocation scheme demonstrates superior performance compared to methods that overlook the spatial correlation of distortions, particularly in terms of spectral efficiency. It is worth highlighting that incorporating distortion correlation in the power allocation problem is essential for optimizing the system's overall performance and ensuring accurate power allocation decisions.

6.3 System Model and Distortion Characterization

6.3.1 System Model

This chapter considers a cellular system where a BS with M antennas is serving U single-antenna users. A frequency-selective block fading channel is assumed such that channel realization within each block is stationary. In this chapter, we denote signals in the discrete domain; however, to approximate the continuous-time operations,

we consider oversampled signals with oversampling factor μ. According to these assumptions, the transmitted orthogonal frequency-division multiplexing (OFDM) modulated signal can be expressed in the time domain as

$$\mathbf{x}[n] = \sum_{k=-N_s/2+1}^{N_s/2} \sum_{u=1}^{U} \mathbf{w}_{u,k} a_{u,k} e^{j\frac{2\pi}{\mu N_s} kn}, \tag{6.1}$$

where $\mathbf{w}_{u,k} \in \mathbb{C}^{M\times 1}$ is the precoding vector at the k^{th} subcarrier for u^{th} user, $\{a_k\}$'s are the QAM digital information signals, which are non-zero for $k = -N_s/2 + 1, \ldots, N_s/2$, and $N = \mu N_s$ is the symbol duration.

In order to suppress inter-user interference, zero-forcing (ZF) type precoding is employed and the precoder vector $\mathbf{w}_{u,k}$ can be expressed as

$$\mathbf{w}_{u,k} = \sqrt{\rho_u}\bar{\mathbf{w}}_{u,k}, \tag{6.2}$$

where ρ_u is the power allocation coefficient of u^{th} user, and

$$\bar{\mathbf{w}}_{u,k} = \alpha_u \left[\mathbf{H}_k \left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\right)^{-1}\right]_{:,u}, \tag{6.3}$$

where $\mathbf{H}_k = [\mathbf{h}_{1,k}, \ldots, \mathbf{h}_{U,k}]$ is the channel matrix, whose columns are the channel vectors of each user in the frequency domain with the auto-correlation matrix $\mathbb{E}\left[\mathbf{h}_{u,k}\mathbf{h}_{u,k}^{\mathrm{H}}\right] \triangleq \mathbf{R}_u$, and $[A]_{:,k}$ denotes the k^{th} column of $\mathbf{A}$ matrix. The large-scale fading coefficient is included in $\mathbf{R}_u$ such that $\frac{1}{M}\mathrm{Tr}(\mathbf{R}_u) = \beta_u$. The coefficient α_u is the scaling factor, which normalizes the average input power of each PA as $\mathbb{E}\left[\|\bar{\mathbf{w}}_{u,k}\|^2\right] = M$, and the average power constraint can be used to obtain a closed-form expression for α_u as

$$\begin{aligned}\mathbb{E}\left[\|\bar{\mathbf{w}}_{u,k}\|^2\right] &= \alpha_u^2 \mathbb{E}\left[\left[\left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\right)^{-1}\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\right)^{-1}\right]_{u,u}\right] \\ &= \alpha_u^2 \mathbb{E}\left[\left[\left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\right)^{-1}\right]_{u,u}\right].\end{aligned} \tag{6.4}$$

In order to find α_u, firstly, we focus on the term $\mathbb{E}\left[\left[\left(\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k\right)^{-1}\right]_{u,u}\right]$, and we can express the correlated channel by using the eigenvalue decomposition of $\mathbf{R}_u = \mathbf{B}_u\mathbf{\Lambda}_u\mathbf{B}_u^{\mathrm{H}}$, where $\mathbf{B}_u$ is unitary and $\mathbf{\Lambda}_u = \mathrm{diag}\{\lambda_{u,1}, \ldots, \lambda_{u,M}\}$ as

$$\mathbf{h}_{k,u} = \mathbf{B}_u \mathbf{\Lambda}_u^{\frac{1}{2}} \tilde{\mathbf{h}}_{k,u}, \tag{6.5}$$

where $\tilde{\mathbf{h}}_{k,u} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. Then, we can express the $(i,r)^{\text{th}}$ element of the matrix $\mathbf{H}_k^{\text{H}}\mathbf{H}_k$ as

$$\left[\mathbf{H}_k^{\text{H}}\mathbf{H}_k\right]_{(i,r)} = \tilde{\mathbf{h}}_{k,i}^{\text{H}} \mathbf{\Lambda}_i^{\frac{1}{2}} \mathbf{B}_i^{\text{H}} \mathbf{B}_r \mathbf{\Lambda}_r^{\frac{1}{2}} \tilde{\mathbf{h}}_{k,r}^{\text{H}}. \tag{6.6}$$

In order to obtain a closed-form expression of the expectation in (6.4), the spatially uncorrelated channel model will be considered to exploit the properties of a Wishart matrix, where $\mathbf{\Lambda}_u = \beta_u \mathbf{I}_M$ and $\mathbf{B}_u = \mathbf{I}_M$. Then, (6.4) can be re-written as

$$\begin{aligned} \mathbb{E}\left[\|\bar{\mathbf{w}}_{u,k}\|^2\right] &= \alpha_u^2 \mathbb{E}\left[\left[\left(\mathbf{H}_k^{\text{H}}\mathbf{H}_k\right)^{-1}\right]_{u,u}\right] = \alpha_u^2 \mathbb{E}\left[\left[\left(\mathbf{\Lambda}\tilde{\mathbf{H}}_k^{\text{H}}\tilde{\mathbf{H}}_k\mathbf{\Lambda}\right)^{-1}\right]_{u,u}\right] \\ &= \alpha_u^2 \left[\mathbf{\Lambda}^{-1}\mathbb{E}\left[\left(\tilde{\mathbf{H}}_k^{\text{H}}\tilde{\mathbf{H}}_k\right)^{-1}\right]\mathbf{\Lambda}^{-1}\right]_{u,u}, \end{aligned} \tag{6.7}$$

where $\mathbf{H}_k = \tilde{\mathbf{H}}_k\mathbf{\Lambda}$ and $\mathbf{\Lambda} = \text{diag}(\sqrt{\beta_1}, \ldots, \sqrt{\beta_U})$. By using the identity in [96], the scaling coefficient can be determined as $\alpha_u = \sqrt{\beta_u M(M-U)}$ to satisfy the constraint $\mathbb{E}\left[\|\bar{\mathbf{w}}_{u,k}\|^2\right] = M$. Then the overall precoding vector for user u can be expressed as

$$\bar{\mathbf{w}}_{u,k} = \sqrt{\beta_u M(M-U)} \sum_{u'=1}^{U} \mathbf{h}_{u',k} \bar{h}_{u',u}, \tag{6.8}$$

where $\bar{h}_{u',u} = \left[\left(\mathbf{H}_k^{\text{H}}\mathbf{H}_k\right)^{-1}\right]_{u',u}$. For each antenna element m, we can define the scalar precoding coefficient as

$$\bar{w}_{u,k}^{(m)} = \sqrt{\beta_u M(M-U)} \sum_{u'=1}^{U} h_{u',k}^{(m)} \bar{h}_{u',u}, \tag{6.9}$$

and the transmit power constraint for each antenna element can be re-formulated as

$$\mathbb{E}\left[\sum_{u'=1}^{U} h_{u',k}^{(m)} \bar{h}_{u',u}^{(m)} \sum_{u''=1}^{U} h_{u'',k}^{(m)} \bar{h}_{u'',u}\right] = \frac{1}{\beta_u M(M-U)}. \tag{6.10}$$

For analytical simplicity, we will assume that $\bar{h}_{u',u}$'s are deterministic and $\bar{h}_{u,u} \gg \bar{h}_{u',u}$ for $u \neq u'$ for large M. This is also a valid approximation for the spatially

correlated channels since user channels become nearly orthogonal if users are located at different angular sectors for large M. By using this approximation, (6.10) can be re-written as

$$\mathbb{E}\left[\sum_{u'=1}^{U} h_{u',k}^{(m)} \bar{h}_{u',u}^{(m)} \sum_{u''=1}^{U} h_{u'',k}^{(m)} \bar{h}_{u'',u}\right] \approx \mathbb{E}\left[|h_{u,k}^{(m)}|^2\right] \bar{h}_{u,u}^2 = \frac{1}{\beta_u M(M-U)}. \quad (6.11)$$

Then $\bar{h}_{u,u}$ can be approximated as

$$\bar{h}_{u,u} = \frac{1}{\sqrt{\beta_u^2 M(M-U)}}. \quad (6.12)$$

Before radiating through the air, the modulated signal is amplified via PAs to increase the signal power to a sufficient level. However, the nonideality of the PA hardware causes nonlinear distortion. For the sake of generality, we represent the nonlinearity for m^{th} antenna element as $\tilde{x}^{(m)}[n] = \Psi(x^{(m)}[n])$. The transmitted signal from the antenna array becomes $\tilde{\mathbf{x}}[n] = [\tilde{x}^{(1)}[n], \ldots, \tilde{x}^{(M)}[n]]^T$. Then the received signal at the single-antenna user terminal can be expressed as

$$y_u[n] = \sum_{l=0}^{L_u-1} \mathbf{h}_u^{\text{H}}[\tau_{u,l}] \tilde{\mathbf{x}}[n - \tau_{u,l}] + \eta_u[n], \quad (6.13)$$

where L_u is the number of significant taps of user u, and $\mathbf{h}[\tau_{u,l}]$ is the channel vector in the time domain, which can be equivalently represented in the frequency domain as

$$\mathbf{h}_k = \sum_{l=0}^{L_u-1} \mathbf{h}[\tau_{u,l}] e^{-j\frac{2\pi}{\mu N_s} k \tau_{u,l}}, \quad (6.14)$$

for $k = -N/2 + 1, \ldots, N/2$, and $\eta_u[n]$ is the additive white Gaussian noise with variance σ_η^2, and $\tau_{u,l}$ is the discrete time index for the l^{th} multipath component of u^{th} user.

6.3.2 Spectral characterization of the nonlinear distortion under frequency-selective channels

It is known that if the number of subcarriers is high enough, the distribution of OFDM modulated signals can be approximated by the circularly symmetric complex Gaussian distribution thanks to the central limit theorem. By exploiting this fact, we can

adopt a complex Itô-Hermite polynomial representation to acquire the second-order statistics of the nonlinear distortion in the frequency domain [23]. Consequently, the nonlinearly transmitted signal is re-written by focusing on only the third-order nonlinearity as

$$\tilde{x}^{(m)}[n] = \underbrace{\alpha_{1m}x^{(m)}[n]}_{\text{Desired signal : } z^{(m)}[n]} + \underbrace{\alpha_{3m}\sigma^3_{x^{(m)}}H_3\left(\frac{x^{(m)}[n]}{\sigma_{x^{(m)}}}\right)}_{\text{Distortion signal : } d^{(m)}[n]}, \tag{6.15}$$

where $H_3(\cdot)$ is the third-order Hermite basis function

$$H_3(x) \triangleq \sum_{i=0}^{1}(-1)^i i! \binom{2}{i}\binom{1}{i} x|x|^{2-2i}, \tag{6.16}$$

and $\sigma_{x^{(m)}}$ is the standard deviation of $x^{(m)}$.

This representation has the following useful property, which enables spectral analysis. The desired and the distortion signals are uncorrelated $\mathbb{E}\left[z^{(m)}[n](d^{(m')}[n-n'])^*\right] = 0\ \forall n', m, m'$, so that we can write the cross-spectral density as the summation of two uncorrelated signal terms as

$$S_{\tilde{x}^{(m)}\tilde{x}^{(m')}}[k] = \alpha_{1m}\alpha^*_{1m'}S_{x^{(m)}x^{(m')}}[k] + S_{d^{(m)}d^{(m')}}[k], \tag{6.17}$$

where $S_{x^{(m)}x^{(m')}}[k] = \mathcal{F}\left\{R_{x^{(m)}x^{(m')}}[n-n']\right\}$ is the cross-spectral density for the desired term, which is defined by the Fourier transform of cross-correlation function $R_{x^{(m)}x^{(m')}}[n-n'] = \mathbb{E}\left[x^{(m)}[n]x^{(m')}[n']\right]$. Similarly, $S_{d^{(m)}d^{(m')}}[k]$ is the Fourier transform of the cross-correlation function of the distortion term, which can be expressed by exploiting the Gaussianity of the PA input signals as

$$R_{d^{(m)}d^{(m')}}[n-n'] = 2\alpha_{3m}\alpha^*_{3m'}R_{x^{(m)}x^{(m')}}[n-n']\left|R_{x^{(m)}x^{(m')}}[n-n']\right|^2. \tag{6.18}$$

By using the expression in (6.18), cross-spectral density $S_{d^{(m)}d^{(m')}}[k]$ can be obtained as

$$S_{d^{(m)}d^{(m')}}[k] = 2\alpha_{3m}\alpha^*_{3m'}\left(S_{x^{(m)}x^{(m')}}[k] * S_{x^{(m)}x^{(m')}}[k] * S^*_{x^{(m)}x^{(m')}}[-k]\right), \tag{6.19}$$

where $*$ denotes the convolution operation. The cross-spectral density for the desired term $S_{x^{(m)}x^{(m')}}[k]$ can be written for the given channel state information (CSI) as

$$S_{x^{(m)}x^{(m')}|\mathbf{H}_k}[k] = \sum_{u=1}^{U} \rho_u \bar{w}_{u,k}^{(m)} \left(\bar{w}_{u,k}^{(m')}\right)^* \tag{6.20}$$

Cross-spectral density for the distortion term can be written by using the cross-spectral density for the desired term as

$$\begin{aligned} S_{d^{(m)}d^{(m')}|\mathbf{H}_k}[k] = \frac{2}{N_s^2}|\alpha_3|^2 \sum_{k'=-N/2+1}^{N/2} \sum_{k''=-N/2+1}^{N/2} \\ \left[\left(\sum_{u=1}^{U} \rho_u \bar{w}_{u,k'}^{(m)} \left(\bar{w}_{u,k'}^{(m')}\right)^* \right) \left(\sum_{u'=1}^{U} \rho_{u'} \bar{w}_{u',k''}^{(m)} \left(\bar{w}_{u',k''}^{(m')}\right)^* \right)^* \right. \\ \left. \left(\sum_{u''=1}^{U} \rho_{u''} \bar{w}_{u'',k'+k''-k}^{(m)} \left(\bar{w}_{u'',k'+k''-k}^{(m')}\right)^* \right) \right] \end{aligned} \tag{6.21}$$

for $|k'| \leq N_S/2$, $|k''| \leq N_S/2$ and $|k' + k'' - k| \leq N_S/2$.

6.4 Achievable Rate Analysis

In this section, an achievable information rate expression is derived. To this end, coherent distortion beamforming is taken into account and the received distortion power is obtained to calculate the SDNR. In addition, the derived expression contains power allocation coefficients so that they will further be utilized to optimize the spectral efficiency of the massive MIMO system.

The symbol estimates can be expressed by using Itô-Hermite polynomial representation as

$$y_{\tilde{u}}[n] = \sum_{l=0}^{L_{\tilde{u}}-1} \mathbf{h}_{\tilde{u}}^{\mathrm{H}}[\tau_{\tilde{u},l}]\mathbf{x}[n-\tau_{\tilde{u},l}] + \sum_{l=0}^{L_{\tilde{u}}-1} \mathbf{h}_{\tilde{u}}^{\mathrm{H}}[\tau_{\tilde{u},l}]\mathbf{d}[n-\tau_{\tilde{u},l}] + \eta_{\tilde{u}}[n], \tag{6.22}$$

where $\mathbf{d}[n] = \left[d^{(1)}[n], \ldots, d^{(M)}[n]\right]^{\mathrm{T}}$, which can be represented in frequency domain as

$$y_{\tilde{u},k} = \underbrace{\mathbf{h}_{\tilde{u},k}^{\mathrm{H}}\mathbf{w}_{\tilde{u},k}a_{\tilde{u},k}}_{\text{Desired symbol}} + \underbrace{\sum_{u=1,u\neq\tilde{u}}^{U} \mathbf{h}_{\tilde{u},k}^{\mathrm{H}}\mathbf{w}_{u,k}a_{u,k}}_{\text{Multi-user interference}} + \underbrace{\mathbf{h}_{\tilde{u},k}^{\mathrm{H}}\mathbf{d}_k}_{\text{Received nonlinear distortion}} + \eta_{\tilde{u},k}. \tag{6.23}$$

By using the received signal expression in (6.23) and use-and-then-forget (UatF) bound [97], spectral efficiency for user $\tilde{u}$ $\mathrm{SE}_{\tilde{u}}$ can be given by

$$\mathrm{SE}_{\tilde{u}} = \log_2\left(1 + \mathrm{SDNR}_{\tilde{u}}\right), \tag{6.24}$$

where signal-to-distortion-plus-noise ratio $\mathrm{SDNR}_{\tilde{u}}$ is calculated as

$$\mathrm{SDNR}_{\tilde{u}}[k] = \frac{\rho_{\tilde{u}} \left|\mathbb{E}_{\mathbf{H}_k}\left[\mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \bar{\mathbf{w}}_{\tilde{u},k}\right]\right|^2}{\sum_{u=1}^{U} \rho_u c_{u,k} - \rho_{\tilde{u}} \left|\mathbb{E}_{\mathbf{H}_k}\left[\mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \bar{\mathbf{w}}_{\tilde{u},k}\right]\right|^2 + \mathbb{E}_{\mathbf{H}_k}\left[\left|\mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \mathbf{d}_k\right|^2\right] + \sigma_\eta^2}, \tag{6.25}$$

where $c_{u,k} = \mathbb{E}_{\mathbf{H}_k}\left[\left|\mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \bar{\mathbf{w}}_{u,k}\right|^2\right]$.

6.4.1 Spectral analysis of the received distortion

In this section, we will derive the spectral characteristics of the received distortion signal. The received power by the intended user $\tilde{u}$ at the subcarrier k is

$$S_{yy|\mathbf{H}_k}^{(\tilde{u})}[k] = \sum_{m=1}^{M} \sum_{m'=1}^{M} h_{\tilde{u},k}^{(m)} S_{\tilde{x}^{(m)}\tilde{x}^{(m')}|\mathbf{H}_k}[k] \left(h_{\tilde{u},k}^{(m')}\right)^*. \tag{6.26}$$

By utilizing the decomposition of $S_{\tilde{x}^{(m)}\tilde{x}^{(m')}}^{(\tilde{u})}[k]$ in (6.17), we can rewrite the PSD of the received signal as

$$S_{yy|\mathbf{H}_k}^{(\tilde{u})}[k] = S_{zz|\mathbf{H}_k}^{(\tilde{u})}[k] + S_{dd|\mathbf{H}_k}^{(\tilde{u})}[k], \tag{6.27}$$

where $S_{zz}[k]$ is the PSD of the desired term, and $S_{dd|\mathbf{H}_k}^{(\tilde{u})}[k] = \mathbb{E}\left[\left|\mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \mathbf{d}_k\right|^2 |\mathbf{H}_k\right]$.

The received distortion signal at $\tilde{u}$ can be written as

$$\begin{aligned} S_{dd|\mathbf{H}_k}^{(\tilde{u})}[k] =& \frac{2}{N_s^2}|\alpha_3|^2 \sum_{k'=-N/2+1}^{N/2} \sum_{k''=-N/2+1}^{N/2} \sum_{u=1}^{U}\sum_{u'=1}^{U}\sum_{u''=1}^{U} \rho_u \rho_{u'} \rho_{u''} \\ & \sum_{m=1}^{M}\sum_{m'=1}^{M} \left[\left(h_{\tilde{u},k}^{(m)}\right)^* \bar{w}_{u,k'}^{(m)} \left(\bar{w}_{u',k''}^{(m)}\right)^* \bar{w}_{u'',k'+k''-k}^{(m)}\right] \\ & \left[\left(h_{\tilde{u},k}^{(m')}\right)^* \bar{w}_{u,k'}^{(m')} \left(\bar{w}_{u',k''}^{(m')}\right)^* \bar{w}_{u'',k'+k''-k}^{(m')}\right]^* \end{aligned} \tag{6.28}$$

for $|k'| \leq N_S/2$, $|k''| \leq N_S/2$ and $|k' + k'' - k| \leq N_S/2$.

From (6.28), it is observed that the received distortion signal has three components. The first component is the result of the summation for $\tilde{u} = u = u' = u''$. In this case,

the distortion is coherently combined at the receiver location. The second component also yields a coherent combination of summation indices $\tilde{u} = u$, $u' = u''$, $u \neq u'$ and $\tilde{u} = u'$, $u = u''$, $u \neq u''$, which also yields coherent distortion combining. The last component includes the remaining terms, which correspond to an uncoherent combination of the distortion. Besides, this case can be considered as isotropic radiation of the distortion.

6.4.2 Calculation of distortion power for different components

In order to find a closed form expression for the received distortion power, spatially uncorrelated channel assumption will be exploited. One should note that this assumption yields accurate approximations if the angular spread is large enough for the one-ring scatter model for large antenna arrays, where the eigenvalue spread of the autocorrelation function is also large enough. Furthermore, the validity of the obtained closed-form expression in the case of spatially correlated channels will be shown numerically via simulations.

6.4.2.1 Power of the distortion term due to the intended user

The distortion power expression for this case is identical to the single-user scenario with frequency selectivity. The distortion power for this component $S^{(\tilde{u})}_{dd,1}[k]$ can be expressed as

$$S^{(\tilde{u})}_{dd,1|\mathbf{H}_k}[k] = \frac{2}{N_s^2}|\alpha_3|^2 \sum_{k'=-N/2+1}^{N/2} \sum_{k''=-N/2+1}^{N/2} \rho_{\tilde{u}}^3 S^{(\tilde{u})}_{dd,1a}[k, k', k''] \tag{6.29}$$

where

$$S^{(\tilde{u})}_{dd,1a}[k, k', k''] \triangleq \left| \sum_{m=1}^{M} \left[\left(h^{(m)}_{\tilde{u},k} \right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{\tilde{u},k''} \right)^* \bar{w}^{(m)}_{\tilde{u},k'+k''-k} \right] \right|^2 \tag{6.30}$$

for $|k'| \leq N_S/2$, $|k''| \leq N_S/2$ and $|k'+k''-k| \leq N_S/2$. However, we can exploit the massive spatial diversity achieved when having a large number of antennas to obtain a deterministic closed-form expression for the received signal and distortion power. Consider $S_{dd,1a}[k, k', k'']$ term

$$S^{(\tilde{u})}_{dd,1a|\mathbf{H}_k}[k,k',k''] = \sum_{m=1}^{M} \left(h^{(m)}_{\tilde{u},k}\right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{\tilde{u},k''}\right)^* \bar{w}^{(m)}_{\tilde{u},k'+k''-k}, \tag{6.31}$$

which is the summation of many independent random variables; hence, the law of large numbers (LLN) implies that

$$\frac{S^{(\tilde{u})}_{dd,1a|\mathbf{H}_k}[k,k',k'']}{M} \to \mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k}\right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{\tilde{u},k''}\right)^* \bar{w}^{(m)}_{\tilde{u},k'+k''-k}\right] \tag{6.32}$$

when $M \to \infty$. For a large but finite number of antennas, we can use the LLN to approximate the random variable $S_{dd,1a}[k,k',k'']$ by a single deterministic scalar as

$$S^{(\tilde{u})}_{dd,1a|\mathbf{H}_k}[k,k',k''] \approx M\mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k}\right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{\tilde{u},k''}\right)^* \bar{w}^{(m)}_{\tilde{u},k'+k''-k}\right], \tag{6.33}$$

which can be re-written by expanding the pre-coding vector $\bar{w}^{(m)}_{u,k}$ as

$$\begin{aligned} &S_{dd,1a|\mathbf{H}_k}[k,k',k''] \approx M\left(\beta_{\tilde{u}} M(M-U)\right)^{\frac{3}{2}} \\ &\mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k}\right)^* \left(\sum_{v=1}^{U} h^{(m)}_{v,k'} \bar{h}_{v,\tilde{u}}\right) \left(\sum_{v'=1}^{U} h^{(m)}_{v',k''} \bar{h}_{v',\tilde{u}}\right)^* \left(\sum_{v''=1}^{U} h^{(m)}_{v'',k'+k''-k} \bar{h}_{v'',\tilde{u}}\right)\right] \end{aligned} \tag{6.34}$$

By using the approximation in (6.11), $S_{dd,1a|\mathbf{H}_k}[k,k',k'']$ can be simplified as

$$S_{dd,1a|\mathbf{H}_k}[k,k',k''] \approx M \frac{\left(\beta_{\tilde{u}} M(M-U)\right)^{\frac{3}{2}}}{\left(\beta^2_{\tilde{u}} M(M-U)\right)^{\frac{3}{2}}} \underbrace{\mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k}\right)^* h^{(m)}_{\tilde{u},k'} \left(h^{(m)}_{\tilde{u},k''}\right)^* h^{(m)}_{\tilde{u},k'+k''-k}\right]}_{\mathcal{E}_{k,k',k''}}, \tag{6.35}$$

Calculation of this expectation is not straightforward and depends on the channel's frequency selectivity. To find the correlation between the subcarriers' channels in (6.35), consider the equivalent representation of the expectation as

$$\begin{aligned} \mathcal{E}_{k,k',k''} &= \mathbb{E}\left[h^{(m)}_{k} \left(h^{(m)}_{k'}\right)^* \left(h^{(m)}_{k''}\right)^* h^{(m)}_{k'+k''-k}\right] \\ &= \sum_{\substack{\tau_l,\tau_{l'},\\ \hat{\tau}_l,\hat{\tau}_{l'}\in\mathcal{L}^{\tilde{u}}}} \mathbb{E}\left[h^{(m)}_{\tilde{u}}[\tau_l] \left(h^{(m)}_{\tilde{u}}[\tau_{l'}]\right)^* \left(h^{(m)}_{\tilde{u}}[\hat{\tau}_l]\right)^* h^{(m)}_{\tilde{u}}[\hat{\tau}_{l'}]\right] \\ &\qquad e^{-j\frac{2\pi}{N}k\tau_l} e^{j\frac{2\pi}{N}k'\tau_{l'}} e^{j\frac{2\pi}{N}k''\hat{\tau}_l} e^{-j\frac{2\pi}{N}(k'+k''-k)\hat{\tau}_{l'}}, \end{aligned} \tag{6.36}$$

where $\mathcal{L}^{\tilde{u}} \subset \mathbb{Z}^{L^{\tilde{u}}}$ is the set of MPC indices of the intended user, $[\mathcal{L}^{\tilde{u}}]_l \in [0, \tau_{\tilde{u},\max}]$, and for the sake of notational simplicity we have removed subscript u for time delays.

Note that the different multipath taps are uncorrelated; therefore, only a subset of the terms in (6.36) are non-zero, which can be expressed as

$$\begin{aligned}\mathcal{E}_{k,k',k''} &= \sum_{\tau_l \in \mathcal{L}^{\tilde{u}}} \mathbb{E}\left[\left|h_{\tilde{u}}^{(m)}[\tau_l]\right|^4\right] \\ &+ \sum_{\tau_l \in \mathcal{L}^{\tilde{u}}} \sum_{\substack{\tau_{l'} \in \mathcal{L}^{\tilde{u}} \\ l \neq l'}} \left(\mathbb{E}\left[\left|h_{\tilde{u}}^{(m)}[\tau_l]\right|^2\right] e^{-j\frac{2\pi}{N}(k-k')\tau_l}\right)\left(\mathbb{E}\left[\left|h_{\tilde{u}}^{(m)}[\tau_{l'}]\right|^2\right] e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}}\right) \\ &+ \sum_{\tau_l \in \mathcal{L}^{\tilde{u}}} \sum_{\substack{\tau_{l'} \in \mathcal{L}^{\tilde{u}} \\ l \neq l'}} \left(\mathbb{E}\left[\left|h_{\tilde{u}}^{(m)}[\tau_l]\right|^2\right] e^{-j\frac{2\pi}{N}(k-k'')\tau_l}\right)\left(\mathbb{E}\left[\left|h_{\tilde{u}}^{(m)}[\tau_{l'}]\right|^2\right] e^{-j\frac{2\pi}{N}(k''-k)\tau_{l'}}\right).\end{aligned} \tag{6.37}$$

From (6.37), it can be seen that the expectation depends on the difference between the subcarriers $k - k'$. Therefore, one can interpret $\mathcal{E}_{k,k',k''}$ as a correlation between the subcarriers. Furthermore, the expression of $\mathcal{E}_{k,k',k''}$ can be simplified as

$$\begin{aligned}\mathcal{E}_{k,k',k''} =& \frac{\beta_{\tilde{u}}^2}{L_{\tilde{u}}^2} \underbrace{\sum_{\tau_l \in \mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k-k')\tau_l} \sum_{\tau_{l'} \in \mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}}}_{L_{\tilde{u}}^2 f[k-k';\mathcal{L}^{\tilde{u}}]} \\ &+ \frac{\beta_{\tilde{u}}^2}{L_{\tilde{u}}^2} \underbrace{\sum_{\tau_l \in \mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k-k'')\tau_l} \sum_{\tau_{l'} \in \mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k''-k)\tau_{l'}}}_{L_{\tilde{u}}^2 f[k-k'';\mathcal{L}^{\tilde{u}}]},\end{aligned} \tag{6.38}$$

which yields

$$\mathcal{E}_{k,k',k''} = \beta_{\tilde{u}}^2 f[k-k';\mathcal{L}^{\tilde{u}}] + \beta_{\tilde{u}}^2 f[k-k'';\mathcal{L}^{\tilde{u}}]. \tag{6.39}$$

Note that $f[k-k';\mathcal{L}^{\tilde{u}}]$ is a positive real number since it is the result of the multiplication of a conjugate pair. It can be seen that $f[k-k';\mathcal{L}^{\tilde{u}}]$ is a function of the temporal distribution of multipaths, which causes $f[k-k';\mathcal{L}^{\tilde{u}}]$ to be a random variable. By using (6.39), the distortion power for a given MPC realization can be calculated as

$$S_{dd,1|\mathbf{H}_k}^{(\tilde{u})}[k] \approx 2M^2 \frac{|\alpha_3|^2 \rho_u^3 \beta_{\tilde{u}}}{N_S^3} \sum_{k',k''} (f[k-k';\mathcal{L}^{\tilde{u}}] + f[k-k'';\mathcal{L}^{\tilde{u}}])^2, \tag{6.40}$$

for $|k'| \leq N_S/2$, $|k''| \leq N_S/2$ and $|k'+k''-k| \leq N_S/2$.

As mentioned above, we can treat the MPC distribution as a random variable; hence, We can find the average distortion power by taking the expectation with respect to the MPC distribution $S_{dd,1}^{(\tilde{u})}[k] = \mathbb{E}_{\mathbf{H}_k}\left[S_{dd,1|\mathbf{H}_k}^{(\tilde{u})}[k]\right]$ as

$$S_{dd,1}^{(\tilde{u})}[k] \approx 2M^2 \frac{|\alpha_3|^2 \rho_u^3 \beta_{\tilde{u}}}{N_S^3} \sum_{k',k''} \epsilon_{k,k'}^{f2} + 2\epsilon_{k,k',k''}^{ff} + \epsilon_{k,k''}^{f2} \tag{6.41}$$

for $|k'| \leq N_S/2$, $|k''| \leq N_S/2$ and $|k' + k'' - k| \leq N_S/2$, where variables are defined as $\epsilon_{k,k'}^{f2} \triangleq \mathbb{E}_{\mathbf{H}_k}\left[f^2[k-k';\mathcal{L}^{\tilde{u}}]\right]$, and $\epsilon_{k,k',k''}^{ff} \triangleq \mathbb{E}_{\mathbf{H}_k}\left[f[k-k';\mathcal{L}^{\tilde{u}}]f[k-k'';\mathcal{L}^{\tilde{u}}]\right]$. We can evaluate $\epsilon_{k,k',k''}^{ff}$ analytically, assuming that the time delays τ_l have a particular random distribution. We obtain the expression (6.42) by assuming that the time delays are uniformly distributed independent random variables $\mathcal{U}[0, \tau_{\max}]$, where $\tau_{\max}$ is the maximum delay spread.

$$\epsilon_{k,k',k''}^{ff} = \frac{1}{L_{\tilde{u}}^4} \sum_{l=0}^{L_{\tilde{u}}-1} \sum_{l'=0}^{L_{\tilde{u}}-1} \sum_{\hat{l}=0}^{L_{\tilde{u}}-1} \sum_{\hat{l}'=0}^{L_{\tilde{u}}-1} \mathbb{E}_{\mathbf{H}_k}\left[e^{-j\frac{2\pi}{N}(k-k')\tau_l} e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}} e^{-j\frac{2\pi}{N}(k-k'')\tau_{\hat{l}}} e^{-j\frac{2\pi}{N}(k''-k)\tau_{\hat{l}'}}\right]. \tag{6.42}$$

Note that for the set $\mathcal{L}_s^{\tilde{u}} = \{l = l' \text{ and } \hat{l} = \hat{l}' \big| \forall l, l', \hat{l}, \hat{l}'\}$, the term in the expectation in (6.42) becomes 1; hence, we can rewrite this expectation by only focusing on the dominant terms, which results in (6.43)

$$\begin{aligned}\epsilon_{k,k',k''}^{ff} \approx \frac{1}{L_{\tilde{u}}^2} + \frac{1}{L_{\tilde{u}}^4} \sum_{l,l',\hat{l},\hat{l}' \notin \mathcal{L}_s^{\tilde{u}}}^{L-1} & \mathbb{E}_{\mathcal{L}}\left[e^{-j\frac{2\pi}{N}(k-k')\tau_l}\right] \mathbb{E}_{\mathcal{L}}\left[e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}}\right] \\ & \mathbb{E}_{\mathcal{L}}\left[e^{-j\frac{2\pi}{N}(k-k'')\tau_{\hat{l}}}\right] \mathbb{E}_{\mathcal{L}}\left[e^{-j\frac{2\pi}{N}(k''-k)\tau_{\hat{l}'}}\right]\end{aligned} \tag{6.43}$$

We also note that $\mathbb{E}_{\mathbf{H}_k}\left[e^{-j\frac{2\pi}{N}(k-k')\tau_l}\right] \mathbb{E}_{\mathbf{H}_k}\left[e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}}\right]$ is the multiplication of two complex conjugate numbers, so we can simplify (6.43) as

$$\begin{aligned}\epsilon_{k,k',k''}^{ff} &\approx \frac{1}{L_{\tilde{u}}^2} + \frac{1}{L_{\tilde{u}}^4} \sum_{l,l',\hat{l},\hat{l}' \notin \mathcal{L}_s}^{L-1} \left| \underbrace{\mathbb{E}_{\mathbf{H}_k}\left[e^{-j\frac{2\pi}{N}(k-k')\tau}\right]}_{\xi_{k-k'}} \right|^2 \left| \underbrace{\mathbb{E}_{\mathbf{H}_k}\left[e^{-j\frac{2\pi}{N}(k-k'')\tau}\right]}_{\xi_{k-k''}} \right|^2, \\ &\approx \frac{1}{L_{\tilde{u}}^2} + \frac{L_{\tilde{u}}^4 - L_{\tilde{u}}^2}{L_{\tilde{u}}^4} \left|\xi_{k-k'}\right|^2 \left|\xi_{k-k''}\right|^2.\end{aligned} \tag{6.44}$$

Then the expectation for $\xi_{k-k'}$ can be evaluated as

$$\begin{aligned}\xi_{k-k'} &= \sum_{\tau=0}^{\tau_{\tilde{u},\max}-1} \frac{1}{\tau_{\tilde{u},\max}} e^{-j\frac{2\pi}{N}(k-k')\tau} \\ &= \frac{1}{\tau_{\tilde{u},\max}} \frac{e^{-j\frac{2\pi}{N}(k-k')\frac{\tau_{\tilde{u},\max}}{2}}}{e^{-j\frac{2\pi}{N}(k-k')\frac{1}{2}}} \frac{\sin\left(\frac{2\pi}{N}(k-k')\frac{\tau_{\tilde{u},\max}}{2}\right)}{\sin\left(\frac{2\pi}{N}(k-k')\frac{1}{2}\right)}.\end{aligned} \tag{6.45}$$

Similarly, we can approximate $\epsilon_{k'}^{f2} = \mathbb{E}_{\mathbf{H}_k}\left[f^2[k-k';\mathcal{L}]\right]$ as

$$\epsilon_{k,k'}^{f2} \approx \frac{1}{L_{\tilde{u}}^2} + \frac{L_{\tilde{u}}^4 - L_{\tilde{u}}^2}{L_{\tilde{u}}^4} \frac{1}{\tau_{\tilde{u},\max}^4} \frac{\sin^4\left(\frac{2\pi}{N}(k-k')\frac{\tau_{\tilde{u},\max}}{2}\right)}{\sin^4\left(\frac{2\pi}{N}(k-k')\frac{1}{2}\right)}. \tag{6.46}$$

To sum up, the received distortion power component for this case can be expressed as

$$S_{dd,1}^{(\tilde{u})}[k] \approx \rho_{\tilde{u}}^3 \beta_{\tilde{u}} \bar{S}_{dd,1}^{(\tilde{u})}[k] \tag{6.47}$$

where $\bar{S}_{dd,1}^{(\tilde{u})}[k] = 2M^2\frac{|\alpha_3|^2}{N_S^3}\sum_{k',k''}\epsilon_{k,k'}^{f2} + 2\epsilon_{k,k',k''}^{ff} + \epsilon_{k,k''}^{f2}$ is the distortion power, which independent of the power allocation coefficients.

As can be interpreted from (6.41), the distortion power depends on the frequency selectivity via $\epsilon_{k,k'}^{f2}$ and $\epsilon_{k,k',k''}^{ff}$ terms, which have the form of $\sin(\tau_{\max}x)/\sin(x)$ (sinc-type behavior). The terms in (6.41) are larger for the main lobe of the sinc-type function, which makes $\tau_{\max}$ another factor that impacts the received distortion. If $\tau_{\max}$ is large, then the distortion power is dominated by the constant term $1/L_{\tilde{u}}^2$ in (6.41) since the main lobe is narrow. Consequently, as L increases, the received distortion power reduces. On the other hand, if $\tau_{\tilde{u},\max}$ is small (the main lobe is wide), then the in-band distortion power, where $|k-k'|$ is small, is roughly proportional to $L_{\tilde{u}}^4/L_{\tilde{u}}^4$, which corresponds to the case that the received distortion remains the same.

6.4.2.2 Power of the distortion stemming from coherently combined distortion due to interfering users

We will derive the distortion power expression for the first sub-case $\tilde{u} = u$, $u' = u''$, $u \neq u'$, for which the distortion is also coherently combined at the received location. Consider the received distortion component for this sub-case $S_{dd,2a|\mathbf{H}_k}^{(\tilde{u})}[k,k',k'']$ as

$$S^{(\tilde{u})}_{dd,2a|\mathbf{H}_k}[k,k',k''] = \rho_{\tilde{u}} \sum_{u'=1,u'\neq\tilde{u}}^{U} \rho^2_{u'} \left| \underbrace{\sum_{m=1}^{M} \left[\left(h^{(m)}_{\tilde{u},k} \right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{u',k''} \right)^* \bar{w}^{(m)}_{u',k'+k''-k} \right]}_{S^{(\tilde{u},u')}_{dd,2aa|\mathbf{H}_k}[k,k',k'']} \right|^2 \tag{6.48}$$

Similar to the first case, LLN will be used to approximate the random variable, which is denoted by $S_{dd,2aa|\mathbf{H}_k}[k,k',k'']$ by a single deterministic scalar as

$$S^{(\tilde{u},u')}_{dd,2aa|\mathbf{H}_k}[k,k',k''] \approx M\mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k} \right)^* \bar{w}^{(m)}_{\tilde{u},k'} \left(\bar{w}^{(m)}_{u',k''} \right)^* \bar{w}^{(m)}_{u',k'+k''-k} \right] \tag{6.49}$$

which can be expressed under the assumption in (6.10) as

$$S^{(\tilde{u},u')}_{dd,2aa|\mathbf{H}_k}[k,k',k''] \approx M\frac{\sqrt{\beta_{\tilde{u}}}\beta_{u'}}{\beta_{\tilde{u}}{\beta_{u'}}^2} \underbrace{\mathbb{E}\left[\left(h^{(m)}_{\tilde{u},k} \right)^* h^{(m)}_{\tilde{u},k'} \left(h^{(m)}_{u',k''} \right)^* h^{(m)}_{u',k'+k''-k} \right]}_{\mathcal{Q}_{k,k',k''}}. \tag{6.50}$$

Since the user channels are independent, the expectation $\mathcal{Q}_{k,k',k''}$ can be decomposed as

$$\mathcal{Q}_{k,k',k''} = M\mathbb{E}\left[h^{(m)}_{\tilde{u},k} \left(h^{(m)}_{\tilde{u},k'} \right)^* \right] \mathbb{E}\left[\left(h^{(m)}_{u',k''} \right)^* h^{(m)}_{u',k'+k''-k} \right] \tag{6.51}$$

where $\mathbb{E}\left[h^{(m)}_{\tilde{u},k} \left(h^{(m)}_{\tilde{u},k'} \right)^* \right]$ can be evaluated by using the Fourier transform representation as

$$\begin{aligned} \mathbb{E}\left[h^{(m)}_{\tilde{u},k} \left(h^{(m)}_{\tilde{u},k'} \right)^* \right] &= \frac{\beta_{\tilde{u}}}{L_{\tilde{u}}} \sum_{\tau_l,\tau_{l'}\in\mathcal{L}^{\tilde{u}}} \mathbb{E}\left[h^{(m)}_{\tilde{u},k}[\tau_l] \left(h^{(m)}_{\tilde{u},k'}[\tau_{l'}] \right)^* \right] e^{-j\frac{2\pi}{N}k\tau_l} e^{j\frac{2\pi}{N}k'\tau_{l'}} \\ &= \frac{\beta_{\tilde{u}}}{L_{\tilde{u}}} \sum_{\tau_l\in\mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k-k')\tau_l} \end{aligned} \tag{6.52}$$

similarly $\mathbb{E}\left[h^{(m)}_{u',k''} \left(h^{(m)}_{,u'k'+k''-k} \right)^* \right]$ can be evaluated as

$$\begin{aligned} \mathbb{E}\left[h^{(m)}_{u',k''} \left(h^{(m)}_{u',k'+k''-k} \right)^* \right] &= \frac{\beta_{u'}}{L_{u'}} \sum_{\tau_l,\tau_{l'}\in\mathbb{L}_{u'}} \mathbb{E}\left[h^{(m)}_{u',k''}[\tau_l] \left(h^{(m)}_{u',k'+k''-k}[\tau_{l'}] \right)^* \right] e^{-j\frac{2\pi}{N}k''\tau_l} \\ &\qquad e^{j\frac{2\pi}{N}(k'+k''-k)\tau_{l'}} \\ &= \frac{\beta_{u'}}{L_{u'}} \sum_{\tau_l\in\mathbb{L}_{u'}} e^{-j\frac{2\pi}{N}(k'-k)\tau_l}. \end{aligned} \tag{6.53}$$

Consequently, for a given channel realization the distortion power for the first sub-case of the second component becomes

$$S^{(\tilde{u},u')}_{dd,2aa|\mathbf{H}_k}[k,k',k''] \rightarrow M\frac{\sqrt{\beta_{\tilde{u}}}}{L^2_{\tilde{u}}}\sum_{\tau_l\in\mathcal{L}^{\tilde{u}}} e^{-j\frac{2\pi}{N}(k-k')\tau_l}\sum_{\tau_{l'}\in\mathcal{L}^{u'}} e^{-j\frac{2\pi}{N}(k'-k)\tau_{l'}}. \tag{6.54}$$

Then the average distortion power can be calculated by using the independence of the channels of different users $S^{(\tilde{u})}_{dd,2a}[k,k',k''] = \mathbb{E}\left[\rho_{\tilde{u}}\sum_{\substack{u'=1\\u'\neq\tilde{u}}}^{U}\rho^2_{u'}\left|S^{(\tilde{u})}_{dd,2aa|\mathbf{H}_k}[k,k',k'']\right|^2\right]$ as

$$S^{(\tilde{u},u')}_{dd,2a|\mathbf{H}_k}[k,k',k''] \approx M^2\frac{\beta_{\tilde{u}}\rho_{\tilde{u}}}{L^2_{\tilde{u}}L^2_{u'}}\Gamma_{\tilde{u}}[k-k']\sum_{\substack{u'=1\\u'\neq\tilde{u}}}^{U}\rho^2_{u'}\Gamma_{u'}[k-k'], \tag{6.55}$$

where $\Gamma_u[k-k'] \triangleq \mathbb{E}\left[\left|\sum_{\tau_l\in\mathbb{L}_u} e^{-j\frac{2\pi}{N}(k-k')\tau_l}\right|^2\right]$, and it is evaluated as in (6.56).

$$\mathbb{E}_{\mathbb{L}_u}\left[\left|\sum_{\tau_l\in\mathbb{L}_u} e^{-j\frac{2\pi}{N}(k-k')\tau_l}\right|^2\right] = L_u + L_u(L_u-1)\left|\frac{1}{\tau_{\max}}\frac{\sin^2\left(\frac{2\pi}{N}(k-k')\frac{\tau_{\max}}{2}\right)}{\sin^2\left(\frac{2\pi}{N}(k-k')\frac{1}{2}\right)}\right|. \tag{6.56}$$

Also, note that the value of $S^{(\tilde{u},u')}_{dd,2a}[k,k',k'']$ only depends on the different $k-k'$; hence, we can denote it as $S^{(\tilde{u},u')}_{dd,2a}[k-k']$. Similarly for the second sub-case $\tilde{u}=u''$, $u=u'$, $u\neq u'$, the received distortion component $S^{(\tilde{u},u')}_{dd,2b}[k,k',k''] = S^{(\tilde{u},u')}_{dd,2b}[k-k'']$ can be obtained as

$$S^{(\tilde{u})}_{dd,2b}[k-k''] = M^2\frac{\beta_{\tilde{u}}\rho_{\tilde{u}}}{L^2_{\tilde{u}}}\Gamma_{\tilde{u}}[k-k'']\sum_{u=1,u\neq\tilde{u}}^{U}\frac{\rho^2_u}{L^2_u}\Gamma_u[k-k'']. \tag{6.57}$$

Consequently, the overall received distortion power for the second component is represented as

$$S^{(\tilde{u})}_{dd,2}[k] = \frac{\beta_{\tilde{u}}\rho_{\tilde{u}}}{L^2_{\tilde{u}}}\sum_{u=1,u\neq\tilde{u}}^{U}\frac{\rho^2_u}{L^2_u}\Gamma^{\tilde{u}}_2[k], \tag{6.58}$$

where

$$\begin{aligned}\Gamma^{\tilde{u}}_2[k] =& M^2\frac{2|\alpha_3|^2}{N_s^2}\sum_{k'=-N/2+1}^{N/2}\sum_{k''=-N/2+1}^{N/2}\\ &(\Gamma_{\tilde{u}}[k-k'']\Gamma_u[k-k''] + \Gamma_{\tilde{u}}[k-k']\Gamma_u[k-k'])\end{aligned} \tag{6.59}$$

for $|k'|\leq N_S/2$, $|k''|\leq N_S/2$ and $|k'+k''-k|\leq N_S/2$.

6.4.2.3 Distortion power due to isoropic radiation

In this section, the distortion components for $\tilde{u} \neq u$ and $\tilde{u} \neq u''$ cases will be analyzed. For this case, it can be observed that distortion is not coherently combined since phases of the distortion terms are not aligned. Therefore, the received distortion for $\tilde{u} \neq u$ and $\tilde{u} \neq u''$ can be expressed as

$$\begin{aligned} S_{dd,3a}^{(\tilde{u})}[k] = \sum_{m,m'} & \left[h_{k,\tilde{u}}^{(m)} \bar{w}_{u,k'}^{(m)} \left(\bar{w}_{u',k''}^{(m)} \right)^* \bar{w}_{u'',k'+k''-k}^{(m)} \right] \\ & \left[h_{k,\tilde{u}}^{(m')} \bar{w}_{u,k'}^{(m')} \left(\bar{w}_{u',k''}^{(m')} \right)^* \bar{w}_{u'',k'+k''-k}^{(m')} \right]. \end{aligned} \tag{6.60}$$

which can be expressed by using the assumption in (6.10) as

$$\begin{aligned} S_{dd,3a}^{(\tilde{u})}[k] &\approx M \frac{\beta_u \beta_{u'} \beta_{u''}}{\beta_u^2 \beta_{u'}^2 \beta_{u''}^2} \mathbb{E}\left[|h_{\tilde{u},k}^{(m)}|^2 \right] \mathbb{E}\left[|h_{u,k}^{(m)}|^2 \right] \mathbb{E}\left[|h_{u',k}^{(m)}|^2 \right] \mathbb{E}\left[|h_{u'',k}^{(m)}|^2 \right] \\ &= M \beta_{\tilde{u}}. \end{aligned} \tag{6.61}$$

It is observed from (6.61) that distortion power scales with M, and it is independent of the frequency bin. Hence, this component is incoherently combined at the receiver, and its effect is negligible if the number of antennas is large. The overall distortion power due to all uncoherent components is expressed as

$$S_{dd,3}^{(\tilde{u})}[k] = \beta_{\tilde{u}} \Gamma_3^{\tilde{u}}[k] \varrho_{\tilde{u}} \tag{6.62}$$

where

$$\Gamma_3^{\tilde{u}}[k] \triangleq M \frac{2|\alpha_3|^2}{N_s^2} \sum_{k'=-N/2+1}^{N/2} \sum_{k''=-N/2+1}^{N/2} 1. \tag{6.63}$$

and $\varrho_{\tilde{u}} \triangleq \sum_{\substack{u,u',u'' \\ \tilde{u} \neq u, \tilde{u} \neq u''}} \rho_u \rho_{u'} \rho_{u''}$ and it is simplified as $\varrho_{\tilde{u}} = 1 - 2\rho_{\tilde{u}} \sum_{u=1}^{U} \rho_u^2 + \rho_{\tilde{u}}^3$. As a result, we can express the received distortion power $S_{dd}^{(\tilde{u})}[k] = \mathbb{E}_{\mathbf{H}_k}\left[\left| \mathbf{h}_{\tilde{u},k}^{\mathrm{H}} \mathbf{d}_k \right|^2 \right]$ as

$$S_{dd}^{(\tilde{u})}[k] \approx \rho_{\tilde{u}}^3 \beta_{\tilde{u}} \bar{S}_{dd,1}^{(\tilde{u})}[k] + \frac{\beta_{\tilde{u}} \rho_{\tilde{u}}}{L_{\tilde{u}}^2} \sum_{u=1, u \neq \tilde{u}}^{U} \frac{\rho_u^2}{L_u^2} \Gamma_2^{\tilde{u}}[k] + \beta_{\tilde{u}} \Gamma_3^{\tilde{u}}[k] \varrho_{\tilde{u}}. \tag{6.64}$$

6.5 Power Allocation via Geometric Programming

In this section, a power allocation framework based on the derived distortion power expression is developed. Based on the expressions given by (6.25) and (6.64), it is evident that the frequency dependency is exclusively present in the received distortion power. Furthermore, it can be observed from the convolution operations in (6.19) that the maximum distortion power is concentrated at the center frequency, specifically $k = 0$. Hence, to obtain power allocation coefficients that are independent of subcarrier index, we will consider the peak distortion power in the optimization problem. As a result, the sum-rate R_{sum} can be expressed as

$$R_{\text{sum}} = \sum_{u=1}^{U} \log_2(1 + \text{SDNR}_u[0]). \tag{6.65}$$

Then the following optimization problem is solved to find the optimal user powers as

$$\begin{aligned} &\underset{\boldsymbol{\rho}}{\text{maximize}} \quad R_{\text{sum}} \\ &\text{subject to} \quad \sum_{u=1}^{U} \rho_u \leq 1. \end{aligned} \tag{6.66}$$

It is obvious from (6.66) that the optimization problem does not have a standard form to be solved efficiently. Nevertheless, it is feasible to approximate the problem, thereby rendering it equivalent to a form that can be efficiently solved. By using the assumption $\text{SDNR}_u[0] \gg 1$ [94], (6.66) can be approximated as

$$\begin{aligned} (\text{P1}): \quad &\underset{\boldsymbol{\rho}}{\text{maximize}} \quad \prod_{u=1}^{U} \text{SDNR}_u[0] \\ &\text{subject to} \quad \sum_{u=1}^{U} \rho_u \leq 1. \end{aligned} \tag{6.67}$$

It can be observed from (6.67), denominator of the objective function consists of multiplication of the posynomials and the numerator contains a monomial.Therefore, P1 can be formulated such that it can be solved via geometric programming. Note that

maximization of $\prod_{u=1}^{U} \mathrm{SDNR}_u[0]$ is equivalent to minimization of $\prod_{u=1}^{U} \frac{1}{\mathrm{SDNR}_u[0]}$; hence, P1 can be expressed interms of minimization. Furthermore, since we are employing zero-forcing type precoding, we can remove the effects of inter-user interference for the optimization problem since its effects are negligible compared to the nonlinear distortion power.

$$\begin{aligned} (\mathrm{P2}): \quad & \underset{\boldsymbol{\rho}}{\text{minimize}} \quad \prod_{u=1}^{U} \frac{\rho_u^3 \beta_u \bar{S}_{dd,1}^{(u)}[0] + \frac{\beta_u \rho_u}{L_{\tilde{u}}^2} \sum_{\substack{u'=1,\\ u' \neq u}}^{U} \frac{\rho_{u'}^2}{L_{u'}^2} \Gamma_2^u[0] + \beta_u \Gamma_3^u[0] \varrho_u + \sigma_\eta^2}{\rho_u \mathbb{E}_{\mathbf{H}_0}\left[\left|\mathbf{h}_{u,0}^{\mathrm{H}} \bar{\mathbf{w}}_{u,0}\right|^2\right]} \\ & \text{subject to} \quad \sum_{u=1}^{U} \rho_u \leq 1. \end{aligned} \tag{6.68}$$

By dividing both the numerator and denominator of the objective function by the factor $\prod_{u=1}^{U} \rho_u$, the problem turns into the standard form of geometric programing as

$$\begin{aligned} (\mathrm{P3}): \quad & \underset{\boldsymbol{\rho}}{\text{minimize}} \quad \prod_{u=1}^{U} \frac{c_u \bar{S}_{dd,1}^{(u)}[0] + \frac{\beta_u}{L_{\tilde{u}}^2} \sum_{\substack{u'=1,\\ u' \neq u}}^{U} \frac{\rho_{u'}^2}{L_{u'}^2} \Gamma_2^u[0] + \beta_u \Gamma_3^u[0] \varrho_u \rho_u^{-1} + \sigma_\eta^2 \rho_u^{-1}}{\mathbb{E}_{\mathbf{H}_0}\left[\left|\mathbf{h}_{u,0}^{\mathrm{H}} \bar{\mathbf{w}}_{u,0}\right|^2\right]} \\ & \text{subject to} \quad \sum_{u=1}^{U} \rho_u \leq 1. \end{aligned} \tag{6.69}$$

where $c_u = \rho_u^2 \beta_u$. Since the objective function to be minimized is the multiplication of U posynomials, power allocation coefficients can be obtained via employing general-purpose solvers such as CVX [98].

6.6 Simulation Results

In this section, the derived distortion power expression at the user location is verified, and the performance of the proposed power allocation scheme is tested for different scenarios. In the simulations, the effects of several factors, such as the number of users U, the number of antennas M, and the frequency selectivity, is investigated.

For the simulations, a memoryless polynomial PA model in (2.4) with $W_{\mathrm{PA}} = 1$, whose parameters are given in [99], is employed. In order to provide a practical investigation, the bandwidth of the signal is assumed to be 20 MHz, and the sampling rate is 80 MHz, which corresponds to oversampling factor $\mu = 4$. The noise variance is taken as -174 $\mathrm{dBm/Hz}$, and the maximum total power that the BS can supply is 43 dBm. The path loss parameter β_u in dB scale can be modeled as

$$\beta_{u,\mathrm{dB}} = -30.5 - 36.7 \log_{10}\left(\frac{d_u}{1}\right) \ \mathrm{dB}, \tag{6.70}$$

where d_u is the distance of the u^{th} user to the BS. The users are randomly generated in a region of $1\mathrm{km} \times 1\mathrm{km}$ area.

6.6.1 The effects of the number of users on nonlinear distortion

Firstly, the distortion behavior exhibits variations in response to the number of users being served. In order to specifically investigate the impacts of multi-user diversity, we focus our analysis on a noise-free scenario characterized by flat fading. By adopting this approach, we isolate and examine the influence of multi-user diversity, thereby excluding the confounding effects of additional impairments. For this purpose, the PSD of the received distortion, which is normalized by the total received power, is shown in Figure 6.1 to show the relative distortion power to the useful signal for $M = 64$ antennas. It can be observed that as the number of served users increases, the received distortion power decreases more rapidly to distortion power since the number of the directions where the distortion is beamformed can be expressed by U^3, which means that distortion power decreases by $\frac{1}{U^3}$ for $U \ll M$. However, the received power for the useful part is reduced by $\frac{1}{U}$. Therefore, up to some threshold, a reduction in the effective distortion power is observed. However, after that threshold, even if the number of users increases, the received distortion power remains the same since distortion is radiated isotropically. Hence, the relative distortion power seems to increase since the power of the useful part is decreased to serve more users, while the total received distortion power remains the same. In addition, it can be seen that the analytical distortion PSD is well in line with the numerical results, which verifies the performed analysis.

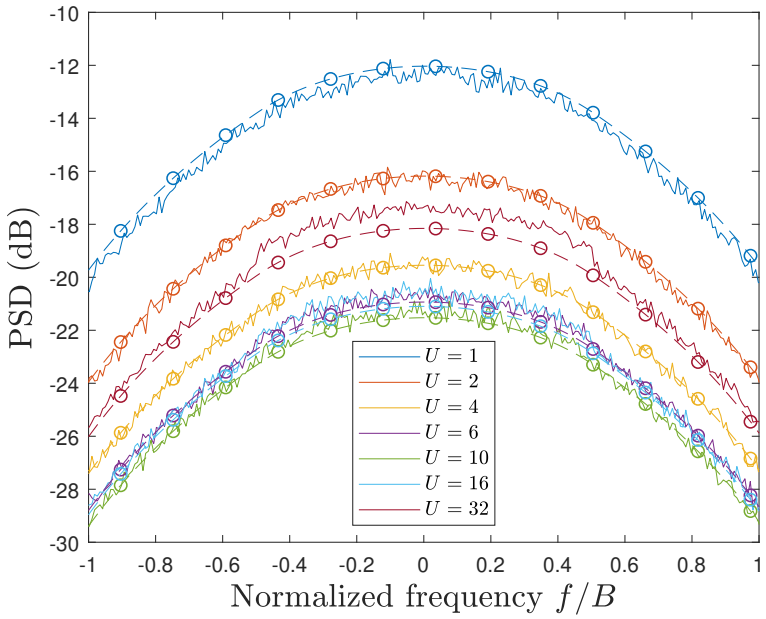


Figure 6.1: PSD of the received distortion term for the different number of users.

Besides, to show the significance of the distortion correlation, the estimated PSD for the uncorrelated distortion assumption is shown in Figure 6.2. It can be seen that the estimated distortion power under the assumption of uncorrelated distortion is much less than the exact distortion power. However, the proposed analytical framework is able to take the distortion correlation into account.

6.6.2 The effects of frequency-selectivity on nonlinear distortion

In Fig. 6.3, normalized PSDs of the received signals, for which the received signal power is scaled to be 1, are shown to examine the effects of both the number of MPCs and the delay spread. It can be observed from Fig. 6.3 that the received distortion power tends to decrease as the number of MPCs increases; however, different behaviors are observed for the in-band and OOB distortion characteristics. The decay rate of the OOB distortion power is higher compared to that of in-band distortion, which is also verified by the analytical expression in (6.41) since (6.41) can indicate the distinctive characteristics for the in-band and OOB distortion as can be seen from In Fig. 6.3, and the obtained analytical distortion curves overlap with the numerical ones. The difference between the decay characteristics stems from the variation of

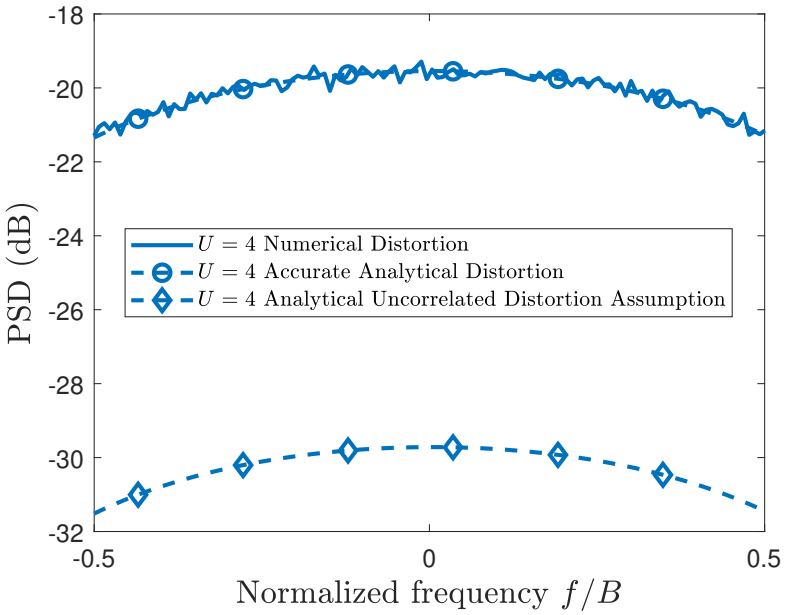


Figure 6.2: The effects of distortion correlation on PSD of the received distortion term for $U = 4$ and $M = 64$.

the correlation between the subcarriers contained in $\epsilon^{f2}_{k,k'}$ and $\epsilon^{ff}_{k,k',k''}$ terms, which are described by the *sinc*-type behavior. The reason is that terms in the summation (6.41) are only non-zero for larger $k-k'$, $k-k''$ terms, which correspond to side-lobes of the *sinc*-type function. However, in the case of in-band distortion, the terms are also non-negative for the neighborhood of k, which corresponds to the main lobe of the *sinc* function. Eventually, the received distortion power is larger for the in-band distortion than for the OOB distortion. As the number of MPCs increases, the significance of the *sinc*-related term becomes dominant over the constant term (L or L^2). This yields a substantial difference between the in-band and OOB distortion powers. Fig. 6.3 also shows the PSD of the received signal for $M = 1000$ to examine the nearly asymptotic case. It can be seen that as M increases, the numerically PSD converges to the analytical PSD for higher MPC numbers since the approximations in (6.35) are tighter. Therefore, it can be concluded that the obtained analytical received distortion expression provides an accurate asymptotic deterministic approximation. In addition, the analytical received distortion expression is a good approximation for many practical scenarios. It provides close approximations in practical parameter ranges.

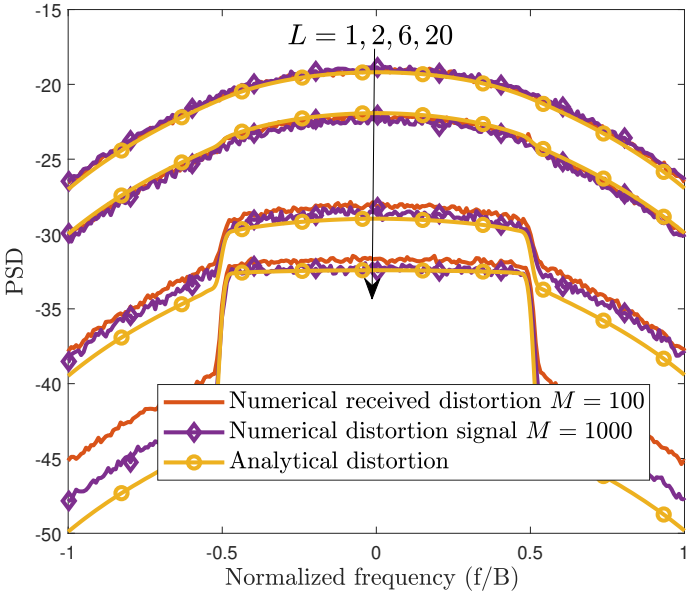


Figure 6.3: PSD of the received signal for different number of MPCs for $\tau_{\max}/\mu = 25$

.

The impact of the maximum delay spread is demonstrated in Fig. 6.4. It can be observed that as the delay spread increases, the in-band distortion power converges to the OOB distortion power since the main lobe of the *sinc*-type function gets narrower. A narrower main lobe reduces the number of significant terms in the summation (6.41), which yields a reduction in the received distortion power.

6.6.3 Sum SE Evaluations

The performance evaluation of the proposed power allocation scheme is conducted under various scenarios, encompassing both frequency-flat and frequency-selective channels. Firstly, we examine the sum spectral efficiency (SE) performance of massive MIMO systems employing $M = 64$ antennas with the presence of $U = 4$ users. As a benchmark for comparison, we consider a power allocation scheme that disregards the distortion correlation across the antenna arrays. In Figure 6.5, CDF of the sum SE is shown for the proposed power allocation scheme (Prop. Pow. All.), and it is compared with the power allocation based on uncorrelated distortion assumption (Unc. Pow. All.) for both flat and frequency-selective channels. Furthermore,

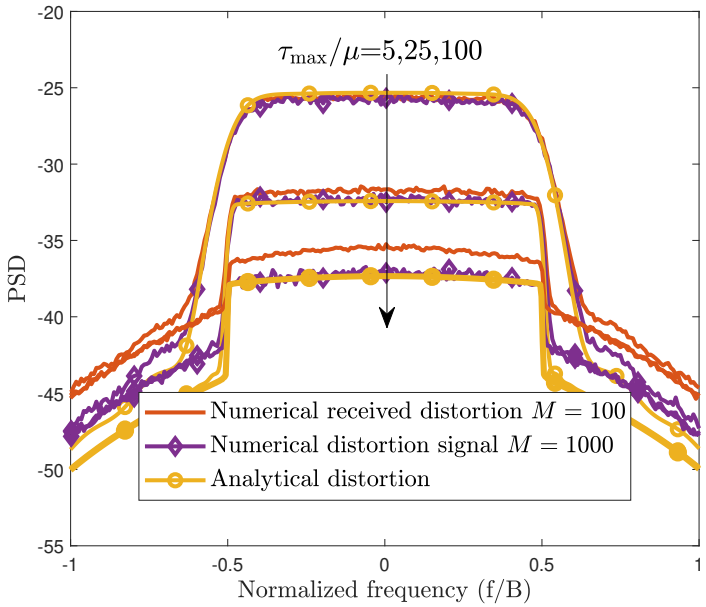


Figure 6.4: PSD of the received signal for different delay spread for $L = 20$.

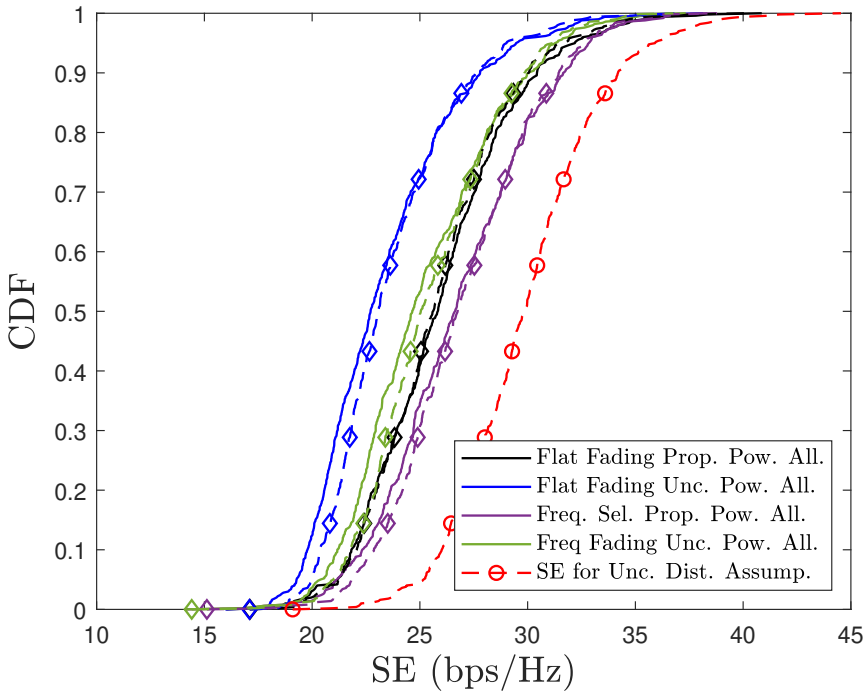
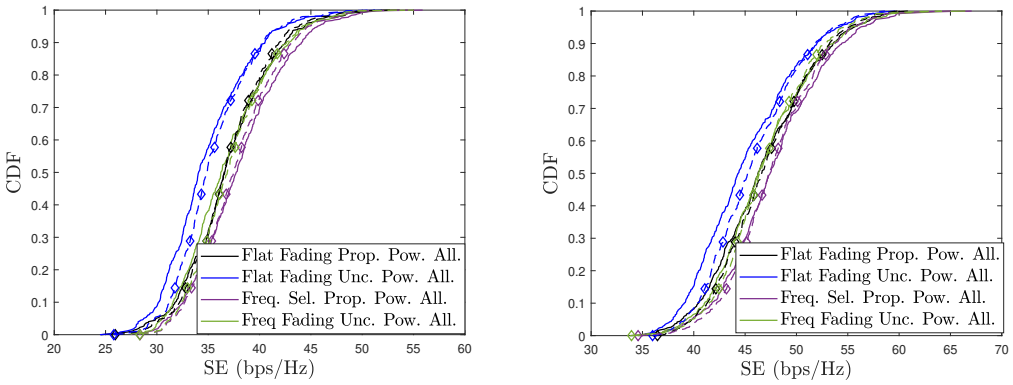


Figure 6.5: CDF of the sum SE for $M = 64$ and $U = 4$.

the CDFs of the sum SE obtained from the analytical framework are illustrated using dashed lines accompanied by diamond markers. The results demonstrate that the developed analytical expression for SDNR accurately captures the SE behavior. However, when SE analysis is based on the assumption of uncorrelated distortion (referred to as SE for Unc. Dist. Assump.), it leads to a significantly flawed evaluation of system performance as it overestimates the achieved spectral efficiency. In addition, it is observed that by taking the spatial correlation of the distortion into account, one could improve the system performance significantly. Besides, it can be concluded that the achievable performance of the systems improves with the frequency-selectivity since less number of subcarrier channels contributes distortion, yielding smaller received distortion power. Hence, higher SE can be achieved under frequency-selective channels.

To investigate the behavior of nonlinear distortion, we examined various scenarios where the number of served users was systematically increased. Figure 6.6 displays the results, revealing that the enhancement achieved by incorporating distortion correlation across the antenna array diminishes as the number of users grows. This outcome is in line with expectations, as the escalating number of served users causes the antenna input to decorrelate, resulting in uncorrelated distortion. Consequently, the disparity in performance between methods considering and disregarding spatial distortion correlation diminishes proportionally to the system's load. These findings contribute to a comprehensive understanding of the impact of user quantity on the effectiveness of distortion correlation schemes.

To further investigate the influence of the number of antennas, we extended our investigation to consider a scenario with $M = 100$ antennas. The objective was to assess the effects of antenna quantity on system performance. Figure 6.7 illustrates the spectral efficiency (SE) results obtained for $M = 100$ antennas and $U = 4$ users. It is evident from the figure that an increase in the number of antennas leads to more pronounced performance improvements. To provide a comprehensive analysis, we included results for scenarios involving $U = 6$ and $U = 8$ users. The results indicate that performance improvement becomes more significant when considering distortion correlation. This is primarily due to the fact that as the number of antennas increases, the correlated distortion power relative to the uncorrelated distortion power becomes



(a) $M = 64, U = 6$ (b) $M = 64, U = 8$

Figure 6.6: CDFs of the sum SE for $M = 64$, $U = 6$ and $U = 8$

more substantial. Thus, accounting for distortion correlation becomes increasingly advantageous.

Lastly, the effects of the number of antennas and the number of users are investigated. In Figure 6.9 (a), the sum SE efficiency is shown for different numbers of antennas, where the number of users is fixed to $U = 8$. For the smaller number of antennas, both the proposed distortion correlation-aware power allocation scheme and power allocation based on uncorrelated distortion have similar performances since distortion becomes uncorrelated for a smaller M/U ratio. However, as the number of antennas increases distortion correlation starts to become significant, and ignoring distortion correlation degrades the system performance. In Figure 6.9 (b), the effects of the number of users are shown by examining per-user average SE performance. It can be seen that when the number of users is increased from $U = 4$ to $U = 6$, the per-user SE increases since the relative received distortion power decreases, which means that the effect of noise power is negligible compared to the distortion power. Also, it can be seen that the performance gap is larger for the lower number of users since the M/U ratio is higher, and distortion correlation is more effective.

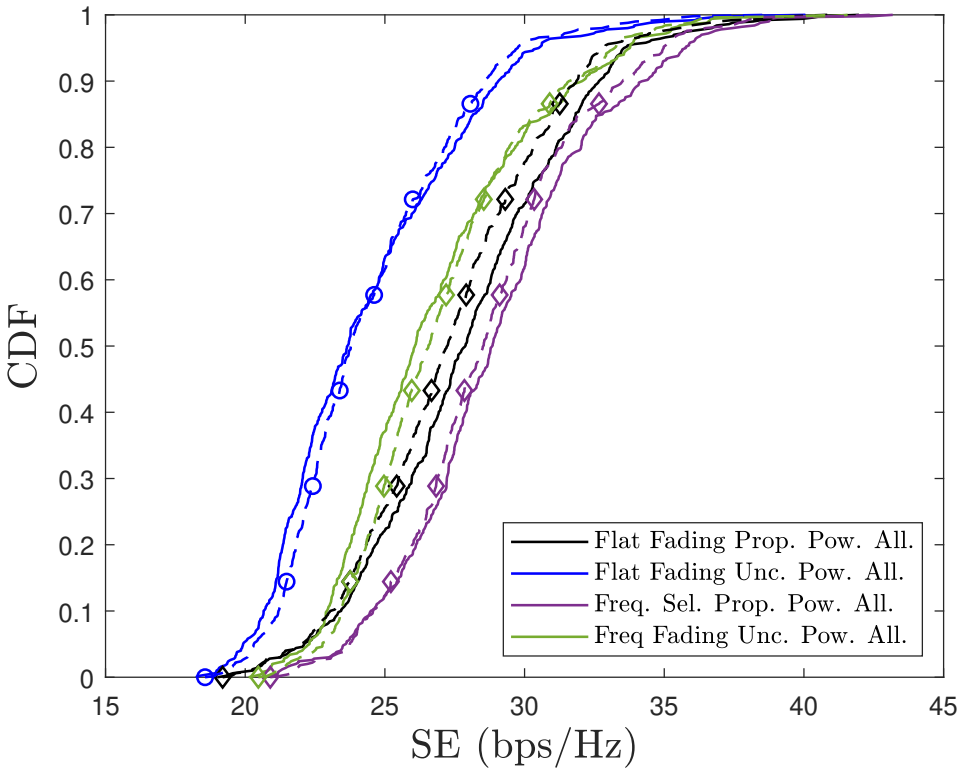
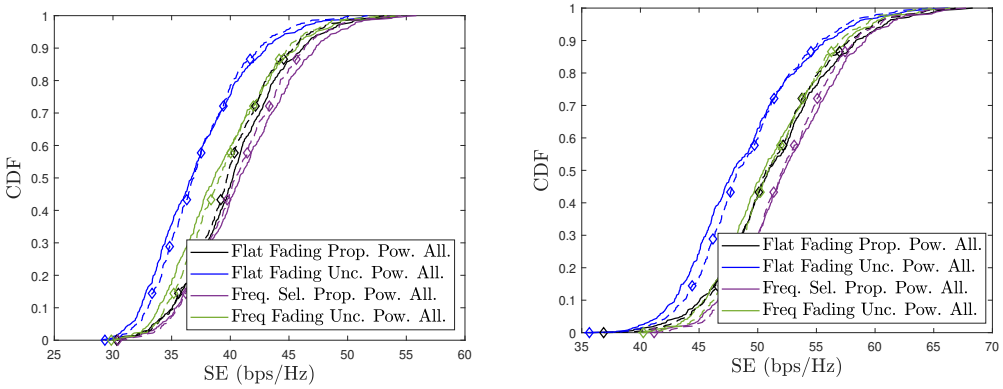


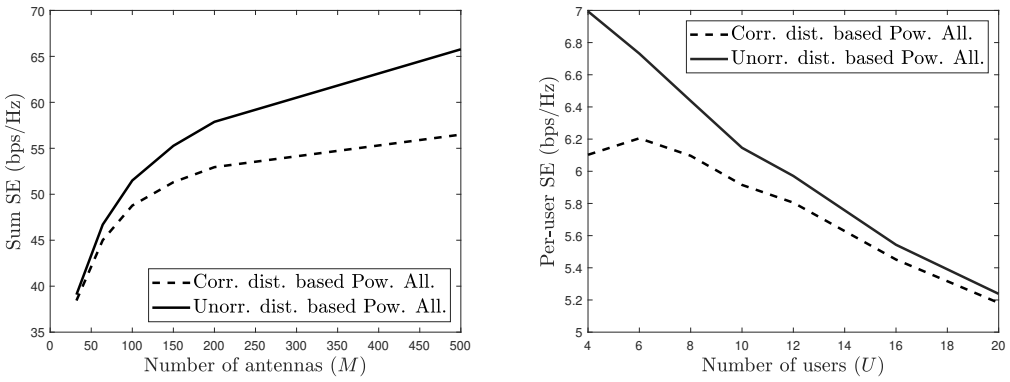
Figure 6.7: CDF of the sum SE for $M = 100, U = 4$



(a) CDF of the sum SE for $M = 100, U = 6$

(b) CDF of the sum SE for $M = 100, U = 8$

Figure 6.8: CDFs of the sum SEs for $M = 100$, $U = 6$ and $U = 8$



(a) (b)

Figure 6.9: (a) Average sum SE with respect to number of antennas for $U = 8$ users (b) average SE per user with respect to number of users for $M = 100$ antennas

6.7 Conclusion

This chapter presents a comprehensive analysis of the received distortion in multi-user massive multiple-input-multiple-output (MIMO) systems with frequency-selective channels. The analysis accurately characterizes the received distortion power by considering the spatial correlation of the distortion and leveraging slowly varying large scale parameters. By incorporating these factors, the proposed analysis provides an in-depth understanding of the received distortion behavior.

The results of the analysis reveal important insights. It is observed that the relative distortion power exhibits a decreasing trend as the number of users increases, up to a certain threshold. However, beyond this threshold, the relative distortion power starts to increase. This finding suggests that increasing the number of users can initially mitigate the effects of nonlinear distortion, but there is a point where the distortion power begins to dominate and adversely affect system performance.

Furthermore, the influence of frequency selectivity on the distortion signal across the antenna array is investigated. The analysis demonstrates that as the degree of frequency selectivity increases, the distortion signals become decorrelated across the antenna array. This decorrelation effect contributes to a reduction in the received distortion power, ultimately improving system performance.

Lastly, based on the insights gained from the analysis, a power allocation framework is proposed. The aim of this framework is to dynamically adjust the power allocated to each user, thereby mitigating the adverse effects of nonlinear distortion. By optimizing the power allocation, the proposed framework enables a more effective management of the nonlinear distortion in the system.

CHAPTER 7

REDUCED COMPLEXITY CORRELATION-BASED MULTI-STREAM DPD FOR HYBRID MASSIVE MIMO

7.1 Introduction

Hybrid beamforming for massive MIMO is a cost-effective solution to reduce the number of RF chains, which are prohibitively expensive, especially for mmWave scenarios, while enabling high-performance beamforming capabilities. To achieve further energy efficiency and cost reduction, use of highly non-linear and low-cost power amplifiers operating close to saturation can be necessary when the number of antennas is large. This necessitates the employment of high-performance, yet computationally low-complex, digital predistortion (DPD) that can work effectively for hybrid massive MIMO structures. In this study, we propose a vector digital predistortion algorithm for fully-connected massive MIMO. Our results demonstrate that we can outperform the existing DPD solutions for fully-connected hybrid MIMO in terms of out-of-band emissions, the error-vector magnitude of the received signals, and computational complexity.

7.2 Literature Review and Contributions

Extensive research has been conducted in the field of linearization of power amplifiers, and it has been established through various studies that the implementation of digital pre-distortion (DPD) techniques can effectively meet the necessary linearity constraints [38]. Particularly for single PA linearization, inverse learning architecture (ILA) is the state-of-the-art approach, which calculates DPD parameters by estimat-

ing the inverse of the PA characteristics [38]. However, the design of the DPD unit for hybrid beamforming systems is a nontrivial as there is no direct path between the digital front-end and PA. Moreover, the effectiveness of DPD is closely tied to the connection structure of the analog beamformer [3]. For instance, when using sub-array connected analog beamformers, a dedicated DPD system per RF chain is feasible [3, 99, 100], as each antenna is linked solely to one RF chain, allowing the implementation of a DPD system that is specifically tailored to the characteristics of that particular RF chain [3, 99, 100]. However, individual DPD per RF chain does not provide sufficient linearization for fully connected analog beamformers since each antenna input is formed by the linear combination of the signals of all RF chains, and in order to achieve the required linearization performance, vector DPD methods are recently proposed [76, 101–103].

In the studies [101–103], a multi-stream DPD method, where the basis functions for DPD implementation are generated as the products of the signals on all RF chains, is proposed. However, the computational complexity of the multi-stream DPD methods in [101] and [102] escalates exponentially with the increasing number of data streams, posing a significant challenge. The complexity of [103] increases with the number of transmit antennas, which must be avoided for large arrays. In this letter, we propose a correlation-based vector DPD method whose complexity is independent of the number of antennas, and grows almost linearly with the number of users. The complexity reduction of the proposed algorithm relies on a greedy virtual DPD branch selection, based on the correlation between the transmit antennas and a two-stage linear transformation onto the reduced dimensional space to achieve a similar predistortion output compared to a full complexity vector DPD solution. The proposed vector DPD (vDPD) achieves a remarkable complexity reduction and outperforms the benchmark DPD techniques for fully-connected hybrid MIMO systems.

7.3 System Model

7.3.1 Multi-carrier Downlink Transmission

In this study, we consider a mmWave massive MIMO system with hybrid beamforming (BF) architecture, where a base station with M antennas and D RF chains serves $U \leq D$ single-antenna users simultaneously. Data symbol at k^{th} subcarrier for user u is denoted by $a_u[k]$, and digital precoder $\mathbf{F}_{\text{D}}[k] \in \mathbb{C}^{D\times U}$ is applied to each subcarrier. The precoded signal $\mathbf{s}[k] \in \mathbb{C}^{D\times 1}$ at k^{th} subcarrier is expressed as

$$\mathbf{s}[k] = \mathbf{F}_{\text{D}}[k]\mathbf{a}[k], \quad k = -N_A/2, \ldots, N_A/2 - 1, \tag{7.1}$$

where $[\mathbf{a}[k]]_u = a_u[k]$, and N_{A} is the number of active subcarriers. The time domain signal can be obtained via inverse Fourier transform for the d^{th} RF chain as

$$\hat{s}_{d,n} = \frac{1}{N_{\text{A}}} \sum_{k=-N_{\text{A}}/2}^{N_{\text{A}}/2-1} s_d[k] e^{j\frac{2\pi}{N}kn}, \tag{7.2}$$

where $N > N_{\text{A}}$ is the total number of subcarriers, and $s_d[k] = [\mathbf{s}[k]]_d$. After proper cyclic prefix insertion and windowing on $\hat{s}_{d,n}$, the time domain signal $s_{d,n}$ at time n is obtained in Figure 7.1. The analog beamforming output can be represented as

$$\mathbf{x}_n = \mathbf{F}_{\text{RF}}\mathbf{s}_n, \tag{7.3}$$

where $\mathbf{s}_n \triangleq [s_{1,n}, \ldots, s_{D,n}]^{\text{T}}$. Prior to transmission, the beamformed signal undergoes amplification via PAs, yielding nonlinear distortion. For brevity, we represent the nonlinear PA output for m^{th} antenna element as $y_{m,n} = \psi(x_{m,n})$, $x_{m,n} \triangleq [\mathbf{x}_n]_m$. Then, the received signal for u^{th} user becomes

$$r_n^u = \sum_{l=0}^{L-1} \mathbf{h}_{u,l}^{\text{H}} \mathbf{y}_{n-l}, \tag{7.4}$$

where $\mathbf{h}_{u,l}$ is the l^{th} multipath component of u^{th} user channel. To have a practical wideband mmWave channel model, we use the Saleh–Valenzuela channel model [104]. Then, the channel response vector $\mathbf{h}_{u,l}$ can be expressed as

$$\mathbf{h}_{u,l} = \sum_{c=1}^{C} \sum_{r=1}^{R} g_{u,cr} \rho(l - l_{u,c} - l_{u,r}) \mathbf{a}(\theta_{u,cr}), \tag{7.5}$$

where $g_{u,cr}$ is the complex gain for r^{th} ray of c^{th} cluster with Rician distribution $g_{u,cr} \sim \mathcal{CN}(\kappa_{u,cr}, \sigma^2_{u,cr})$, $\rho(\cdot)$ is the band-limitation function, $\theta_{u,cr} = \vartheta_{u,c} + \frac{r-1}{R}\Delta_{u,c} - \frac{\Delta_{u,c}}{2}$ is the angle of arrivals (AoA) of r^{th} ray of c^{th} cluster, $\vartheta_{u,c}$ is the mean AoA, $\Delta_{u,c}$ is the angular spread, and $\mathbf{a}(\theta)$ is the array manifold vector with m^{th} element $[\mathbf{a}(\theta)]_m = \frac{1}{\sqrt{M}} e^{j\frac{2\pi}{\lambda_c} d_{\text{a}} m \sin\theta}$, where λ_c is the carrier wavelength and d_{a} is antenna spacing.

7.3.2 Nonlinear PA and Inverse PA Models

In order to have an accurate model, we use a generalized memory polynomial (GMP) to represent nonlinear PA characteristics $y_{m,n} = \psi(x_{m,n})$, which is given as [38]

$$y_{m,n} = \sum_{k=0}^{K-1} \sum_{l=0}^{L_{\text{PA}}-1} \sum_{w=-W+1}^{W-1} a_{k,l,w} x_{m,n-l} |x_{m,n-l+w}|^k, \tag{7.6}$$

where $a_{k,l,w}$ are the weight coefficients of the PA model. To achieve linearity in the PA output, it is desired to employ an appropriate DPD technique that can accurately model and compensate for the nonlinear behavior of the amplifiers. For this purpose, we have employed DPD structure based on GMP basis functions, and inverse learning architecure is used to obtain the weight coefficients for the per-antenna DPD model $\phi(\cdot)$, which can be represented as $\phi(\cdot) = \psi^{-1}(\cdot)$ [38].

7.4 Vector DPD (vDPD) for Fully Connected Hybrid Beamforming Architecture

7.4.1 Full Complexity Array Projection Based vDPD

The core of this method is in the linearization of each PA individually. Therefore, to find multi-stream DPD, the analog beamforming should be replicated in the digital domain to generate the *desired* transmitted signal as $\mathbf{x}_n = \mathbf{F}_{\text{RF}} \mathbf{s}_n$. The main objective is to find the DPD output vector $\tilde{\mathbf{s}}_n$ in Figure 7.1, which minimizes $||\mathbf{x}_n -$

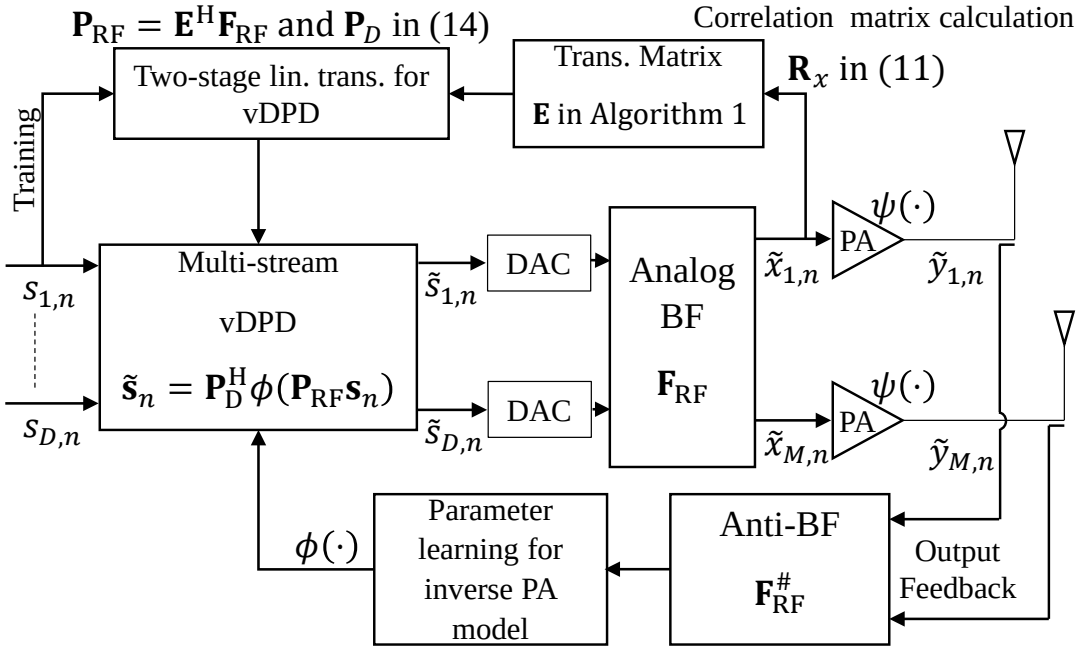


Figure 7.1: Proposed reduced complexity correlation based two-stage vDPD architecture

$\psi\left(\mathbf{F}_{\mathrm{RF}}\tilde{\mathbf{s}}_n\right)||_2^2$. Instead of obtaining optimal $\tilde{\mathbf{s}}_n$ directly from this minimization, it is more convenient to re-formulate the problem in an *indirect* fashion by employing *virtual* per-antenna DPD outputs $\tilde{\mathbf{y}}_n^{\mathrm{v}} \triangleq \phi(\mathbf{x}_n) = \psi^{-1}(\mathbf{x}_n)$ as

$$\underset{\tilde{\mathbf{s}}_n}{\text{minimize}}\ ||\tilde{\mathbf{y}}_n^{\mathrm{v}} - \mathbf{F}_{\mathrm{RF}}\tilde{\mathbf{s}}_n||_2^2. \tag{7.7}$$

Then, the least-squares (LS) solution of (7.7) yields the vDPD output in D-dimensional space:

$$\tilde{\mathbf{s}}_n = \mathbf{P}_{\mathrm{F}}\phi\left(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_n\right), \tag{7.8}$$

where $\mathbf{P}_{\mathrm{F}} = \left(\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{F}_{\mathrm{RF}}\right)^{-1}\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}$ is the linear transformation matrix, which is obtained to have the desired transmitted vector at the antenna array output as a solution of the *equivalent* optimization problem (given digital precoder output $\mathbf{s}_n$):

$$\underset{\mathbf{P}_{\mathrm{F}}}{\text{minimize}}\ ||\phi(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_n) - \mathbf{F}_{\mathrm{RF}}\mathbf{P}_{\mathrm{F}}\phi(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_n)||_2^2. \tag{7.9}$$

It should be noted that the synthesized vDPD output after analog beamforming, namely $\tilde{\mathbf{x}}_n = \mathbf{F}_{\mathrm{RF}}\tilde{\mathbf{s}}_n$, can be expressed as $\tilde{\mathbf{x}}_n = \mathbf{F}_{\mathrm{RF}}\left(\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{F}_{\mathrm{RF}}\right)^{-1}\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}\phi\left(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_n\right)$, where the term $\mathbf{P}_{\mathrm{P}} \triangleq \mathbf{F}_{\mathrm{RF}}\left(\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{F}_{\mathrm{RF}}\right)^{-1}\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}$ is a projection matrix. It projects the full dimensional desired virtual vDPD output, namely $\tilde{\mathbf{y}}_n^{\mathrm{v}} = \phi\left(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_n\right)$, onto the range space of $\mathbf{F}_{\mathrm{RF}}$ by reducing the signaling dimension from M to D.

7.4.2 Proposed Low Complexity Correlation-Based vDPD

It is known that nonlinear distortion in antenna arrays is correlated when antenna inputs are correlated, which is especially noticeable in massive MIMO systems with a large number of antennas relative to users (where $M \gg U$) [23].

In order to reduce the high complexity of (7.8), which requires the generation of $\tilde{y}^{\mathrm{v}}_{m,n} = \phi\left(x_{m,n}\right)$ for each antenna element, we propose a method to form the virtual PA inputs required for vDPD synthesis. Due to the high correlation among the antenna elements, the required number of virtual per-antenna DPD outputs in (7.8) can be significantly reduced by using a linear transformation matrix, with a size independent of M. This implies a *two-stage* processing for reduced complexity vDPD synthesis summarized in Figure 7.1.

In the *first stage*, virtual analog beamforming output is generated via the following transformation matrix $\mathbf{P}_{\mathrm{RF}} \in \mathbb{C}^{P\times D}$:

$$\mathbf{P}_{\mathrm{RF}} = \mathbf{E}^{\mathrm{H}}\mathbf{F}_{\mathrm{RF}}, \tag{7.10}$$

where $P < M$ is the number of per-antenna DPD operations to be performed for reduced complexity vDPD construction. In (7.10), $\mathbf{E} \in \mathbb{C}^{M\times P}$ is the antenna selection matrix, whose columns are composed of elementary vectors. The index of non-zero element in each column of $\mathbf{E}$ specifies the selected row of $\mathbf{F}_{\mathrm{RF}}$. For a given P, matrix $\mathbf{E}$ is to be constructed such that the maximum value among the entries of $\mathbf{E}^{\mathrm{H}}\mathbf{R}_x\mathbf{E}$ is minimized, where $\mathbf{R}_x \triangleq \mathbb{E}\{\mathbf{x}_n\mathbf{x}_n^H\}$ for $\mathbf{x}_n = \mathbf{F}_{\mathrm{RF}}\mathbf{s}_n$ is the correlation matrix for the actual array output (without vDPD). It can be approximated as in the following proposition:

Proposition 1. $\mathbf{R}_x \triangleq \mathbb{E}\{\mathbf{x}_n\mathbf{x}_n^H\}$ *can be approximated as*

$$\mathbf{R}_x \approx \mathbf{F}_{\mathrm{RF}}\mathbf{B}^{-1}\mathbf{F}_{\mathrm{RF}}^{\mathrm{H}} \tag{7.11}$$

where $\mathbf{B} = \mathrm{diag}\left\{\beta_1, \ldots, \beta_U\right\}$.

Proof. See Appendix B. □

In order to form the antenna selection matrix $\mathbf{E}$, the antenna index set $\mathcal{S}$ with cardinality P, namely $|\mathcal{S}| = P$, is constructed. The $\mathcal{S}\{p\}^{\mathrm{th}}$ entry of the p^{th} column of $\mathbf{E}$

is set to 1, i.e. $[\mathbf{E}]_{(\mathcal{S}\{p\},p)} := 1$ for $p = 1, \ldots, P$. By using these definitions, matrix $\mathbf{E}$ is constructed according to Algorithm 1. Consequently, the new element to be added to the set $\mathcal{S}$ is determined such that the maximum pairwise correlation among the selected elements is minimized.

Algorithm 1 Construction of $\mathbf{E}$

Input: P, $\mathbf{F}_{\mathrm{RF}}$, $\mathbf{B}$

1: Set $\mathcal{S}\{1\} := 1$ {Since $\mathbf{R}_x$ is Toeplitz, one can always start with the first antenna.}

2: $p := 2$

3: **while** $p \leq P$ **do**

4: $\quad \mathcal{S}\{p\} = \underset{n, n \neq m}{\operatorname{argmin}} \underset{m}{\max} \left| [\mathbf{R}_x]_{(m,n)} \right|$ for $m \in \mathcal{S}\{1 : p-1\}$

5: **end while**

After obtaining $\mathbf{P}_{\mathrm{RF}}$, the *virtual* array output in reduced dimension is formed as $\tilde{\mathbf{x}}_n^{\mathrm{v}} = \mathbf{P}_{\mathrm{RF}}\mathbf{s}_n$ at the *first stage* of the proposed vDPD algorithm in Figure 7.1. Then, per-antenna DPD is performed over each element of $\tilde{\mathbf{x}}_n^{\mathrm{v}}$ as $\tilde{y}_{p,n}^{\mathrm{v}} = \phi(\tilde{x}_{p,n}^{\mathrm{v}})$, where $\tilde{x}_{p,n}^{\mathrm{v}} = [\tilde{\mathbf{x}}_n^{\mathrm{v}}]_p$ for $p = 1, \ldots, P$. Finally, the *virtual* per-antenna DPD output $\tilde{\mathbf{y}}_n^{\mathrm{v}} = [\tilde{y}_{1,n}^{\mathrm{v}}, \ldots, \tilde{y}_{P,n}^{\mathrm{v}}]^{\mathrm{T}}$ is projected onto D-dimensional space via a transformation matrix $\mathbf{P}_{\mathrm{D}}$ at the *second stage* of the proposed two-stage vDPD as follows:

$$\tilde{\mathbf{s}}_n = \mathbf{P}_{\mathrm{D}}^{\mathrm{H}}\tilde{\mathbf{y}}_n^{\mathrm{v}} = \mathbf{P}_{\mathrm{D}}^{\mathrm{H}}\phi(\tilde{\mathbf{x}}_n^{\mathrm{v}}) = \mathbf{P}_{\mathrm{D}}^{\mathrm{H}}\phi(\mathbf{P}_{\mathrm{RF}}\mathbf{s}_n). \tag{7.12}$$

The linear transformation matrix $\mathbf{P}_{\mathrm{D}} \in \mathbb{C}^{P \times D}$ can be acquired during a training phase with a training sequence $\{\mathbf{s}_n\}_{n=0}^{L_{\mathrm{T}}-1}$ of length L_{T}, and the observation sequence can be generated by using the full complexity projected vDPD output in (7.8). Then, the matrix $\mathbf{P}_{\mathrm{D}}$ can be found such that it minimizes the Frobenius norm of the difference between the full complexity vDPD in (7.8) and the proposed vDPD method in (7.12) as

$$J = \left|\left| \mathbf{P}_{\mathrm{F}}\tilde{\mathbf{Z}} - \mathbf{P}_{\mathrm{D}}^{\mathrm{H}}\tilde{\mathbf{Y}}^{\mathrm{v}} \right|\right|_{\mathrm{F}}^2, \tag{7.13}$$

where $\tilde{\mathbf{Z}} = [\phi(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_0), \ldots, \phi(\mathbf{F}_{\mathrm{RF}}\mathbf{s}_{L_{\mathrm{T}}-1})]$, and $\tilde{\mathbf{Y}}^{\mathrm{v}} = \left[\tilde{\mathbf{y}}_0^{\mathrm{v}}, \ldots, \tilde{\mathbf{y}}_{L_{\mathrm{T}}-1}^{\mathrm{v}}\right]$ for $\tilde{\mathbf{y}}_n^{\mathrm{v}} = \phi(\mathbf{P}_{\mathrm{RF}}\mathbf{s}_n)$. By using LS method, the linear transformation matrix $\mathbf{P}_{\mathrm{D}}$ can be found as

$$\mathbf{P}_{\mathrm{D}} = \left(\left[\tilde{\mathbf{Y}}^{\mathrm{v}} \right]^{\mathrm{H}} \right)^{\dagger} \tilde{\mathbf{Z}}^{\mathrm{H}}\mathbf{P}_{\mathrm{F}}^{\mathrm{H}}, \tag{7.14}$$

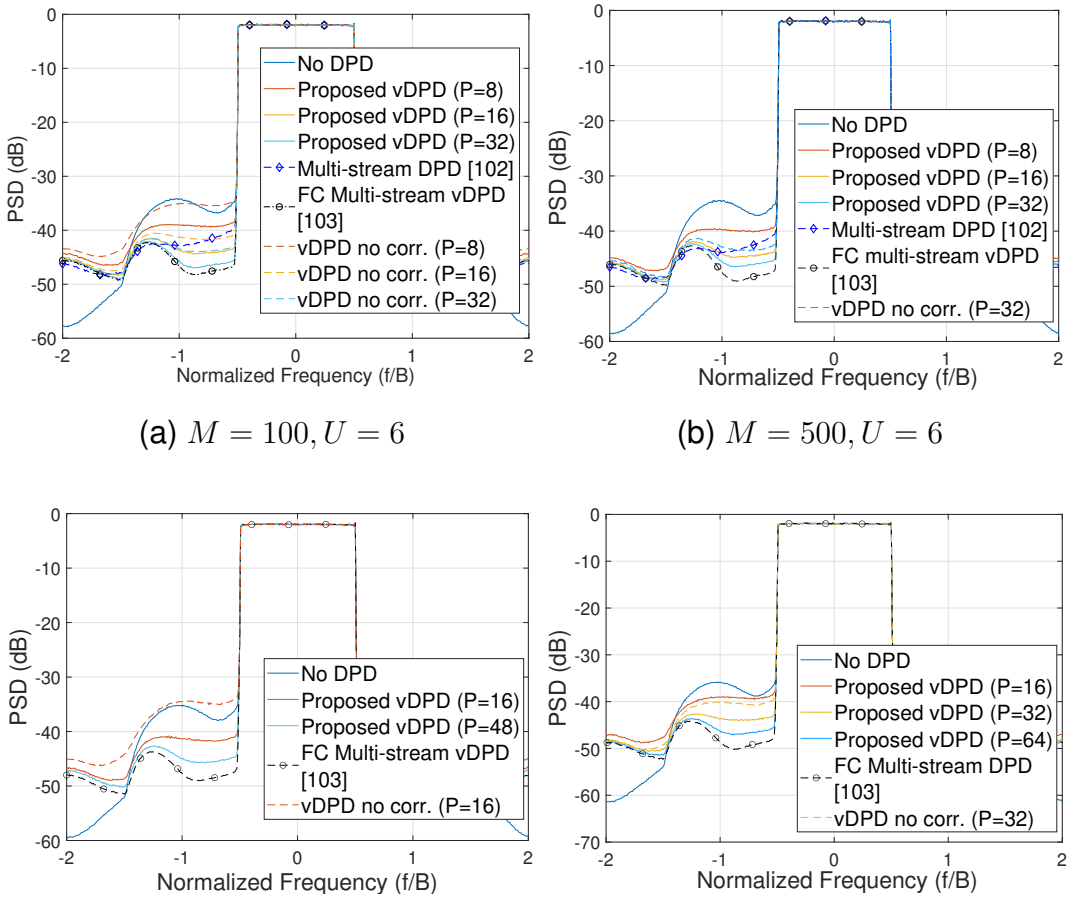
where † represents pseudo-inverse operation.

7.5 Computational Complexity

The computational complexity of the proposed algorithm can be assessed by investigating (7.12) and per-antenna DPD operation $\tilde{y}^{\mathrm{v}}_{p,n} = \phi(\tilde{x}^{\mathrm{v}}_{p,n})$. The required number of complex multiplications for virtual beamforming and back-projection operations defined in (7.12) is $\mathcal{O}(PD)$, individually. For each virtual antenna branch, the required number of complex computations is $\mathcal{O}_{\mathrm{GMP}} \triangleq \mathcal{O}(KL_{\mathrm{PA}}) + 2\mathcal{O}(KL_{\mathrm{PA}}W)$, which yields a total of $P\mathcal{O}_{\mathrm{GMP}}$ complex multiplications for virtual DPD outputs. Then, the total computational complexity can be expressed as $\mathcal{O}_{\mathrm{C}} = P\mathcal{O}_{\mathrm{GMP}} + 2\mathcal{O}(PD)$. The full complexity multi-stream DPD solution in [103]; on the other hand, has the complexity of $\mathcal{O}_{\mathrm{C}} = M\mathcal{O}_{\mathrm{GMP}} + 2\mathcal{O}(MD)$. It is readily observed the multi-stream DPD in [103] has higher computational complexity compared to the proposed method by the factor of M/P. On the other hand, the required number of multiplications is tremendous for the method proposed in [102] since the basis functions are generated by multiplexing signals of D data streams. For instance, for the same GMP parameters given in Section 7.6, the required number of complex multiplications for the proposed algorithm for $U = 6$ users, where P is selected as 32, is 1280. The number of complex multiplications for the multi-stream DPD in [102] is 2448, and that of multi-stream DPD in [103] is 20000 for $M = 500$ and 4000 for $M = 100$ antennas, respectively. It is important to note that the required complexity of the proposed vDPD is independent of the number of antennas. However, for $U = 12$, the required number of complex multiplications for the proposed algorithm, where P is chosen as 64, is 2560; however, multi-stream DPD in [102] requires 16236 complex multiplications.

7.6 Numerical Results

The performances of the DPD methods are evaluated via several performance criteria. The proposed vDPD method is compared with the state-of-the-art methods that are multi-stream DPD methods in [102] and [103]. We consider an OFDM waveform with $N_{\mathrm{A}} = 1024$ active subcarriers and oversampling factor 4 ($N = 4096$). User symbols are selected from the 64 QAM alphabet. PA models are extracted from



(a) $M = 100, U = 6$ (b) $M = 500, U = 6$

(c) $M = 100, U = 12$ (d) $M = 500, U = 12$

Figure 7.2: PSD for $M = 100/500$, $U = 6/12$.

the measurements from RF Weblab by using 50 MHz OFDM signal [2] with GMP parameters $K = 4$, $L = 3$, $M = 3$.

In order to assess the performance of the proposed reduced complexity vDPD method, power spectral density (PSD) of the received signal is evaluated for the proposed and benchmark methods. Various system configurations are considered to observe the effects of the system parameters. For this purpose, we select the number of antennas as $M = 100$ or $M = 500$, and the number of users as $U = 6$ or $U = 12$, and the user locations are generated randomly in a 1×1 km^2 region, with MPC distribution having 2° angular spread and 10 dB Rician factor. In Figure 7.2, the PSD of the received signal of the edge user is demonstrated, which is more critical from an inter-cell interference perspective. In Figure 7.2 (a), $M = 100$ antennas and $U = 6$ users

case is studied, where the proposed vDPD approach is evaluated for different number of virtual DPD branches, namely P. It can be observed from Figure 7.2 (a) that the proposed method outperforms the multi-stream DPD method in [102], despite having lower computational complexity. In addition, the necessity of the correlation-based antenna selection algorithm is shown by choosing the virtual antenna elements randomly without considering the antenna correlations. It can be seen that virtual array design, without considering the input correlation, yields significant performance degradation, especially for lower P values. To show that the required complexity for the proposed vDPD method is independent of the number of antennas, we consider $M = 500$ antennas at the BS. From Figure 7.2 (b), it can be observed that by using the same P value, the proposed reduced complexity vDPD algorithm achieves a similar performance with the $M = 100$ scenario. Moreover, if the maximum PSD values at the out-of-band (OOB) section are compared, the proposed vDPD provides similar performance compared to the full complexity DPD solution in [103], despite having much less complexity.

In order to observe the effects of the number of users, a scenario with $U = 12$ users is also considered, where the number of antennas is chosen as $M = 100$ and $M = 500$, as in the previous setup. It can be observed from Figure 7.2 (c) and Figure 7.2 (d), the number of required virtual DPDs increases for larger number of users since the distortion correlation between antennas is reduced, which necessitates using a larger number of virtual DPD units. Furthermore, for the same P value, similar vDPD performance is achieved for different numbers of antennas. Moreover, comparing $U = 6$ and $U = 12$ cases, the number of necessary vDPD branches P scale gracefully (almost linearly) with the number of users for the proposed vDPD method, implying a linear complexity increase of the proposed vDPD method with the number of users, rather than an exponential increase. The performance of multi-stream DPD in [102] is not included for $U = 12$ for conciseness since it suffers from a significant performance loss. The reason is that the required number of basis functions are tremendous for $U = 12$, thus singularity problems occur in matrix inversions to find the weight coefficients for the DPD proposed in [102].

In addition to PSD performances at a specific direction, the angular distribution of the OOB distortion is also obtained. To simplify the visualization, we consider a

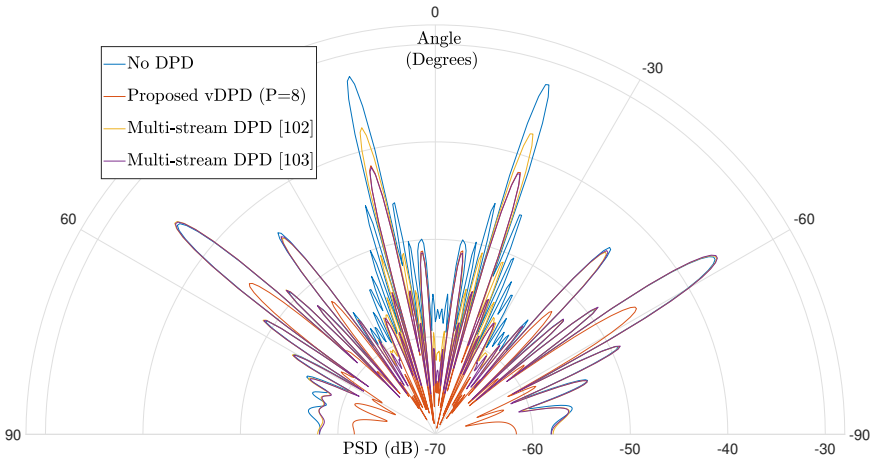
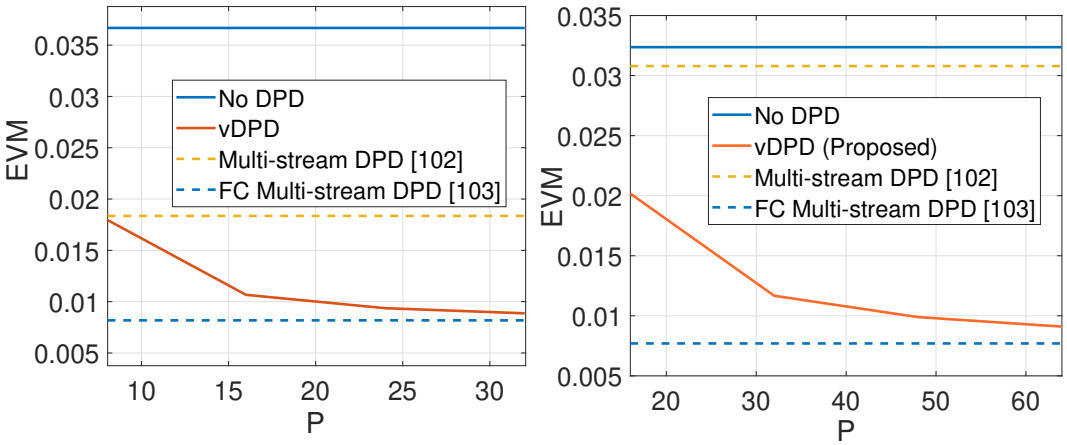


Figure 7.3: Angular distribution of the OOB distortion.

scenario with $M = 50$ antennas at the BS, serving $U = 2$ users. In Figure 7.3, the radiation pattern of the total OOB distortion power on the adjacent frequency band is shown. It is observed that state-of-the-art DPD methods for fully connected array architecture can only mitigate the nonlinear distortion along the intended user directions. However, we also achieve OOB suppression along the *spatial harmonic* angles, in which the nonlinear distortion is coherently combined, by extending the analog beamformer to direct beams along the spatial harmonic angles. It is also observed that the multi-stream DPD method in [102] suffers from a significant performance loss compared to the proposed vDPD algorithm. Moreover, the proposed reduced complexity vDPD provides sufficient OOB suppression, which is very close to the full complexity DPD method in [103] in the intended user directions. It is important to note that the proposed DPD method can mitigate the unwanted OOB radiation along angles other than the intended user direction, especially on the spatial harmonic angles, unlike [102], [103].

The received signal quality is also assessed for the intended users by measuring the error vector magnitude (EVM) for the received data symbol estimates. For this purpose, the scenarios for $M = 500$ antennas and $U = 6/12$ user are considered. It is seen that for $U = 6$ and $U = 12$ case, the proposed algorithm provides almost optimal performance with $P = 32$ and $P = 64$, respectively, which means that it provides almost the same performance with multi-stream DPD in [103] with more than 80% complexity reduction. Moreover, the performance of the multi-stream DPD in [102]



(a) $U = 6$ (b) $U = 12$

Figure 7.4: EVM vs. P for $M = 500$, $U = 6/12$.

is very inferior (especially for $U = 12$) compared to the proposed vDPD.

7.7 Conclusion

Hybrid MIMO architecture is a key enabler for cost efficient implementation of massive MIMO systems with superior beamforming capabilities. A further reduction in system cost can be through employment of low-cost power amplifiers with high-nonlinearity, whose negative impact on the system performance can be mitigated by effective digital predistortion methods. In this study, a vector DPD algorithm for fully-connected hybrid MIMO architectures is proposed. The proposed algorithm outperforms the existing DPD algorithms for fully-connected hybrid MIMO arrays in terms of OOB emissions, EVM and computational complexity.

CHAPTER 8

CONCLUSIONS

In conclusion, this thesis has made significant contributions to the field of signal processing for addressing the challenges posed by power amplifier nonlinearities in communication systems. The research presented in this work has explored various aspects of nonlinear distortion and proposed effective mitigation strategies, thereby advancing the understanding and practical implementation of communication systems with non-ideal power amplifiers.

Firstly, the thesis investigated the effects of nonlinear distortion in both uplink and downlink communication scenarios, considering the integration of non-ideal power amplifiers in user equipment and base stations. Unlike previous studies, this research acknowledged the implications of pulse shaping filters and aliasing due to downsampling in the SC modulation scheme used in uplink communication. By considering the nonlinear inter-symbol interference caused by amplifier nonlinearity, the thesis developed nonlinear post-distortion schemes and distortion modeling techniques to mitigate the effects of nonlinear distortion at the receiver side.

Furthermore, the thesis extended its scope to explore full-duplex communication systems, providing insights into the complexities and potential solutions associated with simultaneous bidirectional transmission. It introduced the concept of distortion amplification, which arises due to the mismatch between the effective channels in the downsampled domain of the desired and distortion signals. To mitigate this effect and elevate diversity in the temporal domain, a distortion-aware receiver structure based on fractional sampling was developed.

Moreover, the thesis addressed the challenges posed by nonlinear distortion in multi-

user MIMO systems with hybrid beamforming hardware structures. By employing nonlinear distortion cancellation and the distortion estimation method proposed earlier, a sophisticated distortion-aware receiver was designed, significantly reducing complexity while maintaining performance. Additionally, the thesis conducted an extensive analysis of signal-to-distortion-plus-noise (SDNR) and investigated the effects of distortion amplification analytically.

The research also considered the radiation of nonlinear distortion from multi-antenna base stations, which had not been extensively explored in the literature. It investigated the validity of the assumption of uncorrelated distortion across the antenna array in massive MIMO systems and developed a method for finding optimal power allocation coefficients considering distortion correlation. These findings contribute to a better understanding of the quantitative analysis of received distortion power and provide insights for future system design.

Lastly, the thesis addressed the challenges of designing digital pre-distortion (DPD) algorithms for fully connected hybrid beamforming systems, which integrate analog and digital components. It proposed a correlation-based reduced complexity DPD design that avoids redundant signal processing operations in correlated antenna branches, thereby achieving satisfactory performance while reducing complexity.

In summary, this thesis has made significant contributions to the field of signal processing for communication systems with power amplifier nonlinearities. The proposed frameworks, mitigation strategies, and receiver structures offer valuable insights and practical solutions for improving system performance and managing nonlinear distortion effects in various communication scenarios. These contributions pave the way for future research and advancements in the field, ultimately leading to more efficient and robust communication systems in the presence of power amplifier nonlinearities.

There are several promising avenues for extending the proposed frameworks to further enhance performance and reduce computational complexity. Firstly, an online learning approach could be developed to continuously track the time-varying behavior of PAs. This would allow for real-time adaptation to changing PA characteristics, leading to improved system performance. Additionally, the length of the ST sequence

required for distortion modeling can be reduced by employing decision feedback techniques, effectively optimizing the trade-off between accuracy and computational overhead.

Secondly, to mitigate distortion amplification, modifications to the spectral content of the transmitted signal can be explored. By strategically shaping the frequency band used for data transmission, it is possible to avoid regions where distortion amplification occurs, leading to improved signal quality at the receiver.

Thirdly, investigating the effects of the distortion amplification phenomenon on OFDM modulated signals holds great potential for understanding its impact on modern communication systems. By analyzing the interplay between distortion amplification and OFDM, new insights can be gained and novel solutions devised.

Lastly, for downlink communication, an integrated approach can be pursued, combining distortion-aware precoding with the proposed power allocation scheme. This joint design would optimize the transmission process, resulting in enhanced received signal quality and overall system performance.

In conclusion, exploring these academic research directions can lead to valuable advancements in signal processing for communication systems, enabling better management of nonlinear distortion effects and paving the way for more efficient and robust communication technologies.

REFERENCES

[1] Ericsson, "Further elaboration on PA models for NR," tech. rep., Ericsson AB, Stockholm, Sweden, Aug. 2016.

[2] P. N. Landin, S. Gustafsson, C. Fager, and T. Eriksson, "WebLab: A Web-Based Setup for PA Digital Predistortion and Characterization [Application Notes]," *IEEE Microw. Mag. MultiMedia*, vol. 16, no. 1, pp. 138–140, 2015.

[3] A. Brihuega, L. Anttila, M. Abdelaziz, T. Eriksson, F. Tufvesson, and M. Valkama, "Digital Predistortion for Multiuser Hybrid MIMO at mmWaves," *IEEE Trans. Signal Process.*, vol. 68, pp. 3603–3618, 2020.

[4] S. Sullivan, A. Brighente, S. A. P. Kumar, and M. Conti, "5G Security Challenges and Solutions: A Review by OSI Layers," *IEEE Access*, vol. 9, pp. 116294–116314, 2021.

[5] U. Gustavsson, P. Frenger, C. Fager, T. Eriksson, H. Zirath, F. Dielacher, C. Studer, A. Pärssinen, R. Correia, J. N. Matos, D. Belo, and N. B. Carvalho, "Implementation Challenges and Opportunities in Beyond-5G and 6G Communication," *IEEE J. Microw.*, vol. 1, no. 1, pp. 86–100, 2021.

[6] I. Ross, "Wireless network directions," *IEEE Commun. Mag.*, vol. 29, no. 2, pp. 40–42, 1991.

[7] J. Qiao, X. S. Shen, J. W. Mark, Q. Shen, Y. He, and L. Lei, "Enabling device-to-device communications in millimeter-wave 5G cellular networks," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 209–215, 2015.

[8] H.-h. Chen, M. Guizani, and W. Mohr, "Evolution toward 4G wireless networking [Guest Editorial]," *IEEE Network*, vol. 21, no. 1, pp. 4–5, 2007.

[9] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What Will 5G Be?," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[10] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, 2016.

[11] A. Liu and V. Lau, "Hierarchical Interference Mitigation for Massive MIMO Cellular Networks," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4786–4797, 2014.

[12] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 106–112, 2014.

[13] "3GPP TS36.141 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Conformance Testing (Release 15)," tech. rep., 3GPP Std., Rev. V15.3.0, 2018.

[14] R. W. Heath, G. Laus, T. Q. Quek, S. Talwar, and P. Zhou, "Signal Processing for the 5G Revolution [From the Guest Editors]," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 12–13, 2014.

[15] P. K. Singya, N. Kumar, and V. Bhatia, "Mitigating NLD for Wireless Networks: Effect of Nonlinear Power Amplifiers on Future Wireless Communication Networks," *IEEE Microw. Mag.*, vol. 18, no. 5, pp. 73–90, 2017.

[16] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 30–38, 2006.

[17] C. Dehos, J. L. Gonzalez, A. D. Domenico, D. Ktenas, and L. Dussopt, "Millimeter-wave access and backhauling: The solution to the exponential data traffic increase in 5G mobile communications systems.," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 88–95, 2014.

[18] W. Hong, K. Baek, Y. Lee, Y. Kim, and S. Ko, "Study and prototyping of practically large-scale mmWave antenna systems for 5G cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63–69, 2014.

[19] H. Dabag, B. Hanafi, O. D. Gurbuz, G. M. Rebeiz, J. F. Buckwalter, and P. M. Asbeck, "Transmission of Signals With Complex Constellations Using Millimeter-Wave Spatially Power-Combined CMOS Power Amplifiers and Digital Predistortion," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 7, pp. 2364–2374, 2015.

[20] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.

[21] T. L. Marzetta, "Massive MIMO: An Introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[22] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson, "Out-of-band radiation measure for MIMO arrays with beamformed transmission," in *IEEE International Conf. on Commun. (ICC)*, pp. 1–6, 2016.

[23] C. Mollén, U. Gustavsson, T. Eriksson, and E. G. Larsson, "Spatial Characteristics of Distortion Radiated From Antenna Arrays With Transceiver Nonlinearities," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6663–6679, 2018.

[24] D. Korpi, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Digital self-interference cancellation under nonideal RF components: Advanced algorithms and measured performance," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, pp. 286–290, 2015.

[25] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-Band Full-Duplex Wireless: Challenges and Opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, 2014.

[26] D. Kim, H. Lee, and D. Hong, "A Survey of In-Band Full-Duplex Transmission: From the Perspective of PHY and MAC Layers," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2017–2046, 2015.

[27] P. P. Campo, L. Anttila, D. Korpi, and M. Valkama, "Adaptive Cancellation of Nonlinear Self-Interference in Wireless Full-Duplex: Cascaded Spline-

Interpolated Methods," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, pp. 1265–1271, 2020.

[28] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive Digital Predistortion of Wireless Power Amplifiers/Transmitters Using Dynamic Real-Valued Focused Time-Delay Line Neural Networks," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 95–104, 2010.

[29] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "Augmented Real-Valued Time-Delay Neural Network for Compensation of Distortions and Impairments in Wireless Transmitters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 242–254, 2019.

[30] M. Rawat, K. Rawat, and F. M. Ghannouchi, "Adaptive Digital Predistortion of Wireless Power Amplifiers/Transmitters Using Dynamic Real-Valued Focused Time-Delay Line Neural Networks," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 95–104, 2010.

[31] F. Perez-Cruz, J. J. Murillo-Fuentes, and S. Caro, "Nonlinear channel equalization with Gaussian processes for regression," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5283–5286, 2008.

[32] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE J. Sel. Areas Commun.*, vol. 1, no. 1, pp. 57–62, 1983.

[33] M. B. Salman and G. M. Guvensen, "An Efficient QAM Detector via Nonlinear Post-distortion based on FDE Bank under PA Impairments," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7108–7120, 2021.

[34] A. Kurt, M. B. Salman, U. B. Sarac, and G. M. Guvensen, "An Adaptive-Iterative Nonlinear Interference Cancellation in Time-Varying Full-Duplex Channels," *IEEE Trans. Veh. Technol.*, vol. 72, no. 2, pp. 1862–1878, 2023.

[35] M. B. Salman, G. M. Guvensen, and T. Ciloglu, "Low complexity nonlinear detection for multiuser hybrid mimo systems and performance analysis," *Physical Communication*, p. 102126, 2023.

[36] M. B. Salman, E. Björnson, Guvensen, G. M., and T. Ciloglu, "Analytical Nonlinear Distortion Characterization for Frequency-Selective Massive MIMO Channels," in *Proc. ICC*, pp. 1–6, 2023.

[37] A. A. M. Saleh, "Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, 1981.

[38] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A Generalized Memory Polynomial Model for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Signal Process.*, vol. 54, pp. 3852–3860, Sep. 2006.

[39] J. Van De Beek and F. Berggren, "Out-of-Band Power Suppression in OFDM," *IEEE Commun. Lett.*, vol. 12, no. 9, pp. 609–611, 2008.

[40] G. Colavolpe and A. Piemontese, "Novel SISO Detection Algorithms for Nonlinear Satellite Channels," in *Proc. IEEE GLOBECOM*, pp. 1–5, 2011.

[41] Z. Gulgun and A. O. Yilmaz, "Detection schemes for high order M -Ary QAM under transmit nonlinearities," *IEEE Trans. Commun.*, vol. 67, pp. 4825–4834, Mar. 2019.

[42] D. Yoda and H. Ochiai, "Decision region optimization and metric-based compensation of memoryless nonlinearity for APSK systems," *IEEE Trans. Broadcast.*, vol. 64, no. 2, pp. 281–292, 2018.

[43] J. Choi, "Single-Carrier Index Modulation for IoT Uplink," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 6, pp. 1237–1248, 2019.

[44] S. Chen, X. Hong, E. F. Khalaf, A. Morfeq, N. D. Alotaibi, and C. J. Harris, "Single-Carrier Frequency-Domain Equalization With Hybrid Decision Feedback Equalizer for Hammerstein Channels Containing Nonlinear Transmit Amplifier," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3341–3354, 2017.

[45] W. Xu, Z. Zhong, Y. Be'ery, X. You, and C. Zhang, "Joint neural network equalizer and decoder," in *Proc. Int. Symp. Wireless Commun.Syst.*, pp. 1–5, 2018.

[46] X. Hong, S. Chen, C. J. Harris, and E. F. Khalaf, "Single-carrier frequency domain equalization for Hammerstein communication systems using complex-valued neural networks," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4467–4478, 2014.

[47] X. Hong and S. Chen, "Modeling of Complex-Valued Wiener Systems Using B-Spline Neural Network," *IEEE Trans. Neural Netw.*, vol. 22, no. 5, pp. 818–825, 2011.

[48] J. Guerreiro, R. Dinis, and P. Montezuma, "Low-Complexity SC-FDE Techniques for Massive MIMO Schemes With Low-Resolution ADCs," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2368–2380, 2019.

[49] J. Madeira, J. Guerreiro, and R. Dinis, "Iterative Frequency-Domain Detection for MIMO Systems with Strong Nonlinear Distortion Effects," in *2018 Int. Congress on Ultra Modern Telecommun. and Control Syst. and Workshops*, pp. 1–5, 2018.

[50] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, pp. 55–67, Jan. 1982.

[51] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai Shitz, "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, 1994.

[52] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, 2000.

[53] G. M. Guvensen, Y. Tanik, and A. O. Yilmaz, "On the carrier frequency offset estimation for frequency hopping burst mode mobile radio," in *Proc. 2010 MILCOM*, pp. 1244–1249.

[54] T. Araújo and R. Dinis, *Analytical Evaluation of Nonlinear Distortion Effects on Multicarrier Signals*. Boca Raton, FL, USA: CRC Press, 2015.

[55] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, 2002.

[56] F. Perez-Cruz, S. Van Vaerenbergh, J. J. Murillo-Fuentes, M. Lazaro-Gredilla, and I. Santamaria, "Gaussian processes for nonlinear signal processing: An overview of recent advances," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 40–50, 2013.

[57] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*. MIT press Cambridge, MA, 2006.

[58] S. S. Haykin, *Neural Networks and Learning Machines.* New York: Prentice Hall, 2009.

[59] O. T. Demir and E. Bjornson, "The Bussgang Decomposition of Nonlinear Systems: Basic Theory and MIMO Extensions [Lecture Notes]," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 131–136, 2021.

[60] O. T. Demir and E. Bjornson, "Channel Estimation in Massive MIMO Under Hardware Non-Linearities: Bayesian Methods Versus Deep Learning," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 109–124, 2020.

[61] A. Modenini, F. Rusek, and G. Colavolpe, "Optimal transmit filters for ISI channels under channel shortening detection," *IEEE Trans. Commun.*, vol. 61, pp. 4997–5005, Dec. 2013.

[62] M. Failli, "Digital Land Mobile Radio Communications. COST 207," tech. rep., European Commission, Luxembourg, 1989.

[63] F. Gregorio, S. Werner, T. I. Laakso, and J. Cousseau, "Receiver Cancellation Technique for Nonlinear Power Amplifier Distortion in SDMA–OFDM Systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2499–2516, 2007.

[64] J. Tellado, L. Hoo, and J. Cioffi, "Maximum-likelihood detection of nonlinearly distorted multicarrier symbols by iterative decoding," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 218–228, 2003.

[65] J. Madeira, J. Guerreiro, and R. Dinis, "Iterative frequency-domain detection and compensation of nonlinear distortion effects for MIMO systems," *Phys. Commun.*, vol. 37, p. 100869, 2019.

[66] S. Chen, S. X. Ng, E. F. Khalaf, A. Morfeq, and N. D. Alotaibi, "Multiuser Detection for Nonlinear MIMO Uplink," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 207–219, 2020.

[67] Y. Zhang, Z. Wang, Y. Huang, W. Wei, G. F. Pedersen, and M. Shen, "A Digital Signal Recovery Technique Using DNNs for LEO Satellite Communication Systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 7, pp. 6141–6151, 2021.

[68] R. Mei, Z. Wang, and W. Hu, "Robust blind equalization algorithm using convolutional neural network," *IEEE Signal Process. Lett.*, vol. 29, pp. 1569–1573, 2022.

[69] Y. Yang, B. Zhang, D. Guo, W. Wang, J. Nie, Z. Xiong, R. Xu, and X. Zhou, "Stochastic geometry-based age of information performance analysis for privacy preservation-oriented mobile crowdsensing," *IEEE Trans. Veh. Technol.*, pp. 1–14, 2023.

[70] Y. Yang, B. Zhang, D. Guo, R. Xu, C. Su, and W. Wang, "Age of information optimization for privacy-preserving mobile crowdsensing," *IEEE Trans. Emerg. Topics Comput.*, pp. 1–14, 2023.

[71] H. Peng, Z. Tian, X. Li, W. Wang, G. Nauryzbayev, K. Rabie, and T. R. Gadekallu, "Covert communication for cooperative noma with two phases detection," *Alexandria Engineering Journal*, vol. 67, pp. 39–49, 2023.

[72] M. B. Salman and G. M. Guvensen, "A Nonlinear Detector for Uplink SC-FDE mm-Wave Hybrid Massive MIMO under Hardware Impairments," in *Proc. ICC*, pp. 1–6, 2021.

[73] G. M. Guvensen and E. Ayanoglu, "A Generalized Framework on Beamformer Design and CSI Acquisition for Single-Carrier Massive MIMO Systems in Millimeter Wave Channels," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, pp. 1–7, 2016.

[74] H. Lin, F. Gao, S. Jin, and G. Y. Li, "A New View of Multi-User Hybrid Massive MIMO: Non-Orthogonal Angle Division Multiple Access," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2268–2280, 2017.

[75] A. Adhikary, E. Al Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint Spatial Division and Multiplexing for mm-Wave Channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1239–1255, 2014.

[76] M. B. Salman and G. M. Guvensen, "On the Effects of PA Nonlinearities for Hybrid Beamforming Based Wideband Massive MIMO Systems," in *Proc. Int. Conf. Commun. (ICC)*, pp. 1–7, Jun. 2020.

[77] O. Özdoğan, E. Björnson, and E. G. Larsson, "Massive MIMO With Spatially Correlated Rician Fading Channels," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3234–3250, 2019.

[78] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[79] D.-S. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, 2000.

[80] H. Hemesi, A. Abdipour, and A. Mohammadi, "Analytical Modeling of MIMO-OFDM System in the Presence of Nonlinear Power Amplifier with Memory," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 155–163, 2013.

[81] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Tech. Rep. 216, Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.

[82] I. Iofedov and D. Wulich, "MIMO–OFDM With Nonlinear Power Amplifiers," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4894–4904, 2015.

[83] S. Teodoro, A. Silva, R. Dinis, F. M. Barradas, P. M. Cabral, and A. Gameiro, "Theoretical Analysis of Nonlinear Amplification Effects in Massive MIMO Systems," *IEEE Access*, vol. 7, pp. 172277–172289, 2019.

[84] N. D. Lahbib, M. Cherif, M. Hizem, and R. Bouallegue, "Ber analysis and cs-based channel estimation and hpa nonlinearities compensation technique for massive mimo system," *IEEE Access*, vol. 10, pp. 27899–27911, 2022.

[85] J. Guerreiro, R. Dinis, P. Montezuma, and M. M. da Silva, "On the Achievable Performance of Nonlinear MIMO Systems," *IEEE Commun. Lett.*, vol. 23, no. 10, pp. 1725–1729, 2019.

[86] J. Guerreiro, R. Dinis, and P. Montezuma, "Analytical Performance Evaluation of Precoding Techniques for Nonlinear Massive MIMO Systems With Channel Estimation Errors," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1440–1451, 2018.

[87] Z. Mokhtari and R. Dinis, "Sum-rate of cell free massive mimo systems with power amplifier non-linearity," *IEEE Access*, vol. 9, pp. 141927–141937, 2021.

[88] E. Björnson, L. Sanguinetti, and J. Hoydis, "Hardware Distortion Correlation Has Negligible Impact on UL Massive MIMO Spectral Efficiency," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1085–1098, 2019.

[89] S. R. Aghdam, S. Jacobsson, U. Gustavsson, G. Durisi, C. Studer, and T. Eriksson, "Distortion-aware linear precoding for massive MIMO downlink systems with nonlinear power amplifiers," *CoRR*, vol. abs/2012.13337, 2020.

[90] S. R. Aghdam and T. Eriksson, "On the performance of distortion-aware linear receivers in uplink massive mimo systems," in *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 208–212, 2019.

[91] H. Moazzen, A. Mohammadi, and M. Majidi, "Performance Analysis of Linear Precoded MU-MIMO-OFDM Systems With Nonlinear Power Amplifiers and Correlated Channel," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6753–6765, 2019.

[92] S. V. Zavjalov, D. K. Fadeev, and S. V. Volvenko, "Influence of input power backoff of nonlinear power amplifier on ber performance of optimal sefdm signals," in *2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 447–450, 2016.

[93] B. Liu, F. Rottenberg, and S. Pollin, "Power allocation for distributed massive los mimo with nonlinear power amplifiers," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, pp. 1–5, 2022.

[94] J. Jee, G. Kwon, and H. Park, "Precoding Design and Power Control for SINR Maximization of MISO System With Nonlinear Power Amplifiers," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14019–14024, 2020.

[95] J. Jee, G. Kwon, and H. Park, "Joint Precoding and Power Allocation for Multiuser MIMO System With Nonlinear Power Amplifiers," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 8883–8897, 2021.

[96] E. G. L. Thomas L. Marzetta, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[97] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need?," *IIEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.

[98] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1." `http://cvxr.com/cvx`, Mar. 2014.

[99] M. Abdelaziz, L. Anttila, A. Brihuega, F. Tufvesson, and M. Valkama, "Digital predistortion for hybrid MIMO transmitters," *IEEE J. Sel. Topics Signal Process.*, vol. 12, pp. 445–454, Apr. 2018.

[100] J. Yan, H. Wang, and J. Shen, "Novel Post-Weighting Digital Predistortion Structures for Hybrid Beamforming Systems," *IEEE Commun. Lett.*, vol. 25, pp. 3980–3984, Jul. 2021.

[101] H. Yan and D. Cabric, "Digital predistortion for hybrid precoding architecture in millimeter-wave massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, pp. 3479–3483, Mar. 2017.

[102] X. Liu, W. Chen, J. Chu, F. M. Ghannouchi, and Z. Feng, "Multi-Stream Spatial Digital Predistortion for Fully-Connected Hybrid Beamforming Massive MIMO Transmitters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, pp. 2998–3011, Apr. 2021.

[103] L. Zhao, S. Wan, K. Kang, W. Feng, and H. Qian, "A General Digital Predistortion Linearization Scheme for Hybrid Beamforming System," *IEEE Trans. Veh. Technol.*, vol. 72, pp. 654–663, Jan. 2023.

[104] A. Saleh and R. Valenzuela, “A Statistical Model for Indoor Multipath Propagation,” *IEEE J. Sel. Areas Commun.*, vol. 5, pp. 128–137, Feb. 1987.

Appendix A

EQUIVALENT CHANNEL REPRESENTATIONS

In this section, equivalent channels for the linear desired signal and the received distortion channels are analyzed. In order to investigate the linear and nonlinear parts of the received signal, we employ generalized Bussgang decomposition [59] for the nonlinearly distorted transmitted signal as

$$\tilde{x}_n^{(u)} = \alpha_x^{(u)} x_n^{(u)} + \gamma_n^{(u)}, \tag{A.1}$$

where $\alpha_x^{(u)}$ is the Bussgang coefficient, and $\gamma_n^{(u)}$ is the orthogonal distortion term. To reveal the mismatch between the linear and nonlinear channels, we investigate the noise free received signal for u^{th} user $y_n^{(u)}$ in (5.30), after MF, which can be written in the frequency domain as

$$\begin{aligned} Y^{(u)}(e^{j\omega}) = &\sum_{u'=1}^{U} \alpha_x^{(u')} X^{(u')}(e^{j\omega}) H^{(u',u)}(e^{j\omega}) P^*(e^{j\omega}) \\ &+ \sum_{u'=1}^{U} \Gamma^{(u')}(e^{j\omega}) H^{(u',u)}(e^{j\omega}) P^*(e^{j\omega}), \end{aligned} \tag{A.2}$$

where $X^{(u)}(e^{j\omega}) = P(e^{j\omega}) A^{(u)}(e^{j\mu\omega})$ and $A^{(u)}(e^{j\omega}) \triangleq \sum_{m=0}^{N_D-1} a_m^{(u)} e^{-jm\omega}$. In (A.2), $H^{(u',u)}(e^{j\omega}) \triangleq \sum_{n=0}^{L-1} h_n^{(u',u)} e^{-jn\omega}$ and $P(e^{j\omega}) \triangleq \sum_{m=-\infty}^{\infty} p_m e^{-jm\omega}$ are discrete time Fourier transforms (DTFT) of $h_n^{(u',u)}$ and p_n.

Bussgang decomposition decouples the signal into uncorrelated terms; however, it does not provide any information about the spectral characteristics of the nonlinearly modified signal. Hence, inter-modulation product analysis is also used to identify the frequency domain characteristics of the distortion signal, similar to the approach in [48]. The DTFT of the distortion $\Gamma^{(u)}(e^{j\omega}) \triangleq \sum_{n=0}^{\mu N_D-1} \gamma_n^{(u)} e^{-jn\omega}$ is related to the

power spectral density of the distortion term as

$$S_{\Gamma}^{(u)}(e^{j\omega}) = \lim_{N_D \to \infty} \mathbb{E}\left[\frac{1}{\mu N_D}|\Gamma^{(u)}(e^{j\omega})|^2\right], \tag{A.3}$$

which can be approximated as in [54]

$$S_{\Gamma}^{(u)}(e^{j\omega}) \approx \sum_{s=1}^{\infty} p_{\Gamma,s}^{(u)} \underbrace{S_X^{(u)}(e^{-j\omega}) * ... * S_X^{(u)}(e^{-j\omega})}_{\text{s terms}} * \underbrace{S_X^{(u)}(e^{j\omega}) * ... * S_X^{(u)}(e^{j\omega})}_{\text{s+1 terms}}, \tag{A.4}$$

where $p_{\Gamma,s}^{(u)}$ is the power of the s^{th} order nonlinearity.

After downsampling, the received signal for branch i in (5.43) can be obtained in frequency domain from (A.2) as

$$\begin{aligned} Y^{(u),i}(e^{j\omega}) = &\sum_{u'=1}^{U} H^{(u',u),i}(e^{j\omega}) A^{(u')}\left(e^{j\omega}\right) \\ &+ \sum_{u'=1}^{U}\sum_{m=0}^{\mu-1} \bar{H}_m^{(u',u),i}\left(e^{j\omega}\right) \Gamma^{(u')}\left(e^{j\frac{(\omega-2\pi m)}{\mu}}\right). \end{aligned} \tag{A.5}$$

In (A.5), $H^{(u',u),i}(e^{j\omega})$ is the frequency domain effective channel of the i^{th} branch for the desired term, which is given in terms of upsampled channel $H^{(u',u)}(e^{j\omega})$ as

$$H^{(u',u),i}(e^{j\omega}) = \frac{\alpha_x^{(u')}}{\mu} \sum_{m=0}^{\mu-1} H^{(u',u)}\left(e^{j\frac{(\omega-2\pi m)}{\mu}}\right) \left|P\left(e^{j\frac{(\omega-2\pi m)}{\mu}}\right)\right|^2 e^{j\frac{i(\omega-2\pi m)}{\mu_B}}, \tag{A.6}$$

and $\bar{H}_m^{(u',u),i}$ is the effective distortion channel of branch i for the nonlinear term

$$\bar{H}_m^{(u',u),i}(e^{j\omega}) = \frac{1}{\mu} H^{(u',u)}\left(e^{j\frac{(\omega-2\pi m)}{\mu}}\right) P^*\left(e^{j\frac{(\omega-2\pi m)}{\mu}}\right) e^{j\frac{i(\omega-2\pi m)}{\mu_B}}. \tag{A.7}$$

It can be observed from (A.5)-(A.7) that channels corresponding to the linear and nonlinear terms are different. Hence, there is a mismatch between the channel, which is used for nIC in (5.44), and the channel in (A.7), which the distortion experiences. Eventually, mismatched channels may yield significant performance degradation due to *distortion amplification*.

From (A.6), it can be inferred that employing fractional sampling provides diversity since the effective channel is obtained as the summation of the same terms in different phases due to $e^{j\frac{i(\omega-2\pi m)}{\mu_B}}$. Therefore, there may be a sampling instant, which

may prevent nulls in the frequency response, and thus, undermine the effects of the aforementioned mismatch.

The received signal in (A.5) can be re-written in DFT domain, in which IB-DFE operations are performed, by sampling the DTFT at $\omega = 2\pi k/N_D$, as

$$Y_k^{(u),i} = \sum_{u'=1}^{U} H_k^{(u',u),i} A_k^{(u')} + \sum_{u'=1}^{U} \tilde{D}_{k,\text{all}}^{(u',u),i}, \tag{A.8}$$

where $H_k^{(u',u),i} = H^{(u',u),i}\left(e^{j\frac{2\pi k}{N_D}}\right)$, $A_k^{(u)} = A^{(u)}\left(e^{j\frac{2\pi k}{N_D}}\right)$, the total nonlinear distortion $\tilde{D}_{k,\text{all}}^{(u',u),i} = \sum_{m=0}^{\mu-1} \bar{H}_{m,k}^{(u',u),i}\Gamma_{m,k}^{(u')}$, $\bar{H}_{m,k}^{(u',u),i} = \bar{H}_m^{(u',u),i}\left(e^{j\frac{2\pi k}{N_D}}\right)$ is the distortion channel for branch i and $\Gamma_{m,k}^{(u)} = \Gamma^{(u)}\left(e^{j\frac{2\pi(k-N_D m)}{\mu N_D}}\right)$.

As can be noted from Section 5.4.3.1, the symbol-rate nonlinear model in (5.37) is obtained for the flat channel during ST training, and acquired model output is passed through the estimated symbol-rate channel of the i^{th} branch in (A.6). Hence, defining the distortion term $d_n^{(u)}$ for u^{th} user for the flat channel at symbol rate is necessary, and it can be expressed for the k^{th} DFT bin as

$$D_k^{(u)} = \frac{1}{\alpha_x^{(u)}} \sum_{m=0}^{\mu-1} P^*\left(e^{j\frac{2\pi(k-N_D m)}{\mu N_D}}\right)\Gamma^{(u)}\left(e^{j\frac{2\pi(k-N_D m)}{\mu N_D}}\right). \tag{A.9}$$

By using (A.9), we can decompose the distortion signal in (A.8) into two components. The first component is the modeled term, which is given as $H_k^{(u',u),i} D_k^{(u')}$, and the remaining is the not-modelable term $\tilde{D}_k^{(u',u),i} \triangleq \tilde{D}_{k,\text{all}}^{(u',u),i} - H_k^{(u',u),i} D_k^{(u')}$ that we refer to as the *mismatch distortion*. Consequently, the received signal of user u in (A.8) can be written for the i^{th} branch by using the modeled and not-modelable distortion terms as

$$Y_k^{(u),i} = \sum_{u'=1}^{U} \left[H_k^{(u',u),i}\left(A_k^{(u')} + D_k^{(u')}\right) + \tilde{D}_k^{(u',u),i}\right], \tag{A.10}$$

and the mismatched distortion, $\tilde{D}_k^{(u',u),i}$, can be explicitly written from (A.7)-(A.9) as

$$\tilde{D}_k^{(u',u),i} = \sum_{m=0}^{\mu-1} \tilde{H}_{m,k}^{(u',u),i}\Gamma_{m,k}^{(u')}, \tag{A.11}$$

where $\tilde{H}_{m,k}^{(u',u),i} \triangleq \bar{H}_{m,k}^{(u',u),i} - \frac{1}{\alpha_x^{(u')}} P_{m,k}^0 H_k^{(u',u),i}$, is the resultant mismatched channel, and

$$P_{m,k}^i \triangleq P^*\left(e^{j\frac{2\pi(k-N_D m)}{\mu N_D}}\right) e^{j\frac{i(2\pi(k-N_D m))}{\mu_B N_D}}. \tag{A.12}$$

$$\mathbf{R}_d^{(u',u)}[i,r] = \frac{1}{N_D}\sum_{k=0}^{N_D-1}\mathbb{E}\Bigg\{\left(F_k^{(u),i}\tilde{D}_k^{(u',u),i}\right)\left(F_k^{(u),r}\tilde{D}_k^{(u',u),r}\right)^* + F_k^{(u),i}\mathcal{V}_k^{(u),i}\left(F_k^{(u),r}\mathcal{V}_k^{(u),r}\right)^*\delta_{uu'}\Bigg\}$$
$$= \frac{1}{N_D}\sum_{k=0}^{N_D-1}F_k^{(u),i}\left(F_k^{(u),r}\right)^*\sum_{m=0}^{\mu-1}\Bigg[\tilde{H}_{m,k}^{(u',u),i}\left(\tilde{H}_{m,k}^{(u',u),r}\right)^* S_\Gamma^{(u')}\left(e^{j\frac{2\pi(k-N_Dm)}{\mu N_D}}\right) + P_{m,k}^i\left(P_{m,k}^r\right)^* N_o\delta_{uu'}\Bigg] \tag{A.13}$$

By using (A.11)-(A.12), $(i,r)^{th}$ entity of $\mathbf{R}_d^{(u',u)}$ can be obtained as in (A.13).

Appendix B

PROOF OF PROPOSITION 1

To obtain the autocorrelation matrix of the beamformed signal $\mathbf{R}_x$, we firstly define the effective channel matrix in frequency domain, namely $\tilde{\mathbf{H}}[k]$, as

$$\tilde{\mathbf{H}}[k] = \sum_{l=0}^{N-1} \mathbf{F}_{\mathrm{RF}}^{\mathrm{H}} \mathbf{H}_l e^{-j\frac{2\pi}{N}kl}, \tag{B.1}$$

where $\mathbf{H}_l \triangleq [\mathbf{h}_{1,l}, \ldots, \mathbf{h}_{U,l}]$. Since the beamforming matrix $\mathbf{F}_{\mathrm{RF}}$ is designed to illuminate the strongest multi-path component (MPC), that has Rician distribution, of the intended user, the effective channel vector of u^{th} user in frequency domain can be expressed as $\tilde{\mathbf{h}}_u[k] \approx \left[\sum_{l=0}^{N-1} \alpha_{u,l} e^{-j\frac{2\pi}{N}kl}\right] \mathbf{e}_u$ for $D = U$, where $\alpha_{u,l} = \sum_{r=1}^{R} \rho(l - l_{u,1} - l_{u,r}) g_{u,1r}$, $\mathbf{e}_u$ is all zero vector, except its u^{th} element, equal to 1. Assuming that MPCs belonging to the same cluster are not resolvable and Nyquist-1 pulse is used, which implies $\rho(l - l_{u,1} - l_{u,r}) = \delta_{l-l_{u,1}}$, $\tilde{\mathbf{h}}_u[k]$ can be further approximated as

$$\tilde{\mathbf{h}}_u[k] \approx e^{-j\frac{2\pi}{N}kl_{u,1}} \left(\sum_{r=1}^{R} g_{u,1r}\right) \mathbf{e}_u. \tag{B.2}$$

Then, $\tilde{\mathbf{h}}_{u_1}^{\mathrm{H}}[k]\tilde{\mathbf{h}}_{u_2}[k] \approx \delta_{u_1-u_2} \left|\sum_{r=1}^{R} g_{u_1,1r}\right|^2$ and it can be further approximated as $\tilde{\mathbf{H}}[k] \approx \mathrm{diag}\left\{\sum_{r=1}^{R} g_{u,1r} e^{-j\frac{2\pi}{N}kl_{u,1}}\right\}_{u=1}^{U}$. By using these approximations, one can obtain $\mathbf{R}_x$ as

$$\begin{aligned} \mathbf{R}_x &= \frac{1}{N}\mathbf{F}_{\mathrm{RF}}\mathbb{E}_{\mathbf{h}_{u,l}} \left[\sum_{k=-N_A/2}^{N_A/2-1} \mathbf{F}_{\mathrm{D}}[k]\mathbf{F}_{\mathrm{D}}[k]^{\mathrm{H}}\right] \mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}, \\ &\approx \frac{1}{N}\mathbf{F}_{\mathrm{RF}}\mathbb{E}_{\mathbf{h}_{u,l}} \left[\sum_{k=-N_A/2}^{N_A/2-1} \mathrm{diag}\left\{\left|\sum_{r=1}^{R} g_{u,1r}\right|^{-2}\right\}\right] \mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}, \end{aligned} \tag{B.3}$$

where $\mathbf{F}_D[k] = \tilde{\mathbf{H}}[k]\left(\tilde{\mathbf{H}}^{\mathrm{H}}[k]\tilde{\mathbf{H}}[k]\right)^{-1}$. Therefore, we can approximate $\mathbf{R}_x$ as

$$\mathbf{R}_x \approx \mathbf{F}_{\mathrm{RF}} \left(\mathrm{diag} \left\{ \beta_u \right\}_{u=1}^{U} \right)^{-1} \mathbf{F}_{\mathrm{RF}}^{\mathrm{H}}, \tag{B.4}$$

where $\beta_u \triangleq \frac{N}{N_A} \mathbb{E} \left[\left| \sum_{r=1}^{R} g_{u,1r} \right|^2 \right]$.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Salman, Murat Babek
Nationality: Turkish (TC)

EDUCATION

Degree	Institution	Year of Graduation
M.S.	Middle East Technical University	2018
B.S.	Middle East Technical University	2015

PROFESSIONAL EXPERIENCE

Year	Place	Enrollment
2015-2017	ASELSAN Inc. 1	Digital design engineer
2017-2023	Middle East Technical University	Reseach Assistant

PUBLICATIONS

International Journal Publications

1. M. B. Salman and G. M. Guvensen, "An Efficient QAM Detector via Nonlinear Post-Distortion Based on FDE Bank Under PA Impairments," in IEEE Transactions on Communications, vol. 69, no. 10, pp. 7108-7120, Oct. 2021, doi: 10.1109/TCOMM.2021.3095974.

2. A. Kurt, M. B. Salman, U. B. Sarac and G. M. Guvensen, "An Adaptive-Iterative Nonlinear Interference Cancellation in Time-Varying Full-Duplex Channels," in IEEE Transactions on Vehicular Technology, vol. 72, no. 2, pp. 1862-1878, Feb. 2023, doi: 10.1109/TVT.2022.3208766.

3. M. B. Salman, G. M. Guvensen, and T. Ciloglu, “Low complexity nonlinear detection for multiuser hybrid mimo systems and performance analysis,” Physical Communication, p. 102126, 2023.

International Conference Publications

1. M. B. Salman and G. M. Guvensen, "On the Effects of PA Nonlinearities for Hybrid Beamforming Based Wideband Massive MIMO Systems," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-7, doi: 10.1109/ICC40277.2020.9149193.

2. M. B. Salman and G. M. Guvensen, "A Nonlinear Detector for Uplink SC-FDE mm-Wave Hybrid Massive MIMO under Hardware Impairments," ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICC42927.2021.9500524.

3. M. B. Salman, E. Björnson, Guvensen, G. M., and T. Ciloglu, “Analytical Nonlinear Distortion Characterization for Frequency-Selective Massive MIMO Channels,” in Proc. ICC, pp. 1–6, 2023.