UNIVERSITY OF NAIROBI

FACULTY OF ENGINEERING

Reduction of Peak-to-Average Power Ratio in Orthogonal Frequency Division Multiplexing Radio Systems

By

Stephen Ngure Kiambi

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Electrical and Electronic Engineering
in the Department of Electrical and Information Engineering
in the University of Nairobi.

March 2023

DECLARATION OF AUTHORSHIP

This thesis is my original work and has not been presented for a degree in any other university.

Name: __Stephen Ngure Kiambi__   Registration Number: __F80/98749/2015__

Signature: _____   Date: __12ᵗʰ March 2023__
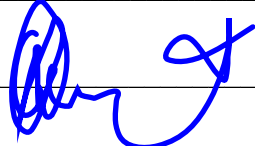
Supervisors' approval:

This thesis has been submitted for examination with our approval as university supervisors.

Name: __Prof. Elijah Mwangi__

Signature: _____   Date: __14ᵗʰ March 2023__

Name: __Prof. George Kamucha__

Signature: _____   Date: __14/03/2023__

ACKNOWLEDGEMENT

ABSTRACT

Orthogonal frequency division multiplexing (OFDM) has continued to be the most preferred signal-multiplexing scheme for the current and future standards for high-speed transmission of data over the wireless channel. This is due to its ability to provide a high spectral efficiency and to combat both intersymbol interference and multipath fading. However, an OFDM signal tends to have a high peak-to-average power ratio (PAPR) when the number of subcarriers is large. Subsequently, the high power amplifier in the transmitter nonlinearly amplifies the signal resulting in-band and out-of-band radiations in the output of the device. In turn, these radiations give rise to bit-error rate (BER) degradation and adjacent channel interferences. Therefore, it is necessary to reduce the PAPR to avoid these two effects. This thesis proposes five new PAPR reduction methods for OFDM systems with a contribution to the class of additive signal methods. Three of the proposed methods follow the conventional tone-reservation (TR) approach; these are second-order cone program TR (SOCP-TR), linear program TR (LP-TR) and iterative re-weighted least squares TR (IRLS-TR). The remaining two methods, namely low-complexity signal addition (LCSA) and low-complexity additive signal mixing (LCASM), extend the transmit time to cater for the transmission of nonzero samples of the peak-reduction signal. The SOCP-TR and LP-TR are optimal schemes, while the remaining three are suboptimal and are preferable for practical realization due to their fast convergence rates and low computational complexities. The suboptimal methods can achieve similar or better PAPR reductions than the optimal schemes. From the perspective of the reservation of PAPR reduction resources, in either frequency or time domain, the IRLS-TR and LCSA methods offer the best performances. For example, with 5% of the transmit resources reserved in a system with 256 subcarriers, the IRLS-TR and LCSA methods achieve reductions of 5.84 and 5.69 dB, at the expense of only a small change in transmit power of 0.36 and -0.10 dB, respectively. Additionally, the five proposed methods reduce PAPR without degrading the BER. The proposed methods also perform better than other methods proposed in literature. In addition to offering new PAPR reduction methods, this thesis has established that the type of subcarrier modulation and signal structure, real or complex, insignificantly affect the amount of PAPR reduction.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ACE	Active Constellation Extension
ACI	Adjacent Channel Interferences
ACPR	Adjacent Channel Power Ratio
ACT	Adaptive Clipping Technique
ADC	Analogue-to-Digital Converter
AM	Amplitude Modulation
AMC	Adaptive Modulation and Coding
AWGN	Additive White Gaussian Noise
BER	Bit-Error Rate
BPSK	Binary Phase-Shift Keying
CCDF	Complementary Cumulative Distribution Function
CDF	Cumulative Distribution Function
CF	Curve Fitting
CF-TR	CF-Tone Reservation
CP	Cyclic Prefix
DAB	Digital Audio Broadcasting
DAC	Digital-to-Analogue Converter
dB	Decibel
dBc	dB relative to carrier
DFT	Discrete Fourier Transform
DMT	Discrete Multitone
DMB	Digital Multimedia Broadcasting
DPD	Digital Predistortion
DVB-T2	Second generation of Terrestrial Digital Video Broadcasting
ECC	Error-Correction Coding
EER	Envelope Elimination and Restoration

ET	Envelope Tracking
EVM	Envelope Vector Magnitude
FDM	Frequency Division Multiplexing
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FIR	Finite-Impulse Response
HPA	High Power Amplifier
IBO	Input Back-Off
ICI	Inter-Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
IIR	Infinite-Impulse Response
IM	Intermodulation
IMD	IM Distortion
IRLS	Iterative Re-Weighted Least Squares
IRLS-TR	IRLS-Tone Reservation
ISE	Instantaneous Squared-Error
ISI	Inter-Symbol Interference
IVO	Initial Value Optimization
IVO-TR	IVO-Tone Reservation
kW	Kilowatt
KWh	Kilowatt-hour
LCASM	Low-Complexity Additive Signal Mixing
LCSA	Low-Complexity Signal Addition
LMS	Least Mean Square
LOS	Line-of-Sight
LP	Linear Program
LP-TR	LP-Tone Reservation
LS	Least Squares
LSA	Least-Squares Approximation

LSA-TR	LSA-Tone Reservation
M-QAM	M-ary Quadrature Amplitude Modulation
MATLAB	MATrix LABoratory
MIMO	Multiple-Input Multiple-Output
$MtCO_2$	Metric tons of Carbon Dioxide
MISO	Multiple-Input Single-Output
MSE	Mean Square Error
MW	Mega Watt
NLOS	Non-Line-of-Sight
OFDM	Orthogonal Frequency Division Multiplexing
OOB	Out-of-Band
P/S	Parallel-to-Serial Conversion
PA	Power Amplifier
PAPR	Peak-to-Average Power Ratio
pdf	Probability Density Function
PM	Phase Modulation
PRCs	Peak-Reduction Coefficients
PRS	Peak-Reduction Signal
PRT	Peak-Reduction Tone
PSD	Power Spectral Density
PTS	Partial Transmit Sequences
QAM	Quadrature Amplitude Modulation
QCQP	Quadratically Constrained Quadratic Program
QP	Quadratic Program
QPSK	Quadrature Phase-Shift Keying
RLS	Recursive Least-Squares
RF	Radio Frequency
S/P	Serial-to-Parallel Conversion
SFBC	Space-Frequency Block Code
S-TR	Selective Tone Reservation
SCS-SLM	Selective Codeword Shift and Selective Mapping

SER	Symbol-Error Rate
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SLM	Selective Level Mapping
SNR	Signal-to-Noise Ratio
SOCP	Second-Order Cone Program
SOCP-TR	SOCP- Tone Reservation
SSPA	Solid-State Power Amplifier
STBC	Space-Time Block Code
TR	Tone Reservation
TWT	Travelling-Wave Tube
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

LIST OF FIGURES

xiv

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This chapter introduces the research work in this thesis. The chapter begins with an overview of orthogonal frequency division multiplexing (OFDM) technique highlighting its competitive advantages and disadvantages. The specific drawback of OFDM technology that is addressed in this thesis is outlined in the section on problem statement. The chapter also provides the main and specific objectives of the research and the scope of work involved in this thesis. Towards the end of the chapter, a list of publications that have been realized during the course of this work is provided followed by a presentation of how the rest of the thesis is organized.

1.1 Background

The never-ending demand for high-speed transmission of data especially over the wireless medium has required a continuous evolution of wireless communication networks from one generation to the next. This evolution has been geared towards adopting new or enhanced technologies that aid in meeting latest and future demands for high data-rate over the hostile wireless medium. One of the key enabling technologies for high data-rate transmission over the wireless medium is the OFDM technique [1].

The OFDM technique can be viewed as a digital multicarrier transmission technique that achieves high data-rate through having many data-bearing symbols modulate an equivalent number of subcarriers simultaneously. In this regard, each subcarrier constitutes a transmission subchannel. The overall data-rate can be enhanced further by applying different modulation schemes on different subcarriers depending on the quality of each channel. In this case, those subchannels with high signal-to-noise ratio (SNR) are used to carry more bits per symbol through employment of high-order modulation schemes. On the contrary, subchannels with low SNR carry a smaller number of bits via symbols drawn from low-order modulation schemes.

In addition, several other important advantages of OFDM technique have made it very attractive for use in wireless communication. These include the

ability to suppress inter-symbol interference (ISI) and being resilient to multipath fading, the capability to attain high spectral efficiency, its efficient implementation using fast Fourier transform algorithm and the simplicity in the receiver design.

The ISI is eliminated by having long symbol duration, typically much greater than the maximum expected multipath delay, together with guard interval in-between consecutive symbols. The long symbol duration also helps to combat multipath fading. In particular, the subchannel bandwidth, determined by the number of subcarriers and the total available bandwidth, is less than the coherence bandwidth; hence each subchannel has a flat-fading frequency response. This in turn simplifies the receiver design since only a single-tap equalizer is needed to recover a transmitted symbol.

Additionally, a high spectral efficiency in each OFDM symbol is achieved by having mutually orthogonal subcarriers transmit data over the allocated bandwidth. The orthogonality of subcarriers allows many subcarriers to be tightly packed in the available bandwidth, with adjacent ones overlapping one another, hence increasing the data-rate.

In addition to the above-mentioned advantages, OFDM can be combined with multiple-input multiple-output (MIMO) technology to add spatial diversity and/or spatial multiplexing. Through spatial diversity, the MIMO technique can improve the reliability of a wireless medium, and thus reduce the bit-error rate (BER) of a communication system. On the other hand, spatial multiplexing increases the data rate of a system. In addition, MIMO technique can be used to shape the transmit beam to increase coverage or to reduce the required transmit power.

Despite having the above highlighted advantages, OFDM technique has a few drawbacks that remain unresolved. Two major ones are the sensitivity to both nonlinear amplification and frequency offsets. The first one is the focus of this thesis and is exacerbated by the occurrence of high peak-to-average power ratio (PAPR) in OFDM signals. This implies that the peak power of the transmit signal is much greater than the average power as illustrated in Figure 1.1. Essentially, a high PAPR indicates that a signal has a wide dynamic range of signal amplitudes with a few of them being very high. The occurrence of high PAPR originates from the nature of the composition of the transmit signal as a sum of a large number of modulated signals.

The high PAPR causes two problems when the signal is processed through an OFDM transreceiver. The first one is the requirement to equip the digital-to-analogue converter (DAC) in the transmitter section with a large number of bits to cover the broad dynamic range of signal amplitudes. This leads to a significant increase in the cost of the

DAC. The same requirement and the associated cost apply similarly to the analogue-to-digital converter (ADC) in the receiver section.

The second problem resulting from high PAPR is the nonlinear amplification of the signal in the high power amplifier (HPA) in the transmitter section. The nonlinear amplification results in two types of distortions, namely out-of-band and in-band radiations. The out-of-band radiations causes the spectrum of the transmit signal to spread out of the allocated bandwidth thus leading to interferences in the adjacent channels. Mitigating this problem requires the use of highly selective filters, which are costly to design, or a wideband guard between OFDM channel and adjacent channels and this will lead to poor utilisation of the frequency spectrum. On the other hand, the in-band radiations can degrade the BER of the system to an extent that the transmitted information cannot be recovered.

Traditionally, a workaround solution to avoid nonlinear amplification of transmit signal and the two undesirable effects has been to back-off the HPA from the 1-dB compression point to a linear point of operation depending on the PAPR of the input signal. However, this backing-off operation reduces the efficiency of the HPA, thereby causing it to consume and dissipate a lot of power. The high power consumption increases the cost of the HPA in addition to the device requiring a costly design.
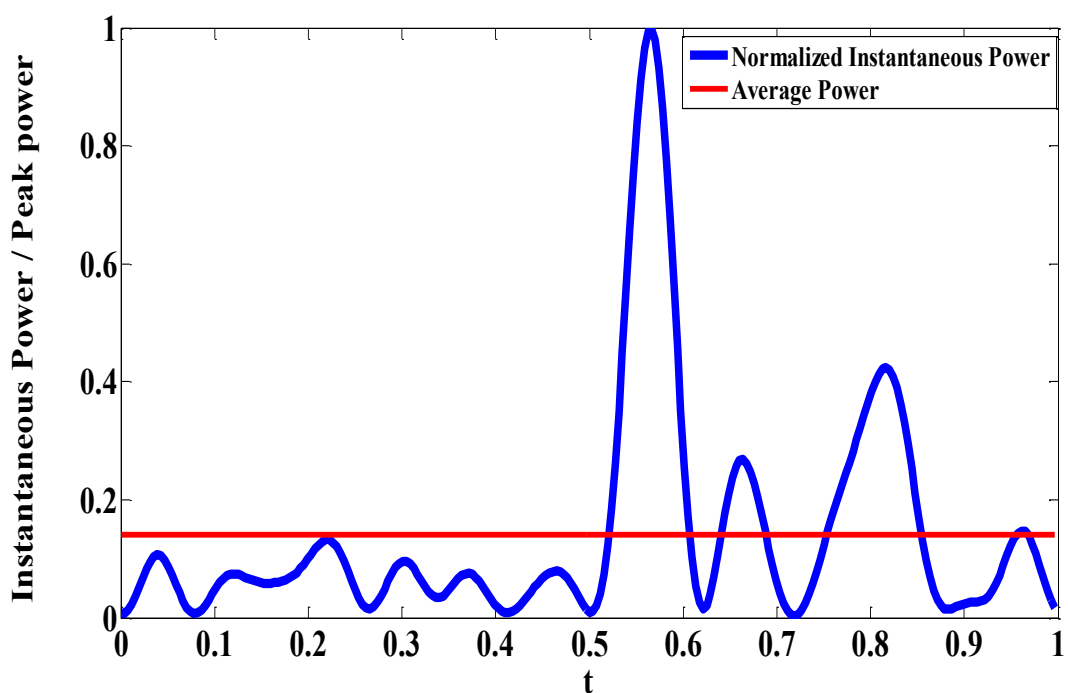
Figure 1.1. OFDM signal power fluctuations

Additionally, the lifetimes of battery powers in user terminals, such as handheld devices, are reduced significantly [2]. Therefore, the solution of backing-off the HPA is not satisfactory. This has motivated many researchers, up-to-date, to attempt to seek better solutions for avoiding nonlinear amplification of high-PAPR signals without interfering with the rated efficiency of the HPA.

Recently, several methods have been proposed in literature to address the problem of high PAPR and the attendant nonlinear amplification in OFDM transmitters. The proposed methods are targeted at either reducing PAPR, extending linear operation or enhancing efficiency of HPA. The reduction of PAPR has been found to be much more convenient than the other two approaches because it is of lower computational complexity and does not require additional analogue circuitry. This has led to a lot more of research being directed towards PAPR reduction than in the other two approaches. In some cases, some methods combining PAPR reduction and linearization or efficiency enhancement have also been proposed. However, combining different techniques may increase the complexity and the cost of an OFDM system.

The techniques proposed for reducing PAPR of input signals to the HPA are typically aimed at minimizing the ratio to a suitable level to enable the device to operate at the upper end of the linear region near the rated maximum efficiency without the need for a large input back-off. Up to now, the proposed PAPR reduction methods have been found to suffer from some shortcomings that hinder their practical realizations. The methods fail to have majority of the attributes desired for implementations in real-time systems. These include

i)   good PAPR reduction,

ii) low computational complexity,

iii) non-degradation of BER,

iv) minimal loss in data-rate and the

v)   preservation of the transmit power

Therefore, the aim of this thesis was to present practically realizable methods for reducing PAPR in wireless communication systems that use OFDM technique.

1.2 Problem Statement

OFDM technique has of late been receiving a lot of attention as the best solution for sustaining the demand for high data rate, spectral efficiency and quality of service in wireless communication networks. However, an OFDM signal can have a high level of PAPR that can lead to degradation of quality of service, increase in the cost of transmitter devices and reduced battery lifetime in user terminals.

Recently, several methods have been proposed in literature for reducing PAPR in wireless communication systems that use OFDM technique. However, these methods do not have one or more of the attributes desired for their implementation in real-time systems. The proposed methods are known to introduce one or more undesirable effects in the communication system; mainly a loss in data rate, an increase in transmit power, degradation in BER, an increase in computational complexity and additional analogue circuitry. To this end, there is still a need to develop new PAPR reduction methods that can be practically implemented in the current and future generations of OFDM-based wireless networks.

1.3 Research Objectives

The main objective of this research work was to develop and implement suitable algorithms for PAPR reduction in wireless communication systems that utilize OFDM technique with the aim of eliminating signal distortions that arise in the HPA in the transmitter due to nonlinear amplification of signals having high PAPR. The specific objectives were as follows:

   i)   To develop optimal PAPR reduction algorithms, based on tone reservation, for benchmarking the performance of suboptimal algorithms.

   ii)   To investigate whether the type of subcarrier modulation employed in an OFDM system affects PAPR reduction.

  iii)   To investigate if the type of signal structure, real or complex, can affect PAPR reduction.

   iv)   To develop low-complexity and reliable suboptimal algorithms for PAPR reduction based on tone reservation and other signal addition methods.

   v)   To compare the performance of proposed algorithms and other related algorithms.

1.4 Scope of Research

The focus of this thesis was on the reduction of PAPR in wireless communication networks that typically employ OFDM in the transmission of signals. All simulations were carried out in MATLAB to investigate the PAPR reduction capability of the developed algorithms and the associated effects on the underlying network using OFDM system configurations with different subcarrier-modulation schemes and numbers of subchannels.

1.5 Publications

This section gives a list of publications in refereed journals and conference proceedings that were achieved during the course of this research. Six international peer-reviewed publications were accomplished from the findings of this research. Two of these were published in conference proceedings of the Institute of Electrical and Electronics Engineers (IEEE), while the remaining four were in peer-reviewed journals.

The following are the papers published in peer-reviewed IEEE conference proceedings:

i) The first one was titled "Effect of OFDM Signal Structure and Subcarrier Modulation on the Reduction of the Signal Peak Power". It was published in the proceedings of *IEEE AFRICON 2017* conference that took place from 18 to 20 September 2017 in Cape Town, South Africa. It appears on page 262-266 of the proceedings [3].

ii) The second paper titled "Use of Preset Reserved Tones in Reduction of PAPR in OFDM Systems" was published in *IEEE AFRICON 2019 Proceedings* of the conference held from 25 to 27 September 2019 in Accra, Ghana [4].

The list of the four articles published in refereed journals is as follows:

i) The first article was titled "An Iterative Re-Weighted Least-Squares Tone Reservation Method for PAPR Reduction in OFDM Systems". It was published in the journal of *Transactions on Communications* of *World Scientific and Engineering Academy and Society* (WSEAS) on page 153-161 of volume 18 of 2019 [5].

ii) The second journal paper titled "A Low-Complexity Signal Addition Method for

PAPR Reduction in OFDM Systems" was published in the *International Journal of Computer Application* (IJCA). It is on page 21-34 of issue number 10 of volume 5 of 2020 [6].

iii)   The third article titled "PAPR Reduction in MIMO-OFDM Systems Using Low-Complexity Additive Signal Mixing" was published in the *Journal of Communications* (JCM) on page 468-478 of the November issue of volume 16 of the 2021 publications [7].

iv)   The fourth article was titled "Reducing PAPR of OFDM Signals Using a Tone Reservation Method Based on ℓ_∞-Norm Minimization". It was published in the *Journal of Electrical Systems and Information Technology* (JESIT) and it appears as article number 12 in volume 9 of 2022 [8].

1.6 Organization of Thesis

The rest of the thesis is organized as follows. Chapter 2 offers a review of published literature on PAPR reduction in OFDM system that is related to this thesis together with alternative approaches, which are HPA linearization and efficiency enhancement techniques. Additionally, the chapter provides a highlight of the existing knowledge gaps. This is followed in Chapter 3 by coverage of OFDM technique including the cause of high PAPR, nonlinear amplification distortions and the effects they have in a communication system. In Chapter 4, important concepts of convex optimization and adaptive signal processing are presented. These concepts are essential to the development of the PAPR reduction methods proposed in Chapter 5. Chapter 5 details the materials used in the research and the algorithms of the proposed PAPR reduction methods. In Chapter 6, simulation results from the proposed methods are presented, discussed and compared to results of other promising methods. Chapter 7 provides conclusion and recommendations for future work. This is followed by a list of references and appendices. The appendices start by providing the notations used in this thesis in Appendix A. This is then followed by published works, which appear in Appendix B to G, and lastly by MATLAB codes for various simulations including those of the proposed methods in Appendix H.

CHAPTER 2

LITERATURE REVIEW

This chapter first presents a review of related work that has been proposed for the reduction of PAPR in OFDM and MIMO-OFDM systems. This is then followed in the second part by a brief coverage of linearization and efficiency enhancement techniques as applied to high power amplifiers. A linearization or an efficiency enhancement operation is considered as an alternative to PAPR reduction or as a supplement when it is combined with the latter with the intention to improve the performance. In the last part of the chapter, the knowledge gap is identified. This research is envisioned to fill this gap.

2.1 Review of PAPR Reduction Methods

In a communication system employing OFDM technique to transmit signals, the possibility of a transmit signal exhibiting a high PAPR remains a plaguing concern particularly so when the number of subcarriers involved is large. A signal with high PAPR is nonlinearly amplified in the high power amplifier (HPA) and as a result the BER of the system is degraded due to in-band radiations and interferences occur in adjacent channels owing to out-of-band radiations.

In order to deal with the above-mentioned problems from nonlinear amplification of signals, different methods for reducing PAPR have recently been proposed in literature. These include signal clipping [9], [10], companding [11], selective mapping [12] - [14], partial transmit sequence [15], [16], coding [17], tone reservation [18], [19] and hybrid schemes [20] - [22].

In general, based on how the above PAPR reduction methods alter the transmit signal, they can be classified into three main categories [23], namely:

i)   Signal distortion techniques,

ii) Signal scrambling techniques and

iii) Hybrid techniques.

In the following subsections, some key methods in the three categories are briefly reviewed together with a highlight of their advantages and disadvantages.

2.1.1 Signal Distortion Techniques

In the signal distortion group, the reduction of PAPR is achieved through an irreversible modification of the OFDM signal to reduce the highest peaks prior to amplification in the HPA. The techniques under this category, generally, introduce both in-band and out-of-band interferences and also increase the complexity of the underlying OFDM system. Some of the methods in this class are discussed here below.

A. Clipping and Filtering

The clipping of high amplitudes in a signal is the simplest way to reduce the PAPR [24] - [26]. In its basic form, the clipping method simply trims the amplitudes of a signal $x(t)$ to a desired threshold before passing it to the HPA. Taking $\varphi(t)$ to be the instantaneous phase angles corresponding to the signal amplitudes, the clipped signal is given as

$$x'(t) = \begin{cases} x(t) \quad, & |x(t)| < A \\ Ae^{j\varphi(t),} & |x(t)| \geq A \end{cases} \qquad (2.1)$$

where A is the clipping threshold.

Since the clipping operation is nonlinear, the PAPR of the signal is reduced at the expense of in-band distortion and out-of-band radiations; the latter is also referred to as adjacent channel emissions [27]. The in-band distortion cannot be eliminated by filtering and thus end up degrading the BER. Obviously, the PAPR reduction and BER degradation depend on the level of clipping. A higher clipping threshold results in both a lower PAPR reduction and a lower BER degradation and vice versa.

The out-of-band radiations may lead to a reduction in spectral efficiency since a larger frequency guard would be required between adjacent channels to avoid interferences. The filtering of a signal after a clipping operation can reduce out-of-band radiation but it may also cause peak regrowth at some time instants. At those time instances, the signal amplitudes exceed the clipping level after the filtering operation. To eliminate peak regrowth, a repeated clipping-and-filtering operation can be performed [28] but generally requires too much iteration to reach a desired amplitude level.

An adaptive clipping technique (ACT) was proposed in [29] with an aim to

eliminate the drawbacks of the traditional clipping technique in Alamouti encoded MIMO-OFDM systems. However, the amount of PAPR reduction by the technique had to be regulated to avoid an intolerably high BER degradation.

B. Peak Windowing

As a solution to lower the out-of-band radiations associated with hard clipping, the peak widowing technique multiplies the largest signal peak with a window function [30]. The typical window functions employed in the technique are Gaussian, cosine, Hanning, Kaiser and Hamming windows. The resulting spectrum is a convolution of the spectrum of the original OFDM signal and that of the applied window function.

Ideally, the spectrum of the window function should be as narrow as possible but not too long in the time-domain so as not to affect many OFDM signal samples and thereby increase the BER. However, a narrow window function in the time-domain implies wide out-of-band radiations and hence a reduction in spectral efficiency.

C. Envelope Scaling

An envelope-scaling algorithm was proposed in [31] to reduce PAPR by scaling the envelopes of some subcarriers in the input signal before it is sent for IFFT processing in the transmitter. The aim of the scheme is to achieve almost the same envelope in the subcarriers akin to PSK modulation. The combination of envelopes that gives the lowest PAPR is forwarded to the IFFT section for onward processing and transmission.

The phases of the resulting transmit signal are the same as for the original signal but the two signals have different envelopes. Therefore, the received data symbols can be decoded at the receiver without requiring any side information in the case of PSK modulation. However, for QAM-modulated subcarriers, the envelope-scaling operation can result in serious BER degradation. This can only be mitigated by transmitting an unwarranted large amount of side information, especially when the number of subcarriers is large, to the receiver for the decoding of the transmitted symbols.

D. Random Phase Updating

In [32], a random phase updating algorithm was proposed where random phases are generated and assigned to each subcarrier. The phase update is done iteratively together with PAPR calculation until a minimum threshold level or a maximum iteration number is reached. The phase shifts have to be known to both the transmitter and the receiver in order to avoid BER degradation.

A large amount of side information has to be transmitted for the receiver to have information about all the effected phase changes in the transmitter and this reduces the data-rate of the system. In addition, the efficiency of the algorithm was found to depend on the selected threshold level and the maximum iteration number. Although quantization and grouping of phases together with an adaptive threshold could improve the performance of the algorithm, all this would occur at the expense of an increased computational complexity.

E. Peak-Reduction Carrier

In this technique [33], a number of peak-reduction carriers are employed to reduce PAPR of OFDM signals. The technique finds an optimal amplitude and phase of each peak-reduction carrier for a given OFDM symbol. To get optimal peak-reduction carriers, an exhaustive search for all possible combinations of amplitudes and phases is needed. The exhaustive search is computationally intensive especially when a large number of subchannels and peak-reduction carriers are involved. Many peak-reduction carriers, greater than 25% of the total number of subcarriers, are generally required to attain a significant PAPR reduction. Consequently, the large number of peak-reduction carriers significantly reduces the data-rate of the system.

In the implementation of the technique, a higher order modulation scheme is used to represent a lower order modulation symbol. This forces the amplitude and phase of peak-reduction carriers to be confined within the constellation region of the data symbol. However, the use of higher order modulations increases the probability of error when detecting the transmitted symbols and hence a degradation in BER. Therefore, a trade-off between PAPR reduction level and BER performance must be considered when selecting a digital modulation scheme and the number of the peak-reduction carriers.

F. Companding

The companding technique is highly used in processing of speech signals. Such signals are composed of low signal amplitudes and infrequent occurrences of high signal peaks. A similar characteristic is observed in OFDM signals. In the companding technique, larger quantization step-sizes are used to code high amplitudes in comparison to lower ones; a process achieved through employment of a companding law such as μ-law or A-law. This way, the technique improves the quantization resolution of low signal amplitudes at the expense of high amplitudes. Because of the infrequent occurrences of

high amplitudes, the speech quality is barely affected by the companding operation. In addition, the companding operation reduces the number of quantization levels and therefore the required number of bits per sample and the transmission bandwidth.

Similar companding transforms to the ones used in speech applications have been proposed [34], [35] for reducing PAPR in OFDM signals. After the IFFT operation at the transmitter, the signal is first companded using μ-law and then quantized before digital-to-analogue conversion. With a μ-law compression, a discrete-time companded signal can be expressed as

$$x_{comp}(n) = \frac{x_p \, \text{sgn}(x(n)) \ln\left(1 + \mu \left|\frac{x(n)}{x_p}\right|\right)}{\ln(1 + \mu)} \tag{2.2}$$

where x_p is the maximum peak of the signal $x(n)$ before companding, μ is a parameter for setting the nonlinearity of the companding function, $\text{sgn}(.)$ and $\ln(.)$ are the standard signum and natural logarithmic functions, respectively.

At the receiver, after analogue-to-digital conversion, the received signal is expanded via an operation that attempts to invert the companding transform in the transmitter. The expanded discrete-time signal can be expressed as

$$y_{expand}(n) = \frac{x_p \exp\left(\frac{y(n)}{x_p \, \text{sgn}(y(n))} \ln(1 + \mu)\right) - 1}{\mu \, \text{sgn}(x(n))} \tag{2.3}$$

where $y(n)$ is the received discrete-time signal.

Due to the companding operation, quantization errors in high signal amplitudes can be significantly large, thus leading to BER degradation. In addition, the μ-law companding does not change the peak power of the original OFDM signal but instead, through the control of μ, it increases the average power in order to reduce PAPR. Subsequently, the increase in the average power leads to a further degradation in BER. Therefore, the reduction in PAPR through companding is achieved at the expense of BER degradation.

2.1.2 Signal Scrambling Techniques

Methods in this class can either generate many alternatives of the original OFDM signal followed by transmission of the one with the minimum PAPR or modify the signal by introducing phase shifts, adding peak-reduction tones or altering constellation points.

An overview of the main approaches in this category is given here below.

A. Coding

The coding scheme is motivated by the fact that the PAPR of a given OFDM symbol is dependent on the modulation symbols. Some combinations of modulation symbols in an OFDM symbol may result in higher PAPRs than others. The worst case PAPR can occur if the same data symbol modulates all subcarriers. An example is the case when BPSK modulation symbol $X(k) = +1$, for $k = 0, 1, ..., N - 1$, is transmitted on all the N subchannels, thus after IFFT resulting in a discrete-time signal with a PAPR equal to N.

In [36], a block coding approach was employed to avoid the transmission of identical modulation symbols. Another coding approach to avoid high PAPR is to search for a combination of optimal codes that yield low PAPR. For example, in [37] Golay complementary sequences were applied to reduce PAPR. In addition, the idea of integrating PAPR reduction and error-correction coding has been presented in several papers [38]. However, it is important to note that any form of coding reduces the useful data rate and increases computational complexity of the system.

B. Tone Reservation

This can be regarded as one of the additive signal methods for reducing PAPR in OFDM systems. The general idea behind a signal addition method is the creation of a dummy signal, also referred to as a peak-cancelling or a peak-reduction signal, which upon its addition to OFDM signal results into a signal of a much lower PAPR. In general, signal addition methods can achieve high PAPR reductions but their implementation may require the use of complex optimization algorithms to obtain the dummy signal.

Tone-reservation methods [39] - [42] need to set aside a subset of subcarriers, referred to as peak-reduction tones, for creating the dummy signal through a search for a combination of peak-reduction weights in the frequency domain. The dummy signal is found from the inverse discrete Fourier transform (IDFT) of the peak-reduction weights. In other words, a tone reservation method attempts to solve a minimax problem

$$\text{minimize } \max(|\mathbf{x} + \mathbf{c}|) \tag{2.4}$$

where $\mathbf{x}$ is the OFDM signal vector whose PAPR is to be reduced through addition of a dummy signal, denoted here by vector $\mathbf{c}$, and given by

$$\mathbf{c} = \text{IFFT}(\mathbf{C}) \qquad (2.5)$$

where vector $\mathbf{C}$ contains the peak-reduction coefficients in the reserved locations and zeros in the rest of locations in an OFDM symbol.

At the receiver, the reserved tones are ignored. Thus, a tone-reservation method does not require any side information to be transmitted and does not affect the BER of the underlying system. However, the average power of the transmitted signal increases and the overall bit-rate is reduced since the reserved subcarriers do not carry user data. Attempts to address these two drawbacks in addition to reducing the computational complexity in the finding of the peak-reduction coefficients, while at the same time attaining good PAPR reduction, have continued to attract a lot of research.

Depending on the methodology of finding the peak-reduction weights, a tone-reservation method can be said to be optimal or suboptimal. A tone-reservation method that seeks optimal solution to the peak-reduction weights generally employs an optimization technique such as linear programming [43], second-order cone programming [44] or quadratically constrained quadratic programming [45] to solve the problem in equation (2.4). Thus, an optimal tone-reservation method is characterised by both a high computational complexity and a slow convergence rate.

Due to the shortcomings of the optimal tone-reservation approaches, several suboptimal algorithms have been proposed in literature. Among the first suboptimal methods, hereafter referred to as K-TR due to the use of a time-domain kernel signal, is the scheme that was proposed in 1997 by Tellado and Cioffi [46]. It employs a gradient algorithm in which the peak-reduction signal at each iteration is an appropriately scaled and phase-rotated kernel signal that is circularly shifted to have its peak coincide with the peak of the signal whose PAPR is being reduced in that particular iteration. The kernel signal is a pulse waveform that approximates the Dirac delta function and is given by

$$\mathbf{p} = \frac{\sqrt{N}}{L}\text{IFFT}(\mathbf{P}) \qquad (2.6)$$

where N and L denote the total number of subcarriers and the number of reserved subcarriers, respectively, and $\mathbf{P}$ is a frequency-domain vector that has peak-reduction weights set to one at reserved positions and zero for all the other elements.

At the ith iteration, the peak-reduced transmit signal can be expressed as

$$\mathbf{x}^i = \mathbf{x}^{i-1} - \alpha_i \mathbf{p}[((n - m_i))_N] \qquad (2.7)$$

where m_i is the index of the peak of signal $\mathbf{x}^{i-1}$, $((n - m_i))_N$ denotes a circular shift by m_i samples and α_i is a complex scaling factor given by

$$\alpha_i = \frac{x^{i-1}(m_i)}{\max\limits_{0 \leq n \leq N-1}|x^{i-1}(n)|}\left(\max\limits_{0 \leq n \leq N-1}|x^{i-1}(n)| - x_{CL}\right) \tag{2.8}$$

where x_{CL} is a predetermined clipping level.

The effective peak-cancelling signal at the ith iteration is obtained as follows:

$$\mathbf{c} = -\sum_i \alpha_i \mathbf{p}[((n - m_i))_N] \tag{2.9}$$

and the peak-reduction weights vector is given by $\mathbf{C} = \text{FFT}(\mathbf{c})$, while the peak-reduced transmit signal is the combined signal $(\mathbf{x} + \mathbf{c})$. The algorithm stops either when the targeted PAPR or the maximum iteration number is reached.

The K-TR scheme can achieve good PAPR reductions. However, a major drawback of the scheme is the likelihood that a new peak is generated when equation (2.7) is on the executed. This problem, normally referred to as peak regrowth or regeneration, occurs because the kernel signal may have nonzero entries at other positions apart from the location of the peak amplitude. The peak regeneration destabilises the trajectory of the algorithm, hence slowing down its convergence. Additionally, the use of K-TR method increases the average transmit power.

Another old suboptimal method is the clipping control tone-reservation (CC-TR) method that was proposed in 1997 by Gatherer and Polley [47]. It also employs a gradient algorithm to find peak-reduction weights. The method starts by hard clipping the multicarrier signal $\mathbf{x}$ to obtain a clipped signal $\mathbf{y}$ with the entries given by

$$y(n) = \begin{cases} x(n), & |x(n)| < x_{CL} \\ x_{CL}e^{j\varphi(n)}, & |x(n)| > x_{CL} \end{cases} \tag{2.10}$$

where $\varphi(n)$ is the phase of $x(n)$.

From the clipped signal, the so-called correction signal is obtained as follows:

$$\tilde{\mathbf{c}} = \mathbf{y} - \mathbf{x} \tag{2.11}$$

On this signal, a FFT is applied to obtain a frequency-domain vector containing peak-reduction weights as follows:

$$\tilde{\mathbf{C}} = \text{FFT}(\tilde{\mathbf{c}}) \tag{2.12}$$

However, to conform to the requirements of tone-reservation, the weights on the non-reserved positions are set to zero, thus giving the elements of the actual peak-reduction weights vector **C** as

$$C(k) = \begin{cases} \tilde{C}(k), & k \in i_{PRC} \\ 0, & k \notin i_{PRC} \end{cases} \tag{2.13}$$

where i_{PRC} denotes the set of indices of the reserved-subcarrier locations. From this equation, the peak-reduction signal is obtained by IFFT using equation (2.5).

Every iteration in the method results in a time-domain signal **c**, which is added to the input signal to yield a peak-reduced signal. Thus, at the ith iteration, the peak-reduced signal is given by

$$\mathbf{x}^i = \mathbf{x}^{i-1} + \mu\mathbf{c}^i \tag{2.14}$$

where μ is a weighting factor and this together with the maximum iteration number are the parameters of the algorithm.

The CC-TR method can attain a modest PAPR reduction with a slight degradation of BER performance. However, the maximum peak of the peak-reduction signal during the first iteration of the method is extremely smaller than the maximum peak of the original correction signal. This disorients the method, thus making it to require too much iteration to converge, thereby increasing its computational complexity.

A tone reservation method based on least-squares approximation (LSA-TR) in [48] utilizes the CC-TR algorithm to find peak-reduction weights and least-squares approximation to obtain an optimal weighting factor μ. Although the method converges fast and has a small increase in average power, it exhibits poor PAPR reduction performance. In addition, the method has a high computational complexity.

In [49], a suboptimal tone-reservation method based on curve-fitting, abbreviated as CF-TR method, was proposed. The CF-TR method creates a peak-reduction signal by applying curve-fitting optimization to another signal referred to as a clipping noise signal, which is similar to the correction signal in equation (2.11). The reduction in PAPR by this method is strictly determined by the clipping threshold and is greatly diminished when the number of reserved subcarriers is less than the number of nonzero elements in the clipping noise vector. Although the method achieves a moderate PAPR reduction, it evaluates the computationally intensive Moore-Penrose pseudoinverse that gives inaccurate results in some instances when the coefficient matrix is close to being

16

singular. In addition, the method is prone to peak regeneration and it increases the average power of the resultant transmit signal.

The paper in [50] proposed a scaling signal-to-clipping noise ratio tone-reservation (SSCR-TR) method. A clipping noise signal is scaled down by an optimal factor to obtain a kernel peak-reduction signal. The scaling factor is found by solving a least-squares approximation problem with peak-regeneration constraints. The method converges fast but it is still disposed to peak regeneration and has the PAPR reduction performance strongly determined by the clipping threshold employed.

A suboptimal weighted tone-reservation method in [51] solves a weighted least-squares approximation problem to generate peak-reduction weights. Although the method is aimed at suppressing peak re-growth problem during PAPR reduction, it is not able to eliminate it. In addition, the method is not able to find the optimal peak-reduction weights and has poor PAPR reduction performance.

The work in [52] proposed a tone reservation method, which combines artificial neural network and initial-value optimization (IVO-TR). It was aimed at reducing the runtime computational complexity through an offline generation of peak-cancelling signals. At runtime, the method classifies an OFDM signal and then searches for an appropriate peak-cancelling signal from among those stored in a pre-work table. The CC-TR algorithm is employed to generate peak-reduction signals, which become the training targets for feedforward neural network. Consequently, the reduction in PAPR strongly depends on the CC-TR performance. In addition, the training process is computationally intensive, requires costly and powerful servers, and takes a lot of time due to the large set of data involved to generate near optimal peak-reduction signals. Moreover, the prediction process at runtime still has high computational complexity and the average transmission power is increased.

An extreme learning machine tone-reservation (ELM-TR) scheme was proposed in [53]. It is based on an online sequential extreme learning machine. The training time is reduced by having a single hidden layer in the feedforward neural network. Peak-reduction signals generated by the CC-TR scheme are used to train the neural network. Thus, the CC-TR implementation determines the PAPR reduction performance of the method. In addition, the generation of training and test data sets and the training process is computationally intensive and a huge storage space is needed to store the input data and several trainable parameters.

A sub-optimal tone-reservation referred to as selective TR (S-TR) was proposed

in [54] for use to reduce PAPR in MIMO-OFDM systems employing space-frequency block codes (SFBC). The proposed algorithm reduces the peak power by adding a kernel peak-reduction signal to the transmit signal with the maximum PAPR. Although this method has lower computational complexity than optimal schemes, it is prone to re-generation of new peaks during iterations, which slow down the convergence, and it increases the average power of the transmit signal. In addition, it has a poor PAPR reduction performance.

C. Tone Injection

The tone injection [55] method makes use of the so-called expanded constellation to reduce PAPR. An original constellation point is mapped onto one of the several points in the expanded constellation. The possible mapping points are spaced at the same distance from the original point. All the original symbols are mapped onto one of their possible alternatives in the expanded constellation to yield a combination of symbols that result in reduced PAPR.

The operation of replacing an original constellation point with another one in the expanded constellation is similar to injecting a tone with an appropriate frequency and phase into the OFDM signal; thus the name of "tone injection". Similar to the tone reservation, this method too does not require transmission of side information and has no effect on the BER of the underlying system. However, these advantages are outwitted by the high computational complexity involved to find a new set of modulation symbols and the rise in the average transmit power owing to the use of the expanded constellation.

D. Adaptive Symbol Selection / Selective Level Mapping

Adaptive symbol selection or selective level mapping (SLM) [56] - [58] is based on the creation of P different signals from the original input data vector. All signal alternatives bear the same information. After the IFFT, the variant that yields the smallest PAPR is transmitted. Generally, two approaches are used to create signal alternatives. The first approach, illustrated in Figure 2.1, applies $(P-1)$ interleavers on the input bit stream to generate P OFDM symbols X_i, $i = 0,1,\dots,P-1$, while the second one in Figure 2.2 involves multiplying the original symbol X with different phase sequences ψ_i, $i = 0,1,\dots,P-1$, of the same length as X.

Depending on the choice of the SLM approach, the extent of PAPR reduction is highly influenced by the method of interleaving and the number of interleavers or by the

number and mode of generation of phase sequences. A large number of interleavers or phase sequences are required to attain a substantial PAPR reduction. The optimization process of selecting the signal with the lowest PAPR out of P signals may be computationally intensive if the number of subcarriers and P are large.

Moreover, the knowledge about the method chosen for interleaving or phase sequencing needs to be sent to the receiver for the demodulation of symbols. The transmitted side information about interleaving and phase sequence reduces the data-rate of the system. In case of errors during the detection of the side information, the recovery of the data symbols may be impossible. Blind SLM schemes [59], [60] have been proposed to avoid transmission of side information but the complexity of demodulation at the receiver is greatly increased.
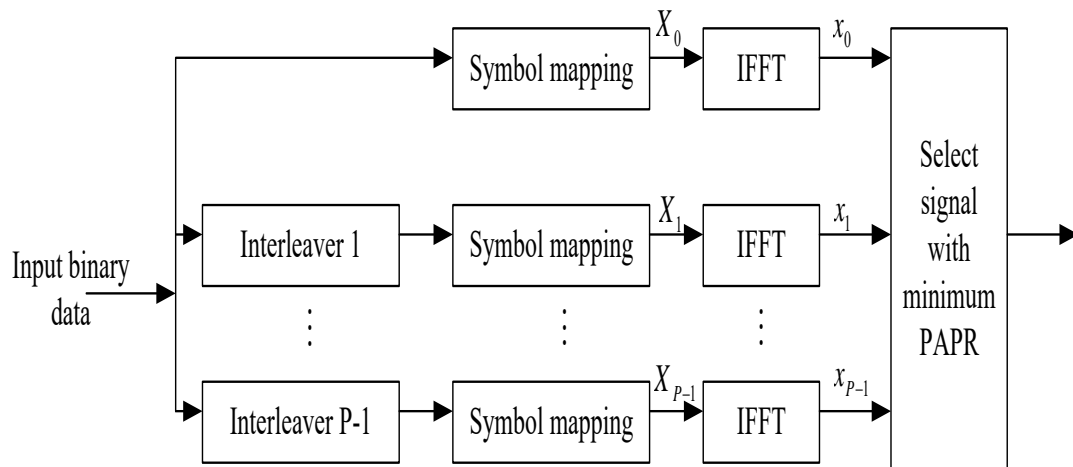
Figure 2.1. Block diagram of SLM using interleavers [56]
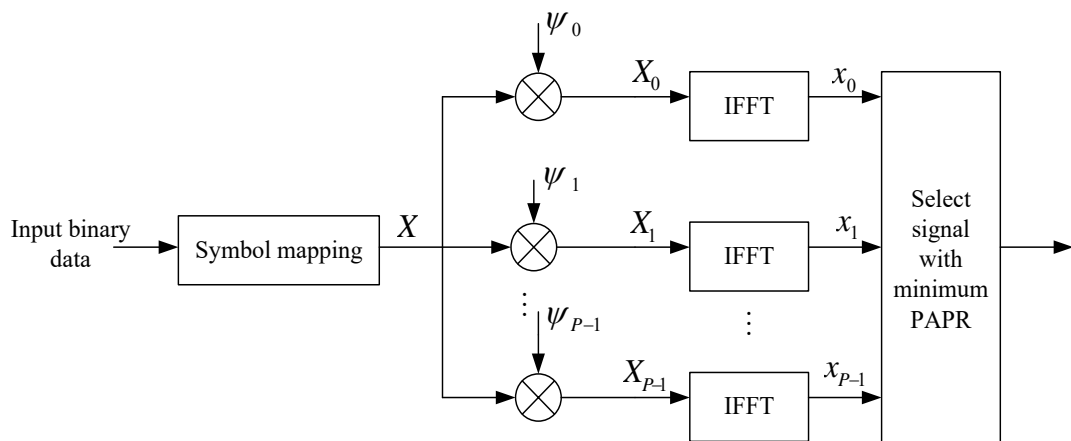
Figure 2.2. Block diagram of SLM using phase sequences [56]

A. Partial Transmit Sequences

In partial transmit sequences (PTS) method [61] - [63], the input symbol to the IFFT is split into P disjoint sub-blocks X_i, $i = 0, 1, ..., P - 1$, as illustrated in Figure 2.3. The IFFT operation is then performed, not on the entire N-length OFDM symbol, but on the separate sub-blocks. The output of the ith sub-block is then multiplied by the so-called complex rotation factor b_i. The values of b_i, $i = 0, 1, ..., P - 1$, for all sub-blocks, are optimized to yield a combined signal, $x = \sum_{i=0}^{P-1} b_i x_i$, with minimum PAPR. This optimization has to be performed in real time for each input OFDM symbol to the IFFT and hence leads to increased transmitter complexity.

Moreover, the information about the complex rotation factors has to be sent as side information to the receiver. This transmission of the side information reduces the useful data rate. In addition, as demonstrated in [64], although the PAPR reduction increases with the number of sub-blocks, the effectiveness of PTS schemes are limited.

B. Active Constellation Extension

In this method, abbreviated as ACE, the exterior constellation points in the set of modulation symbols constituting one OFDM symbol are extended in the outward direction from their original positions in the constellation map to reduce PAPR [65], [66]. The motivation behind this method is to reduce PAPR by moving some points of the modulation symbols farther away from the decision boundaries without degrading the system BER.
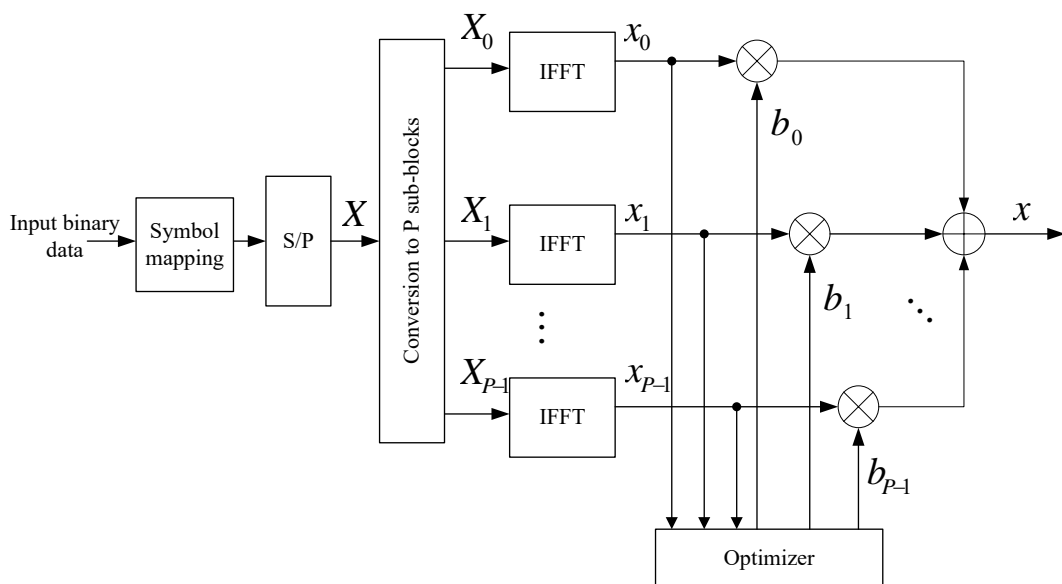
Figure 2.3. Block diagram of OFDM transmitter with PTS scheme [64]

20

The ACE modification process is illustrated in Figure 2.4 for QPSK and 16-QAM modulation schemes. Each of the shaded regions represents the feasible points of a possible extension of the corresponding original point located at the innermost corner of region. For the 16-QAM constellation, the possible movement for the outer non-corner points are the straight lines from the point towards infinity.

The ACE method guarantees that those points in one feasible region and another have a minimum Euclidean distance less than that between the corresponding original points. This ensures that only a small degradation of BER performance occurs. The constellation extension operation is transparent to the receiver. This implies that the ACE method does not need transmission of any side information and thus does not affect the system data rate.

However, although the ACE technique can be applied to any digital modulation scheme, a large constellation size has a small number of possible extension points and this limits PAPR reduction. One consequence of constellation extension is the increase in transmit power that needs to be managed when the ACE method is employed. In addition, the complexity of the ACE technique strongly depends on the method used to extend the constellation points.

(a)   (b)
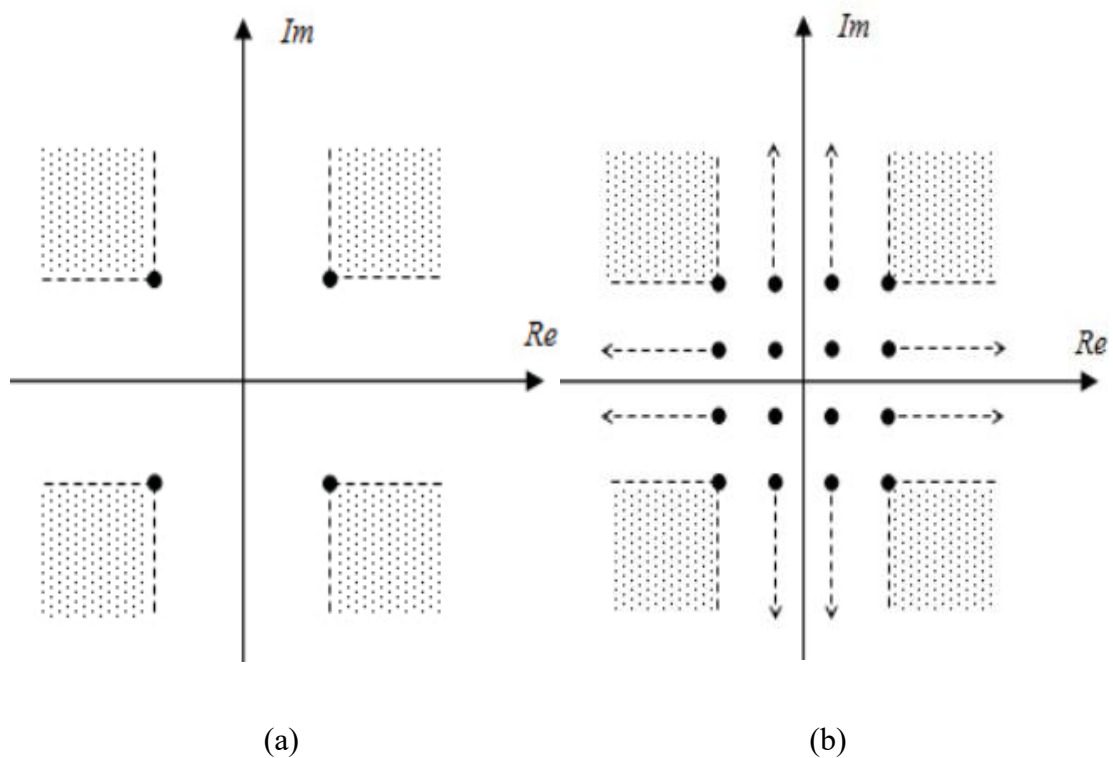
Figure 2.4. Illustration of ACE method [65] (a) QPSK (b) 16-QAM

2.1.3 Hybrid PAPR Reduction techniques

For this class of methods, two or more techniques are combined in order to improve PAPR reduction and reap from the advantages of different schemes. However, the hybrid schemes are more computationally complex [67] to implement in both hardware and software than discrete methods. The main aims of combining different methods have been to improve PAPR reduction, lower computational complexity and reduce BER. For example, in [68] a method combining PTS and error-correction coding (ECC) was proposed. This PTS-ECC technique offers better PAPR reductions than the ordinary PTS scheme especially when using pseudo-random sub-blocking partitions to reduce the complexity of implementing the PTS algorithm.

Similarly, methods combining SLM and error-correction codes, and referred to as SLM-ECC schemes, have been proposed in [69] and [70]. These methods are built on the concatenated SLM scheme, which employs label insertion and scrambling to avoid transmission of side information and an error correction scheme together with interleaving block to improve BER performance. Typical error correction codes such as linear block codes, convolutional, turbo or cyclic codes can be used. However, the BER performance improvement is at the expense of reduced PAPR reduction performance and increased system complexity.

Another example of a hybrid scheme was proposed in [71]. It combines iterative clipping and filtering technique and enhanced nonlinear companding scheme. The enhanced companding scheme does not require an expansion operation at the receiver to reverse signal compression at the transmitter. The hybrid method has a computational complexity somewhere between that of the enhanced nonlinear companding scheme and that of the iterative clipping and filtering technique. In addition, the hybrid method achieves better PAPR reduction capability and BER performance than the iterative clipping and filtering scheme alone. Additionally, the BER performance of the hybrid method is better than that of the enhanced nonlinear companding scheme when considering the same amount of PAPR reduction.

The work in [72] proposed a hybrid technique that combines selective codeword shift and selective level mapping schemes (SCS-SLM) with aim to improve PAPR reduction in Alamouti coded MIMO-OFDM systems. The hybrid scheme offers a better performance than that offered by the individual SCS and SLM schemes; however, PAPR

reduction is still poor.

A hybrid PAPR reduction method combining and optimizing three techniques, namely convolutional code, successive suboptimal cross-antenna rotation and inversion and iteratively modified companding and filtering, here abbreviated as CSC, was proposed for Alamouti encoded MIMO-OFDM systems in [73]. The hybrid scheme can achieve good PAPR reductions but at the cost of a degraded BER and increase in system complexity.

2.2 Review of Alternatives to PAPR Reduction Techniques

Typically, the PAPR of an OFDM signal ranges between 10 and 12 dB. Considering the lower value, if there is no PAPR reduction it would require the HPA to be operated with a 10 dB power back-off from the 1-dB compression (P1 dB) point to linearly amplify all signal amplitudes. However, the 10 dB power back-off significantly reduces the efficiency of the HPA. With typical efficiency values being in the range of 40-70% at the P1dB point, the reduction could be to levels below 5% [74]. Consequently, applications based on OFDM technique are likely to cause poor battery lifetime in mobile terminals and high electrical power consumption in base stations.

Two alternatives to PAPR reduction for avoiding operating the HPA in low power-efficiency point are either to extend the linear region of operation using linearization techniques or to increase the power efficiency through efficiency enhancement methods. Linearization techniques are applicable to a HPA that is operated near the saturation point where the efficiency approaches the maximum value but the gain is compressed and nonlinear distortion is likely to occur.

On the other hand, efficiency enhancement techniques are applicable to a HPA that has been backed-off to have a linear performance and thus have low power efficiency. Therefore, a solution to improve linearity or efficiency ends up combining a linearizer and an efficiency enhancement technique. Several techniques proposed under the two approaches are discussed briefly in the following two subsections.

2.2.1 Linearization Techniques for Power Amplifiers

Linearization techniques attempt to correct the distortion from nonlinear amplification by providing an inversion function or by generating an appropriate correction signal. These techniques aim to extend the linear range of the power amplifier

to reduce signal distortion, specifically the intermodulation products, to below -60 dBc.

The three most common linearization techniques are the digital predistortion, feedforward and feedback methods. The first two have lately attracted a lot of attention and are undergoing intensive research and development for use in wireless communication applications [75].

A. Digital Predistortion

Predistortion generally refers to any technique that linearizes the HPA through an appropriate modification of amplitude and phase responses of the input signal. This is typically realized by passing an input signal through a predistorter system whose transfer function is the inverse of the HPA as illustrated in Figure 2.5. As a result, the combination of the predistorter and HPA yields a transfer function that allows for linear amplification of all the amplitudes of the input signal $v_i(t)$.

Predistortion can be performed either at the baseband or passband level. However, in order to utilise the immense potential of digital signal processing, predistortion is performed at baseband frequencies and is thus generally referred to as digital predistortion. Two well-known drawbacks of the digital predistorters are their high sensitivity to analogue circuitry imperfections and their use, with some parameters, can result in intolerably high adjacent channel emissions [76].

(a)

(b)
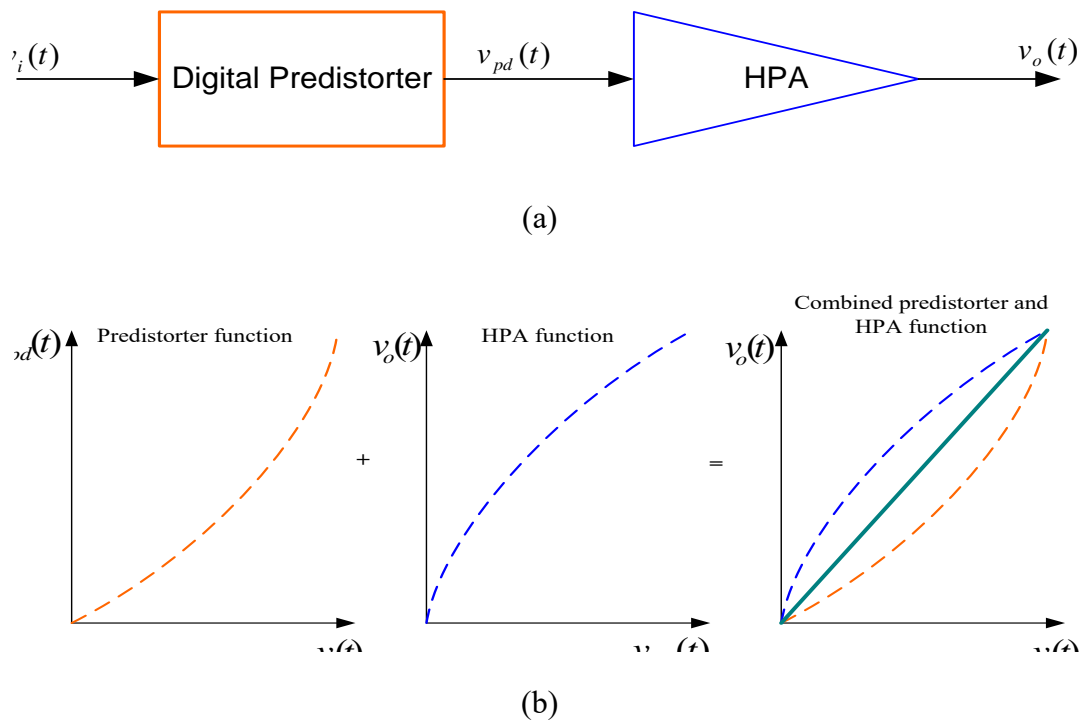
Figure 2.5. Digital predistortion (a) Device connection (b) Operation

B. Feedforward technique

The feedforward method is the dominant linearization technique in multicarrier transmission systems. It attempts to cancel the nonlinear distortion of the power amplifier by applying a corrective signal on the output as illustrated in Figure 2.6. The input signal $v_i(t)$ is applied to two channels. In the first channel, the signal is amplified through the main PA. In the second channel, the signal is appropriately delayed in the block labelled t_a to give a reference signal for comparison with the output of the attenuator. The block t_a compensates for the delay incurred during the processing of the input signal through the main PA and attenuator.

The output of the main PA can be considered to consist of the undistorted input signal and the distortion signal, i.e.

$$v_a(\text{t}) = A_v v_i(t) + v_d(t) \tag{2.15}$$

For simplicity of notation, the delay through the main PA is not indicated. The output of the main PA is simultaneously passed to the attenuator and to the delay block t_e. The output of the attenuator is given by

$$v_c(\text{t}) = v_i(t) + \frac{v_d(t)}{A_v} \tag{2.16}$$

Assuming perfect delay compensation in the block t_a, the output of the first summation block is an error signal given by
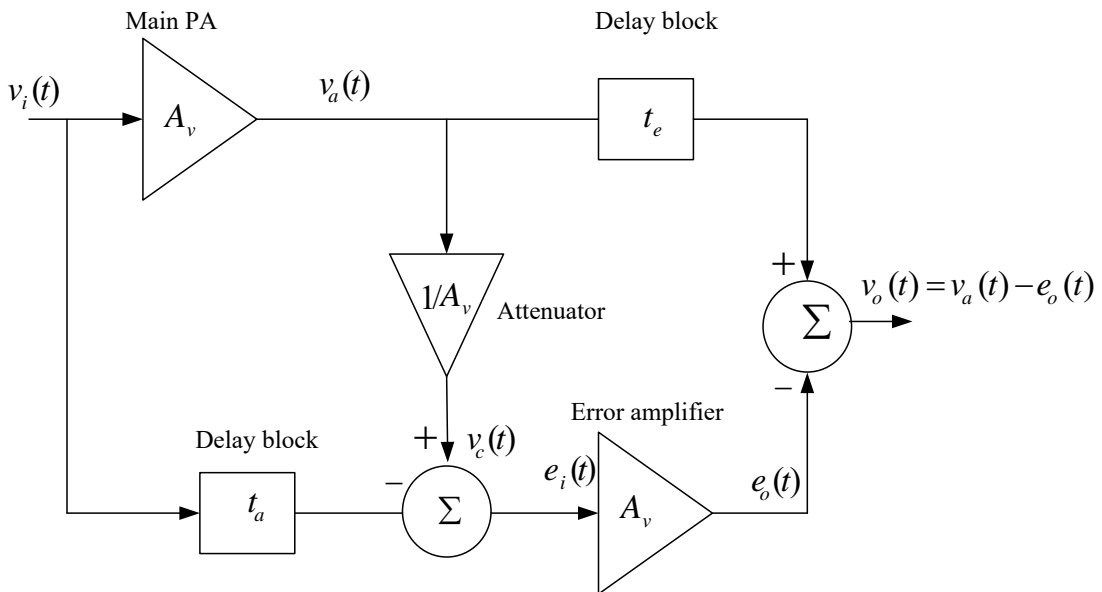
Figure 2.6. Feedforward linearization technique [75]

25

$$e_i(t) = v_i(t) - v_c(t) = \frac{v_d(t)}{A_v} \tag{2.17}$$

Signal $e_i(t)$ is amplified through the error amplifier to give a distortion signal

$$e_o(t) = v_d(t) \tag{2.18}$$

Finally, the distortion signal is subtracted from the output of the main PA, after the latter has passed through the delay block t_e. The passage of the main PA output through block t_e compensates for the processing time in the error amplifier. Again assuming perfect delay compensation, the output of the second summation is

$$v_o(t) = v_a(t) - e_o(t) = A_v v_i(t) \tag{2.19}$$

which is the undistorted part of the signal in equation (2.15).

Therefore, the feedforward technique is able to cancel the nonlinear component in the output of the main PA and this is in addition to other known benefits of feedforward mechanism such as the suppression of internally generated noise and bandwidth extension. These benefits make the technique very attractive for use in multicarrier wireless applications.

However, use of feedforward linearization may result in both gain and phase mismatches in the two signals at the input of the summation blocks. These mismatches affect the linearization that can be realised using the technique. In addition, although delay blocks can be realized using passive elements or transmission lines to alleviate the mismatch problems, the blocks may reduce the amplifier efficiency owing to power dissipation and they are difficult to design especially for wideband applications.

C. Feedback technique

The feedback technique can also be considered as a form of input signal correction or predistortion [77]. In this technique, a feedback network is added to pass a part of the output signal to the input terminal as shown in Figure 2.7. This linearization technique is based on the principle that a large change in open-loop gain causes only a small change in the closed-loop gain of a negative feedback amplifier. This is normally expressed using the following relationship between the two gains:

$$\frac{dA_f}{A_f} = \frac{1}{1 + \beta A_v} \frac{dA_v}{A_v} \tag{2.20}$$

where A_f, A_v and β are the closed-loop gain, open-loop gain and feedback factor,

respectively.

The output signal can be expressed in the form

$$v_o(\text{t}) = A_v v_i(t) + v_d(t) - \beta A_v v_o(t) \tag{2.21}$$

where $v_d(t)$ represents the distortion introduced by the nonlinear amplification of the input signal though the HPA.

Equation (2.21) can be rearranged to give the output signal as

$$v_o(\text{t}) = \frac{A_v}{1 + \beta A_v} v_i(t) + \frac{v_d(t)}{1 + \beta A_v} \tag{2.22}$$

where the distortion is now reduced by a factor $1 + \beta A_v$ due to the feedback network.

From equation (2.22), it is clear that the overall gain on the input signal is reduced by the negative feedback from A_v to $A_f = A_v/(1 + \beta A_v)$. This means that a feedback amplifier needs a large drive in order to give the same response level as the amplifier with no feedback.

In addition, the reduction in nonlinear distortion is pegged on the loop gain βA_v being large. At RF levels, a large loop gain may not be achieved. Additionally, the feedback loop may affect the stability of the system due to the introduction of a large number of poles by intrinsic components. This large number of poles may result in an extreme phase shift near 180°, thus making the system to oscillate.
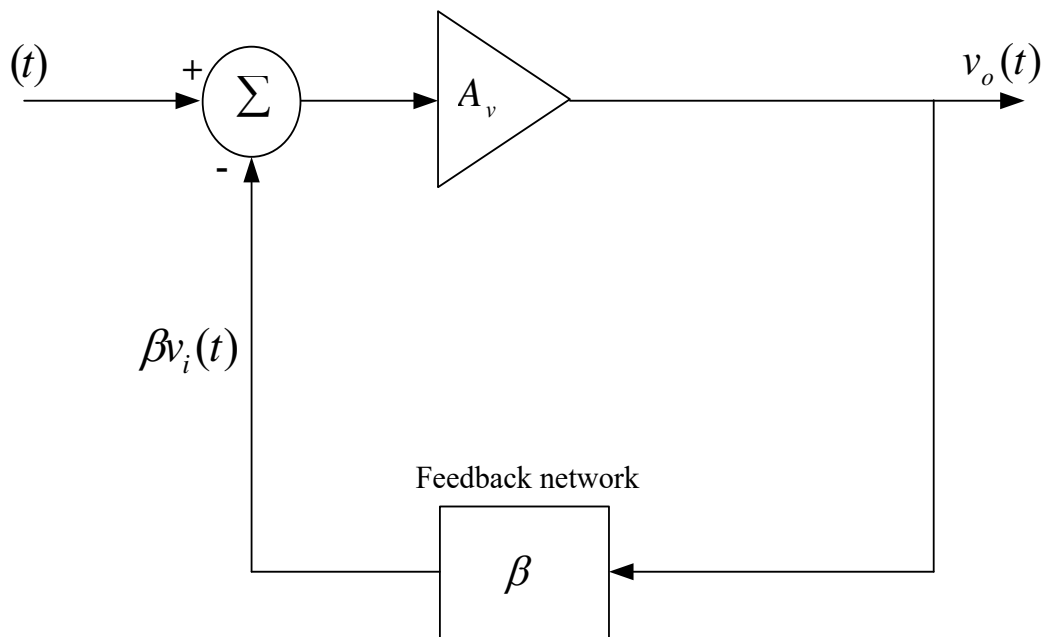
Figure 2.7. Feedback linearization technique

2.2.2 Efficiency Enhancement Techniques for Power Amplifiers

One of the inherent problems of amplifying the power of amplitude-modulated RF signals is that the circuit of a conventional PA is designed to provide maximum efficiency at a single power point typically near the rated maximum power for the device. When the PA is backed-off away from this level, its efficiency is significantly reduced and this is accompanied by increased power dissipation, even though the output power is reduced.

To avoid the operation of PA at low power efficiency level, several efficiency enhancement techniques have been proposed. The four most common approaches are dynamic loading (Doherty amplifier), dynamic biasing, envelope elimination and restoration (EER) and envelope tracking (ET) [78]. These techniques work either by varying the load or by biasing the PA according to the amplitude characteristics of the input signal. These techniques are briefly reviewed here below.

A. Doherty power amplifier

The Doherty power amplifier, originally proposed in 1936 [79], employs dynamic loading to sustain high efficiency operation over a broad range of the input voltage in a linear PA supplied with high PAPR amplitude-modulated signals. The amplifier consists of two power amplifiers, referred to as main (or carrier) and auxiliary (or peak) PA as shown in Figure 2.8. The transistors in the two amplifiers are operated as dependent current sources. The main device is Class B (or Class AB) power amplifier, while the auxiliary is of Class C. The quarter-wave transmission line introduces a 90° phase shift on the signals passing through it.

The final maximum RF output power of the system is the sum of the powers of the two amplifiers. In a typical implementation, as the main PA saturates at high input power, the auxiliary amplifier turns on. When the input drive decreases by about 6 dB from the maximum rated power, the auxiliary amplifier turns off and stops drawing DC input power. Thus the two amplifiers operate together only at high input power levels. The turning-off of the auxiliary PA helps to improve the power efficiency of the system at low input power levels. In addition, under correct impedance matching conditions, the main PA can remain near the maximum efficiency during the top 6 dB of the power range.

The high efficiency together with a constant voltage gain over a broad range of input power is achieved by presenting the main amplifier with a load that dynamically

varies with the current of the auxiliary amplifier. The load variation is based on the principle that the impedance seen by one current source can be changed by supplying a current from another source.

The quarter-wave transmission line between the main PA and the load R_L acts as an impedance transformer. Assuming a typical implementation using MOSFET transistors, the transformed load seen by the main PA is the dynamic resistance given by

$$R_m = \frac{z_0^2}{R_L \left(1 + \frac{I_a}{I_m}\right)} \qquad (2.23)$$

where z_0 is the characteristic impedance of the transmission line and I_a and I_m are the maximum values of the fundamental components of the drain currents of the auxiliary PA and main PA, respectively.

From equation (2.23), if current I_a decreases from the maximum value of I_m to zero with the gate-to-source voltage $v_i(t)$, the load resistance R_m increases from $z_0^2/2R_L$ to z_0^2/R_L. Therefore, as I_m decreases, the maximum value of the fundamental component of the drain-to-source voltage of the main amplifier given by $V_m = R_m I_m$ remains a constant. This is the condition for achieving high efficiency over the dynamic range of the input signal for a fixed voltage supply V_{DD}. Thus, the efficiency of the main PA is

$$\eta = \frac{\frac{1}{2}V_m I_m}{V_{DD} I_D} = \frac{1}{2}\left(\frac{I_m}{I_D}\right)\left(\frac{V_m}{V_{DD}}\right) \qquad (2.24)$$

where I_D is the drain current.
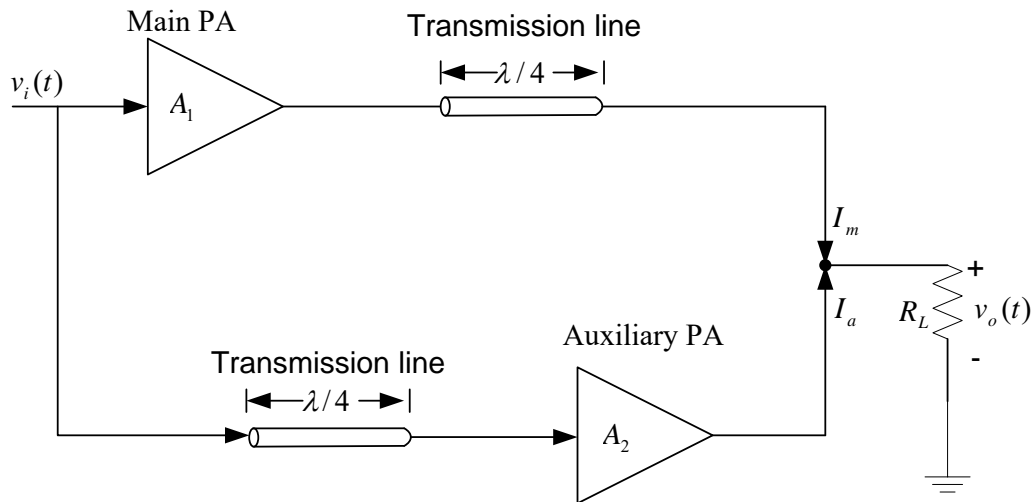
Figure 2.8. Block diagram of Doherty amplifier [75]

Since the ratio I_m/I_D is constant, owing to the fact that the drain current has a conduction angle that is approximately constant, a high efficiency can be achieved if the ratio V_m/V_{DD} is fixed at a value close to 1, a condition satisfied when V_m is constant. Thus, when the auxiliary amplifier becomes active, it pulls the load impedance of the main amplifier to hold the maximum output voltage at a constant. This in turn achieves maximum efficiency over a wide range of the input power levels. The main drawback of the Doherty power amplifier is its weak linearity and this can only be improved by an additional analogue circuitry.

B. Envelope elimination and restoration technique

The envelope elimination and restoration (EER) technique, initially proposed by Kahn in 1952 [80], is illustrated in Figure 2.9. The input signal $v_i(t)$ is simultaneously fed to two signal-processing paths. One path comprises of an envelope detector followed by a low-frequency PA. The second path contains an RF limiter and a nonlinear high efficiency switched-mode PA.

The input signal to the envelope detector can be considered to be of the form

$$v_i(t) = e(t) \cos[2\pi f_c t + \theta(t)] \tag{2.25}$$

where $e(t)$ is the signal envelope or the amplitude modulation part and $\theta(t)$ is the phase modulation part.

The envelope detector separates the signal envelope from the input signal, thus giving the output as the envelope $e(t)$. This signal envelope is then amplified by the low-frequency PA to give $A_v e(t)$, which then modulates the DC power supply of the switched-mode RF PA to yield $V_{DD} A_v e(t)$.

On the hand, the limiter gives a constant-envelope signal $\vartheta(t)$, which is fed to the gate of the switched-mode PA as the input. Since the switched-mode RF PA is saturated, the output voltage amplitude is proportional to the modulated DC supply voltage. Therefore, the envelope of the input signal $v_i(t)$ is restored back at the output of the switched-mode PA.

One important advantage of the EER technique is that it avoids the use of a linear power amplifier, which has very low power efficiency that also decreases with the level of the drive voltage. Highly efficient switched-mode PAs such as those in class D, E, DE or F have efficiencies that are not affected by the level of the drive voltage and can be used in the EER technique [81].
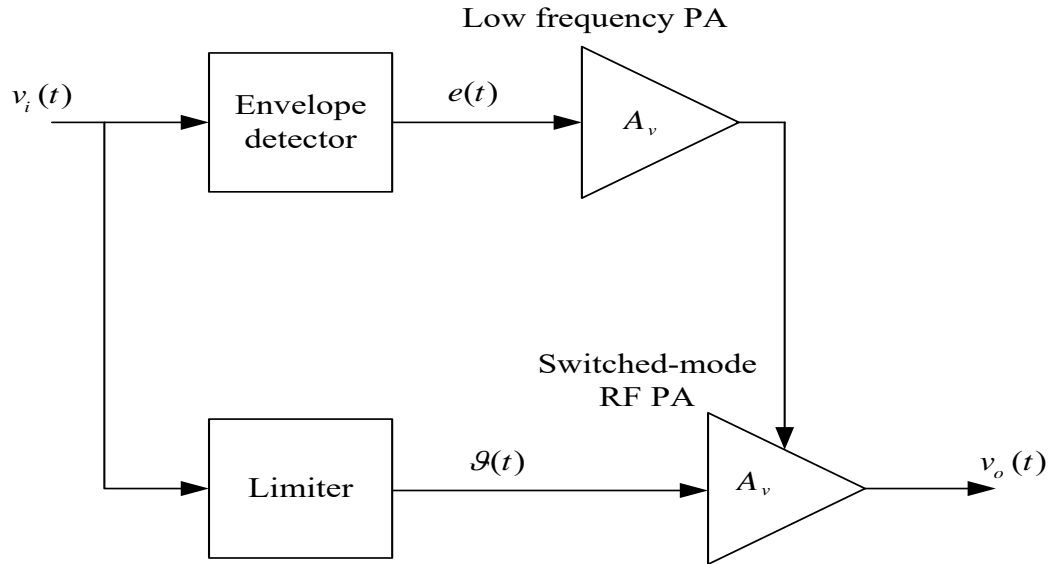
Figure 2.9. Block diagram of envelope elimination and restoration scheme

However, the EER system suffers from three major drawbacks. Firstly, there is a large phase mismatch between the two signal-processing paths. This is because the phase shift in the low-frequency path is greater than in the RF path. Secondly, since the limiter is built from nonlinear components, it introduces a phase distortion that corrupts the phase of the input signal. Thirdly, the amplitude modulation of the DC supply in the switched-mode PA causes the nonlinear transistor capacitance to introduce a phase distortion.

C. Envelope tracking technique

This efficiency enhancement method applies an envelope-derived modulation similar to the EER technique to the DC power supply of a linear RF power amplifier as illustrated in Figure 2.10 [78]. The level of the supply voltage is varied in proportion to the drive voltage in such a way as to ensure that the ratio $V_m/V_{DD} \approx 1$ in equation (2.24) in order to achieve a highly efficient amplification.

One major disadvantage of the envelope tracking method is the introduction of gain and phase variations, which significantly reduces the linearity of the PA. To compensate for these variations a complex digital predistortion circuitry is required.

2.3 Knowledge Gaps

Based on the review of the PAPR reduction methods in the literature, it can be concluded that the techniques that have been proposed up to now fail to meet the majority

31

of the attributes that are desired for practical implementation especially in real-time systems. Several attributes have been cited by different sources [82], [83] as the basic requirement for a practically realizable PAPR reduction method. The most important attributes are the non-degradation of BER, low computational complexity, simplicity in implementation, high PAPR reduction capability and that the method should not affect the transmitted power and data-rate.

In addition, the review shows that linearization techniques extend the linear operation of the power amplifier at the expense of an increased complexity in digital signal processing especially for digital predistortion or a complex analogue circuitry in the case of feedforward and feedback techniques.

Similarly, the efficiency enhancement techniques are practically difficult to realize especially for high volume applications. This is because of the required additional analogue circuitry and an exact time synchronization mechanism between the carrier and the envelope of the modulating signal.

From the above-mentioned drawbacks of linearization and efficiency enhancement techniques, the reduction of PAPR using a practically realizable method stands out as the best way to avoid signal distortion in a HPA owing to nonlinear amplification. For this reason, the aim of this work is to develop new PAPR reduction methods that have the desirable attributes for practical implementation and use in wireless communication systems employing OFDM technique.
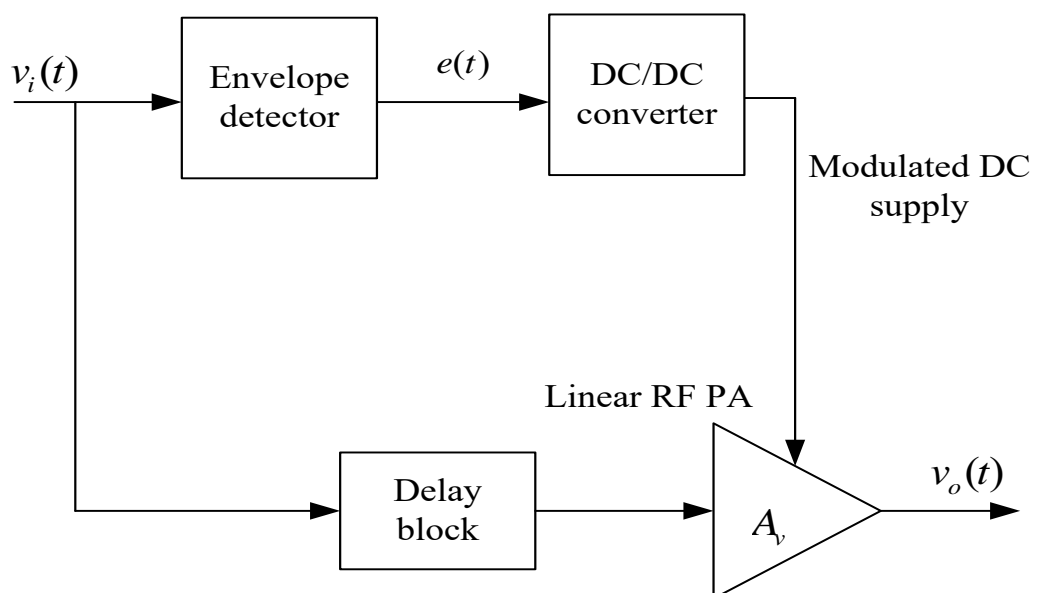
Figure 2.10. Block diagram of envelope tracking

CHAPTER 3

ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING

Orthogonal frequency division multiplexing (OFDM) is a multicarrier modulation technology that continues to be widely employed in signal broadcasting and wireless communication networks. Some of the applications include digital audio broadcasting (DAB), 1st and 2nd generations of terrestrial digital video broadcasting (DVB-T1 and DVB-T2), digital multimedia broadcasting (DMB), 4th and 5th generations (4G and 5G) of mobile communication network, IEEE 802.16 worldwide interoperability for microwave access (WiMAX), IEEE 802.11 wireless local area network (WLAN), power-line and optical communications [84] - [86]. In addition, OFDM is being considered as one of the candidate technology that will be combined with other technologies to meet the expanded requirements and use cases for radio access network in the future 6G mobile communication network [86] - [89].

The use of OFDM in several wire-line and wireless communication applications is propelled by several advantages. The three main ones that make the technology very competitive are the high spectral efficiency, its robustness against frequency-selective fading and intersymbol interferences. This chapter is intended to cover the fundamentals of OFDM transmission together with the advantages and disadvantages of the technique.

3.1 Basic Principles of OFDM

The OFDM technique divides the total available channel bandwidth into many equally spaced narrowband subchannels. Each subchannel bandwidth, also referred to as *subcarrier spacing*, is given by

$$\Delta f = \frac{W}{N} \tag{3.1}$$

where W is the bandwidth of the system and N is the total number of subchannels.

The signal carriers on the subchannels are mutually orthogonal to one another, thus allowing adjacent spectra to overlap without interferences. Each subcarrier signal can be considered as a product of a single-tone signal and a rectangular pulse. The

spectrum of the rectangular pulse is a *sinc* function, as illustrated in Figure 3.1, which has a zero-closing bandwidth of $2/T$. Thus, a typical OFDM signal spectrum consists of N adjacent *sinc* functions, which are shifted by Δf in the frequency axis as illustrated in Figure 3.2 for $N = 4$ subcarriers. The peak of each subcarrier occurs at a point where the amplitude of all the other subcarriers is zero. This overlapping of adjacent spectra significantly increases the spectral efficiency of the system.

(a)   (b)
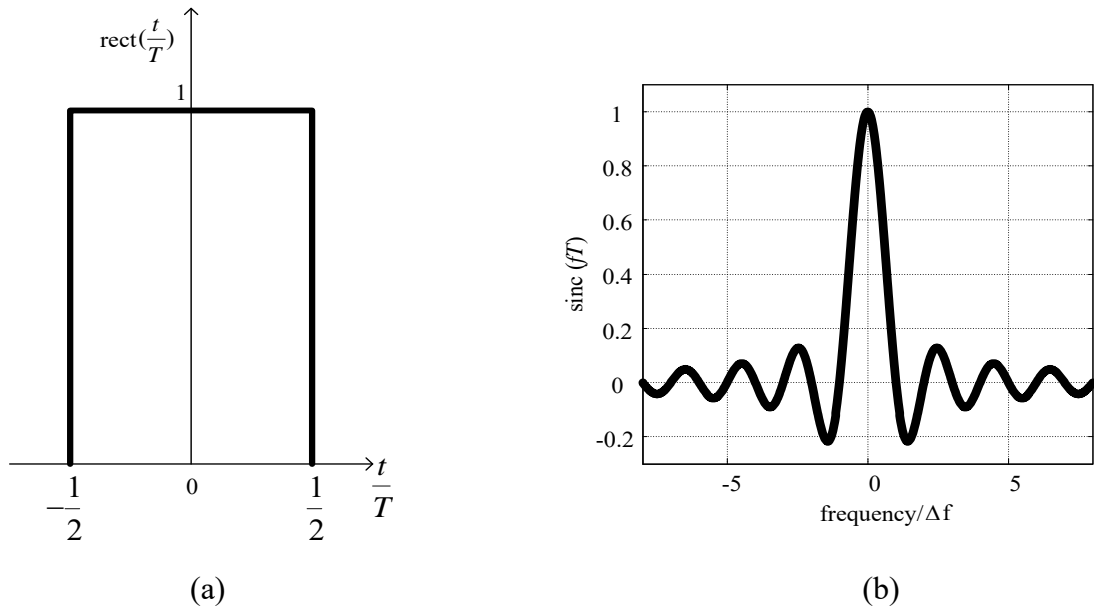
Figure 3.1. Fourier spectrum of a rectangular pulse (a) rect(*t/T*) pulse (b) sinc(*fT*) pulse
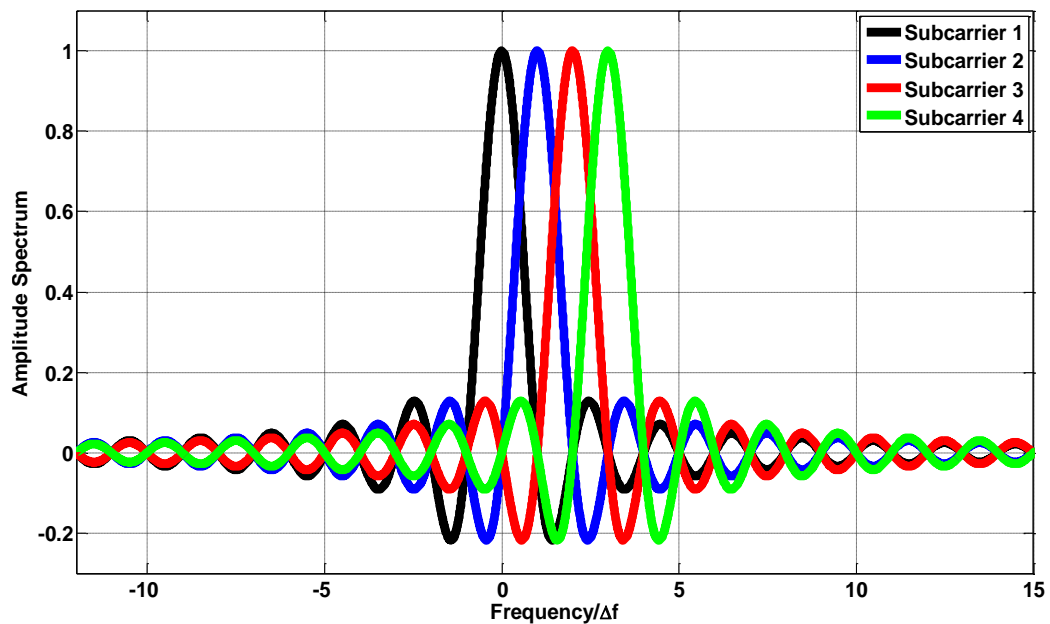
Figure 3.2. Arrangement of OFDM subcarriers

34

The arrangement of subcarriers in OFDM contrasts that in the conventional frequency division multiplexing (FDM), which does not use orthogonal subcarriers but rather separates neighbouring channels with a frequency guard in an attempt to eliminate ACI. Due to the use of frequency guards, an FDM system has a much lower spectral efficiency than an OFDM system for the same number of subcarriers.

Majority of subcarriers in an OFDM system are employed to carry user information by way of being modulated by data-bearing symbols. Some of the remaining subcarriers are reserved for transmitting training (pilot) symbols, while others are designated as null subcarriers. Each modulation symbol represents a sequence of information bits and could be from one of the commonly used digital modulation schemes, namely the binary phase-shift keying (BPSK), quadrature phase-shift keying (QPSK) or an M-ary quadrature amplitude modulation (M-QAM) scheme.

The OFDM baseband signal over one symbol duration is a superimposition of N modulated subcarrier signals and can be expressed as

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi k \Delta f t}, \qquad 0 \leq t \leq T \qquad (3.2)$$

where $X(k)$ is the kth modulation symbol, Δf is the subcarrier spacing and T is the symbol duration. The division by the term $\sqrt{N}$ ensures that both time and frequency domains have the same signal power.

Compared with a single-carrier system of the same symbol rate, the symbol duration of an OFDM system is N times longer. This implies that for the same transmission rate, if one symbol is transmitted during a duration T_s in the single-carrier system, then in the OFDM system, the symbol duration $T = NT_s$.

The use of a long symbol duration in OFDM helps to reduce intersymbol interferences (ISI) that is known to occur in a time-dispersive channel owing to a long delay spread caused by multipath propagation. In a multipath scenario, transmitted signals may take different radio paths from the transmitter to the receiver, which results in different transmission delays and frequency-selective fading. The *delay spread*, which by definition is the difference in propagation time between the longest and shortest path, is used to characterise a multipath transmission environment. In this characterisation, only the paths that have significant signal energies are considered.

Some signals in a multipath propagation environment may follow the line-of-sight

(LOS) path, while others may follow different non-line-of-sight (NLOS) paths. The NLOS paths are due to reflections by objects between the transmitter and receiver and have different radio propagation times as illustrated in Figure 3.3. In the figure, the propagation delays for the three NLOS paths are denoted by τ_1, τ_2 and τ_3. Assuming that the propagation delays are such that $\tau_1 < \tau_2 < \tau_3$, the *delay spread* of the multipath channel is then

$$\tau_m = \tau_3 - \tau_1 \tag{3.3}$$

The delay spread of a channel determines the *coherence bandwidth B_c*; an indicator of how quickly a transfer function changes with frequency. Specifically, the coherence bandwidth is the range of frequencies in which the channel response is constant and is approximately equal to the reciprocal of the delay spread, i.e. $B_c \cong 1/\tau_m$.

A channel in which the coherence bandwidth is much greater than the bandwidth of the input signal, or equivalently the delay spread is considerably less than the symbol duration, is referred to as a *flat-fading* channel. In other words, a flat-fading channel has a constant frequency response, i.e. it is frequency-non-selective. Such a channel can be fully characterised by a *single-tap* impulse response using a scaled Dirac-delta function as $h(t) = \beta\delta(t)$, where β is a scalar.

In case of a large number of obstructions between the transmitter and receiver, such that the delay spread is considerably greater than the symbol duration, the multipath channel becomes *frequency-selective* and the impulse response can be modelled as an n-tap Rayleigh-fading channel [90] using the equation
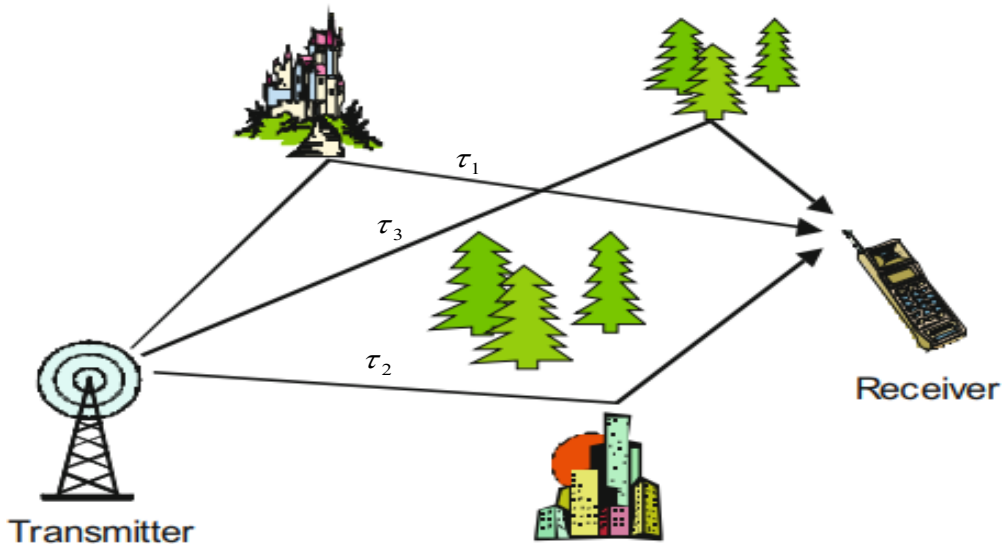
Figure 3.3. Multipath propagation scenario [91]

$$h(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} c_i(t) \tag{3.4}$$

where $c_i(t)$ denotes the complex coefficient of the ith tap given by

$$c_i(t) = \alpha_i(t)e^{-j\theta_i(t)} \tag{3.5}$$

In equation (3.5), $\alpha_i(t)$ and $\theta_i(t) = 2\pi f_c \tau_i$ are the attenuation and phase of the ith path, respectively, and f_c is the carrier frequency. The division by the term $\sqrt{n}$ normalizes the average channel power over the paths to unity.

The real and imaginary components of each tap weight $c_i(t)$ are independent and identically distributed Gaussian random variables with zero mean and variance σ^2 of 0.5. Consequently, the tap magnitudes $z = |\alpha_i(t)|$, $i = 1, 2, \dots, n$ are characterised by a Rayleigh random variable Z with a probability density function

$$p_Z(z) = \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}} \tag{3.6}$$

and the phases $\theta_i(t)$, $i = 1, 2, \dots, n$, have a uniform distribution over the range between 0 and 2π and they are also independent of one another.

In a typical design of an OFDM system for a wireless communication that takes into account multipath propagation and ensures that the channel is flat fading, the symbol duration is set sufficiently longer than the expected delay spread of the underlying environment, i.e.

$$T \gg \tau_m \tag{3.7}$$

or equivalently the subcarrier spacing is such that

$$\Delta f \ll B_c \tag{3.8}$$

A typical rule of thumb usually employed has $T \geq 4\tau_m$ [91].

In addition to the setting in equation (3.7), the orthogonality among the subcarriers is guaranteed if the relationship between the symbol duration and subcarrier spacing is governed by the equation

$$T = \frac{1}{\Delta f} \tag{3.9}$$

With this setting, the following minimum condition for mutual orthogonality among subcarriers is fulfilled over the duration of each OFDM symbol:

37

$$\int_0^T e^{j2\pi m\Delta ft} e^{-j2\pi n\Delta ft} dt = \begin{cases} T, & m = n \\ 0, & m \neq n \end{cases} \qquad (3.10)$$

where m and n are integers in the range 0 to $N - 1$.

In addition, it can be inferred from equations (3.1) and (3.8) that by selecting the number of subcarriers N such that

$$N \gg \frac{W}{B_c} \qquad (3.11)$$

renders the ISI to be very small.

However, the size of the subchannel bandwidth Δf is limited mainly by two system design constraints. The first one requires that Δf be sufficiently small so that the subchannel frequency response is essentially flat across each sub-band and the transmission rates are close to the channel capacity owing to high SNR. Because a flat frequency response corresponds to a scaled Dirac-delta impulse response in the time-domain, each subchannel can be modelled as a linear time-invariant (LTI) system.

Since the input signal to a given subchannel is a complex exponential, which happens to be an eigenfunction of a LTI system, the output signal is simply the input scaled by the frequency response of the channel, i.e.

$$y_k(t) = H_k(f_k)X(k)e^{j2\pi f_k t} \qquad (3.12)$$

where $f_k = k\Delta f$ is the kth subcarrier frequency and $H_k(f_k)$ is the frequency response of the kth subchannel evaluated at f_k. The subchannel frequency response $H_k(f_k)$ is in general a complex quantity and can be known by initially transmitting a pilot signal consisting of a known modulation symbol or an un-modulated subcarrier and observing the received signal.

From equation (3.12), it is clear that the transmission over the radio channel changes only the amplitude and phase of the modulation symbols. In other words, the transmission does not affect subcarrier frequencies. This implies that the mutual orthogonality between subchannels is maintained all through to the receiver as long as the subchannels are linear and time-invariant and have constant frequency responses.

The second constraint on the subcarrier spacing is imposed by the expected Doppler spread that may exist in the case of a time-variant radio channel when there is a relative motion between the transmitter and the receiver. The relative motion causes a difference in frequencies between the transmitted one and the received one. The received

frequency can be found as

$$f_r = \left(1 \pm \frac{v}{c}\right) f_c = f_c + f_D \tag{3.13}$$

where v is the relative velocity of the receiver with respect to the transmitter, c is the speed of light, f_c is the transmitted frequency and the differential frequency f_D, referred to as the *Doppler shift*, is given by

$$f_D = \pm \frac{v}{c} f_c \tag{3.14}$$

The Doppler frequency shift may lead to an increase or a decrease in a subcarrier frequency and this can likely result into inter-carrier interferences (ICI) and a loss of orthogonality between subcarriers. In order to maintain the ICI at a tolerable level, the subcarrier spacing in OFDM is set considerably greater than the maximum expected Doppler shift, also referred to as *Doppler spread*, denoted here by f_{Dm}, i.e.

$$\Delta f \gg f_{Dm} \tag{3.15}$$

with typical values of $\Delta f \geq 33 f_{Dm}$ [91].

The Doppler power spectrum, which is the statistical variation of the channel power with respect to frequency, occurs over a bandwidth equal to $2f_{Dm}$ i.e. from $-f_{Dm}$ to $+f_{Dm}$. Another important quantity related to the *Doppler shift* is the *coherence time, T_c*, which indicates how quickly an impulse response of a channel changes over time. The coherence time T_c is inversely proportional to f_{Dm} and is simply the period over which the impulse response remains constant.

Depending on the magnitude of the coherence time in relation to the symbol duration, a channel can be said to be *slow* or *fast fading*. A slow-fading channel has $T \ll T_c$, while a fast-fading one has $T \gg T_c$. From equations (3.9) and (3.15), it is clear that an OFDM channel has $T \ll T_c$, thus has slow-fading characteristics. This implies that the impulse responses of subchannels remain literally unchanged during the symbol period.

3.2 Implementation of OFDM Using IFFT/FFT

In practical OFDM systems, signals are digitally processed using FFT algorithms. The sampling of the OFDM baseband signal is typically done at a rate greater than the Nyquist rate. Given that the frequency components of the baseband signal in equation

(3.2) are in the range $-0.5N\Delta f$ to $+0.5N\Delta f$, the Nyquist sampling rate is $N\Delta f$ and a signal sampled at this rate can be expressed as

$$x(nT_s) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k)e^{j2\pi k\Delta f nT_s}, \qquad 0 \le n \le N-1 \tag{3.16}$$

where $T_s = 1/N\Delta f$ is the sampling time interval. On substituting for T_s in equation (3.16), the sampled signal becomes

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k)e^{\frac{j2\pi kn}{N}}, \qquad 0 \le n \le N-1 \tag{3.17}$$

Equation (3.17) is equivalent to the IDFT of $X(k)$ and can thus be evaluated by the computationally proficient FFT algorithms.

In case the sampling is at a frequency greater than the Nyquist rate, say by a factor $\mathcal{L}$, the oversampled signal can be expressed as

$$x_s(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N\mathcal{L}-1} \acute{X}(k)e^{\frac{j2\pi km}{N\mathcal{L}}}, \qquad 0 \le m \le N\mathcal{L}-1 \tag{3.18}$$

where $\acute{X}(k)$ is obtained by zero-padding $X(k)$ with $N(\mathcal{L}-1)$ zeros as follows:

$$\acute{X}(k) = \begin{cases} X(k), & 0 \le k \le N-1 \\ 0, & N \le k \le N\mathcal{L}-1 \end{cases} \tag{3.19}$$

The IFFT can also be used to generate the oversampled discrete-time signal in equation (3.18). However, the $N(\mathcal{L}-1)$ zeros should be inserted in the midpoint of $X(k)$ rather than at the end since the first half of the samples relates to the positive frequencies of the IFFT and the last half to the negative frequencies. Therefore, $\acute{X}(k)$ should be formed as follows:

$$\acute{X}(k) = \begin{cases} X(k), & 0 \le k \le \dfrac{N}{2}-1 \\ 0, & \dfrac{N}{2} \le k \le N\mathcal{L}-\dfrac{N}{2}-1 \\ X(k), & N\mathcal{L}-\dfrac{N}{2} \le k \le N\mathcal{L}-1 \end{cases} \tag{3.20}$$

and in this case the oversampled signal is given by

$$x_s(m) = \mathcal{L} \times \text{IFFT}\{\acute{X}(k)\} \tag{3.21}$$

In addition, the oversampling operation can be implemented directly in the time

domain by interpolation, which is done in two steps as shown in Figure 3.4. The input signal is sampled at the Nyquist rate f_s. The interpolation starts with the insertion of $(\mathcal{L} - 1)$ zeros between the samples of $x(n)$ to obtain $x_u(m)$ followed by a low-pass filtering operation to generate the oversampled signal $x_s(m)$ containing the original samples of $x(n)$ and $(\mathcal{L} - 1)$ interpolated samples in between them. The two most common impulse responses of the low pass filter are the rectangular and root-raised cosine pulses.

3.3 Cyclic Prefix

Although the use of long symbol duration greatly reduces ISI, it may not be fully eliminated. A guard time interval between consecutive OFDM symbols is still needed to eradicate ISI. The guard interval is inserted between consecutive symbols in the form of a *cyclic prefix* (CP) or through padding with either zeros or known samples. However, the insertion of CP is the most common and involves a cyclic extension of an OFDM symbol through copying of a few tail-end samples to the front as demonstrated in Figure 3.5.

By denoting the length of the guard interval by T_{gi}, the effective duration of an OFDM symbol after the insertion of a CP is

$$T_{eff} = T_{gi} + T \tag{3.22}$$

The length of T_{gi} depends on the number of samples spanned by the delay spread of a channel. For example, if τ_m spans l samples, the CP is made up of $\{x_{N-l}, x_{N-l+1}, \dots, x_{N-1}\}$ samples from an OFDM block of N samples $\{x_0, x_1, \dots, x_{N-1}\}$ as in equation (3.17). The guard interval is found as

$$T_{gi} = \frac{l}{N} T \tag{3.23}$$
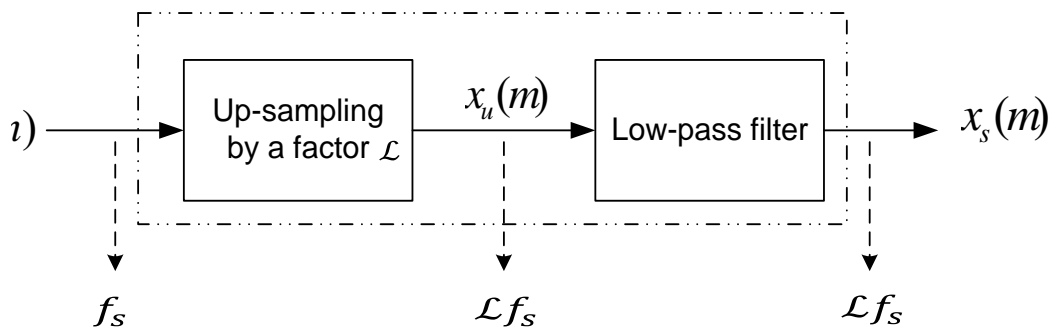
Figure 3.4. Block diagram of oversampling using $\mathcal{L}$-times interpolator

41
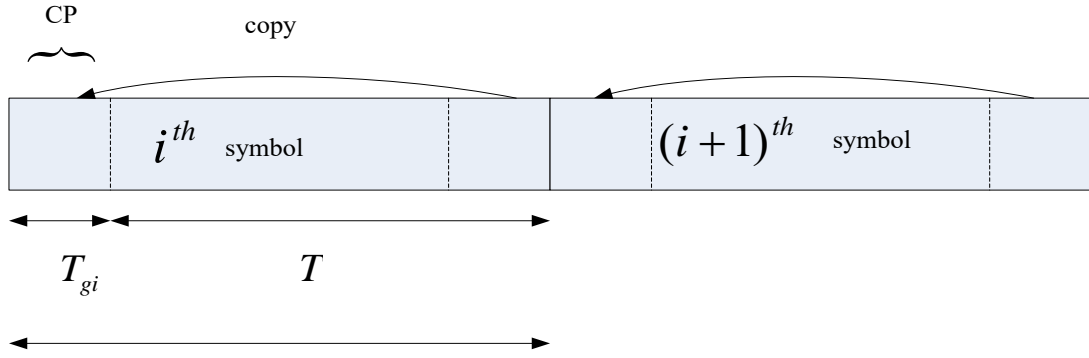
Figure 3.5. ODFM symbol extension by cyclic prefix

Once the CP is inserted, the new OFDM block consists of $N + l$ samples, i.e. $\{x_{N-l}, x_{N-l+1}, \dots, x_{N-1}, x_0, x_1, \dots, x_{N-1}\}$. These samples are sent in series, one at time, within the duration T_{eff}. In wireless applications, T_{gi} is typically set to $1/4T$, thus making the CP to contain a quarter of the total number of samples. In a typical implementation, T_{gi} is set greater than τ_m to confine the ISI of the current symbol within the guard interval of the next symbol in order not to affect the FFT processing taken over the duration T.

The CP is directly removed at the receiver after a time synchronization procedure prior to the demodulation of the received signal. The use of long symbol duration makes the radio channel a LTI system within a period of a few OFDM symbols. This in turn ensures that the orthogonality of all subcarriers is completely preserved in the receiver even under frequency-selective radio channels. Thus, the CP helps to combat both ISI and ICI when the channel is non-ideal. The maintenance of orthogonality between subcarriers along the entire transmission path allows the modulated signals to be split directly into different data symbols using FFT and a simple single-tap equalization applied on each subchannel.

Although the addition of CP to an OFDM symbol comes with the above highlighted benefits, it has two drawbacks. The first one is the reduction in SNR, which typically requires a boost of the transmitted signal energy. The SNR reduction can be explained from the consideration that the bit energy E_b represents the energy in the useful bits, while the symbol energy E_s is spread across the entire extended OFDM symbol.

Assuming b bits per symbol, the total bit energy is NbE_b and the total symbol energy before cyclic extension is NE_s. Since energy is conserved, the total symbol energy

is equal to the total bit energy, i.e. $NE_s = NbE_b$ and thus $E_s = bE_b$, which gives the SNR before the insertion of CP as

$$\frac{E_s}{N_0} = \frac{bE_b}{N_0} \tag{3.24}$$

where N_0 is the noise power spectral density.

When the OFDM symbol is extended by CP to length $(N + l)$, the total bit energy remains the same, while the total symbol energy becomes $(N + l)E_s$ [92]. Because of the energy conservation principle, $(N + l)E_s = NbE_b$ and the symbol energy becomes

$$E_s = \frac{N}{N + l} bE_b \tag{3.25}$$

while the overall SNR after the addition of CP is

$$\frac{E_s}{N_0} = \frac{N}{N + l} \frac{bE_b}{N_0} \tag{3.26}$$

From equations (3.24) and (3.26), the reduction in SNR from the original value due to the CP is given by

$$\begin{aligned} SNR_{red} &= \frac{bE_b}{N_0} - \frac{N}{N + l} \frac{bE_b}{N_0} \\ &= \frac{l}{N + l} \frac{bE_b}{N_0} \end{aligned} \tag{3.27}$$

This SNR reduction is due to the fraction of the total power allocated for the transmission of CP and hence not available for communicating user data. Thus, after CP extension, the transmitted symbol energy, in equation (3.25), needs to be boosted by multiplying it by a factor $N/(N + l)$ to restore it to its original value of bE_b.

Secondly, the addition of a CP causes a loss in data rate. Without the CP, N modulation symbols are transmitted within the duration T, i.e. at a symbol rate equal to N/T. However, with the CP, the N modulation symbols are transmitted over the extended symbol time T_{eff}, hence at a rate of $N/(T + T_{gi})$. Thus, there is a reduction in data rate given by

$$\begin{aligned} R_{loss} &= \frac{N}{T} - \frac{N}{T + T_{gi}} \\ &= \frac{T_{gi}}{(T + T_{gi})} \frac{N}{T} \end{aligned} \tag{3.28}$$

By substituting for T_{gi} from equation (3.23) in (3.28), the data-rate loss becomes

$$R_{loss} = \frac{l}{(N+l)} \frac{N}{T} \qquad (3.29)$$

Considering equations (3.27) and (3.29), both the SNR reduction and data-rate loss can be made small by selecting N and l such that $l \ll N$.

3.4 Guard Band

The combination of many frequency-shifted *sinc* waveforms in an OFDM signal, as was illustrated in Figure 3.2, leads to a large out-of-band (OOB) power and the occurrence of ACI. Thus in an OFDM system, a guard frequency band is necessary to mitigate the ACI outside the allocated bandwidth. One practical way to completely reduce the OOB emissions and hence combat the ACI is to employ a raised-cosine pulse shaping, rather than a rectangular windowing, in combination with a number of null or virtual subcarriers at the two edges of the transmission band [93]. The null or virtual subcarriers are what constitute the guard band.

However, the OOB power is dependent on the value of the roll-off parameter of the raised-cosine pulse employed. As the value of the roll-off increases from 0 to the maximum value of 1, the OOB radiations are reduced. However, a large roll-off factor results in a further increase in the guard interval. Therefore, a trade-off is needed between the reduction in OOB radiations and the elongation of the guard interval.

Moreover, virtual subcarriers do not carry any information but only offer a band for the OFDM spectrum to decay smoothly to zero at the symbol boundary. Although these unused subcarriers require no additional processing, they should be limited to a small number because they reduce the overall spectral efficiency of the system by a factor N_u/N, where N_u is the number of data-bearing subcarriers and N is the total number of subcarriers in the system.

3.5 OFDM Signal Processing Blocks

In Figure 3.6, the main signal processing blocks in an OFDM system from the transmitter to the receiver are shown. The upper path represents the transmitter, while the lower one is the receiver. The binary input data is first encoded through addition of a forward error-correction code. The encoded binary data is then interleaved and/or

punctured before being mapped onto BPSK or M-QAM symbols. This is followed by the insertion of pilot symbols and serial-to-parallel (S/P) conversion.

After the mapping of data and pilot symbols, the IFFT is applied on the symbols to generate discrete-time samples. This sequence of samples is then cyclically extended and a raised-cosine windowing applied to make the signal waveform transit smoothly to zero at the symbol boundary. The resulting samples are then converted into a continuous-time signal followed by carrier modulation at radio frequencies, power-amplification and lastly the signal is sent to the antenna for RF transmission.

In the receiver, reverse operations to those in the transmitter are performed. The received RF signal is first down-converted to baseband frequencies and then to a discrete-time signal by the ADC block. This is followed by digital signal processing, which starts with a training phase to determine symbol timing and frequency offset. The cyclic prefix is then removed and FFT performed to recover transmitted modulation symbols. The recovered symbols are corrected using a single-tap equalizer and then mapped into bits, which are then de-interleaved and decoded to produce a binary data stream.

Two important operations of the receiver are the recovery of the transmitted symbols via FFT and channel equalization, which are now briefly explained. Using equation (3.12) and assuming sampling at Nyquist rate, the received signal is given by

$$y(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{\infty} H_k(f_k)X(k)e^{\frac{j2\pi kn}{N}} + z(n) \qquad (3.30)$$

The term $z(n)$ represents additive noise from the transmission channel.
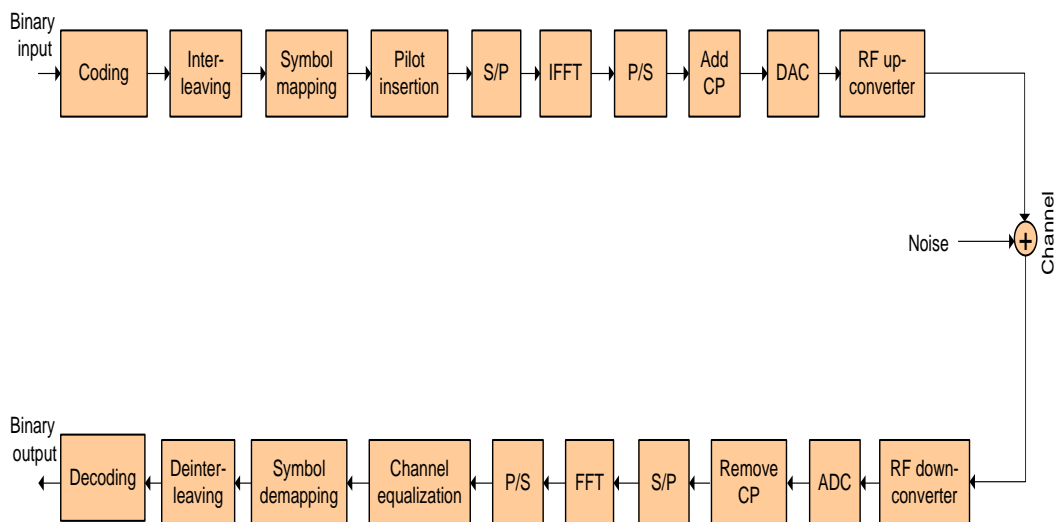
Figure 3.6. Block diagram of OFDM system

The discrete-time signal in equation (3.30) can be considered as the actual output of the block labelled "Remove CP" in Figure 3.6. On applying FFT on $y(n)$ after parallel-to-serial (P/S) conversion, the received symbol from each subchannel is given by

$$Y(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \left(\frac{1}{\sqrt{N}} \sum_{m=0}^{\infty} H_m(f_m) X(m) e^{\frac{j2\pi nm}{N}} + z(n) \right) e^{-\frac{j2\pi kn}{N}} \qquad (3.31)$$

This equation can be re-written as follows:

$$Y(k) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\sum_{m=0}^{\infty} H_m(f_m) X(m) e^{\frac{j2\pi n(m-k)}{N}} \right) + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z(n) e^{-\frac{j2\pi kn}{N}} \qquad (3.32)$$

Noting that in the first term of this equation, we have

$$\sum_{n=0}^{N-1} \left(\sum_{m=0}^{\infty} H_m(f_m) X(m) e^{\frac{j2\pi n(m-k)}{N}} \right) = \begin{cases} N H_m(f_m) X(m), & m = k \\ 0 & , & m \neq k \end{cases} \qquad (3.33)$$

the received symbols are obtained as follows:

$$Y(k) = H_k(f_k) X(k) + Z(k), \qquad 0 \leq k \leq N - 1 \qquad (3.34)$$

where $Z(k)$ denote the DFT of the noise signal $z(n)$.

Single-tap equalization is then performed on each subchannel by dividing the received symbol by the frequency response to obtain an estimate of the transmitted symbol

$$\tilde{Y}(k) = \frac{Y(k)}{H_k(f_k)} = X(k) + \frac{Z(k)}{H_k(f_k)}, \qquad 0 \leq k \leq N - 1 \qquad (3.35)$$

Using the constellation diagram of the modulation scheme employed during the mapping operation in the transmitter, the actual transmitted symbol is chosen as the one nearest to an estimated symbol.

3.6 Peak-to-Average Power Ratio

The peak-to-average power ratio (PAPR) measures the amplitude fluctuations of a signal in terms of how much the maximum instantaneous power exceeds the average power. The PAPR for an analogue signal is defined as

$$\text{PAPR}\{x(t)\} = \frac{\max\{|x(t)|^2\}}{\text{E}\{|x(t)|^2\}} \qquad (3.36)$$

where E{.} is the expectation operator, which gives the average value of the argument, i.e.

$$E(|x(t)|^2) = \lim_{T \to \infty} \frac{1}{T} \int_0^T |x(t)|^2 \, dt \tag{3.37}$$

If the statistical distribution of signal amplitudes is known, it is usually possible to know the expected amount of PAPR of that signal. For the OFDM signal in equation (3.2), the N complex modulated signals can be considered statistically independent and identically distributed. The signal can be written as a sum of real and imaginary signals as follows:

$$x(t) = x_R(t) + jx_I(t) \tag{3.38}$$

where the amplitudes of each of the two components are Gaussian distributed with a mean of zero and a variance of 0.5. Thus, the signal envelope

$$|x(t)| = \sqrt{|x_R(t)|^2 + |x_I(t)|^2} \tag{3.39}$$

has Rayleigh-distributed amplitudes when N is large in accordance with the central limit theorem [94]. As an example, the normal and Rayleigh distributions of the amplitudes of the real, imaginary and envelope of an OFDM signal with 16 subcarriers are shown Figure 3.7.

The Rayleigh distribution of the amplitudes of the envelope $|x(t)|$ implies that some of them could be well above the mean value as shown in Figure 3.8. In the figure, the real and imaginary signal parts of the 16 modulated signals are shown in black, while their respective combined signals are in blue colour. The signal envelope at the bottom of the figure has some peaks that are greater than the average value indicated by the red horizontal line and this occurs at the time instants when the modulated signals add constructively.

Although it is very appropriate to compute PAPR from the continuous-time signal, owing to the use of digital signal processing, it is in practice obtained from the oversampled discrete-time signal in equation (3.18) as follows:

$$PAPR\{x_s(n)\} = \frac{\max_{0 \le n \le N\mathcal{L}-1} \{|x_s(n)|^2\}}{E\{|x_s(n)|^2\}} \tag{3.40}$$

However, a sufficiently high oversampling factor typically greater than 4, i.e. $\mathcal{L} \ge 4$ [95], is needed in order have the discrete-time PAPR closely estimate the

continuous one given by equation (3.36). The high oversampling factor ensures that the oversampled signal has as many samples as possible so that the peak amplitude of the continuous-time signal is present in the discrete-time signal.
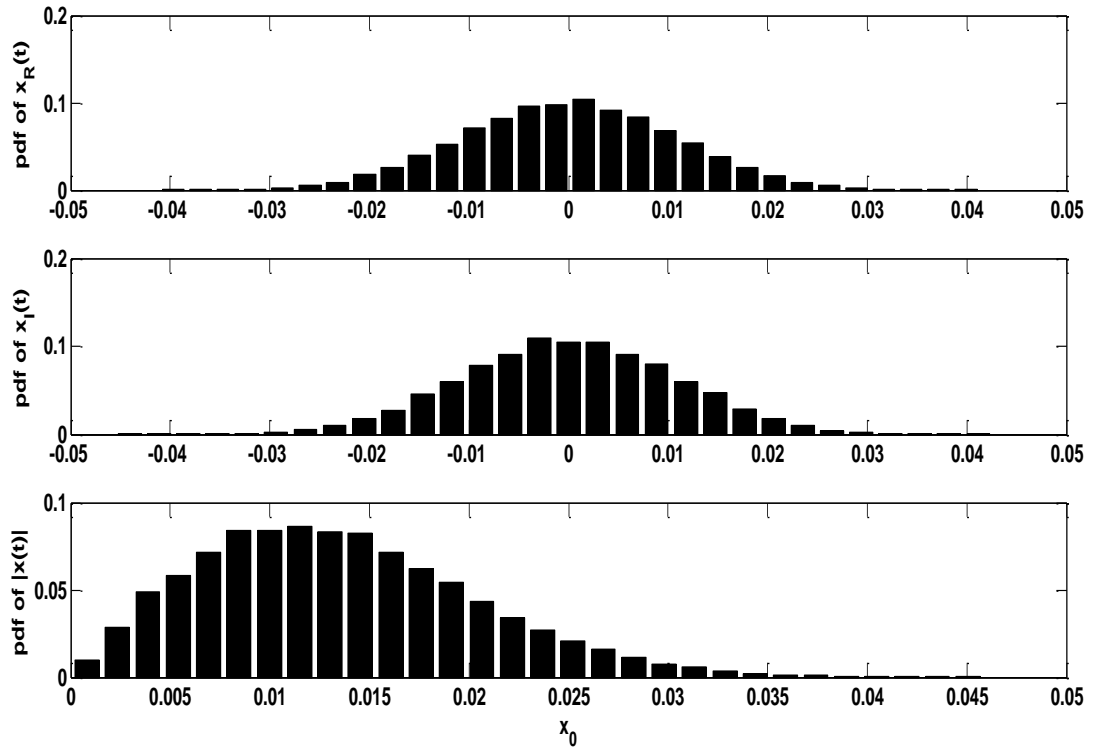
Figure 3.7. Distribution of OFDM signal amplitudes with $N = 16$ subcarriers
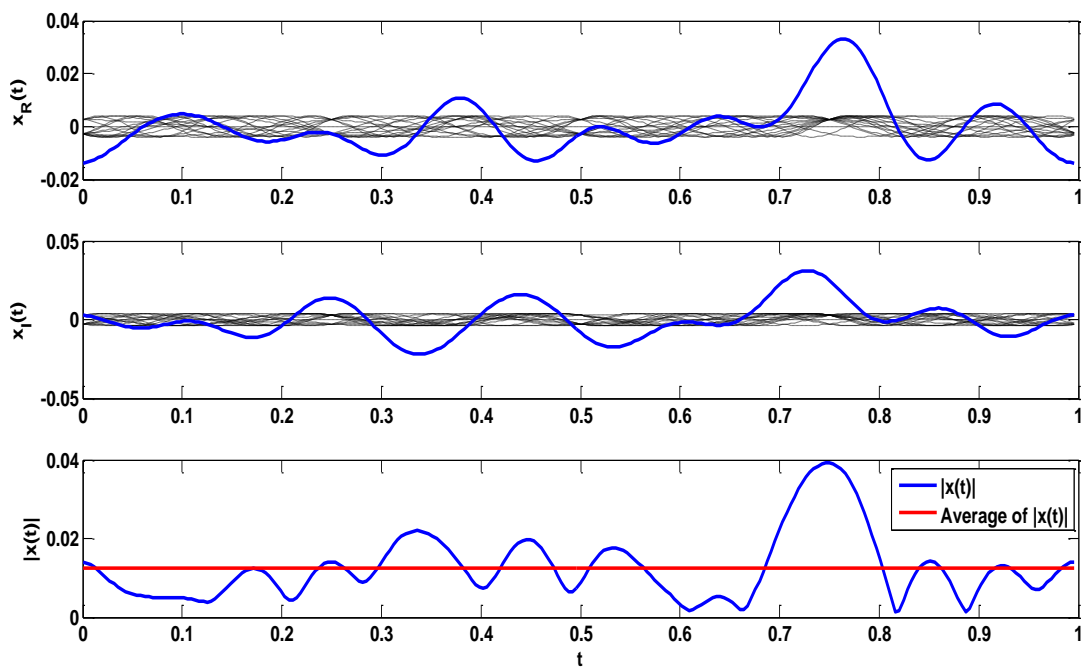
Figure 3.8. Envelope of an OFDM signal with $N = 16$ subcarriers

3.6.1 Complementary Cumulative Distribution Function

The magnitude of the PAPR can be measured by complementary cumulative distribution function (CCDF), which gives the probability of the PAPR being greater than a specified value. This CCDF for a discrete-time signal can be found conveniently by considering that a discrete-time signal $x(n)$ can be decomposed into its real and the imaginary parts by expressing it as $x(n) = x_R(n) + jx_I(n)$.

The absolute amplitudes of the signal are

$$|x(n)| = \sqrt{x_R^2(n) + x_I^2(n)} \qquad (3.41)$$

They are independent and identically Rayleigh distributed, regardless of the underlying distribution of the real and imaginary parts. However, for the case of OFDM signals, both the real and imaginary components are Gaussian distributed.

The CCDF of $|x(n)|$ can be derived as follows. Let X be the Rayleigh random variable of the distribution of the signal amplitudes $|x(n)|$. Then, the probability density function (pdf) of X is given by

$$f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \qquad (3.42)$$

where σ^2 is the variance of the Gaussian distribution of the real and imaginary parts of $x(n)$ and x is a variable for the values of the signal amplitudes $|x(n)|$.

From equation (3.42), the cumulative distribution function (CDF) of X can be found as follows:

$$F_X(x) = \Pr(X \leq x)$$
$$= \int_0^x \frac{u}{\sigma^2} e^{-\frac{u^2}{2\sigma^2}} \, du \qquad (3.43)$$

where $\Pr(X \leq x)$ denotes the probability that X is less than or equal to x. By changing the variable u through letting $y = u^2/(2\sigma^2)$, the CDF becomes

$$F_X(x) = \int_0^{\frac{x^2}{2\sigma^2}} e^{-y} dy$$
$$= 1 - e^{-\frac{x^2}{2\sigma^2}} \qquad (3.44)$$

The maximum value that the random variable X can have is referred to as the

49

crest factor and is simply the square root of the PAPR. Denoting the crest factor by x_{cf} and considering that all possible values X can take are less than or equal to x_{cf}, the CDF $\Pr(X \le x_{cf})$ can be expressed as

$$\Pr(X \le x_{cf}) = \prod_{i=0}^{N-1} \Pr(X_i \le x_{cf}) = \left(1 - e^{-\frac{x^2}{2\sigma^2}}\right)^N \tag{3.45}$$

The magnitude of the crest factor can be measured by the complementary CDF (CCDF), which is the probability of the random variable exceeding the crest factor and is given by

$$\Pr(X > x_{cf}) = 1 - \left(1 - e^{-\frac{x^2}{2\sigma^2}}\right)^N \tag{3.46}$$

Considering that $x^2/(2\sigma^2)$ is a ratio of the power of signal x over the average power, equation (3.46) can be rewritten in a form that allows for the direct assessment of the magnitude of PAPR as follows:

$$\Pr\{\text{PAPR}(x(n)) > \gamma\} = 1 - (1 - e^{-\gamma})^N \tag{3.47}$$

Here, $\gamma = x^2/(2\sigma^2)$ is the threshold value against which the comparison is done to check the highness of PAPR. This equation also shows that the CCDF, and hence the PAPR, increases with the number of subcarriers employed in an OFDM system. To further demonstrate this, examples of CCDFs for N = 64, 128, 256, 512 and 1024 based on equation (3.47) are plotted in Figure 3.9 together with the simulated CCDFs.
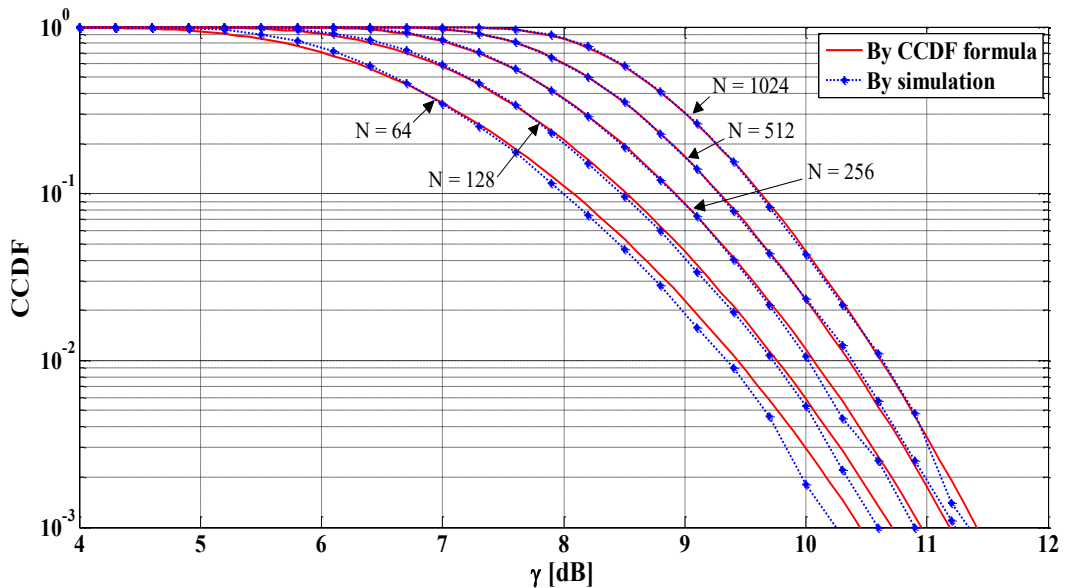
Figure 3.9. CCDF of OFDM signals with different number of subcarriers

Additionally, equation (3.47) shows that if the CCDF value on the left-hand side is kept constant, then for a fixed number of subcarriers, a large value of γ points to a higher PAPR and vice versa. This in turn implies that the difference between any two values of γ at the same value of CCDF may be used to measure the reduction in PAPR and to evaluate the performance of a proposed method.

3.6.2 The Upper Bound on PAPR

The statistical analysis of OFDM signal presented in Subsection 3.6.1 has established that the summation of many modulated subcarrier signals as given by equation (3.17) is likely to increase the PAPR. The increase was found to be due to constructive addition of signals and is proportional to the number of subcarriers.

Following the preceding analysis, an upper limit or the maximum possible PAPR can be determined for an OFDM signal by considering a case where all subcarriers are modulated by the same data symbol, i.e. $X(k) = X_d$ for $0 \leq k \leq N - 1$. The peak power of the corresponding time-domain signal $x(n)$ is given by

$$
\begin{aligned}
\max_{0 \leq n \leq N-1} (|x(n)|^2) &= \max_{0 \leq n \leq N-1} \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{\frac{j2\pi kn}{N}} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X^*(k) e^{\frac{-j2\pi kn}{N}} \right) \\
&= \max_{0 \leq n \leq N-1} \left(\frac{1}{N} \sum_{k=0}^{N-1} \sum_{k=0}^{N-1} |X_d|^2 \right) \\
&= N|X_d|^2
\end{aligned}
\tag{3.48}
$$

Similarly, the average power of the time-domain signal is given by

$$
\begin{aligned}
\mathrm{E}(|x(n)|^2) &= \mathrm{E} \left(\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{\frac{j2\pi kn}{N}} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X^*(k) e^{\frac{-j2\pi kn}{N}} \right) \\
&= |X_d|^2
\end{aligned}
\tag{3.49}
$$

Therefore, from equations (3.48) and (3.49), the upper limit of PAPR is

$$
\mathrm{PAPR}_{ul} = \frac{N|X_d|^2}{|X_d|^2} = N
\tag{3.50}
$$

This again shows that the maximum possible PAPR in an OFDM system is proportional to the number of subcarriers.

3.7 PAPR of passband OFDM Signal

The passband OFDM signal can be easily related to the baseband signal given in equation (3.2). This is because the passband signal is obtained through amplitude modulation of a carrier signal, at RF frequency f_c, with the baseband signal, i.e.

$$s(t) = \text{Re}\{x(t)e^{j2\pi f_c t}\}$$
$$= x_R(t)\cos(2\pi f_c t) - x_I(t)\sin(2\pi f_c t) \tag{3.51}$$

This signal can also be expressed in the form

$$s(t) = |x(t)|\cos(2\pi f_c t + \theta(t)) \tag{3.52}$$

where $x(t) = \sqrt{(x_R^2(t) + x_I^2(t))}$ is the signal envelope and $\theta(t) = \tan^{-1}(x_I(t)/x_R(t))$ is the phase angle.

The envelope in equation (3.52) is still Rayleigh distributed and the phase is uniformly distributed over the range 0 to 2π. It can be inferred from equation (3.52) that the RF passband signal has the same envelope as the baseband signal but is of half its average power due to the multiplication by the cosine function. This reduction of the average power by a half makes the PAPR of the passband signal to be twice the baseband PAPR.

3.8 Power Amplification

An OFDM signal is an amplitude-modulated signal that has a non-constant envelope. This requires the signal to be linearly amplified using linear power amplifiers, typically class A, B or AB amplifiers. These power amplifiers are operated as current-dependent sources in the active region where the output signal varies linearly with the input. Considering a sinusoidal input to the three classes of power amplifiers, the conduction angles, which are determined by the location of the quiescent point, are 360°, 180° and between 180° and 360° for the class A, class B and class AB power amplifiers, respectively.

In terms of linearity, class A power amplifier is the most linear of the three classes and gives an output waveform similar to the input, while class B is the least linear and its output has crossover effects and harmonic distortions. The linearity performance of class AB is in between that of class A and class B and the output waveform has barely noticeable distortions compared to the input waveform.

However, in terms of power efficiency, class A is the worst with a maximum efficiency of 50%, while class B has the best maximum efficiency of 78.5%. The power efficiency of Class AB is between 50% and 78.5%.

Although, the three classes are operated in the linear region, they may be driven by large signal amplitudes that, depending on the quiescent point of the underlying power amplifier, may lead to nonlinear amplification of the input signal. This section considers the different signal distortions that arise from nonlinear amplification of a signal and how they are measured.

3.8.1 Nonlinear Amplification and Distortion Measurements

A high power amplifier (HPA) as a nonlinear device can be modelled as shown in Figure 3.10. The baseband nonlinearity of the device can be described by the equation

$$g\{x(t)\} = G\{|x(t)|\}e^{j(\theta(t)+\Psi\{|x(t)|\})} \tag{3.53}$$

Here, functions $G\{|x(t)|\}$ and $\Psi\{|x(t)|\}$ are the AM/AM and AM/PM characteristics that operate on the input signal amplitude to produce the amplitude and phase angle characteristics in the output, respectively.

Due to different AM/AM and AM/PM characteristics, several HPA models exist. The models are either for the travelling wave tube amplifier (TWTA) or for the solid-state power amplifier (SSPA). The TWTA is essentially used in microwave receivers as a low noise RF amplifier and in amplifying wideband signals especially in satellite communication, while the SSPA is used with narrowband signals in wireless communication. The two types of amplifiers are modelled in different ways.

The most common model for the TWTA is the Saleh model given in [96], while the SSPA is usually modelled using Ghorbani [97] or Rapp [98] models. However, the Rapp model is the more commonly used model for the SSPA owing to its simplicity and as such, it will be used in the simulation work in this thesis.

$$x(t)$$

$$=|x(t)|e^{j\theta(t)}$$

HPA
$$g\{.\}$$

$$g\{x(t)\}$$

$$= G\{|x(t)|\}e^{j(\theta(t)+\Psi\{|x(t)|\})}$$
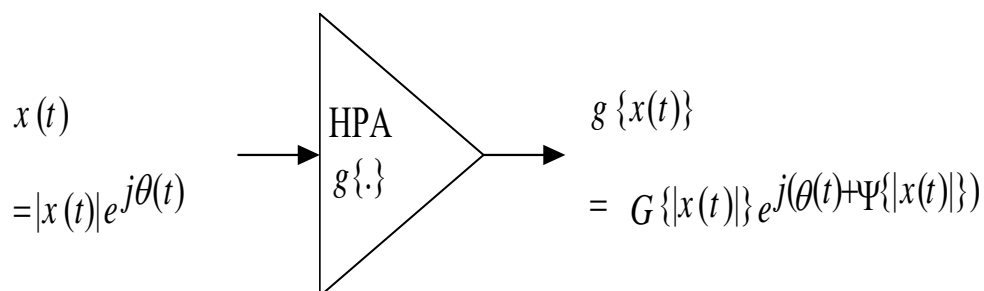
Figure 3.10. General HPA model

The Rapp model for a SSPA has the following AM/AM and AM/PM characteristics:

$$G\{|x(t)|\} = \frac{|x(t)|}{\left(1 + \left(\frac{|x(t)|}{A_{sat}}\right)^{2p}\right)^{\frac{1}{2p}}} \tag{3.54}$$

$$\Psi\{|x(t)|\} = 0$$

where p is a smoothness parameter and A_{sat} is the output saturation level of the HPA at the 3-dB point.

Typical values of the smoothness parameter p are in the range between 2 and 3. The parameter is employed to render smooth amplification changeover from the linear to the saturation region. Thus, the smaller the parameter is, the smoother is the transition. The Rapp model does not introduce phase distortion; hence, its output signal is simply

$$g\{x(t)\} = G\{|x(t)|\}e^{j\theta(t)}$$

$$= \frac{|x(t)|}{\left(1 + \left(\frac{|x(t)|}{A_{sat}}\right)^{2p}\right)^{\frac{1}{2p}}} e^{j\theta(t)} \tag{3.55}$$

A typical AM/AM characteristic curve for a Class A power amplifier is illustrated in Figure 3.11. For maximum output power and efficiency, it is desired that the HPA be operated near the saturation point, typically defined at the 1-dB compression point. However, at that point, the gain of the HPA is reduced and the input is nonlinearly amplified, which results in signal distortions in the output of the device.

There are mainly two types of nonlinear distortions, namely harmonic and intermodulation distortions. Harmonic distortion refers to the case where the output has new signal components at frequencies, which are integer multiples of the fundamental frequencies in the input signal. This distortion can occur regardless of whether the input signal consists of a single tone or multiple tones.

On the other hand, intermodulation (IM) products can only occur when the input signal contains multiple tones. When different frequency components are mixed in the power amplifier, they give rise to new frequency components in the output of the device consisting of sums and differences of the frequencies in the input.

To analyse nonlinear distortions from a power amplifier, the output is commonly modelled using a Taylor series
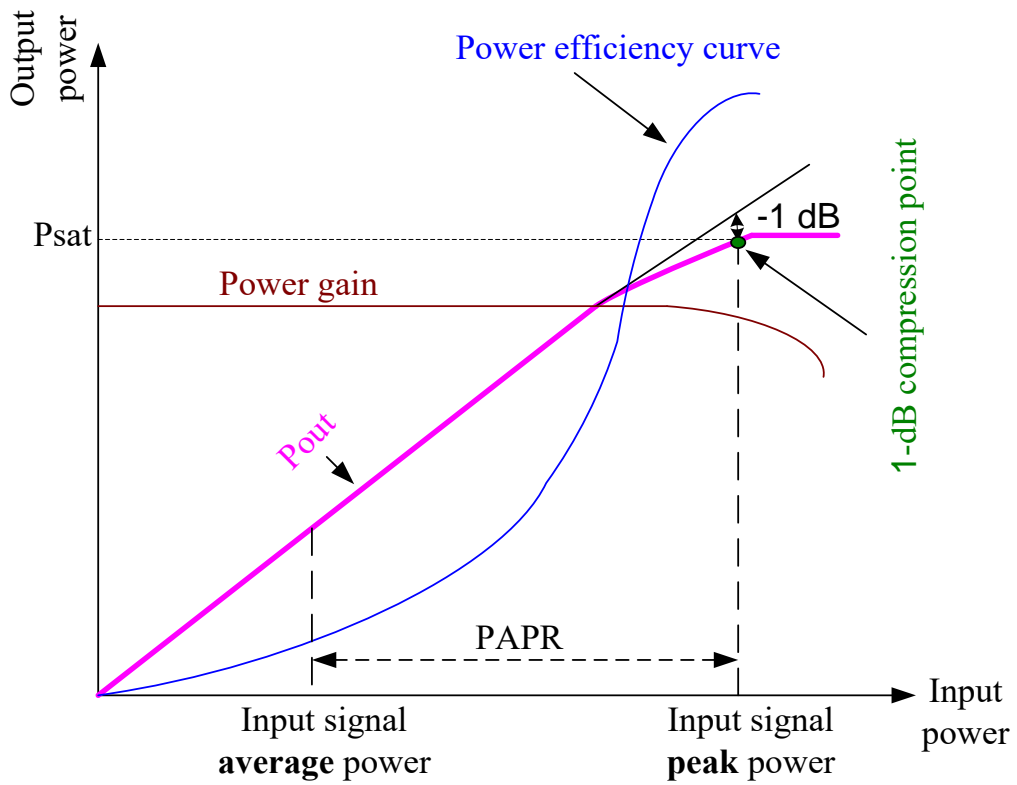
Figure 3.11. AM/AM characteristic of Class A power amplifier [99]

$$y(t) = \sum_{k=1}^{\infty} a_k \big(x(t)\big)^k \tag{3.56}$$

where a_k are the coefficients of the series and $x(t)$ is the input signal.

The input to the amplifier is taken to be a two-tone test signal $x(t) = A_1 cos(\omega_1 t) + A_2 cos(\omega_2 t)$. Using this signal, the frequencies of the harmonic and IM products in the output signal can be expressed as

$$\omega = m\omega_1 \pm n\omega_2 \tag{3.57}$$

where m and n are integers in the range from 0 to ∞.

If either m or $n = 0$ in equation (3.57), there is a harmonic distortion component otherwise it is an IM term. The IM distortion terms are commonly referred to by their order, typically specified as i^{th}-order IM, where i is the sum of the values of m and n, i.e.

$$i = m + n \tag{3.58}$$

To simply finding the output signal from a nonlinear amplifier, it is assumed that $A_1 = A_2 = A$ and that $a_3 \ll a_1$. The output signal from equation (3.56) up to the third

55

power can obtained as

$$
\begin{aligned}
y(t) \approx\ & a_1 A[\cos(\omega_1 t) + \cos(\omega_2 t)] \\
& + a_2 A^2 \big[\cos\big((\omega_1 - \omega_2)t\big) + cos\big((\omega_1 + \omega_2)t\big)\big] \\
& + \frac{3}{4} a_3 A^3 \big[\cos\big((2\omega_1 - \omega_2)t\big) + cos\big((2\omega_1 + \omega_2)t\big) \\
& + \cos\big((2\omega_2 - \omega_1)t\big) + cos\big((2\omega_2 + \omega_1)t\big)\big] + \cdots
\end{aligned}
\tag{3.59}
$$

Equation (3.59) reveals that the amplitude of an ith-order IM component is proportional to the ith power of the amplitude of the fundamental component. This in turn implies that for a 1 dB increase in the fundamental level, the ith-order IM component increases by i dB.

The ith-order IM distortion (IMD_i) can be measured by the ratio of the power in the ith-order IM component to that in the fundamental component. For example, the intermodulation distortion for the 3rd-order IM component at the frequency $2\omega_1 - \omega_2$ is obtained as follows:

$$
\begin{aligned}
IMD_3 &= 10 \log\left(\frac{p(2\omega_1 - \omega_2)}{p(\omega_1)}\right) \\
&= 20 \log\left(\frac{3}{4}\frac{a_3}{a_1} A^2\right)
\end{aligned}
\tag{3.60}
$$

where $p(\omega)$ denotes the average power of a component at frequency ω.

Generally, the IM components of greater concern are those that have frequencies very close to those of the input signal since they may be captured in the receiver's bandwidth, thus leading to degradation of the BER of the received signal. Given that the difference between ω_1 and ω_2 is generally small, the IM components of greater concern correspond to the odd-order terms having frequencies

$$
\omega = m\omega_1 - n\omega_2
\tag{3.61}
$$

The 3rd-order IM products corresponding to these frequencies dominate the intermodulation distortions at small signal levels. However, at large signal levels, the 5th- and 7th-order terms may become significant and should not be ignored.

For real signals like the two-tone test signal, the figure of merit for measuring the IM distortion given in equation (3.60) is appropriate. However, it is insufficient when strong linearity is required and complex digital modulations are involved e.g. in OFDM

56

signals. The two figures of merit best suited for these latter cases are the envelope vector magnitude (EVM) and adjacent channel power ratio (ACPR) [100], [101].

The EVM measures the extent of in-band distortions in a received signal. It is defined as the magnitude of the error vector between the original or reference constellation point and the respective measured point at the receiver as illustrated in Figure 3.12 for a 64-QAM constellation. It is computed over a number of symbols and is given by the formula

$$EVM = \sqrt{\frac{\sum_{k=1}^{K}|Y_k - X_k|^2}{\sum_{k=1}^{K}|X_k|^2}} \tag{3.62}$$

where X_k and Y_k are the vectors corresponding to the transmitted and received symbols and K is the number of symbols used in the EVM computation.

The ACPR ratio measures OOB radiations or the spread of a signal spectrum into adjacent channels due to IM distortions. It is given by the ratio of the power that leaks into an adjacent channel to the power in the main channel, i.e.

$$ACPR = 10\log\left(\frac{\int_{B_{ach}} S(f)\, df}{\int_{B_{mch}} S(f)\, df}\right) \text{ [dBc]} \tag{3.63}$$

where B_{ach} and B_{mch} are, respectively, the bandwidth of the adjacent and the main channel as illustrated in Figure 3.13 and $S(f)$ is the power spectral density of the signal amplified through the power amplifier.

(a)   (b)
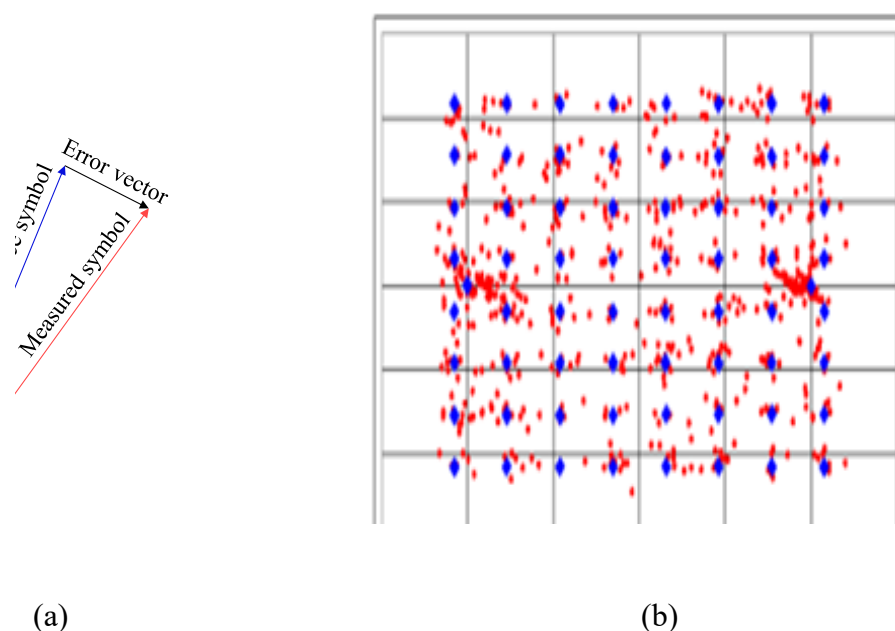
Figure 3.12. EVM (a) Error vector (b) Reference (blue) and received (red) constellation points [102]
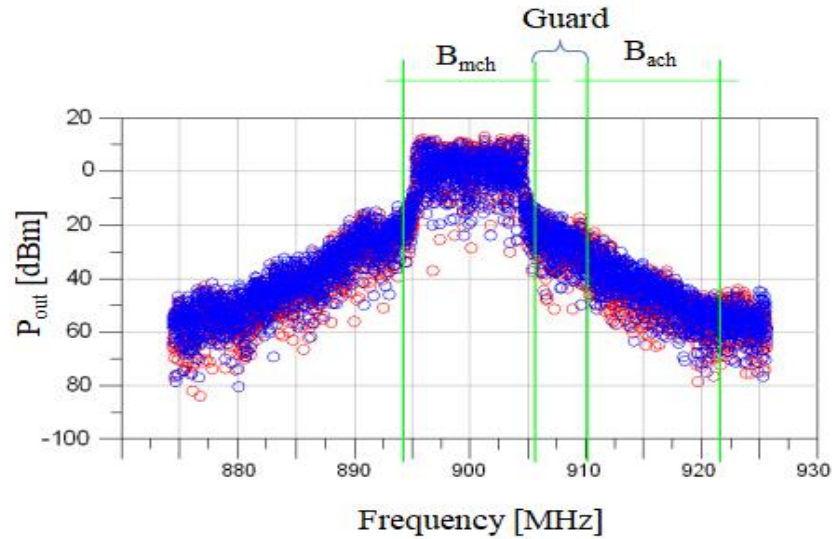
Figure 3.13. Power leakage into adjacent channel [102]

3.8.2 Operating Point of Power Amplifier

To avoid nonlinear distortions, the operating point of a power amplifier can be set in the linear region away from the saturation point depending on the envelope fluctuations of the input signal. Practically, the operating point is determined by the value of the input power back-off (IBO), which is the ratio of the saturation (maximum peak) power to the average power of a signal, i.e.

$$\text{IBO} = 10 \log \left(\frac{P_{\text{in,sat}}}{P_{\text{av}}} \right) \ [\text{dB}] \tag{3.64}$$

Thus, it is equal to the PAPR of the signal.

From equation (3.64), it is clear that a signal that has a low PAPR requires a small IBO, and hence can be amplified near the saturation point with high power efficiency. This is illustrated in Figure 3.14, where due to the reduction in PAPR, it is shown that the power fluctuations in the signal are reduced, thus making the operating point to move from low to high efficiency region owing to reduced IBO.

Without PAPR reduction, a high PAPR causes a power amplifier to be operated at low efficiencies, even below 5%. For example, a typical OFDM signal has a PAPR ranging from 10 to 12 dB, thus requires a legacy linear power amplifier to be operated with an IBO of between 10 and 12 dB from the 1-dB compression point in order to amplify linearly all the signal amplitudes. Since a linear power amplifier has an efficiency of between 40 and 70%, an IBO of 10 to 12 dB forces the device to operate with a much lower efficiency of around 5%.

58
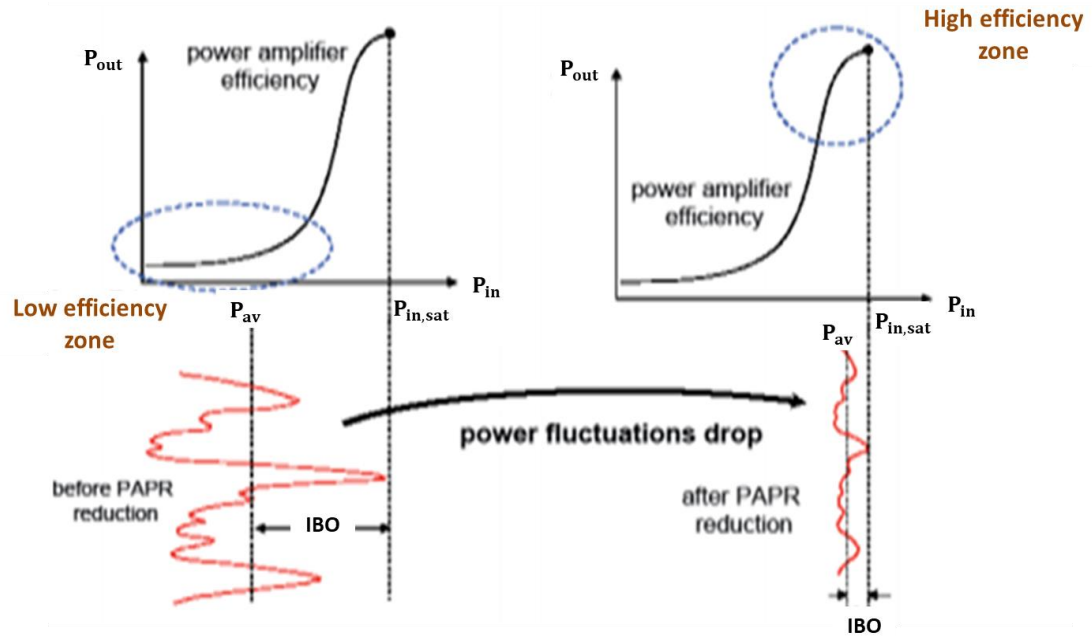
Figure 3.14. Power amplifier operation at a high efficiency region [103]

The low efficiency operation leads to a high power consumption and a high thermal power dissipation and hence an increase in the cost of the device due to the need of an advanced cooling system. High power consumption may translate to a high utility bill that can be extreme especially with base stations, which are normally in large numbers in a wireless communication network. This is in addition to air pollution and carbon emissions from the electricity generating stations that power the base stations.

For example, a single-carrier tri-sector 3G macro-site base station, that is commonly used to cover highly populated areas, when using a power amplifier that has been linearized and the efficiency enhanced to 15%, may consume power of approximately 1.45 kW [104]. The power amplifier in the base station consumes about 55% (0.8 kW) of this power.

At the current cost of KSh 19.70 per kWh in Kenya, the total cost of power will be KSh 246,801.60 per year per a base station. An operator having 2,500 base stations would consume 3.625 MW of power at a cost of KSh 617 million per year at a carbon footprint of approximately 0.02 $MtCO_2$ per year.

In addition, in the case of portable devices such as mobile phones, operating a power amplifier at a low efficiency zone owing to a large IBO depletes the battery charge very fast. Thus, the battery requires constant or frequent charging, which greatly inconveniences the user.

3.8.3 Bit-Error Rate Performance

Although the EVM given in equation (3.62) is a good measure of in-band distortions in a signal, the most important performance metric in a digital communication system is the BER. By definition, BER is the ratio of the number of erroneous bits to the total number of bits received. However, it is commonly measured by the probability of a bit error (P_b) at a given ratio of bit energy to noise power spectral density (E_b/N_0). The ratio E_b/N_0 can be related to the SNR by expressing it as follows:

$$\frac{E_b}{N_0} = \frac{\dfrac{S}{R}}{\dfrac{N}{W}}$$

$$= \frac{S}{N}\frac{W}{R}$$

(3.65)

where the ratio S/N is the SNR, W is the signal detection bandwidth at the receiver and R is the bit rate. Because of the relationship between the two ratios, E_b/N_0 is commonly referred to as SNR per bit.

It is normally desired that the required SNR per bit be as small as possible for a given value of P_b. However, this is not always the case. A typical plot of P_b values against E_b/N_0 has a "waterfall-like" shape indicating a trade-off between the two as illustrated in Figure 3.15. For each curve, P_b decreases nonlinearly with the increase in E_b/N_0. The arrow in the figure indicates the direction of improved BER performance. Thus, curve 2 has a better BER performance than curve 1 because for the same value of P_b e.g. at $P_b = p_0$, curve 2 requires a lower E_b/N_0 than curve 1.

The required BER depends on the underlying application. For a given application, the BER of the received signal is expected to be equal to or less than the one specified for that particular application. For example, vocoded speech requires a BER between 10^{-2} and 10^{-3}, data transmission over wireless channel aims for a BER between 10^{-5} and 10^{-6}, video transmission needs a BER in the range of 10^{-7} to 10^{-12}, while financial data transmission requires a BER of 10^{-11} and below [105].

In an OFDM system, the P_b can be found by considering the modulation scheme employed together with the underlying channel characteristics. For the general M-ary QAM scheme, the symbol-error rate (SER) probability function for an additive white Gaussian noise (AWGN) channel or a Raleigh-fading channel is sufficient for analysing the BER performance of a system.
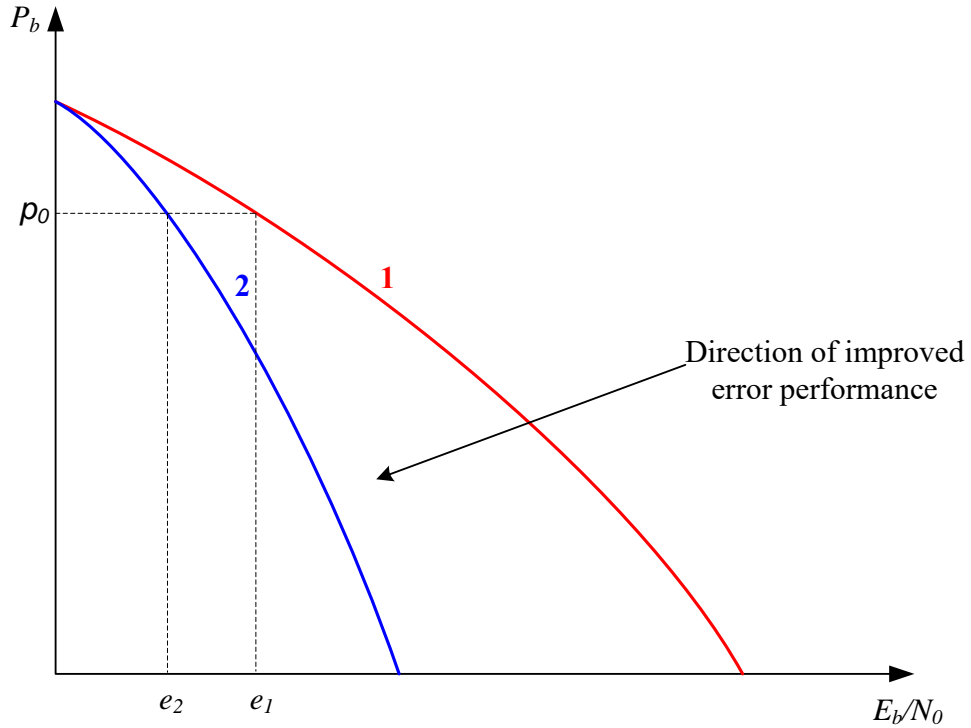
Figure 3.15. Illustration of bit-error rate probabilities

In an OFDM wireless channel characterised by AWGN and employing M-ary QAM, the probability of occurrence of a symbol error is given [105] by

$$P_s = 1 - \left[1 - \frac{2(\sqrt{M} - 1)}{\sqrt{M}} Q\left(\sqrt{\frac{6 \log_2(\sqrt{M})}{(M - 1)} \frac{E_b}{N_0}}\right)\right]^2 \qquad (3.66)$$

where $Q()$ is the well-known Q-function defined as

$$Q(x) = 1 - F_X(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy \qquad (3.67)$$

where $F_X(x)$ is the CDF of a random variable X.

The bit-error probability can be obtained from the symbol-error probability in equation (3.66) as

$$P_b = \frac{P_s}{\log_2 \sqrt{M}} \qquad (3.68)$$

rather than by $\log_2 M$ because an M-ary QAM symbol can be considered to consist of an in-phase and a quadrature component, each of them representing $\log_2 \sqrt{M}$ bits, and

61

modulating an in-phase carrier $\cos(2\pi f_c t)$ and a quadrature carrier $\sin(2\pi f_c t)$, respectively.

In the case of a transmission through a Rayleigh-fading channel, the probability of a symbol error is obtained as follows [93]:

$$P_s = \frac{\left(\sqrt{M} - 1\right)}{\sqrt{M}} \left[1 - \frac{\beta}{\beta + 1}\right] \tag{3.69}$$

where

$$\beta = \frac{6 \log_2\left(\sqrt{M}\right)}{2(M - 1)} \frac{E_b}{N_0} \tag{3.70}$$

Similarly, the bit-error probability is given by equation (3.68).

The two symbol-error rates obtained using equation (3.66) and (3.69) offers a theoretical lower-limit or the best performance that can be used for benchmarking simulated SER performances for a given communication system.

3.9 Multiple-Antenna Techniques in OFDM System

OFDM can be combined with multiple antennas in the transmitter and receiver to boost the overall channel capacity and reliability of a communication system. Such a system is referred to as multiple-input multiple-output OFDM, i.e. MIMO-OFDM system. Since OFDM technique has been covered in the preceding sections of this chapter, this section will briefly cover MIMO technology and its benefits in the first part. This will then be followed by a short coverage of one of the most common transmit-antenna diversities referred to as Alamouti space-time block coding that was used in the simulation of MIMO-OFDM systems in this work.

3.9.1 MIMO Antenna Technique

In a MIMO system, multiple antennas are employed at both the transmitter and the receiver. The antennas are sufficiently spaced, typically by a distant of more than 10λ, to avoid interferences. Denoting the number of transmit antennas by N_t and the receive antennas by N_r in a general MIMO configuration, three special cases of antenna arrangements that can be derived. These are the *single-input single-output* (SISO) when $N_t = N_r = 1$, *single-input multiple-output* (SIMO) in which $N_t = 1$ and *multiple-input single-output* (MISO) antenna configuration when $N_r = 1$.

The aim of combining multiple transmit and receive antennas in a communication system is to achieve spatial diversity and/or spatial multiplexing. Spatial diversity improves the BER performance, hence increasing reliability of the system. On the other hand, spatial multiplexing boosts the transmitted data rate. Combining spatial diversity and spatial multiplexing in a communication system improves both system reliability and bit rate.

The spatial diversity of a MIMO system can be measured by what is referred to as diversity order or gain, which by definition is the number of uncorrelated transmission channels between the transmitter and the receiver and is given by

$$N_d = N_t \times N_r \tag{3.71}$$

Spatial diversity greatly improves system reliability because the probability of having a low SNR on all the N_d channel paths is very small. In particular, the average bit-error probability in a MIMO system decays exponentially as diversity order [106] increases according to the relation

$$P_b = k \left(\frac{S}{N}\right)^{-N_d} \tag{3.72}$$

where k is a constant that depends on the modulation scheme used in the system and S/N is the received SNR.

In addition to providing spatial diversity, multiple transmit antennas can be used for beamforming, i.e. to shape the beam of the transmit signal towards the target receiver, when the relative phases of the downlink channels are known at the transmitter. Beamforming can increase the strength of received signal by up to N_t times, thus allowing the coverage area to be increased or the transmit power to be reduced.

Additionally, an increase in the signal strength can lead to a linear increase in SNR and a proportional increase in the channel capacity as illustrated in Figure 3.16. This is according to Shannon-Hartley channel capacity theorem [107], which can be expressed as

$$C_N = \log_2 \left(1 + \frac{S}{N}\right) \tag{3.73}$$

where $C_N = C/W$ is the normalized channel capacity, i.e. channel capacity divided by the bandwidth, and S and N denote the average power of signal and noise, respectively.

However, the stated increase in channel capacity happens only for channels that

have low SNR, in which case $C_N \approx S/N$. Because of low SNR, such channels have a low data rate. As the diversity is increased, the channel capacity becomes proportional to $\log_2(S/N)$ at high values of SNR and eventually saturates unless the bandwidth is also increased.

A substantial increase in data rate at higher SNRs can be achieved through spatial multiplexing where multiple antennas send multiple uncorrelated data streams. This allows the channel capacity to increase linearly with the number of antennas without saturating. In a power-limited system, it is possible to have $N_m = \min(N_t, N_r)$ parallel channels each with N_m times lower SNR due to splitting of the available power among the channels [108]. The normalized channel capacity for one channel is then given by

$$C_N^{(i)} = \log_2\left(1 + \frac{N_r}{N_m}\frac{S}{N}\right), \qquad i = 0, 1, \dots, N_m \tag{3.74}$$

while that for the whole system is

$$C_N = N_m \log_2\left(1 + \frac{N_r}{N_m}\frac{S}{N}\right) \tag{3.75}$$

By setting $N_t = N_r$ in equation (3.75), the total channel capacity for a MIMO system with spatial multiplexing becomes

$$C_N = N_m \log_2\left(1 + \frac{S}{N}\right) \tag{3.76}$$

Thus, it increases linearly with the number of antennas.
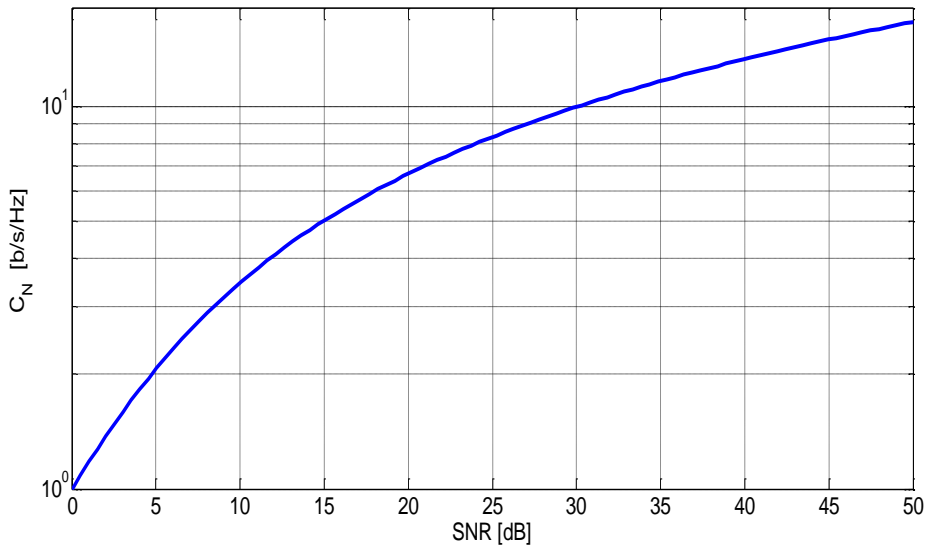
Figure 3.16. Normalized channel capacity variation with SNR

3.9.2 Transmit Diversity by Space-Time Block Coding

The Alamouti space-time coding was the first space-time block coding (STBC) to provide transmit diversity. The Alamouti STBC delivers a space-time rate of one without ISI. A space-time rate of one means that the encoder input has the same symbol rate as the one available at each transmit antenna, thus offering bandwidth utilization equal to one.

The Alamouti STBC is a complex orthogonal space-time coding scheme that specialise on just two transmit antennas. It can also be extended to a higher number of antennas but for complex modulation schemes, such as M-ary QAM, a space-time rate of one without ISI can only be achieved with two transmit antennas.

The encoding of symbols in an Alamouti STBC is done on two consecutive symbols, X_1 and X_2, using a space-time codeword matrix [109]:

$$X_{Alamouti} = \begin{bmatrix} X_1 & -X_2^* \\ X_2 & X_1^* \end{bmatrix} \tag{3.77}$$

where the superscript $*$ signifies complex conjugation.

The signal from the Alamouti-STBC encoder is sent over two symbol periods from two transmit antennas. The first and the second antenna, respectively, transmits symbols X_1 and X_2 simultaneously in the first symbol duration. In the second symbol duration, the first and the second transmit antenna essentially resends the same symbols in the form $-X_2^*$ and X_1^*, respectively. The receiver needs to implement an Alamouti STBC decoder for recovering the transmitted symbols.

There are two realizations of receivers for decoding Alamouti STBC signals. The first one has one receive antenna, while the other has two, respectively forming 2×1 and 2×2 multiple-antenna systems. Since over the two symbol durations, only two symbols are essentially transmitted, the two implementations of Alamouti STBC receivers have the same channel capacity as a SISO configuration.

However, the 2×2 MIMO scheme has a diversity gain of 4, which is twice that of a 2×1 MISO configuration and four times that of SISO system. Similarly, from equation (3.72), the 2×2 MIMO system has a better BER performance than the other two systems.

Figure 3.17 illustrates the Alamouti STBC encoding scheme for the 2×2 configuration. The impulse responses h_{11}, h_{12}, h_{21} and h_{22} of the four channel paths are assumed to be complex and constant over two symbol periods, hence can be expressed in

65

the form

$$h_{ij} = |h_{ij}|e^{\theta_{ij}} \tag{3.78}$$

where $|h_{ij}|$ and $e^{\theta_{ij}}$ are the amplitude and phase of the impulse response of the path from antenna j to i over two symbol durations. Subscripts i and j can be either 1 or 2.

At each receive antenna, the received signal is given by the convolution of the channel impulse responses and the symbols transmitted from the two antennas. Noting that a convolution of a signal and a constant is simply the product of the two, the signals received by the top receive antenna in Figure 3.17 in the first and second symbol intervals are given, respectively, by

$$Y_{11} = h_{11}X_1 + h_{12}X_2 + n_{11}$$
$$Y_{12} = -h_{11}X_2^* + h_{12}X_1^* + n_{12} \tag{3.79}$$

Similarly, the received signals at the second receive antenna are

$$Y_{21} = h_{21}X_1 + h_{22}X_2 + n_{21}$$
$$Y_{22} = -h_{21}X_2^* + h_{22}X_1^* + n_{22} \tag{3.80}$$

The four additive terms n_{11}, n_{12}, n_{21} and n_{22} are the AWGN terms of the four channels.

On complex conjugating the two signals received during the second interval, the system equation for an Alamouti STBC 2×2 MIMO system can be written as

$$\begin{bmatrix} Y_{11} \\ Y_{12}^* \\ Y_{21} \\ Y_{22}^* \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^* & -h_{11}^* \\ h_{21} & h_{22} \\ h_{22}^* & -h_{21}^* \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} n_{11} \\ n_{12}^* \\ n_{21} \\ n_{22}^* \end{bmatrix}. \tag{3.81}$$
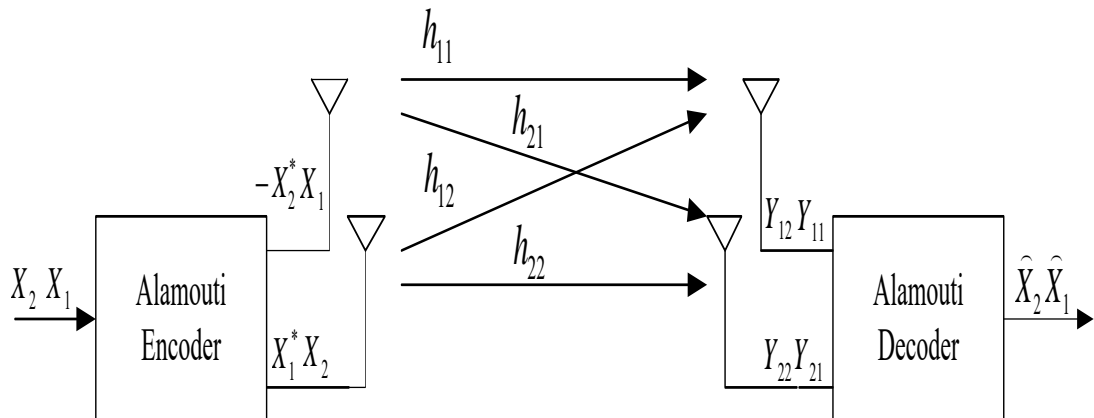
Figure 3.17. Alamouti scheme for 2×2 MIMO system

66

and has the form

$$\mathbf{Y} = H\mathbf{X} + \mathbf{n} \tag{3.82}$$

where H is the channel matrix given by

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^* & -h_{11}^* \\ h_{21} & h_{22} \\ h_{22}^* & -h_{21}^* \end{bmatrix} \tag{3.83}$$

Because the two columns of H are orthogonal, the estimates of the transmitted symbols can be found by multiplying the two sides of equation (3.82) by the conjugate transpose of H given by

$$H^H = \begin{bmatrix} h_{11}^* & h_{12} & h_{21}^* & h_{22} \\ h_{12}^* & -h_{11} & h_{22}^* & -h_{21} \end{bmatrix} \tag{3.84}$$

The transmitted symbols can then be estimated as follows:

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \end{bmatrix} = \frac{H^H \mathbf{Y}}{|h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2} \tag{3.85}$$

3.9.3 Space-Time Block-Coded MIMO-OFDM System

The 2×2 STBC MIMO system in Figure 3.17 can be combined with OFDM to have a 2×2 MIMO-OFDM system. This can be accomplished by adding blocks for processing OFDM signals after the STBC encoder as illustrated in Figure 3.18 [110], [111].
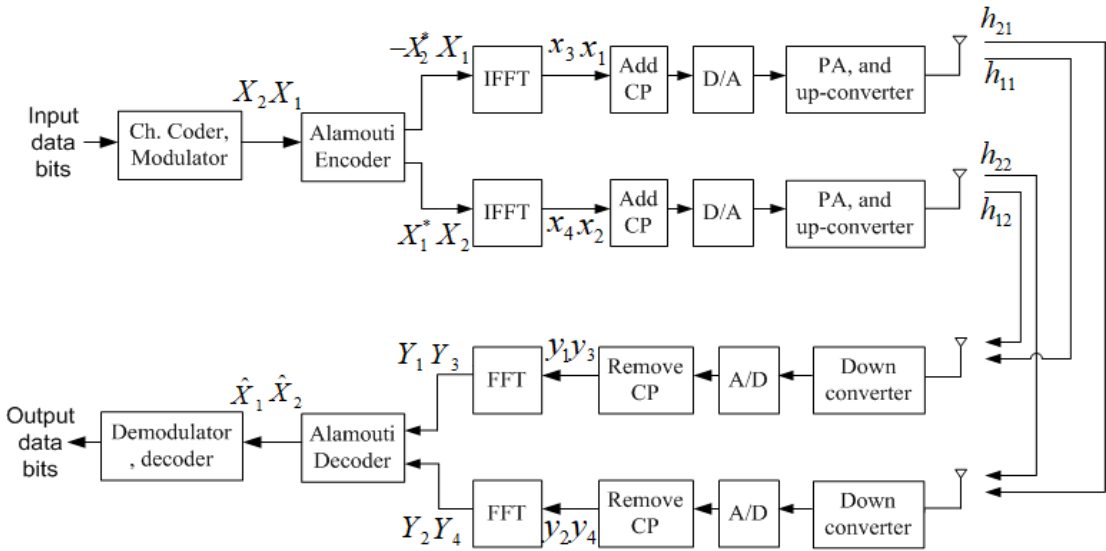
Figure 3.18. MIMO-OFDM system employing 2×2 Alamouti STBC

67

Therefore, a general MIMO-OFDM transmitter has N_t independent transmit branches, with each of them similar to a SISO-OFDM transmission path. Each transmit branch performs the same sequence of signal-processing operations presented in Section 3.5, which include serial-to-parallel conversion, pilot insertion, N-point IFFT, cyclic extension and digital-to-analogue conversion before the final signal is amplified and up-converted to RF level for transmission. In some spatial multiplexing systems [112], channel encoding and digital modulation can also be done differently on each of the N_t transmit branches.

The receiver in a MIMO-OFDM system reverses the signal processing operations done in the transmitter in a similar manner described earlier in Section 3.5 for the SISO-OFDM system. From the foregoing presentation of MIMO-OFDM system, it is clear that each of the N_t transmit signals has the same characteristics as those for the OFDM signal, including a high PAPR. Therefore, due to high PAPR, some or all of the N_t transmit signals may experience nonlinear amplification resulting in the undesirable effects presented in Section 3.8.

3.9.4 PAPR of a MIMO-OFDM System

In a MIMO-OFDM system, there are N_t transmit branches and hence N_t parallel OFDM signals, each with its own PAPR. Considering the same specifications for HPAs in all the N_t branches, the interest would then be to limit the maximum PAPR in order to avoid nonlinear distortions in the system. The maximum PAPR for the system can obtained as

$$PAPR_{MIMO} = \max\big(\text{PAPR}\{x_1(n)\}, \text{PAPR}\{x_2(n)\}, \dots, \text{PAPR}\{x_{N_t}(n)\}\big) \quad (3.86)$$

where $x_i(n)$ is a sufficiently oversampled discrete-time signal corresponding to the ith transmit path during the first symbol duration.

The maximum PAPR in the second symbol duration can be found in a similar manner. However, since the same symbols are essentially retransmitted in the second period, the maximum PAPR will be the same for the two symbol durations.

CHAPTER 4

FUNDAMENTALS OF CONVEX OPTIMIZATION AND ADAPTIVE SIGNAL PROCESSING

This chapter covers the basic concepts that form the foundation of this work, specifically for the development of the proposed algorithms. These concepts are mainly drawn from convex optimization and adaptive signal processing. The first part focuses on the essential concepts of convex functions and optimization. Under this part, a special coverage is given to three classes of convex optimization problems, namely least-squares minimization, linear programming and second-order cone programming, which form the basis of the PAPR reduction algorithms proposed in this thesis. In the second part, important concepts of adaptive signal processing that are used to characterise the proposed algorithms are covered in details.

4.1 Basics of Convex Functions and Optimization

Optimization problems have become a common occurrence in signal processing applications e.g. in the design of filters, signal estimations and noise cancellation. An optimization problem normally involves finding an optimum solution to a cost function defined over the input signal and some coefficients. The general form of an optimization problem is given by a set of equations

$$
\begin{aligned}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & g(x) \leq 0 \\
& h(x) = 0 \\
& x_l \leq x \leq x_u
\end{aligned}
\tag{4.1}
$$

where x is the optimization variable of the problem, $f(x)$ is the objective function and $g(x)$ and $h(x)$ are the inequality and equality constraint functions, respectively. The limits x_l and x_u are the lower and upper bounds of the optimization variable, respectively [113].

Many optimization problems exist but they can be classified according to the characteristics of their objective and constraint functions. Of interest to this thesis is the class of convex optimization problems in which both the objective and constraint functions are convex. This means that the values of the functions satisfy the following inequality:

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2) \tag{4.2}$$

for every real number α in the range $0 < \alpha < 1$ and every pair of points (x_1, x_2) in a convex set $\mathcal{R}_c$.

The set $\mathcal{R}_c$ is said to be convex if a third point formed as

$$x_3 \leq \alpha x_1 + (1 - \alpha)x_2 \tag{4.3}$$

from any two points in the set is also located in the same set. In other words, if a straight line connects any two points (x_1, x_2) in the set $\mathcal{R}_c$, then $\mathcal{R}_c$ is said to be convex if every point on the line segment between the two points is a member of the same set.

On the left-hand side of equation (4.2), $f(x)$ is evaluated on the line segment joining points x_1 and x_2, whereas on the right-hand side, it is approximated using a linear interpolation. Thus, the function $f(x)$ is said to be convex over the set $\mathcal{R}_c$ if the linear interpolation of the function between any two points in the set overestimates the function.

Since all functions in a convex optimization problem are convex, the local minimum of function $f(x)$ is the global minimum. In other words, an optimal solution is always guaranteed for a convex optimization problem. Unfortunately, there is in general no analytical method for finding optimal solutions of convex optimization problems. However, the interior-point algorithms [114] have of late proved to be very effective in solving the problems in a number of iterations between 10 and 100.

An interior-point method finds the best solution by traversing the interior of the feasible region, itself a convex set, encoded by a self-concordant barrier function. This follows from the fact that any convex optimization problem can be transformed into a minimization (or a maximization) of a linear function over a convex set by converting it into an epigraph form. The barrier function is a convex function with the property that its value at a point tends to infinity as the point approaches the boundary of the feasible region of an optimization problem.

The interior-point method approximates a constrained convex optimization problem

$$\text{minimize} \quad f_0(x)$$

$$\text{subject to} \quad f_i(x) \leq 0, \quad i = 1, 2, \dots, m \tag{4.4}$$

where the inequality constraints $f_i(x)$, $i = 1, 2, \dots, m$, are convex and smooth (i.e. at least twice continuously differentiable), by another smooth unconstrained convex optimization problem of the form

$$\text{minimize} \quad f_0(x) + \frac{1}{\mu} \mathcal{B}(x) \tag{4.5}$$

where μ is a small positive scalar, i.e. $\mu > 0$ and $\mathcal{B}(x)$ is a barrier function for the feasible set of the original problem in equation (4.4).

Parameter μ is used to control how deep inside the feasible set one wants to be with respect to how much the objective function need to be minimized. In other words, the parameter sets the accuracy of the optimization. A high value of μ gives a good approximate solution of the original problem in equation (4.4).

The basic idea of using the barrier function is to implicitly put the inequality constraints in the objective function and allow Newton's method, discussed in Section 4.2, to solve the optimization problem. Thus, the barrier function should be continuous and at least twice differentiable. A typical barrier function for the problem is the *logarithmic barrier* function given by

$$\mathcal{B}(x) = -\sum_{i=1}^{m} \log(-f_i(x)) \tag{4.6}$$

The domain of this function is the set of points that strictly satisfy all the inequality constraints.

Although the quality of approximation of the solution of the problem in equation (4.5) to the one of (4.4) improves with increasing value of μ, it is difficult to minimize the former when the parameter is large, since the Hessian of $f_0(x) + (1/\mu)\mathcal{B}(x)$ changes quickly close to the boundary of the feasible set. To avoid this difficulty, the interior-point algorithm finds the optimal solution by solving a sequence of unconstrained convex problems with increasing values of μ.

One important advantage of interior-point algorithms is that the number of iterations is determined by a solver's stopping criteria rather than the problem size and starting point. The stopping criteria may include several tolerances that can be set by the

user depending on the required precision in a problem. A tolerance is simply a threshold that when crossed stops the iterations of a solver.

Typical tolerances for interior-point algorithms include the function tolerance that gives the lower bound on the change in the value of the objective function from one iteration to the next and the maximum allowed number of iterations. Others are the constraint tolerance that gives an upper bound of the absolute value of any of the constraints, and the step tolerance that sets a lower bound of the step-size.

From the foregoing discussion, it is clear that the computational effort in an iteration of an interior-point algorithm determines the differences of computational complexities and convergence times of different optimization problems.

4.1.1 Norm Approximation

In most signal processing applications, the optimization problem involves estimating a specified or desired response. The problem requires minimization of the difference between the actual and desired response, commonly referred to as approximation or residual error, given by

$$\mathbf{r} = \mathbf{b} - A\mathbf{x} \tag{4.7}$$

where A is a matrix of m rows and n columns containing real coefficients, i.e. $A \in \mathbb{R}^{m \times n}$, $\mathbf{x}$ is a real vector of n components, i.e. $\mathbf{x} \in \mathbb{R}^n$, and the product $A\mathbf{x}$ constitutes the actual response, while vector $\mathbf{b} \in \mathbb{R}^m$ represents the desired response.

If $\mathbf{b}$ does not lie in the space spanned by the columns of A or $m \neq n$, there is no exact solution to equation (4.7), i.e. $\mathbf{r}$ cannot be $\mathbf{0}$. The case when an exact solution does not exist corresponds to the system of linear equations $A\mathbf{x} = \mathbf{b}$ being overdetermined if $m > n$ or underdetermined if $m < n$. For these two cases, in order to have $\mathbf{r} \approx \mathbf{0}$, an approximation problem is posed to minimize an objective function defined as a norm of some type on vector $\mathbf{r}$.

The chosen norm is simply a function that assigns a positive length to vector $\mathbf{r}$ in the respective vector space. Using a general ℓ_p-norm, an objective function $f(x)$ can be expressed as

$$f(x) = \|\mathbf{r}\|_p = \left(\sum_{i=1}^{m} |r_i(x)|^p \right)^{\frac{1}{p}} \tag{4.8}$$

72

or simply as

$$f(x) = \|\mathbf{r}\|_p^p = \sum_{i=1}^{m} |r_i(x)|^p \tag{4.9}$$

where p is an positive integer and $r_i(x)$ denotes the ith component of $\mathbf{r}$.

Two special cases of ℓ_p-norm, one when $p = 2$ and the other when $p = \infty$, corresponding to the *Euclidean* and *Chebyshev* norms are discussed next together with their respective optimization problems.

4.1.2 Least-Squares Problem

If $p = 2$ in equation (4.8), the *Euclidean* or ℓ_2-norm given by

$$f(x) = \|\mathbf{r}\|_2 = \left(\sum_{i=1}^{m} |r_i(x)|^2 \right)^{\frac{1}{2}} \tag{4.10}$$

is minimized.

On squaring the *Euclidean* norm, the objective function can be written as

$$f(x) = \|\mathbf{r}\|_2^2 = \sum_{i=1}^{m} |r(i)|^2 \tag{4.11}$$

but still minimizes the same norm.

The optimization problem in equation (4.11) is referred to as a *least-squares problem*. In vector form, it can be expressed as follows:

$$f(x) = (\mathbf{b} - A\mathbf{x})^T (\mathbf{b} - A\mathbf{x})$$
$$= \mathbf{b}\mathbf{b}^T - 2\mathbf{b}^T A\mathbf{x} + \mathbf{x}^T A^T A\mathbf{x} \tag{4.12}$$

Here, without loss of generality, matrix A and vectors $\mathbf{r}$, $\mathbf{x}$ and $\mathbf{b}$ are assumed to be real-valued and $m > n$. Equation (4.12) is quadratic in $\mathbf{x}$, and therefore convex, and can be solved by setting the first derivative, given by

$$\frac{\partial f(x)}{\partial x} = -2A^T \mathbf{b} + 2A^T A\mathbf{x} \tag{4.13}$$

to zero. This is equivalent to solving a system of linear equations

$$(A^T A)\mathbf{x} = A^T \mathbf{b} \tag{4.14}$$

normally referred to as *normal equations*, whose closed form solution is

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b} \tag{4.15}$$

The computational complexity of obtaining the least-squares solution is dependent on the dimensions of the coefficient matrix A and is found to be $O(mn^2)$.

4.1.3 Weighted Least-Squares Problem

In some cases of the application of least-squares minimization, there may be a need to emphasize the minimization of some critical residual error components, while de-emphasizing others. This can be achieved through solving a *weighted least-squares problem*, where different weights are applied on the error components to yield an objective function of the form

$$f(x) = \|W\mathbf{r}\|_2^2 = \sum_{i=1}^{m} |w_i r_i(x)|^2 \tag{4.16}$$

where $W \in \mathbb{R}^{m \times m}$ is a diagonal matrix of relative weights $w_i, \ i = 1, 2, \dots, m,$ i.e.

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_m \end{bmatrix} \tag{4.17}$$

Here, w_i denotes the weight applied on the ith component of $\mathbf{r}$.

The weighted least-squares problem in equation (4.16) can be readily cast into a standard least-squares problem with the following set of normal equations:

$$(A^T W^T W A)\mathbf{x} = A^T W^T W \mathbf{b} \tag{4.18}$$

hence leading to a closed form solution

$$\mathbf{x} = (A^T W^T W A)^{-1} A^T W^T W \mathbf{b} \tag{4.19}$$

After the minimization of the problem in equation (4.16), the final residual errors in the weighted error vector tend to have approximately the same magnitude, say equal to δ, i.e. , i.e. $w_i r_i(x) \cong \delta$ for $i = 1, 2, \dots, m$. Therefore, in case a large weight, say w_k, was placed on the component $r_k(x)$ during the minimization, the corresponding residual error on this component in the un-weighted error vector $\mathbf{r}$ is δ/w_k and is therefore very small, i.e. it was highly minimized owing to the large weighting.

4.1.4 Iterative Re-Weighted Least Squares

The iterative re-weighted least squares (IRLS) is an optimization procedure that iteratively solves the weighted least-squares problem in equation (4.16) in order to solve the ℓ_p-norm approximation problem of the objective function in equation (4.9). The objective function for the minimization of the ℓ_p-norm can be expressed as an equivalent weighted least-squares function as follows:

$$f(x) = \|\mathbf{r}\|_p^p = \sum_{i=1}^m \left| r_i^{\frac{p-2}{2}}(x) r_i(x) \right|^2 \tag{4.20}$$

which can be rewritten as

$$f(x) = \|\mathbf{r}\|_p^p = \sum_{i=1}^m |w_i r_i(x)|^2 \tag{4.21}$$

where the components of the residual error are weighted by

$$w_i = |r_i(x)|^{\frac{p-2}{2}}, \qquad i = 1, 2, \dots, m \tag{4.22}$$

The weighted least-squares objective function in equation (4.21) has the same analytical solution given in equation (4.19). However, because the optimal weights that will give the minimum value of the ℓ_p-norm are unknown, the IRLS algorithm solves the weighted least-squares problem iteratively until convergence is achieved when the residual error does not change any more [115].

The algorithm starts by setting all weights to unity, i.e. $w_i = 1$, for $i = 1, 2, \dots, m$, and then finding the initial solution using equation (4.19). At every other iteration, the algorithm calculates a new error vector using equation (4.7) and new weights from equation (4.22). In order to have a proportionate weighting on the components of the residual error, the new weights are normalized by the sum $\sum_{i=1}^m w_i$. Using the normalized weights, the algorithm finds a new solution $\mathbf{x}$. These steps are repeated until a possible convergence to an optimal solution for the ℓ_p-norm approximation problem.

4.1.5 Linear Programming Problem

Linear Programming is a class of optimization problems in which the objective and all constraint functions are linear functions of the optimization variable, i.e. they satisfy the condition

$$f(\alpha x_1 + (1 - \alpha)x_2) = \alpha f(x_1) + (1 - \alpha)f(x_2) \tag{4.23}$$

On comparing equation (4.23) and equation (4.2), it can be observed that, like the least-squares problem, the linear program falls under the class of convex optimization problems.

A general linear program (LP) can be formulated using the following set of equations:

$$
\begin{aligned}
\text{minimize} \quad & \mathbf{c}^T \mathbf{x} \\
\text{subject to} \quad & A_{neq}\mathbf{x} \leq \mathbf{b}_{neq} \\
& A_{eq}\mathbf{x} = \mathbf{b}_{eq} \\
& \mathbf{x}_l \leq \mathbf{x} \leq \mathbf{x}_u
\end{aligned}
\tag{4.24}
$$

Here $\mathbf{c}^T \mathbf{x}$ is the inner product of vectors $\mathbf{c}$ and $\mathbf{x}$, and $\mathbf{x}_l$ and $\mathbf{x}_u$ are the lower and upper bounds of vector $\mathbf{x}$, respectively.

Unlike the least-squares problem, the linear program in equation (4.24) does not have a closed form solution. Nevertheless, there exist several efficient algorithms for solving the program such as interior-point, simplex and active-set algorithms [116].

In practice, the computational complexity of linear program solvers such as the interior-point method is known to be $O(n^2 m)$, where n denotes the number of variables, i.e. the length of vector $\mathbf{x}$, and m is the number of constraints. Many problems, which in their original formulation are not in the standard linear program form in equation (4.24), can be readily transformed to an equivalent linear program and solved using an appropriate linear program solver.

4.1.6 Chebyshev Approximation Problem

When $p = \infty$ in equation (4.8), the objective function is referred to as *Chebyshev* or ℓ_∞-norm and is given as

$$f(x) = \| \mathbf{r} \|_\infty = \lim_{p \to \infty} \left(\sum_{i=1}^{m} |r_i(x)|^p \right)^{\frac{1}{p}} \tag{4.25}$$

If we let f_{max} denote the maximum of the absolute values of the components of $\mathbf{r}$, i.e. $f_{max} = \max_{1 \leq i \leq m} |r_i(x)|$, then equation (4.25) can be written as follows:

$$f(x) = f_{max} \lim_{p \to \infty} \left(\sum_{i=1}^{m} \left| \frac{r_i(x)}{f_{max}} \right|^p \right)^{\frac{1}{p}} \tag{4.26}$$

Because all the terms $|(r_i(x))/f_{max}|$ in the sum in equation (4.26) are less than unity except for the maximum term, which is equal to 1, they will tend to zero when raised to a large power and hence the summation in the bracket will be equal to 1.

Therefore, an objective function defined in terms of ℓ_∞-norm is simply the maximum absolute value of the components of vector $\mathbf{r}$ and can be written as

$$f(x) = \|\mathbf{r}\|_\infty = \max\{|r_1(x)|, |r_2(x)| \dots, r_m(x)\} \tag{4.27}$$

An optimization problem to minimize this function is referred to as *Chebyshev approximation problem* or *minimax approximation problem* and is stated as follows:

$$\text{minimize } \max\{|r_1(x)|, |r_2(x)| \dots, r_m(x)\} \tag{4.28}$$

Unfortunately, there is no closed form solution to the *minimax problem* but upon casting it into a linear program, a solution can be found in a finite number of steps. The linear program for the *Chebyshev approximation problem* has an epigraph form

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & |r_i(x)| - t \leq 0 \\ & -|r_i(x)| - t \leq 0 \end{aligned} \tag{4.29}$$

for $i = 1, 2, \dots, m$ and t and x are the optimization variables.

4.1.7 Second-Order Cone Programming

A second-order cone programming (SOCP) is an optimization problem in which the constraints are second-order cones. Both the objective function and the second-order cone constraints are convex, thus making a SOCP a convex optimization problem. The standard form of a SOCP has the following set of equations:

$$\begin{aligned} \text{minimize} \quad & \mathbf{f}^T \mathbf{x} \\ \text{subject to} \quad & \|A_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i \end{aligned} \tag{4.30}$$

for $i = 1, 2, \dots m$. Here $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, while $\mathbf{f} \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{(k_i-1) \times n}$, $\mathbf{b}_i \in \mathbb{R}^{k_i-1}$, $\mathbf{c}_i \in \mathbb{R}^n$ and $d_i \in \mathbb{R}$ are the problem parameters.

The constraint $\|A_i \mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T \mathbf{x} + d_i$ is referred to as a *second-order cone*

constraint because it requires the affine functions $A_i\mathbf{x} + \mathbf{b}_i$ and $\mathbf{c}_i^T\mathbf{x} + d_i$ to lie in a unit second-order cone of dimension k_i defined as follows:

$$\mathcal{C}_{k_i} = \left\{ \begin{bmatrix} v \\ t \end{bmatrix} \in \mathbb{R}^{k_i} \mid \|v\|_2 \leq t \right\} \tag{4.31}$$

where $v = A_i\mathbf{x} + \mathbf{b}_i$ and $t = \mathbf{c}_i^T\mathbf{x} + d_i$.

A SOCP can be viewed as a generalization of several convex optimization problems such as the linear program, quadratic program (QP) and quadratically constrained quadratic program (QCQP). This is due to the reason that many convex optimization problems can be reformulated as the program in equation (4.30) through a slight modification of the second-order cone constraints.

For example, if $A_i = 0$ and $\mathbf{b}_i = 0$, for $i = 1, 2, \dots m$, the SOCP reduces to a general linear program

$$\begin{aligned} \text{minimize} \quad & \mathbf{f}^T\mathbf{x} \\ \text{subject to} \quad & 0 \leq \mathbf{c}_i^T\mathbf{x} + d_i \end{aligned} \tag{4.32}$$

Similarly, if $\mathbf{c}_i = 0$ for $i = 1, 2, \dots m$, the second-order cone constraints are reduced to $\|A_i\mathbf{x} + \mathbf{b}_i\|_2 \leq d_i$, which can be expressed as quadratic constraints

$$\|A_i\mathbf{x} + \mathbf{b}_i\|_2^2 \leq d_i^2, \qquad i = 1, 2, \dots m \tag{4.33}$$

hence the SOCP reducing to a QCQP.

4.2 Fundamental Concepts in Adaptive Signal Processing

In many signal-processing applications, there is need for digital filters to adapt their coefficients to time-varying characteristics of the input signal, noise, and/or the physical system in order to produce the desired output. This kind of signal processing is referred to as adaptive filtering.

There are many different adaptive filtering schemes but they differ only in the way the desired signal is obtained. Adaptive filtering has been widely used in signal processing to solve many problems [117] e.g., in channel equalization, signal prediction, interference cancellation, adaptive antenna arrays and system identification.

An adaptive filter system consists of two basic elements, namely a digital filter and an adaptation algorithm. The digital filter gives an output in response to an input while the adaptation algorithm is responsible for adjusting the filter coefficients. An

adaptive filter can be described as a digital filter that self-adjusts its coefficients in order to minimize a cost function.

A block diagram of an adaptive filter in the discrete-time domain is illustrated in Figure 4.1. The input signal to the filter is denoted by $x(n)$, the desired (or reference) signal by $d(n)$, the filter coefficients by $w(n)$ and the output signal by $y(n)$. The arrow through the block of filter coefficients, indicates their time adaptation with the error signal $e(n)$.

The error signal $e(n)$ is the difference between the output and the desired signal, i.e.

$$e(n) = d(n) - y(n) \tag{4.34}$$

The adaptation or the update of the coefficients is aimed at minimizing a certain objective function, which is normally defined as a specific norm of the error signal. The minimization of the error is supposed to yield a filter whose output approximates the desired signal.

Different applications have different requirements on the adaptation algorithm. The most common measures used to assess the suitability of an adaptation algorithm for a given application are the following:

i) Rate of convergence

This is the number of iterations an adaptation algorithm undergoes to converge to the optimal solution for a given objective function. A fast rate of convergence means that an algorithm is well suited for adjusting to rapid changes in the signal or system characteristics.
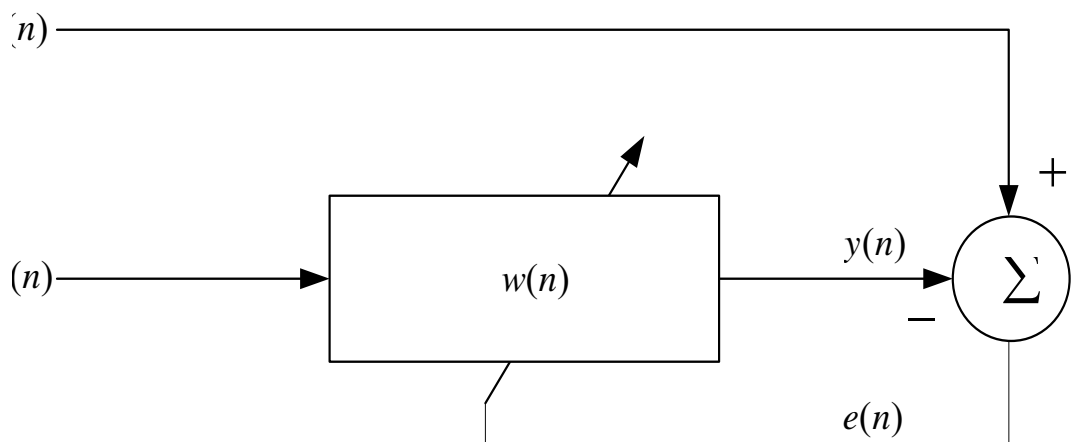
Figure 4.1. Block diagram of adaptive filter

ii)   Computational complexity

This refers to the number of multiplications and additions the adaptation algorithm performs to find the optimal filter coefficients. Since for every sample of the input signal, the algorithm must complete the computation of filter coefficients, a lower computational effort means that the signal can be sampled at a higher sampling rate and vice versa.

iii)   Robustness

This is the ability of an adaptation algorithm to converge under the effects of disturbances or that they just lead to small estimation errors. For example, a robust and numerically stable adaptive algorithm is insensitive to quantization errors resulting from finite-length registers in digital signal processors.

iv)   Misadjustment

This is an indicator of how far an obtained solution is from the optimal solution. It is defined as

$$m = \frac{f(\mathbf{w}^*) - f(\mathbf{w}_{opt})}{f(\mathbf{w}_{opt})} \tag{4.35}$$

where $f(\mathbf{w}^*)$ and $f(\mathbf{w}_{opt})$ are, respectively, the values of the objective function at the final adaptive filter coefficients and the optimal weights. For example, if the objective function is given by the mean square value of the error signal, then $f(\mathbf{w}_{opt})$ will be the optimal value given by the Wiener filter.

The adaptation process in the filter is largely determined by the choice of the filter structure. On the other hand, the structure is application-specific as the resulting filtering transfer function determines the relationship between the input and the output signals. Typical filter structures include the transversal (or tapped-delay line) finite-duration impulse response (FIR), lattice FIR and infinite-duration impulse response (IIR) filters.

The choice of the filter structure greatly affects the computational complexity of a given adaptation algorithm and the speed of convergence of the adaptation process. Because of its simplicity and efficiency, the transversal FIR structure in Figure 4.2, is by far the most widely used structure. Each of the blocks labelled z^{-1} represents a delay element. The output of the FIR filter is given by

$$y(n) = w_o x(n) + w_1 x(n-1) + \cdots + w_{N-1} x(n-N+1)$$
$$= \mathbf{w}^T \mathbf{x}(n)$$

(4.36)

where N is the filter order and $\mathbf{x}(n)$ and $\mathbf{w}$ are the vectors containing the input signal samples and filter coefficients $w(n)$ respectively, i.e.

$$\mathbf{x}(n) = [x(n)\, x(n-1) \ldots x(n-N+1)]^T$$

(4.37)

and

$$\mathbf{w} = [w_0\, w_1 \ldots w_{N-1}]^T$$

(4.38)

In case $\mathbf{x}(n)$ is complex, the output signal is given by

$$y(n) = \mathbf{w}^H \mathbf{x}(n)$$
$$= \mathbf{x}^H(n)\mathbf{w}$$

(4.39)

where the superscript H denotes the Hermitian or conjugate transpose.

An optimal filter design problem can be represented as shown Figure 4.3. The aim of optimal filter design is to find the filter coefficients that optimize a particular objective function of the error signal given the input and desired signals. Several criteria exist for optimizing the objective function depending on the purpose of the filter.
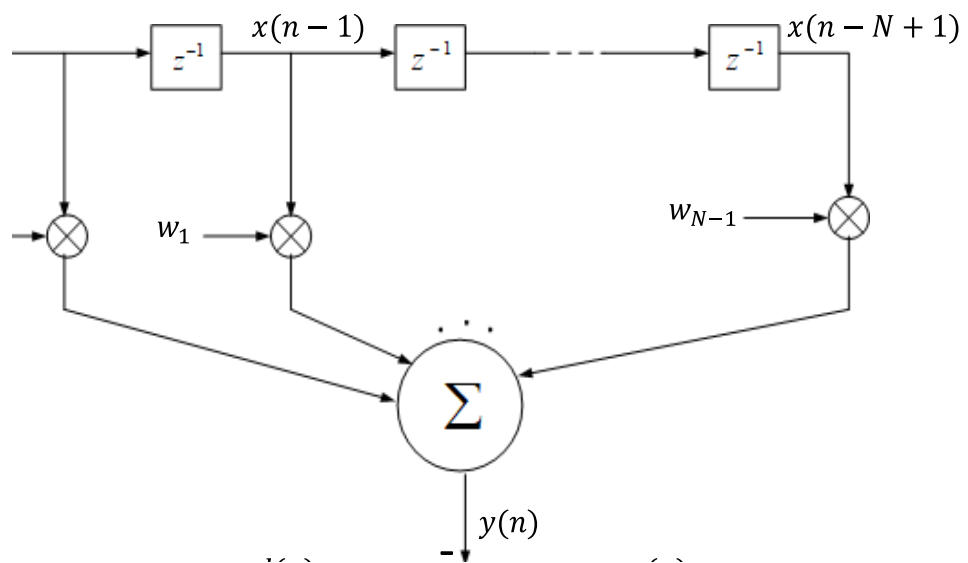
Figure 4.2. Block diagram of transversal FIR filter
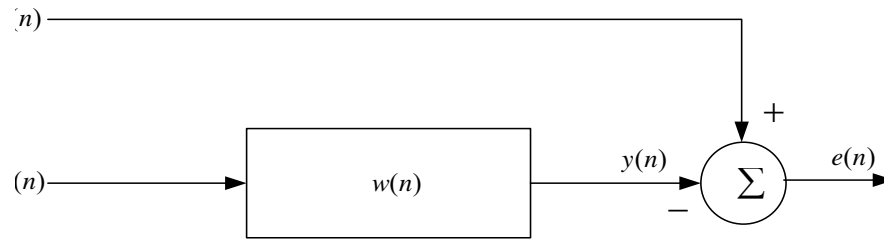
Figure 4.3. Block diagram of optimal filter

The most common criteria used to optimize the objective function of the optimal filter are:

i)   Minimum mean square error (MSE) criterion, which finds the filter coefficients **w** that give the minimum value of the average of the square of the error signal, usually denoted by the expectation of the squared signal as $E\{e^2(n)\}$. This corresponds to the Wiener filtering problem.

ii) Least-squares criterion, which finds the filter coefficients **w** that gives the minimum value of the sum of squared errors up to the present sample, i.e. $\sum_{i=0}^{N-1} e^2(i)$.

iii) Least absolute sum criterion that finds the filter coefficients **w** that give the minimum of the sum of absolute errors of the error signal given by $\sum_{i=0}^{N-1} |e(i)|$.

iv) Minimum of the maximum absolute-error criterion, which determines the filter coefficients **w** that yield the minimum of the maximum value of $(|e(i)|)$. This corresponds to the optimal Chebyshev filter design problem.

v)   Minimum mean absolute error that finds **w**, which minimizes $E\{|e(n)|\}$.

The first and the second optimization criteria are the most commonly used standards because of their relatively low computation, robust performance and analytical tractability that leads to closed form solutions.

4.2.1 Wiener Filtering

The optimal Wiener filter [117] is the basis of adaptive filtering and is the fundamental filter that most adaptation algorithms attempt to achieve. Wiener filtering aims to minimize the MSE objective function

$$f(\mathbf{w}) = E\{e^2(n)\}$$

$$= E\{[d(n) - y(n)]^2\} \tag{4.40}$$

$$= E\{d^2(n) - 2d(n)y(n) + y^2(n)\}$$

with respect to the tap weights $\mathbf{w}$.

Since $y(n)$ is a sum of products, and therefore a scalar quantity, from equation (4.39), the term $y^2(n)$ can be written as

$$y^2(n) = \mathbf{w}^T \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w} \tag{4.41}$$

Substituting for $y^2(n)$ from (4.41) in equation (4.40) and using the fact that the expectation operator is linear and distributive over the terms in the curly brackets, the MSE function becomes

$$f(\mathbf{w}) = E\{d^2(n)\} - 2E\{d(n)\mathbf{w}^T\mathbf{x}(n)\} + E\{\mathbf{w}^T\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}\} \tag{4.42}$$

Because the filter coefficient vector is not a random variable, the above equation can be rewritten as

$$f(\mathbf{w}) = E\{d^2(n)\} - 2\mathbf{w}^T\mathbf{p} + \mathbf{w}^T\mathbf{R}\mathbf{w} \tag{4.43}$$

where $\mathbf{p}$ is the correlation vector of the desired signal and the input signal, and $\mathbf{R}$ is the input signal correlation matrix, with the two defined as follows:

$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\}$$

$$= [p_0 \ p_1 \cdots p_{N-1}]^T \tag{4.44}$$

and

$$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$$

$$= E\left\{\begin{bmatrix} r_{0,0} & r_{0,1} & \cdots & r_{0,N-1} \\ r_{1,0} & r_{1,1} & \cdots & r_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N-1,0} & r_{N-1,1} & \cdots & r_{N-1,N-1} \end{bmatrix}\right\} \tag{4.45}$$

where each element of $\mathbf{p}$ is given by $p_i = d(n)x(n-i)$ and that of $\mathbf{R}$ is given by $r_{i,j} = x(n-i)x(n-j)$ for integers i and j in the range 0 to $N-1$.

In equation (4.43), the second and the third terms are linear and quadratic with $\mathbf{w}$. This makes the MSE function to be quadratic and therefore a convex function. The optimal minimum point of the function can be found by taking the gradient with respect

to $\mathbf{w}$ and setting it to zero. The gradient can be found by partially differentiating the function $f(\mathbf{w})$ with respect to the tap-weights and can be given as a vector

$$\nabla f(\mathbf{w}) = -2\mathbf{p} + 2\mathbf{R}\mathbf{w} \tag{4.46}$$

where the gradient operator ∇ defined as

$$\nabla = \left[\frac{\partial}{\partial w_0} \; \frac{\partial}{\partial w_1} \cdots \frac{\partial}{\partial w_{N-1}} \right]^T \tag{4.47}$$

Setting the gradient in equation (4.46) to zero, gives the famous *Wiener-Hopf* equation

$$\mathbf{R}\mathbf{w}_{opt} = \mathbf{p} \tag{4.48}$$

where the subscript opt denotes optimal coefficients. From this equation and assuming that the inverse of $\mathbf{R}$ exists, the optimum Wiener filter, which yields minimum MSE can be obtained as follows:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{p} \tag{4.49}$$

By substituting the optimal tap-weights in equation (4.43), the minimum MSE is obtained as

$$\begin{aligned} f(\mathbf{w}_{opt}) &= E\{d^2(n)\} - \mathbf{w}_{opt}^T\mathbf{p} \\ &= E\{d^2(n)\} - \mathbf{w}_{opt}^T\mathbf{R}\mathbf{w}_{opt} \end{aligned} \tag{4.50}$$

Although theoretically, the Wiener filter is the optimal solution for the minimum MSE function, it is not suitable for practical implementation for two main reasons. First, computing the inverse of the correlation matrix, i.e. $\mathbf{R}^{-1}$, is computationally intensive and therefore not appealing to most real-time applications. Secondly, the Wiener filter design is based on statistical formulation of $\mathbf{R}$ and $\mathbf{p}$, thus it requires a prior knowledge of the mean and correlation statistics of the signals being processed in order to compute the filter coefficients.

For a reliable estimation of the required statistics, a large number of realizations of signal sequences are required. This is not feasible because practical situations deal with single realizations of signals. Therefore, $\mathbf{R}$ and $\mathbf{p}$ are not available and have to be somehow estimated, for example, by assuming that the underlying signals are ergodic and therefore stationary with the same ensemble and time averages. This assumption can then allow time averages to be used in the Wiener filter design.

Nevertheless, the estimation of the required statistics can be avoided by using an adaptation algorithm that updates the filter coefficients iteratively while attempting to attain the optimal Wiener solution with a minimum number of computations per iteration. In particular, the iterative adaptation process has been found to be more advantageous than the non-iterative approach for the following main reasons [117], [118]:

i)   The iterative approach does not require a lot of memory like that needed in the optimal design for storage of signal samples for computing the signal statistics.

ii)   The iterative design approach eliminates the processing delay in the optimal design that comes from the accumulation of signal samples for computing the required signal statistics.

iii)   Using an iterative adaptation enables the filter to have a tracking ability and therefore is able to work in a non-stationary environment because the solution provided can adapt to new signal statistics.

iv)   In general, it is easier to code and/or implement an iterative design in hardware than a corresponding non-iterative design.

4.2.2 Gradient-Based Iterative Search Algorithms for Wiener Filter

In the last subsection, it was shown that the optimal coefficients of a transversal Wiener filter can be obtained by solving the *Wiener–Hopf* equation (4.49) provided that $\mathbf{R}$ and $\mathbf{p}$ of the signals to be processed are available. However, the computation of the wiener filter required the inversion of $\mathbf{R}$, which renders the equation hard to implement in real time.

A better alternative approach is to use an iterative search algorithm to approximate the Wiener filter iteratively by adjusting the filter coefficients vector $\mathbf{w}$ at each iteration k, until the resulting solution $\mathbf{w}(k)$ converges to the optimal solution $\mathbf{w}^*$, possibly in a few iterations. An iterative search algorithm starts with an arbitrary initial tap-weight vector $\mathbf{w}(k = 0)$ and progressively moves towards the optimum vector in steps through adjusting of the filter coefficients at each of the iterations. Each step is chosen so that the value of the cost function is reduced at the next iteration.

If the cost function is convex, like the MSE function in equation (4.43), the iterative adaptation algorithm is guaranteed to converge to the optimal minimum point of the function at the optimum filter tap-weights. Some common gradient-based iterative

search algorithms for approximating the Wiener filter will now be presented. They include the Newton's, steepest-descent, least mean-square error (LMS) and recursive least-squares (RLS) algorithms.

A. Newton's Algorithm

This algorithm has a fast speed of convergence to the Wiener filter. The update equation [118] for the tap-weights at the kth iteration is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \frac{\mu}{2}\mathbf{R}^{-1}\nabla f\big(\mathbf{w}(k)\big) \tag{4.51}$$

where the step-size $\mu\mathbf{R}^{-1}$ is matrix-based, and $\nabla f\big(\mathbf{w}(k)\big)$ is the gradient vector at the kth iteration given by

$$\nabla f\big(\mathbf{w}(k)\big) = 2\mathbf{R}\mathbf{w}(k) - 2\mathbf{p} \tag{4.52}$$

On substituting for $\nabla f\big(\mathbf{w}(k)\big)$ in equation (4.51), the filter adaptation equation becomes

$$\mathbf{w}(k+1) = (1 - \mu)\mathbf{w}(k) + \mu\mathbf{R}^{-1}\mathbf{p} \tag{4.53}$$

which by using $\mathbf{p} = \mathbf{R}\mathbf{w}_{opt}$ from equation (4.48), it further reduces to

$$\mathbf{w}(k+1) = (1 - \mu)\mathbf{w}(k) + \mu\mathbf{w}_{opt} \tag{4.54}$$

Subtracting $\mathbf{w}_{opt}$ from both sides of equation (4.54) gives the adaptation equation,

$$\mathbf{w}(k+1) - \mathbf{w}_{opt} = (1 - \mu)\big(\mathbf{w}(k) - \mathbf{w}_{opt}\big) \tag{4.55}$$

Defining a filter coefficients error vector

$$\boldsymbol{\varepsilon}(k) = \mathbf{w}(k) - \mathbf{w}_{opt} \tag{4.56}$$

and on substituting it in equation (4.55), the filter adaptation equation becomes

$$\mathbf{w}(k+1) = \mathbf{w}_{opt} + (1 - \mu)\boldsymbol{\varepsilon}(k) \tag{4.57}$$

If at the initial point, i.e. at $k = 0$, the initial coefficients error vector,

$$\boldsymbol{\varepsilon}(0) = \mathbf{w}(0) - \mathbf{w}_{opt} \tag{4.58}$$

then at the kth iteration, the filter coefficients vector is given by

$$\mathbf{w}(k) = \mathbf{w}_{opt} + (1 - \mu)^k\boldsymbol{\varepsilon}(0) \tag{4.59}$$

Similarly, the error-performance function at the kth iteration is given by

$$f\big(\mathbf{w}(k)\big) = f(\mathbf{w}_{opt}) + (1 - \mu)^{2k} f(\mathbf{w}(0)) \tag{4.60}$$

where $f\big(\mathbf{w}(k)\big)$ denotes the value of the MSE function for filter coefficients $\mathbf{w}(k)$.

Equations (4.59) and (4.60) show that the convergence of the Newton's algorithm to the Wiener filter $\mathbf{w}_{opt}$ is solely determined by the step-size parameter μ. In order to guarantee for the convergence and the stability of the algorithm, the step-size parameter should be selected to satisfy the condition $|1 - \mu| < 1$, or equivalently put

$$0 < \mu < 2 \tag{4.61}$$

The conventional Newton's method has a step-size $\mu = 1.0$ in order to make the algorithm converge to $\mathbf{w}_{opt}$ in a single step. Despite this fast convergence, the algorithm must perform the computationally intensive inversion operation on matrix $\mathbf{R}$, which may render it prohibitive for real-time applications.

B. Steepest-Descent Algorithm

This algorithm avoids the inversion operation performed on matrix $\mathbf{R}$ in the Newton's algorithm by replacing the matrix-based step-size $\mu\mathbf{R}^{-1}$ with just a scalar step-size parameter [119]. The algorithm starts at an arbitrary point on the error-performance surface of the objective function and then takes a small step in the direction in which the objective function decreases fastest. At k^{th} iteration, the filter coefficients are updated as follows:

$$\mathbf{w}(k + 1) = \mathbf{w}(k) - \frac{\mu}{2}\nabla f\big(\mathbf{w}(k)\big) \tag{4.62}$$

where μ is a scalar step-size parameter greater than zero and $\nabla f\big(\mathbf{w}(k)\big)$ is the gradient of the objective function at the iteration. This update is repeated until convergence to the minimum value of the objective function at which $\mathbf{w}(k + 1) = \mathbf{w}_{opt}$.

The stepwise search at each iteration corresponds to a step in the direction along the steepest-descent slope of the error-performance surface, i.e. following the opposite direction to the gradient vector. This is illustrated in Figure 4.4 for the case of a two-tap adaptive filter. The rate of convergence of the algorithm towards $\mathbf{w}_{opt}$ depends on the value of the step-size parameter.

A small step-size leads to a slow convergence, while a large one makes the adaptation algorithm to converge within a few steps although it may also cause the recursive equation in (4.62) to diverge. To investigate the effect of the step-size on the

convergence rate of the algorithm, equation (4.46) is substituted in equation (4.62) to yield

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\mathbf{R}\mathbf{w}(k) + \mu\mathbf{p} \qquad (4.63)$$

This equation can be re-arranged to have

$$\mathbf{w}(k+1) = (\mathbf{I} - \mu\mathbf{R})\mathbf{w}(k) + \mu\mathbf{p} \qquad (4.64)$$

where $\mathbf{I}$ is an $N \times N$ identity matrix.

Substituting for $\mathbf{p}$ from equation (4.48) and subtracting $\mathbf{w}_{opt}$ from both sides of equation (4.64) gives

$$\mathbf{w}(k+1) - \mathbf{w}_{opt} = (\mathbf{I} - \mu\mathbf{R})\mathbf{w}(k) + \mu\mathbf{R}\mathbf{w}_{opt} - \mathbf{w}_{opt} \qquad (4.65)$$

which reduces to

$$\mathbf{w}(k+1) - \mathbf{w}_{opt} = (\mathbf{I} - \mu\mathbf{R})(\mathbf{w}(k) - \mathbf{w}_{opt}) \qquad (4.66)$$

By using the definition of the coefficients error vector $\boldsymbol{\varepsilon}(k)$ in equation (4.56), equation (4.66) can be expressed in the form

$$\boldsymbol{\varepsilon}(k+1) = (\mathbf{I} - \mu\mathbf{R})\boldsymbol{\varepsilon}(k) \qquad (4.67)$$
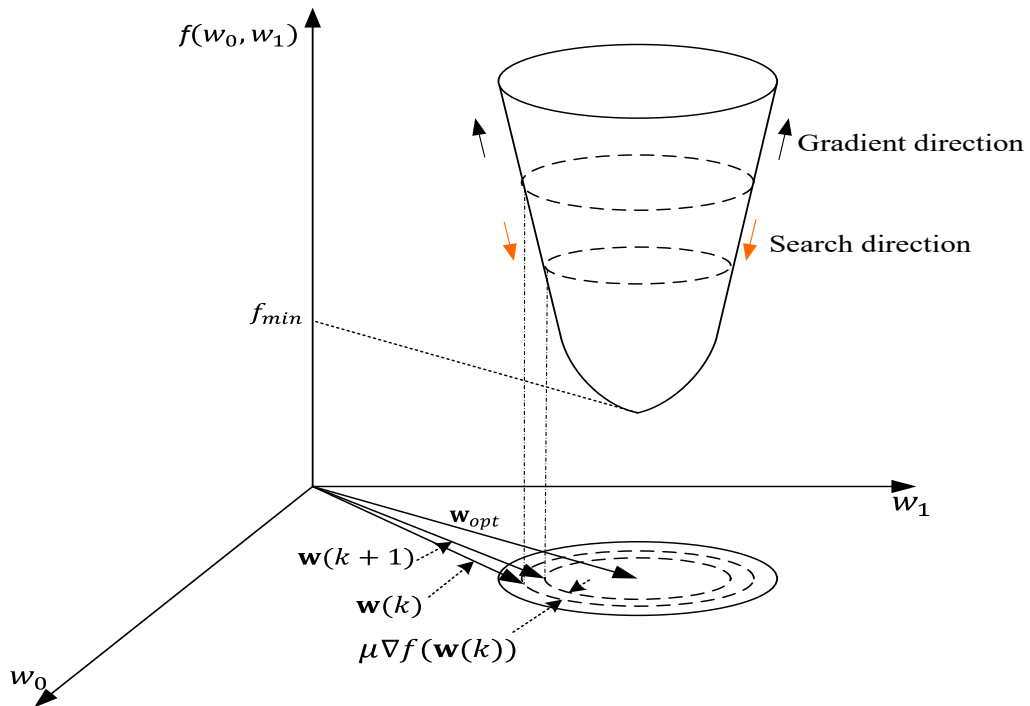
Figure 4.4. Error-performance surface of steepest descent algorithm

Because the correlation matrix $\mathbf{R}$ is Hermitian, i.e. $\mathbf{R}^H = \mathbf{R}$, it can be diagonalised or equivalently transformed using a unitary matrix $\mathbf{U}$, having the property $\mathbf{UU}^T = \mathbf{I}$, as follows

$$\mathbf{R} = \mathbf{U\Lambda U}^T \tag{4.68}$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_0 \ \lambda_1 \ ... \ \lambda_{N-1})$ is a diagonal matrix containing the eigenvalues λ_i ($i = 0, 1, N - 1$) of $\mathbf{R}$ in the diagonal and $\mathbf{U}$ is a unitary matrix containing the corresponding eigenvectors in the columns.

Now, substituting equation (4.68) in equation (4.67), the coefficients error vector changes according to

$$\mathbf{\varepsilon}(k + 1) = (\mathbf{I} - \mu \mathbf{U\Lambda U}^T)\mathbf{\varepsilon}(k) \tag{4.69}$$

which by pre-multiplying with $\mathbf{U}^T$ on both sides and using $\mathbf{UU}^T = \mathbf{I}$ becomes

$$\mathbf{U}^T\mathbf{\varepsilon}(k + 1) = \mathbf{U}^T\mathbf{\varepsilon}(k) - \mu \mathbf{\Lambda U}^T\mathbf{\varepsilon}(k) \tag{4.70}$$

By defining $\hat{\mathbf{\varepsilon}}(k)$ to represent the transformed filter coefficients error vector as

$$\hat{\mathbf{\varepsilon}}(k) = \mathbf{U}^T\mathbf{\varepsilon}(k) \tag{4.71}$$

equation (4.70) can be written as follows:

$$\hat{\mathbf{\varepsilon}}(k + 1) = (\mathbf{I} - \mu \mathbf{\Lambda})\hat{\mathbf{\varepsilon}}(k) \tag{4.72}$$

Each element of the transformed error vector in this equation is given by

$$\hat{\varepsilon}_i(k + 1) = (1 - \mu \lambda_i)\hat{\varepsilon}_i(k), \qquad i = 0, 1, ..., N - 1 \tag{4.73}$$

Since both μ and λ_i are constants, if the initial i^{th} element in the transformed coefficients error vector is $\hat{\varepsilon}_i(0)$ at $k = 0$, then at the kth iteration it becomes

$$\hat{\varepsilon}_i(k) = (1 - \mu \lambda_i)^k \hat{\varepsilon}_i(0), \qquad i = 0, 1, ..., N - 1 \tag{4.74}$$

Similarly, the filter weights and the error-performance function at the kth iteration relative to the optimal point are given by

$$\mathbf{w}(k) = \mathbf{w}_{opt} + \sum_{i=0}^{N-1} \hat{\varepsilon}_i(k) \, \mathbf{v}_i \tag{4.75}$$

and

$$f\big(\mathbf{w}(k)\big) = f(\mathbf{w}_{opt}) + \sum_{i=0}^{N-1} \lambda_i \hat{\varepsilon}_i^2(k) \tag{4.76}$$

respectively, where $\mathbf{v}_i$ is the eigenvector of $\mathbf{R}$ corresponding to the eigenvalue λ_i for $i = 0, 1, \dots, N - 1$.

Equations (4.74) and (4.75) imply that the filter coefficients $\mathbf{w}(k)$ will converge to the optimal values $\mathbf{w}_{opt}$ if and only if $\hat{\varepsilon}_i(k)$ converges to zero, as k increases towards infinity. This is only possible if the step-size parameter satisfies the condition

$$|1 - \mu\lambda_i| < 1, \qquad i = 0, 1, \dots, N - 1 \tag{4.77}$$

In addition, this condition is a guarantee for the stability of the algorithm since the error vector decays exponentially towards zero as k increases.

Again, since $\mathbf{R}$ is Hermitian, it is also positive semidefinite and all the eigenvalues are real and nonnegative, i.e. $\lambda_i \geq 0$, $i = 0, 1, \dots, N - 1$. Therefore, the condition in equation (4.77) can equivalently be written as

$$0 < \mu < \frac{2}{\lambda_i}, \qquad i = 0, 1, \dots, N - 1 \tag{4.78}$$

Since this condition must be satisfied for all the eigenvalues, the condition for the convergence of the steepest-descent algorithm requires the step-size to be set as follows:

$$0 < \mu < \frac{2}{\lambda_{max}} \tag{4.79}$$

where λ_{max} is the largest eigenvalue of $\mathbf{R}$.

From the foregoing discussion, it has been established that the speed of convergence of the steepest-descent algorithm is highly influenced by the range of the eigenvalues of the correlation matrix. In equation (4.77), each λ_i can be seen as contributing to one mode of convergence; the fastest convergence being achieved when $\lambda_i = \lambda_{max}$ and the slowest at $\lambda_i = \lambda_{min}$. For each mode of convergence, the fastest convergence would be achieved if the step-size were selected as

$$\mu = \frac{1}{\lambda_i} \tag{4.80}$$

so as to have $|1 - \mu\lambda_i| = 0$ and thereby achieve convergence in one iteration.

However, for the whole algorithm, the speed of convergence is determined by the smallest eigenvalue and can be indicated by the factor

$$\xi = 1 - \mu\lambda_{min} \tag{4.81}$$

for a given step-size parameter μ.

Considering the full range of the eigenvalues, the optimal step-size can be found at the point where the two extreme lines; $|1 - \mu\lambda_{min}|$ and $|1 - \mu\lambda_{max}|$, corresponding to the minimum and maximum eigenvalues meet. The two extreme convergences are illustrated in Figure 4.5. In this figure, all other convergence modes will fall between the two lines. At the optimal step-size μ_{opt}, the two lines satisfies the condition

$$(1 - \mu_{opt}\lambda_{min}) = -(1 - \mu_{opt}\lambda_{max}) \tag{4.82}$$

Thus, yielding the optimal step-size as

$$\mu_{opt} = \frac{2}{\lambda_{min} + \lambda_{max}} \tag{4.83}$$

Substituting equation (4.83) in equation (4.81) gives ξ as a function of the eigenvalue spread $a = \lambda_{max}/\lambda_{min}$, i.e.

$$\xi = \frac{a - 1}{a + 1} \tag{4.84}$$

Based on this equation, the convergence factor ξ can only have values between 0 and 1. It is equal to 0 when $\lambda_{max} = \lambda_{min}$, which means that the algorithm converges to the optimal solution in one step. On the other hand, a value of 1, corresponding to the case when $\lambda_{max} \gg \lambda_{min}$, implies a very slow convergence.
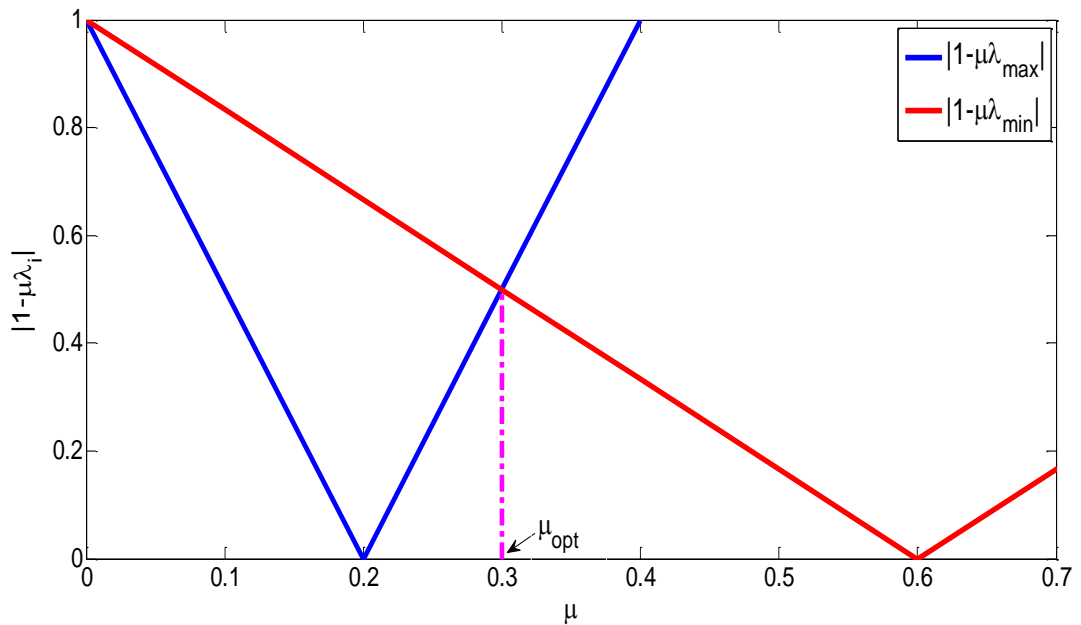
Figure 4.5. Optimal step-size of the steepest descent algorithm when $\lambda_{min} = 1.67$ and $\lambda_{max} = 5.00$

The foregoing analysis shows that the steepest-descent algorithm has a good performance only when the eigenvalue spread is small. This corresponds to a signal having a nearly flat power spectral density similar to the white noise.

C. Least Mean Square Algorithm

The least mean square (LMS) algorithm, proposed by Widrow and Hoff in 1960 [120], is in practice the most widely used adaptive filtering algorithm. This is due to its simplicity in implementation and low computational complexity. The algorithm simply approximates the steepest-descent algorithm by replacing the MSE function with instantaneous squared-error (ISE) function. The ISE cost function is given by

$$f(\mathbf{w}) = e^2(n)$$

$$= [d(n) - y(n)]^2 \tag{4.85}$$

$$= d^2(n) - 2d(n)y(n) + y^2(n)$$

On substituting equations (4.36) and (4.41) in (4.85), the objective function becomes

$$f(\mathbf{w}) = d^2(n) - 2d(n)\mathbf{w}^T\mathbf{x}(n) + \mathbf{w}^T\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w} \tag{4.86}$$

The gradient of this ISE function with respect to filter coefficients is

$$\nabla f(\mathbf{w}) = -2d(n)\mathbf{x}(n) + 2\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}$$

$$= -2\,\mathbf{x}(n)e(n) \tag{4.87}$$

This derivation shows that the LMS algorithm uses the estimates of $\mathbf{p}$ and $\mathbf{R}$, from equations (4.44) and (4.45) of the steepest-descent algorithm, as follows:

$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\} \cong d(n)\mathbf{x}(n) \tag{4.88}$$

and

$$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$$

$$\cong \mathbf{x}(n)\mathbf{x}^T(n) \tag{4.89}$$

Using equation (4.87), the per iteration filter adaptation equation is given by

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{\mu}{2}\,\nabla f(\mathbf{w}) \tag{4.90}$$

and thus

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\,\mathbf{x}(n)e(n) \tag{4.91}$$

where μ is the step-size parameter.

Considering equation (4.91), the simplicity of implementing the LMS algorithm can be easily determined. For a filter with N taps, the equation requires $N + 1$ multiplications and N additions. The computation of the filter output $y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$ requires N multiplications and $N - 1$ additions, while the error $e(n) = d(n) - y(n)$ needs only one addition. In total, the algorithm requires $2N + 1$ multiplications and $2N$ additions. Since additions demand less computations than multiplications, the computational complexity of the algorithm is $2N + 1 \cong 2N$, i.e. $O(N)$ and hence, it is proportional to the length of the filter.

Although the LMS algorithm has much lower computational complexity compared to the steepest-descent algorithm, it has the same problem of slow convergence, which is dependent on the eigenvalue spread of the underlying input signals. The convergence of the algorithm can be analysed starting with the subtraction of the optimal filter coefficients from both sides of equation (4.90) to give

$$\mathbf{w}(n+1) - \mathbf{w}_{opt} = \mathbf{w}(n) - \mathbf{w}_{opt} + \mu\,\mathbf{x}(n)[d(n) - \mathbf{x}^T(n)\mathbf{w}(n)] \qquad (4.92)$$

Using the definition of the tap-weight error in equation (4.56), we have

$$\boldsymbol{\varepsilon}(n+1) = \boldsymbol{\varepsilon}(n) + \mu[\mathbf{x}(n)d(n) - \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n)] \qquad (4.93)$$

Taking the expectation of both sides yields gives

$$E\{\boldsymbol{\varepsilon}(n+1)\} = E\{\boldsymbol{\varepsilon}(n)\} + \mu[E\{\mathbf{x}(n)d(n)\} - E\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n)\}] \qquad (4.94)$$

At this point, it is assumed that at time n, the filter coefficients vector $\mathbf{w}(n)$ is independent of the input vector $\mathbf{x}(n)$ and the desired signal $d(n)$. This is referred to as the independence assumption and is practically valid for small values of μ. Using this assumption, equation (4.94) can be rewritten as follows:

$$E\{\boldsymbol{\varepsilon}(n+1)\} = E\{\boldsymbol{\varepsilon}(n)\} + \mu[\mathbf{p} - \mathbf{R}E\{\mathbf{w}(n)\}] \qquad (4.95)$$

which by substituting for $\mathbf{p} = \mathbf{R}\mathbf{w}_{opt}$ becomes

$$E\{\boldsymbol{\varepsilon}(n+1)\} = (\mathbf{I} - \mu\mathbf{R})E\{\boldsymbol{\varepsilon}(n)\} \qquad (4.96)$$

Except for the expectation operator, the above equation is similar to equation (4.67) that was used to derive the conditions for the convergence of the steepest-descent algorithm. Therefore, following the same analysis as before, the stability and the convergence of the LMS algorithm is guaranteed if the mean of the tap-weight error

vector $\boldsymbol{\varepsilon}(n)$ goes to zero as n is increased. This is possible only if the values of the step-size parameter lie in the range:

$$0 < \mu < \frac{2}{\lambda_{max}} \tag{4.97}$$

The tendency of the mean of the tap-weight error towards zero implies that the expected value of the tap-weights converges to the Wiener solution, i.e. $E\{\mathbf{w}(n)\} \rightarrow \mathbf{w}_{opt}$. This shows that the use of an LMS adaptation algorithm causes the adaptive filter to converge in the mean.

Practically, the eigenvalues of the correlation matrix $\mathbf{R}$ are not known beforehand; thus, equation (4.97) cannot be used to set the step-size parameter. A practical way is to estimate the upper limit of μ directly from the power of the input signal. This is justified by the following derivation that relates the sum of the eigenvalues of the correlation matrix to the power of the input signal.

From the transformation $\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ in equation (4.68), since $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, then $\mathbf{R}$ can be diagonalised using unitary transformation to give $\boldsymbol{\Lambda}$, i.e.

$$\mathbf{U}^T\mathbf{R}\mathbf{U} = \boldsymbol{\Lambda} \tag{4.98}$$

For any two matrices $\mathbf{A}$ and $\mathbf{B}$, the following holds on the trace, i.e. the sum of the diagonal elements, of their product:

$$\text{tr}\{\mathbf{A}\mathbf{B}\} = \text{tr}\{\mathbf{B}\mathbf{A}\} \tag{4.99}$$

Using this relationship, the trace of the matrix product on the left of equation (4.98) is given by

$$\text{tr}\{\mathbf{U}^T\mathbf{R}\mathbf{U}\} = \text{tr}\{\mathbf{U}^T(\mathbf{R}\mathbf{U})\}$$

$$= \text{tr}\left\{\mathbf{R}\underbrace{\mathbf{U}\mathbf{U}^T}_{\mathbf{I}}\right\} \tag{4.100}$$

$$= \text{tr}\{\mathbf{R}\}$$

Since $\boldsymbol{\Lambda} = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{N-1})$, then $tr\{\boldsymbol{\Lambda}\} = \sum_{i=0}^{N-1} \lambda_i$, which from equation (4.98) implies that

$$\text{tr}\{\mathbf{R}\} = \text{tr}\{\boldsymbol{\Lambda}\}$$

$$= \sum_{i=0}^{N-1} \lambda_i \tag{4.101}$$

But since for the LMS algorithm, $\mathbf{R} \cong \mathbf{x}(n)\mathbf{x}^T(n)$, then

$$
\begin{aligned}
\text{tr}\{\mathbf{R}\} &= x^2(n) + x^2(n-1) + \cdots + x^2(n-N+1) \\
&= \|\mathbf{x}(n)\|_2^2
\end{aligned}
\tag{4.102}
$$

and this is equal to the power of $\mathbf{x}(n)$ which is usually known or can be estimated *a priori*. Because λ_{max} is the maximum of the elements in the set $\{\lambda_0, \lambda_1, \ldots, \lambda_{N-1}\}$, it is obvious that

$$
\begin{aligned}
\lambda_{max} &\leq \sum_{i=0}^{N-1} \lambda_i \\
&= \|\mathbf{x}(n)\|_2^2
\end{aligned}
\tag{4.103}
$$

and

$$
\frac{1}{\lambda_{max}} \geq \frac{1}{\|\mathbf{x}(n)\|_2^2}
\tag{4.104}
$$

Therefore, for convergence of the LMS algorithm, the necessary limits for the step-size parameter in equation (4.97) can be set firmer and more realistic as follows:

$$
0 < \mu < \frac{2}{\|\mathbf{x}(n)\|_2^2}
\tag{4.105}
$$

A larger step-size leads to a faster convergence of the LMS algorithm but this occurs at the expense of increased misadjustment error [121], which is given by

$$
m_{LMS} = \sum_{k=0}^{N-1} \frac{\mu\lambda_k}{2 - \mu\lambda_k}
\tag{4.106}
$$

Therefore, the value of the step-size should be chosen in a way that allows a trade-off between the rate of convergence and the steady-state misadjustment error.

D. Recursive Least-Squares Algorithm

The recursive least-squares (RLS) filtering algorithm is based on the method of least-squares which is a deterministic framework for finding FIR filter tap-weights at time $n > 0$ that minimize the objective function of the sum of n weighted squared errors up to the current time, i.e.,

$$
f_n(\mathbf{w}) = \sum_{k=1}^{n} \gamma^{n-k} e^2(k)
\tag{4.107}
$$

where the weighting (or the forgetting) factor γ is limited to the range $0 < \gamma \leq 1$ and is used for emphasizing the contribution of the current data, rather than the past data, to the cost function.

In equation (4.107), $k = 1$ denotes the starting time of the algorithm, $k = n$ is the present time, and $e(k)$ denotes the error sample at any time instant k for $k = 1, 2, \ldots, n$. Therefore, the least-squares method optimizes filter coefficients using past and present statistics.

However, when $\gamma < 1$, the weights γ^{n-k}, $k = 1, 2, \ldots, n$, in equation (4.107) give more weightage to the most current error samples than the past samples. In other words, a value of $\gamma < 1$ emphasizes more on the recent samples, while tending to forget the past ones. The lower the value of γ, the more the emphasis on latest samples and vice versa.

The typical values of the γ are close to 1 [116]. However, the forgetting factor can be chosen with an aim to achieve a small misadjustment. The misadjustment error is a function of the forgetting factor and is given [117] by

$$m_{RLS} = \frac{1 - \gamma}{1 + \gamma} N \qquad (4.108)$$

where N is filter length. From this equation, the forgetting factor is found to be

$$\gamma = \frac{N - m_{RLS}}{N + m_{RLS}} \qquad (4.109)$$

Thus, for a given desired value of misadjustment and the number of filter coefficients, the value of the forgetting factor can be obtained using equation (4.109).

In order to find the RLS algorithm, the objective function in equation (4.107) is expressed in quadratic form using equation (4.34) for the error $e(k)$, as follows:

$$
\begin{aligned}
f_n(\mathbf{w}) = \sum_{k=1}^{n} \gamma^{n-k} & [d^2(k) - 2d(k)\mathbf{w}^T(n)\mathbf{x}(k) \\
& + \mathbf{w}^T(n)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(n)]
\end{aligned}
\qquad (4.110)
$$

$$= \sigma_d^2 - 2\mathbf{w}^T(n)\mathbf{p}(n) + \mathbf{w}^T(n)\mathbf{R}(n)\mathbf{w}(n)$$

where

$$\sigma_d^2(n) = \sum_{k=1}^{n} \gamma^{n-k} d^2(k) \qquad (4.111)$$

$$\mathbf{p}(n) = \sum_{k=1}^{n} \gamma^{n-k} d(k)\mathbf{x}(k) \tag{4.112}$$

$$\mathbf{R}(n) = \sum_{k=1}^{n} \gamma^{n-k} \mathbf{x}(k)\mathbf{x}^T(k) \tag{4.113}$$

The quadratic cost function in equation (4.110) has a gradient

$$\nabla f_n(\mathbf{w}) = -2\mathbf{p}(n) + 2\mathbf{R}(n)\mathbf{w}(n) \tag{4.114}$$

which is set to zero to yield the optimal solution for the weighted least-squares problem as

$$\mathbf{w}(n) = \mathbf{R}^{-1}(n)\mathbf{p}(n) \tag{4.115}$$

To avoid matrix inversion, a recursive least squares (RLS) algorithm is used to solve equation (4.115). The RLS algorithm avoids the intensive computation of $\mathbf{p}(n)$ and $\mathbf{R}(n)$ using equations (4.112) and (4.113) by adopting recursive updates

$$\begin{aligned}\mathbf{p}(n) &= \sum_{k=1}^{n} \gamma^{n-k} d(k)\mathbf{x}(k) \\ &= \gamma\mathbf{p}(n-1) + d(n)\mathbf{x}(n)\end{aligned} \tag{4.116}$$

and

$$\begin{aligned}\mathbf{R}(n) &= \sum_{k=1}^{n} \gamma^{n-k} \mathbf{x}(k)\mathbf{x}^T(k) \\ &= \gamma\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n)\end{aligned} \tag{4.117}$$

Substituting equation (4.116) into equation (4.115) gives a recursive update for the tap-weights

$$\begin{aligned}\mathbf{w}(n) &= \mathbf{R}^{-1}(k)[d(n)\mathbf{x}(n) + \gamma\mathbf{p}(n-1)] \\ &= \mathbf{R}^{-1}(k)[d(n)\mathbf{x}(n) + \gamma\mathbf{R}(n-1)\mathbf{w}(n-1)] \\ &= \mathbf{R}^{-1}(k)[d(n)\mathbf{x}(n) - \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n-1) \\ &\quad + \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{w}(n-1) + \gamma\mathbf{R}(n-1)\mathbf{w}(n-1)] \\ &= \mathbf{R}^{-1}(k)\{ \mathbf{x}(n)[d(n) - \mathbf{x}^T(n)\mathbf{w}(n-1)] + \mathbf{R}(n)\mathbf{w}(n-1)\}\end{aligned} \tag{4.118}$$

$$= \mathbf{w}(n-1) + \mathbf{R}^{-1}(k)\mathbf{x}(n)\hat{e}(k)$$

where

$$\hat{e}(k) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n-1) \qquad (4.119)$$

is called *a priori* error estimate.

A recursive formula for finding the matrix inverse $\mathbf{R}^{-1}(n)$ can be derived from equation (4.117) by using the following lemma for inverting a matrix [121]:

$$[\mathbf{A} + \mathbf{BCD}]^{-1} = \frac{\mathbf{A}^{-1}\mathbf{BDA}^{-1}}{\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B}} \qquad (4.120)$$

By setting $\mathbf{A} = \gamma\mathbf{R}(n-1)$, $\mathbf{B} = \mathbf{x}(n)$, $\mathbf{C} = 1$ and $\mathbf{D} = \mathbf{x}^T(n)$, the inverse $\mathbf{R}^{-1}(n)$ is obtained as follows:

$$\mathbf{R}^{-1}(n) = \frac{\mathbf{R}^{-1}(n-1)}{\gamma + \mathbf{x}^T(n)\mathbf{R}^{-1}(n-1)\mathbf{x}(n)} \qquad (4.121)$$

The basic RLS algorithm starts at $n = 0$ with the initial filter weights $\mathbf{w}(0) = 0$ and the inverse $\mathbf{R}^{-1}(0) = (1/\delta)\mathbf{I}$, where δ is a small positive constant. At each iteration n the algorithm computes the product $\mathbf{R}^{-1}(n)$ and the error $\hat{e}(k)$ using equations (4.121) and (4.119), then updates the tap-weights using equation (4.118). This is repeated until $n \geq N$ in order to allow for the complete formation of matrix $\mathbf{R}$ and vector $\mathbf{p}$ and for convergence to be achieved, which occurs when the mean of tap-weights equals to the optimal tap-weights.

CHAPTER 5

MATERIALS AND METHODS

This chapter presents the materials that were used to carry out the research work in this thesis and an in-depth coverage of the methods that were developed for reducing PAPR in OFDM systems. The chapter has two main sections: Section 5.1 and Section 5.2. The chapter starts with Section 5.1, which gives key materials used during the research period. This is followed by Section 5.2, which opens with the description of tone-reservation concept in relation to PAPR reduction in OFDM systems. The second part of Section 5.2 presents five PAPR reduction methods that were originated during the course of the research together with an analysis of their computational complexities.

5.1 Materials

The list of key materials used during the course of this work especially for simulating OFDM systems and for programming and testing of the proposed methods is as follows:

i) Dell desktop OptiPlex 7010, Windows 10 Professional, Intel core i5-3470 @3.20 GHz 3.20 GHz, 4.00 GB RAM, 64-bit operating system, x64-based processor.

ii) Math Works, MATLAB version 8.1.0.604 (R2013a), for 64-bit Windows operating system, February 15, 2013 [122]. This is an interactive, matrix-based system for scientific and engineering numeric computation and visualization. In addition to the basic MATLAB product, the three toolboxes for signal processing, communications system and optimization, all of which have inbuilt functions for digital signal processing, were used to develop and test the proposed algorithms during the research period.

iii) CVX MATLAB software for disciplined convex programming, version 2.2, which was available at http://cvxr.com/cvx by January 2020. CVX is a software package for modelling convex optimization problems in MATLAB. The package allows MATLAB to act as a modelling language, thus allowing cost

functions and constraints to be specified using the standard syntax of MATLAB.

5.2 Proposed PAPR Reduction Methods

In this section, a detailed description of the tone-reservation concept and the methods developed at different stages of the research are presented. These are:

i) SOCP-based tone reservation method

ii) LP-based tone reservation

iii) IRLS-based tone reservation method

iv) Low-complexity signal addition method

v) Low-complexity additive signal mixing method

The first three methods directly utilise the tone reservation concept where some frequency resources are reserved for use in the reduction of PAPR. The remaining two methods also apply a similar technique of reserving some transmission resources for PAPR Reduction but they do so in the time-domain. In addition, both the first and the second methods seek an optimal solution to the minimax problem but through solving equivalent SOCP and LP problems, respectively.

Therefore, the first two methods are expected to have slow convergence rates and high computational complexities. However, the second method has a lower computational complexity than the first one because of a simple generation of the desired peak-reduction signal for a given OFDM signal, which is then estimated by the optimization process. The two optimal methods, (i) and (ii), were needed for providing results for benchmarking the performances of methods (iii) – (v), which utilise suboptimal schemes to find peak-reduction signal.

The third method was built around the second technique but utilizing a fast IRLS algorithm to solve the minimax problem to find the peak-reduction signal. The development of the last two methods was motivated by the need to reduce further the computational complexity and convergence rate, while at the same time aiming at achieving similar or higher PAPR reduction performance than the third technique.

In the following sections, the general tone reservation concept upon which the proposed methods are founded is presented followed by a detailed description of the proposed methods.

5.2.1 Tone Reservation Concept

The concept of reserving some tones to carry peak-reduction signal for reducing PAPR in multicarrier transmission systems was first proposed by Tellado and Cioffi in 1997 [123]. The concept is just about allocating a small number of subcarriers to carry a combination of coefficients that are converted via an IFFT operation into a peak-reduction signal. This means that if there N total number of subcarriers available for transmitting user data, L subcarriers out of that, where $L \ll N$, are assigned for carrying peak-reduction coefficients.

Those subcarriers reserved for transmitting peak-reduction coefficients are referred to as peak-reduction tones (PRTs). Similarly, the coefficients carried by PRTS are referred to as peak-reduction coefficients (PRCs). The time-domain signal arising from the modulation of the PRTS by the PRCs, found by applying IFFT on the PRCs, is the peak-reduction signal (PRS) that is used to reduce the PAPR of OFDM signal.

In Figure 5.1, the concept of tone reservation is illustrated. The inputs to the upper N-IFFT block are the modulation symbols, $X(i)$, $i = 0, 1, \ldots, N - 1$, which are all non-zeros except in the positions of the L- reserved subcarriers. Similarly, the inputs to the lower N-IFFT block are the PRCs, $C(i)$, $i = 0, 1, \ldots, N - 1$, which are all non-zeros except in the $(N - L)$ positions of subcarriers for carrying user data.

Since all PRCs are placed on subcarriers different from those for carrying user data, the transmitted modulation symbols $X(i)$, $i = 0, 1, \ldots, N - 1$, can be extracted from the received signal by FFT without any distortion. This is can be accomplished by choosing the set of outputs corresponding to the $(N - L)$ positions of data-bearing subcarriers in the FFT output.
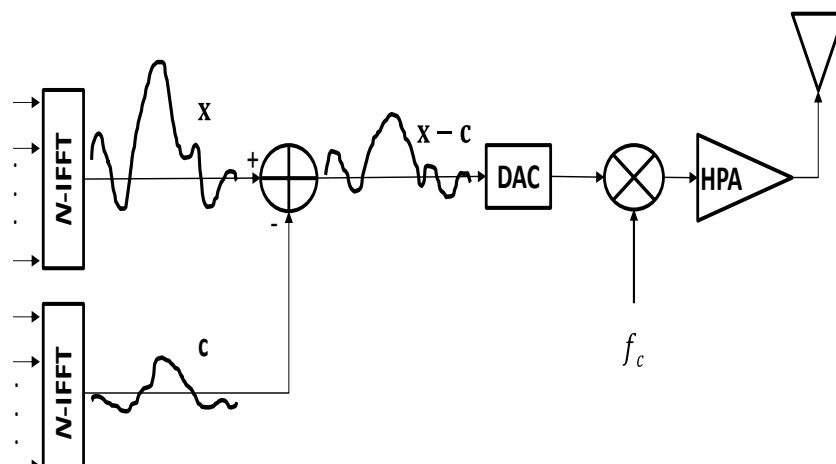
Figure 5.1. Tone reservation concept

101

The reservation of L subcarriers in an OFDM symbol can be done in three ways, namely in a contiguous block, equally spaced or in a random manner. However, for the random allocation, the information about the locations of reserved subcarriers must be sent to the receiver for correct demodulation of the data-bearing modulation symbols.

To reduce the PAPR of the OFDM signal, it is necessary to select carefully the PRCs whose combination with the available data symbols, through an IFFT operation, yields a signal of lower PAPR than when the latter are transmitted alone. This in the time domain implies that with a good selection of PRCs, the PAPR of signal $(\mathbf{x} - \mathbf{c})$ from the IFFT block in the transmitter should be less than the PAPR of the input signal $\mathbf{x}$.

The IFFT of the combined signal can be described by the equation

$$\mathbf{x} - \mathbf{c} = \mathbf{x} - \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} C_k e^{\frac{j2\pi kn}{N}} \tag{5.1}$$

This equation can equivalently be written as

$$\mathbf{x} - \mathbf{c} = \mathbf{x} - Q\mathbf{C} \tag{5.2}$$

with $\mathbf{x} \in \mathbb{C}^N$, $\mathbf{c} \in \mathbb{C}^N$, $\mathbf{C} \in \mathbb{C}^N$ and $Q \in \mathbb{C}^{N \times N}$ is the IFFT matrix whose elements are given by $\left(1/\sqrt{N}\right) \exp(j2\pi kn/N)$ and k and n are integers in the range 0 to N-1.

The time-domain vector $\mathbf{x} = [x(0), x(1), \dots, x(N-1)]^T$ contains the time-samples of the OFDM signal while $\mathbf{C} = [C(0), C(1), \dots, C(N-1)]^T$ is a frequency-domain vector whose nonzero components comprises the PRCs. Since, vector $\mathbf{C}$ contains $(N - L)$ zeros, the complexity in the computation of the discrete-time signal $\mathbf{c}$ can be significantly reduced if only the nonzero components of $\mathbf{C}$ and the L columns of the IFFT matrix Q corresponding to the reserved subcarriers are used.

Accordingly, letting a vector $\hat{\mathbf{C}} \in \mathbb{C}^L$ to contain only the nonzero elements of $\mathbf{C}$ and a matrix $\hat{Q} \in \mathbb{C}^{N \times L}$ to be made up of only the L columns of Q corresponding to locations of reserved subcarriers, the peak-reduction signal can be expressed as

$$\mathbf{c} = \hat{Q}\hat{\mathbf{C}} \tag{5.3}$$

Following this development, the problem of finding the PRCs that result in PAPR reduction can be framed as a *minimax* optimization problem of the form

$$\min_{\hat{\mathbf{C}}} \max\left(\left|\mathbf{x} - \hat{Q}\hat{\mathbf{C}}\right|\right) \tag{5.4}$$

The optimization variable for this problem is $\hat{\mathbf{C}} \in \mathbb{C}^L$ and the problem parameters are

102

$\mathbf{x} \in \mathbb{C}^N$ and $\hat{Q} \in \mathbb{C}^{N \times L}$.

The effectiveness of a method founded on the tone-reservation concept to reduce PAPR can be determined by comparing the peak power of the combined signal $(\mathbf{x} - \mathbf{c})$ and the mean power of the original signal $\mathbf{x}$. To enable such an assessment, a convenient measure of PAPR of the combined signal is given as

$$\text{PAPR}\{\mathbf{x} - \mathbf{c}\} = \frac{\max_{0 \leq n \leq N-1}\{|x(n) - c(n)|^2\}}{E\{|x(n)|^2\}} \tag{5.5}$$

Although, the application of the tone reservation concept may be very effective in reducing PAPR, two major drawbacks arise from its use. The very first and obvious one is the loss in data-rate due to the reserved subcarriers, which cannot be used to carry user data. Such a loss can be measured by the ratio of the number of the reserved subcarriers to the total number of subcarriers initially meant for the transmission of user data, i.e.

$$R_f = \frac{L}{N} \tag{5.6}$$

To minimize data-rate loss, it is desirable that L be much less than N, i.e. $L \ll N$. If $L = 0$, there is not PAPR reduction and $R_f = 0$. On the other hand, $L = N$ corresponds to a trivial case where no user-data is transmitted, thus there is 100% data-rate loss. A given value of L limits the number of available PRCs combinations to cancel all the highest peaks of a signal and thus the amounts of PAPR reductions that can be achieved.

Obviously, PAPR reduction will increase as the value of L increases, but this is at the expense of an increase in the value of R_f. Therefore, a method realizing tone-reservation concept should aim to achieve a high PAPR reduction with the smallest number of reserved tones possible. In particular, the minimum value of L is the one that can achieve PAPR reduction sufficient to avoid nonlinear amplification of signals in the HPA.

The second drawback associated with the tone-reservation approach is the increase in the average power of the transmitted signal, in this case the power of the combined signal $(\mathbf{x} - \mathbf{c})$. This is caused by the minimax nature of the optimization problem, which is known to yield signal amplitudes that are virtually uniformly distributed. A possible mitigation of this drawback requires the search for the PRCs to be done in a manner allowing for only a minimal increase in the average transmission power, while at the same time adhering to the maximum power rating of the HPA.

5.2.2 SOCP-based Tone Reservation Method

The SOCP-based Tone Reservation Method (SOCP-TR) was developed with the aim to reduce PAPR through finding an optimal peak-reduction vector $\hat{\mathbf{C}}$ and to establish whether the type of subcarrier modulation affect the amount of PAPR reduction. The knowledge on the influence or lack of it, of the type of subcarrier modulation on the level of PAPR reduction, once established, would be essential in choosing one modulation scheme over another or in generalising the results of PAPR reduction from just one modulation scheme.

The third quest for the SOCP-TR method was to find out whether using a real signal as is normally the case with discrete multi-tone (DMT) technique, the baseband variant of OFDM in wireline twisted-pair channel, or a complex OFDM signal as employed in wireless communication channel influences the extent of PAPR reduction.

In addition, since the method finds optimal PRCs, its PAPR reduction results could be used as references for benchmarking the performance of a method that finds suboptimal PRCs and therefore help to judge whether it is well designed.

A. Proposed Algorithm

The *minimax* problem in equation (5.4) is formulated as a *second-order cone program*. As described in Subsection 4.1.7 of Chapter 4, an SCOP is a convex optimization problem that involves minimization of a linear function over the intersection of an affine set and product of second-order cones. Considering equation (5.2) and using a time-domain vector $\mathbf{s}$ to represent the PAPR-reduced transmit signal as follows:

$$\mathbf{s} = \mathbf{x} - \hat{Q}\hat{\mathbf{C}} = [s(0), s(1), \dots, s(N-1)]^T \tag{5.7}$$

the *infinity* or ℓ_∞-norm of vector $\mathbf{s}$ can be expressed as

$$\|\mathbf{s}\|_\infty = \max\{|s(0)|, |s(1)|, \dots, |s(N-1)|\} \tag{5.8}$$

The components of vector $\mathbf{s}$, i.e. $s(i)$, $i = 0, 1, \dots, N-1$, are complex and each of them can be considered to be a complex number of the form

$$r_i = x_i - \mathbf{q}_i^T \hat{\mathbf{C}}, \qquad i = 0, 1, \dots, N-1 \tag{5.9}$$

Here superscript T denotes transpose and $\mathbf{q}_i \in \mathbb{C}^L$ is a column vector equal to the transpose of the ith row of IFFT submatrix $\hat{Q} \in \mathbb{C}^{N \times L}$.

An optimization problem to minimize the ℓ_∞-norm of $\mathbf{s}$ can be formulated as a

linear program as follows [123]:

$$\text{minimize} \quad t$$
$$\text{subject to} \quad \left| x_i - \mathbf{q}_i^T \hat{\mathbf{C}} \right| \le t, \quad i = 0, 1, \dots, N - 1 \tag{5.10}$$

Here the optimization variables are $\hat{\mathbf{C}} \in \mathbb{C}^L$ and $t \in \mathbb{R}$ and the input parameters are $x_i \in \mathbb{C}$ and $\mathbf{q}_i \in \mathbb{C}^L$.

Letting $x_i - \mathbf{q}_i^T \hat{\mathbf{C}} = r_i$, the inequality constraints in equation (5.10) require the magnitude of each complex number r_i to be less or equal to t. Since the magnitude $|r_i|$ can be written as

$$|r_i| = (r_i r_i^*)^{\frac{1}{2}}, \qquad i = 0, 1, \dots, N - 1 \tag{5.11}$$

where the superscript * denotes complex conjugate, then each inequality constraint in equation (5.10) can be expressed as a 3-dimensionsional second-order cone as

$$C_3 = \begin{bmatrix} r_i \\ t \end{bmatrix} = \begin{bmatrix} x_i - \mathbf{q}_i^T \hat{\mathbf{C}} \\ t \end{bmatrix}, \qquad i = 0, 1, \dots, N - 1 \tag{5.12}$$

By decomposing vector $\hat{\mathbf{C}}$ into two vectors, $\mathbf{y}$ and $\mathbf{z}$, made up of the real and imaginary parts such that $\hat{\mathbf{C}} = \mathbf{y} + j\mathbf{z}$, the problem formulation in equation (5.10) can be expressed in the standard form of SOCP as follows:

$$\text{minimize} \quad \begin{bmatrix} \mathbf{0}_{1L} \\ \mathbf{0}_{1L} \\ 1 \end{bmatrix}^T \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \\ t \end{bmatrix}$$

$$\text{subject to} \quad \left\| -\begin{bmatrix} Re\{\mathbf{q}_i^T\} & -Im\{\mathbf{q}_i^T\} & 0 \\ Im\{\mathbf{q}_i^T\} & Re\{\mathbf{q}_i^T\} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \\ t \end{bmatrix} + \begin{bmatrix} Re\{x_i\} \\ Im\{x_i\} \end{bmatrix} \right\|_2 \le \begin{bmatrix} \mathbf{0}_{1L} \\ \mathbf{0}_{1L} \\ 1 \end{bmatrix}^T \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \\ t \end{bmatrix} \tag{5.13}$$

In this formulation, $\|.\|_2$ denotes the ℓ_2-norm, $\mathbf{0}_{1L}$ is a column vector of L zeros and index $i = 0, 1, \dots, N - 1$. The problem has $(2L+1)$ optimization variables, which are $t \in \mathbb{R}$, $\mathbf{y} = Re\{\hat{\mathbf{C}}\} \in \mathbb{R}^L$ and $\mathbf{z} = Im\{\hat{\mathbf{C}}\} \in \mathbb{R}^L$.

Because the SOCP optimization problem in equation (5.13) is convex, it yields an optimal solution for the $(2L+1)$ variables for any input signal $\mathbf{x}$, and hence the required PRCs and PRS. The main steps of the just described SOCP-TR algorithm are given in Table 5.1. A software package referred to as CVX, which is a modelling system for forming and finding solutions of disciplined convex programs in MATLAB [125], can be directly employed to solve the problem.

Table 5.1. SOCP-based tone reservation algorithm

SCOP-TR algorithm
i.   Set number of subcarriers N, allowed data-rate loss R_f and maximum allowed $PAPR_{max}$
ii.   Generate OFDM signal **x**
iii.   Calculate PAPR of **x**
iv.   If calculated PAPR $< PAPR_{max}$, transmit **x** and terminate the program, else go to step (v)
v.   Generate IFFT submatrix $\hat{Q} \in \mathbb{C}^{N \times L}$
vi.   Use SOCP optimization function formulated in CVX to find $\hat{\mathbf{C}} = \mathbf{y} + j\mathbf{z}$
vii.   Compute peak-reduction vector $\mathbf{c} = \hat{Q}\hat{\mathbf{C}}$
viii.   Compute $\mathbf{s} = \mathbf{x} - \mathbf{c}$ and transmit
ix.   End

A. Computational Complexity

In CVX, a SOCP can be solved using primal-dual interior-point algorithm by SeDuMi [126], SDPT3 [127] or MOSEK [128] software packages. The three packages can solve a SOCP problem in about 10 to 100 iterations, each with a polynomial complexity of $O(n^2 m)$, where n and m are the number of variables and constraints, respectively. Thus, to find $n = 2L + 1$ variables subject to N constraints in equation (3.13), the computational complexity per iteration is approximately $O(NL^2)$.

5.2.3 LP-based Tone reservation method

This method, abbreviated as LP-TR, was invented with the motivation that an ideal peak-reduction signal should be equal to the signal the HPA clips off from an original OFDM signal. Following this, an ideal peak-reduction signal can be represented analytically by vector $\mathbf{d} = [d(0), d(1), \dots, d(N-1)]^T$, whose components are given by

$$d(n) = \begin{cases} \dfrac{x(n)}{|x(n)|}(|x(n)| - x_{th}), & |x(n)| > x_{th} \\ 0 & , \quad |x(n)| \leq x_{th} \end{cases} \tag{5.14}$$

where x_{th} is the clipping threshold.

The clipping threshold should be ideally greater than the average of the magnitudes of the complex signal amplitudes. It is possible to find this threshold for a given OFDM signal if the maximum PAPR allowed by the HPA is provided using the following equation:

$$x_{th} = \sqrt{PAPR_{max}E\{|x(n)|^2\}} \tag{5.15}$$

where $E\{.\}$ is the expectation operator that outputs the average of the argument.

A. Proposed Algorithm

For efficient reduction of PAPR, the actual peak-reduction signal needs to be approximately equal to the ideal one in equation (5.14). In other words, the residual error signal between the desired and actual peak-reduction signals given by

$$\mathbf{r} = \hat{Q}\hat{\mathbf{C}} - \mathbf{d} \tag{5.16}$$

where $\mathbf{r} = [r(0), r(1), ..., r(N-1)]$, should be as small as possible if not zero.

The process of generating the PRCs can be viewed in a similar manner to the adaptive filtering scheme illustrated in Figure 5.2. For every OFDM signal vector $\mathbf{x}$, a new vector $\hat{\mathbf{C}}$ containing a new set of PRCs is computed. However, in the process of generating the PRCs, the desired signal is not fixed like in the adaptive filtering since it varies with the incoming OFDM signal.

The search for the peak-reduction vector that yields components of $\mathbf{r}$ near zero can be accomplished by minimizing a specially chosen measure of the residual error. Since the aim here is to reduce the PAPR caused by the peak amplitude of $x(n)$, the largest magnitude of the residual error should be minimized. This requires the minimization of the Chebyshev norm of the residual error.

Thus, to find a vector $\hat{\mathbf{C}}$ that minimizes $\mathbf{r}$, a Chebyshev approximation problem expressed as

$$\underset{\hat{\mathbf{C}}}{\text{minimize}} \quad ||\hat{Q}\hat{\mathbf{C}} - \mathbf{d}||_\infty \tag{5.17}$$

where $||.||_\infty$ denotes ℓ_∞-norm, needs to be solved.
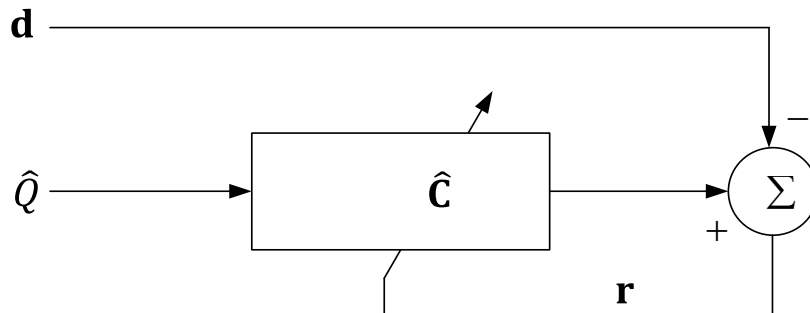
Figure 5.2. Generation of peak-reduction coefficients

The Chebyshev approximation problem in equation (5.17) is known not to have a closed form solution but it can be casted into a LP and solved using readily available LP solvers. The LP for the problem is of the form

$$\text{minimize} \quad t$$

$$\text{subject to} \quad \widehat{\mathbf{q}}_n^T \widehat{\mathbf{C}} - t \leq d(n) \tag{5.18}$$

$$-\widehat{\mathbf{q}}_n^T \widehat{\mathbf{C}} - t \leq -d(n)$$

with $t \in \mathbb{R}$ and $\widehat{\mathbf{C}} \in \mathbb{C}^L$ being the optimization variables and $\widehat{\mathbf{q}}_n \in \mathbb{C}^L$ and $d(n) \in \mathbb{C}$, for $n = 0, 1, 2, \ldots, N - 1$, are the problem parameters. Here, it should be noted that vector $\widehat{\mathbf{q}}_n$ is equal to the transpose of the nth row of matrix $\widehat{Q}$.

After solving for the PRCs vector, the PRS and the peak-reduced signal can be found from equations (5.3) and (5.7), respectively. A sequence of steps for the just described LP-TR algorithm is given in Table 5.2.

The use of the LP-TR method to reduce PAPR may increase the average power of the transmit signal. However, because the majority of elements in the desired PRS given in equation (5.14) are zero, a similar majority of the PRS samples generated by the method would be very small nearing zero; hence, the power increase is expected to be very small.

Table 5.2. LP-based tone reservation algorithm

LP-TR algorithm
i.   Set the number of subcarriers N, allowed data-rate loss R_f and maximum allowed $PAPR_{max}$
ii.   Generate OFDM signal $\mathbf{x}$
iii.   Compute PAPR of signal $\mathbf{x}$
iv.   If PAPR $< PAPR_{max}$, transmit $\mathbf{x}$ and end the program, otherwise go to step (v)
v.   Compute clipping threshold $x_{th} = \sqrt{PAPR_{max} E\{|x(n)|^2\}}$
vi.   Generate desired PRS $\mathbf{d}$
vii.   Generate IFFT submatrix $\widehat{Q} \in \mathbb{C}^{N \times L}$
viii.   Apply interior-point algorithm on $||\widehat{Q}\widehat{\mathbf{C}} - \mathbf{d}||_\infty$ to solve for $\widehat{\mathbf{C}}$
ix.   Compute PRS $\mathbf{c} = \widehat{Q}\widehat{\mathbf{C}}$
x.   Compute and transmit the peak-reduced signal $\mathbf{s} = \mathbf{x} - \mathbf{c}$
xi.   End

A. Computational Complexity

In practice, there is no simple analytical method to solve a LP. This makes it difficult to establish exactly the required number of arithmetic operations needed to find a LP solution. However, one of the three solvers for LPs, namely interior-point method, active-set method or simplex method, can be employed to solve for the PRCs vector $\hat{\mathbf{C}}$ for the problem in equation (5.18) depending on the problem size.

The interior-point method is the most superior of the three solvers and can solve the LP in less than 85 iterations each of maximum complexity $O(NL^2)$, where N and L are, respectively, the row and column dimensions of matrix $\hat{Q}$. In addition, since matrix $\hat{Q}$ is derived from the IFFT matrix, which bears a special structure, the computational complexity of the LP could be reduced to $O(N \log_2 N)$ [129].

5.2.4 IRLS-based Tone Reservation Method

This method is abbreviated as IRLS-TR. It employs the IRLS algorithm presented in Subsection 4.1.4 of Chapter 4 to find the required PRCs for a given OFDM signal.

A. Proposed Algorithm

The method starts by generating the desired peak-reduction signal and the residual error signal using equations (5.14) and (5.16), respectively. The problem of minimizing l_∞-norm in equation (5.17) can be framed as a problem of minimizing ℓ_p-norm when p is large, i.e.

$$\underset{\hat{\mathbf{C}}}{\text{minimize}} \quad ||\hat{Q}\hat{\mathbf{C}} - \mathbf{d}||_p^p, \qquad p \to \infty. \tag{5.19}$$

In practice, the solution that minimizes ℓ_p-norm is found to approximate that which minimizes ℓ_∞-norm when $p \geq 10$ [130]. However, there does not exist an analytical procedure to find the optimal solution that minimizes any other norm except ℓ_2-norm. Thus, it is necessary to transform the problem in equation (5.19) into a form of a least-squares problem, specifically to a weighted least-squares formulation, in order to find its closed form solution.

An equivalent weighted least-squares form of problem (5.19) can be expressed as

$$\underset{\hat{\mathbf{C}}}{\text{minimize}} \quad ||W(\hat{Q}\hat{\mathbf{C}} - \mathbf{d})||_2^2 \tag{5.20}$$

where the weighting matrix W is a real diagonal matrix of size $N \times N$

$$W = \begin{bmatrix} w(0) & 0 & \cdots & 0 \\ 0 & w(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w(N-1) \end{bmatrix} \qquad (5.21)$$

The purpose of this matrix is to apply heavy weights on the largest amplitudes of the residual error to emphasize on their minimization.

The problem in equation (5.20) can be cast into a standard least-squares problem and if the diagonal weights are known, the peak-reduction vector is given by the closed form solution

$$\hat{\mathbf{C}} = \left(\hat{Q}^H W^T W \hat{Q}\right)^{-1} \hat{Q}^H W^T W \mathbf{d} \qquad (5.22)$$

Here the superscripts T and H denote the transpose and conjugate transpose of a matrix, respectively.

The two problems in equation (5.19) and (5.20) can be made equal by assigning the residual errors $r(n)$ to the diagonal elements of the weighting matrix according to

$$w(n) = |r(n)|^{\frac{p-2}{2}}, \qquad n = 0, 1, \ldots, N - 1 \qquad (5.23)$$

Thus, by substituting equation (5.23) in (5.20), the weighted least-squares problem becomes identical to the ℓ_p-norm minimization problem in (5.19).

The weighted least-squares problem can be solved analytically but not in a single step because the weights that give an optimal approximation solution when they are substituted in equation (5.22) are unknown *a priori*. For this reason, the weighted least-squares problem is solved iteratively us an IRLS algorithm. The starting point of the algorithm is to find a solution of $\hat{\mathbf{C}}$ using equation (5.22) when all the weights are set to 1, i.e.

$$w(n) = 1, \qquad n = 0, 1, \ldots, N - 1 \qquad (5.24)$$

This is followed by finding a residual error using equation (5.16) and a computation of new weights for the next iteration using equation (5.23).

A new weighted least-squares solution $\hat{\mathbf{C}}$ is then computed from the new set of weights using equation (5.22). The steps from the computation of the residual error are repeated until the algorithm converges to a literally unchanging value of the ℓ_p-norm of the error, measured by a tolerance parameter of the required precision, during several iterations or the maximum number of iterations is reached.

The IRLS algorithm, in its basic form given above, has two main drawbacks that need to be avoided. First, it may not converge due to numerical instabilities experienced when the algorithm is minimizing the ℓ_p-norm for some values of p. Secondly, the basic algorithm converges in linear manner, thus making it slow and unsuitable for real-time practical applications. The two drawbacks can be evaded by transforming the basic IRLS algorithm to have an iterative behaviour similar to the Newton's method presented in Subsection 4.2.2 of Chapter 4 by only partially updating the solution at each iteration as follows:

$$\hat{\mathbf{C}}(k) = \mu_k \hat{\mathbf{C}}(k) + (1 - \mu_k)\hat{\mathbf{C}}(k - 1) \tag{5.25}$$

where $\mu_k = 1/(p_k - 1)$ is the partial-update parameter at the kth iteration [131], [132].

However, as it is common with most of the methods based on the Newton's algorithm, the partial update in equation (5.25) causes the IRLS algorithm to become sensitive to initial approximations. This sensitivity to initial approximations affects the convergence rate at the very early phases of the algorithm. Fortunately, the convergence in the initial phases can be improved by gradually varying the value of p from the initial value of 2 to the one of the target ℓ_p-norm chosen to approximate the ℓ_∞-norm.

The gradual increase of p from a starting value to a final value can be considered similar to a homotopic deformation [133], [134] and can be done iteratively through multiplication of p and a convergence-rate controlling factor α of between 1 and 2. The convergence-rate controlling factor can be determined from the number of iterations and the value of p of the target ℓ_p-norm according to the following equation:

$$\alpha = 10^{\log(P)/k_{max}} \tag{5.26}$$

where k_{max} denotes the maximum iteration number allowed and P is the value of p of the target ℓ_p-norm. Following this, the value of p at the kth iteration can be determined from the equation

$$p_k = \min(P, \alpha p_{k-1}) \tag{5.27}$$

The sequences of steps for the proposed IRLS-TR method are summarized in Table 5.3.

From equation (5.27), it can be found that p_k increases gradually from an initial value of 2 but after several iterations it saturates at the final value of P. For example, if the problem of ℓ_p-norm minimization is to be considered for $P = 20$ over 15 iterations with $\alpha = 1.22$, then p will increase from 2 to 20 as shown in Figure 5.3. This means that

at different iterations, different ℓ_p-norm minimizations determined by the value of p are performed starting with ℓ_2-norm at the first iteration up to ℓ_{20}-norm at the thirteenth iteration, which is then minimized for the remaining number of iterations.

Following the two settings in equations (5.26) and (5.27), a reliable quadratic convergence to the optimal approximation solution is guaranteed in just a small number of iterations. For example, minimizing ℓ_{20}-norm for the ℓ_∞-norm, the algorithm converges in about 9 iterations as illustrated in Figure 5.4. A fast convergence is achieved because the difference between consecutive solutions of the minimization of the gradually changing ℓ_p-norm gets smaller as p progressively increases towards the value of the target ℓ_P-norm at which the algorithm iterates for numerous times.

Table 5.3. IRLS-based tone reservation algorithm

	IRLS-TR algorithm
i.	Set number of subcarriers N, allowed data-rate loss R_f and maximum allowed $PAPR_{max}$
ii.	Set maximum iteration number k_{max}, ℓ_p-norm parameter P and convergence parameter α
iii.	Initialize diagonal weights $w(n) = 1, n = 0, 1, \ldots, N-1$ and set function tolerance ε_{th}
iv.	Generate OFDM signal **x** and calculate PAPR
v.	If PAPR $< PAPR_{max}$, transmit **x** and end the program, otherwise go to step (vi)
vi.	Compute clipping threshold $x_{th} = \sqrt{PAPR_{max}E\{|x(n)|^2\}}$
vii.	Generate desired PRS **d**
viii.	Generate IFFT submatrix $\hat{Q} \in \mathbb{C}^{N \times L}$
ix.	Initialize iteration counter $k = 0$ and set initial ℓ_p-norm parameter $p_0 = 2$
x.	Compute initial weighted least-squares solution using $\hat{\mathbf{C}}(0) = \left(\hat{Q}^H W^T W \hat{Q}\right)^{-1} \hat{Q}^H W^T W \mathbf{d}$
xi.	Compute residual error $\mathbf{r}_k = \hat{Q}\hat{\mathbf{C}}_k - \mathbf{d}$
xii.	Compute new diagonal weights $w_k(n) = |r_k(n)|^{(p_k-2)/2}, n = 0, 1, \ldots, N-1$
xiii.	Compute current weighted least-squares solution $\hat{\mathbf{C}}(k) = \left(\hat{Q}^H W^T W \hat{Q}\right)^{-1} \hat{Q}^H W^T W \mathbf{d}$
xiv.	Calculate new update parameter $\mu_k = 1/(p_k - 1)$
xv.	Partially update current weighted least-squares solution as $\hat{\mathbf{C}}(k) = \mu_k \hat{\mathbf{C}}(k) + (1 - \mu_k)\hat{\mathbf{C}}(k-1)$
xvi.	Compute ℓ_p-norm of error $||\boldsymbol{\varepsilon}_k||_p = \left(\sum_{n=0}^{N-1}|r_k(n)|^p\right)^{1/p}$
xvii.	Set $k = k + 1$. If $k < k_{max}$ or $||\boldsymbol{\varepsilon}_k||_p > \varepsilon_{th}$, update ℓ_p-norm parameter $p_k = \min(P, \alpha p_{k-1})$ and go to step (xi); else transmit $\mathbf{s} = \mathbf{x} - \hat{Q}\hat{\mathbf{C}}$
xviii.	End
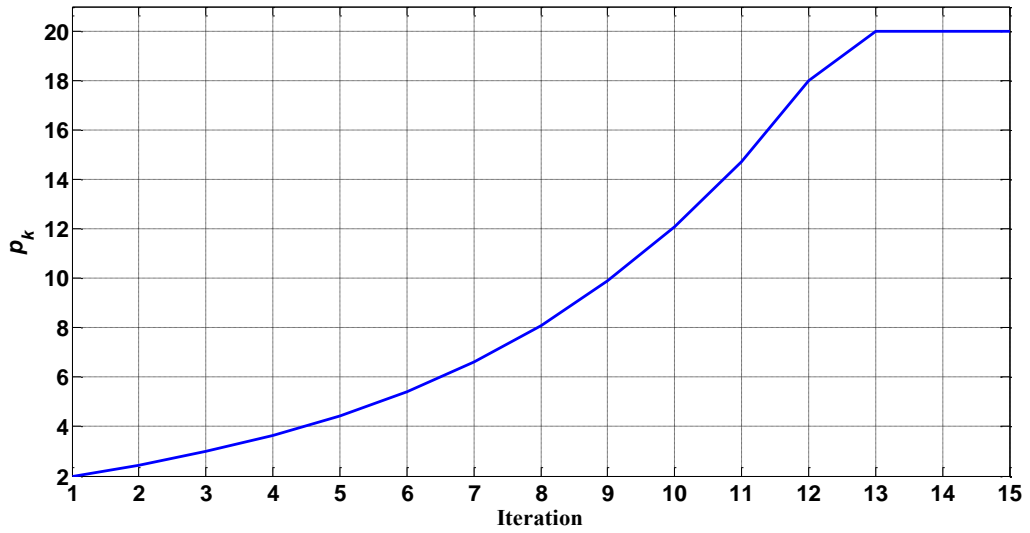
Figure 5.3. Homotopic variation of p
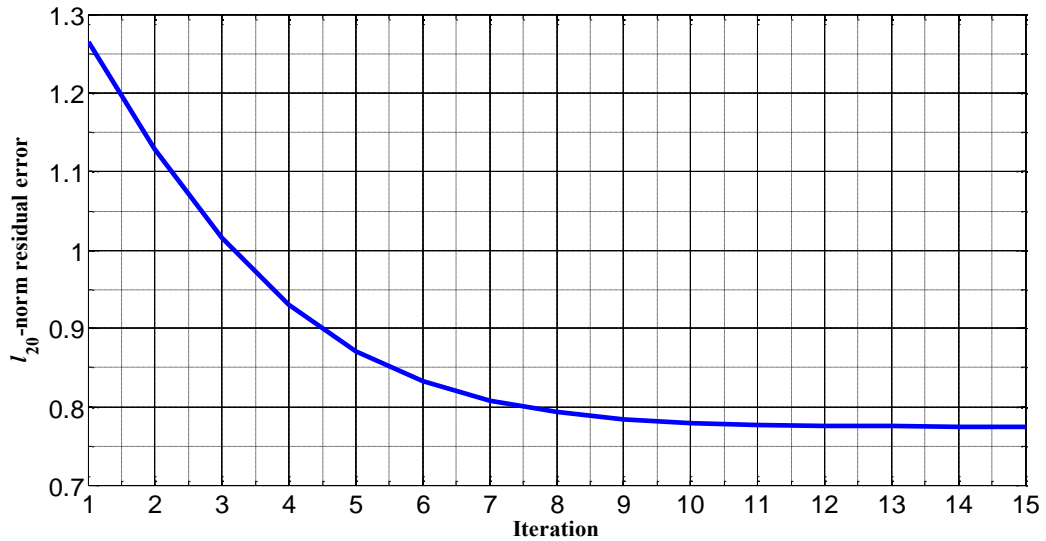
Figure 5.4. Convergence characteristic curve of IRLS algorithm

B. Computational Complexity

Much of the computational effort is spent in step (xiii) in Table 5.3 to compute the weighted least-squares solution from the normal equations

$$\left(\hat{Q}^H W^T W \hat{Q}\right)\hat{\mathbf{C}} = \hat{Q}^H W^T W \mathbf{d} \tag{5.28}$$

In order to determine the computational complexity of finding the weighted least-squares solution, equation (5.28) can be rewritten in the standard form

$$(A^H A)\hat{\mathbf{C}} = \mathbf{b} \tag{5.29}$$

where $A = W\hat{Q}$ and $\mathbf{b} = A^H W \mathbf{d}$. In addition, only multiplication operations are

113

considered for the evaluation of complexity because they are computationally more demanding than additions.

Obtaining matrix A with dimensions $N \times L$ requires $2NL$ real multiplications. Since one complex multiplication requires four real multiplications, then for the complex matrix product $(A^H A)$ of dimensions $L \times L$, a total of $2NL + 4NL^2$ real multiplications are needed. Moreover, to obtain the L complex elements of vector $\mathbf{b}$, $4NL$ real multiplications are required. In addition, solving for $\hat{\mathbf{C}}$ from the system of linear equations in (5.29), here rewritten as

$$B\hat{\mathbf{C}} = \mathbf{b} \tag{5.30}$$

needs the formation of matrix $B = A^H A$, which requires $4L^2$ real multiplications.

Adding all the complexities together, the computation of the weighted least-squares solution needs a total of $2NL + 4NL^2 + 4NL + 4L^2 \cong 4NL^2$ real multiplications. Therefore, the computational complexity per iteration of the IRLS algorithm is proportional to (NL^2) and can be stated as $O(NL^2)$. Since this complexity increases with the square of the value of L, it is necessary to reserve only a small number of subcarriers to achieve a PAPR reduction sufficient to avoid nonlinear amplification of a signal.

5.2.5 Low-Complexity Signal Addition Method

This method, abbreviated as LCSA, reserves resources for reducing PAPR in the time-domain by way of extending the transmit signal by a few samples. The main objective of this method is to create a peak-reduction signal that is an exact copy of the desired signal in equation (5.14) through a very low computational complexity process and without a significant change on the transmission power. To achieve these goals simultaneously, the method follows a very different approach of generating the actual peak-reduction signal from the one employed by the IRLS-TR method in Subsection 5.2.6.

A. Proposed Algorithm

The peak-reduction signal from the IRLS algorithm does not match the desired signal amplitudes as demonstrated by Figure 5.5. For example, at the positions where the desired signal $\mathbf{d}$ has zero entries, the actual signal $\mathbf{c}$ may have nonzero entries. This may lead to an increase in the average power of the transmit signal.

114

The proposed LCSA method utilizes only the nonzero components of $\mathbf{d}$ to generate the actual PRS. If the desired PRS has M nonzero entries, a vector containing only these components can be expressed as

$$\hat{\mathbf{c}} = [\hat{c}(0), \hat{c}(1), \dots, \hat{c}(M-1)]^T$$
$$= [d(i_0), d(i_1), \dots, d(i_{M-1})]^T \tag{5.31}$$

The elements of $\hat{\mathbf{c}}$ correspond to a set of indices $\{i_0, i_1, \dots, i_{M-1}\}$ of vector $\mathbf{d}$ in the range 0 to N-1 corresponding to the positions of the nonzero entries.

Generating a time-domain PRS signal $\mathbf{c}$ matching the desired signal $\mathbf{d}$ by using a frequency-domain product, i.e. by having $\mathbf{c} = \hat{Q}\hat{\mathbf{C}}$ in the tone reservation approach, is not feasible because $\hat{Q} \in \mathbb{C}^{N \times L}$ is not square, given that $L \ll N$ and the value of N must be always the same as the length of $\mathbf{d}$. A simpler way to circumvent this difficulty is to confine the formation of the actual PRS in the time-domain by having it equal to the desired signal, i.e.

$$\mathbf{c} = \mathbf{d} \tag{5.32}$$

However, generating the peak-reduced signal using $\mathbf{s} = \mathbf{x} - \mathbf{c}$ will clip some signal amplitudes and subsequently lead to severe BER degradation. Thus, after transmitting signal $\mathbf{s}$, it is necessary to restore back the clipped amplitudes at the receiver in order to guarantee a successful recovery of the transmitted modulation symbols. This can be accomplished by transmitting the nonzero samples of the actual PRS vector together with the samples of peak-reduced signal so that at the receiver they can be recovered and used to reconstruct the clipped samples back to their original amplitudes.
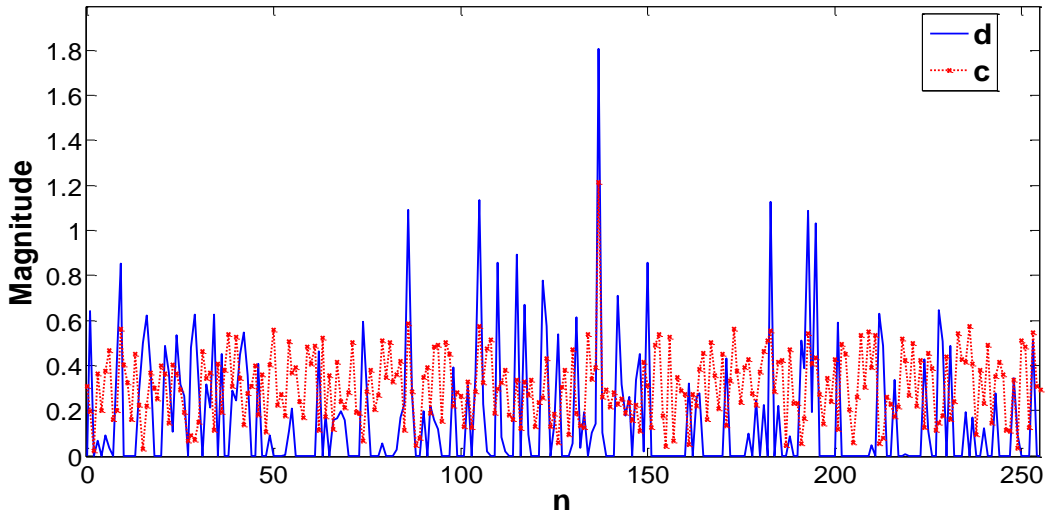
Figure 5.5. Desired and actual peak-reduction signals for the IRLS-TR method

Since the number of nonzero samples will in practice be small, they can be transmitted by appending them to the tail end of the transmit signal, thus yielding a composite signal $\hat{\mathbf{s}}$ as shown in Figure 5.6. In the figure, the signal part in blue is the peak-reduced signal $\mathbf{s}$, while in red are the nonzero samples of the PRS. As can be observed, the waveform of the composite signal is essentially the same as the waveform of the peak-reduced signal, thus the average powers of the two signals should be approximately the same.

However, the actual effect of the proposed method on the transmit power can be analysed by comparing the average power of composite signal to the power of the original OFDM signal. This can be done by comparing the total power in the clipped samples and the appended PRS samples to the total power in the corresponding unclipped samples in the original OFDM signal. To this end, let $i = 0, 1, \ldots, M - 1$ be the index of a clipped sample in the original OFDM signal and $\hat{c}(i)$ and $x(i)$ be, respectively, the nonzero PRS and original unclipped signal samples. The total power in the clipped and nonzero PRS samples satisfies the condition

$$\sum_{i=1}^{M} (x_{th}^2 + \hat{c}^2(i)) \leq \sum_{i=1}^{M} x^2(i) \tag{5.33}$$

The inequality condition in equation (5.33) implies that the power of the composite signal will always be less than or equal to that of the original OFDM signal. However, because the amplitudes of the nonzero PRS samples are very small and all the sample magnitudes under consideration are between 0 and 1, the power of the composite signal will practically be equal to the power of the original OFDM signal.
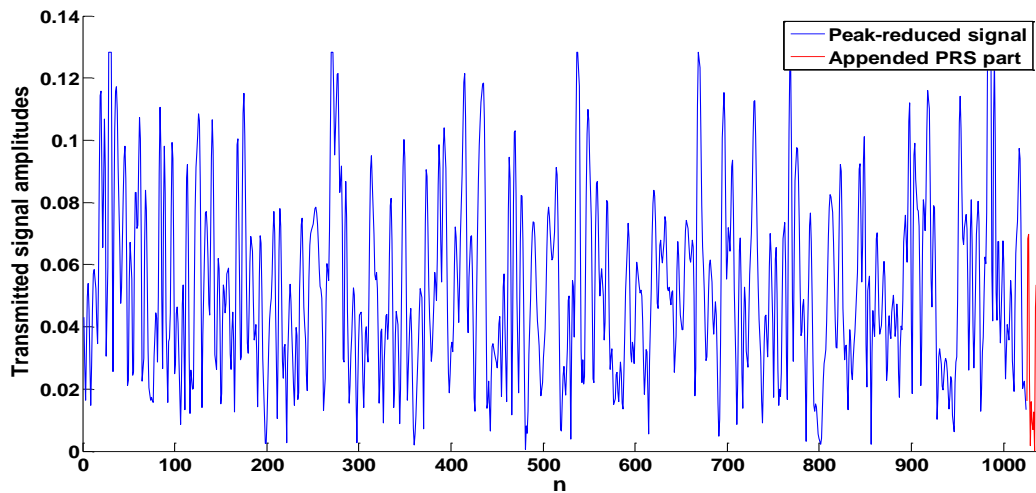
Figure 5.6. Composite transmit signal

The foregoing analysis shows that the LCSA method does not affect the transmit power. However, the extension of the transmit signal by some peak-reduction samples reduces the data-rate of the system because during the period the PRS samples are transmitted no user data can be sent. This reduction in data rate can be measured by the ratio of the number of nonzero samples in the PRS to the total number of samples in the composite transmit signal, i.e.

$$R_t = \frac{M}{N + M} \tag{5.34}$$

Obviously, for $M = L$, this data-rate loss is slightly less than the loss due to the tone reservation in equation (5.6).

The data-rate loss in equation (5.34) is a function of M and can be controlled by altering the values of M. However, this can only be achieved by adjusting the clipping threshold. When the clipping threshold is set, M can be found easily from the statistical distribution of the amplitudes of an OFDM signal, which from Chapter 3 are known to be Raleigh-distributed with a cumulative distribution function

$$F_X(x) = 1 - e^{-\frac{x^2}{2\sigma^2}} \tag{5.35}$$

where σ^2 is the variance of the normal distribution of the real and imaginary parts of the OFDM signal.

From equation (5.35), the probability of the magnitude of a signal amplitude being higher than a specified clipping threshold is given by

$$\Pr(X > x_{th}) = e^{-\frac{x_{th}^2}{2\sigma^2}} \tag{5.36}$$

and is simply equal to the ratio of the number of nonzero entries in the PRS to the total number of samples in the signal, i.e.

$$\frac{M}{N} = e^{-\frac{x_{th}^2}{2\sigma^2}} \tag{5.37}$$

The clipping threshold can be related to the mean of the distribution in equation (5.35) by

$$x_{th} = \lambda \bar{x} \tag{5.38}$$

where $\bar{x}$ is the average value and λ is a clipping level control parameter ranging from 1 up to a maximum value given by $\max(|x(n)|) / \bar{x}$.

The mean of the Rayleigh distribution in equation (5.35) is given by

$$\overline{x} = \sigma\sqrt{\frac{\pi}{2}} \tag{5.39}$$

On substituting equation (5.39) in (5.38), the clipping threshold becomes

$$x_{th} = \lambda\sigma\sqrt{\frac{\pi}{2}} \tag{5.40}$$

The number of signal amplitudes above the clipping threshold can be found upon the substitution of equation (5.40) in (5.37) as

$$M = Ne^{-\frac{\pi}{4}\lambda^2} \tag{5.41}$$

Equation (5.1) shows that M falls exponentially with the square of the clipping-level control parameter. It has a maximum value of $0.46N$ at $\lambda = 1$, i.e. when the clipping threshold is equal to the average value of the signal. On the other hand, M has a minimum value of zero when the clipping threshold is equal to the maximum peak value of the signal. It is desirable to have a high clipping threshold close to the maximum peak value in order to have a small vale of M and hence to reduce the loss in data rate.

However, the reduction in PAPR decreases with M and approaches zero when the clipping threshold tends toward the maximum peak of the signal. Thus, while setting the value of M, it is necessary to take into account the trade-off between PAPR reduction and data-rate loss. If for a given OFDM system the maximum acceptable data-rate loss is specified, the value of M can be found from equation (5.34) as follows:

$$M = \frac{NR_t}{1 - R_t} \tag{5.42}$$

Consequently, the optimal clipping threshold can be obtained from equation (5.37) as follows:

$$x_{th} = \overline{x}\sqrt{\frac{4}{\pi}\ln\left(\frac{N}{M}\right)} \tag{5.43}$$

This optimal clipping level can then be used to find the desired PRS and to reduce PAPR. On the other hand, if the OFDM system is not sensitive to data-rate loss, the value of M can be set in a flexible manner to control the amount of PAPR reduction.

One simple way of controlling the amount of PAPR reduction using M is by setting a clipping threshold equal to one of the highest samples of a signal. The amplitudes of the signal can first be arranged in a descending order and if the position of

the sample designated as the clipping threshold is n, then M is simply equal to $n-1$. By choosing different samples as the clipping threshold and hence different M values, PAPR reduction can be controlled easily. Similarly, the value of M that gives a desired PAPR can be found without difficulty.

Following the above discussion, the main steps of the proposed LCSA algorithm can be summarized as shown in Table 5.4. A transmitter implementing the method is illustrated in Figure 5.7. In the figure, **X** is a vector containing modulation symbols, **x, c** and **s** are the discrete-time OFDM signal, peak-reduction signal and peak-reduced signal, respectively. The nonzero samples of **c** are added to the peak-reduced signal in the block "Append **ĉ**" to give a composite transmit signal **ŝ**.

In the "Append CP" block, the OFDM signal is extended with a cyclic prefix to create a time-guard interval for eliminating ISI. At the receiver, after the removal of the cyclic prefix, the M PRS samples are then recovered. These samples are added to the clipped samples in the received signal to recreate them to their original level before the demodulation operation in the FFT block.

Table 5.4. LCSA PAPR reduction algorithm

	LCSA algorithm
i.	Set number of subcarriers N, data-rate loss R_t or maximum allowed $PAPR_{max}$
ii.	Generate OFDM signal **x** and calculate PAPR
iii.	If PAPR< $PAPR_{max}$ transmit **x** and terminate the program, otherwise go to step (iv)
iv.	Set clipping threshold x_{th}
v.	Generate desired PRS **d**
vi.	Assign **d** to be the actual peak-reduction signal **c**
vii.	Generate vector **ĉ** to contain nonzero samples of **c**
viii.	Generate peak-reduced signal $\mathbf{s} = \mathbf{x} - \mathbf{c}$
ix.	Append **ĉ** to **s** and pass the composite signal **ŝ** for on-ward processing
x.	End
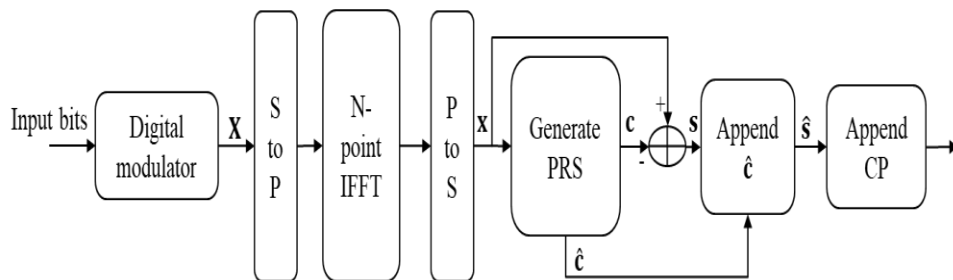
Figure 5.7. Transmitter implementation of LCSA method

B. Computational Complexity

From the foregoing description of the proposed LCSA method, the two main operations that are computationally demanding are the generation of a peak-reduction signal using equation (5.14) and the subtraction of this signal from an OFDM signal to reduce the PAPR. The generation of a peak-reduction signal is done in a single step, which requires M real additions and $2M$ real multiplications. The operation of reducing PAPR requires $2M$ real additions.

Therefore, the algorithm executes $2M$ real multiplications and $2M$ real additions in the two operations. Considering only multiplications, because of being more computationally intensive than additions, the computational complexity of the LCSA algorithm increases with M, thus it can be stated as $O(M)$. Given that $M \ll N$, the proposed method will barely increase the computational complexity of an OFDM system since $O(M)$ is much less than $O(N \log_2 N)$, the complexity of the FFT, which happens to be the most computationally expensive operation in the system.

Compared to the IRLS-TR method proposed in Subsection 5.2.4 for $L \cong M$, the proposed LCSA is by far much faster because it requires only 1 iteration of complexity $O(M)$ against 9 iterations of the former, each requiring a complexity of $O(NL^2)$.

5.2.6 Low-Complexity Additive Signal Mixing Method for MIMO-OFDM System

This section presents a *low-complexity additive signal mixing* (LCASM) algorithm that can be employed in MIMO-OFDM systems to reduce PAPR. In general, the LCASM algorithm is quite similar to the LCSA technique in presented in Table 5.4, except for a few amendments to cater for simultaneous reduction of PAPR in more than one transmission paths.

Without loss of generality, only an algorithm for a 2×2 Alamouti-encoded MIMO-OFDM system, which was presented in Subsection 3.9.3, will be given. The system can be considered as having two parallel OFDM systems, each with its own transmit signal. Thus, reduction of PAPR must be performed on the two transmit signals, here denoted by $\mathbf{x}_1$ and $\mathbf{x}_2$, from the two parallel transmission paths. However, of the two signals, more reduction should be performed on the one with higher PAPR during a given symbol duration in order to effectively reduce the PAPR of the whole system, which is given by

$$PAPR_{MIMO} = \max(PAPR\{\mathbf{x}_1\}, PAPR\{\mathbf{x}_2\}) \qquad (5.44)$$

A. Proposed Algorithm

From the allowed data-rate loss in a MIMO-OFDM system, the number of nonzero samples in peak-reduction signals can be determined using equation (5.42). Two clipping threshold, x_{th1} and x_{th2}, for each transmit signals can then be obtained using equation (5.43). The effective clipping threshold for the system should be the maximum of the two thresholds, i.e.

$$x_{th} = \max(x_{th1}, x_{th2}) \qquad (5.45)$$

The required peak-reduction signals, $\mathbf{c}_1$ and $\mathbf{c}_2$, for the two transmit branches are then generated using equation (5.14). Additionally, two vectors $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{c}}_2$, for holding the nonzero elements of the peak-reduction signals, $\mathbf{c}_1$ and $\mathbf{c}_2$, respectively, are needed. If one of the vectors, $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{c}}_2$, has less number of elements than the other, it can be padded with zeros to have the same length on the two signals.

The peak-reduced signals for the two MIMO transmit branches can be created as follows:

$$\mathbf{s}_1 = \mathbf{x}_1 - \mathbf{c}_1$$

$$(5.46)$$

$$\mathbf{s}_2 = \mathbf{x}_2 - \mathbf{c}_2$$

Signals $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{c}}_2$ are then appended to the two peak-reduced signals, $\mathbf{s}_1$ and $\mathbf{s}_2$, to form the composite transmit signals $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$, respectively. Two typical composite transmit signals are shown in Figure 5.8. From the figure, it is evident that as long as the number of nonzero PRS samples appended to the transmit signal is small, the waveforms of composite transmit signals, $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$, are essentially the same as the ones for the peak-reduced signals $\mathbf{s}_1$ and $\mathbf{s}_2$. Therefore, similar to the proposed LCSA method, the transmission of nonzero PRS samples together with peak-reduced MIMO-OFDM signals barely affects the transmission power.

B. Computational Complexity

The proposed LCASM algorithm is presented in Table 5.5. Since the sequence and the steps of the algorithm are similar to those of the LCSA method in Table 5.4, except for the accommodation of a second transmit signal, the LCASM method has a similar implementation and computational complexity to the LCSA method.
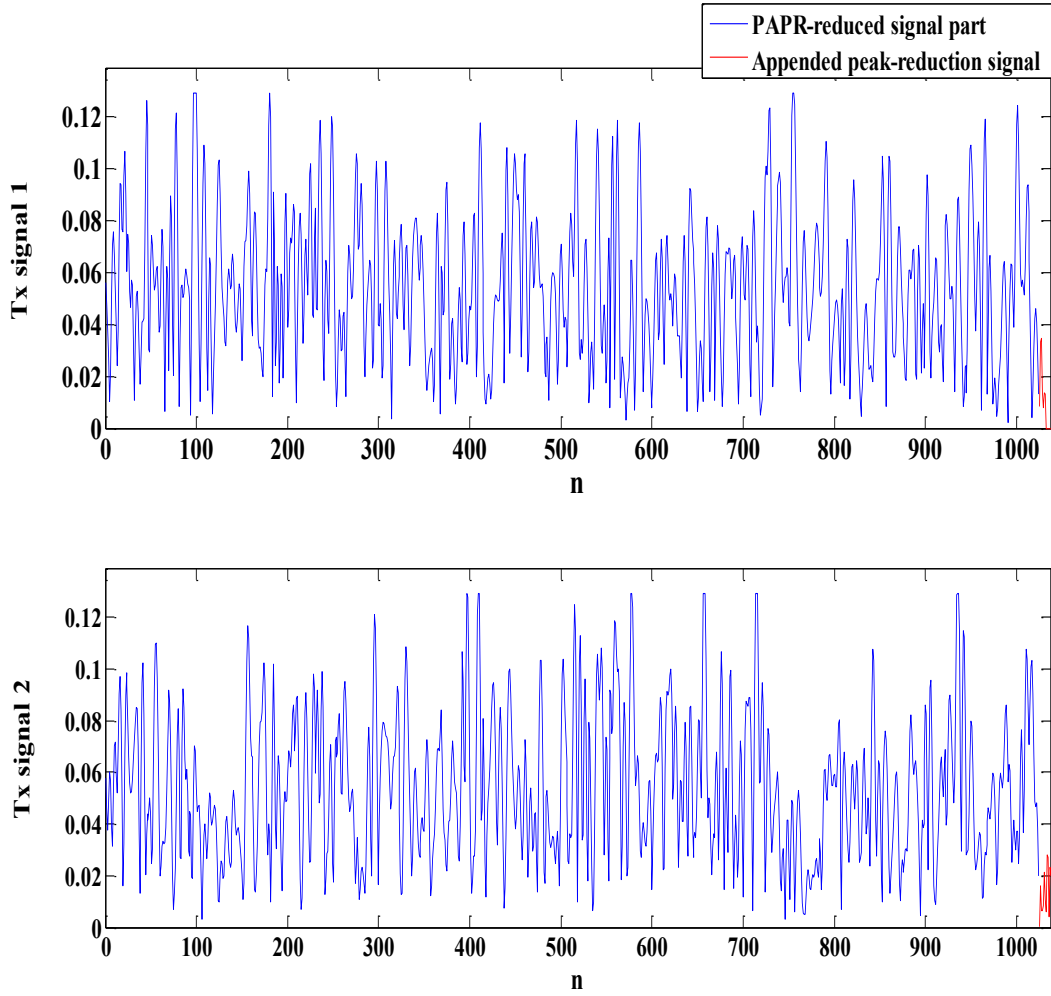
Figure 5.8. Peak-reduced transmit signals for 2×2 Alamouti encoded MIMO-OFDM system

Table 5.5. LCASM PAPR reduction algorithm for MIMO-OFDM system

LCASM algorithm

i. Set the number of subcarriers N, data-rate loss R_t and $PAPR_{max}$

ii. Generate signals $\mathbf{x}_1$ and $\mathbf{x}_2$ for the two MIMO branches

iii. Find $PAPR_{MIMO}$

iv. If $PAPR_{MIMO} < PAPR_{max}$, transmit $\mathbf{x}_1$ and $\mathbf{x}_2$, end the algorithm, otherwise go to step (v)

v. Calculate clipping threshold x_{th}

vi. Generate peak-reduction signals $\mathbf{c}_1$ and $\mathbf{c}_2$

vii. Generate signals $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{c}}_2$ containing nonzero elements of $\mathbf{c}_1$ and $\mathbf{c}_2$

viii. Generate peak-reduced signals $\mathbf{s}_1 = \mathbf{x}_1 - \mathbf{c}_1$ and $\mathbf{s}_2 = \mathbf{x}_2 - \mathbf{c}_2$

ix. Append $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{c}}_2$ to $\mathbf{s}_1$ and $\mathbf{s}_2$ and transmit the composite signals

x. End

122

5.2.7 Computational complexity Comparison

Table 5.6 gives the computational complexities of the proposed LCSA, LCASM, IRLS-TR, LP-TR and SOCP-TR methods and some other methods relevant to this research and which were briefly reviewed in Chapter 2, namely LSA-TR [48], CF-TR [49], SSCR-TR [50], IVO-TR [52] and ELM-TR [53]. The number of iterations required to obtain the PRS at runtime for the proposed IRLS-TR, LP-TR and SOCP-TR methods is denoted by I_{irls}, I_{lp} and I_{socp}, respectively. For the LSA-TR, CF-TR and SSCR-TR methods, the number of iterations at runtime is denoted by I_{lsa}, I_{cf} and I_{sscr}, respectively.

The parameter M in the complexities of LCSA, LCASM, CF-TR and LSA-TR denotes the number of nonzero entries in the clipping noise signal. The parameters N_s, N_i, N_h and N_o in the complexities of the IVO-TR and ELM-TR methods are the size of the training data set and the number of neurons in the input layer, hidden layer and output layer of the neural network, respectively.

The LCSA, LCASM, IRLS-TR, LP-TR, SOCP-TR, SSCR-TR, CF-TR and LSA-TR methods have only runtime operations. However, the computational complexity of the two machine-learning algorithms, IVO-TR and ELM-TR, is in three parts; these are in the CC-TR [47] iterations to generate training targets, the training process and the runtime prediction procedure.

Except for the number of iterations, the computational complexities of the proposed IRLS-TR, LP-TR and SOCP-TR methods are similar. However, the IRLS-TR method converges faster in about 9 iterations against about 85 iterations of the SOCP-TR and LP-TR methods.

In addition, the minimization of the ℓ_∞-norm through an iterative computation of the weighted least-squares solution in IRLS-TR method is rather more direct and faster than using the rather complicated primal-dual interior-point method to solve a linear program and a second-order cone program in LP-TR and SOCP-TR methods, respectively.

Compared to the IRLS-TR method, the LCSA method is much faster as it requires only one operation of complexity $O(M)$ to generate the PRS against 9 iterations each of complexity $O(NL^2)$ by the former, for approximately the same number of reserved subcarriers and number of nonzero samples, i.e. $L \cong M$ and $M \ll N$. In addition, the LCSA method has a much lower complexity than the other methods in the table.

Table 5.6. Computational complexity of different PAPR reduction methods

PAPR reduction method	Complexity during training	Complexity at runtime
Proposed LCSA	Not applicable	$O(M)$
Proposed LCASM	Not applicable	$O(M)$
Proposed IRLS-TR	Not applicable	$I_{irls} \times O(NL^2)$
Proposed LP-TR	Not applicable	$I_{lp} \times O(NL^2)$
Proposed SOCP-TR	Not applicable	$I_{socp} \times O(NL^2)$
LSA-TR [48]	Not applicable	$I_{lsa} \times O(M + N \log_2 N)$
CF-TR [49]	Not applicable	$I_{cf} \times O(ML^2 + N \log_2 N)$
SSCR-TR [50]	Not applicable	$I_{sscr} \times O(N + N \log_2 N)$
IVO-TR [52]	$N_s \times O(N_i N_h + N_h N_o + L N_o)$	$O(N_i N_h + N_h N_o)$
ELM-TR [53]	$O(N_i N_h + N_s N_h N_o)$	$O(N_i N_h + N_h N_o)$

CHAPTER 6

RESULTS AND DISCUSSION

This chapter presents results from MATLAB simulation of the proposed methods. The results are for the methods in Chapter 5, which are the three tone-reservation methods based on SOCP, LP and IRLS algorithms and the other two additive signal mixing methods formerly abbreviated as LCSA and LCASM. For each method, results from MATLAB simulations are first presented and discussed followed by a comparison with those from other related and recently published PAPR reduction methods.

The OFDM and MIMO parameters used during simulations are listed in Table 6.1. The simulations were run for 10^4 randomly generated OFDM symbols to enhance the accuracy of the results. For each OFDM symbol, subcarriers were modulated with data from one of the most commonly used digital modulation schemes, namely BPSK, QPSK and M-ary QAM for $M = 16, 64, 256, 1024$ or 2048.

The performance analysis for each method is based on the results of PAPR reduction, BER degradation and increase in average transmit power. The evaluation of PAPR reduction capability was done at the conventionally accepted CCDF value of 10^{-3}. For the measurement of the level of BER degradation, the amount of SNR per bit, i.e. E_b/N_o, required at the BER probability of 10^{-3} was used as a performance indicator in both AWGN and Rayleigh-fading channels.

Table 6.1. Simulation parameters

FFT size, N	16, 32, 64, 128, 256, 512, 1024, 2048
Modulation	BPSK, QPSK, M-ary QAM
Number of symbols per frame	10^4
Time-domain signal structure	Real and complex
Type of subcarrier reservation	Random, block and equidistant
Oversampling factor $\mathcal{L}$	4
HPA model	Rapp model with $p = 2$, IBO = 7, 8
MIMO configuration	2×2 with Alamouti STBC
Channel model	AWGN, Rayleigh flat-fading

6.1 Results from SOCP-based Tone Reservation Method

This section presents results of PAPR reductions achieved by the proposed SOCP-TR method when 5% and 20% of subcarriers were reserved and different modulation schemes were employed with real and complex signals. The two percentages of tone reservations, 5% and 20%, were purposely chosen to give two representative performances of the method when a small and a large number of subcarriers are reserved, respectively.

The proposed method was initially tested to confirm that it minimized the highest peaks of the input signal as required. Since its algorithm is designed for solving a minimax problem to find the required peak-reduction coefficients for a given input OFDM signal, the distribution of the resultant peak-reduced signal amplitudes is expected to follow a uniform distribution. This was found to be the case regardless of the type of subcarrier modulation or the input signal structure, real or complex, as depicted in Figure 6.2 and Figure 6.1, thus verifying the correct functioning of the proposed algorithm.
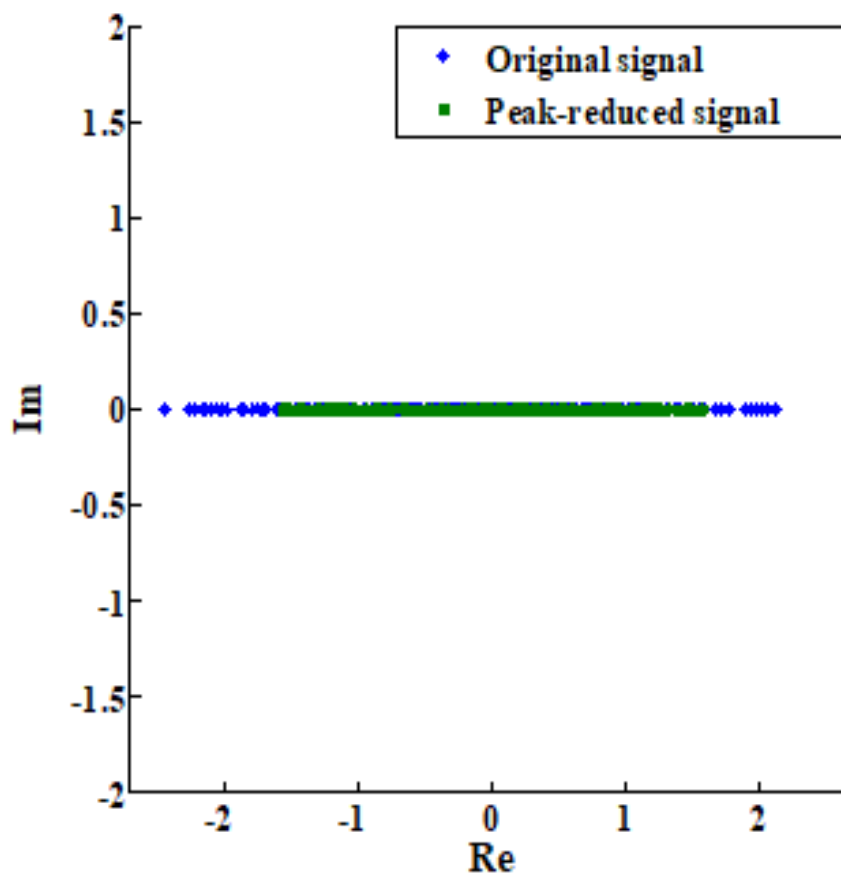
Figure 6.1. Scatter plot of 16-QAM real signal
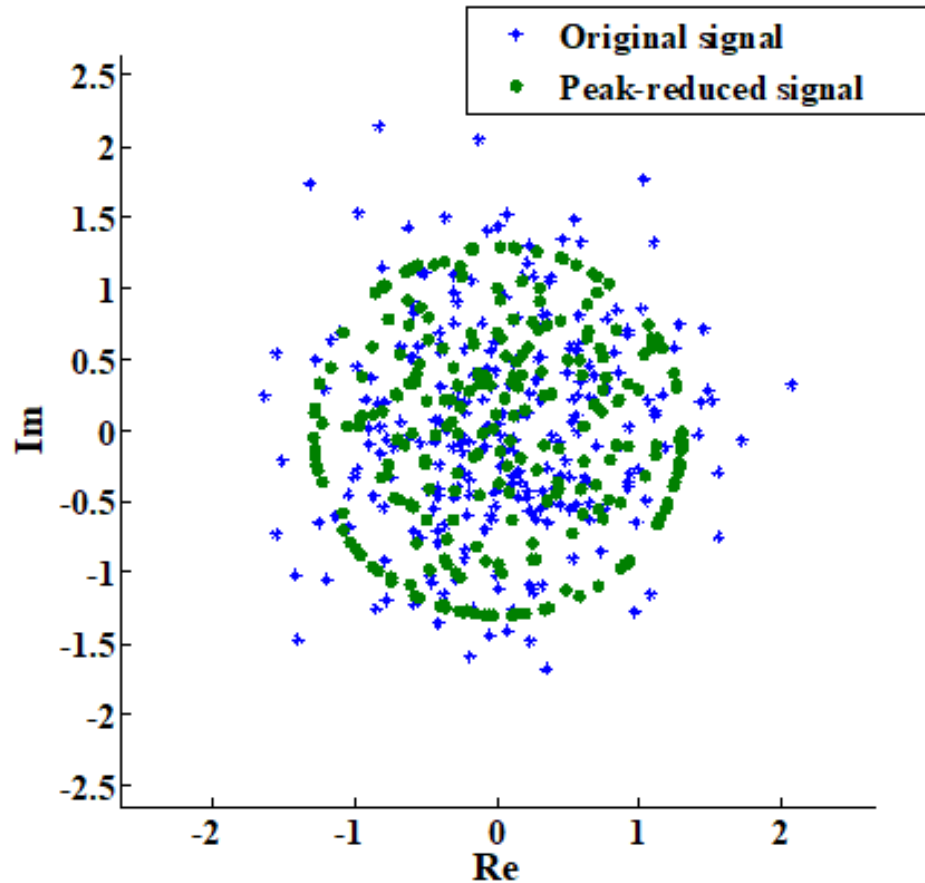
Figure 6.2. Scatter plot of 16-QAM complex signal

After successful testing of the proposed method, it was applied to reduce PAPR in an OFDM system with $N = 256$ subcarriers. The variation of PAPR with the number of bits per symbol when the system employed BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, 1024-QAM and 2048-QAM subcarrier modulations is given in Figure 6.3. A symbol in these modulation schemes carries 1, 2, 4, 6, 8, 10 and 11 bits, respectively. The legends "0% for PRCs", "5% for PRCs" and "20% for PRCs" are indications that 0, 13 and 52 subcarriers out of 256 were reserved for peak reduction, respectively.

From Figure 6.3, two there are two clear observations. The first one is that a real signal has a higher PAPR than a complex signal of the same length, typically by approximately 2.4 dB, regardless of the type of subcarrier modulation scheme employed. This observation was indicated for the three percentages of tone-reservations and could be attributed to the requirement of conjugate symmetry in the formation of a real signal, which can limit destructive addition of amplitudes in the OFDM signal.

The second observation is that the PAPR reduction capability of the proposed method increases with the number of reserved subcarriers. This observation is brought

out clearly by the results of PAPR reductions in Table 6.2 and Figure 6.4 that were achieved with different modulation schemes with 5% and 20% of tones reserved.

For both real and complex signals, the proposed method is capable of attaining good PAPR reductions. Specifically, with 5% of subcarriers reserved for PRCs, the respective average PAPR reduction was 4.37 dB and 4.59 dB for the real and complex signals for the 7 modulation schemes. In the case of reserving 20% of tones, the average reductions were 7.93 dB and 7.71 dB for the real and complex signals, respectively. Thus, the difference in PAPR reduction between the real and complex signals is approximately ±0.2 dB, implying that the structure of the signal has an insignificant effect on the amount of PAPR reduction.

To assess the effect of the type of subcarrier modulation on PAPR reduction, the standard deviation at each level of subcarrier reservation was computed over the seven modulation schemes. The deviations from the mean PAPR reduction for complex signals were 0.20 and 0.21 for 5 and 20% tone reservations, respectively. For the real signals, the respective deviations were 0.21 and 0.14 for the two amounts of tone reservations. These results show that the type of subcarrier modulation does not significantly affect the extent of PAPR reductions.
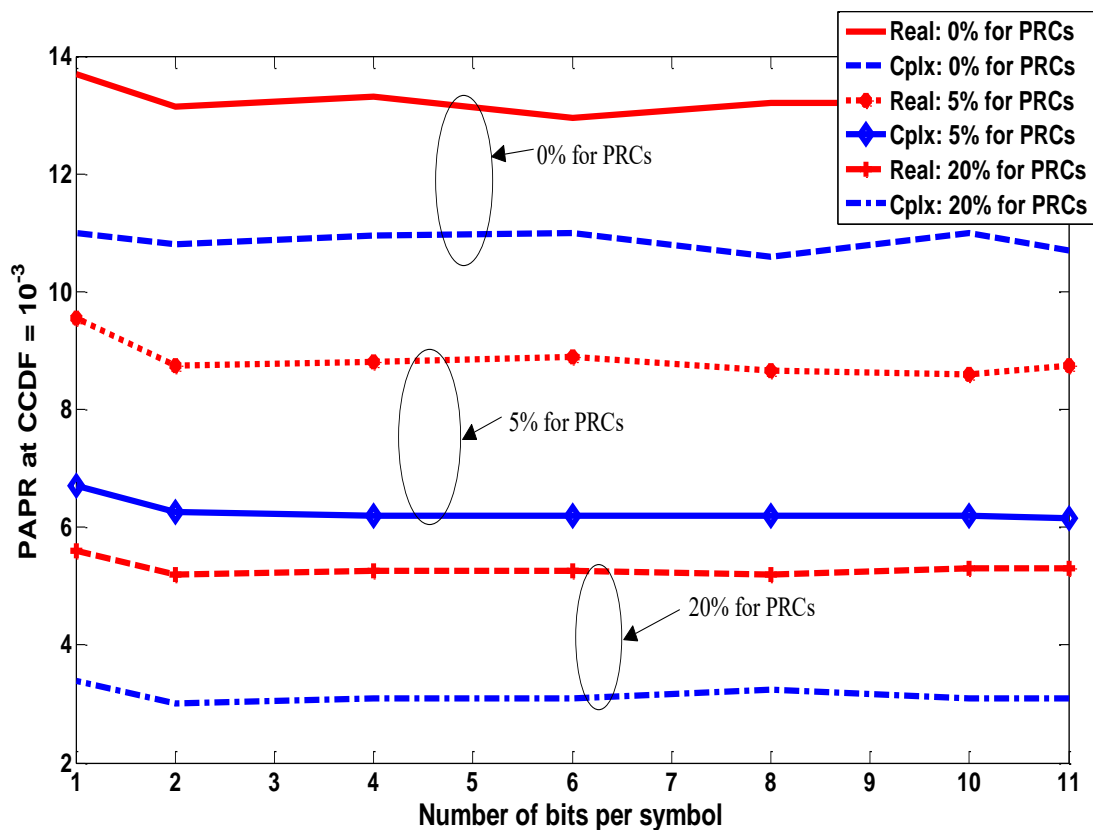
Figure 6.3. PAPR versus modulation bits for the SOCP-TR method

Table 6.2. PAPR reduction by SOCP-TR for real and complex signals for different modulations

Signal structure	Percentage reservation	BPSK	QPSK	16 -QAM	64 -QAM	256 -QAM	1024 -QAM	2048 -QAM
Real signal	5%	4.15	4.40	4.05	4.50	4.60	4.55	4.35
	20%	8.10	7.95	8.05	7.70	8.00	7.90	7.80
Complex Signal	5%	4.30	4.55	4.75	4.80	4.40	4.80	4.55
	20%	7.60	7.80	7.85	7.90	7.35	7.90	7.60
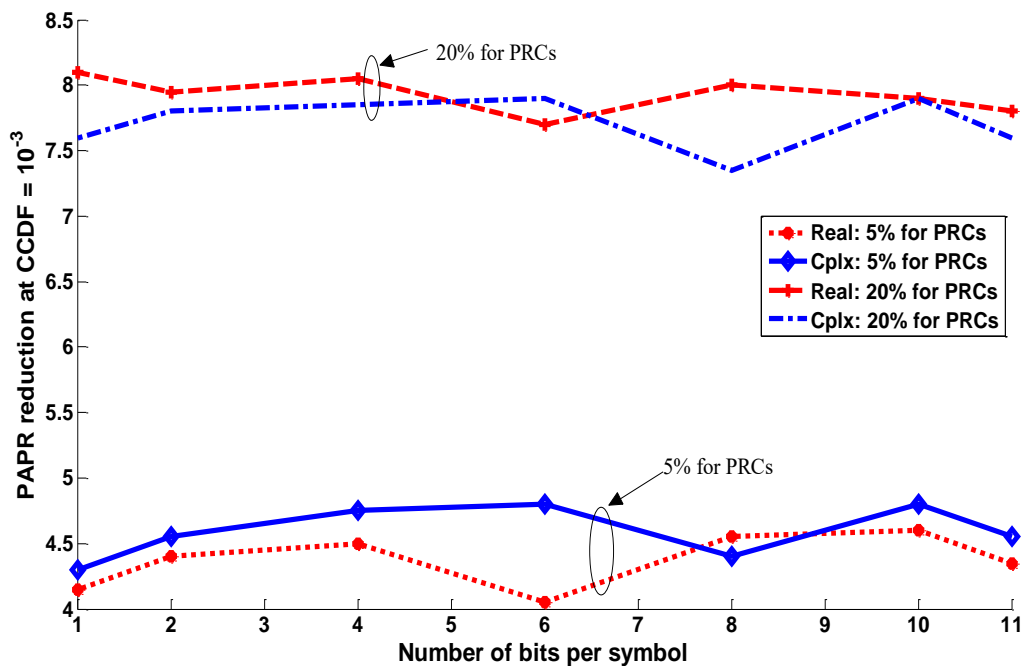
Figure 6.4. PAPR reduction versus number of modulation bits when using SOCP-TR method

6.2 Results from LP-based Tone Reservation Method

In this section, simulation results of the LP-TR method proposed in Subsection 5.2.3 and their analysis are presented. In order to simulate the method in MATLAB, the problem in equation (5.18) was firstly re-formulated to separate real and imaginary inequality constraints. An inbuilt MATLAB *linprog* function for solving linear programs was called with interior-point algorithm to solve the re-formulated problem to find the peak-reduction signal given an OFDM signal.

In addition, an oversampling factor of four was applied to the OFDM signal for ensuring that the PAPR of the discrete-time signal was close to that of the continuous-time signal. Moreover, the Rapp model of HPA with a smoothness parameter value of 2

and an IBO of 8 dB were used during the simulations for BER measurements. This IBO was approximately 1 dB greater than the PAPR of the signal at CCDF = 10^{-3} and ensured that the HPA clipped less than 1% of the input signal amplitudes.

Figure 6.5 shows the capability of the proposed method to reduce PAPR in a system with 64 subcarriers, out of which 4 and 8 were reserved, corresponding to a reservation of 6.25% and 12.5% of tones in a symbol, respectively. The two reservations were intended to give two representative performances of the method when a small and a large number of tones are reserved.

For the two cases of 6.25% and 12.5% of tone reservation, the PAPR reduced by 4.06 and 5.75 dB, respectively. The achieved reductions were at a cost of a slight increase in the average power by 0.46 and 0.19 dB, respectively. On an OFDM system with 256 subcarriers with 5% reserved, the PAPR reduction was 4.66 dB, as shown in Figure 6.6, and the transmit power increased slightly by 0.57 dB.
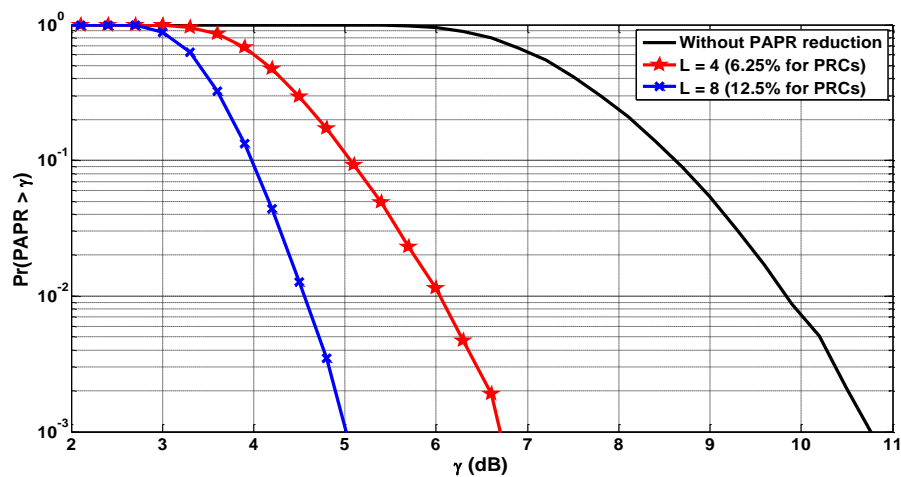
Figure 6.5. PAPR reduction by LP-TR method for 6.25% and 12.5% of tones reserved, for *N* = 64
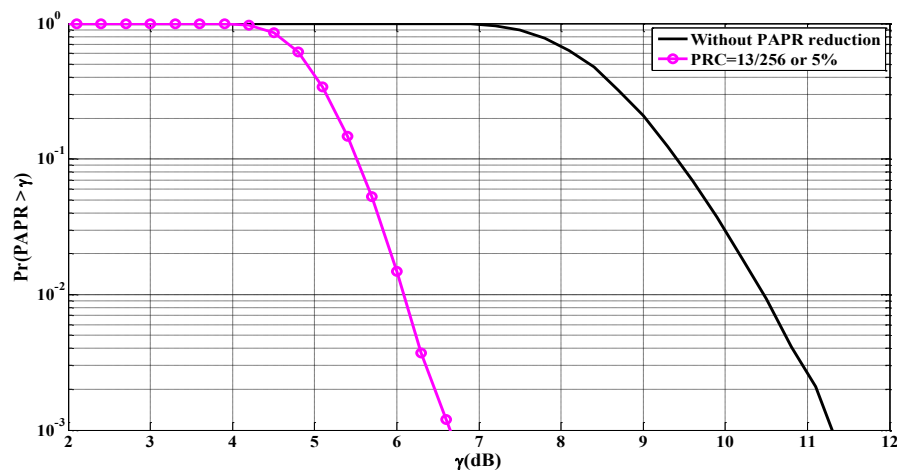
Figure 6.6. PAPR reduction by LP-TR method for 5% of tones reserved, for *N* = 256

The above results demonstrate the excellent capability of the proposed method to reduce PAPR with only a few subcarriers reserved for PRCs. In addition, while the reduction in PAPR will tend to rise with the number of reserved tones, this is conversely true for the increase in the average power of the signal.

Additionally, the proposed LP-TR method was compared to the SOCP-TR method proposed in this thesis and five other related tone reservation methods found in the literature and which were earlier on reviewed and abbreviated as LSA-TR, CF-TR, SSCR-TR, IVO-TR and ELM-TR. The iterations in CF-TR, LSA-TR and SSCR-TR methods are 2, 3 and 5, respectively.

The IVO-TR and ELM-TR methods have $N_o = 100$ peak-cancelling signal classes (or neurons in the output layer) and $N_h = 1000$ neurons in the hidden layer. The size of the training and test data sets is 10^5 and 10^4 of random OFDM input signals, respectively. In addition, each training target peak-reduction signal is generated from 100 iterations of CC-TR algorithm.

The performance comparison was for an OFDM system with 64 subcarriers, out of which 4 are reserved for peak-reduction coefficients, and QPSK modulation was used on all subcarriers. The PAPR reduction performances for the seven methods are depicted in Figure 6.7. The results of PAPR reduction and increase in average power are provided in Table 6.3. From these results, the proposed LP-TR method has better PAPR reduction performance than the proposed optimal SOCP-TR method and the other five methods. The proposed method attains more reduction in PAPR of 0.45, 1.75, 0.90, 0.87, 0.51 and 0.21dB than SOCP-TR, LSA-TR, CF-TR, SSCR-TR, IVO-TR and ELM-TR methods, respectively.
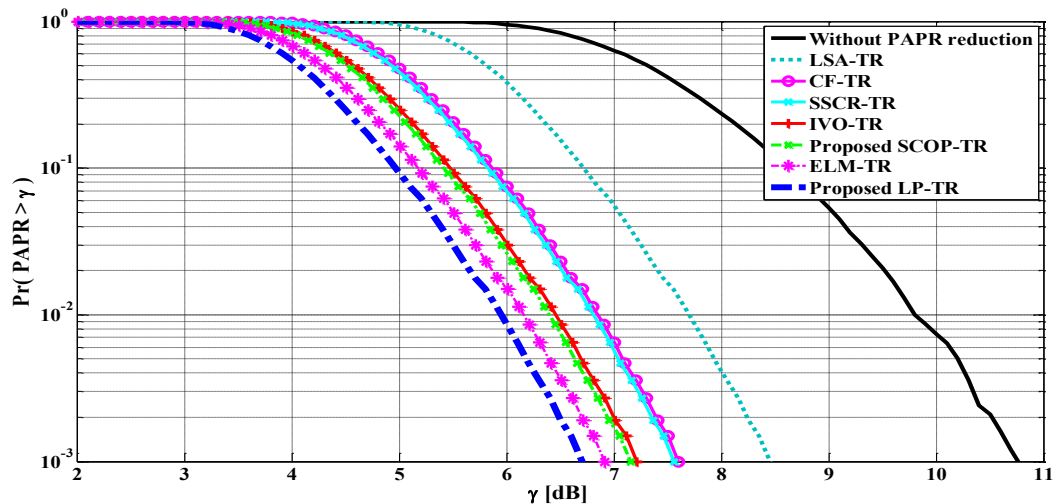
Figure 6.7. PAPR reduction performance comparison for the LP-TR method

131

Table 6.3. Comparison of reduction in PAPR and increase in average power for the LP-TR method

PAPR reduction Method	PAPR reduction [dB]	Average Power increase [dB]
Proposed LP-TR	4.06	0.46
Proposed SCOP-TR	3.61	0.51
ELM-TR [53]	3.85	0.59
SSCR-TR [50]	3.19	0.25
IVO-TR [52]	3.55	0.57
CF-TR [49]	3.16	0.84
LSA-TR [48]	2.31	0.30

In addition to PAPR reduction comparison, the BER performance of each method in an AWGN channel is shown in Figure 6.8. The amounts of the required E_b/N_0 at BER = 10^{-3} for each method under comparison are also listed in Table 6.4. The BER curve with the legend "Theoretical" is the expected BER performance for the QPSK modulation obtained by the formula in equation (3.66).

The performance with the legend "Without PAPR reduction" represents the BER when the signal was transmitted through a HPA, with IBO set at 0 dB, without prior reduction in PAPR; thus, it presents the worst amount of BER degradation. Since all the techniques being compared are tone-reservation methods, they are not expected to degrade the BER, hence their BER performances should be nearly the same. However, due to the value of the IBO used and the higher reduction in PAPR by the LP-TR method, it has a minor BER improvement of 0.02, 0.02, 0.06, 0.03, 0.08 and 0.11 dB over the SOCP-TR, ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR methods, respectively.

6.3 Results from IRLS-based Tone Reservation Method

This section presents results of the IRLS-TR method proposed in Subsection 5.2.4 of Chapter 5. The performance of the method was assessed from MATLAB simulations of OFDM system with different percentages of reserved subcarriers. Reserving of subcarriers in an OFDM symbol was done in either a random or a fixed manner.

The aim of the proposed method is to reduce the convergence rate of the optimal LP-TR method in Section 5.2.3, while at the same time achieving high PAPR reductions by generating a peak-reduction signal with the highest peaks close to the desired ones as

shown in Figure 6.9.

In the figure, the algorithm generated a peak-reduction signal with the largest peak of magnitude of 1.21 at the discrete-time instant $n = 137$ to approximate the maximum peak of magnitude 1.81 in the desired peak-reduction signal at the same time instant. Furthermore, as it can be found from the figure, all the major peaks in the actual PRS are below the highest peaks in the desired PRS.
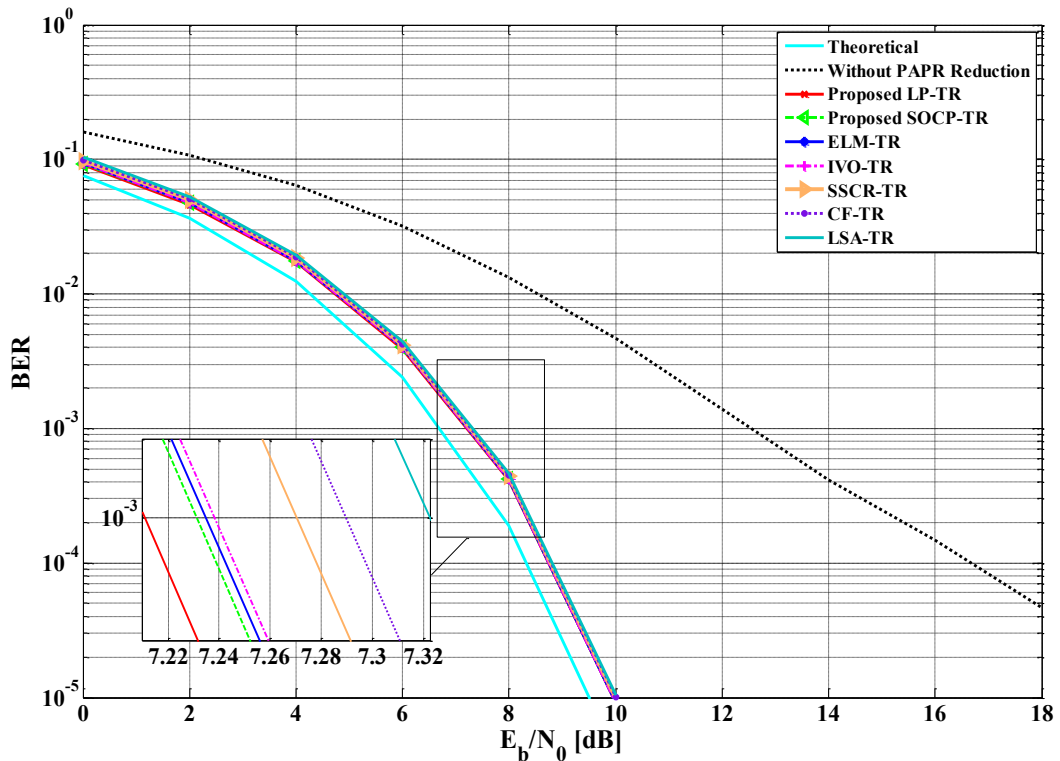
Figure 6.8. Comparison of BER performances over AWGN channel for the LP-TR method

Table 6.4. Comparison of required E_b/N_0 for the LP-TR method

Method	E_b/N_0 [dB] at BER = 10^{-3}
Proposed LP-TR	7.21
Proposed SOCP-TR	7.23
ELM-TR	7.23
SSCR-TR	7.27
IVO-TR	7.24
CF-TR	7.29
LSA-TR	7.32

133
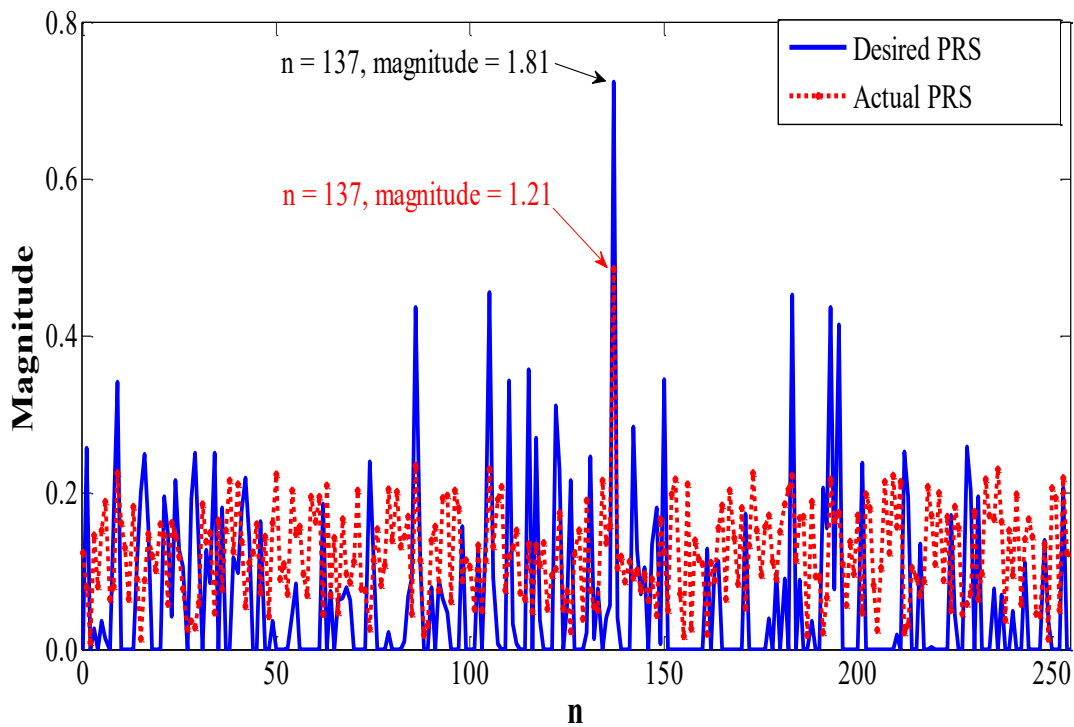
Figure 6.9. Desired and actual peak-reduction signals

Because of the highest peaks in the actual PRS being below those in the desired PRS, on subtracting the former from the OFDM signal, no new peaks are generated at other locations. This avoids the problem of peak re-growth experienced in tone reservation methods that utilize a kernel PRS [135], [136] to cancel the highest peaks, one at a time iteratively.

The ability of the proposed method to reduce PAPR is illustrated in Figure 6.10 for an OFDM system with 256 subcarriers, all QPSK-modulated, out of which 2, 4, 8, 13, 22 and 52 were reserved for PRS. The selection of these numbers of reserved tones is to enable have a representative performance spectrum from small to large amount of tone reservation and bring out the dependency of PAPR reduction on the number of reserved subcarriers.

It is observable from Figure 6.10 that PAPR reduction increases with the number of reserved subcarriers. However, even with a few reserved subcarriers, a good PAPR reduction performance can be realized. This is clearly demonstrated in Table 6.5 for the achieved PAPR reduction at CCDF of 10^{-3} for different numbers of reserved tones. For example, with 4 and 13 tones reserved, which are equivalent to approximately 1.6% and 5% of total tones reserved, 3.77 dB and 5.84 dB reductions were attained, respectively.
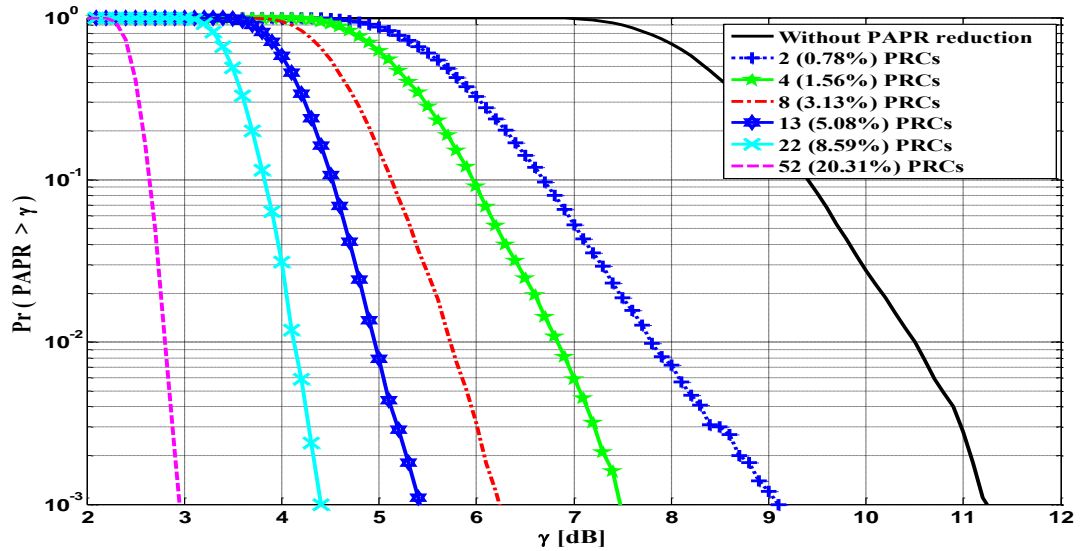
Figure 6.10. PAPR reduction with different number of reserved subcarriers out of 256

Table 6.5. PAPR reduction by IRLS-TR method for different subcarrier reservations

Number of reserved subcarriers, L	Percentage of reserved tones	PAPR Reduction [dB] at CCDF = 10^{-3}
2	0.78 %	2.15
4	1.56 %	3.77
8	3.13 %	5.02
13	5.08 %	5.84
22	8.59 %	6.85
52	20.31 %	8.30

The PAPR reduction performances of the proposed IRLS-TR and LP-TR methods were also compared for an OFDM system with 64 total subcarriers out of which 6.25% of the tones were reserved. The performance by each method is shown in Figure 6.11. The proposed IRLS-TR method has a better PAPR reduction performance of about 1 dB over the LP-TR method. This is in addition to a smaller increase in the average transmit power of 0.19 dB below that of the latter. Furthermore, based on the above results and those reported earlier in Table 6.3, it is clear that the proposed method outperforms other methods proposed in literature.

The results just presented are for randomly reserved subcarrier locations. Such a reservation would require the receiver to be furnished, symbol by symbol, with the information about the indices of reserved subcarriers for the demodulation of symbols.
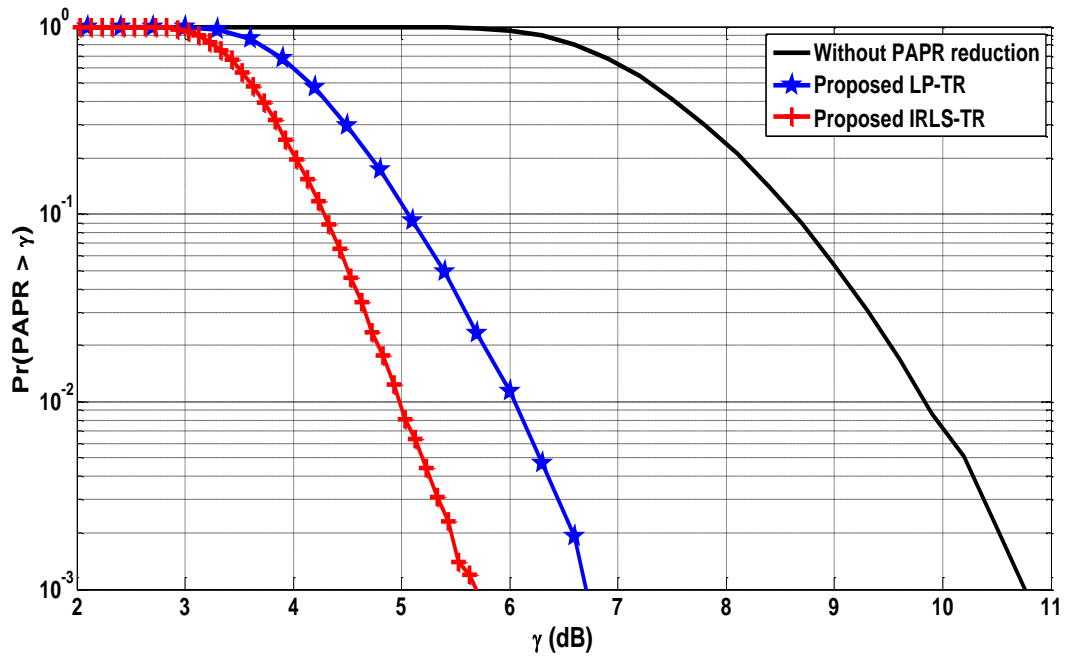
Figure 6.11. Comparison of PAPR reduction by IRLS-TR and other methods

However, the transmission of side information about the locations of reserved subcarriers would additionally lower the data rate. A better way to avoid transmitting the side information about subcarrier reservation is to pre-select and permanently fix the reserved subcarrier locations in all OFDM symbols. This should be the case if pre-set reservation can achieve a comparable PAPR reduction performance to the random reservation. Two common ways of pre-setting locations of reserved subcarriers are either to place them in a block of contiguous positions or to space them equally across the entire spectrum of an OFDM symbol.

The results of an investigation carried out on the PAPR reduction performance of the two types of fixed reservations are given in Table 3.1. For the purpose of performance comparison, the results achieved with randomly reserved tones are also included in the same table. The results in the table show that randomly reserved subcarriers provide better PAPR reduction than the two approaches of fixed positions of reserved subcarriers for an equivalent number of reserved tones in an OFDM symbol.

For example, when 6.25% of subcarriers are reserved out of 64, the random reservation achieves 1.12 dB more of PAPR reduction than the two types of fixed reservations as shown in Figure 6.12. However, the equispaced reservation has the lowest increase in the average power, which is 0.07 and 0.15 dB below the random and contiguous reservations, respectively.

136

Table 6.6. Performance comparison of random, contiguous and equispaced tone reservations

Type of subcarrier reservation	L = 2 (3.125%)		L = 4 (6.25%)		L = 8 (12.50%)	
	PAPR reduction [dB]	Power increase [dB]	PAPR reduction [dB]	Power increase [dB]	PAPR reduction [dB]	Power increase [dB]
Random	3.62	0.27	5.07	0.27	6.46	0.03
Contiguous	3.60	0.32	3.95	0.35	4.91	0.21
Equispaced	3.45	0.27	3.95	0.20	4.04	0.06
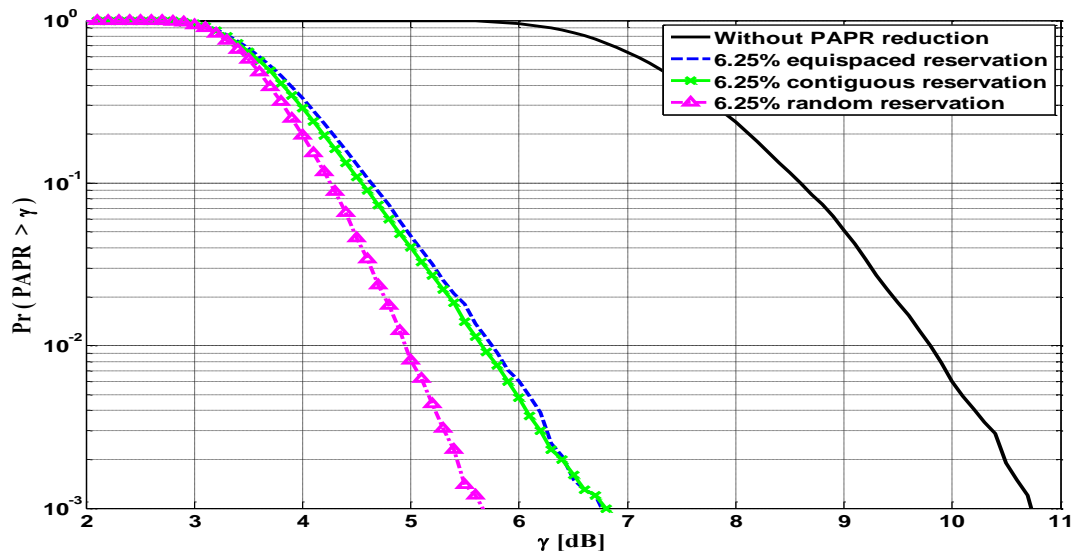
Figure 6.12. PAPR reduction with random, contiguous and equispaced 6.25% reservation

Considering the effect of the number of reserved tones, in this case $L = 2$, 4 and 8, on the PAPR reduction for the two fixed reservations, the PAPR tends to reduce by higher amounts with contiguously reserved subcarriers than with equally spaced ones, although at the expense of increased average transmit power. This is demonstrated in Figure 6.13 for the case of $L = 8$ reserved subcarriers out of 64 and also in Table 3.1, where the contiguously reserved tones have 0.87 dB more reduction in PAPR than the equally-spaced tones but at the expense of an extra 0.15 dB increase in signal power.

In addition to achieving a good PAPR reduction, the proposed IRLS-TR algorithm does not degrade the BER performance because reserved subchannels are dropped during the demodulation procedure to recover data in the receiver. Thus, the BER performance of the method will always be close to the theoretical one given by equation (3.66) and very similar to the performance in Figure 6.8 of the LP-TR method.
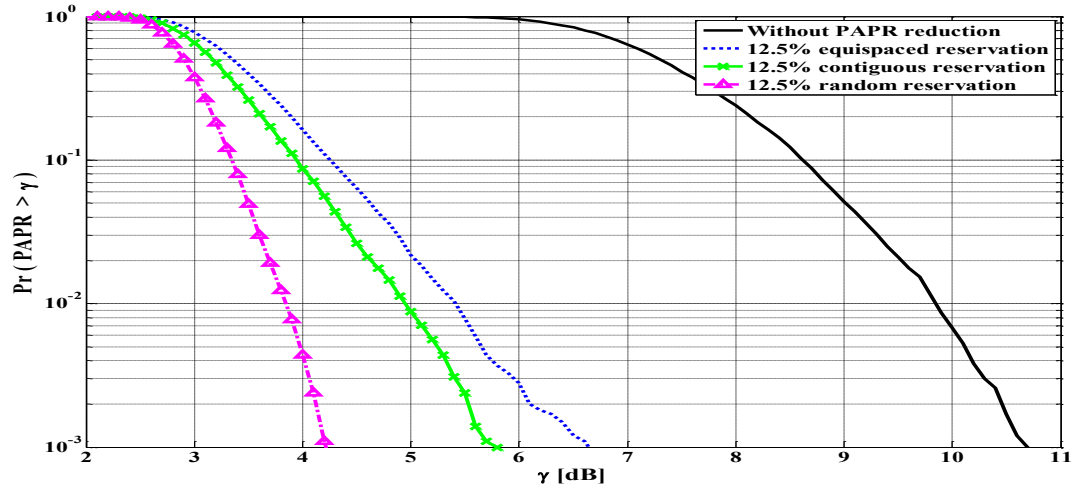
Figure 6.13. PAPR reduction with random, contiguous and equispaced 12.5% reservation

Lastly, an investigation was carried out to establish if there existed any relationship between PAPR reduction and the total number of subcarriers. Figure 6.14 shows the simulation results for this investigation for the same subcarrier modulation when 5%, 10%, 15% and 20% of subcarriers were reserved for PAPR reduction. From these results, it can be observed that PAPR reduction varies almost linearly with the binary logarithm of the total number of subcarriers in a system. For the subcarrier reservations of 5 and 15%, PAPR reductions can be estimated by straight-line equations $0.4 \log_2(N) + 0.6$ and $0.4 \log_2(N) + 1.7$, respectively.

The linear variation of PAPR reduction with $\log_2 N$ can help to estimate the expected PAPR reduction when given both the total number of subcarriers and the percentage of reservation. In addition, since the linear relationships for different percentages of reserved subcarriers have almost the same gradient, it is easy to estimate the possible increase in PAPR reduction at a given N owing to increase of reserved subcarriers. For the two cases of 5 and 15% reservation, the difference of 1.1 dB between the vertical intercepts of the two line equations is the PAPR reduction gained by increasing the percentage of reserved subcarriers from 5 to 15%.

6.4 Results from LCSA Method

This section presents and discusses the results of PAPR reduction and BER performance of the LCSA method proposed in Subsection 5.2.5 of Chapter 5. The section also gives a detailed performance comparison of the LCSA method and the other two methods, LP-TR and IRLS-TR, proposed in this thesis.
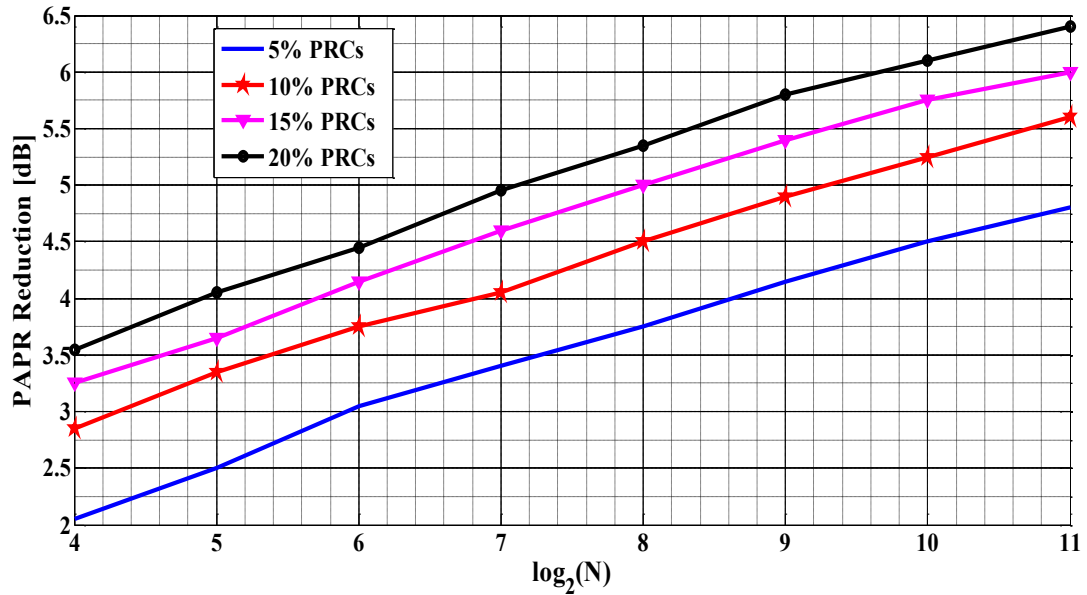
Figure 6.14. Variation of PAPR reduction with number of subcarriers

The proposed LCSA method is designed to generate a PRS signal that exactly matches the desired one. This is illustrated in Figure 6.15 for an OFDM system with 64 total subcarriers and in which the desired PRS is generated in such a way as to have only four nonzero samples with an oversampling factor of four. As can be observed from the figure, the actual and the desired peak-reduction signals are identical.

Without loss of generality, an OFDM system with 256 QPSK-modulated subcarriers was chosen for assessing the PAPR reduction capability of the method. The PAPR of the system was targeted for reduction using different peak-reduction signals having $M = 3, 6, 13, 19$ and 26 nonzero PRS samples, approximately corresponding to 1.2, 2.3, 4.8, 6.9 and 9.2% data-rate losses, respectively.

The CCDFs for different values of M are portrayed in Figure 6.16. From this figure, it is clear that reduction in PAPR increases with M. This trend can be well observed in the results of PAPR reduction in Table 6.7. In addition, these results show that using a peak-reduction signal with only a few nonzero samples, and hence a corresponding small data-rate loss, the proposed method can achieve a significant reduction in PAPR.

In addition, the last row of Table 6.7 demonstrates that the proposed method barely affects the average transmit power. For example, when the PRS has $M = 26$ nonzero PRS samples, which is unlikely in practice, the average power drops by just -0.21 dB or 4.7% from the original value before PAPR reduction. From the same table,

it can be observed that for all the simulations the average transmit power is slightly decreased which translates into a small increase in the power efficiency of the system given that PAPR is also reduced.

In addition, the plots in Figure 6.17 of power spectral density of the original and peak-reduced OFDM signals for different values of M show that using the proposed method does not significantly increase the out-of-band radiations. This is demonstrated by the worst case of PRS with $M = 26$, which presents a relatively small out-of-band radiation of -61.90 dB, which is just 1.63 dB above the radiation of -63.53 dB from the original OFDM signal.
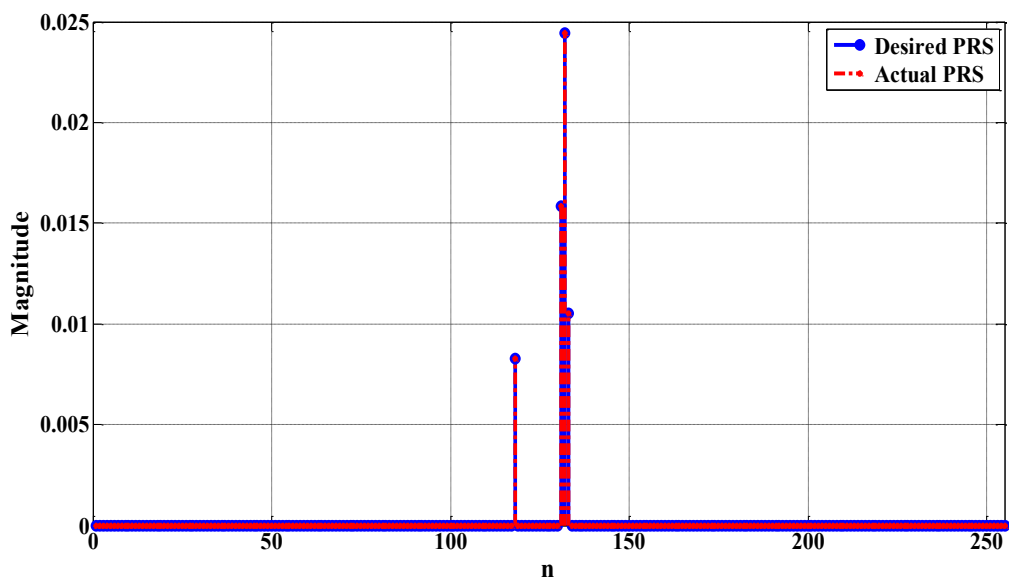
Figure 6.15. Desired and actual peak-reduction signals for the proposed LCSA method
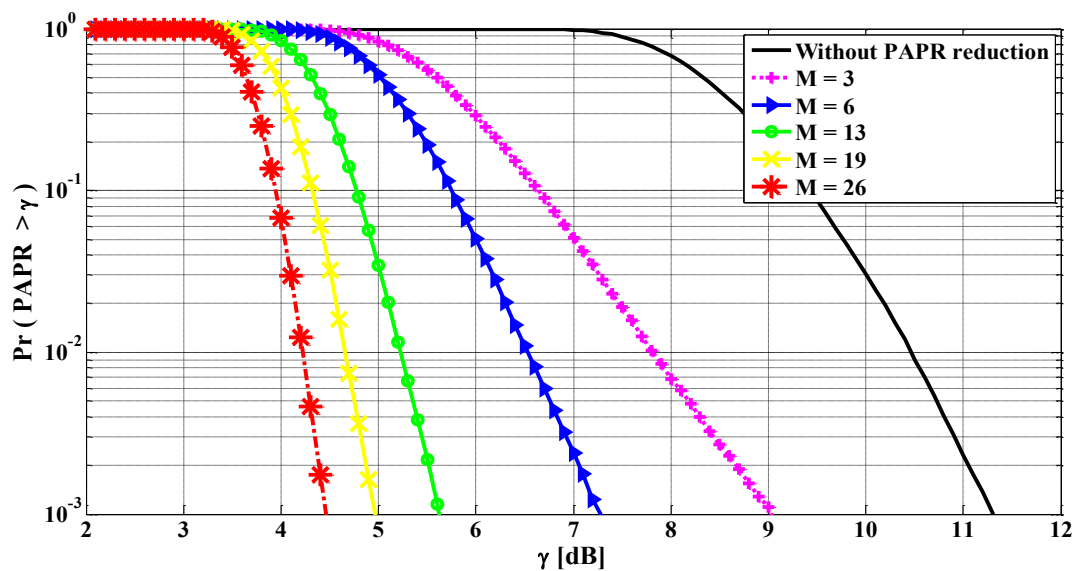
Figure 6.16. PAPR reduction by LCSA method for different PRS signals for N = 256

140

Table 6.7. PAPR reduction at CCDF = 10-3 for different PRS signals for N = 256

Number of nonzero samples in PRS	3	6	13	19	26
Percentage loss in data rate	1.16%	2.29%	4.83%	6.91%	9.22%
PAPR Reduction [dB]	2.28	4.04	5.69	6.35	6.85
Increase in average power [dB]	-0.02	-0.05	-0.10	-0.15	-0.21
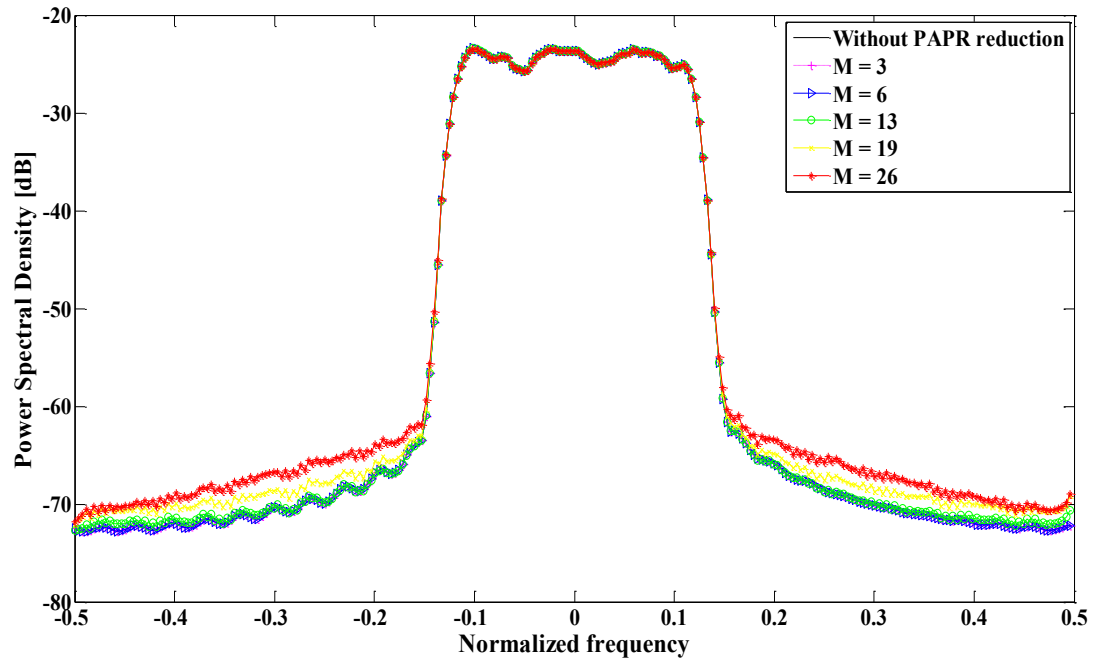
Figure 6.17. Power spectral densities for different peak-reduced signals

Additionally, the proposed method was compared to the previously proposed LP-TR and IRLS-TR methods. A QPSK-modulated OFDM system with 64 subcarriers was employed for this purpose. A PRS with four nonzero samples was used for the proposed LCSA method, while four subcarriers were reserved for peak-reduction coefficients in the case of the two tone-reservation methods.

The CCDFs for the three methods are plotted in Figure 6.18 and their respective PAPR reductions are presented in Table 6.8. Compared to the suboptimal IRLS-TR method, the LCSA method has a lesser PAPR reduction by 0.94 dB. However, the LCSA method slightly reduces the average transmit power by 0.12 dB, compared to a 0.27 dB increase by IRLS-TR method.

In comparison with the optimal LP-TR method, the LCSA technique has a similar performance since their PAPR reductions are separated barely by 0.07 dB. However, the

LP-TR method causes a power increase of 0.58 dB more than the LCSA method. Based on this comparison, it is clear that the proposed LCSA method also outperforms the other methods in Table 6.3.

In addition to the PAPR reduction performance analysis and comparison, the BER performance of the proposed method was evaluated using a Rapp model of HPA with IBO set at 7 dB and parameter p at 2. The setting of the IBO was just above the PAPR of 6.57 dB at CCDF equal 10^{-3} to ensure that the HPA clipped less than 1% of signal amplitudes.

The transmission channel was assumed to have AWGN channel characteristics. At the receiver, the appended samples of peak-reduction signal were added back to positions of clipped-amplitudes in the peak-reduced signal. This was followed by demodulation and de-mapping of modulation symbols to recover transmitted binary data. The recovered binary data was then used to calculate the BER.

The variation of BER with E_b/N_0 is illustrated in Figure 6.19 and the required E_b/N_0 for each method is given in Table 6.9. From the results of the required E_b/N_0, there are minor differences between the LCSA method and the LP-TR and IRLS-TR methods proposed earlier. This shows that the LCSA technique has a similar BER performance to tone reservation methods; specifically it does not interfere with the BER of the underlying system.
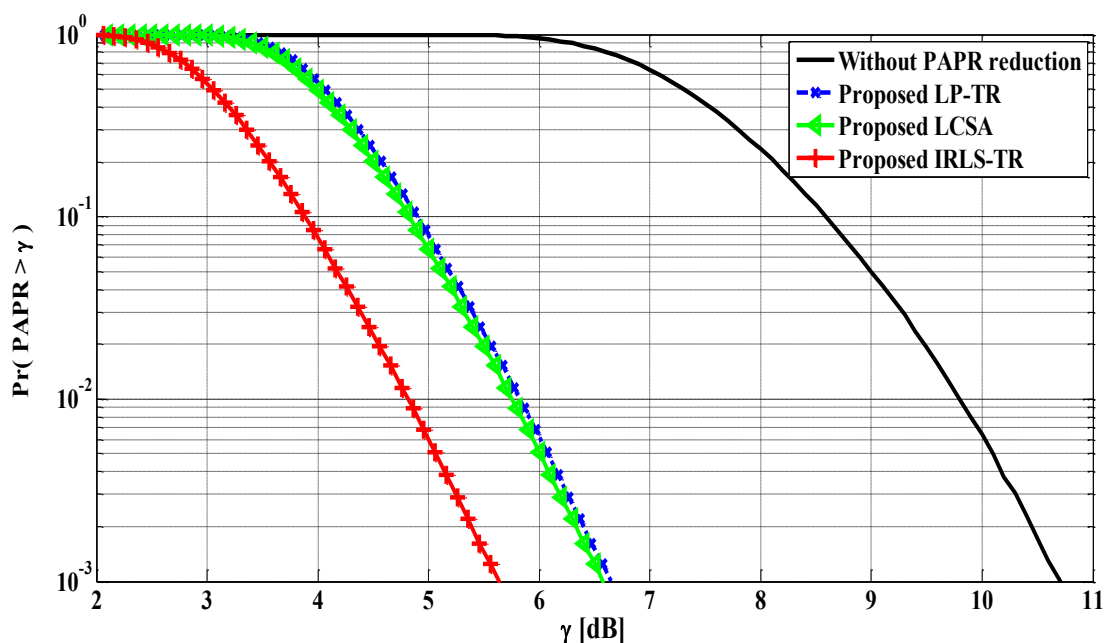
Figure 6.18. PAPR reduction performance comparison for the proposed LCSA method

Table 6.8. PAPR reduction comparison for the proposed LCSA method at CCDF = 10^{-3}

Method	PAPR reduction [dB]	Power increase [dB]	Data-rate loss
Proposed LCSA	4.13	-0.12	5.88%
Proposed IRLS-TR	5.07	0.27	6.25%
Proposed LP-TR	4.06	0.46	6.25%
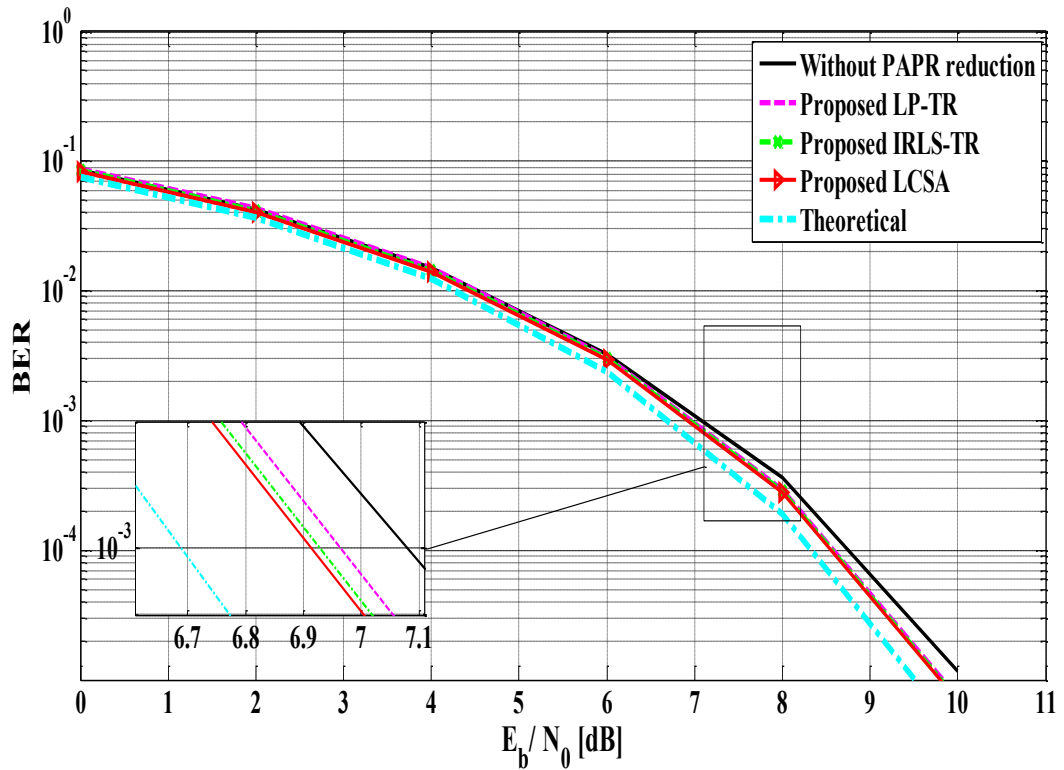
Figure 6.19. BER performance comparison for the proposed LCSA method

Table 6.9. Comparison of required E_b/N_0 for the LCSA method

Method	E_b/N_0 [dB] at BER = 10^{-3}
Proposed LCSA	6.91
Proposed IRLS-TR	6.93
Proposed LP-TR	6.97

6.5 Results from LCASM Method for MIMO-OFDM Systems

This section covers the PAPR reduction performance of the LCASM method, proposed in Subsection 5.2.6, in MIMO-OFDM systems. The system in Figure 3.18, which has OFDM combined with a 2×2 MIMO configuration with an Alamouti space-

143

time block encoder is used in assessing the performance of the method.

The PAPR reduction performance of the LCASM technique is also compared to the performances of four other related methods, which were reviewed in the literature review chapter of this thesis, namely ACT, SCS-SLM, CSC and S-TR. A MIMO-OFDM system with 128 subcarriers and QPSK modulation is used in this comparison. The acceptable loss in data-rate is set at 15% and a peak-reduction signal with 23 nonzero samples is generated.

Figure 6.20 shows the performances of the five methods. The respective PAPR reductions at the CCDF value of 10^{-3} are given in Table 6.10. From these results, it is evident that the PAPR reduction capability of the proposed LCASM method is superior to the other four methods. The proposed method reduces PAPR more than SCS-SLM, S-TR, ACT and CSC methods by 4.40, 3.35, 2.85 and 0.90 dB, respectively.

Additionally, the BER performance of the method was evaluated with a HPA modelled using the Rapp's model with parameter p set at a value of 2. The IBO was fixed at 12.5 dB to guarantee that less than 1% of signal amplitudes were clipped by the HPA since the PAPR of signals for the simulated MIMO-OFDM system was 11.8 dB at the CCDF value of 10^{-3}.

During the simulation, the composite signals derived from the PAPR reduction operation were first amplified through the HPA followed by a transmission through Rayleigh-fading channels with AWGN. At first, an analysis was done on the effect of different peak-reduction signals generated by the LCASM method on the BER. This was followed by a BER performance comparison with the other methods.
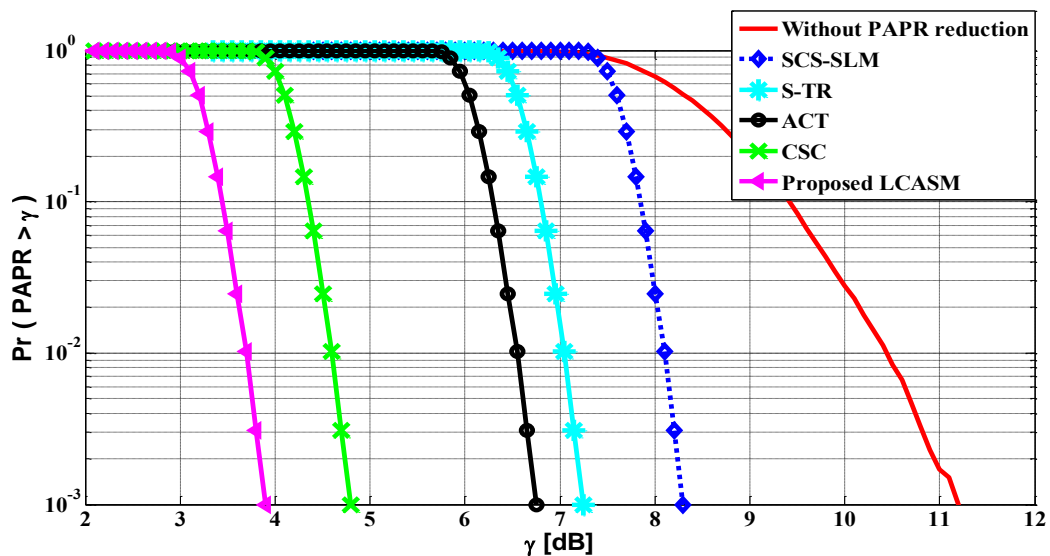
Figure 6.20. PAPR reduction for MIMO-OFDM system by different methods for $N = 128$

Table 6.10. PAPR Reduction comparison for LCSAM for CCDF = 10^{-3} and N = 128

Method	Proposed LCASM	CSC [73]	ACT [29]	S-TR [54]	SCS-SLM [72]
PAPR reduction [dB]	7.32	6.39	4.46	3.97	2.89

The BER performances of the LCASM method when employing different peak-reduction signals are shown in Figure 6.21. The two performances marked "Ideal" represent cases when the transmit signals were not subjected to PAPR reduction. In one of the two cases, a HPA was not used, while in the other it was present but was backed-off with 12.5 dB, which is greater than the expected PAPR; thus the two cases are similar and represent the situation when there is no BER degradation.

The performance with legend "No reduction, IBO=0" represents the worst BER degradation. It corresponds to the performance of a system with the HPA IBO set to 0 dB and there is no PAPR reduction; hence, all amplitudes above the average value are clipped. The rest of the BER performances for various values of M show that the BER is only slightly degraded by an increase in the value of M, or rather an increase in PAPR reduction. All the BER performances for different values of M are near to the ideal case. Therefore, the proposed method can proficiently reduce PAPR to improve the power efficiency of a HPA, without degrading the BER of a MIMO-OFDM system.
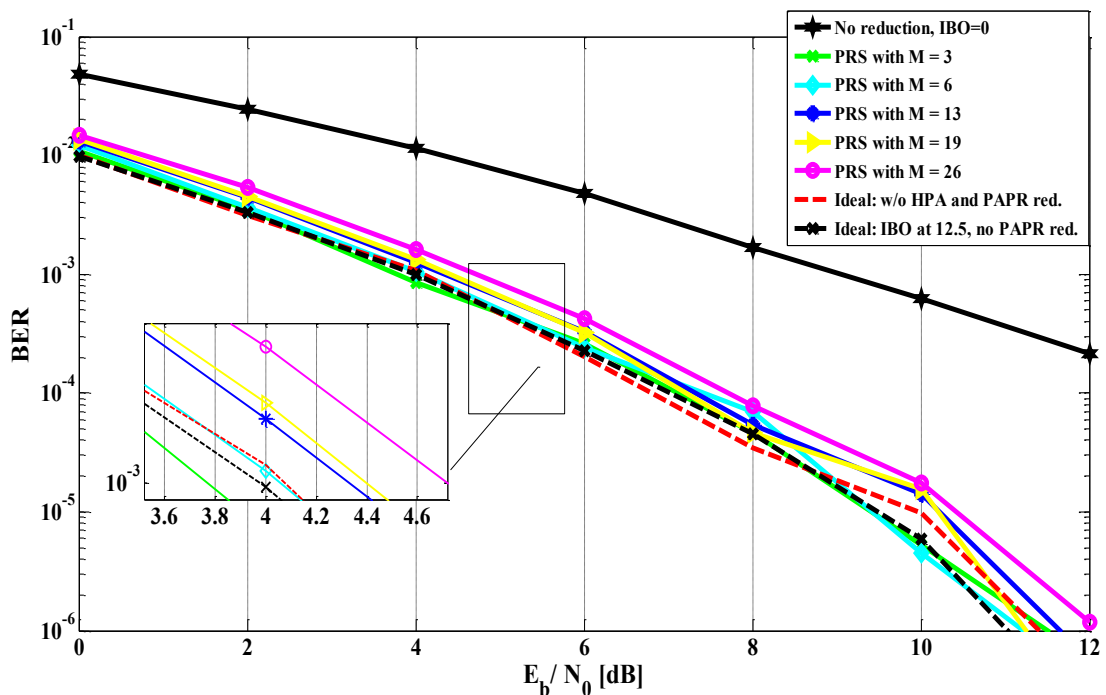
Figure 6.21. BER performance for PRS signal with different number of nonzero samples

145

For the purpose of comparison, the BER performance of the proposed LCASM method together with the performances of other four methods, which are ACT, CSC, SCS-SLM and S-TR methods, are depicted in Figure 6.22. It can be observed from the figure that the BER performances from equation (3.66) and by simulation are the same for a SISO-OFDM system.

The BER performance in Figure 6.22 labelled "MIMO-OFDM ideal" represents the case where a HPA is not used and there is no PAPR reduction, so the BER is not degraded. The performance with the legend "MIMO-OFDM no PAPR red." is for the case where a HPA without IBO is used and PAPR is not reduced, thus leading to a severe BER degradation. The rest of BER performances in the figure are for when the LCASM and the other four PAPR reduction methods listed above were integrated into the system.

One important and noticeable observation in the figure is that the simulated MIMO-OFDM system, with a diversity gain of 4, has a far much better BER performance, even for the worst degradation scenario, than a SISO-OFDM system that has a diversity gain of 1. This concurs with the theoretical postulation in equation (3.72) that the BER probability of a system reduces exponentially with diversity gain.

From the figure, it can also be observed that the LCASM and S-TR methods have similar BER performances near to the ideal performance, while the other three methods exhibit poor performances away. In overall, the proposed LCASM method outperforms the rest. This is clearly demonstrated by Table 6.11, where the LCASM demand for SNR per bit is lower than for the S-TR, SCS-SLM, CSC and ACT methods by 0.29, 2.68, 3.05 and 3.64 dB, respectively.
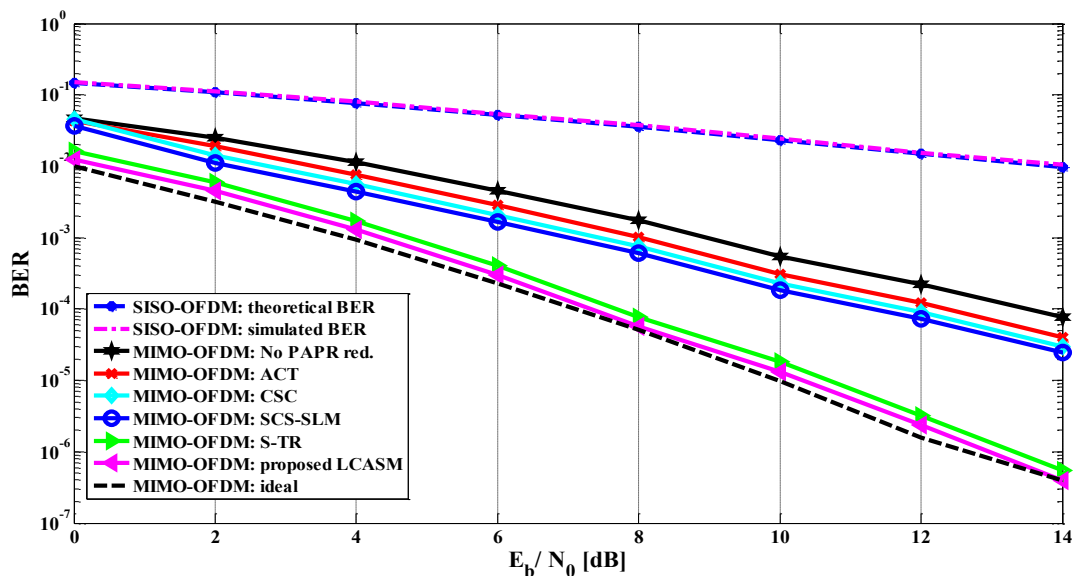
Figure 6.22. BER performance by different methods in MIMO-OFDM system for $N = 128$

Table 6.11. Required E_b/N_o for different methods at BER = 10^{-3} for MIMO-OFDM system

PAPR reduction Method	Proposed LCASM	S-TR	SCS-SLM	CSC	ACT
Required E_b/N_o [dB]	4.36	4.75	7.04	7.41	8.00

CHAPTER 7

CONCLUSION AND FUTURE WORK

This chapter summarizes the findings and the achievements of the research carried out in this thesis together with recommendations that are considered appropriate for future work on PAPR reduction in OFDM systems.

7.1 Conclusions

This thesis has addressed the problem of high PAPR in OFDM signals that causes signal distortions, due to nonlinear amplification, and an increase in the cost of transmission over a communication system employing OFDM technique. In particular, five methods for reducing high PAPR in OFDM systems have been proposed in this thesis.

The proposed methods can achieve high PAPR reductions at the cost of a small loss in data rate caused by a reservation of transmission resources, previously meant for user data, to carry peak-reduction coefficients or time-domain samples. The transmission resources for PAPR reduction are reserved either by setting aside some subcarriers to carry peak-reduction coefficients or by extending the transmit signal by a few of peak-reduction signal samples. In addition, all the five proposed methods were found not to degrade the BER of a system when they were applied to reduce PAPR.

The thesis starts by proposing an optimal method, referred to as SOCP-TR, which employs second-order cone programming to find the required peak-reduction coefficients for a given OFDM signal using the conventional tone-reservation approach of reducing PAPR. Although this method may not be very attractive in real-time OFDM systems, its purpose was firstly to establish whether subcarrier modulation and the structure of OFDM signal, real or complex, affect PAPR reduction. By the end of this investigation, it was found that neither the subcarrier modulation nor the signal structure affects the amount of PAPR reduction. This important discovery allowed for using, without any reservations, of any subcarrier modulation and complex-valued OFDM signals in the rest

of the research work.

The second purpose of developing the optimal SOCP-TR scheme was to find out the standard PAPR reductions that can be achieved by tone-reservation methods. The findings from this investigation helped to design, in this thesis, practical suboptimal methods that are simple, fast and reliable but which have mathematical tractability and low computational complexity and can achieve similar or better PAPR reduction performance than optimal methods.

This thesis approaches the key issue of reducing computational complexity and the time to solution of an optimal method, such as the SOCP-TR, through devising a simple formulation of the desired peak-reduction signal for an OFDM signal. The proposed methods then attempt to generate a replica of the desired peak-reduction signal. There are four such methods proposed in this thesis, namely LP-TR, IRLS-TR, LCSA and LCASM. The first two, LP-TR and IRLS-TR methods, are founded on the tone reservation concept that operates in the frequency-domain, while the remaining two are purely based on signal addition operations in the time-domain.

The LP-TR method utilizes linear programing, while the IRLS-TR employs iterative re-weighted least-squares algorithm in estimating the desired peak-reduction signal. Thus, the LP-TR method is an optimal scheme, and consequently has high computational complexity and slow convergence rate. Its peak-reduction results were used for benchmarking the performance of suboptimal methods. The IRLS-TR method is a suboptimal scheme, which can be viewed as an improvement of LP-TR algorithm to satisfy the requirements of practical realization and to achieves better and higher PAPR reductions than both SOCP-TR and LP-TR methods and others proposed in literature.

For example, in an OFDM system with 64 subcarriers out of which 4 are reserved for peak-reduction coefficients, following the tone-reservation concept, the IRLS-TR method reduces PAPR by 5.07 dB, which is 1 to 3 dB above the reductions by the proposed SOCP-TR and LP-TR methods and other related methods proposed in literature, namely LSA-TR, CF-TR, SSCR-TR, IVO-TR and ELM-TR. The reported PAPR reduction for the proposed IRLS-TR scheme is at the expense of a 0.27 dB increase in average power, which is 0 to 0.57 dB below the increases by the other methods.

The LCSA method is also a suboptimal scheme but it reserves transmission resources for reducing PAPR in terms of the time to transmit nonzero samples of the desired peak-reduction signal. The actual peak-reduction signal is the same as the desired

peak-reduction signal. Therefore, the method converges in a single step and has a simple implementation of very low computational complexity.

Additionally, the desired PAPR reduction can be easily controlled in the LCSA method by fixing the number of nonzero samples in the peak-reduction signal. In other words, given the maximum allowed PAPR for an OFDM system, one only needs to set the number of nonzero samples to fix the clipping threshold and to attain a PAPR reduction up to a required level.

The LCSA method reduces PAPR without significantly affecting the transmit power of an OFDM signal. The method has a high PAPR reduction performance comparable to the proposed IRLS-TR method for OFDM systems with many subcarriers. However, for OFDM systems with a small number of subcarriers, the LCSA method achieves a PAPR reduction slightly less by 1 dB to the IRLS-TR method.

For example, with about 5% of transmission resources reserved for PAPR reduction in an OFDM system with 256 total subcarriers, the LCSA method reduces PAPR by 5.69 dB with a negligible power reduction of 0.10 dB, while the IRLS-TR method attains 5.84 dB with a power increase of 0.36 dB. When the OFDM system has 64 subcarriers and approximately 6% of resources reserved for PAPR reduction, the LCSA method reduces PAPR by 4.13 dB with a negligible power loss of 0.12 dB, while the IRLS-TR method has a reduction of 5.07 dB with a 0.27 dB increase in power. From this, it is obvious that the LCSA method outperforms all the other methods given above.

In addition, the algorithm of the LCSA method was modified to come up with LCASM method for reducing PAPR in MIMO-OFDM systems where multiple signals are transmitted simultaneously. The LCASM method can achieve good PAPR reductions higher than by other methods for MIMO-OFDM systems. For example, in a 2×2 Alamouti space-time encoded MIMO-OFDM system with 128 subcarriers, the LCASM technique reduces PAPR by 7.30 dB. This reduction is more by 0.90, 2.85, 3.35 and 4.40 dB to what is achieved by S-TR, SCS-SLM, CSC and ACT methods, respectively.

7.2 Recommendations for Future Work

The search for practical algorithms to achieve good PAPR reduction in OFDM systems still faces several challenges. Further research can be carried out to address the following identified issues:

i)   An investigation on PAPR reduction in OFDM systems with adaptive modulation

and coding (AMC).

In this thesis, the same type of subcarrier modulation has been assumed as applicable to all subchannels during one symbol duration. However, to maximize system capacity and maintain the desired BER performance, AMC schemes are normally employed during one symbol duration, where different modulations channel coding are used on the subcarriers depending on the quality of subchannels. Therefore, it would be a good idea in future to estimate PAPR and its distribution for an OFDM system utilising mixed modulations based on AMC and develop new algorithms to reduce it or apply the already proposed ones such as those found in this thesis.

ii) Restrictions on the increase in transmit power when using tone-reservation methods.

One drawback of all tone-reservation methods is that the average power of the transmit signal after reduction of PAPR is increased. Although the increase is normally small, it cannot be ignored and should be constrained to about 0 dB while reducing the PAPR. This is to ensure that the transmit power budget remains unaltered; otherwise the gain of the power amplifier must be dynamically adjusted according to the power of the transmit signal and this can increase the cost of the hardware. Therefore, future research may consider integrating a constraint for power conservation in the tone-reservation algorithms proposed in this thesis.

iii) Implementation of the proposed methods in a testbed platform with power amplifiers and other equipment.

Simulations in this thesis were carried out in MATLAB and the results thereof used to analyse the performance of the proposed PAPR reduction techniques. In future, it will be advisable to implement the proposed algorithms in a testbed with real-time signal processing equipment to perform measurements and compare the experimental results with the simulations.

iv) Development of specific PAPR reduction solutions for different future generations of OFDM systems.

The OFDM technique is currently employed in several communication systems, which include broadcasting, mobile communication and wireless internet and has been proposed for use in future generations of these networks. In this thesis, the proposed methods are for reducing PAPR in a general OFDM system. In future

works, the proposed methods can be tuned or modified, if need be, to fit for specific wireless application, e.g. 5G or 6 G mobile networks, taking into account the unique features of each system. The proposed methods could also be applied for PAPR reduction in OFDM with index modulation (OFDM-IM), which combines subcarrier and index modulations to increase spectral efficiency.

v)   Development of PAPR reduction method based on machine learning utilizing the proposed algorithms to generate training targets for neural network.

Because of the low-computational complexity and good PAPR reduction performance, especially by the proposed IRLS-TR, LCSA and LCASM algorithms, it will be of much interest to use these schemes to generate peak-reduction signals for training the neural network in PAPR reduction schemes employing deep learning.

REFERENCES

[1] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, 3G Evolution: HSPA and LTE for Mobile Broadband, London: Elsevier Ltd, 2007.

[2] Y. Louët and J. Palicot, "A classification of methods for efficient power amplification of signals," *Annals of Telecom.*, vol. 63, No. 7-8, 2008, pp. 351-368.

[3] S. Kiambi, E. Mwangi and G. Kamucha, "Effect of OFDM Signal Structure and Subcarrier Modulation on the Reduction of the Signal Peak Power," *IEEE AFRICON 2017 Proceedings*, Cape Town, South Africa, Sept. 18-20, 2017, pp. 262-266, https://doi.org/10.1109/AFRCON.2017.8095492.

[4] S. Kiambi, E. Mwangi and G. Kamucha, "Use of Preset Reserved Tones in Reduction of PAPR in OFDM Systems," *IEEE AFRICON 2019 Proceedings*, Accra, Ghana, Sept. 25-27, 2019, pp. 1-5, https://doi.org/10.1109/AFRICON46755.2019.9133894.

[5] S. Kiambi, E. Mwangi and G. Kamucha, "An Iterative Re-Weighted Least-Squares Tone Reservation Method for PAPR Reduction in OFDM Systems," *WSEAS Transactions on Communications*, E-ISSN: 2224-2864, vol. 18, pp. 153–161, 2019, http://www.wseas.org /multimedia/journals/communications/2019/a425104-1078.php.

[6] S. Kiambi, E. Mwangi and G. Kamucha, "A Low-Complexity Signal Addition Method for PAPR Reduction in OFDM Systems," *International Journal of Computer Application*, ISSN: 2250-1797, Issue 10, vol. 5, pp. 21–34, 2020, doi: https://dx.doi.org/10.26808/rs.ca.i10v5.03.

[7] S. Kiambi, E. Mwangi and G. Kamucha, "PAPR Reduction in MIMO-OFDM Systems Using Low-Complexity Additive Signal Mixing," *Journal of Communications*, JCM vol. 16, no. 11, pp. 468-478, November 2021. http://www.jocm.us/show-261-1699-1.html.

[8] S. Kiambi, E. Mwangi and G. Kamucha, "Reducing PAPR of OFDM Signals Using a Tone Reservation Method Based on ℓ_∞-Norm Minimization," *Journal of Electrical Systems and Information Technology* (JESIT), vol. 9, article no. 12, pp. 1-15, 2022, doi: https://doi.org/10.1186/s43067-022-00055-0.

[9] H. Ochiai and H. Imai, "Performance analysis of deliberately clipped OFDM signals," *IEEE Transactions on Communications*, vol. 50, no. 1, pp. 89–101, January 2002.

[10] R. Zayani, H. Shaiek and D. Roviras, "PAPR-aware massive MIMO-OFDM downlink," *IEEE Access*, vol. 7, pp. 25474-25484, 2019.

[11] Vallavaraj, B. G. Stewart and D. K. Harrison, "An evaluation of modified μ-law companding to reduce the PAPR of OFDM systems," *AEU-International Journal of Electronics and Communications*, vol. 64, no. 9, pp. 844-857, 2010.

[12] H. B. Jeon, J. S. No and D. J. Shin, "A low-complexity SLM scheme using additive mapping sequences for PAPR reduction of OFDM signals," *IEEE Transactions on Broadcasting*, vol. 57, no. 4, pp. 866-875, 2011.

[13] T. Jiang, C. Ni and L. Guan, "A novel phase offset SLM scheme for PAPR reduction in Alamouti MIMO-OFDM systems without side information," *IEEE Signal Processing Letters,* vol. 20, no. 4, pp. 383-386, 2013.

[14] X. Wang, Q. Zhang, R. Gao, W. Xishuo, X. Xin, T. Feng, Q. Tian, W. Yongjun, Z. Li, D. Guo and H. Huan, "Constellation Reshaping Method for PAPR Reduction of SIM-OFDM based on SLM algorithm," *2021 Asia Communications and Photonics Conference (ACP)*, 2021, pp. 1-3.

[15] Lahcen, A. Saida and A. Adel, "Low computation complexity PTS scheme for PAPR reduction of MIMO-OFDM systems," *10th International Conference Interdisciplinarity in Engineering, Procedia Eng.* 181, 2017, pp. 876-883.

[16] H. Xu, Y. Chen, Y. Lu, J. Liu and F. Hu, "Improved Elitist Genetic Algorithm Optimization based PTS to PAPR Reduction for Multi-Carrier Signals," *2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2022, pp. 1746-1750.

[17] N. Telagam, S. Lakshmi and K. Nehru, "PAPR reduction of GFDM system using Parallel concatenation of LDPC codes," *2022 IEEE Fourth International Conference on Advances in Electronics, Computers and Communications (ICAECC)*, 2022, pp. 1-5.

[18] Y. Ni and T. Jiang, "A novel adaptive tone reservation scheme for PAPR reduction in large-scale multiuser MIMO-OFDM systems," *IEEE Wireless Communications Letters*, vol. 5, no. 5, pp. 480-483, 2016.

[19] X. Wang, N. Jin and J. Wei, "A Model-Driven DL Algorithm for PAPR Reduction in OFDM System," in *IEEE Communications Letters*, vol. 25, no. 7, pp. 2270-2274, July 2021.

[20] H. Abdelali, H. Ghennioui, M. El Kamili and M. Firdaoussi, "New PAPR reduction technique combining the Tone Reservation based on PCG algorithm with Clipping," *2019 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2019, pp. 1-6.

[21] B. Bakkas, I. Chana and H. Ben-Azza, "PAPR reduction in MIMO-OFDM based on Polar Codes and Companding," *International Conference on Advanced Communication Technologies and Networking* (*CommNet*), 2019, pp. 1-6.

[22] Beena, S. Pillai and N. Vijayakumar, "Hybrid PTS-clipping scheme for PAPR reduction in MIMO-OFDM systems," *International Journal of Applied Engineering Research*, ISSN 0973-4562, vol. 13, no. 11, pp. 9924-9928, 2018.

[23] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Communication Surveys and Tutorials*, vol. 15, no. 4, pp. 1567-1592, 2013.

[24] X. Li and L. Cimini, "Effect of clipping and filtering on the performance of OFDM,"

IEEE Communications Letters, vol. 2, no. 5, pp. 131-133, May 1998.

[25] V. -N. Tran and T. -H. Dang, "New Clipping-and-Filtering Method for Peak-to-Average Power Ratio Reduction in OFDM," *2021 International Conference Engineering and Telecommunication (En & T)*, 2021, pp. 1-5.

[26] Rajini and G. Sikri, "Reducing peak- to-average by classical clipping over BPSK and QPSK in OFDM system," *Proceedings of International Conference on Recent Advances and Future Trends*, 2012, pp. 26-28.

[27] H. Ochiai and H. Imai, "On the clipping for peak power reduction of OFDM signals," *in Proceedings of IEEE Global Communications Conference* (*GLOBECOM*), San Francisco, USA, 2000, pp. 731–735.

[28] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *IEEE Electronics Letters*, vol. 38, no. 8, pp. 246-247, February 2002.

[29] S. Singh and A. Kumar, "Performance analysis of adaptive clipping technique for reduction of PAPR in Alamouti coded MIMO-OFDM systems," *Procedia Comp. Sci. 93*, 609 – 616, 2016.

[30] G. Chen, R. Ansari and Y. Yao, "Improved Peak Windowing for PAPR Reduction in OFDM," *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, 2009, pp. 1-5.

[31] P. Foomooljareon and W. Fernando, "Input sequence envelope scaling in PAPR reduction of OFDM," *IEEE 5th International Symposium on Wireless Personal Multimedia Communications*, vol. 1, October 2002.

[32] H. Nikookar and K. Lidsheim, "Random phase updating algorithm for OFDM transmission with low PAPR," *IEEE Transactions on Broadcasting*, vol.48, no. 2, pp. 123-128, July 2002.

[33] C. Tan and I. Wassell, "Data bearing peak reduction carriers for OFDM systems," Fourth International Conference on Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia, 2003, pp. 854-858.

[34] X. Wang and T. Tjhung, and C. Ng, "Reduction of peak to average power ratio of OFDM system using a companding technique," *IEEE transactions on broadcasting*, vol. 25, no. 3, pp. 303-308, September 1999.

[35] W. Al-Azzo and B. Ali, "Adaptive square-rooting companding technique for PAPR reduction in OFDM systems," *WASET International Journal of Electronics and Communication Engineering*, vol. 5, no. 3, pp. 283-287, 2011.

[36] T. Wilkinson and A. Jones, "Minimization of the peak-to-mean envelope power ratio of multicarrier transmission schemes by block coding," *IEEE Vehicular Technology Conference*, July 1995, pp. 825-829.

[37] B. Popovic, "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE transactions on Communications*, vol. 39, no. 7, pp. 1031-1033, July

1991.

[38]   A. Abouda, "PAPR reduction of OFDM signal using turbo coding and selective mapping," *Proceedings of the 6th Nordic Signal Processing Symposium NORSIG-2004*, Espoo, Finland, June 2004.

[39]   B. Horvath and B. Botlik, "Optimization of Tone Reservation-Based PAPR Reduction for OFDM Systems", *Radio Engineering*, vol. 26, no. 3, pp. 791-797, September 2017.

[40]   F. Hu, Y. Lu, L. Jin, J. Liu, Z. Xia, G. Zhang and J. Xiao, "Hybrid Energy Efficiency Friendly Frequency Domain TR Algorithm Based on PSO Algorithm Evaluated by Novel Maximizing HPA Efficiency Evaluation Criteria," *Energies*, vol. 15, issue 3, pp. 1-19, 2022.

[41]   Z. Liu, W. Liu, L. Hu, L. Zhang and F. Yang, "A Low Complexity Improved Tone Reservation Method Based on ADMM for OFDM Systems' PAPR Reduction," *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2021, pp. 1-5.

[42]   S. Sumitra and G. Ramprabhu, "Reduction of Peak-to-Average Power Ratio in OFDM System using SCR based Tone Reservation Technique," *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2020.

[43]   J. Tellado and J. M. Cioffi, "Efficient algorithms for reducing PAR in multicarrier systems," *Proceedings of 1998 IEEE International Symposium on Information Theory (Cat. No.98CH36252)*, 1998, p. 191.

[44]   S. Zabre, J. Palicot, Y. Louet and C. Lereau, "SOCP approach for ofdm peak-to-average power ratio reduction in the signal adding context," *Proceedings of IEEE International Symposium on Signal Processing and Information Technology*, 2006, pp. 834–839.

[45]   D. Guel J. Palicot and Y. Louet, "Tone reservation technique based on geometric method for orthogonal frequency division multiplexing peak-to-average power ratio reduction," *IET Communications*, vol. 4, issue 17, pp. 2065–2073, 2010.

[46]   J. Tellado and J. Cioffi, "PAR reduction in multicarrier transmission systems," *ANSI Document, T1E1.4 Technical Subcommittee*, no. 97-367, 98-083 and *ITU*, Q4/15, no. D-150 (WP 1/15).

[47]   A. Gatherer and M. Polley, "Controlling clipping probability in DMT transmission," *in Proceedings of 31st Asilomar Conference on Signals, Systems and Computers,* vol. 1, November 1997, pp. 578–584.

[48]   H. Li, T. Jiang and Y. Zhou, "An improved tone reservation scheme with fast convergence for PAPR reduction in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 57, no. 4, pp. 902-906, December 2011.

[49]   T. Jiang, C. Ni, C. Xu, and Q. Qi, "Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems," *IEEE Communication Letters*, vol. 18, No. 5, pp. 805–808, May 2014.

[50]   J. Wang, X. Lv and W. Wu, "SCR-based tone reservation schemes with fast

convergence for PAPR reduction in OFDM system," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 624–627, April 2019.

[51] L. Xin and W. Yi, "A New Weighted Tone Reservation Method for PAPR Reduction in OFDM Systems," *Journal of Communications*, vol. 9, no. 12, pp. 980-986, 2014.

[52] H. Li, J. Wei and N. Jin, "Low-complexity tone reservation scheme using pre-generated peak-cancelling signals," *IEEE Communication Letters*, vol. 23, issue: 9, pp. 1586-1589, September 2019.

[53] Z. Li, N. Jin, X. Wang and J. Wei, "Extreme Learning Machine-Based Tone Reservation Scheme for OFDM Systems," in *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 30-33, Jan. 2021.

[54] M. Habibi and M. Naeiny, "Selective tone reservation method for PAPR reduction in SFBC-OFDM systems," *Journal of Communication Engineering,* vol. 3, no. 2, pp. 109-122, 2014.

[55] S. Han, J. Ciofi and J. Lee, "Tone injection with hexagonal constellation for peak-to-average power ratio reduction in OFDM," *IEEE Communication Letters*, vol. 10, no. 9, pp. 646–648, September 2006.

[56] H. Jeon, J. No and D. Shin, "A Low-complexity SLM scheme using additive mapping sequences for PAPR reduction of OFDM signals," *IEEE Transactions on Broadcasting*, vol. 57, no. 4, pp. 866-875, June 2011.

[57] S. Carcangiu, A. Fanni and A. Montisci, "A closed form selected mapping algorithm for PAPR reduction in OFDM multicarrier transmission," *Energies*, vol. 15, issue 15, pp. 1-23, 2022.

[58] D. Lim, J. No, C. Lim and H. Chung, "A new SLM OFDM scheme with low complexity for PAPR reduction," *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 93-96, February 2005.

[59] H. Breiling, S. Muller-Weinfurtner and J. Huber, "SLM peak power reduction without explicit side information," *IEEE Communication Letters*, vol. 5, no. 6, pp. 239–241, June 2001.

[60] S. Goff, S. Al-Samahi, B. Khoo, C. Tsimenidis and B. Sharif, "Selected mapping without side information for PAPR reduction in OFDM," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 3320–3325, July 2009.

[61] H. Mäuller and J. Huber, "A novel peak power reduction scheme for OFDM," *Proceedings of the International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC97*, September 1997, Helsinki, Finland, pp. 1090-1094.

[62] F. Hu, Y. Lu, L. Jin, J. Liu, Z. Xia, G. Zhang and J. Xiao, "Hybrid-domain evaluation PTS with adaptive selection methods for PAPR reduction," *Energies*, vol. 15, issue 8, pp. 1-21, 2022.

[63] H. Chen and G. Pottie, "An orthogonal projection-based approach for PAR reduction in OFDM," *IEEE Communications Letters*, vol. 6, no. 5, pp. 169-171, May 2002.

[64] S. Han and J. Lee, "PAPR reduction of OFDM signals using a reduced complexity PTS technique," *IEEE Signal Processing Letters,* vol. 11, no. 11, pp. 887-890, November 2004.

[65] B. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Transactions on Broadcasting*, vol. 49, no. 3, pp. 258-268, September 2003.

[66] Z. Wang, X. Chen, Z. Sun and X. Ning, "On the Performance of Nonlinear Corrective Active Constellation Expansion in OTFS Systems," in *IEEE Communications Letters,* 2022, doi: 10.1109/LCOMM.2022.3165191.

[67] F. Sandoval, G. Poitau and F. Gagnon, "Hybrid peak-to-average power ratio reduction techniques: review and performance Comparison," *IEEE Access*, vol. 5, pp. 27145-27161, 2017.

[68] A. Ghassemi and T. A. Gulliver, "PAPR reduction of OFDM using PTS and error-correcting code sub-blocking ," *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 980–989, March 2010.

[69] A. Abouda, "PAPR reduction of OFDM signal using turbo coding and selective mapping," *in Proceedings of the 6th Nordic Signal Processing Symposium*, June 2004, pp. 248-251.

[70] H. Chen and H. Liang, "A modified selective mapping with PAPR reduction and error correction in OFDM systems," *in 2007 IEEE Wireless Communications and Networking Conference*, March 2007, pp.1329–1333.

[71] B. Tang, K. Qin and H. Mei, "A hybrid approach to reduce the PAPR of OFDM signals using clipping and companding," *IEEE Access*, vol. 8, pp. 18984-18994, 2020.

[72] Abdullah and M. Hidayat, "SCS-SLM PAPR reduction technique in STBC MIMO-OFDM systems," *in Proc.7th IEEE ICCSCE, Penang, Malaysia,* 2017, pp. 104-109.

[73] F. Sandoval, G. Poitau and F. Gagnon, "On optimizing the PAPR of OFDM signals with coding, companding and MIMO," *IEEE Access*, vol. 7, pp. 24132-24139, 2019.

[74] M. Yarlequé, *RF Power Amplifiers for Wireless Communications*, PhD-thesis, Catholic University, Leuven, Belgium, 2008.

[75] M. Kazimierczuk, *RF Power Amplifiers*, John Wiley & Sons Ltd, West Sussex, United Kingdom, 2008.

[76] E. Dalakta, A. Dweik and C. Tsimenidis, "Efficient BER reduction technique for nonlinear OFDM transmission using distortion prediction," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 5, pp. 230-36, June 2012.

[77] R. Marsalek, P. Jardin and G. Baudoin, "From post-distortion to pre-distortion for power amplifiers linearization," *IEEE Communications Letters*, vol. 7, pp. 308-310, August 2003.

[78] S. Cripps, *RF Power Amplifiers for Wireless Communications*, 2nd Edition, Artech

House, Massachusetts, USA, 2006.

[79] W. Doherty, "A new high-efficiency power amplifier for modulated waves," *The Bell System Technical Journal*, vol. 15, no. 3, pp. 469-475, July 1936.

[80] L. Kahn, "Single-sideband transmission by envelope elimination and restoration," *Proceedings of IRE*, vol. 40, pp. 803–806, July 1952.

[81] A. Diet, C. Berland, M. Villegas and G. Baudoin, "EER architecture specifications for OFDM transmitter using a class E amplifier," *IEEE Microwave and Wireless Component Letters*, vol. 14, no. 8, pp. 389-391, 2004.

[82] T. Jiang and Y Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM signals," *IEEE Transactions on Broadcasting*, vol.54, no.2, pp. 257-268, June.2008.

[83] S. Han and J. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56–65, 2005.

[84] R. Prasad, *OFDM for Wireless Communications Systems*, Artech House, London, UK, 2004.

[85] R. Bahai, R. Saltzberg and M. Ergen, *Multi-carrier digital communications: theory and applications of OFDM*, Springer-Verlag Inc, New York. UK, 2004.

[86] ETSI EN 302 755 V.1.3.1. Digital Video Broadcasting (DVB); Frame Structure Channel Coding and Modulation for a Second Generation Digital Terrestrial Television Broadcasting System (DVB-T2). Available at: https://www.dvb.org/standards.

[87] https://www.celplan.com/technologies/2g-to-6g.

[88] NTT DOCOMO Inc., *White paper 5G Evolution and 6G*, January 2020.

[89] F. Juwono and R. Reine, "Future OFDM-based Communication Systems Towards 6G and Beyond: Machine Learning Approaches," *Green Intelligent Systems and Applications*, vol. 1(1), pp. 19-25, 2021.

[90] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, first edition, Cambridge University Press, 2004.

[91] H. Rohling, *OFDM: Concepts for Future Communication Systems*, Signals and Communication Technology, Springer-Verlag, Berlin Heidelberg, Germany, 2011.

[92] M. Viswanathan, *Simulation of Digital Communication systems using Matlab*, second Edition, Amazon Kindle, 2013.

[93] Y. Cho, J. Kim, W. Yang and C. Kang, *MIMO-OFDM Wireless Communications with MATLAB,* John Wiley & Sons (Asia) Pte Ltd, Singapore, 2010.

[94] S. Rao, *Engineering optimization Theory and Practice*, 3rd edition, New Age International, 2013.

[95] M. Sharif, M. Gharavi-Alkhansari and B. Khalaj, "New Results on the Peak Power of OFDM Signals Based on Oversampling," *in Proceedings of IEEE ICC*, vol. 2, 2002, pp. 866–871.

[96] A. Saleh, "Frequency-independent and frequency-dependent nonlinear models for TWT amplifiers," *IEEE Transactions on Communications*, vol. 29, pp. 1715–1720, Nov. 1981.

[97] A. Ghorbani and M. Sheikhan, "The effect of solid state power amplifiers (SSPAs) nonlinearities on MPSK and M-QAM signal transmission," *in Proceedings of the 6th International Conference on Digital Processing of Signals in Communications*, Sept. 1991, pp. 193–197.

[98] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM signal for a digital sound broadcasting signal," *in proceedings of 2nd European conference on satellite communication*, Liège, Belgium, Oct. 1991, pp. 179-184.

[99] Y. Louët and J. Palicot, "A classification of methods for efficient power amplification of signals," *Annals of Telecommunications*, vol. 63, No. 7-8, 2008, pp. 351-368.

[100] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Second edition, Cambridge University Press, December 2003.

[101] R. Gilmore and L. Besser, *Practical RF Circuit Design for Modern Wireless Systems*, vol. II, Artech House, 2003.

[102] M. Yarlequé, *RF power amplifiers for wireless communications*, PhD thesis, Catholic University, Leuven, Belgium, June 2008.

[103] J. Palicot, Y. Louet and S. Hussain, "Power Amplification Issues Related to Dynamic Spectrum Access in the Cognitive Radio Systems," in Cognitive Radio Systems, London, UK, IntechOpen, 2009, https://www.intechopen.com/chapters/ 8827.

[104] C. Forster et al, Understanding the environmental impact of communication system, Ofcom, UK, April 2009.

[105] S. Haykin and M. Moher, *Introduction to Analog and Digital Communications*, second edition, John Wiley & Sons, Inc, New Jersey, USA, 2007.

[106] J. Andrews, A. Ghosh and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*, Pearson Education, Inc., USA, 2007.

[107] B. Sklar, *Digital communications: fundamentals and applications*, Second edition, Prentice Hall, New Jersey, USA, 2001.

[108] J. Proakis and M. Saleh, *Digital communications*, Fifth edition, McGraw-Hill, New York, USA, 2008.

[109] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal of Selected Areas in Communications*, vol. 16, no. 8, pp. 1451-1458, 1998.

[110] I. Barhumi, G. Leus and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Transaction on Signal Processing*, vol. 51, no. 6, pp. 1615-1624, June 2003.

[111] A. Zelst and T. Schenk, "Implementation of a MIMO OFDM based Wireless LAN system," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 483-494, February 2004.

[112] X. Li, H. Huang, G. Foschini and R. A. Valenzuela, "Effects of iterative detection and decoding on the performance of BLAST," *IEEE Proceedings of Global Telecommunications Conference*, San Francisco, CA, USA, 2000, pp. 1061-1066.

[113] A. Antoniou and W. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer Science + Business Media, LLC, New York, USA, 2007.

[114] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, USA, 2004.

[115] S. Burrus, J. Barreto and I. Selesnick, "Iterative re-weighted least-squares design of FIR filters," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 2926-2936, 1994,

[116] A. Messac, Optimization in Practice with MATLAB for Engineering Students and Professionals, Cambridge University Press, New York, USA, 2015.

[117] B. Farhang-Boroujeny, *Adaptive filters: theory and applications*, Second edition, John Wiley & Sons Ltd, West Sussex, United Kingdom, 2013.

[118] P. Diniz, *Adaptive filtering: algorithms and practical implementatio*n, Third edition, Springer, New York, USA, 2008.

[119] S. Haykin, *Adaptive filter theory*, third edition, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[120] B. Widrow and S. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Inc., NJ, USA, 1985.

[121] T. Bose, *Digital signal and image processing*, John Wiley & Sons Inc., USA, 2004.

[122] *MATLAB Version 8.1.0.604* (R2013a), Math works, February 15, 2013.

[123] J. Tellado, *Multicarrier Modulation with Low PAR: Applications to DSL and Wireless,* first edition, Kluwer Academic Publishers, Norwell, MA (USA), 2002, ISBN: 978-0-7923-7988-1.

[124] M. S. Loboa, L. Vandenberghe, S. Boyd and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, Issues 1–3, pp. 193–228, November 1998.

[125] M. Grant and S. Boyd, *CVX: MATLAB software for disciplined convex programming*, version 2.2, http://cvxr.com/cvx, January 2020.

[126] J. Sturm, "Using SEDUMI 1.02, a MATLAB toolbox for optimization over symmetric

cones," *Optimization Methods and Software*, 11-12: 625–653, 1999.

[127] R. Tütüncü, K. Toh, and M. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, Series B, 95: 189-217, 2003.

[128] MOSEK ApS, "The MOSEK Optimization Tools," *User's manual and reference*, 2002.

[129] S. Boyd, L. Vandenberghe and M. Grant, "Efficient convex optimization for engineering design," *in proceedings of IFAC Symposium on Robust Control Design*, Rio De Janeiro, Brazil, September 1994.

[130] R. Byrd and D. Pyne, "Convergence of the iteratively re-weighted least squares algorithm for robust regression," *Technical report 313*, Dept. of Math. Sci, Johns Hopkins University, Baltimore, Maryland, 1979.

[131] S. Kahng, "Best L_p approximation," *Mathematics of Computation*, vol. 26, no. 118, 1972, pp. 505-508.

[132] R. Fletcher, J. Grant, and M. Hebden, The calculation of linear best L_p approximations, *Computer Journal*, Vol.14, 1971, pp. 27-279.

[133] V. Stonick and S. Alexander, "Globally optimal rational approximation using homotopy continuation methods," *IEEE Transactions on Signal Processing*, vol.40, no.9, pp. 2358-2361, 1992.

[134] J. Nocedal and S. Wright, *Numerical Optimization*, Springer series in operations research, Springer-Verlag, New York, USA, 1999.

[135] J. Song and H. Ochiai, "Performance analysis for OFDM signals with peak cancellation," *IEEE Transactions on Communications.* vol. 64, no. 1, 2016, pp. 261–270.

[136] B. Li, L. Hu, F. Yang, L. Ding and T. Song, "Tone reservation ratio optimization for papr reduction in OFDM systems," *IEEE Wireless Communication and Networking Conference (WCNC)*, 2018, pp. 1-6.

[137] H. Yin, R. Yang, X. Luo, L. Jiang and L. Zhu , *Weighted tone reservation for OFDM PAPR reduction*, U.S. Patent 7796498, 2008.

APPENDICES

The appendices are organized as follows. Appendix A gives the main notations adopted in the thesis. In Appendix B to G, a list of publications accomplished in the course of the research is provided. Lastly, Appendix H presents MATLAB codes for the main simulations carried out during the research period including those for the methods proposed for PAPR reduction in OFDM systems.

Appendix A: Notations

The following notations are used throughout this thesis. A matrix is denoted by an italicised uppercase letter e.g. A, a vector by a bolded lowercase letter e.g. **a,** and a scalar by an italicised lowercase letter e.g. a.

A matrix A consisting of m rows and n columns has $m \times n$ elements, which can be real or complex numbers, is written as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{A.1}$$

and is said to be of size $m \times n$. A is said to be a *square* matrix if $m = n$, otherwise it is a *rectangular* matrix. A rectangular matrix is said to be tall if $m > n$ and wide if $m < n$.

A vector **a** consisting of m elements (real or complex numbers) means it is a column vector, which can be equivalently expressed as an $m \times 1$ matrix as follows:

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = [a_1 \ a_2 \ ... \ a_m]^T \tag{A.2}$$

where T is for the transpose operation.

The set containing $m \times n$ real or complex matrices is denoted as $\mathbb{R}^{m \times n}$ or $\mathbb{C}^{N \times L}$. A matrix A of size $m \times n$ having all real or all complex elements is indicated by $A \in \mathbb{R}^{m \times n}$ or $A \in \mathbb{C}^{N \times L}$, respectively. Similarly, a vector containing real or complex m elements is denoted as $\mathbf{a} \in \mathbb{R}^m$ or $\mathbf{a} \in \mathbb{C}^m$, respectively. In case of a real and complex scalar, the notations $a \in \mathbb{R}$ or $a \in \mathbb{C}$ are used, respectively.

A discrete-time signal consisting of N samples, real or complex, is represented by $x(n)$ or by a column vector $\mathbf{x}$, where

$$\mathbf{x} = [x(0), x(1), \dots, x(N-1)]^T \tag{A.3}$$

A sample in the discrete-time signal is referenced by specifying its index in the parenthesis e.g. the second sample in $x(n)$ is indicated by $x(2)$.

A square matrix A is said to be a $m \times m$ diagonal matrix if all the elements are zero except the diagonal elements a_{ii}, for $i = 1, 2, \dots, m$, and is commonly written as

$$A = \text{diag}\{a_{11}\ a_{22}\ \dots\ a_{mm}\} \tag{A.4}$$

The complex conjugate of a matrix A is denoted as A^*, and the transpose as A^T. If the matrix has complex elements, its conjugate transpose, denoted as A^H, is formed by first conjugating every element followed by the transpose operation or vice versa, i.e.

$$A^H = (A^*)^T = (A^T)^* \tag{A.5}$$

A square matrix A of size $m \times m$ is said to be symmetric if $A^T = A$. If the square matrix has complex elements, it is said to be Hermitian if $A^H = A$. A square matrix whose diagonal elements are unity and off-diagonal elements are zero is referred to as identity matrix.

The inverse (if it exists) of a square matrix is denoted by A^{-1} and has the property

$$A^{-1}A = AA^{-1} = I_m \tag{A.6}$$

where I_m is an $m \times m$ identity matrix. If the inverse of A does not exist, A is said to be *singular* otherwise it is *nonsingular*.

The *trace* of an $m \times m$ square matrix A is denoted by $\text{tr}(A)$ and is the sum of the diagonal elements, i.e.

$$\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{mm} \tag{A.7}$$

If matrix A is of size $m \times n$, its rank is the maximum number of linearly independent columns or rows it has; it does not matter whether one takes rows or columns. A matrix is said to be of *full rank* if its rank is equal to the number of rows or columns; whichever is smaller.

Appendix B: First Publication

Effect of OFDM Signal Structure and Subcarrier Modulation on the Reduction of the Signal Peak Power

Stephen Kiambi, Elijah Mwangi, George Kamucha
School of Engineering, University of Nairobi, Kenya
skiambi@students.uonbi.ac.ke

Abstract—The problem of high ratio of peak-to-average power experienced in orthogonal-frequency-division-multiplexed signals remains a major challenge when it comes to practical realization of transmitters for such signals. This is mainly due to the requirement on high power amplifiers to operate with large input power back-offs and thus at undesirable low power efficiency region in order to avoid clipping of such signals and the subsequent degradation of bit error rate and spectral interference to adjacent channels. In this paper, an efficient peak power reducing technique is first proposed, its performance verified, and then used to investigate the influence of OFDM signal structure and subcarrier modulation on the peak power reduction capability of peak power reduction techniques. The proposed technique involves carefully selecting some subcarrier signals that when added to information-modulated subcarriers yield a combined signal with an acceptable lower ratio of peak-to-average power. This technique is based on the application of second order cone convex programming to the minimization of maximum of norms. Simulation results show that the proposed technique reduces peak-to-average power ratio (PAPR) by 4.5 dB and 7.8 dB while utilizing 5% and 20% of subcarriers for peak power reduction. Real OFDM signals are found to exhibit a PAPR that is higher than that of complex signals by 2.4 dB. In addition, the OFDM signal structure and the type of subcarrier modulation are found to have minimal impact on the peak power reduction capability since the variation in PAPR reductions for the modulation schemes used in the research is within ±0.75 dB, and the difference in PAPR reduction for both real and complex signals is on average within ±0.22 dB.

Keywords—Complementary cumulative distribution function (CCDF); high power amplifier (HPA); orthogonal frequency division multiplexing (OFDM); peak-to-average power ratio (PAPR); bit error rate (BER)

I. INTRODUCTION

OFDM has been and still remains the multiplexing technique of choice for high data rate transmission in radio systems whose performance can get significantly limited by multipath interference. Such systems include Digital Audio Broadcasting (DAB), Digital Video Broadcasting (DVB), IEEE 802.11a Wireless Local Area Networks (WLAN), IEEE 802.16a Wireless Metropolitan Area Networks (WMAN), High Performance Radio Local Area Network Type-2 (HIPERLAN-2), 4G and 5G mobile networks. Despite its competitive advantages, OFDM generated signals are characterized by very high peak-to-average power ratio (PAPR) [1]. This makes the signals to be susceptible to nonlinear effects of the analogue components in the transmitter; of critical influence being the high power amplifier (HPA).

In order to achieve high power efficiency, a HPA is designed to operate near its saturation zone. However, in this region, the output is not proportional to the amplifier input. This nonlinear amplification gives rise to in-band, and out-of-band radiations. The in-band radiations degrade the bit error rate of the system,

while the out-of-band radiations interfere with the adjacent channels. One simple solution to avoid the operation in the saturation region is to provide the amplifier with a large enough input power back-off (IBO) dependent on the peak power of the input signal. However this large IBO lowers the power efficiency of the amplifier, and thus raising the total cost of the system [2]. For these reasons, a better solution is to reduce the high ratio of peak-to-average power especially so for practical OFDM systems employing a large number of subcarriers [3] where the ratio can be unacceptably high.

Many methods aimed at reducing PAPR have been proposed in the literature but each method has its own advantages and disadvantages. This then necessitates different applications to choose different methods that best suit them when the benefits outweigh the drawbacks [4]. In this regard it is then important to consider the OFDM signal structure as a multicarrier transmission technique and its expected PAPR and possible reduction capability if such PAPR is intolerably high. Considerations of PAPR level for real, and complex signals, and possible reduction levels while taking into account different subcarrier modulations will be of the main interest in this paper.

The rest of the paper is organized as follows. Section II gives the statistical description of PAPR. In section III, an outline of the general concept of tone reservation techniques for PAPR reduction is presented. Section IV presents the proposed technique. Simulation and results analysis are presented in section V, while section VI gives conclusion and suggestions for future work.

II. DISTRIBUTION OF PAPR

A statistical description of PAPR can be directly obtained from the structure of the OFDM signal. The signal is a sum of N overlapping but mutually orthogonal subcarrier signals and its baseband representation, over one symbol duration, is given by the equation,

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{j2\pi f_k t}, \qquad (1)$$

where $x_k(t)$ is the kth subcarrier modulated signal with frequency $f_k = k\Delta f$ and X_k as the modulating symbol. The modulating symbols come from set of symbols in the constellation of any of the following modulation schemes: binary phase-shift keying (BPSK), quadrature phase-shift keying (QPSK) or M-ary quadrature amplitude modulation (M-QAM) [5]. Mutual orthogonality among the subcarriers is achieved by setting the frequency spacing between any two adjacent subcarriers to $\Delta f = 1/T$, with T being the symbol duration of interest. Equation (1) resembles that of the standard inverse Fourier transform (IFFT) and thus any OFDM-based system can

be easily implemented using the well-known fast Fourier transform (FFT) algorithms.

The major concern here is to find the probability that the maximum instantaneous power of $x(t)$ is out of the linear range of HPA. PAPR will be used to measure temporal power fluctuations of the signal and by definition it is the ratio of maximum instantaneous power to average signal power i.e.

$$\text{PAPR}\{x(t)\} = \frac{\max\{|x(t)|^2\}}{\text{E}\{|x(t)|^2\}}, \qquad (2)$$

where $\text{E}\{\cdot\}$ denotes the expectation operation.

The statistical distribution of amplitudes of the time-domain signal $x(t)$ is found by considering that the subcarrier signals $x_k(t)$ are statistically independent and that N approaches infinity. By the central limit theorem, both the real and imaginary parts of $x(t)$ will have Gaussian distributions and accordingly, the amplitudes of $x(t)$ will be Rayleigh distributed [6]. This in turn means that $x(t)$ could have some high amplitude values that are well above the average of the signal amplitudes.

Equation (2) is for the continuous-time PAPR. However, in practice, PAPR is calculated from the samples of the signal $x(t)$ since OFDM signal processing is done in the discrete-time domain. Assuming that the signal is normalised with its root-mean-square value, its complex samples can be represented by the set $\{x_n = x(nT_s),\ 0 \le n \le N-1\}$. Let X_n be a random variable representing the magnitudes of the complex samples. Then X_n is a Rayleigh random variable with the probability density function,

$$f_{X_n}(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} = 2x e^{-x^2}, \qquad (3)$$

since $2\sigma^2 = E\{X_n^2\} = 1$ due to the normalization. The maximum amplitude X_{max} is then the crest factor and is equal to the square root of the discrete-time PAPR defined by the formula

$$\text{PAPR}\{x_n\} = \frac{\max_{0 \le n < N}\{|x_n|^2\}}{E\{|x_n|^2\}}, \qquad (4)$$

and has cumulative distribution function

$$\Pr(X_{max} \le x)$$
$$= \Pr(X_0 \le x)\Pr(X_1 \le x)\dots\Pr(X_{N-1} \le x) \qquad (5)$$
$$= \left(1 - e^{-x^2}\right)^N.$$

where $\Pr(X_n \le x) = \int_0^x f_{X_n}(u)\, du = \int_0^{x^2} e^{-y} dy = 1 - e^{-x^2}$.

The complementary cumulative distribution function of X_{max} is then the probability $\Pr(X_{max} > x)$, or equivalently $\Pr(\text{PAPR}\{x_n\} > x^2)$ and is therefore given by

$$\Pr\left(\text{PAPR}\{x_n\} > \gamma\right) = 1 - \left(1 - e^{-\gamma}\right)^N, \qquad (6)$$

where $\gamma = x^2$ is power of the chosen threshold amplitude.

Equation (4) will give a PAPR value that approaches the one given by (2) if the sampling rate is chosen sufficiently higher than the Nyquist rate in order to avoid missing the peak of the

continuous-time signal. Since the Nyquist rate is equal to the occupied bandwidth, which in turn is the product of the number of occupied subcarriers and the subcarrier spacing, the oversampled signal may be described by the equation,

$$x_n = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} X_k e^{\frac{j2\pi kn}{\alpha N}} = \frac{1}{\sqrt{\alpha N}}\sum_{k=0}^{\alpha N-1} X_k' e^{j2\pi kn/\alpha N}, \qquad (7)$$

where

$$X_k' = \begin{cases} X_k, & 0 \le k < N \\ 0, & N \le k < \alpha N \end{cases}, \qquad (8)$$

and α is an oversampling factor greater than one such that the product αN is a power of two equal to the FFT size. Values of $\alpha \ge 4$ have been found to be sufficient enough [7] for the discrete-time PAPR in (4) to approximate that in (2) and therefore equation (4) will be used throughout in this work.

III. TONE RESERVATION TECHNIQUES

PAPR reduction schemes that are classified as tone reservation techniques employ the same basic idea: reserve and utilize L subcarriers, out of N subcarriers, to carry carefully selected signals that when added to the $(N-L)$ data-bearing subcarrier signals, through an IFFT operation, yield a combined signal with a lower PAPR than the original signal comprising of only $(N-L)$ data-bearing signals [8]. The PAPR reduction operation is implemented in the transmitter section and restricts the modulating symbols and peak-reducing signals to two disjoint frequency subspaces. The operation therefore, does not distort the modulating symbols. At the receiver, the modulating symbols are extracted from the received signal by choosing the set of $(N-L)$ subcarriers at the output of the FFT block.

The general tone reservation model is illustrated in Fig. 1. In the figure, X represents the frequency domain vector of modulating symbols and has non-zero values in all subcarriers except in the reserved L subcarriers. The frequency domain vector C represents the set of peak-reducing signals and has non-zero values in all subcarriers except in the $(N-L)$ data-bearing subcarriers [9]. The time domain signals, x and c are combined to produce a time domain signal $x + c$ with a PAPR less than that of signal x i.e. $\text{PAPR}\{x+c\} < \text{PAPR}\{x\}$. The IFFT signal processing operation can therefore be described by the equations,

$$x + c = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}(X_k + C_k)e^{j2\pi kn/N}$$
$$\qquad\qquad\qquad (9)$$
$$= x + \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} C_k e^{j2\pi kn/N},$$

or in matrix notation,

$$x + c = x + QC, \qquad (10)$$

where $x \in \mathbb{C}^N$, $c \in \mathbb{C}^N$, $C \in \mathbb{C}^N$, and $Q \in \mathbb{C}^{N \times N}$ is the IFFT matrix with elements $\frac{1}{\sqrt{N}}e^{j2\pi kn/N}$. By taking the advantage of $(N-L)$ zeros in the matrix of C, the computational complexity of (10) can be further reduced by considering that,

$$c = \hat{Q}\hat{C}, \qquad (11)$$

263

166
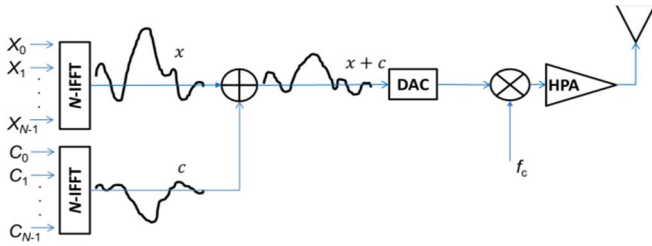
Figure 1: Concept of tone reservation technique

where $\hat{C} \in \mathbb{C}^L$ is a vector containing nonzero values of C, and $\hat{Q} \in \mathbb{C}^{N\times L}$ is a submatrix of Q formed by choosing L columns corresponding to the reserved subcarriers. The problem of finding elements of c, which will minimize the PAPR of x then reduces to that of searching for elements of $\hat{C}$, and can be expressed as,

$$\min_{\hat{C}} ||x + \hat{Q}\hat{C}||_\infty , \qquad (12)$$

where $\hat{C} \in \mathbb{C}^L$ is the optimization variable, $x \in \mathbb{C}^N$ and $\hat{Q} \in \mathbb{C}^{N\times L}$ are the problem parameters, and $||.||_\infty$ denotes the infinity norm [10].

The reduction in PAPR will obviously increase the average transmit power. The search for $\hat{C}$ should therefore ensure that the increase in average power is minimal so as to ensure compatibility with the power amplifier specifications and to avoid transforming the signal into a deterministic one, with PAPR equal to one, and thus not usable for carrying user data [10]. In addition, since the L reserved subcarriers do not carry any user information, the overall user data rate can be reduced, and thus the size of L should be such that there is no significant loss in data rate.

IV. PROPOSED TECHNIQUE

The optimization problem in (12) is that of minimizing maximum-of-norms and can be formulated as a convex optimization problem. The proposed method formulates the problem as a second order cone program (SOCP) which has been found easier to solve than the semi-definite program or the quadratically-constrained quadratic program [11, 13]. Second order cone program falls under the class of convex optimization, and minimizes a linear function over the intersection of an affine set and the product of second-order cones. Considering (12) and letting

$$s = x + \hat{Q}\hat{C} = [s_1, s_2, s_3, \dots, s_N]^T, \qquad (13)$$

yields,

$$||s||_\infty = \max\{|s_i| : i = 1, \dots, N\}. \qquad (14)$$

The ith complex number s_i, is given by $s_i = x_i + q_i^T \hat{C}$, where $q_i \in \mathbb{C}^L$ is the transpose of the ith row vector of $\hat{Q}$. Minimization of $||s||_\infty$ will have the epigraph formulation [11]:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & |x_i + q_i^T \hat{C}| \le t, \ i = 1, \dots, N \end{aligned} \qquad (15)$$

where $\hat{C} \in \mathbb{C}^L$ and $t \in \mathbb{R}$ are the optimization variables, and $q_i \in \mathbb{C}^L$ are the problem parameters. Since the complex number s_i has a magnitude $|s_i| = (s_i s_i^*)^{1/2}$, then each of the ith inequality

constraint in (15) is a second order cone of dimension three [12, 13] represented by

$$\mathcal{C}_3 = \begin{bmatrix} x_i + q_i^T \hat{C} \\ t \end{bmatrix}, \qquad i = 1, \dots, N. \qquad (16)$$

The formulation in (15) can be rewritten in the standard SOCP form by letting $\hat{C} = y + jz$ as:

$$\text{minimize} \quad [0_{2L} \ ; \ 1]^T [y; z; t]$$

$$\text{subject to} \quad \left\| \begin{bmatrix} Re\{q_i^T\} & -Im\{q_i^T\} & 0 \\ Im\{q_i^T\} & Re\{q_i^T\} & 0 \end{bmatrix} \begin{bmatrix} y \\ z \\ t \end{bmatrix} + \begin{bmatrix} Re\{x_i\} \\ Im\{x_i\} \end{bmatrix} \right\|_2 \qquad (17)$$

$$\le [0_{2L}; 1]^T [y; z; t], \ i = 1, \dots, N$$

where $t \in \mathbb{R}$, $y = Re\{\hat{C}\} \in \mathbb{R}^L$ and $z = Im\{\hat{C}\} \in \mathbb{R}^L$ are the optimization variables, $||.||_2$ denotes the Euclidean norm and 0_{2L} is a column vector of 2L zeros.

The optimization problem in (17) is still convex and can be solved exactly for the $(2L+1)$ variables for both real and complex OFDM signals. The number of peak-reducing subcarriers L can be set based on the OFDM system with the aim of utilizing free subcarriers that by system design are not used to carry any user information or system-related information such as pilot signals, and thereby avoid lowering the data rate of the whole system.

V. RESULTS AND DISCUSSION

The performance of the PAPR reduction technique proposed in section IV and the effect of signal structure and subcarrier modulation on peak power reduction were evaluated through simulations in Matlab for both real-valued and complex-valued time domain signals for OFDM systems with 256 subcarriers and modulated by BPSK, QPSK, 16QAM, 64QAM, 256QAM, 1024QAM, and 2048QAM. All the simulations were carried out for 10^4 OFDM blocks. The simulation parameters are listed in Table 1.

Table 1: OFDM Parameters

FFT size, N	256
Modulation	BPSK, QPSK, M-ary QAM
Number of subcarriers	256
Number of symbols	10,000
Structure of time domain signal	real, complex
Number of peak-reducing subcarriers, L	12, 50
Peak-reducing subcarriers' positions	random

Different values of L were used during the simulation for each modulation scheme. The results are as illustrated in Fig. 2. In the figure's legend, the labels 0% PRS, 5% PRS and 20% PRS indicate the percentage of peak-reducing subcarriers out of the total subcarriers. The peak reduction capability, for all modulation schemes, was evaluated at the same probability, of 10^{-3}, of PAPR being greater than a given threshold. Therefore, a lower PAPR threshold indicates a lower PAPR.

The ability of the proposed technique to reduce PAPR is well demonstrated in Fig. 2. From the figure, it can be observed that the higher the number of peak-reducing signals the lower is the PAPR. Also for both real and complex signals, with the use of only 5% of subcarriers to carry peak-reducing signals, the average

decrease in PAPR is approximately 4.5 dB, while with 20% of subcarriers reductions of approximately 7.8 dB are achieved across the seven modulation schemes.

Another notable observation from Fig. 2 is that real-valued time domain OFDM signals have higher PAPR than complex-valued ones by approximately 2.4 dB and this could be due to the conjugate symmetric properties of the former, thus giving limited magnitude and phase cancellations in the combined signal.
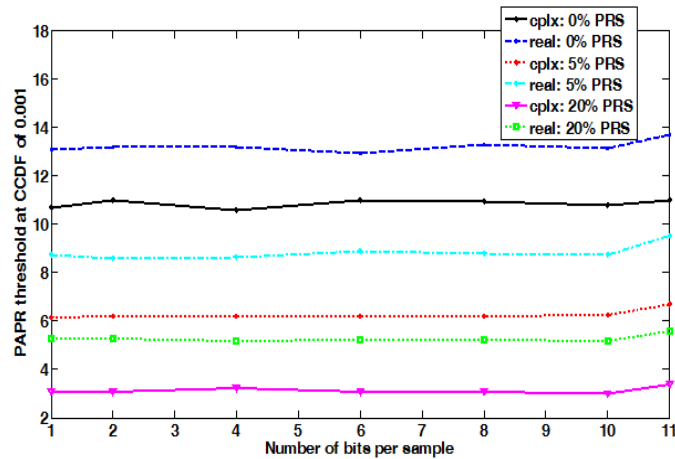
Figure 2: PAPR reduction versus modulation bits

Further investigation of the simulation results reveals that PAPR reduction capability does not strictly depend on the modulation scheme or on the structure (real or complex) of the OFDM signal. This observation is illustrated in Fig. 3 in which for a given number of peak-reducing subcarriers, the variation of CCDF reduction across the seven modulation schemes is within ± 0.75 dB while the difference in the CCDF reduction for real and complex signals is on average within ± 0.22 dB.
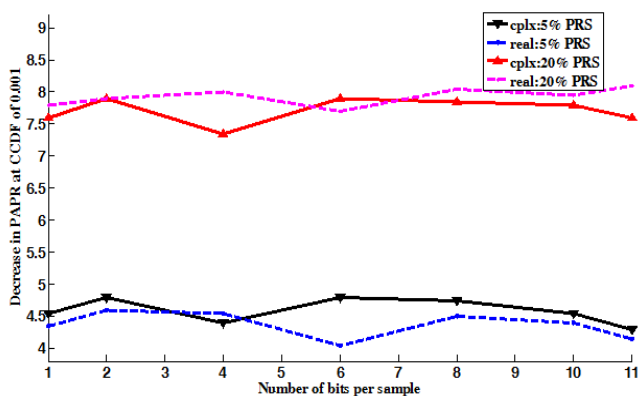
Figure 3: PAPR reduction versus modulation bits

The proposed method achieves better PAPR reductions than reported in the literature for OFDM system with the same parameters. In [14] the maximum PAPR reduction is derived as 6.02 dB and the average reduction as 1.47 dB, but these values can be slightly changed by the number of peak-reducing signals employed in the system. Figure 4 illustrates PAPR reductions for the case of QPSK modulation on 256 subcarriers. At CCDF of 0.01, the proposed method achieves a PAPR reduction of 4.2 dB with 12 peak-reducing signals, while [14] and [15] reports reductions of 1.4 dB and 3.8 dB respectively.
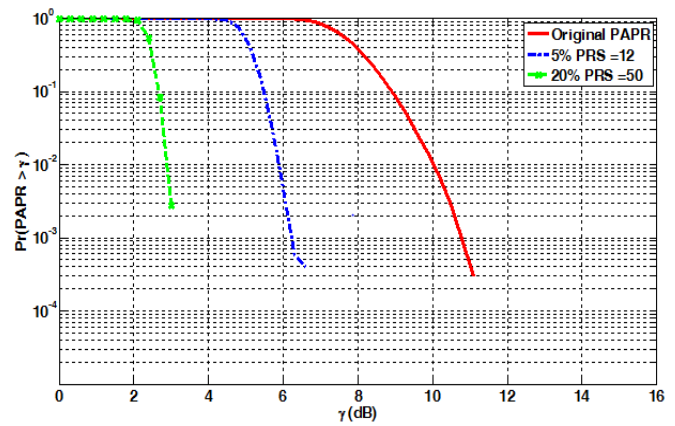
Figure 4: CCDF for different number of peak-reducing signals

The correct working of the proposed technique can be corroborated by considering that the formulation in (17) is a minimax problem and as such should yield peaks that are almost uniformly distributed [16]. This is evidenced by the scatter plots of the original signal x and the peak-reduced signal s for each modulation scheme, such as those depicted in Fig. 5 and Fig. 6 for 16-QAM; therefore confirming the correct functioning of the proposed technique.

From the scatter plots in Fig. 5 and Fig. 6, it is obvious that the technique reduces PAPR at the expense of slight increase in the average transmit power and this remains a challenge for all tone reservation techniques [17]. However, an application specific upper limit on the power increment can be set, considering the HPA specifications, and included in the optimisation as a constraint.

From the foregoing analysis it is also clear that the higher the number (L) of reserved subcarriers the higher the PAPR reduction but this implies a possible decrease in data rate especially so if all subcarriers are already allocated for data or control information transmission. However, many OFDM systems have some free subcarriers not allocated for data or system control and can be utilized for peak power reduction and thus help to mitigate data rate loss. In addition, a target PAPR that is system specific can be included as a constraint in the optimisation so as to lessen data rate loss.
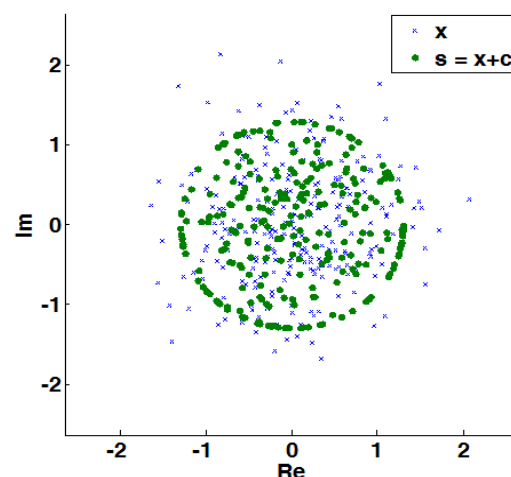
Figure 5: Scatter plot for 16-QAM complex signal

265

168
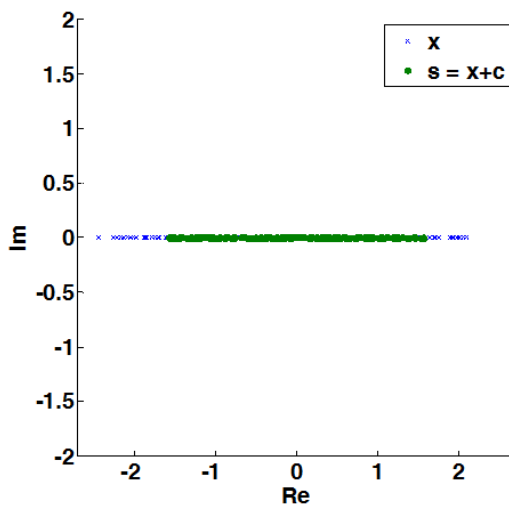
Figure 6: Scatter plot for 16-QAM real signal

VI. CONCLUSION

In this research work, a PAPR reduction technique which falls under the group of tone reservation techniques has been proposed and simulated. An investigation into its peak power reducing capability has been carried out in a 256-subcarrier OFDM system and has been found that with use of only 5% and 20% of the total subcarriers for peak reduction, the technique achieves significant PAPR reduction of 4.5 dB and 7.8 dB respectively. In addition it has been found that real time-domain OFDM signals have a PAPR higher by 2.4 dB above that of complex time-domain signals.

The simulations were carried out using several modulation schemes and with both real and complex OFDM signals. The results also show that the peak reducing capability is marginally affected by the modulation scheme or the structure of the OFDM signal, as for a fixed number of subcarriers the variations in PAPR reductions were within less than one decibel. The correct working of the proposed technique was verified by observation of scatter plots for both the original signal and the peak-reduced signal for each modulation scheme.

In this research, subcarrier modulation was assumed to be the same for all subcarriers. A suggestion for future investigation is to consider PAPR level, and its reduction in the case of adaptive subcarrier modulation when taking into account channel quality.

REFERENCES

[1] T. Jiang, Y Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM signals," IEEE Transactions on Broadcasting, Vol.54, No.2, June.2008, pp. 257-268

[2] Y. Louët and J. Palicot, "A classification of methods for efficient power amplification of signals," Annals of Telecommunications, vol. 63, no. 7-8, pp. 351–368, 2008

[3] ETSI, "Digital Video Broadcasting (DVB); Implementation guidelines for a second digital terrestrial television broadcasting system (DVBT2)," ETSI TR 102 831 v0.9.6, January 2009

[4] Z. Wang, E. Sun, Y. Zhang, "An overview of peak-to-average power ratio reduction techniques for OFDM Signals," International Journal of Mobile Network Communications & Telematics (IJMNCT) Vol. 6, No.3, June 2016

[5] E. Dahlman, S. Parkvall, J. Sköld, P. Beming, "3G Evolution: HSPA and LTE for Mobile Broadband (first edition)," London: Elsevier Ltd, 2007

[6] H. Rohling, "OFDM: Concepts for Future Communication Systems, Signals and Communication Technology," DOI: 10.1007/978-3-642-17496-4_3, Springer-Verlag Berlin Heidelberg, 2011.

[7] M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "New Results on the Peak Power of OFDM Signals Based on Oversampling," in Proc. IEEE ICC, vol. 2, Apr. 2002, pp. 866–871.

[8] J. Tellado, "Peak to average power reduction for multi-carrier modulation," Ph.D. thesis, Stanford University, Stanford, Calif, USA, September 1999

[9] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56–65, 2005.

[10] R. V. Nee, R. Prasad, "OFDM for wireless multimedia communications," Artech House, 2000

[11] S. Rao, "Engineering optimization Theory and Practice," New Age International, 3rd edition 2013

[12] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.

[13] M. S. Loboa, L. Vandenberghe, S. Boyd, H. Lebret, "Applications of second-order cone programming, Linear Algebra and its Applications," Volume 284, Issues 1–3, 15 November 1998, Pages 193–228.

[14] S. Hu, G. Wu, Y. Guan, C. Law, S. Li, "Analysis of Tone Reservation Method for WiMAX System," IEEE 2006 international symposium on communication and information technologies, October, 2006.

[15] L. Amhaimar, S. Ahyoud, A. Asselman, Elkhaldi Said, "Peak-to-Average Power Ratio Reduction based Varied Phase for MIMO-OFDM Systems," (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 9, 2016.

[16] A. Antoniou, W. Lu, "Practical optimization: algorithms and engineering applications," Springer Science and Business Media, LLC, 2007

[17] Y. Rahmatallah, S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," IEEE Communications Surveys and Tutorials, Vol. 15, No. 4, pp 1567-1592, 2013.

Appendix C: Second Publication

An Iterative Re-Weighted Least-Squares Tone Reservation Method for PAPR Reduction in OFDM Systems

STEPHEN KIAMBI, ELIJAH MWANGI, GEORGE KAMUCHA
Department of Electrical and Information Engineering
University of Nairobi
P. O. Box 30197, 00100 Nairobi
KENYA
skiambi@students.uonbi.ac.ke

Abstract: - In OFDM systems, the problems associated with a high ratio of peak-to-average power still exist. A search for a simple and practical method to reduce the ratio continues. In this paper, a robust sub-optimal tone reservation method based on iterative re-weighted least-squares minimization of infinity norm is proposed. The method is simple and has a fast quadratic convergence and per iteration complexity $O(LP)$ lower than that of the FFT, where L and P are, respectively, the number of reserved subcarriers, and nonzero elements in the desired peak-reducing signal. In addition, the method does not experience peak re-growth problems and achieves high PAPR reductions of 3.9 dB and 5.6 dB for 1.6% and 5% reserved subcarriers respectively. For 20% reserved subcarriers, the method reaches 7.4 dB PAPR reductions. These reductions are at a small cost of 0.6 dB increase in the average transmitted power. The PAPR reductions from the proposed method compare well with the highly slow and complex optimal tone reservation methods but are far much higher than from sub-optimal methods reported in literature. Simulation results also show that the method has PAPR reductions that are linear with the binary logarithm of the number of subcarriers, and this can help to predict PAPR reductions for different OFDM systems with different number of reserved subcarriers.

1 Introduction

Multicarrier transmission techniques employ parallel low data rate streams to achieve high data rate on aggregation, and to avoid multipath interference. OFDM is one such technique, which in addition, has mutually orthogonal subcarriers. Orthogonality of subcarriers makes OFDM spectrally efficient and if preserved over the radio channel, only a simple single-tap equalizer is required at the receiver to restore each of the subcarrier signals. In order to eliminate both inter-subcarrier and inter-symbol interferences, the inter-subcarrier spacing and symbol duration values are normally set above the maximum Doppler spread and multipath delay. These noble properties of OFDM have made it the preferred multiplexing technique for high data rate transmissions in many radio systems including Digital Audio Broadcasting (DAB), Digital Video Broadcasting (DVB), IEEE 802.11 Wireless Local Area Networks (WLAN), IEEE 802.16a Wireless Metropolitan Area Networks (WMAN), 4G, and 5G mobile communication networks. Despite the numerous competitive advantages, OFDM signals tend to exhibit high peak-to-average power ratio (PAPR) [1] . Distortionless processing of high PAPR signals through the transmitter

section requires the nonlinear transmitter devices mainly the digital-to-analogue converter and high power amplifier (HPA) to have a costly wide dynamic range in order to accommodate all the signal amplitudes.

In addition, the high PAPR signal affects the point of operation of the HPA. Ideally, the HPA should be operated near the saturation region in order to have high power efficiency. However, this will cause nonlinear amplification of the high signal amplitudes and in turn results to in-band and out-of-band radiations, and consequently the degradation of the bit-error rate (BER) and frequency interference in the adjacent channels. To avoid the nonlinear amplification, the HPA can be forced to operate deep in the linear region by providing it with an input power back-off as determined by the PAPR of the input signal. However, this will lower the power efficiency and in turn raise the power consumption and hence the cost of the transmitter in addition to reducing the lifetime of the battery power at the user terminals [2] . For these above reasons, it is desirable to reduce the PAPR to suitable levels, more so for OFDM systems with large number of subcarriers as they are more susceptible to unacceptably high PAPR.

Recently, different methods for PAPR reduction in

OFDM systems have been proposed in literature. They include signal coding and companding [3] , [4] , selective mapping (SLM) [4] , signal scaling [5] , partial transmit sequence (PTS) [7] and tone reservation [8] . The focus of the research is now more on the development of simple practical techniques that have low computational complexity, fast convergence rate, and high PAPR reduction.

The tone reservation (TR) approach is quite promising as it involves reserving a few subcarriers, referred to as peak reduction tones, for use to generate and carry a peak-reducing signal that reduces PAPR. Since the user data and the peak-reducing signal are on separate subcarriers, there is no distortion on the data and hence no degradation of BER. In addition, the technique does not require transmission of any side information because for demodulation, the receiver needs only the locations of the data-bearing subcarriers.

Depending on the derivation of the peak-reducing signal, tone reservation methods are either optimal or sub-optimal. Optimal methods such as the linear programmed TR (LP-TR) [8] and second order cone programmed TR (SOCP-TR) [9] have high computational complexity and slow convergence rate but can achieve high PAPR reduction that can help to benchmark the performance of the sub-optimal methods. The sub-optimal curve fitting TR (CF-TR) [11] iteratively solves a least-squares approximation (LSA) problem to generate the peak-reducing signal. However, in the CF-TR, the PAPR reduction depends on the clipping threshold and degrades if the number of reserved subcarriers is set lower than the number of zeros in the clipping noise. The scaling signal-to-clipping noise ratio (S-SCR and MS-SCR) TR [12] methods utilizes a time-domain kernel together with LSA optimized scaling factor and peak regeneration constraints. However, the two S-SCR and MS-SCR methods are still prone to peak re-growth and their PAPR reduction performance depends on the clipping ratio in use. The sub-optimal weighted TR (WTR) method [13] solves a weighted least squares optimization to generate the peak-reducing signal. However, the WTR method has difficulties finding the optimal weights, experiences peak re-growth and has poor PAPR reduction performance.

In this paper, a fast iterative re-weighted least-squares based tone reservation (IRLS-TR) method that offers high PAPR reductions in OFDM systems is proposed. The method generates the required peak-reducing signal by utilizing a robust iterative re-weighted least-squares algorithm for minimization of the infinity norm to approximate the desired peak-reducing signal. This IRLS-TR technique has low computational complexity and fast quadratic convergence in addition to offering better PAPR

reductions than the CF-TR, MS-SCR and WTR methods. The technique has great potential for practical implementations in the current and future multicarrier transmission techniques.

The organization of the rest of the paper is as follows. Section II describes the OFDM signal and statistical distribution of PAPR. Section III outlines the general concept of tone reservation techniques. Section IV describes the proposed technique while simulation results, analysis, and comparison with other techniques are in section V. Section VI concludes the paper and gives suggestion for future work.

2 OFDM Signal and PAPR

The OFDM signal is a superimposition of N mutually orthogonal subcarrier signals. At the baseband level, and during the symbol duration T, the signal has the following analytical expression:

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k)\, e^{j2\pi f_k t}. \quad (1)$$

Here, $x_k(t)$ is the kth modulated subcarrier signal with frequency $f_k = k\Delta f$, and $X(k)$ is the subcarrier-modulating symbol. The modulating symbols are either binary phase-shift keying or M-ary quadrature amplitude modulation symbols [14] . To achieve mutual orthogonality between subcarriers, the subcarrier spacing Δf is set to $1/T$. The relationship given in (1) conveniently ensures the same signal power in the time and frequency domains. Since (1) is similar to the standard inverse discrete Fourier transform (IDFT) equation, signal processing of OFDM signal is via the well-known fast Fourier transform algorithms. Due to the addition of N subcarrier signals in (1), the OFDM signal experiences envelope fluctuations that may have profound effect on the linear processing by the nonlinear devices in the transmitter. The nonlinear device of concern in this work is the high power amplifier (HPA).

Our main interest then, is to find the probability of the maximum instantaneous power of $x(t)$ being out of the linear range of the HPA. A good measure of temporal power fluctuations of a signal is the PAPR, and is defined as the ratio of the maximum instantaneous power to the average power i.e.

$$\text{PAPR}\{x(t)\} = \frac{\max\{|x(t)|^2\}}{\mathrm{E}\{|x(t)|^2\}} \quad (2)$$

where E $\{\cdot\}$ denotes the expectation operation.

Considering that the subcarrier signals $x_k(t)$ are statistically independent and assuming that N is large, then from the central limit theorem, the real and imaginary parts of $x(t)$ are Gaussian distributed and accordingly, the amplitudes of $x(t)$ are Rayleigh distributed [15] . This in turn implies that $x(t)$ could have high PAPR or in other words, has some high amplitude values that are well above the average value of the signal amplitudes.

In practice, the signal $x(t)$ is processed digitally. Therefore, there is need to approximate the continuous-time PAPR in (2) with its discrete-time counterpart given by

$$\text{PAPR}\{x\} = \frac{\max\limits_{0 \le n < N} \{|x(n)|^2\}}{E\{|x(n)|^2\}} \qquad (3)$$

where $x = [x(0), x(1), \dots, x(N-1)]^T$, is the discrete-time signal obtained after sampling of signal $x(t)$.

The signal samples $x(n)$ have magnitudes that are still Rayleigh distributed. Equation (3) is valid if signal $x(t)$ is sampled at a rate sufficiently greater than the Nyquist rate by a factor of at least four in order to avoid missing the peak value [16] . Another important measure of PAPR is the complementary cumulative distribution function (CCDF), which is the probability that PAPR exceeds a certain threshold. The derivation of CCDF has been well treated in [9] and is given by

$$\Pr(\text{PAPR}\{x\} > \gamma) = 1 - (1 - e^{-\gamma})^N \qquad (4)$$

where γ is the threshold PAPR, N is the total number of subcarriers, and $\Pr(.)$ denotes the probability operator. The CCDF metric is a performance tool widely used to measure how well a proposed method reduces the PAPR.

3   Tone Reservation Concept

All tone reservation methods follow the same concept of adding one signal, referred here to as peak-reducing signal, to another signal having high PAPR in order to reduce the PAPR. Figure 1 is an illustration of the concept and has $X(k)$ and $C(k)$ as the frequency-domain subcarrier modulating data symbols and peak-reducing coefficients respectively. The modulating symbols form the data vector $X = [X(0), X(1), \dots, X(N-1)]^T$ with all nonzero values except in L positions reserved for peak-reducing signal. Similarly, the peak-reducing coefficients form the vector $C = [C(0), C(1), \dots, C(N-1)]^T$ with all zero values except in L positions reserved for peak power reduction. After the IDFT operation on X and C, the resulting time signals $x = [x(0), x(1), \dots, x(N-1)]^T$ and

$c = [c(0), c(1), \dots, c(N-1)]^T$, are combined to generate a low-PAPR signal i.e. PAPR$\{x - c\}$ < PAPR$\{x\}$.
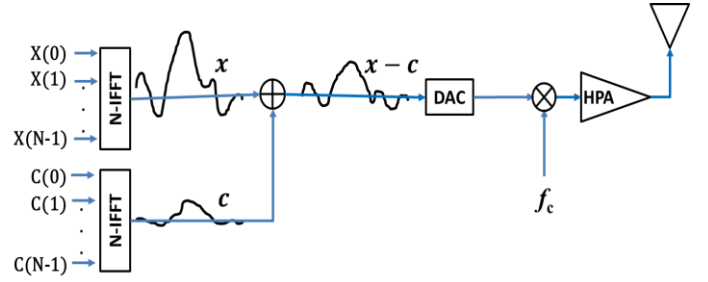
Figure 1: Tone reservation concept

From the foregoing discussion, the signal processing operations can be described by the equation

$$x(n) - c(n) = x(n) - \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} C(k) e^{j2\pi kn/N} \qquad (5)$$

or in matrix notation by

$$x - c = x - QC \qquad (6)$$

where $Q \in \mathbb{C}^{N \times N}$ is the IDFT matrix with elements $\frac{1}{\sqrt{N}} e^{j2\pi kn/N}$ and all $x, c, C \in \mathbb{C}^N$. Since C has $N - L$ zeros, the computational complexity of (6) can be reduced further by expressing the peak-reducing signal in the form

$$c = \widehat{Q}\widehat{C}. \qquad (7)$$

Here $\widehat{C} \in \mathbb{C}^L$ contains only the nonzero values of C and $\widehat{Q} \in \mathbb{C}^{N \times L}$ is a submatrix of Q formed by choosing the L columns corresponding to the reserved subcarriers.

The task now is to find the peak-reducing coefficients or vector $\widehat{C}$ that minimizes the PAPR of x. This problem can be formulated as a minimax [17] optimization problem of the form

$$\min_{\widehat{C}} \max |x - \widehat{Q}\widehat{C}|. \qquad (8)$$

Here, $\widehat{C} \in \mathbb{C}^L$ is the optimization variable while $x \in \mathbb{C}^N$ and $\widehat{Q} \in \mathbb{C}^{N \times L}$ are the problem parameters. To date, sub-optimal algorithms to solve the problem in (8) at a reduced computational complexity and short convergence time and at the same time achieve high PAPR reductions continue originating.

Although tone reservation methods offer quite an attractive approach for reducing PAPR, they are not without drawbacks. First, the average power of the PAPR-reduced signal increases. This calls for one to limit the increase to a minimum as the algorithm executes to ensure compatibility with HPA specifications. Second, the number of reserved subcarriers, L, reduces the data rate because the

reserved subcarriers do not carry any user data. Therefore, to avoid a high data rate loss, L should be small as can be practically possible.

4 Proposed Method

This paper proposes a novel sub-optimal tone reservation method, hereafter referred to as the iterative re-weighted least-squares tone reservation (IRLS-TR) method, which performs fast approximation of peak-power reducing signals. The method approximates a desired peak-reducing signal with a signal designed in accordance with the frequency allocation constraints as imposed by the tone reservation concept discussed in section III.

4.1 Algorithm

Any PAPR reducing method attempts to reduce the peak power to a value close to the average power. With this consideration, the desired peak-reducing signal that is required to cancel the high peaks of the OFDM signal can be posed as the OFDM signal amplitudes above the average value. Subsequently, the desired peak-reducing signal has the analytical equation

$$d(n) = \begin{cases} \dfrac{x(n)}{|x(n)|}(|x(n)| - \bar{x})), & |x(n)| > \bar{x} \\ 0, & |x(n)| \le \bar{x} \end{cases} \quad (9)$$

$$n = 0,1, \dots, N - 1.$$

Here, $\bar{x}$ is the average value of the discrete-time OFDM signal. The elements of $\widehat{C}$ are then determined to have the time-domain peak-reducing signal in (7) equal to the signal in (9) by solving the matrix equation

$$\widehat{Q}\widehat{C} = d. \quad (10)$$

However, since $\widehat{Q} \in \mathbb{C}^{N \times L}$, $\widehat{C} \in \mathbb{C}^L$ and $d \in \mathbb{C}^N$, the system of linear equations in (10) is rectangular and hence overdetermined. For such a system, there is in general no exact solution to $\widehat{C}$ and therefore, the time-domain peak-reducing signal $c = \widehat{Q}\widehat{C}$ can only be determined to approximate the desired signal vector d i.e.

$$\widehat{Q}\widehat{C} \approx d, \quad (11)$$

by minimizing the residual error

$$\epsilon = \widehat{Q}\widehat{C} - d \quad (12)$$

using some norm as a measure of the error size.

In addition to the minimization of the error in (12), for the problem at hand, the high peaks of the designed peak-reducing signal, $c = \widehat{Q}\widehat{C}$, should approximate those of the

signal d in (9) as practically as is possible in order to cancel all the high peaks in the original OFDM signal and thereby reduce the PAPR. The design of the peak-reducing signal will therefore involve finding the elements of $\widehat{C}$ that minimizes the L_∞ norm of the error. This is equivalent to the minimization of the L_p norm of the error for a large value of p [18] as given by the equation

$$\min_{\widehat{C}} ||\widehat{Q}\widehat{C} - d||_p, \quad p \to \infty. \quad (13)$$

In practice, the solution to the L_p problem approximates that of L_∞ if (13) is solved for $p \ge 10$ [19]. However, there is no analytical method for finding the optimal approximation solution for any norm other than the L_2 norm. For this reason, it is necessary to transform the L_p problem into an equivalent simple weighted least squares (WLS) problem that can be solved analytically. This WLS problem has the form

$$\min_{\widehat{C}} ||W(\widehat{Q}\widehat{C} - d)||_2 \quad (14)$$

where W is a real $N \times N$ diagonal weighting matrix that applies large weights on the high signal peaks to emphasize their minimization. If the diagonal weights in W are known, then equation (14) can be solved by a simple method that has the closed form solution [20]

$$\widehat{C} = [\widehat{Q}^* W^T W \widehat{Q}]^{-1} \widehat{Q}^* W^T W d. \quad (15)$$

Here, the superscripts T and $*$ denote matrix transpose and conjugate transpose respectively.

In order to make (14) to be equivalent to (13), there is need for careful selection of the diagonal weights $w(n)$ of the weighting matrix W. Rewriting the L_2 norm of the weighted error, $\varepsilon = W(\widehat{Q}\widehat{C} - d)$, in (14) as

$$||\varepsilon||_2 = (\textstyle\sum_{n=0}^{N-1} w^2(n)|\epsilon(n)|^2)^{1/2}, \quad (16)$$

and assigning the error to the weights according to

$$w(n) = |\epsilon(n)|^{(p-2)/2}, \quad (17)$$

then (16) becomes the L_p norm of the error as in (13), i.e.

$$||\varepsilon||_p = \left(\sum_{n=0}^{N-1} |\epsilon(n)|^{(p-2)} |\epsilon(n)|^2 \right)^{1/p}. \quad (18)$$

Therefore, solving the WLS problem in (14) is identical to solving the L_p problem in (13).

However, solving problem (14) cannot be done in one-step because one needs to find the weights that give the optimal approximation. To this end, this work employs a robust iterative re-weighted least squares (IRLS) algorithm

to find the solution to the WLS problem. The algorithm builds from the analytical solution in (15) but with per iteration re-weighting until convergence to the large L_p norm of (13). In its most basic form, the IRLS algorithm starts by solving for $\widehat{C} \in \mathbb{C}^L$ from (15) with all initial weights set to one i.e. $w(n) = 1$. Then, it computes the error vector from (12), followed by new weights from (17) that are for use in the next iteration. Using the new weights, the algorithm finds a new solution $\widehat{C}$ and this process repeats until convergence when the L_p norm of the error is quite small e.g. less than 10^{-4} or until the number of iterations reaches a predetermined maximum iteration number.

The basic IRLS algorithm presented above has two concerns that need to be addressed. Firstly, the algorithm may not converge and/or is numerically unstable for some L_p norms. Secondly, it has linear convergence and therefore if it converges and is numerically stable, it does so very slowly as to be of any practical use. To overcome these two shortcomings, the basic algorithm is transformed into a form of the Newton's method [21] in which the solution is only partially updated at each iteration. With this update, the solution at ith iteration becomes

$$\widehat{C}_i = \mu_i \widehat{C}_i^c + (1 - \mu_i)\widehat{C}_{i-1} \qquad (19)$$

where $\mu_i = 1/(p_i - 1)$ is the update parameter [22] [23] and $\widehat{C}_i^c \in \mathbb{C}^L$ is the current WLS solution. However, as is common with most Newtonian methods, this modification makes the algorithm sensitive to initial approximations and this may affect the initial convergence rate.

In order to improve on the initial convergence rate, the value of p is increased gradually from its initial value of two to the final value of the L_p norm that is being used to approximate the L_∞ norm. This modification is quite similar to the homotopy [24] [25] and is done iteratively by multiplying p with a convergence parameter, α, of between one and two depending on the required rate of convergence. The value of p at the i-th iteration is then determined by

$$p_i = \min(p, \alpha p_{i-1}). \qquad (20)$$

In addition, the convergence parameter is determine by the value of p in the L_p norm and the maximum iteration number according to

$$\alpha = 10^{\log(p)/M} \qquad (21)$$

where p and M are, respectively, the L_∞ approximation norm and maximum iteration number. The settings in (20) and (21) guarantee reliable convergence of the algorithm to the optimal approximation solution. This is because as p progressively increases from the initial value, the

difference from one L_p solution to the next is small, and when the algorithm approaches the neighbourhood of the desired L_∞ approximation norm it iterates several times at the same p.

The IRLS-TR algorithm can be summarized as follows.

Algorithm: IRLS-TR
1. Set the L_∞ approximation norm p, initial weights to $w = 1$ and the L_p norm error threshold ε_{th}.
2. Set the maximum iteration number M and convergence parameter $\alpha = 10^{\log(p)/M}$.
3. Choose reserved subcarrier locations and compute the IFFT submatrix $\widehat{Q}$.
4. Generate the original OFDM time-domain signal x.
5. Calculate the desired peak-reducing signal d from the signal x.
6. Initialize the iteration counter $i = 0$ set the initial p-norm value $p_0 = 2$.
7. Calculate the initial LS solution $\widehat{C}_0 = \widehat{Q}^{-1}d$.
8. Calculate the error vector $\epsilon_i = \widehat{Q}\widehat{C}_i - d$.
9. Calculate the new weighting matrix elements using $w_i(n) = |\epsilon_i(n)|^{(p_i-2)/2}$.
10. Calculate the LS solution $\widehat{C}_i^c = [\widehat{Q}^* W_i^T W_i \widehat{Q}]^{-1} \widehat{Q}^* W_i^T W_i d$.
11. Calculate the new update parameter $\mu_i = 1/(p_i - 1)$.
12. Update the LS solution to $\widehat{C}_i = \mu_i \widehat{C}_i^c + (1 - \mu_i)\widehat{C}_{i-1}$.
13. Calculate the L_p norm error $\|\varepsilon_i\|_p = (\sum_{n=0}^{N-1}|\epsilon_i(n)|^p)^{1/p}$.
14. Set $i = i + 1$. If $i < M$ or $\|\varepsilon_i\|_p > \varepsilon_{th}$, update the p-norm value to $p_i = \min(p, \alpha p_{i-1})$ and go to step 8. Otherwise, transmit $x - \widehat{Q}\widehat{C}_i$ and terminate algorithm.
15. End

4.2 Convergence

With the modifications in (19), (20) and (21), the algorithm quadratically converges in just a few iterations as illustrated in Fig. 2. The figure depicts a typical plot of the convergence curve of the error in (18) at each iteration as the algorithm was executed. As shown in the figure, the algorithm converged after about 10 iterations.

4.3 Computational Complexity

At each iteration, the algorithm spends most of the computational time to find the solution in (15) for the WLS problem

$$W\widehat{Q}\widehat{C} = Wd. \qquad (22)$$

Since the matrices $W \in \mathbb{R}^{N \times N}$, $\widehat{Q} \in \mathbb{C}^{N \times L}$, and $\widehat{C} \in \mathbb{C}^{L}$, the computational complexity of the algorithm per iteration is then given by $O(NL)$. This complexity can be reduced further if the algorithm computes the solution from only the non-zero elements in vector d. Denoting the number of non-zeros elements in d by P, the computational complexity of the algorithm will then become $O(LP)$. Now, since P is always less than $N/2$, with a few reserved subcarriers for PAPR reduction, the algorithm's complexity of $O(LP)$ is less than $O(N \log_2 N)$ of the fast Fourier transform algorithms.
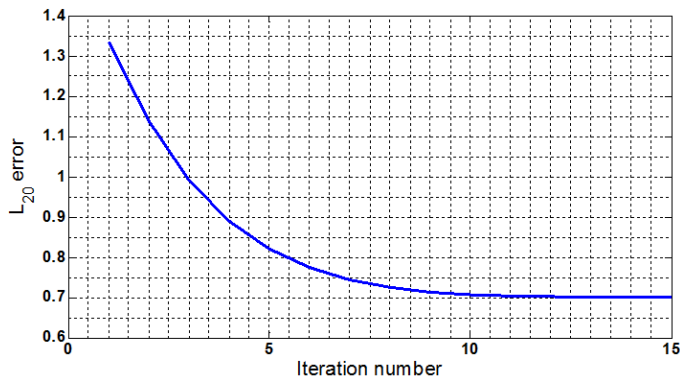
Figure 2: Convergence curve

5 Results and Discussion

The IRLS-TR method proposed in section IV yields the frequency-domain vector $\widehat{C} \in \mathbb{C}^{L}$. The product of this vector and the IDFT submatrix $\widehat{Q}$ is the time-domain signal used to reduce the PAPR of the OFDM signal. Matlab simulations using the OFDM parameters in Table 1 assessed the performance of the method when applied to reduce the PAPR in OFDM systems comprising different subcarrier modulations and different numbers of reserved subcarriers. The locations of the reserved subcarriers were randomly selected in all OFDM symbols.

Table 1: OFDM Parameters

FFT size, N	16, 32, 64, 128, 256, 512, 1024, 2048
Subcarrier modulation	QPSK, 16-QAM, 64-QAM, 256-QAM, 1024-QAM
Number of OFDM symbols	10,000
Number of reserved subcarriers, L	5%, 10%, 15%, and 20% of N
Reserved subcarrier locations	random

To achieve high PAPR reductions, the method generates a peak-reducing discrete-time signal with the highest peaks

approximating those of the desired peak-reducing signal as illustrated in Fig. 3. For example, at $n = 137$ the magnitude of the highest peak of the generated peak-reducing signal is 1.214 and this approximates 1.806 of the desired signal. Thus, subtracting the generated signal from the original high PAPR OFDM signal reduces the PAPR. In addition, the peaks of the designed signal are below all the major spikes of the desired signal. This then avoids the peak re-growth problem where the cancellation of one peak regenerates a new peak at a different position as reported in methods that utilize peak-cancelling kernels [7] [26] [27] to reduce PAPR.

The proposed IRLS-TR method achieves high PAPR reduction by utilizing only a small number of reserved subcarriers to generate and carry the PEAK-REDUCING SIGNAL. The CCDF curves in Fig. 4 illustrates this point for the cases where out of 256 total number of subcarriers, 2, 4, 8, 13, 20, and 52 subcarriers were reserved. For example, with 4 and 13 reserved subcarriers, corresponding to 1.6% and 5% of the total subcarriers, the method achieved 3.9 dB and 5.6 dB PAPR reductions at CCDF = 10^{-3}. Compared to the SOCP-TR by Kiambi et al and LP-TR by Tellado, with 5% reserved subcarriers, the reduction of 5.6 dB from the proposed IRLS-TR method is better than the 4.6 dB and 5.3 dB from the two methods respectively.

For the case with 20% reserved subcarriers, the IRLS-TR method proposed here achieved a PAPR reduction of 7.4 dB, and this again is close to the 8.0 dB from the optimal methods.
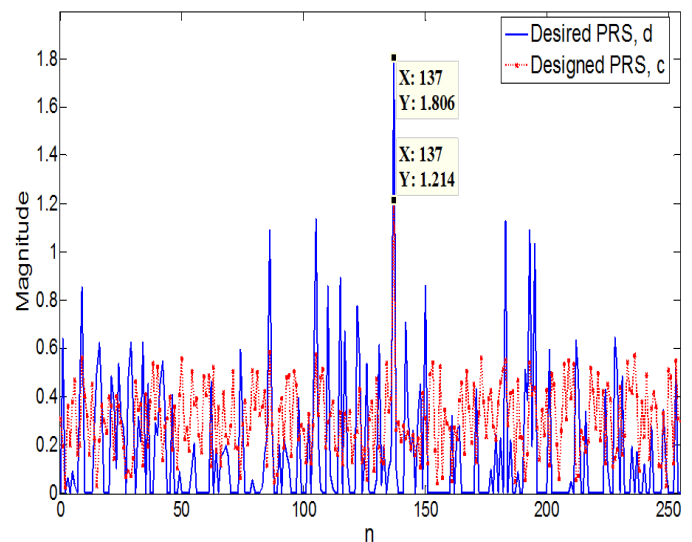
Figure 3: Desired and designed peak-reducing discrete-time signals

To compare the performance of the proposed IRLS-TR method to the sub-optimal methods, the actual results reported in CF-TR by Jiang et al, MS-SCR by Wang et al, and WTR by Xin and Yi were considered. From the results

for the PAPR reduction reported by Jiang et al for the CF-TR, it is clear that the proposed IRLS-TR has a better performance. For example, at CCDF=10^{-2} and 6.25% reserved subcarriers in a 16-QAM system, the CF-TR method achieved 4.4 dB that is lower than the 5.6 dB from the proposed IRLS-TR method. In addition, in the proposed IRLS-TR, the number of reserved subcarriers can be freely set unlike in the CF-TR in which PAPR reduction degrades if the number is set to a value lower than the number of zeros, which varies per iteration, in the clipping noise.

The MS-SCR by Wang et al generates a kernel signal to cancel multiple peaks and has fast convergence but still can exhibit peak re-growth and has lower PAPR reductions than the proposed IRLS-TR. For example, for a 16-QAM system with 12.5% reserved subcarriers at CCDF=10^{-3}, the MS-SCR could only manage 4.7 and 5.2 dB PAPR reductions when using high clipping ratios of 2 and 3 respectively, which reductions are quite low compared to the 6.6 dB from the proposed IRLS-TR method for the same system.

For the new WTR method by Xin and Yi, which also tries to mitigate the peak re-growth problem in the traditional WTR method [28] , the PAPR reduction performance is quite poor. The WTR method reported a PAPR reduction of 2.7 dB at CCDF of 10^{-3} for 5% reserved subcarriers in QPSK-modulated OFDM system. This reduction is less than half of the 5.6 dB given by the proposed IRLS-TR method for the same OFDM system.

In summary therefore, based on the results reported for the CF-TR, MS-SCR and new WTR methods, the proposed IRLS-TR method has better and improved PAPR reduction performance than the three methods. In addition, the proposed IRLS-TR method does not experience peak re-growth problem and the required number of reserved subcarriers can be freely set.
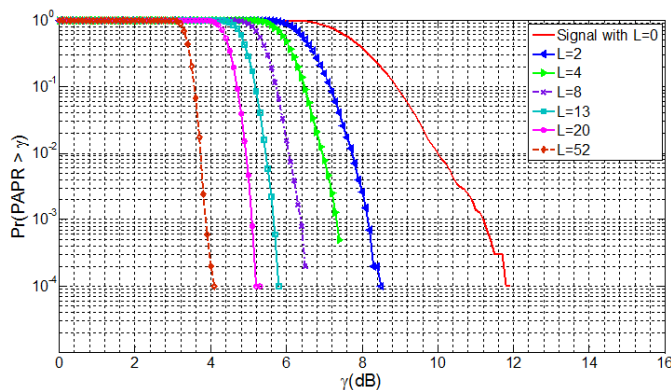
Figure 4: CCDF curves for QPSK-modulated OFDM with 256 subcarriers

On the issue of the increase in the average transmitted power that is expected with all tone reservation methods,

the proposed IRLS-TR method exhibits only a small increase of about 0.6 dB. The sub-optimal CF-TR by Jiang et al method reported the same value while the optimal LP-TR method by Tellado had an increase of 1 dB. Due to the high PAPR reductions achieved by the proposed IRLS-TR method, the small increase of 0.6 dB in the transmitted power is tolerable.

Lastly, the simulation results for the same subcarrier modulation but with different number of OFDM subcarriers revealed that the PAPR reduction from the proposed IRLS-TR method increases almost linearly with $\log_2 N$ as illustrated in Fig. 5. The figure shows PAPR reductions for 5%, 10%, 15% and 20% reserved subcarriers against the binary logarithm of the number of OFDM subcarriers. Since all the plots resemble straight-line graphs, each of them can be estimated analytically by a linear function. This makes it easier to estimate the PAPR reductions for different total number of OFDM subcarriers for a fixed number of reserved subcarriers.

Similarly, since the lines have almost the same slope, one can estimate the increase in the PAPR reduction accruing from an increase in the number of reserved subcarriers. For example, for the 5% and 15% reserved subcarriers, the PAPR reductions are respectively given by the linear equations $0.4 \log_2(N) + 0.6$, and $0.4 \log_2(N) + 1.7$. The difference of 1.1 dB between the two equations is the PAPR reduction achieved due to the increase of the number of reserved subcarriers from 5% to 15%.
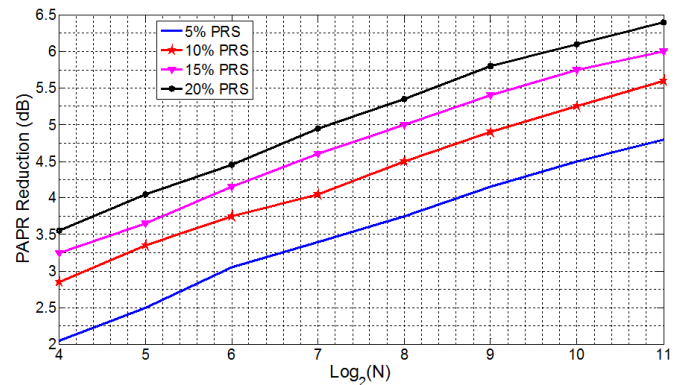
Figure 5: PAPR reduction variation with number of subcarriers

6 Conclusion

In this paper, a new sub-optimal tone reservation technique for PAPR reduction in OFDM systems, referred to as IRLS-TR, is proposed. Investigation into its PAPR reduction capability has found that with only 1.6% and 5% of total subcarriers reserved for PAPR reduction, it respectively achieves 3.9 dB and 5.6 dB of PAPR reduction at a small

cost of 0.6 dB increase in the average transmitted power. The method has a fast quadratic convergence and a low computational complexity per iteration of $O(LP)$ which is less than $O(N \log_2 N)$ of FFT. Here, L is the number of reserved subcarriers, N is the total number of subcarriers, and P is the number of nonzero elements in the desired peak-reducing signal.

In addition, the proposed method does not experience the peak re-growth problem associated with TR methods that cancel signal peaks using a time-domain kernel signal and the traditional weighted tone reservation methods. Further investigations revealed that the method has PAPR reductions that are linear with $\log_2 N$ and therefore one can predict PAPR reductions for different OFDM systems with different number of reserved subcarriers.

The effect of the choice of the desired peak-reducing signal on the PAPR reduction capabilities of the proposed method shall be investigated in future work.

References:

[1]   Y. Rahmatallah and S. Mohan, Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy, *IEEE Commun. Surveys and Tutorials*, Vol. 15, No. 4, 2013, pp. 1567-1592.

[2]   Y. Louët and J. Palicot, A classification of methods for efficient power amplification of signals, *Annals of Telecom.*, vol. 63, No. 7-8, 2008, pp. 351–368.

[3]   S. Takizawa and H.Ochiai, PAPR Reduction of OFDM with Trellis Shaping based on p-Norm Minimization, *IEEE Wireless Communication Letters,*DOI 10.1109/LWC.2019.2903247, 2019.

[4]   F. Sandoval, G. Poitau, and F. Gagnon, On Optimizing the PAPR of OFDM Signals With Coding, Companding, and MIMO, *IEEE Access*, Vol. 7, 2019.

[5]   M. I. Al-Rayif, H. Seleem, A. Ragheb, and S. Alshebeili, A Novel Iterative-SLM Algorithm for PAPR Reduction in 5G Mobile Fronthaul Architecture, *IEEE Photonics Journal*, Vol. 11, No. 1, 2019.

[6]   A. M. Rateb and M. Labana, An Optimal Low Complexity PAPR Reduction Technique for Next Generation OFDM Systems, *IEEE Access*, Vol. 7, Feb. 2019.

[7]   Jawhar, L. Audah, M. Taher, K. Ramli, N. Shah, M. Musa, and M. Ahmed, A Review of Partial Transmit Sequence for PAPR Reduction in the OFDM Systems, *IEEE Access*, Vol. 7, 2019.

[8]   H. Boche, J. Mönich, and E. Tampubolon, Complete Characterization of the Solvability of PAPR Reduction for OFDM by Tone Reservation, *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2023-2027.

[9]   J. Tellado, Peak to average power reduction for multicarrier modulation, Ph.D. dissertation, Dept. Elect. Eng., Stanford University, Stanford, CA, USA, 2000.

[10]   S. Kiambi, E. Mwangi, and G. Kamucha, Effect of OFDM Signal Structure and Subcarrier Modulation on the Reduction of the Signal Peak Power, *IEEE Africon 2017 Proceedings*, 2017, pp. 262–266.

[11]   T. Jiang, C. Ni, C. Xu, and Q. Qi, Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems, *IEEE Comm. Letters*, vol. 18, no. 5, 2014, pp. 805–808.

[12]   J. Wang, Xin Lv, and W. Wu, SCR-Based Tone Reservation Schemes With Fast Convergence for PAPR Reduction in OFDM System, *IEEE Wireless Communications Letters*, Vol. 8 , Issue 2, Apr. 2019, pp. 624 – 627.

[13]   L. Xin and W. Yi, A New Weighted Tone Reservation Method for PAPR Reduction in OFDM Systems, *Journal of Communication*, Vol. 9, No. 12, 2014, pp. 980-986.

[14]   E. Dahlman, S. Parkvall, and J. Sköld, P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband (first edition),* London: Elsevier Ltd, 2007.

[15]   H. Rohling, *OFDM: Concepts for Future Communication Systems, Signals and Communication Technology*, Springer-Verlag Berlin Heidelberg, 2011.

[16]   M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, New Results on the Peak Power of OFDM Signals Based on Oversampling, *in Proc. IEEE ICC*, Vol. 2, 2002, pp. 866–871.

[17]   S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, USA, 2004.

[18]   S. Burrus, J. A. Barreto, and I. W. Selesnick, Iterative re-weighted least-squares design of FIR filters, *IEEE Trans. on Signal Proc.*, Vol. 42, No. 11, 1994, pp. 2926-2936.

[19]   R. H. Byrd and D. A. Pyne, Convergence of the iteratively re-weighted least squares algorithm for robust regression, Technical report 313, Dept. of Math. Sci, Johns Hopkins University, Baltimore, Maryland, 1979.

[20]   Adi Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications,* Springer-Verlag,

New York, 2nd ed, 2003.

[21] R. L. Burden and J. D. Faires, *Numerical Analysis*, Brooks/Cole, Cengage Learning, Boston, USA, 9th ed., 2011.

[22] S. W. Kahng, Best L_p approximation, *Mathematics of Computation*, vol. 26, No. 118, 1972, pp. 505-508.

[23] R. Fletcher, J. A. Grant, and M. D. Hebden, The calculation of linear best L_p approximations, *Computer Journal*, Vol.14, 1971, pp. 27-279.

[24] V. L. Stonick and S. T. Alexander, Globally optimal rational approximation using homotopy continuation methods, *IEEE Trans. on Signal Proc.*, Vol.40, No.9, 1992, pp. 2358-2361.

[25] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer series in operations research. New York, NY: Springer-Verlag, 1999.

[26] J. Song and H. Ochiai, Performance analysis for OFDM signals with peak cancellation, *IEEE Trans. Commun.* vol. 64, no. 1, 2016, pp. 261–270.

[27] Liz , L. Huy, F. Yang, L. Ding, and T. Song, Tone Reservation Ratio Optimization for PAPR Reduction in OFDM Systems, *IEEE Wireless Comm. and Networking Conference (WCNC)*, 2018.

[28] H. J. Yin, R. Z. Yang, X. L. Luo, L. Jiang, and L. J. Zhu , Weighted tone reservation for OFDM PAPR reduction, U.S. Patent 7796498, 2008.

178

Appendix D: Third Publication

Use of Preset Reserved Tones in Reduction of PAPR in OFDM Systems

Stephen Kiambi, Elijah Mwangi, George Kamucha
School of Engineering
University of Nairobi, Kenya
skiambi@students.uonbi.ac.ke

Abstract—**In this paper, a proposed sub-optimal iterative re-weighted least-squares tone reservation method is applied to reduce PAPR in OFDM systems with fixed locations of reserved tones. Three ways of placing the reserved tones are investigated and compared: random, contiguous, and equal-spaced placements in OFDM symbols. The proposed method has a fast quadratic convergence and low per iteration complexity of $O(NL)$ where N is the number of reserved subcarriers, and L is the number of the reserved tones. The method was found to achieve, respectively, high PAPR reductions of 5.3 dB for the random and 4.5 dB for both the contiguous and equal-spaced placements when only 5% of the total number of subcarriers is reserved for PAPR reduction. These PAPR reductions are higher than from other related optimal and sub-optimal methods reported in literature.**

Keywords—**Orthogonal Frequency Division Multiplexing; Peak-to-Average Power Ratio; Tone Reservation**

I. INTRODUCTION

OFDM transmission technique utilizes parallel low data rate streams to achieve high data rate and reduce inter-symbol interference. In addition, the technique has mutually orthogonal subcarriers and this makes it efficient in terms of bandwidth utilization. These OFDM characteristics have made it the multiplexing and transmission technique of choice for high data rate transmissions in many radio systems including Digital Video Broadcasting (DVB), IEEE 802.11 Wireless Local Area Networks (WLAN), IEEE 802.16a Wireless Metropolitan Area Networks (WMAN), 4G, and 5G mobile communication networks. However, OFDM signals tend to exhibit a high ratio of peak power to average power abbreviated as PAPR [1].

To pass a high PAPR signal through the high power amplifier (HPA) in the transmitter without any distortion, the HPA must be forced to operate deep into the linear region by providing it with a large enough input power back-off as determined by the PAPR of the input signal. However, this lowers the power efficiency and therefore raises the power consumption and hence the cost of the transmitter in addition to reducing the battery power lifetime at user terminals. It is therefore considered necessary to reduce the high PAPR in OFDM systems to comfortable levels especially so for systems with large number of subcarriers as they tend to exhibit unacceptably high PAPR.

Recently, various methods for reducing PAPR in OFDM systems have been proposed in literature. They include signal coding [2], selective mapping (SLM) [3], signal scaling [4], partial transmit sequence (PTS) [5], and tone reservation (TR) [6]. The TR approach, due to the simplicity, has of late proved quite promising. This approach reserves a few subcarriers, called peak-reducing tones, which are utilized to generate and carry a peak-reducing signal that is used to reduce the PAPR of the original OFDM signal. Because the user data and the peak-reducing signal are designated to separate subcarriers, no distortion is introduced on the data and hence no BER degradation. In addition, the TR approach does not require transmission of any side information to the receiver if the locations of the reserved subcarriers are predetermined and fixed in all OFDM symbols.

Most of the TR methods [7] [8] [9] [10] proposed in literature have randomly reserved the locations of the peak-reducing tones despite the fact that preset locations would be preferred for practical transmission systems. The sub-optimal kernel-based TR (KTR) in [11] has utilized a contiguous and equal-spaced set of reserved tones but the method suffers from peak-regrowth that slows the convergence. In [12], a dynamic tone allocation in a standard optimal TR (DTA-OTR) method is proposed where a L_1-norm minimization is used to find the reserved subcarriers set that achieves a good PAR reduction. However, in addition to the DTA-OTR method's high computational complexity and slow convergence, side information must be transmitted or the receiver has to estimate in a complex way the locations of the reserved tones. The work in [13] explored random, block and equal-spaced placement of reserved tones in a standard optimal TR (OTR) method, but the method is again slow, has high computational complexity, and achieves low PAPR reductions.

In this paper, a fast sub-optimal iterative re-weighted least-squares based tone reservation (IRLS-TR) method that offers good PAPR reductions in OFDM systems with fixed locations of reserved tones is proposed. The method fast estimates the peak-reducing signal by utilizing a robust iterative re-weighted least-squares algorithm for minimization of L_∞-norm. This IRLS-TR technique has low computational complexity and converges quadratically in addition to achieving better PAPR reductions than the K-TR, DTA-OTR and OTR methods when the reserved subcarrier locations in OFDM symbols are fixed.

The order of the rest of the paper is as follows. Section II discusses PAPR of OFDM signals. In Section III, a review of the tone reservation concept is given. Section IV outlines the proposed method while simulation results and comparison are in section V. In Section VI is the conclusion and suggestion for future work.

II. PAPR OF OFDM SIGNAL

The OFDM signal is an aggregation of N mutually orthogonal subcarrier signals. At the baseband level, and during one symbol duration, the discrete-time signal can be expressed as

$$x[\text{n}] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k)\, e^{\frac{j2\pi nk}{N}}, \tag{1}$$
$$n = 0,1,2, \dots N-1$$

where $X(k)$ is the subcarrier-modulating symbol. These symbols are derived from signal constellations of either binary phase-shift keying or M-ary quadrature amplitude modulations [14]. The normalization by $\sqrt{N}$ ensures that the signal power in the time-domain is equal to that in the frequency-domain. Equation (1) is quite similar to the standard inverse discrete Fourier transform (IDFT) equation and therefore an OFDM system utilizes the fast Fourier transform for signal processing. Due to the addition of N subcarrier signals, the OFDM signal may exhibit peak power fluctuations that may limit the linear processing by the nonlinear HPA in the transmitter section especially so if some of the high peaks cross into the saturation region.

The fluctuation in the signal amplitudes is measured by peak-to-average power ratio (PAPR) given by

$$\text{PAPR}\{\boldsymbol{x}\} = \frac{\max\limits_{0 \le n < N}\{|x(n)|^2\}}{E\{|x(n)|^2\}} \tag{2}$$

where the vector $\boldsymbol{x} = [x(0), x(1), \dots, x(N-1)]^T$, and $E\{.\}$ is the expectation operator.

The interest in most of the times is to determine whether the PAPR exceeds a predetermined threshold of the HPA. This is measured by the complementary cumulative distribution function (CCDF) of the PAPR, which is basically the probability that the PAPR exceeds the threshold. The derivation of CCDF has been well treated in [15] and is given by

$$\Pr(\text{PAPR}\{\boldsymbol{x}\} > \gamma) = 1 - (1 - e^{-\gamma})^N \tag{3}$$

where γ is the threshold PAPR, and $\Pr(.)$ denotes the probability operator.

III. TONE RESERVATION

All tone reservation methods add a peak-reducing signal to the original high-PAPR OFDM signal in order to reduce the PAPR. This is illustrated in Fig. 1, where $X(k)$ and $C(k)$ are, respectively, the subcarrier-modulating data symbol and peak-reducing coefficient.
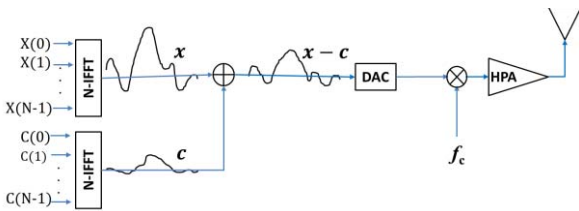
Figure 1: Tone reservation concept

To avoid distorting the user data, the data symbols and peak-reducing coefficients are allocated separate subcarriers i.e. the data vector $\boldsymbol{X} = [X(0), X(1), \dots, X(N-1)]^T$ has all nonzero values except in L positions reserved for peak-reducing signal, while the peak-reducing vector $\boldsymbol{C} = [C(0), C(1), \dots, C(N-1)]^T$ has all zero values except in the L reserved positions. After the

IDFT of the frequency-domain vectors, $\boldsymbol{X}$ and $\boldsymbol{C}$, the resulting time-domain vectors $\boldsymbol{x} = [x(0), x(1), \dots, x(N-1)]^T$ and $\boldsymbol{c} = [c(0), c(1), \dots, c(N-1)]^T$ are combined to generate a low-PAPR signal i.e. $\text{PAPR}\{\boldsymbol{x} - \boldsymbol{c}\} < \text{PAPR}\{\boldsymbol{x}\}$.

From the foregoing, the signal processing in tone reservation methods can be described by the equation,

$$\boldsymbol{x} - \boldsymbol{c} = \boldsymbol{x} - \boldsymbol{QC}, \tag{4}$$

where $\boldsymbol{Q} \in \mathbb{C}^{N \times N}$ is the IDFT matrix with elements $\frac{1}{\sqrt{N}} e^{j2\pi nk/N}$ and all $\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{C} \in \mathbb{C}^N$. Considering that $\boldsymbol{C}$ has only L nonzero values, the peak-reducing signal is simply given by

$$\boldsymbol{c} = \widehat{\boldsymbol{Q}}\widehat{\boldsymbol{C}}. \tag{5}$$

Here $\widehat{\boldsymbol{C}} \in \mathbb{C}^L$ contains only the nonzero values of $\boldsymbol{C}$, and $\widehat{\boldsymbol{Q}} \in \mathbb{C}^{N \times L}$ is a submatrix of $\boldsymbol{Q}$ consisting of only the L columns corresponding to the reserved subcarriers.

The task of the tone reservation technique is then to derive a peak-reducing signal vector $\widehat{\boldsymbol{C}}$ that reduces the PAPR of the original OFDM signal $\boldsymbol{x}$. This can be formulated as a minimax optimization problem of the form

$$\min_{\widehat{\boldsymbol{C}}} \max |\boldsymbol{x} - \widehat{\boldsymbol{Q}}\widehat{\boldsymbol{C}}|. \tag{6}$$

In this problem, the optimization variable is $\widehat{\boldsymbol{C}} \in \mathbb{C}^L$ and the problem parameters are $\boldsymbol{x} \in \mathbb{C}^N$ and $\widehat{\boldsymbol{Q}} \in \mathbb{C}^{N \times L}$. To avoid high computational complexity and slow convergence rate in the optimal minimization of (6), sub-optimal techniques have continued originating.

When solving the problem in (6), all tone reservation techniques face two major drawbacks that need to be mitigated. Firstly, there is an increase in the average transmission power of the PAPR-reduced signal. This then requires one to balance between the PAPR reduction and the power increase in order to conform to the HPA specifications. Secondly, the overall data rate is reduced due to the reserved subcarriers and this requires one to use only a practically small number of reserved subcarriers to efficiently reduce the PAPR.

IV. PROPOSED IRLS-TR METHOD

This paper proposes a robust sub-optimal tone reservation method, which fast approximates the peak-reducing signal in accordance with the frequency allocation constraints imposed by the tone reservation concept described in section III.

A. Algorithm

In order to reduce the PAPR, the signal peak power should be very close to the average power as much as possible. Therefore, the desired peak-reducing signal can analytically be described by the equation

$$d(n) = \begin{cases} \dfrac{x(n)}{|x(n)|}(|x(n)| - \bar{x}), & |x(n)| > \bar{x} \\ 0, & |x(n)| \le \bar{x} \end{cases} \tag{7}$$

$$n = 0,1, \dots, N-1$$

where $\bar{x}$ is the average value of the OFDM signal.

2

However, due to the tone reservation constraints in the frequency-domain, the above desired peak-reducing signal $d = [d(0), d(1), \ldots, d(N-1)]^T$ can only be approximated by the peak-reducing signal c in (5) i.e.

$$\widehat{Q}\widehat{C} \approx d \qquad (8)$$

by minimizing the residual error

$$\epsilon = \widehat{Q}\widehat{C} - d, \qquad (9)$$

where $\epsilon = [\epsilon(0), \epsilon(1), \ldots, \epsilon(N-1)]^T$, by using an appropriate norm as a measure of the error.

Since the interest here, is to design a signal c that approximates all the major high peaks in the desired peak-reducing signal while avoiding regeneration of new ones, a phenomenon normally referred to as peak-regrowth, the best approach is to find a vector $\widehat{C}$ that minimizes the L_∞-norm of the error by solving

$$\min_{\widehat{C}} ||\widehat{Q}\widehat{C} - d||_p^p, \quad p \to \infty. \qquad (10)$$

In practice, it has been found that if the value of $p \geq 10$, the minimization of the L_p-norm closely approximates that of the minimization of the L_∞-norm [16].

However, no analytical method exists to solve (10) for any norm other than the L_2-norm. For this reason, the problem is transformed into a simple weighted least squares (WLS) problem

$$\min_{\widehat{C}} ||W(\widehat{Q}\widehat{C} - d)||_2^2 \qquad (11)$$

where W is a diagonal weighting matrix that emphasizes the minimization of the high peaks. The problem in (11) has the closed form solution [17]

$$\widehat{C} = [\widehat{Q}^* W^T W \widehat{Q}]^{-1} \widehat{Q}^* W^T W d \qquad (12)$$

where, the superscripts T and $*$ denote matrix transpose and conjugate transpose respectively.

In order to make the problem in (11) to be equivalent to (10), the diagonal weights in W are given by

$$w(n) = |\epsilon(n)|^{(p-2)/2}, \quad n = 0, 1, \ldots, N-1 \qquad (13)$$

Therefore, solving the WLS problem in (11) is identical to solving the L_p minimization problem in (10).

Because the weights that give the optimal approximation solution are unknown, the problem in (11) cannot be solved in one step, but rather iteratively using a re-weighted least squares algorithm. This algorithm starts by solving for $\widehat{C}$ in (12) when all the diagonal weights are set to one i.e. $w(n) = 1$, for $n = 0,1,2,\ldots,N-1$. Then, it finds the error in (9), followed by the weights in (13) for use in the next iteration. The new weights are then used to find a new solution. To make the algorithm stable and converge quadratically for all L_p-norms, the actual solution at each iteration is only partially updated in a manner similar to the Newton's method [18] i.e.

$$\widehat{C}_i = q_i \widehat{C}_i^c + (1 - q_i)\widehat{C}_{i-1} \qquad (14)$$

where $q_i = 1/(p_i - 1)$ is the Newton's parameter [19] [20] and $\widehat{C}_i^c$ is the current WLS solution. This process is repeated until the L_p-norm of the error is below a preset threshold.

However, the Newtonian modification may make the algorithm susceptible to poor initial convergence. This is cured by starting with an L_p-norm with $p = 2$ and gradually increasing the value by multiplying it with a convergence parameter of between one and two until the final value of the L_p-norm being used to approximate the L_∞-norm. This modification is similar to the homotopy [21] [22] and has p at the i-th iteration given by

$$p_i = \min(p, \alpha p_{i-1}). \qquad (15)$$

The preceding algorithm is summarized here below.

Algorithm: IRLS-TR
1. Set initial weights to $w(n) = 1$, initial $p = 2$ and the error threshold ε_{th}.
2. Choose reserved subcarrier locations and compute the IFFT submatrix $\widehat{Q}$.
3. Generate the original OFDM time-domain signal x, and calculate the desired peak-reducing signal d.
4. Initialize the iteration counter $i = 0$ and calculate the initial LS solution $\widehat{C}_0 = \widehat{Q}^{-1}d$.
5. Calculate the error vector $\epsilon_i = \widehat{Q}\widehat{C}_i - d$ and the new weights using $w_i(n) = |\epsilon_i(n)|^{(p_i-2)/2}$.
6. Calculate the LS solution $\widehat{C}_i^c = [\widehat{Q}^* W_i^T W_i \widehat{Q}]^{-1} \widehat{Q}^* W_i^T W_i d$.
7. Calculate the new update parameter $q_i = 1/(q_i - 1)$ and update the LS solution to $\widehat{C}_i = \mu q_i \widehat{C}_i^c + (1 - \mu q_i)\widehat{C}_{i-1}$.
8. Calculate the L_p-norm of the error $||\varepsilon_i||_p = (\sum_{n=0}^{N-1}|\epsilon_i(n)|^p)^{1/p}$.
9. Set $i = i + 1$. If $||\varepsilon_i||_p > \varepsilon_{th}$, update the L_p-norm value to $p_i = \min(p, \alpha p_{i-1})$ and go to step 5. Otherwise, transmit $x - \widehat{Q}\widehat{C}_i$ and terminate algorithm.
10. End

B. Convergency

The algorithm quadratically converges in about 10 iterations as illustrated in Fig. 2, which is the typical convergence curve of the L_p-norm of the error at each iteration.
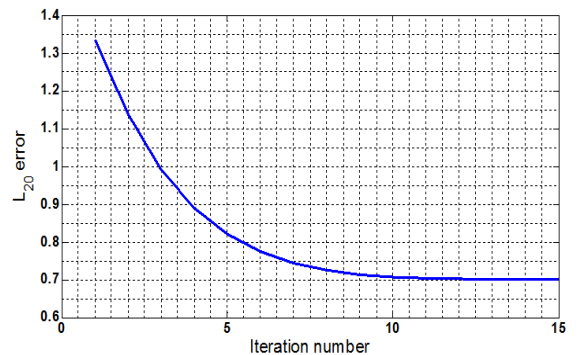
Figure 2: Convergence curve

C. Computational complexity

At each iteration, most of the computational time is on solving the WLS problem in (11), here rewritten as

3

$$W\widehat{Q}\widehat{C} = Wd. \qquad (16)$$

Since $W \in \mathbb{R}^{N \times N}$, $\widehat{Q} \in \mathbb{C}^{N \times L}$, and $\widehat{C} \in \mathbb{C}^{L}$, the computational complexity of the algorithm per iteration is $O(LN)$. This complexity can be further reduced to $O(LP) < O(LN/2)$ if the algorithm computes the WLS solution from only the P nonzero elements in vector d which is always less than $N/2$.

V. RESULTS AND DISCUSSION

In this section, the IRLS-TR method in section IV is applied to reduce PAPR in OFDM systems with different placements of the reserved subcarriers but with a more focus on the fixed or non-random arrangements. Table 1 lists the OFDM system parameters used in the simulations.

Table 1: OFDM Parameters

FFT size, N	64, 256
Subcarrier modulation	16-QAM, 64-QAM
Number of OFDM symbols	10,000
Number of reserved subcarriers, L	3.125%, 5%, 6.25% of N
Reserved subcarrier locations	random, contiguous, equal-spaced

In Fig. 3 and Fig. 4, the simulation results for the 16-QAM and 64-QAM OFDM systems with 5% reserved subcarriers are presented. For the random placement of the reserved tones, which acts as the lower bound for the contiguous and equal-spaced placements, the IRLS-TR method proposed here achieved the same PAPR reductions of 5.30 dB at CCDF of 10^{-3} for the two systems. With the contiguous tone locations at {243, 244,…, 255} and also with equal-spaced locations of the 13 reserved tones, at the same CCDF of 10^{-3}, the proposed method had, for the two systems, the same PAPR reduction of 4.50 dB.

The sub-optimal kernel-based TR (KTR) by Tellado and Cioffi reported 3.25 dB for both the contiguous and equal-spaced tone reservations, which is lower than the 4.5 dB from the proposed IRLS-TR method. In addition, unlike the kernel-based TR methods including the KTR by Tellado and Cioffi [23], the proposed IRLS-TR method does not experience peak-regrowth problems that can lead to slow convergence times.
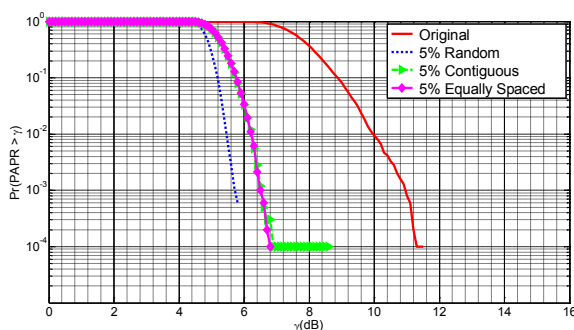
Figure 3: Comparison of different tone positioning for 16-QAM OFDM system with N=256, L=13

Figure 5 illustrates PAPR reductions in a 16-QAM OFDM system in which 3.125 % of tones out of the 256 tones were reserved for PAPR reduction. The set of contiguous locations was {248, 249,…, 255} while for the equal-spaced was {31, 63,

95, 127, 159, 191, 223, 255}. From the figure, the IRLS-TR has PAPR reductions of 2.45 dB and 4.00 dB at CCDF = 10^{-3} for the equally spaced and contiguous tone allocations respectively.
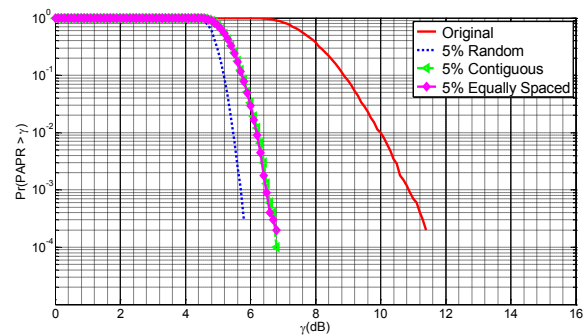
Figure 4: Comparison of different tone positioning for 64-QAM OFDM system with N=256, L=13

These reductions are slightly better than the ones reported by Petersson et al of 2.25 dB and 3.75 dB for the same placements in the complex optimal OTR scheme using the same system parameters.
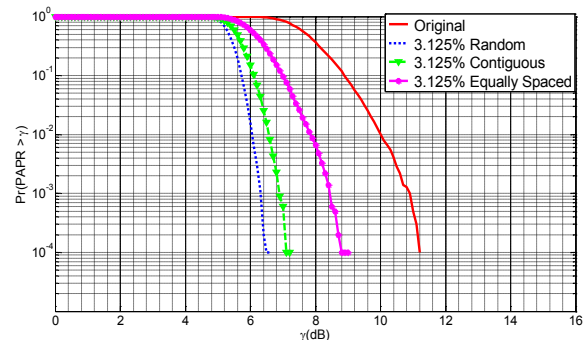
Figure 5: Comparison of different tone positioning for 16-QAM OFDM system with N=256, and L=8

In Fig.6, the simulation results for 16-QAM system with 6.25 % reserved tones out of 64 tones are illustrated. The set of contiguous locations was {60, 61, 62, 63} while that for the equal-spaced was {15, 31, 47, 63}. The proposed IRLS-TR has PAPR reductions of 2.40 and 4.20 dB at CCDF = 10^{-3} for the equal-spaced and contiguous allocations respectively. The contiguous case, which is the better of the two schemes for presetting the reserved tones, has the same PAPR reduction of 4.20 dB at CCDF of 10^{-3} as that reported for the highly complex DTA-OTR method by Malkin et al for the same system parameters. However, one major drawback of the DTA-OTR is that the receiver has to bear the complex burden of deciding which tones were reserved for PAR reduction and on which data was sent; a problem not experienced with the IRLS-TR method proposed here.

On the increase in the average transmitted power, the proposed IRLS-TR method reduces PAPR at the expense of a small increase of about 0.6 dB in the average power. The TR method by Tellado and Cioffi had an increase of 1 dB. Due to the high PAPR reductions achieved by the proposed IRLS-TR method, the increase of 0.6 dB in the transmitted power is expected.

4

In summary therefore, and based on the simulation results, the proposed IRLS-TR method has better and improved PAPR reduction performance than the KTR and OTR methods that have fixed placement of the reserved subcarriers.
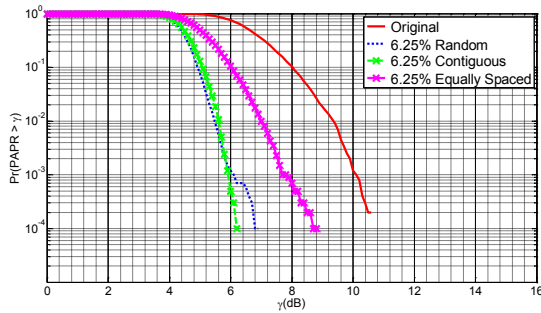
Figure 6: Comparison of different tone positioning for 16-QAM OFDM system with N=64, and L=4

VI. CONCLUSION

In this paper, a new sub-optimal tone reservation technique, IRLS-TR, for PAPR reduction in OFDM systems where the reserved subcarriers are allocated fixed positions in all OFDM symbols has been proposed. Investigation into the PAPR reduction capability found that with 5% of total subcarriers reserved for PAPR reduction, the method reduces PAPR by 4.5 dB at a small cost of 0.6 dB increase in the average transmitted power both equal-spaced and contiguous reserved tone locations respectively.

The method offers fast estimation of the required peak-reducing signal with a quadratic convergence and a low per iteration computational complexity of $O(NL)$ with L as the number of reserved subcarriers and N as the total number of subcarriers.

Further investigation revealed that the method achieves high PAPR reductions for the equal-spaced and contiguous reserved tone locations that are only 0.8 dB lower than the per symbol random reserved tone allocation. This is a good property of the method because for practical transmission systems, fixed allocation of subcarriers is desired. The effect of using only the nonzero elements in the desired peak-reducing signal on the PAPR reduction capabilities of the proposed method shall be investigated in future work.

REFERENCES

[1] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys and Tutorials*, Vol. 15, No. 4, pp 1567-1592, 2013.

[2] F. Sandoval, G. Poitau, and F. Gagnon "On Optimizing the PAPR of OFDM Signals With Coding, Companding, and MIMO," *IEEE Access*, Vol. 7, March 2019.

[3] M. I. Al-Rayif, H. Seleem, A. Ragheb, and S. Alshebeili, "A Novel Iterative-SLM Algorithm for PAPR Reduction in 5G Mobile Fronthaul Architecture," *IEEE Photonics Journal*, Vol. 11, No. 1, Feb. 2019.

[4] A. M. Rateb and M. Labana, "An Optimal Low Complexity PAPR Reduction Technique for Next Generation OFDM Systems," *IEEE Access*, Vol. 7, Feb. 2019.

[5] A. Jawhar, L. Audah, M. Taher, K. Ramli, N. Shah, M. Musa, and M. Ahmed "A Review of Partial Transmit Sequence for PAPR Reduction in the OFDM Systems," *IEEE Access*, Vol. 7, February 2019.

[6] H. Boche, J. Mönich, and E. Tampubolon, "Complete Characterization of the Solvability of PAPR Reduction for OFDM by Tone Reservation," *IEEE International Symposium on Information Theory* (*ISIT*), pp. 2023-2027, June 2017.

[7] S. Kiambi, E. Mwangi, and G. Kamucha, "Effect of OFDM Signal Structure and Subcarrier Modulation on the Reduction of the Signal Peak Power," *IEEE Africon 2017 Proceedings*, pp. 262–266, Sept. 2017.

[8] T. Jiang, C. Ni, C. Xu, and Q. Qi, "Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems," *IEEE Comm. Letters*, vol. 18, no. 5, pp. 805–808, May 2014.

[9] J. Wang, Xin Lv, and W. Wu, "SCR-Based Tone Reservation Schemes With Fast Convergence for PAPR Reduction in OFDM System," *IEEE wireless Commun. Letters*, Vol. 8 , Issue 2, pp. 624 – 627, Apr. 2019.

[10] L. Xin and W. Yi, "A New Weighted Tone Reservation Method for PAPR Reduction in OFDM Systems," *Journal of Communication*, Vol. 9, No. 12, pp. 980-986, Dec. 2014.

[11] J. Tellado and J. M. Cioffi, "PAR reduction in multicarrier transmission systems," ANSI Document, T1E1.4 No. 97-367, Dec. 1997.

[12] M. Malkin, T. Magesacher, and J. Cioffi, "Dynamic allocation of reserved tones for PAR reduction," *In Proc. 13th Internationa. OFDM-Workshop In OWo'08*, pp. 231–235, Hamburg, Germany, Aug. 2008.

[13] N. Petersson, A. Johansson, P. Ödling, and P. Börjesson, "Analysis of tone selection for PAR reduction," *In Proc. International Conference on Information, communications and Signal Processing*, Singapore, Oct. 2001.

[14] E. Dahlman, S. Parkvall, and J. Sköld, P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband (first edition),* London: Elsevier Ltd, 2007.

[15] Y. Cho, J. Kim, W. Yang, and C. Kang, *MIMO-OFDM Wireless Communication with MATLAB*, John Wiley & Sons (Asia) Pte Ltd, Singapore, 2010.

[16] R. H. Byrd and D. A. Pyne, "Convergence of the iteratively re-weighted least squares algorithm for robust regression,"Technical report 313, Dept. of Math. Sci, Johns Hopkins University, Baltimore, Maryland, June 1979.

[17] Adi Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications,* Springer-Verlag, New York, 2nd ed, 2003.

[18] R. L. Burden and J. D. Faires, *Numerical Analysis*, Brooks/Cole, Cengage Learning, Boston, USA, 9th ed., 2011.

[19] S. W. Kahng, "Best L_p approximation," *Mathematics of Computation*, vol. 26, No. 118, pp 505-508, Apr. 1972.

[20] R. Fletcher, J. A. Grant, and M. D. Hebden, "The calculation of linear best L_p approximations," *Computer Journal*, Vol.14, pp. 27-279, 1971.

[21] V. L. Stonick and S. T. Alexander, "Globally optimal rational approximation using homotopy continuation methods," *IEEE Transactions on Signal Processing*, Vol.40, No.9, pp 2358-2361, Sept. 1992.

[22] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer series in operations research. New York, NY: Springer-Verlag, 1999.

[23] J. Song and H. Ochiai, "Performance analysis for OFDM signals with peak cancellation," *IEEE Transactions on Communications,* Vol. 64, No. 1, pp. 261–270, Jan. 2016.

Appendix E: Fourth Publication

A Low-Complexity Signal Addition Method for PAPR Reduction in OFDM Systems

Stephen Kiambi[#1], Elijah Mwangi[#2], George Kamucha[#3]

#1 School of Engineering, University of Nairobi, Nairobi, Kenya,
email: skiambi@uonbi.ac.ke
#2 School of Engineering, University of Nairobi, Nairobi, Kenya,
email: elijah.mwangi@uonbi.ac.ke
#3 School of Engineering, University of Nairobi, Nairobi, Kenya,
email: gkamucha@uonbi.ac.ke

ABSTRACT

OFDM remains the transmission technique of choice for communication systems requiring high-speed transmission of data. However, signals generated via OFDM can exhibit high PAPR; a problem that still needs a practical solution. This paper proposes a fast-converging and low-complexity signal addition method for reducing the PAPR. The resources for transmitting the peak-reducing signal are reserved in the time domain rather than in the frequency domain as is done in tone reservation methods. The method has computational complexity of $O(M)$, where M is the number of nonzero samples in the peak-reducing signal. Therefore, the overall computational complexity of the transmitter is barely increased because $O(M)$ is far much less than $O(N \log_2 N)$ of the IFFT in the transmitter. The proposed method can achieve high PAPR reductions with only a small number of nonzero samples in the peak-reducing signal without affecting the average power of the transmitted signal and at a small data rate loss. For example, with $M = 5\%$ of 256-point IFFT samples, a PAPR reduction of 5.64 dB can be achieved with data rate loss of 4.8%.

Key words: High Power Amplifier (HPA), Orthogonal Frequency Division Multiplexing (OFDM), Peak-to-Average Power Ratio (PAPR), Tone Reservation (TR).

Corresponding Author: Stephen Kiambi

1. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier transmission technique that employs parallel subcarriers to achieve high data rate transmission. The subcarriers are mutually orthogonal and the transmission system has high spectral efficiency. In addition, the system has symbol duration and a guard interval that are much longer than maximum multipath delay in order to eliminate both inter-symbol and inter-carrier interferences, and therefore allow use of a simple single-tap equalizer at the receiver to recover signals. These advantages, among others, have made OFDM the established transmission technique in high data rate applications such as digital audio broadcasting (DAB), digital video broadcasting (DVB), IEEE 802.11 WLAN and IEEE 802.16a WMAN standards, 4th generation and 5th generation mobile communication networks [1].

However, owing to the summation of many modulated subcarrier signals, an OFDM signal may suffer an unacceptably high peak-to-average power ratio (PAPR) especially when the number

21

of subcarriers is large. Distortionless processing of high PAPR signals requires the high power amplifier (HPA) in the transmitter to be input backed-off away from the 1-dB compression point to avoid nonlinear amplification effects, mainly; bit-error rate (BER) degradation and out-of-band radiations.

However, an input backed-off HPA has lower power efficiency and thus high power consumption. This has the negative effect of increasing the cost of the device and reducing the lifetime of battery power at user terminals [2]. A better solution then is to reduce the PAPR to suitable levels before processing of the OFDM signal in the HPA.

Recently, different methods for reducing PAPR have been proposed in literature. These include signal coding [3], clipping and companding [4], selective mapping [5], partial transmit sequence [6] and tone reservation [7], [8]. The aim has always been to develop practically realisable methods with high PAPR reductions and low computational complexities. The tone reservation (TR) method is the most promising signal addition method because it does not distort the user data and hence has minimum BER degradation. Additionally, it does not require transmission of any side information for the demodulation process at the receiver.

Several tone reservation methods have been proposed. In [9], a proposed tone reservation (CF-TR) method designs peak cancelling signals by curve-fitting them to clipping noise signals. However, the method has to evaluate the highly computational intensive Moore-Penrose generalised matrix inversion during each iteration, and the resulting peak-reduced signal has an increased average power. A least-squares approximation TR (LSA-TR) method was proposed in [10]. Although the method has fast convergence, the PAPR reduction performance is poor, and has high computational complexity and leads to increased average transmit power. In [11] a TR method based on artificial neural networks and initial value optimisation (IVO-TR) is proposed to reduce runtime computational complexity by pre-generating peak-cancelling signals. At runtime, the method classifies an OFDM symbol and then searches for an appropriate peak-cancelling signal stored in the pre-work table. However, the method requires a very long pre-work time in order to generate near optimal peak-cancelling signals, and still has high runtime complexity, and results to increased average transmit power.

In this paper, we propose a low-complexity signal addition method with fast convergence. The key idea is to use the desired peak-reducing signal, which is the signal above a predetermined clipping level in an OFDM signal. The method avoids reservation of tones and the high complexity of determining the dummy weights in tone reservation methods by extending the peak-reduced signal by a few samples of the peak-reducing signal. Compared with CF-TR, LSA-TR, and IVO-TR, the method achieves better PAPR reductions while minimally affecting the average power of OFDM signals. Due to the simplicity, reliability, and fast convergence of the method, it can easily be implemented in the OFDM transmitter section.

The rest of the paper is organized as follows: Section 2 defines PAPR and its measurement. Section 3 outlines the general concept of the tone reservation technique. In Section 4, the proposed method is presented. Section 5 provides simulation results and analysis, while conclusions are in Section 6.

2. PAPR OF OFDM

An OFDM signal is a sum of N modulated subcarrier signals. At baseband level, over one symbol duration T, the continuous-time signal has the analytical equation:

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) \, e^{j2\pi f_k t}. \tag{1}$$

Here, $X(k)$ are the modulation symbols obtained from binary phase-shift keying or M-ary quadrature amplitude modulation. The subcarrier frequency f_k is the product of the index k and

22

the subcarrier spacing. The division by $\sqrt{N}$ ensures that the average power in the time and frequency domains are the same. Because the above formulation is quite similar to the standard inverse discrete Fourier transform (IDFT), the well-known fast Fourier transform (FFT) algorithm is used to implement OFDM signal processing.

Due to the formation in (1), an OFDM signal can have high envelope fluctuations that can lead to nonlinear processing in the HPA. The measurement of the extent of the high power fluctuations is given by the probability by which the maximum instantaneous power is above the linear range of the HPA. For this measurement, the PAPR of the signal is needed and is given by:

$$\text{PAPR}\{x(t)\} = \frac{\max\{|x(t)|^2\}}{\mathrm{E}\{|x(t)|^2\}} \tag{2}$$

where, $\mathrm{E}\{\cdot\}$ is the expectation operator.

Assuming a sufficiently large number of subcarriers and that the subcarrier signals are statistically independent, then by the central limit theorem, both the real and imaginary parts of $x(t)$ are Gaussian distributed. Consequently, the instantaneous magnitudes of $x(t)$ are Rayleigh distributed. This in turn means that $x(t)$ can have high PAPR or high amplitude well above the average value.

Because of the digital signal processing of OFDM signals, the continuous-time PAPR is usually estimated in the discrete-time as:

$$\text{PAPR}\{x\} = \frac{\max\limits_{0 \leq n \leq N-1}\{|x(n)|^2\}}{E\{|x(n)|^2\}}. \tag{3}$$

where, the discrete-time signal $x = [x(0), x(1), \ldots, x(N-1)]^T$. For a close estimation of the PAPR, the sampling rate applied on $x(t)$ must be sufficiently higher than the Nyquist rate, typically by a factor ≥ 4, in order to avoid skipping the peak value [12].

The level of PAPR is measured by the complementary cumulative distribution function (CCDF), which by definition is the probability that the PAPR is above a given threshold [13]

$$\Pr\{\text{PAPR}\{x\} > \gamma\} = 1 - (1 - e^{-\gamma})^N. \tag{4}$$

Here, γ is a threshold value, N is the total number of subcarriers, and $\Pr\{.\}$ is the probability operator.

From the CCDF formula, if the CCDF value on the left is fixed, then for a fixed N, a higher value of γ indicates a higher PAPR and vice versa. This implies that the difference between two values of γ at the same CCDF value can be used to measure the level of PAPR reduction and to judge how well a proposed method reduces PAPR.

3. CONCEPT OF TONE RESERVATION

It is the concept of reserving a smaller number of subcarriers, previously meant for carrying user data to carry a peak-reducing signal for a given OFDM signal. The reserved subcarriers are generally referred to as peak reduction tones (PRTs). Because of the reservation, the number of subcarriers for transmitting user data, and hence the data rate, is reduced. Both the peak-reducing signal and the OFDM signal are combined to give a low PAPR transmit signal. The tone reservation concept is illustrated in Fig. 1, where $X(k)$ and $C(k)$ are the modulation data symbols and peak-reducing weights, respectively. At the receiver, only the knowledge of the locations of the data-bearing subcarriers is necessary for the recovery of the user data.
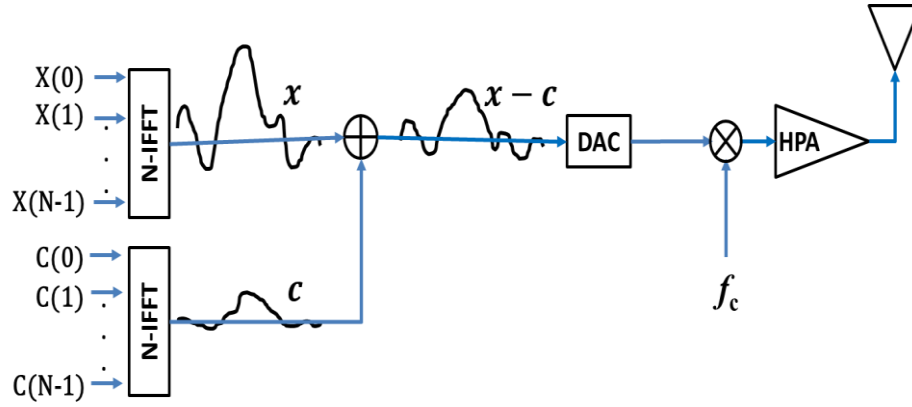
Fig. 1. Tone reservation concept

In order to avoid any distortion on the user data, the data bearing vector and peak-reducing weights must exist in two mutually exclusive frequency subspaces i.e. the data vector $X = [X(0), X(1), \ldots, X(N-1)]^T$ should have all nonzero entries except in the reserved positions and likewise the peak-reducing vector $C = [C(0), C(1), \ldots, C(N-1)]^T$ should have all zero entries except in the reserved positions. When the IDFT is applied to the two vectors, X and C, it gives the discrete-time signals $x(n)$ and $c(n)$, which are, respectively, represented by $x = [x(0), x(1), \ldots, x(N-1)]^T$ and $c = [c(0), c(1), \ldots, c(N-1)]^T$. These two time-domain signals are added to give the transmit signal $s(n)$ as follows:

$$s(n) = x(n) - \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} C(k)e^{j2\pi kn/N}, \qquad n = 0,1, \ldots, N-1. \tag{5}$$

Equation (5) can be given in the matrix notation as:

$$s = x - QC \tag{6}$$

where, $Q \in \mathbb{C}^{N \times N}$ is the IDFT matrix with elements $(1/\sqrt{N})exp(j2\pi kn/N)$, and all the vectors $x, s, C \in \mathbb{C}^N$. If L is the number of reserved subcarriers, then vector C has L nonzero elements, and the peak-reducing signal $c = QC$, can be rewritten in the form:

$$c = \widehat{Q}\widehat{C} \tag{7}$$

where the new peak-reducing vector $\widehat{C} \in \mathbb{C}^L$ contains only the nonzero elements of C, and the submatrix $\widehat{Q} \in \mathbb{C}^{N \times L}$ is made of only the L columns of Q corresponding to the locations of reserved subcarriers.

The generation of the required peak-reducing signal requires finding the elements $\widehat{C}$ that minimises the PAPR of the combined signal so that PAPR$\{s\} <$ PAPR$\{x\}$. To identify that there is reduction of the peak power in the peak-reduced signal compared to the original signal's average power, PAPR$\{s\}$ is defined as follows:

$$PAPR\{s\} = \frac{\max_{0 \le n \le N-1}\{|x(n) + c(n)|^2\}}{E\{|x(n)|^2\}}. \tag{8}$$

The problem of finding $\widehat{C}$ so that PAPR$\{s\} <$ PAPR$\{x\}$ can be presented in the form of a minimax problem [14]

$$\min_{\widehat{C}} \max|x - \widehat{Q}\widehat{C}| \tag{9}$$

24

where $\widehat{C}$ is the optimization variable, and x and $\widehat{Q}$ are the input parameters.

Solving the optimization problem in (9) is computationally intensive and can make implementation of the TR concept impractical especially in real-time systems. In addition, a system applying the concept may experience two major drawbacks. The first one is the increase in the average transmit power of the PAPR-reduced signal, and the second is the reduction in the data rate due to the use of PRTs that do not carry user data.

4. PROPOSED METHOD

This paper proposes a suboptimal PAPR reduction method that is derived from the general concept of tone reservations especially on the part of signal addition. The proposed method is referred to as "a low-complexity signal addition" method; in short LCSA method. Depending on the structure of the OFDM signal, and possibly the maximum allowed data rate loss, the method generates a desired peak-reducing signal (PRS), which it employs to reduce the PAPR. The reduction of PAPR is achieved by cancellation of all the highest peaks in the original OFDM signal. For exact cancellation of the highest peaks without introducing new ones, the ideal PRS should be composed of all the peaks above a threshold value called the clipping level. In addition, the PRS should have a phase spectrum identical to that of the OFDM signal.

Therefore, for a general complex discrete-time OFDM signal $x(n)$, the ideal PRS can be represented as a vector $\boldsymbol{d} = [d(0), d(1), \dots, d(N-1)]^T$ whose elements are given by

$$d(n) = \begin{cases} \dfrac{x(n)}{|x(n)|}(|x(n)| - x_{CL}), & |x(n)| > x_{CL} \\ 0 & , & |x(n)| \le x_{CL} \end{cases}, \qquad n = 0,1,\dots,N-1 \qquad (10)$$

where x_{CL} is the clipping level.

The aim of the proposed method is to create a PRS that matches the desired one in (10). However, to lower the computational complexity during the process of generating the PRS, the proposed method uses a simplified version of (10) that contains only the nonzero entries i.e.

$$\widehat{\boldsymbol{d}} = \left[\hat{d}(0), \hat{d}(1), \dots, \hat{d}(M-1)\right]^T \qquad (11)$$

where M is the number of nonzero entries in $\boldsymbol{d}$.

4.1 LCSA Algorithm

From (7), and considering only the nonzero elements in $\boldsymbol{d}$, the design equation for the peak-reducing signal is simply

$$\hat{c} = \widetilde{Q}\widehat{C} \qquad (12)$$

where $\hat{c} \in \mathbb{C}^M$, $\widehat{C} \in \mathbb{C}^L$ and the new IFFT submatrix $\widetilde{Q} \in \mathbb{C}^{M \times L}$ has only the rows from $\widehat{Q} \in \mathbb{C}^{N \times L}$ corresponding to the M indices of nonzero entries in $\boldsymbol{d}$. With $\hat{c}$ calculated, the actual peak-reducing signal $c \in \mathbb{C}^N$ is formed with all zero components except in M positions whose values are those of vector $\hat{c}$.

The task of designing a peak-reducing signal will then involve finding the elements of $\widehat{C}$ by solving the system of linear simultaneous equations given by:

$$\widetilde{Q}\widehat{C} = \widehat{\boldsymbol{d}} \qquad (13)$$

The solution to this system of equations is strongly dependent on the dimensions of the IFFT submatrix. Since $\widetilde{Q}$ has M orthogonal rows and L orthogonal columns, the IFFT submatrix can have either a full row or column rank and thus there are two possible solutions. For both the square and underdetermined cases i.e. for $M \le L$, the design will give an exact solution in the

25

time domain i.e. $\hat{c} = \hat{d}$, while for the overdetermined case of $M > L$, $\hat{c}$ will only approximate $\hat{d}$. The solution to the system of equations is therefore strongly dependent on the value of M for a fixed value of L. The value of M depends on the clipping that is applied to the OFDM signal.

For a given clipping level, M can analytically be found from the distribution of the signal amplitudes. For this purpose, let X_n denote a Rayleigh random variable representing the signal amplitudes whose cumulative distribution function is given by

$$F_{X_n}(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0 \tag{14}$$

where σ is the scale parameter of the distribution. This distribution has a mean value given by the following equation:

$$\overline{x} = \sigma\sqrt{\frac{\pi}{2}}. \tag{15}$$

The clipping level can be defined as a function of the mean using

$$x_{CL} = \lambda\overline{x} \tag{16}$$

where the clipping level parameter, λ, can have any value from zero to the maximum value of $\max(|x(n)|)/\overline{x}$.

From (14), the probability that some signal amplitudes are higher than the clipping level is given by:

$$P(X_n > x_{CL}) = e^{-\frac{x_{CL}^2}{2\sigma^2}} \tag{17}$$

and since this is equal to M/N, then substituting (16) into (17) gives the number of nonzero entries in d as:

$$M = Ne^{-\frac{\pi}{4}\lambda^2} \tag{18}$$

which is simply the number of peaks above the clipping level.

For the purposes of PAPR reduction, the clipping level should logically be set equal to or greater than the average value of the OFDM signal. From (18), the number of nonzero entries decreases exponentially with the clipping level parameter. The maximum value of M occurs when the clipping level is equal to the average value which corresponds to $\lambda = 1$, and $M = 0.4559N$ or approximately 46% of the total subcarriers. This value of M is obviously much greater than the expected number of reserved subcarriers in any tone reservation method.

When M is greater than L, the designed peak-reducing signal $\hat{c} = \tilde{Q}\hat{c}$ can only approximate the desired peak-reducing signal but can never be equal to it. The difference between the two signals is the residual error given by the equation:

$$\hat{\epsilon} = \tilde{Q}\hat{c} - \hat{d} \tag{19}$$

where the error vector $\hat{\epsilon} = [\hat{\epsilon}(0), \hat{\epsilon}(1), \dots, \hat{\epsilon}(M-1)]$ is composed of complex elements.

The main task of finding the elements of $\hat{c}$ can be done through least-squares minimisation [15], [16], [17]. The minimisation can yield a discrete-time signal that approximates or is equal to the desired peak-reducing signal. In the case $\hat{c}$ is not equal to $\hat{d}$, a good reduction of PAPR may not be achieved and there is also the possibility of having the average power of the PAPR-reduced signal increased. These two issues can be avoided by making the submatrix $\tilde{Q}$ square i.e. to have $L = M$, and therefore obtain a unique optimal peak-reducing signal with the elements of $\hat{c}$ given by

$$\hat{c} = \tilde{Q}^{-1}\hat{d}. \tag{20}$$

26

There are only two possibilities of having $L = M$. One of the options is to fix M and allow sufficient increase on the number of the reserved subcarriers but this has the undesirable effect of reducing the data rate of the system. The second, and the better, option is to fix L and adjust M in (18) using the clipping level parameter. For the different values of M, the optimal clipping level will be theoretically given by the equation:

$$x_{CL} = \overline{x} \sqrt{\frac{4}{\pi} \ln \left(\frac{N}{M} \right)}. \tag{21}$$

However, due to the random nature of the user data and the fixed number of reserved subcarriers, the theoretical clipping level in (21) may not always yield a value of M that is equal to L for all OFDM symbols. In addition, it is important to note that if the clipping level is increased to very high values, M will be less than L and thus making the system $\widetilde{Q}\widehat{C} = \widehat{d}$ underdetermined. Although for this case there are many exact peak-reducing signal solutions, only small PAPR reductions can be achieved. This is because the generated peak-reducing signal will have some few low amplitudes corresponding to a small number of the highest peaks in the OFDM signal.

Based on the above arguments, and depending on the target PAPR reduction, the clipping level can simply be set equal to one of the highest peak of the OFDM signal. This will ensure that we always have a square system of linear equations and therefore an optimal peak-reducing signal solution. For this setting, the peak-reduced signal is given by:

$$\begin{aligned} s &= x - c \\ &= x - d \end{aligned} \tag{22}$$

Because $c = d$, all the high peaks will be cancelled without generating new ones, as found in methods that generate kernel-like peak-cancelling signals [18].

After fixing the clipping level, the system can also be allowed to utilise all the designated data subcarriers for user data transmission without reserving any for PAPR reduction. This is possible if the M nonzero PRS samples are appended to the peak-reduced signal, just before the addition of cyclic prefix, for the recovery of the original OFDM signal at the receiver. The main blocks of the transmitter section of an OFDM system incorporating the PAPR reduction block are illustrated in Fig. 2. In the figure, X represents the modulation symbols while x, c, and s are the discrete-time OFDM signal, peak-reducing signal, and peak-reduced signal, respectively. The CP block represents the process of copying some samples from the tail end of the OFDM symbol to the front end for the creation of a guard interval.

The following is the summary of the proposed LCSA algorithm.

LCSA Algorithm

i)   Set total number of subcarriers N, data rate loss $R_{loss,t}$, and maximum allowed PAPR$_{max}$
ii)   Generate OFDM signal x, and calculate PAPR
iii) If PAPR< PAPR$_{max}$, transmit x and terminate the program, else got to step (iv)
iv)   Set clipping level x_{CL}
v)   Generate desired peak-reducing signal d
vi)   Calculate actual peak-reducing signal c
vii) Generate peak-reducing samples $\widehat{c}$
viii) Generate peak-reduced signal $s = x - c$
ix)   Append $\widehat{c}$ to s and pass the composite signal $\widehat{s}$ for onward processing
x)   End

At the receiver, after removing the cyclic prefix, the M PRS samples are removed and added back to the peak-reduced signal to reconstruct the original OFDM time signal, which is then passed to the FFT block for demodulation.

27
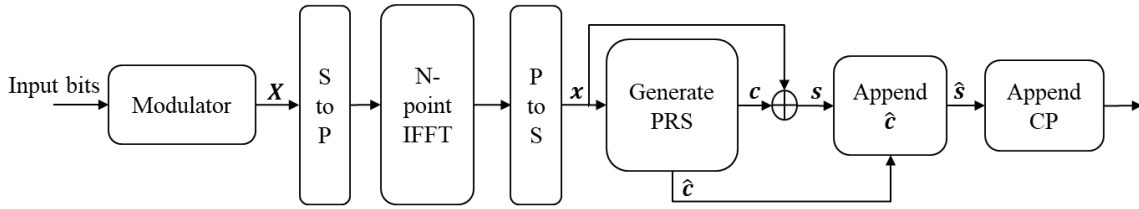
Fig. 2. Transmitter with PAPR reduction block

Considering that in a typical implementation of the algorithm, M will be much less than N, the waveform of the peak-reduced signal will be barely affected by the addition of M samples. A typical composite signal for an OFDM system with $N=256$ and 5% data rate loss is illustrated in Fig. 3 where the peak-reduced signal part is in blue and the peak-reducing signal part in red. As can be seen, the wave pattern generally remains that of the peak-reduced signal.

In the composite signal, all the user data is carried in the peak-reduced signal part. Therefore, time-domain extension by M samples will lead to a reduction in data rate by a factor of:

$$R_{loss,t} = \frac{M}{N + M}. \tag{23}$$

The data rate loss in (23) is less than the one in frequency-domain implemented tone reservation methods, which due to the reserved subcarriers is given by:

$$R_{loss,f} = \frac{M}{N}. \tag{24}$$

In addition to the loss in (23), there will be a small reduction in the average power that also increases with the number of nonzero PRS samples. This can be explained by comparing the power in the clipped samples and the corresponding nonzero PRS samples to the power in the original samples. If we let i be the index of a clipped sample, and $\hat{c}(i)$ and $x(i)$ be, respectively, the corresponding PRS and original signal samples, then the total power in the clipped and PRS samples will be less than or equal to the power in the original samples i.e.

$$\sum_{i=1}^{M} (x_{CL}^2 + \hat{c}^2(i)) \le \sum_{i=1}^{M} x^2(i) \tag{25}$$

Equation (25) implies that if M increases, the average power reduction increases. However, since the sample amplitudes are between zero and one, the power reductions will be very small. Therefore, the power of the original OFDM signal will practically be retained.
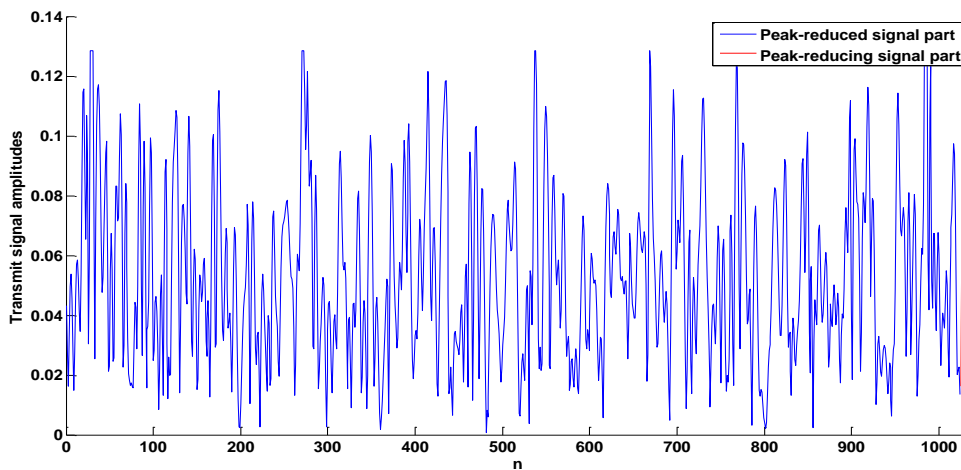
Fig. 3. Composite transmit signal

28

4.2 Computational Complexity

The proposed LCSA algorithm has a very fast convergence rate because computing the required peak-reducing signal from the nonzero entries of the desired peak-reducing signal is done in a single step.

In general, the proposed algorithm has two main operations, which are the generation of peak-reducing signal using (10) and PAPR reduction through the addition of the OFDM signal and peak-reducing signal using (22). The generation of peak-reducing signal requires $2M$ real multiplications and M real additions while the operation of reducing PAPR requires $2M$ real additions.

Therefore, the computational complexity of the proposed algorithm $O(M)$. But because $M \ll N$, the proposed algorithm will add insignificant computational complexity in OFDM systems in which it is deployed to reduce PAPR.

For comparison to IVO-TR, CF-TR and LSA-TR methods, their computational complexities are provided in Table 1. The IVO-TR has many pre-work computations during iteration process, training process and clustering. However, the iteration process, which uses the slow converging CC-TR [18] to generate peak reduction tones for training the self-organizing map of artificial neural networks, dominates the complexity. The algorithm's main parameters are the number of subcarriers N, number of samples N_s, number of iterations I_{ivo}, number of neurons in the input layer N_1 and number of neurons in the hidden layer N_2. The IVO-TR requires a long pre-work time in order to generate peak-reducing tones that are close to the ones generated by CC-TR. At runtime, the algorithm needs a process to recognise OFDM signal and table search for the right peak-cancelling signal.

For the CF-TR and LSA-TR, the number of iterations are, respectively, denoted by I_{cf} and I_{ls}. These two algorithms start by calculating the clipping noise, which is the desired peak-cancelling signal. The LSA-TR then applies least square approximation on the noise to find the actual peak-cancelling signal. On the other hand, the CF-TR applies a generalised matrix inverse approximation on the noise. The two algorithms must in addition perform IFFT/FFT and other computations in each iteration and therefore, have a complexity that increases with the number of iterations.

In order to compare the runtime complexities, which is the most important for real-time systems, we first note that both N_1 and N_2 for the IVO-TR are typically greater than 100. From the tabulated results, it can be observed that LCSA has in general a lower computational complexity and therefore, is more computationally efficient than the IVO-TR, CF-TR, and LSA-TR algorithms.

Table 1. Complexity analysis

Algorithm	Pre-work time complexity	Pre-work space complexity	Runtime complexity	Runtime space complexity
LCSA	None	None	$O(M)$	$O(N)$
IVO-TR	$N_s \times I_{ivo} \times O(N \log_2 N)$	$N_s \times O(N)$	$N_2 \times O(N)$	$(N_1 + N_2) \times O(N)$
CF-TR	None	None	$I_{cf} \times O(N \log_2 N)$	$O(N)$
LSA-TR	None	None	$2I_{ls} \times O(N \log_2 N)$	$O(N)$

29

5. RESULTS AND DISCUSSION

The proposed LCSA method was applied to different OFDM systems, which were simulated in MATLAB. Systems under consideration, whose main parameters are listed in Table 2, were chosen for the purposes of ascertaining the method's performance and comparison with other promising PAPR reduction methods.

Table 2. Simulation inputs

FFT size N	64, 256
Subcarrier modulation	QPSK
Number of OFDM symbols	10^5
Oversampling factor F_s	4
Data rate loss $R_{loss,t}$ in %	diverse
Power amplifier model	Rapp model, $p = 2$, $IBO = 7$ dB

As indicated in the table, during the simulations an oversampling factor, F_s=4, was applied to the discrete-time OFDM signal in order to closely estimate the continuous-time PAPR. All subcarriers were QPSK-modulated and this is sufficient for checking performance and comparing the proposed method with other methods because subcarrier modulation does not affect PAPR reductions. Both PAPR reduction and BER performances were analysed.

From (22), the generated PRS should match the desired signal and this is illustrated in Fig. 4. Without loss of generality, an OFDM system with 256 QPSK-modulated subcarriers was used to ascertain the PAPR reduction capability of the method. For this purpose, different low data rate losses, which are $R_{loss,t} = 1.2\%, 2.3\%, 4.8\%, 6.9\%$, and 9.2% corresponding to $M = 3, 6, 13, 19$, and 26 nonzero PRS samples, respectively, were considered. The CCDF curves for the different values of M are as shown in Fig. 5, and the PAPR reductions at CCDF $= 10^{-3}$ are tabulated in Table 3.

From the results in Table 3, it is evident that although the proposed LCSA method has a PAPR reduction capability that increases with the number of nonzero PRS samples, with only a few samples, high PAPR reductions can be achieved. Furthermore, the results of the average power increase show that the method approximately retains the average power of the original signal. For example with $M = 26$, which seems quite high, the new average power is 95.3% of the original value before PAPR reduction.
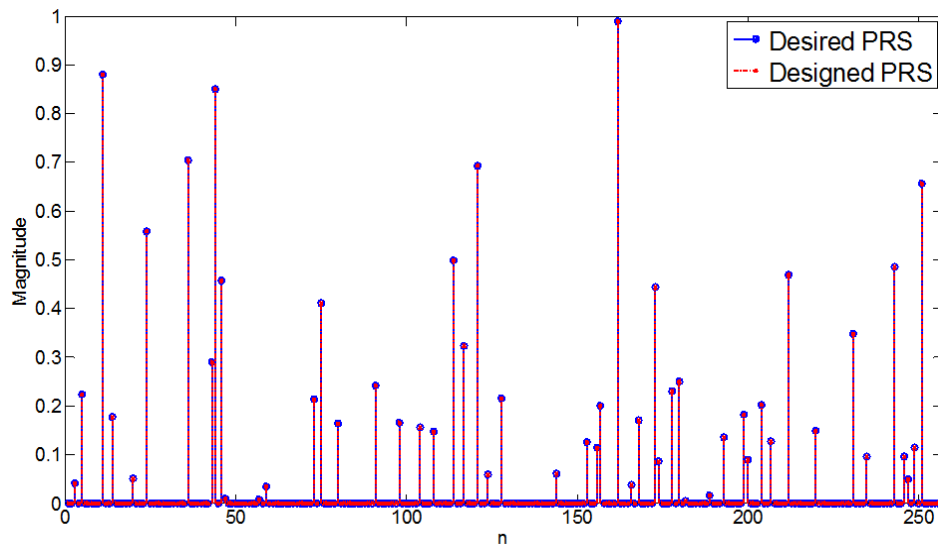
Fig. 4. Desired and designed peak-reducing signals

30
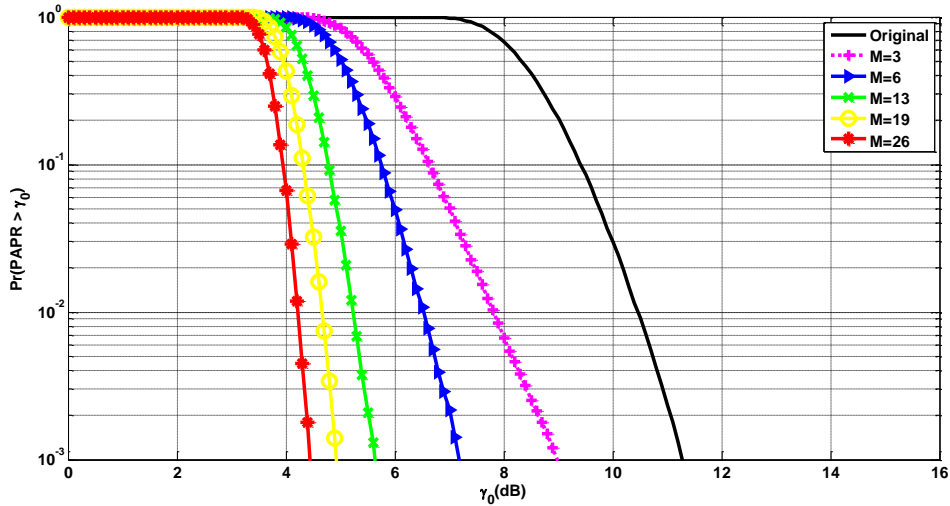
Fig. 5. CCDFs for QPSK-modulated OFDM system with $N = 256$

Table 3. PAPR reductions for different number of PRS samples

Number of PRS samples M	3	6	13	19	26
Data loss rate $R_{loss,t}$	1.2%	2.3%	4.8%	6.9%	9.2%
PAPR Reduction in dB at CCDF $=10^{-3}$	2.28	4.10	5.64	6.35	6.83
Average power increase in dB	-0.02	-0.05	-0.10	-0.15	-0.21

After evaluating the accuracy of the proposed LCSA method, its performance was compared to CF-TR, IVO-TR, and LSA-TR methods. An OFDM system with $N=64$ subcarriers all QPSK-modulated, and of which $M=4$ are reserved, was used for evaluating the performance of each of the four systems. The simulation parameters for IVO-TR were $N_1=100$ neurons for the input layer, $N_2=200$ neurons for the hidden layer, and 10^6 pre-work OFDM signals. The number of iterations was 3 for both the LSA-TR and CF-TR.

The CCDFs for the different methods are shown in Fig. 6 and the PAPR reductions at CCDF$= 10^{-3}$ are presented in Table 4. As can be derived from the table, the proposed method has a PAPR reduction that is 0.6 dB, 1.0 dB and 1.9 dB higher than that of IVO-TR, CF-TR, and LSA-TR, respectively.
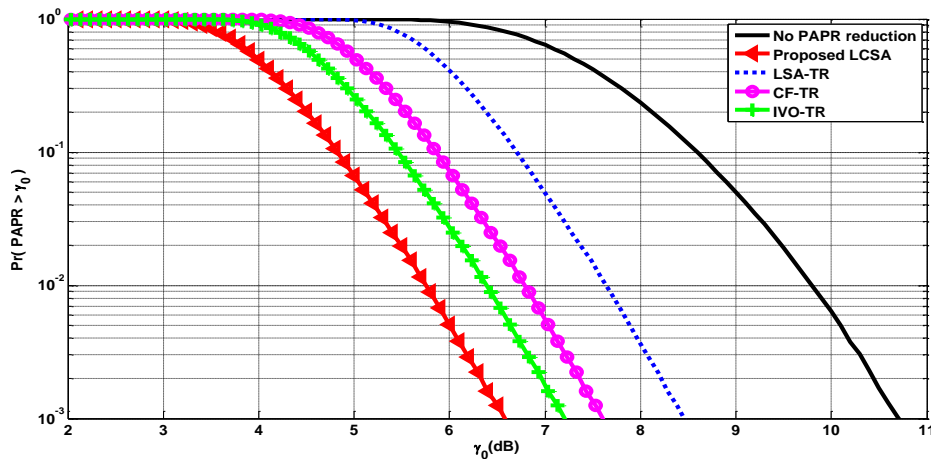
Fig. 6. CCDFs for QPSK-modulated OFDM system with N = 64 and M = 4

31

Table 4. PAPR reduction, average power increase, and data rate losses

Algorithm	PAPR reduction in dB	Average power increase in dB	Data rate loss
LCSA	4.1	-0.12	5.88%
IVO-TR	3.5	0.57	6.25%
CF-TR	3.1	0.84	6.25%
LSA-TR	2.2	0.30	6.25%

Comparing the average power of the PAPR-reduced signal by the different methods, the results show that the proposed method is more power efficient than the rest because it has a power increase that is 0.42, 0.69, and 0.96 dB below that of the LSA-TR, IVO-TR, and CF-TR, respectively. On the data rate loss, the proposed method has the lowest loss of the four methods

Fig. 7 illustrates the simulated BER performances of the different methods. The Rapp model of the solid-state power amplifier (SSPA) was used with the smoothness parameter $p = 2$. The input power back-off (IBO) was set to 7 dB, which is slightly above the proposed method's PAPR of 6.56 dB at CCDF $= 10^{-3}$, in order to ensure that the probability of the clipped OFDM symbols was less than 1%. After amplification through the HPA, the peak-reduced signal was sent over an additive white Gaussian noise (AWGN) channel.

In Fig. 7, the curve labelled "Theoretical" is the lower limit as it is the BER performance given by BER formula for the QPSK. The curve labelled "Without PAPR Red." is the worst case upper limit as it gives the BER performance when the OFDM signal is passed through the HPA without first reducing the PAPR. The required bit energy to noise power spectral density, E_b/N_0, at BER $= 10^{-5}$ of each method is presented in Table 5. Based on the tabulated results, the proposed algorithm has the lowest BER degradation because the required E_b/N_0 is 0.02, 0.05, and 0.07 dB below that of the IVO-TR, CF-TR, and LSA-TR, respectively.
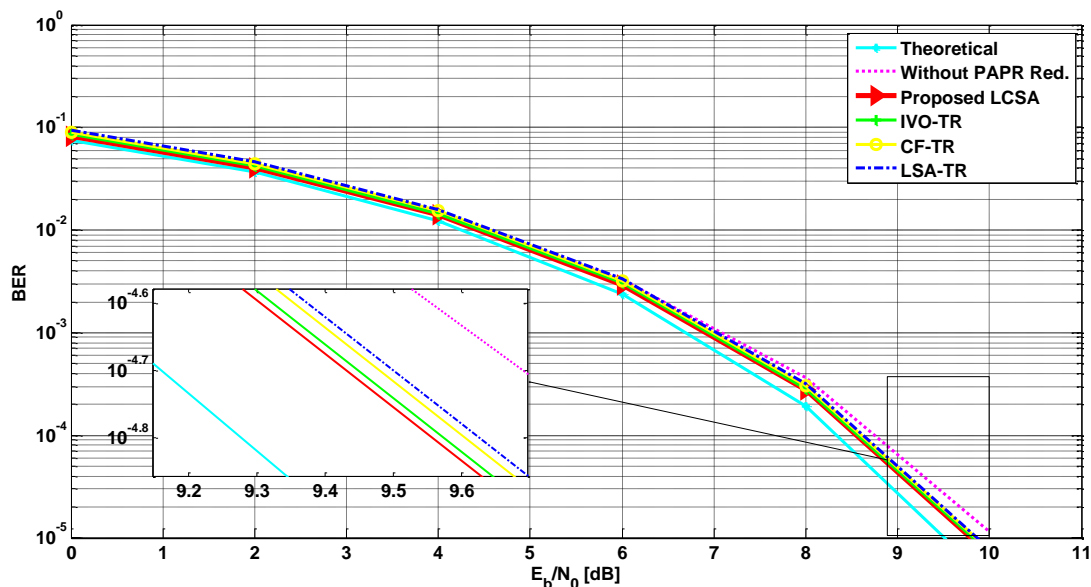
Fig. 7. BER performance of different methods over AWGN channel

Table 5. Required E_b/N_o for the different methods at BER=10^{-5}

Algorithm	LCSA	IVO-TR	CF-TR	LSA-TR
E_b/N_o at BER $=10^{-5}$	9.81	9.83	9.86	9.88

32

6. CONCLUSION

In this paper, we have proposed a new PAPR reduction method for OFDM systems, which we refer to as a low-complexity signal addition (LCSA) method. The method designs a peak-reducing signal in the time-domain and then appends a few samples of the signal to the transmit signal instead of using the traditional way of reserving data subcarriers as done in tone reservation methods. Because of the design approach of the transmitted waveform, the method has a lower data rate loss than the tone reservation methods.

An investigation into the ability of the proposed method to reduce PAPR showed that it could achieve significant PAPR reductions with very low data rate losses while practically maintaining the average power of OFDM signals. The method also has a low runtime computational complexity of $O(M)$ where M is far much less than N, the total number of subcarriers, in an OFDM system.

In comparison with IVO-TR, CF-TR and LSA-TR, the proposed method outperforms the three in terms of PAPR reduction, computational complexity, average transmit power increase, data rate loss and BER performances.

REFERENCES

[1]   Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Comm. Surveys and Tutorials*, Vol. 15, No. 4, 2013, pp. 1567-1592.

[2]   Y. Louët and J. Palicot, "A classification of methods for efficient power amplification of signals," *Annals of Telecom.*, vol. 63, No. 7-8, 2008, pp. 351-368.

[3]   S. Takizawa and H. Ochiai, "PAPR Reduction of OFDM With Trellis Shaping Based on p -Norm Minimization," in *IEEE Wireless Comm. Letters*, vol. 8, no. 4, pp. 988-991, Aug. 2019.

[4]   Tang, K. Qin, and H. Mei "A Hybrid Approach to Reduce the PAPR of OFDM Signals Using Clipping and Companding," *IEEE Access*, Vol. 8, 2020, pp. 18984-18994.

[5]   S. P. Valluri, V. Kishore, and V. M. Vakamulla, "A New Selective Mapping Scheme for Visible Light Systems," *IEEE Access*, Vol. 8, 2020, pp. 18087-18096.

[6]   Jawhar, L. Audah, M. Taher, K. Ramli, N. Shah, M. Musa, and M. Ahmed, "A Review of Partial Transmit Sequence for PAPR Reduction in the OFDM Systems," *IEEE Access*, Vol. 7, 2019, pp. 18021-18041.

[7]   H. Boche and U. J. Mönich, "Peak-to-average power control via tone reservation in general orthonormal transmission systems," *IEEE Trans Signal Process*, vol. 66, no. 13, 2018, pp. 3520-3528.

[8]   J. Tellado, Peak to average power reduction for multicarrier modulation, Ph.D. dissertation, Dept. Elect. Eng., Stanford University, Stanford, CA, USA, 2000.

[9]   T. Jiang, C. Ni, C. Xu, and Q. Qi, "Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems," *IEEE Comm. Letters*, vol. 18, No. 5, pp. 805–808, May 2014.

[10]   H. Li, T. Jiang and Y. Zhou, "An Improved Tone Reservation Scheme With Fast Convergence for PAPR Reduction in OFDM Systems," *IEEE Trans. on Broadcast.*, vol. 57, No. 4, pp. 902-906, Dec. 2011.

[11]   H. Li, J. Wei, and N. Jin, "Low-Complexity Tone Reservation Scheme Using Pre-Generated Peak-Canceling Signals," *IEEE Comm. Letters*, Vol. 23 , Issue: 9, Sept. 2019, pp. 1586-1589.

[12]   M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "New Results on the Peak Power of OFDM Signals Based on Oversampling," *in Proc. IEEE ICC*, Vol. 2, 2002, pp. 866–871.

[13] Y. Cho, J. Kim, W. Yang, and C. Kang, *MIMO-OFDM Wireless Communications with MATLAB*, John Wiley & Sons (Asia) Pte Ltd, Singapore, 2010.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, USA, 2004.

[15] S. Kiambi, E. Mwangi, and G. Kamucha, "An Iterative Re-Weighted Least-Squares Tone Reservation Method for PAPR Reduction in OFDM Systems," WSEAS Trans. on Comm., ISSN / E-ISSN: 1109-2742 / 2224-2864, Vol. 18, 2019, pp. 153–161.

[16] H. J. Yin, R. Z. Yang, X. L. Luo, L. Jiang, and L. J. Zhu , Weighted tone reservation for OFDM PAPR reduction, U.S. Patent 7796498, 2008.

[17] L. Xin and W. Yi, A New Weighted Tone Reservation Method for PAPR Reduction in OFDM Systems, Journal of Communication, Vol. 9, No. 12, 2014, pp. 980-986.

[18] J. Song and H. Ochiai, "Performance analysis for OFDM signals with peak cancellation," *IEEE Trans. on Comm.,* Vol. 64, No. 1, 2016, pp. 261–270.

[19] A. Gatherer and M. Polley, "Controlling clipping probability in DMT transmission," *in Proc. 31st Asilomar Conf. Signals, Syst., Comput*., Vol. 1, Nov. 1997, pp. 578–584.

34

Appendix F: Fifth Publication

PAPR Reduction in MIMO-OFDM Systems Using Low-Complexity Additive Signal Mixing

Stephen Kiambi, Elijah Mwangi, and George Kamucha
School of Engineering, University of Nairobi, Kenya
Email: skiambi@uonbi.ac.ke; elijah.mwangi@uonbi.ac.ke; gkamucha@uonbi.ac.ke

Abstract—A MIMO-OFDM wireless communication technique possesses several advantages accrued from combining MIMO and OFDM techniques such as increased channel capacity and improved BER performance. This has made the technique very amiable to current and future generations of communication systems for high data-rate transmission. However, the technique also inherits the high PAPR problem associated with OFDM signals—a problem still requiring a practical solution. This work proposes a PAPR reduction algorithm for solving the problem of high PAPR in MIMO-OFDM systems. The proposed method uses a low-complexity signal mixing concept to combine the original transmit signal and a generated peak-cancelling signal. The computational complexity of the proposed method is $O(M)$, which is very much less than $O(N \log_2 N)$ of the FFT algorithms. This is because M, which denotes the number of nonzero peak-cancelling samples, is much less than N, the FFT window size. The proposed method was found to achieve high PAPR reductions while utilizing only a few nonzero peak-cancelling samples and it does not significantly change the power of the transmitted signal. For example, with $M = 5\%$ of 256-point IFFT samples, corresponding to a data rate loss of 4.8%, a large PAPR reduction of 5.9 dB could be achieved at a small power loss of 0.09 dB. Compared with other methods proposed in literature, the proposed method was found to outperform them in terms of PAPR reductions and BER performance.

Index Terms—High power amplifier, peak-to-average power ratio, multiple-input multiple-output, orthogonal frequency division multiplexing

I. INTRODUCTION

MIMO-OFDM, as the name suggests, is a technology that takes advantage of the benefits accruing from both Multiple-Input Multiple-Output (MIMO) communication and Orthogonal Frequency Division Multiplexing (OFDM) transmission to increase both data rate and reliability in in a communication system. On one hand, MIMO as a multiple-antenna technique in which multiple antennas are employed at both the transmitter and receiver can be used to bring in spatial diversity and/or spatial multiplexing to, respectively, boost the system reliability i.e. bit-error rate (BER), and increase the achievable data rate by the communication system. In addition, MIMO can also be used for beamforming, to increase coverage, and to reduce

transmit power [1]. On the other hand, OFDM as a multicarrier modulation technique that employs parallel subcarriers to carry user data benefits the MIMO-OFDM communication system mainly with high transmission rates, high spectral efficiency, and the suppression of Inter-Symbol Interference (ISI).

The high spectral efficiency in OFDM is achieved by having all the parallel subcarriers mutually orthogonal to one another. The ISI is eliminated by using a combination of a symbol duration, which is much larger than the expected channel delay spread, and a guard interval between symbols that is simply larger than the delay spread. In addition, in every OFDM symbol, each subcarrier is equivalent to a subchannel. The number of subcarriers is chosen to ensure that each subchannel has a bandwidth less than the coherence bandwidth of the channel so that each of the subchannels experiences a relatively flat fading and therefore at the receiver only a simple single-tap equalizer is required to recover the transmitted data.

The several advantages outlined above have made MIMO-OFDM the key technology for the current and next generations of IEEE 802.16-based worldwide-interoperability for microwave access (WiMAX), 4G and 5G cellular networks, IEEE 802.11-based wireless LAN, wireless Personal Area Network, and broadcasting standards (Digital Audio Broadcasting (DAB), Digital Video Broadcasting (DVB), and Digital Multimedia Broadcasting (DMB)) [1].

Unfortunately, although all MIMO-OFDM systems have all the good attributes of OFDM, they also suffer from the drawbacks they inherit from OFDM. One of the major drawbacks of OFDM, and which is also passed to MIMO-OFDM, is the high peak-to-average power ratio (PAPR) that can occur in the transmit signal. The high PAPR can especially be at unacceptable levels when quite a substantial number or all of the modulated signals on the OFDM subchannels add constructively in a system with a large number of subcarriers. The processing in the High Power Amplifier (HPA) of such high PAPR signals will result to two nonlinear amplification effects, which are BER degradation and out-of-band radiations [2].

In order to avoid the two nonlinear amplification effects and therefore achieve distortionless processing of the high PAPR signals, a simple solution is to input back-off the HPA to a linear region far away from the 1-dB compression point. However, the input backed-off HPA

will have low power efficiency and thus consume more input DC power. This will in turn reduce the lifetime of battery power at user terminals and increase the cost of the transmitter [3]. It is therefore more preferable to reduce any high PAPR in OFDM signals to suitable levels before the signals are passed to the HPA.

In the recent past, different methods have been proposed for PAPR reduction in OFDM and MIMO-OFDM systems. These include signal clipping [4], [5], companding [6], selective mapping [7], [8], partial transmit sequence [9], tone reservation [10], [11], hybrid schemes [12]-[14], etc. As can be noted, the majority of the methods proposed for use in OFDM systems can be re-designed to make them applicable in MIMO-OFDM systems. The proposed methods can generally be classified into four main categories [15]. These are signal distortion techniques, multiple-signalling and probabilistic techniques, coding techniques and hybrid techniques.

Under the signal distortion category, the methods reduce PAPR by distorting the signal before passing it to the HPA. In the multiple-signalling and probabilistic class, the methods both generate numerous alternatives of the OFDM signal and transmit the one with minimum PAPR or they reduce PAPR by modifying the OFDM signal through introduction of phase shifts, or addition of peak reduction tones, or alteration of constellation points. For the coding category, the methods choose the codewords that yield the minimum PAPR while the hybrid class utilises the advantages of different individual techniques and combine two or more schemes to improve PAPR reduction. For all the methods, the bottom line is to achieve significant PAPR reductions and improved BER performance at a minimal change in system complexity.

One of the simplest signal distortion technique employed to reduce PAPR is amplitude clipping, which is normally referred to as conventional clipping. However, the clipping operation can lead to severe BER degradation and high out-of-band radiations. An Adaptive Clipping Technique (ACT) attempting to overcome the shortcomings of the conventional clipping was proposed [16] for Alamouti space-time block code (STBC) MIMO-OFDM systems. However, even with the adaptive clipping, PAPR reductions had to be limited to avoid high BER degradation. In [17], a hybrid technique SCS-SLM combining Selective Codeword Shift (SCS) and selective mapping (SLM) schemes, with the aim of improving PAPR reduction, was proposed for STBC-based MIMO-OFDM systems. Although the hybrid scheme gave better performance than the individual SCS and SLM schemes, the PAPR reduction was still poor. Another hybrid PAPR reduction scheme CSC combining and optimizing three methods, namely convolutional code, successive sub-optimal cross-antenna rotation and inversion, and iterative modified companding and filtering was proposed for STBC-based MIMO-OFDM systems [18]. This method could significantly reduce PAPR but at the expense of poor BER performance and increased system complexity.

PAPR reduction methods that alter the original transmit signal by adding a peak-cancelling signal have been found the most promising because they can achieve both high PAPR reduction and improved BER performance [19]. Examples are methods based on tone reservations, where a subset of data subcarriers are reserved for carrying peak reduction coefficients which on Fourier transformation give the peak-cancelling signal. However, the tone reservation based methods suffer from three major drawbacks: the hard problem of finding the peak-reduction coefficients especially for optimal schemes, data rate loss, and increased transmit power. A sub-optimal Selective Tone Reservation (STR) method [20] was proposed for reducing PAPR in Space-Frequency Block Code (SFBC) MIMO-OFDM systems. The proposed algorithm is based on a time domain kernel, which is added to the signal of the antenna with maximum PAPR to reduce its peak power. Although the method had low complexity compared with an optimal scheme, it was prone to peak re-growth, increased transmit power and had poor PAPR reduction performance.

To overcome the drawbacks in tone reservation methods, in this paper, we propose a low-complexity additive signal mixing method for reducing PAPR in space-time-coded MIMO-OFDM systems. The key idea is to generate a peak-cancelling signal for each MIMO-OFDM branch based on a predetermined clipping threshold for the system at hand. To reduce PAPR, the respective peak-cancelling signals are added to the branch signals. For mitigating the data rate loss and the increase in the average power, only the nonzero samples of the peak-cancelling signals are transmitted together with the PAPR-reduced signals for use to reconstruct clipped amplitudes. In comparison with ACT, SCS-SLM, CSC, and S-TR methods, the proposed method was found to have better performance in PAPR reduction and BER improvement.

The rest of this paper is organized as follows. Section 2 presents the space-time block-coded MIMO-OFDM system and the associated PAPR. In Section 4, the proposed PAPR reduction method is presented. Section 5 provides simulation results and their analysis. Lastly, Section 6 concludes the paper.

II. MIMO-OFDM SYSTEM AND PAPR

A. Space-Time Block Code (STBC)

The very first and well-known STBC that provides transmit diversity is the Alamouti space-time code. The Alamouti code is a complex orthogonal space-time code specialised for the case of two transmit antennas but can be generalised to the case of three antennas or more. In the Alamouti encoder, two consecutive symbols X_1 and X_2 are encoded with the following space-time codeword matrix [21]:

$$X = \begin{bmatrix} X_1 & -X_2^* \\ X_2 & X_1^* \end{bmatrix} \tag{1}$$

where the $*$ denotes complex conjugate.

The Alamouti encoded signal is transmitted from the two transmit antennas over two symbol periods. During the first symbol period, the two symbols X_1 and X_2 are simultaneously transmitted from the first and the second antenna, respectively. In the second symbol duration, the same symbols are essentially re-transmitted in the form of two symbols, $-X_2^*$ and X_1^*, which are simultaneously transmitted from the first and the second antenna, respectively. At the receiver, an Alamouti STBC decoder is implemented.

Systems deploying multiple antenna techniques at the transmitter and the receiver are normally referred by their number of transmit and receive antenna configurations. There are two common implementations of Alamouti STBC receivers: one with one receive antenna and the other with two receive antennas corresponding to 2×1 and 2×2 systems, respectively. The two antenna configurations have the same channel capacity but different spatial diversity gains.

In general, the use of multiple antenna configurations in the transmitter and the receiver is supposed to increase the channel capacity of a single-input single-output (SISO) antenna configuration by a factor of $\min(N_t, N_r)$, where N_t and N_r are the number of transmit and receive antennas, respectively [22]. Since for the Alamouti STBC system, the same symbols are transmitted during two symbol durations, the channel capacity is still equal to the SISO system capacity. However, the Alamouti schemes achieve a diversity gain $N_d = N_t \times N_r$, which by definition is the number of independent channel paths between the transmitter and the receiver. The 2×2 antenna system has a diversity gain of four, which is double that of the 2×1 antenna system.

Since the probability of all the N_d channel paths having low signal-to-noise ratio (SNR) is very small, the diversity gain has profound effect on the system reliability. The average bit-error probability, p_b, of a multiple-antenna system decreases exponentially with the diversity gain [23] according to

$$p_b = k \Upsilon^{-N_d} \qquad (2)$$

where k is a constant that depends on the modulation type and Υ is the received SNR. From this equation, it is clear that a 2×2 MIMO system offers a better BER performance than a multiple-input single-output (MISO) 2×1 system.

The Alamouti scheme for the 2×2 system is illustrated in Fig. 1. The channel impulse responses h_{11}, h_{12}, h_{21} and h_{22} are assumed to be time-invariant over two symbol durations and are of the form $|h_{ij}|e^{\theta_{ij}}$, where $|h_{ij}|$ and $e^{\theta_{ij}}$ denote the amplitude gain and phase alteration over the two symbol periods, and i and j are equal to 1 or 2.

The received signals at the first receive antenna (the upper receive antenna in Fig. 1) during the first and second symbol durations are, respectively, given by

$$Y_{11} = h_{11}X_1 + h_{12}X_2 + n_{11}$$
$$Y_{12} = -h_{11}X_2^* + h_{12}X_1^* + n_{12}. \qquad (3)$$
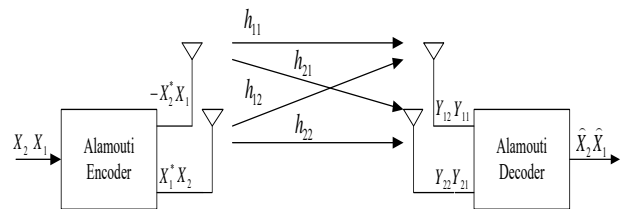
Fig. 1. Alamouti scheme for 2×2 MIMO system.

At the second antenna the received signals for the two symbol durations are:

$$Y_{21} = h_{21}X_1 + h_{22}X_2 + n_{21}$$
$$Y_{22} = -h_{21}X_2^* + h_{22}X_1^* + n_{22}. \qquad (4)$$

The four additive terms n_{11}, n_{12}, n_{21} and n_{22} in (3) and (4) represent additive white Gaussian noise. Taking the complex conjugate of the signals received during the second symbol duration, the system equation for the space-time-coded MIMO system can be written as follows:

$$\begin{bmatrix} Y_{11} \\ Y_{12}^* \\ Y_{21} \\ Y_{22}^* \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^* & -h_{11}^* \\ h_{21} & h_{22} \\ h_{22}^* & -h_{21}^* \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} + \begin{bmatrix} n_{11} \\ n_{12}^* \\ n_{21} \\ n_{22}^* \end{bmatrix}. \qquad (5)$$

This equation is of the form

$$Y = HX + n \qquad (6)$$

where H is the channel matrix given by

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12}^* & -h_{11}^* \\ h_{21} & h_{22} \\ h_{22}^* & -h_{21}^* \end{bmatrix} \qquad (7)$$

Since the two columns of the channel matrix are orthogonal, the system equation in (5) can be decoded to obtain the estimates of the transmitted signals by multiplying through by the Hermitian transpose of the channel matrix given as

$$H^H = \begin{bmatrix} h_{11}^* & h_{12} & h_{21}^* & h_{22} \\ h_{12}^* & -h_{11} & h_{22}^* & -h_{21} \end{bmatrix} \qquad (8)$$

The transmitted signals are then estimated using the equation

$$\hat{X} = \begin{bmatrix} \hat{X}_1 \\ \hat{X}_2 \end{bmatrix} = \frac{H^H Y}{|h_{11}|^2 + |h_{12}|^2 + |h_{21}|^2 + |h_{22}|^2} \qquad (9)$$

B. Space-Time Block-Coded MIMO-OFDM System

A 2×2 MIMO-OFDM system employing space-time coding is simply an extension of the 2×2 MIMO system in Fig. 1 where the blocks of OFDM signal processing are added after the STBC encoder as illustrated in Fig. 2. The

main OFDM signal-processing blocks in the transmitter are Inverse Fast Fourier Transform (IFFT), Cyclic Prefix (CP) addition, Digital-to-Analogue Converter (DAC), HPA, and RF front-end up-converter.

In the receiver section, the signal processing operations in the transmitter section are reversed. Therefore, each branch of a 2×2 MIMO-OFDM system is the same and experiences similar effects of high PAPR as a SISO-OFDM system.

C. PAPR in MIMO-OFDM System

With the consideration that each branch of a MIMO-OFDM system is equivalent to a SISO-OFDM system, the IFFT output during one-symbol duration is the baseband signal given by

$$x_i(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_i(k) \, e^{\frac{j2\pi kn}{N}} \qquad (10)$$

Here, $X_i(k)$ is the modulation symbol from binary phase-shift keying (BPSK) or M-ary quadrature amplitude modulation (M-QAM), N is the total number of subcarriers, and $i = 1$ or 2 is the branch index.

For each branch signal $x_i(n)$, the ratio of the peak power to the average power is given by

$$\text{PAPR}\{x_i(n)\} = \frac{\max_{0 \le n \le N-1}\{|x_i(n)|^2\}}{E\{|x_i(n)|^2\}} \qquad (11)$$

where $E\{.\}$ is the expectation operator. For the MIMO-OFDM system, we are interested with the maximum PAPR among all branches, which for the 2×2 system is given by

$$PAPR_{MIMO} = \max(\text{PAPR}\{x_1(n)\}, \text{PAPR}\{x_2(n)\}) \qquad (12)$$

Because the input to the HPA is a continuous-time signal $x_i(t)$, in the calculation of PAPR, signal $x_i(n)$ should be oversampled with a factor ≥ 4. This avoids skipping the peak value of the continuous-time signal [24] and in turn helps to closely estimate the continuous-time PAPR.

From (10), it is clear that each branch signal is a summation of N signals and therefore can have large amplitude fluctuations resulting from constructive and destructive additions. These amplitude fluctuations can result into high PAPR and nonlinear amplification effects
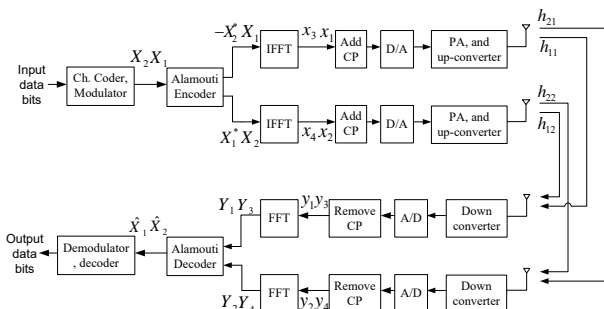
Fig. 2. MIMO-OFDM system deploying 2×2 Alamouti STBC.

when the signal is passed through the HPA. Analytically, the distribution of the magnitudes of the amplitudes, i.e. $|x_i(n)|$, can point as to whether or not a signal has high power fluctuations and therefore high PAPR.

Assuming a sufficiently large N and that the signal amplitudes are statistically independent and identically distributed, by the central limit theorem, both the real and imaginary parts of the amplitudes are Gaussian-distributed and therefore the signal magnitudes $|x_i(n)|$ are Rayleigh-distributed. Consequently, signal $x_i(n)$ can have high PAPR. The highness of a PAPR can be measured using the complementary cumulative distribution function (CCDF), which is the probability that PAPR is above a given threshold γ and is given by the equation

$$\Pr\{\text{PAPR}\{x_i(n)\} > \gamma\} = 1 - (1 - e^{-\gamma})^N \qquad (13)$$

where $\Pr\{.\}$ is the probability operator.

From (13), for fixed values of CCDF and N, a high value of threshold indicates a high PAPR and vice versa. This has the interpretation that for a given CCDF value, the difference between any two-threshold values can be used as a measure of PAPR reduction and to indicate how well any proposed method reduces PAPR.

III. PROPOSED METHOD

In this paper, we propose a low-complexity method that reduces PAPR in MIMO-OFDM systems by generating additive peak-cancelling signals to reduce high signal amplitudes in transmitted signals. The proposed method is referred to as "low-complexity additive signal mixing" or in short form ASM PAPR reduction method. In order to avoid BER degradation due to clipping of high signal amplitudes, a few samples of the peak-cancelling signals are appended to the transmitted signals for the restoration of clipped amplitudes at the receiver. The peak-cancelling signals are derived from MIMO-OFDM transmit signals considering the peak power that the HPA can handle without signal distortion.

A. Proposed Algorithm

For exact cancellation of the highest peaks in a transmit signal without introducing new ones, a peak-cancelling signal should only have samples of signal peaks exceeding a clipping threshold. The clipping threshold can generally be set based on the desired PAPR level in a system. For a given clipping threshold x_{th}, the desired peak-cancelling signal for signal $x_i(n)$ can be generated according to the equation

$$d_i(n) = \begin{cases} \dfrac{x_i(n)}{|x_i(n)|}\left(|x_i(n)| - x_{th}\right), & |x_i(n)| > x_{th} \\ 0 & , \ |x_i(n)| \le x_{th} \end{cases} \qquad (14)$$

In vector form this signal can be expressed as $\boldsymbol{d}_i = [d_i(0), d_i(1), \ldots, d_i(N-1)]^T$. The signal has both zero and nonzero samples. A simplified discrete-time signal $c_i(k)$ containing only the nonzero entries can be written as

$$c_i = [c_i(0), c_i(1), \dots, c_i(M-1)]^T \qquad (15)$$

where M is the number of nonzero samples in $d_i(n)$.

The tone reservation concept in [25] can be used to generate a peak-cancelling signal that estimates the signal in (14). This can be accomplished by reserving L subcarriers and then solving for the frequency-domain peak-cancelling coefficients from the following system of linear equations:

$$\hat{Q}C_i = d_i \qquad (16)$$

where $C_i \in \mathbb{C}^L$ is the peak-cancelling vector with L coefficients, $\hat{Q}_i \in \mathbb{C}^{N \times L}$ is the submatrix made up of L columns, corresponding to the locations of reserved subcarriers in the IDFT matrix $Q \in \mathbb{C}^{N \times N}$ whose elements are given by $(1/\sqrt{N})\exp(j2\pi kn/N)$.

Due to the reservation of L subcarriers that do not carry user data, there is a loss in data rate as indicated by the following ratio:

$$R_{l,f} = \frac{L}{N} \qquad (17)$$

In order to reduce the data rate loss, it is desirable to have a very low value of L that is much smaller than N but this may negatively affect the PAPR reduction capability of the tone-reservation method.

Because $L \ll N$, the system in (16) is overdetermined and can only be solved through the least-squares minimization [26]-[28] of the residual error

$$\epsilon_i = \hat{Q}_i C_i - d_i \qquad (18)$$

resulting in the closed form solution

$$C_i = [\hat{Q}_i^H \hat{Q}_i]^{-1} \hat{Q}_i^H d_i \qquad (19)$$

After finding the frequency-domain coefficients, the time-domain peak-cancelling signal is obtained using the equation

$$\hat{d}_i = \hat{Q}_i C_i \qquad (20)$$

and the PAPR-reduced signal is then given by

$$s_i(n) = x_i(n) - \hat{d}_i(n) \qquad (21)$$

Because of the over-deterministic nature of the system determining the peak-cancelling coefficients, the peak-cancelling signal $\hat{d}_i$ cannot be equal to the desired signal d_i, and has nonzero elements even in positions that had zeros in the desired signal. This can lead to the generation of new signal peaks in the PAPR-reduced signal $s_i(n)$ and in turn result in poor PAPR reduction. The only way to improve the solution in (19), in order to have $\hat{d}_i \cong d_i$, is to sufficiently increase L towards N but this will lead to an unacceptably high data rate loss.

Therefore, the two requirements of high PAPR reduction and low data rate loss compete directly with each other, and this poses a design dilemma in the development of the PAPR reduction method. In order to achieve both a high PAPR reduction and a minimum data rate loss, we propose, in this work, an algorithm that directly employs the desired peak-cancelling signal in (14) to reduce PAPR. The PAPR-reduced signal in this case is given by

$$s_i(n) = x_i(n) - d_i(n) \qquad (22)$$

This ensures that all the highest speaks of the signal $x_i(n)$ are cancelled out without generating new ones. However, the clipping of the signal amplitudes will result in BER degradation. To avoid this, a few samples, the nonzero ones, of the desired peak-cancelling signal will be transmitted together with the PAPR-reduced signal to enable the receiver reconstruct back the clipped amplitudes.

However, the transmission of the nonzero peak-cancelling samples affects the system data rate because they do not carry user data. This data rate loss occurs in the time-domain and is given by

$$R_{l,t} = \frac{M}{N+M}. \qquad (23)$$

But for the same number of reserved tones and nonzero time samples, the data rate loss in the proposed method is less than the one given in (17) for the tone-reservation based methods. However, owing to the choice of the clipping threshold, which must be greater than the average value of the signal, the number of nonzero samples in $d_i(n)$ will be in all cases very small compared to the length of the signal i.e. $M \ll N$ and therefore $R_{l,t} \approx 0$.

Logically, the threshold at which signal $x_i(n)$ is clipped can be expressed as a function of the average value of the signal using the equation:

$$x_{th} = \lambda \mu \qquad (24)$$

where μ is the mean of the signal amplitudes $|x_i(n)|$, and λ is the threshold adjustment parameter in the range $1 < \lambda < \max(|x_i(n)|)/\mu$.

When for a given system, the maximum allowed PAPR is known, the required x_{th}, and hence λ, can be found directly from equation (12). After the determination of the clipping threshold, M can be obtained analytically from the distribution of the signal amplitudes. For such a derivation, let X_n denote a Rayleigh random variable representing the distribution of the signal amplitudes that is given by the equation

$$F_{X_n}(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0 \qquad (25)$$

where σ is the scaling parameter of the distribution. The average value of the distribution is

$$\mu = \sigma \sqrt{\frac{\pi}{2}} \qquad (26)$$

From (25), the probability of signal amplitude being greater than the clipping threshold is given by:

$$P(X_n > x_{th}) = e^{-\frac{x_{th}^2}{2\sigma^2}} \quad (27)$$

and since this is equal to the ratio M/N, then by substituting (24) in (27), the number of nonzero entries in $d_i(n)$ is obtained as follows:

$$M = Ne^{-\frac{\pi}{4}\lambda^2} \quad (28)$$

Equation (28) indicates that the number of nonzero elements in the peak-cancelling signal decreases exponentially with the clipping threshold. The maximum value of M will occur in the trivial case when the clipping threshold is equal to the average value, corresponding to $\lambda = 1$ and $M = 0.46N$. The minimum value of M will occur when $x_{th} = \max(|x_i(n)|)$. Then, in all practical cases, $M \ll N$ and therefore the data rate loss will always be negligible.

The proposed algorithm is flexible in terms of the inputs it can use to process peak-cancelling signals from original transmit signals. Either it can use the maximum acceptable PAPR or the maximum allowed data rate loss. When one of either of the two inputs is given, the algorithm can compute the other and determine whether it is within the acceptable limits. If M or the maximum allowed data rate loss is known, by using (28) the clipping threshold can be obtained as follows:

$$x_{th} = \mu\sqrt{\frac{4}{\pi}\ln\left(\frac{N}{M}\right)} \quad (29)$$

After obtaining x_{th}, the peak-cancelling signal can be found from (14).

Fig. 3 shows two typical peak-cancelling signals for a 2×2 MIMO-OFDM system. It can be observed that for the second transmit antenna, the peak-cancelling signal has 13 nonzero samples while in the first antenna signal they are 10. The highest peak also occurs in the second antenna signal. Consequently, this means that a higher PAPR reduction is required on the transmit signal from the second antenna than on the one from the first antenna. Additionally, because of the oversampling by a factor of 4, N is equal to $1024/4 = 256$ and the data rate loss $R_{l,t} = 0.048$.
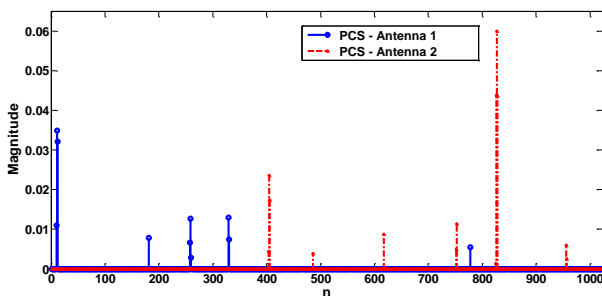
Fig. 3. Peak-cancelling signals in 2×2 MIMO-OFDM system.
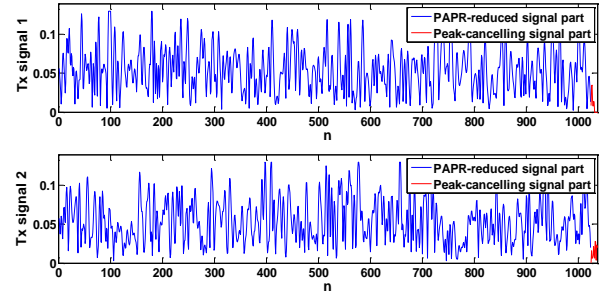
Fig. 4. Composite transmit signals in 2×2 MIMO-OFDM system.

The corresponding composite signals from the two transmit antennas are shown in Fig. 4. From this figure, it is evident that when $M \ll N$ the waveform of the PAPR-reduced signal is barely affected by the addition of M peak-cancelling signal samples. Intuitively, this also means that the transmission of the M samples has no significant effect on the average transmit power.

However, analytically, the effect of transmitting the nonzero peak-cancelling samples on the average transmit power can be established by considering the sum of the powers in the clipped samples and the peak-cancelling clipped samples. For this purpose, let $\tilde{x}_i(k)$ denote the set of non-clipped samples greater than the clipping threshold x_{th} in the ith branch signal $x_i(n)$ in the MIMO-OFDM system. After the amplitude-clipping of the signal $x_i(n)$, all the samples in $\tilde{x}_i(k)$ will have the same signal magnitude equal to x_{th}. Since all the signal magnitudes $|x_i(n)|$ are in the range of 0 to 1, the sum of the powers in the clipped and peak-cancelling samples will always be less than or equal to the total power in the original non-clipped samples i.e.

$$\sum_{k=1}^{M}(x_{th}^2 + |c_i(k)|^2) \leq \sum_{k=1}^{M}|\tilde{x}_i(k)|^2 \quad (30)$$

Equation (30) may suggest that the power of the composite signal can decrease with the number of nonzero samples. However, again, because of the range of values occupied by the signal amplitudes, such power reductions are very small. Therefore, the power of the original transmit signal is practically maintained by the proposed method.

From the foregoing description, the proposed algorithm can be summarized as follows:

ASM Algorithm
i.   Set the number of subcarriers N, data rate loss $R_{l,t}$, and $PAPR_{max}$
ii.   Generate MIMO-OFDM signals $x_i(n)$, $i = 1,2,\ldots N_t$
iii. Find $PAPR_{MIMO}$
iv. If $PAPR_{MIMO} < PAPR_{max}$, transmit $x_i(n)$ and terminate the algorithm, else go to step (v)
v.   Set clipping threshold x_{th}
vi. Generate desired peak-cancelling signals $d_i(n)$
vii. Generate peak-reduced signals $s_i(n) = x_i(n) - d_i(n)$
viii. Determine nonzero samples in $d_i(n)$ and generate signals $c_i(k)$

ix. Append $c_i(k)$ to $s_i(n)$ and transmit combined signal

x. End
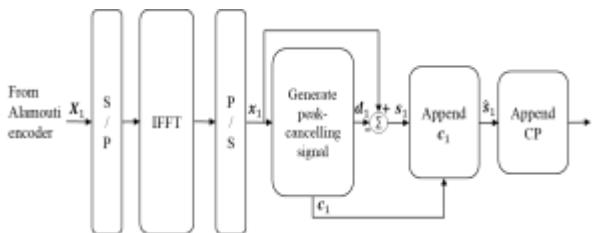
Fig. 5. Transmitter section integrating ASM method.

A MIMO-OFDM system integrating the proposed PAPR reduction is illustrated in Fig. 5. The illustration is for one of the MIMO branches during the first symbol duration. The OFDM symbol X_1 from the Alamouti encoder is first converted from serial to parallel (S/P) format and then passed to the IFFT to obtain a time-domain signal, which is then converted back to serial format to produce signal x_1. The proposed method is then applied to x_1 to produce two outputs: the peak-cancelling signal d_1 and its compressed version c_1 containing only the nonzero samples. Signal d_1 is then subtracted from x_1 to get a PAPR-reduced signal s_1, which is then extended by c_1 to yield the composite transmit signal $\hat{s}_1$.

The composite transmit signal is then cyclically extended in the append CP block with a cyclic prefix of length greater than the expected channel duration. The extension is realized by copying some samples at the tail end of the signal to the front end to provide a time guard interval for the elimination of the ISI. After the CP addition, the composite transmit signal is amplified through the HPA and then passed on for RF transmission.

At the receiver, after removing the cyclic prefix, the M peak-cancelling samples in c_1 are also removed and added back to the clipped samples to reconstruct the original OFDM signal x_1, which is then passed to the FFT block for demodulation and after decoded into binary data.

B. High Power Amplifier Modelling

The Rapp model [29] of the HPA is used in this work. The model has a constant AM/PM characteristic, which makes it suitable for modelling the phase distortion of the solid-state power amplifier (SSPA). On the other hand, the AM/AM characteristic of the model can be easily tuned to simulate various nonlinear characteristics of the HPA. In the model, it is assumed that the HPA has a linear performance up to a point near the saturation point. Near the saturation point, a transition towards a constant saturated output is applied on the input signal.

Since the model does not introduce any phase distortions, the AM/PM relation can be written as

$$\varphi(x(n)) \approx 0 \qquad (31)$$

This means that the HPA does not introduce any phase changes during the amplification of the input signal.

For the amplitude amplification, the general expression for the AM/AM conversion is given by

$$g(x(n)) = \frac{x(n)}{\left(1 + \left(\frac{|x(n)|}{A_{sat}}\right)^{2p}\right)^{\frac{1}{2p}}} \qquad (32)$$

Here, $x(n)$ is the input signal and A_{sat} denotes the output at the 3-dB point and is used to set the HPA saturation level. The smoothness parameter p is used to smoothen the amplification during the transition from the linear to the saturation region. Therefore, a smaller p means a smoother transition and vice versa.

C. Computational Complexity

The proposed algorithm performs two main operations, which are the generation of peak-cancelling signal using (14) and the signal addition operation in (22) to reduce PAPR. The operation of computing the peak-cancelling signal requires $2M$ real multiplications and $2M$ real additions while the operation for reducing PAPR requires $2M$ real additions. At the receiver, the reconstruction of the clipped signals peaks requires $2M$ real additions.

The multiplication operations are more computationally intensive than the additions and are the one that determines the overall complexity of an algorithm. Therefore, in the order of the number of multiplications, the computational complexity of the proposed method is $O(M)$.

Because the main part of the proposed method is to be implemented in the transmitter and therefore form part of the signal processing, there is need to consider any arising increase in the overall computational complexity. For the OFDM signal processing at the transmitter, the most complex operation is the IFFT, which has a computational complexity of $O(N \log_2 N)$. Since $M \ll N$, then $O(M) \lll O(N \log_2 N)$. This means that incorporating the proposed method into a MIMO-OFDM system will not change the overall computational complexity of the system.

IV. RESULTS AND DISCUSSION

The proposed ASM PAPR reduction method was applied to reduce PAPR in MIMO-OFDM systems. Simulations of MIMO-OFDM systems were carried out in MATLAB. The key simulation parameters are listed in Table I. An Alamouti space-time code was used with 2 transmit and 2 receiver antennas over Rayleigh flat-fading channels. The Rapp model of the HPA was used in the simulations. In each simulation scenario, the proposed algorithm was executed for 10^4 symbols.

TABLE I: SIMULATION PARAMETERS

FFT window size	128, 256
Modulation	QPSK
Number of OFDM symbols	10^4
Oversampling factor F_s	4
Power amplifier model	Rapp model, $p = 2$
Guard interval	1/4
Channel model	Rayleigh flat-fading

As it is indicated in the table, all the subcarriers were modulated with QPSK data. This is sufficient for ascertaining the method's performance and for comparison to other methods because the type of modulation does not affect the PAPR reduction performance. Both the PAPR reduction and the BER performances were analysed.

The very first simulations were to help assess the PAPR reduction capability of the proposed method. For this task, a system with $N = 256$ subcarriers was employed. The system was subjected to the following number of peak-cancelling samples: $M = 3$, 6, 13, 19, and 26. These values of M correspond to the following data-rate losses: 1.2%, 2.3%, 4.8%, 6.9%, and 9.2%, respectively. The reduction of PAPR was then evaluated for each case using the CCDF as shown in Fig. 6.
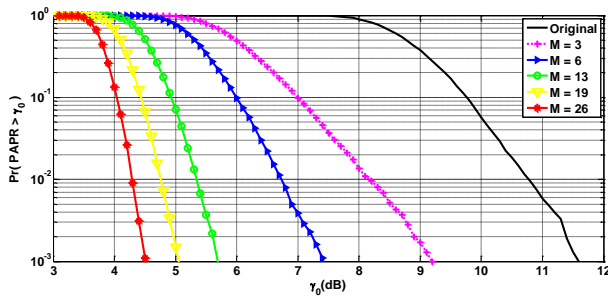
Fig. 6. CCDF for MIMO-OFDM system with QPSK data and $N = 256$.

TABLE II: PAPR REDUCTIONS AT CCDF = 10^{-3}, $N = 256$

M	3	6	13	19	26
$R_{l,t}$ (%)	1.2	2.3	4.8	6.9	9.2
PAPR Reduction (dB)	2.4	4.2	5.9	6.5	7.1
Power Change (dB)	-0.02	-0.04	-0.09	-0.13	-0.18

In Table II, the results for PAPR reductions at CCDF $= 10^{-3}$ and transmit power changes due to the use of different number of peak-cancelling samples are given. From this table, it is evident that the capability of the proposed method to reduce PAPR depends on the number of the peak-cancelling samples employed. However, it can also be observed that with only a small number of peak-cancelling samples, high PAPR reductions can be achieved e.g. with only 6 samples a good reduction of 4.2 dB could be achieved.
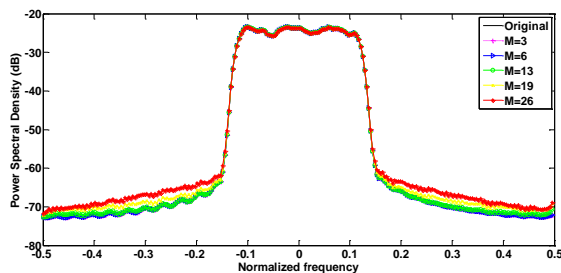
Fig. 7. Power spectral densities for different PAPR-reduced signals.

In addition, on the tabulated results for the average power it is clear that the transmit power of the signal is practically maintained before and after PAPR reduction.

For example, for $M = 26$, the average transmit power is 99.3% of the value before the PAPR reduction. Moreover, from the power spectral densities plots in Fig. 7, it can be observed that due to the small amount of clipping on only a few number of signal amplitudes, the out-of-band radiations are negligible.

After ascertaining the ability of the proposed ASM method to reduce PAPR, it was then compared with four other promising PAPR reduction methods proposed in literature, which were earlier on given the acronyms SCS-SLM, STR, ACT and CSC. For the purpose of this comparison, a MIMO-OFDM system with QPSK-modulated subcarriers and $M = 23$ samples, equivalent to a data rate loss of 15%, was used. The results showing the PAPR reduction performances of the different methods are shown in Fig. 8 and also given in Table III for CCDF $= 10^{-3}$.
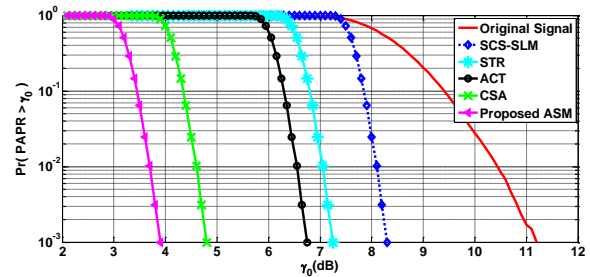
Fig. 8. PAPR reduction by different methods, $N = 128$.

Apparently, the proposed method shows better PAPR reduction capability than the other four methods and this is well demonstrated in the available results. For example, from the results in Table III, the proposed ASM method gives a PAPR reduction that is higher by 0.9, 2.85, 3.35, and 4.4 dB than the corresponding reductions by CSC, ACT, STR and SCS-SLM methods.

TABLE III: PAPR REDUCTIONS AT CCDF $=10^{-3}$, $N = 128$

Method	ASM	CSA	ACT	STR	SCS-SLM
PAPR reduction (dB)	7.30	6.40	4.45	3.95	2.90

The second simulations were used to evaluate the BER performance of MIMO-OFDM systems employing the proposed method. These tests matter a lot because BER performance is the single most important indicator of whether or not a receiver in a communication system can recover transmitted symbols. The Rapp model of HPA was used with an input power back-off (IBO) set just slightly above the PAPR value of signals at CCDF $= 10^{-3}$. This ensured that the number of symbols clipped by the HPA was less than 1%.

After amplification through the HPA, the composite signal was transmitted over Rayleigh flat-fading channels with additive white Gaussian noise. The BER degradation when using different number of nonzero peak-cancelling samples was first considered followed by a BER performance comparison with the other methods.

In Fig. 9, the BER degradations caused by the use of different values of M are shown. There are two curves labelled "Ideal": one for when the system was simulated without the use of any PAPR reduction method and HPA, and the other for the case without PAPR reduction method but with HPA backed-off by 12.5 dB—a value slightly above the maximum PAPR value of 11.8 dB at CCDF $= 10^{-3}$. The two ideal cases are similar and give the best-expected case of BER performance by the system.
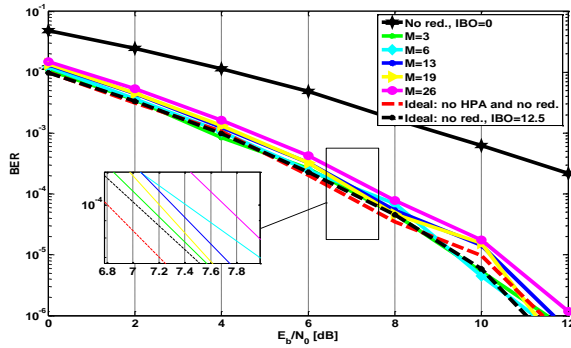
Fig. 9. BER performance by different M-values, $N = 256$.
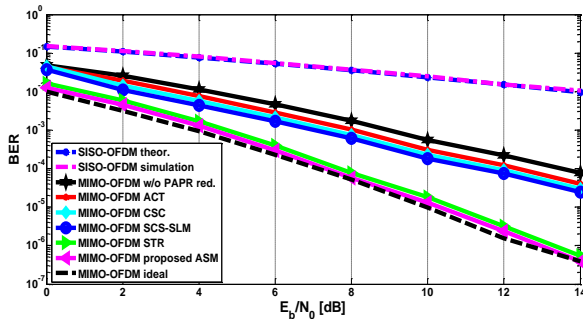
Fig. 10. BER performance by different methods, $N = 128$.

The worst expected BER performance is given by the curve labelled "No red., IBO=0", which corresponds to the case when there was no PAPR reduction and the HPA was not provided with input back-off. For this case, the HPA clipped all the signal amplitudes that had power greater than the average power, and thereby highly degrading the BER.

For the other curves for the different values of M, it can be observed that the degradation of BER slightly increases with M. However, all the curves for the different values of M are tightly close to the ideal BER curves. This shows that the proposed method can effectively reduce PAPR, and thereby improving the HPA efficiency, without significantly degrading the BER performance of the MIMO-OFDM system.

For the comparison with other methods, the BER performances of the various methods are depicted in Fig. 10. In the case of SISO-OFDM system, the BER curve by theoretical formula and that by simulation are merged, thus implying the same BER performance. The simulated case had the proposed method modified and employed to reduce PAPR in the SISO-OFDM system. The curve labelled "MIMO-OFDM ideal" is the lower limit BER

performance as it corresponds to the case when the HPA and PAPR reduction were not used. The curve labelled "MIMO-OFDM w/o PAPR red." is the worst-case and the upper limit performance as it corresponds to the case when no PAPR reduction method was applied to the transmit signal prior to amplification in the HPA.

As it can be observed from the BER curves, the BER performance of all the MIMO cases, even for the worst case, were by far much better than for the two SISO cases. This confirms the theory that was presented earlier that the BER falls exponentially with the diversity order. Additionally, it can be observed that both the proposed ASM method and the STR method have BER degradations close to the ideal case while the SCS-SLM, CSC, and ACT have poor performances.

In the overall, the BER performance of the proposed method is the best. For example, from the results in Table IV of the required SNR per bit, i.e. E_b/N_0, at BER $= 10^{-3}$, the proposed method requires an E_b/N_0 that is smaller by 0.29, 2.68, 3.05, and 3.64 dB to that needed by the STR, SCS-SLM, CSC and ACT methods, respectively.

TABLE IV: REQUIRED E_b/N_0 BY DIFFERENT METHODS AT BER $= 10^{-3}$

Method	ASM	STR	SCS-SLM	CSC	ACT
E_b/N_0 (dB)	4.36	4.75	7.04	7.41	8.00

V. CONCLUSION

In this paper, a new PAPR reduction method for MIMO-OFDM systems has been proposed. The method utilizes a low-complexity additive signal-mixing concept to reduce PAPR by first designing a peak-cancelling signal for each MIMO diversity arm, and then adding it to each arm's transmit signal. To avoid BER degradation due to peak reductions, a few samples of the peak-cancelling signal are appended to the transmit signal to be used for amplitude reconstructions at the receiver.

Therefore, the method reserves peak-cancelling resources in the time domain rather than in the frequency domain. This makes the method to have a lower data rate loss than in a conventional tone reservation method that reserves peak-cancelling resources in the frequency domain.

An investigation into PAPR reduction capability showed that the proposed method could achieve significant PAPR reductions with very low data rate losses while practically maintaining average transmission powers of original MIMO-OFDM signals. In addition, the method has a low computational complexity of $O(M)$, which is by far much less than the FFT complexity of $O(N \log_2 N)$.

Additionally, the method does not degrade the BER of the initial MIMO-OFDM system. In overall, in comparison with four other PAPR reduction methods: ACT, STR, SCS-SLM and CSC, the proposed method gives better PAPR reduction and BER performances.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Y. Cho, J. Kim, W. Yang, and C. Kang, *MIMO-OFDM Wireless Communications with MATLAB*, John Wiley & Sons (Asia) Pte Ltd, Singapore, 2010.

[2] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Comm. Surveys and Tutorials*, vol. 15, no. 4, pp. 1567-1592, 2013.

[3] Y. Louët and J. Palicot, "A classification of methods for efficient power amplification of signals," *Annals of Telecom.*, vol. 63, no. 7-8, pp. 351-368, 2008,.

[4] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *IEE Electronic Lett* , pp. 246–247, 2002.

[5] R. Zayani, H. Shaiek, and D. Roviras, "PAPR-aware massive MIMO-OFDM downlink," *IEEE Access*, vol. 7, pp. 25474-25484, 2019.

[6] Vallavaraj, B. G. Stewart, and D. K. Harrison, "An evaluation of modified μ-law companding to reduce the PAPR of OFDM systems," *AEU-Int. J. Electron. Commun.*, vol. 64, no. 9, pp. 844-857, 2010.

[7] H. B. Jeon, J. S. No, and D. J. Shin, "A low-complexity SLM scheme using additive mapping sequences for PAPR reduction of OFDM signals," *IEEE Trans. Broadcast.*, vol. 57, no. 4, pp. 866-875, 2011.

[8] T. Jiang, C. Ni, and L. Guan, "A novel phase offset SLM scheme for PAPR reduction in alamouti MIMO-OFDM systems without side information," *IEEE Signal Process. Lett.,* vol. 20, no. 4, pp. 383-386, 2013.

[9] Lahcen, A. Saida, and A. Adel, "Low computation complexity PTS scheme for PAPR reduction of MIMO-OFDM systems," in *Proc. 10th Int. Conf. Interdispl. in Eng., Procedia Eng.*, vol. 181, 2017, pp. 876-883.

[10] S. Kiambi, E. Mwangi, and G. Kamucha, "An iterative Re-Weighted least-squares tone reservation method for PAPR reduction in OFDM systems," *WSEAS Trans. on Comm.*, vol. 18, pp. 153–161, 2019.

[11] Ni, Y. Ma, and T. Jiang, "A novel adaptive tone reservation scheme for PAPR reduction in large-scale multiuser MIMO-OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 480-483, 2016.

[12] Tang, K. Qin, and H. Mei "A hybrid approach to reduce the PAPR of OFDM signals using clipping and companding," *IEEE Access*, vol. 8, pp. 18984-18994, 2020.

[13] B. Bakkas, I. Chana, and H. Ben-Azza, "PAPR reduction in MIMO-OFDM based on polar codes and companding," *Proc. IEEE CommNet*, 2019, pp. 1-6.

[14] Beena, S. Pillai, and N. Vijayakumar, "Hybrid PTS-clipping scheme for PAPR reduction in MIMO-OFDM systems," *Int. Journal of Applied Eng. Research*, vol. 13, no. 11, pp. 9924-9928, 2018.

[15] Sandoval, G. Poitau, and F. Gagnon, "Hybrid peak-to-average power ratio reduction techniques: review and performance comparison," *IEEE Access*, vol. 5, pp. 27145-27161, 2017.

[16] S. Singh and A. Kumar, "Performance analysis of adaptive clipping technique for reduction of PAPR in Alamouti coded MIMO-OFDM systems," *Procedia Comp. Sci. 93*, 609 – 616, 2016.

[17] Abdullah and M. Hidayat, "SCS-SLM PAPR reduction technique in STBC MIMO-OFDM systems," in *Proc.7th IEEE ICCSCE, Penang, Malaysia,* 2017, pp. 104-109.

[18] Sandoval, G. Poitau, and F. Gagnon, "On optimizing the PAPR of OFDM signals with coding, companding, and MIMO," *IEEE Access*, vol. 7, pp. 24132-24139, 2019.

[19] B. Horvath and B. Botlik, "Optimization of tone reservation-based PAPR reduction for OFDM systems," *Radio Engineering*, vol. 26, no. 3, pp. 791-797, 2017.

[20] M. Habibi and M. Naeiny, "Selective tone reservation method for PAPR reduction in SFBC-OFDM systems," *J. of Commun. Eng.,* vol. 3, no. 2, pp. 109-122, 2014.

[21] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, 1998.

[22] Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband*, 1st ed., Elsevier Ltd, London, UK, 2007.

[23] J. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*, Pearson Education, Inc., USA, 2007.

[24] M. Sharif, M. Gharavi-Alkhansari, and B. H. Khalaj, "New results on the peak power of OFDM signals based on oversampling," in *Proc. IEEE ICC*, vol. 2, 2002, pp. 866–871.

[25] J. Tellado, "Peak to average power reduction for multicarrier modulation," Ph.D. dissertation, Dept. Elect. Eng., Stanford University, Stanford, CA, USA, 2000.

[26] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications,* 2nd ed., Springer-Verlag, New York, 2003.

[27] R. L. Burden and J. D. Faires, *Numerical Analysis*, Brooks/Cole, Cengage Learning, 9th ed., Boston, USA, 2011.

[28] L. Xin and W. Yi, "A new weighted tone reservation method for PAPR reduction in OFDM systems," *J. of Commun.*, vol. 9, no. 12, pp. 980-986, 2014.

[29] Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal," in *Proc. 2nd Eur. Conf. Satell. Commun.*, Liège, Belgium, Oct. 1991, pp. 179-184.

Stephen Kiambi received the B.Sc. degree in electrical and electronic engineering from University of Nairobi, Nairobi, Kenya, in 2000 and the M.Sc. degree in information and

communication systems from Hamburg University of Technology, Hamburg, Germany, in 2010. He is currently pursuing the Ph.D. degree in electrical and electronic engineering in the University of Nairobi, Kenya. His current research interests include signal processing in wireless and mobile communication networks, OFDM, and PAPR reduction in multicarrier systems.

Elijah Mwangi received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from University of Nairobi, Nairobi, Kenya, in 1978 and 1983, respectively, and the Ph.D. degree in electrical and electronic engineering from Loughborough University of Technology, Loughborough, UK, in 1987. Since 1981, he has been with the School of Engineering, University of Nairobi, Nairobi, Kenya, where he is currently a Professor. His current research interests include digital signal processing, digital image processing, OFDM and PAPR reduction in multicarrier systems.

George Kamucha received the B.Tech. degree in electrical and communications engineering from Moi University, Eldoret, Kenya, in 1990, M.Sc. degree in biomedical engineering from University of Aberdeen, Aberdeen, UK, in 1994, and the Ph.D. degree in electrical engineering from University of Kassel, Kassel, Germany, in 2003. Since 2010, he has been with the School of Engineering, University of Nairobi, Nairobi, Kenya, where he is currently a Senior Lecturer. His current research interests are in biomedical and communication systems.

Journal of Electrical Systems and Information Technology

Reducing PAPR of OFDM signals using a tone reservation method based on ℓ_∞-norm minimization

Stephen Kiambi[*] , Elijah Mwangi and George Kamucha

*Correspondence:
skiambi@uonbi.ac.ke

Department of Electrical and Information Engineering, School of Engineering, University of Nairobi, P.O. Box 30197-GPO, Nairobi, Kenya

Abstract

Orthogonal frequency division multiplexing (OFDM) continues to be the most preferred signal-multiplexing scheme for high-speed data communication. However, OFDM signals are known to have the problem of high peak-to-average power ratio (PAPR), especially when the number of subcarriers is large, which leads to nonlinear amplification in the high power amplifier and consequently to bit-error rate degradation and out-of-band radiation. In this paper, we propose a new optimal tone reservation method for reducing high PAPR in OFDM signals in order to avoid nonlinear amplification effects. The method employs Chebyshev-norm minimization to determine peak-reduction coefficients for OFDM signal. Simulation results show that the proposed method can achieve high PAPR reduction at the expense of a small loss in data rate and a slight increase in average transmit power. For example, with 4 out of 64 subcarriers reserved for peak-reduction coefficients, which represents 6.25% data-rate loss, the method can achieve 4.06 dB of PAPR reduction with only a 0.46 dB increase in average transmit power. Similarly, when 8 subcarriers or 12.5% of the total number of subcarriers are reserved, a PAPR reduction of 5.75 dB is achieved with a paltry 0.19 dB rise in transmit power.

Keywords: High power amplifier, Orthogonal frequency division multiplexing, Peak-to-average power ratio, Tone reservation

Introduction

Recently, OFDM is the most widely used transmission technique in high data-rate applications. For example, the transmission technique is used in digital audio broadcasting (DAB), digital video broadcasting (DVB), IEEE 802.11-based wireless local area network (WLAN), IEEE 802.16-based worldwide interoperability for microwave access (WiMAX), 4th and 5th generations of mobile communication network and is a candidate technology for 6G of the same [1–5].

The extensive use of OFDM technique is due to its several important advantages; among them are high spectral efficiency, simple receiver implementation and robustness against frequency-selective fading. The high spectral efficiency comes from the use of a large number of mutually orthogonal subcarriers. The design of the receiver is simple because only single-tap equalization is needed. This is made possible by the fact that

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 2 of 15

transmitted signals do not experience intersymbol interferences because of the use of sufficiently long symbol duration in conjunction with adequate time guard interval.

However, OFDM signals tend to have high PAPR that is caused by the operation of multiplexing many modulated signals. In some instances, the PAPR can reach unacceptable levels especially when large number of subcarriers is involved. A high PAPR leads to nonlinear amplification of signal by the high power amplifier (HPA) in the transmitter. This nonlinear amplification produces in-band and out-of-band radiations. The in-band radiations degrade the bit-error rate (BER), while the out-of-band radiations result in adjacent channel interferences. A simple way to avoid nonlinear amplification and the associated detrimental effects is to shift the operating point of the HPA away from the 1-dB compression point by a sufficient input back-off (IBO) depending on the expected PAPR of input signals.

However, when the HPA is provided with an IBO to force it to operate deep into the linear region, its power efficiency is reduced and thus it consumes more power. Low power efficiency requires a complex HPA design, which increases the cost of the transmitter. On the other hand, high power consumption leads to a significant reduction in lifetime of battery power in user terminals [6]. Therefore, a far much better solution is to reduce the PAPR to a suitable level before the processing of an OFDM signal in the HPA.

Recently, several PAPR reduction methods have been proposed in literature. Among them are those based on signal coding [7], clipping and filtering [8], companding [9], selective mapping [10], partial transmit sequence [11] and tone reservation [12]. PAPR reduction methods that reserve some tones for peak-cancelling signal have been found to be the most promising because they do not affect user data, and thus, the BER of the underlying system is maintained. In addition, such methods do not require transmission of side information to the receiver to aid in the recovery of user data.

The main objective of this paper is to propose an optimal tone reservation method of a good PAPR reduction performance that marginally increases the average transmission power. The method employs Chebyshev-norm approximation to find peak-reduction coefficients that yield a peak-reduction signal that closely estimates the desired signal. The proposed method can be used to reduce PAPR of signals in the current and future generations of communication networks in which OFDM technique is deployed. The performance of the proposed method is verified via simulation by the results of PAPR reduction and BER of OFDM system, which are also compared to those of other relevant and promising PAPR reduction methods.

The following is the organization of the rest of the paper. The "Related Work" section gives an overview, advantages and disadvantages of existing methods, which are relevant to this study. The section titled "PAPR in OFDM Signals" describes PAPR and its measurement. In the section "Methods", the proposed PAPR reduction algorithm is presented, while in the section headlined "Results and Discussion", simulation results and their analysis are given. Lastly, the "Conclusion" section summarises the paper.

Related work

This section gives an overview of some pertinent PAPR reduction methods and whose performances will be compared to that of the proposed method in this paper. Generally, tone reservation (TR) methods differ by the way the peak-reduction coefficients that

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 3 of 15

are used to modulate the reserved tones to produce peak-reduction signal are generated. The mode of generating peak-reduction coefficients comes with its own advantages and disadvantages, with the aim being to minimize the latter to allow for practical realization.

In [13], a tone reservation method, CF-TR, based on curve fitting was proposed. The method applies curve-fitting optimization technique on a signal referred to as clipping noise to find peak-reduction coefficients for OFDM signal. Although the method has good PAPR reductions, it has to perform the computationally intensive Moore–Penrose matrix inversion in every iteration and the resulting peak-reduced transmit signal has its average power significantly increased above that of the original OFDM signal.

Another tone reservation method, LSA-TR, proposed in [14] employs least squares approximation to find peak-reduction coefficients. Although the method converges fast and the increase in average transmit power is small, its PAPR reduction performance is very poor. A TR method, IVO-TR, based on machine learning feedforward neural network and initial value optimisation is proposed in [15]. The method pre-generates and stores all possible peak-reduction signals in a pre-work table based on the training targets generated by CC-TR method [16]. At runtime, an OFDM symbol is classified and a search is done in the table for an appropriate peak-reduction signal. Although the method attempts to reduce runtime complexity, its PAPR reduction is limited by the performance of the CC-TR method whose convergence rate is affected by initial conditions. In addition, the method requires long pre-work training time to generate near optimal peak-reduction signals and increases the average transmit power.

A scaling signal-to-clipping noise ratio tone-reservation method, referred to as SSCR-TR, is proposed in [17]. The peak-reduction signal is a time-domain kernel signal obtained by scaling down the clipping noise signal by an optimal scaling vector. The LSA algorithm together with peak-regeneration constraints is employed to find the optimal scaling vector. Although the method has fast convergence rate, it is still prone to peak re-growth and the PAPR reduction performance strongly depends on the clipping ratio employed.

Another tone-reservation scheme, ELM-TR, based on online sequential extreme learning machine with a single-hidden layer feedforward neural network is proposed in [18]. The method is trained on peak-cancelling signals generated by the CC-TR scheme. Due to the use of a single hidden layer, the training time is significantly reduced. However, its PAPR reduction capability depends on the performance of the CC-TR method. Furthermore, the training of the neural network is computationally intensive and requires big storage capacity due to massive input data and numerous trainable parameters.

PAPR in OFDM signals

The OFDM signal arises from the summation of N modulated signals. Each modulated signal is basically a subcarrier signal $e^{j2\pi f_k t}$ modulated by a data symbol $X(k)$. In the discrete-time domain, the complex baseband OFDM signal in one symbol duration can be expressed in the form:

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 4 of 15

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{j2\pi \frac{kn}{N}}, \quad n = 0, 1, \ldots, N-1 \tag{1}$$

The modulation symbols are obtained from binary phase-shift keying (BPSK) modulation or any M-ary quadrature amplitude modulation (M-QAM). Because (1) is similar to the inverse discrete Fourier transform (IDFT), OFDM is easily implemented using the well-known fast Fourier transform (FFT) algorithm.

Assuming a large N and that the modulated signals are statistically independent and identically distributed, then from the central limit theorem, both the real and imaginary parts of $x(n)$ have Gaussian distribution. Accordingly, the signal magnitudes $|x(n)|$ are Rayleigh-distributed, which implies that $x(n)$ can have large amplitudes well above the average value. This can in turn lead to nonlinear amplification of the large amplitudes in the HPA, thus giving rise to in-band and out-of-band radiations.

The level of peak power with reference to average power of the continuous-time OFDM signal can be estimated by the peak-to-average power ratio defined as

$$\text{PAPR}\{x(n)\} = \frac{\max_{0 \leq n \leq N-1} \left\{ |x(n)|^2 \right\}}{\text{E}\{|x(n)|^2\}} \tag{2}$$

where $E\{\}$ denotes the expectation operator. In order to avoid missing the highest peak of the continuous-time signal and, therefore, wrongly estimating the PAPR using (2), the discrete-time signal $x(n)$ should be sufficiently oversampled typically by a factor greater than 4 above the Nyquist rate [19].

The level of PAPR is indicated by the complementary cumulative distribution function (CCDF) [20], which is defined as the probability that the PAPR is above a specified threshold γ, i.e.

$$\Pr\left\{\text{PAPR}\{x(n)\} > \gamma\right\} = 1 - \left(1 - e^{-\gamma}\right)^N \tag{3}$$

where Pr{.} represents the probability operator. The CCDF is normally plotted against different threshold values and this produces a curve with a waterfall-like characteristic. Since the number of subcarriers, *N,* in (3) is known, when considering more than one plots of CCDFs, the difference between any two thresholds at the same CCDF value measures the level of PAPR reduction. Therefore, this measurement can be used to judge how well a proposed method reduces PAPR.

Methods

The proposed method utilises the concept of tone reservation [21] in which a smaller number of OFDM subcarriers, which were previously intended for the transportation of user data, are reserved to carry PAPR reduction coefficients. The reserved subcarriers are referred to as peak-reduction tones. Because the reserved subcarriers do not carry user data, a data-rate loss expressed as

$$R_f = \frac{L}{N} \tag{4}$$

is expected in a commrununication system employing a PAPR reduction method based on the tone reservation concept. In (4), L and N denote the number of reserved subcarriers and total number of subcarriers in one OFDM symbol, respectively.

In order to minimize the data-rate loss in (4), the number of reserved subcarriers should be set much smaller than the total numbers of subcarriers, i.e. $L \ll N$. In addition, to avoid distorting the user data due to the introduction of peak-reduction coefficients, the peak reduction tones and the data-bearing subcarriers are made to occupy two disjoint frequency subspaces in every OFDM symbol. Thus, in the reserved subcarrier positions, there are no modulating data symbols, i.e. they are set to zero. Likewise, in the locations allocated for subcarriers for user data, the peak-reduction coefficients are set to zero. At the receiver, the disjoint frequency subspaces allow the transmitted symbols to be recovered from the FFT output without distortion by considering only the locations of data-bearing subcarriers.

The tone reservation concept that has just been described is illustrated in Fig. 1, where $X(k)$ and $C(k)$ are the modulating data symbols and peak-reduction coefficients, respectively. As shown in the figure, after the inverse FFT (IFFT) operation, the resulting combined signal $\mathbf{s} = \mathbf{x} - \mathbf{c}$ has a reduced peak amplitude and hence lower PAPR than the original time signal $\mathbf{x}$. After reduction in PAPR, the combined signal is converted into an analogue signal by a digital-to-analogue converter (DAC) then up-converted to radio frequency f_c before being passed to HPA for power amplification.

The generation of a low PAPR transmit signal $s(n)$ using the tone reservation concept can be described by the equation

$$
\begin{aligned}
s(n) &= x(n) - c(n) \\
&= x(n) - \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} C(k) e^{j2\pi \frac{kn}{N}}
\end{aligned}
\tag{5}
$$

or in matrix notation as

$$
\mathbf{s} = \mathbf{x} - Q\mathbf{C}
\tag{6}
$$

Here, $Q \in \mathbb{C}^{N \times N}$ is the IDFT matrix and contains the elements given by $(1/\sqrt{N}) \exp(j2\pi kn/N)$, $\mathbf{s} = [s(0), s(1), \ldots, s(N-1)]^{\mathrm{T}}$ is the peak-reduced signal vector, $\mathbf{x} = [x(0), x(1), \ldots, x(N-1)]^{\mathrm{T}}$ contains samples of original OFDM signal and $\mathbf{C} = [C(0), C(1), \ldots, C(N-1)]^{\mathrm{T}}$ is a frequency-domain vector of the peak-reduction coefficients.

If we let $\hat{\mathbf{C}} \in \mathbb{C}^L$ denote the vector containing the L nonzero elements of $\mathbf{C}$, the peak-reduction signal $\mathbf{c}$ can be expressed as

$$
\mathbf{c} = \hat{Q}\hat{\mathbf{C}}
\tag{7}
$$

where the IFFT submatrix $\hat{Q} \in \mathbb{C}^{N \times L}$ is made up of L columns of Q corresponding to the locations of the reserved subcarriers.

Proposed method

Ideally, the peak-reduction signal can be considered to consist of samples of the difference signal between the original OFDM samples and the clipped

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol* (2022) 9:12

Page 6 of 15

Fig. 1 Tone reservation concept

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol* (2022) 9:12

Page 7 of 15
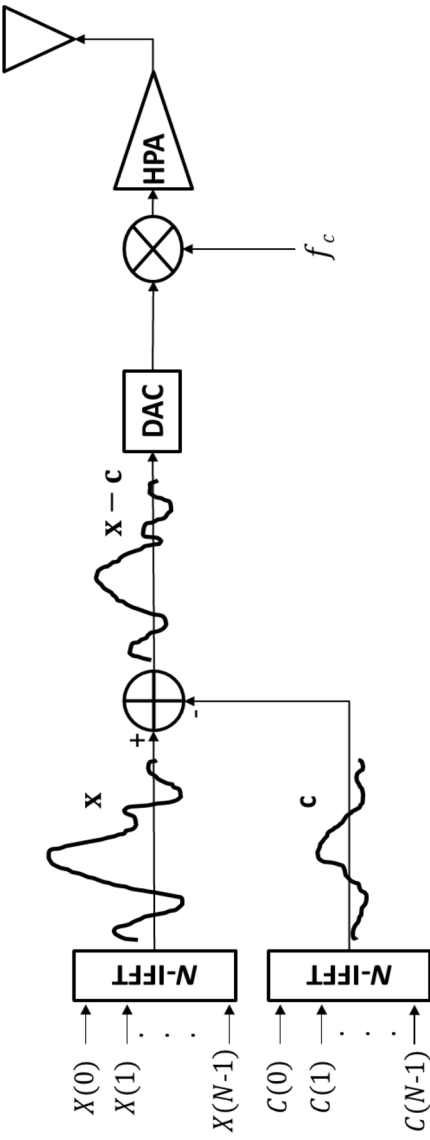
version. The clipped signal is derived by clipping the OFDM signal at a threshold x_{th}. Analytically, the desired peak-reduction signal can be expressed as a vector $\mathbf{d} = [d(0), d(1), \ldots, d(N-1)]^T$ with components given by

$$
d(n) = \begin{cases} \frac{x(n)}{|x(n)|}(|x(n)| - x_{th}), & |x(n)| > x_{th} \\ 0, & |x(n)| \leq x_{th} \end{cases} \tag{8}
$$

The saturation point of the HPA can be used to determine the clipping threshold. Given the maximum PAPR allowed in an OFDM-based communication system, the clipping threshold can be found from (2) as follows:

$$
x_{th} = \sqrt{\text{PAPR}_{max}E\{|x(n)|^2\}} \tag{9}
$$

For effective reduction in PAPR, the actual peak-reduction signal should be close or equal to the desired signal. In other words, the residual error

$$
\mathbf{r} = \hat{Q}\hat{\mathbf{C}} - \mathbf{d} \tag{10}
$$

between the two signals should be as small as possible if not equal to zero. This can be achieved through the minimization of the residual error using an appropriate norm to measure the error level. Noting that the highest peak in $\mathbf{d}$ is the main cause of high PAPR in signal $\mathbf{x}$, it is preferable to minimize the Chebyshev (ℓ_∞) norm of the residual error in order to ensure that the largest error magnitude is minimized.

From the foregoing discussion, the problem of minimizing the residual error can be formulated as the following Chebyshev approximation problem [22]:

$$
\text{minimize} \quad ||\hat{Q}\hat{\mathbf{C}} - \mathbf{d}||_\infty \tag{11}
$$

where $||.||_\infty$ denotes the ℓ_∞-norm.

The Chebyshev approximation problem (11) has no closed form solution but can be solved after casting it into the following linear program:

$$
\begin{aligned}
\text{minimize} \quad & t \\
\text{subject to} \quad & \hat{Q}_n^T\hat{\mathbf{C}} - t \leq d(n) \\
& -\hat{Q}_n^T\hat{\mathbf{C}} - t \leq -d(n)
\end{aligned} \tag{12}
$$

in which $t \in \mathbb{R}$ and $\hat{\mathbf{C}} \in \mathbb{C}^L$ are the optimization variables and $\hat{Q}_n \in \mathbb{C}^L$ and $d(n) \in \mathbb{C}$, for $n = 0, 1, 2, \ldots, N-1$, are the problem parameters. Note that, $\hat{Q}_n$ is a column vector equal to the transpose of the nth row of matrix $\hat{Q}$.

After solving (12), the time-domain peak-reduction signal and the transmit signal are obtained using (7) and (6), respectively. The main steps of the proposed algorithm are listed in Table 1.

The effectiveness of the proposed method in terms of peak-power reduction can be measured by comparing the level of the maximum power of the peak-reduced signal to that of the average power of the original OFDM signal. In order to do such a comparison, the peak-to-average power ratio of the peak-reduced signal is defined as follows:

215

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 8 of 15

Table 1 Proposed TR algorithm

i	Set number of subcarriers N, allowed data rate loss R_f and maximum allowed $PAPR_{max}$
ii	Generate OFDM signal **x** and calculate PAPR
iii	If calculated PAPR < $PAPR_{max}$, transmit **x** and terminate program, else got to step (iv)
iv	Calculate clipping threshold $x_{th} = \sqrt{PAPR_{max}E\{|x(n)|^2\}}$
v	Generate desired peak-reduction signal **d**
vi	Generate IFFT submatrix $\hat{Q} \in \mathbb{C}^{N \times L}$
vii	Use interior-point method to minimize $||\hat{Q}\hat{\mathbf{C}} - \mathbf{d}||_\infty$ and solve for PRCs vector $\hat{\mathbf{C}}$
viii	Compute peak-reduction vector $\mathbf{c} = \hat{Q}\hat{\mathbf{C}}$
ix	Compute $\mathbf{s} = \mathbf{x} - \mathbf{c}$ and transmit
x	End

$$PAPR\{s(n)\} = \frac{\max_{0 \leq n \leq N-1} \left\{ |x(n) - c(n)|^2 \right\}}{E\{|x(n)|^2\}} \tag{13}$$

One drawback of the just proposed method is that it increases the transmit power, i.e. the average power of signal **s** will be higher than the average power of signal **x**—a problem that is inherent to all methods based on the tone reservation concept. However, because the desired peak-reduction signal has most of the components equal to zero, the peak-reduction signal resulting from the Chebyshev approximation is expected to have most of its samples very small, close zero, and thus the increase in the average power will be small.

Computational complexity

Depending on the size of the problem in (12), it can be solved using one of the three linear programming algorithms [23], namely interior-point, active-set and simplex algorithms to find the peak-reduction coefficients. The proposed method will employ the interior-point method to solve (12). Since there is no simple analytical formula for the solution to a linear program, the required number of arithmetic operations cannot exactly be established.

However, in practice, the interior-point method is known to have a complexity $O(NL^2)$, where N and L are number of rows and columns of matrix $\hat{Q}$, respectively [22]. Additionally, since $\hat{Q}$ is a submatrix of the well-structured IDFT matrix, the linear program in (12) can be solved with complexity $O(N \log_2 N)$ [24].

The complexity of the proposed method and those in the related work section are listed in Table 2. In the table, I, I_{sscr}, I_{cf} and I_{ls} denote the respective number of iterations required to find the peak-reduction coefficients at runtime for the proposed method, SSCR-TR, CF-TR and LSA-TR methods. For the ELM-TR and IVO-TR methods, N_s, N_i, N_1 and N_o denote the size of the training-data set and the number of neurons in the input, hidden and output layers, respectively.

Results and discussion

The proposed method was employed to reduce PAPR in OFDM systems with $N = 64$ subcarriers. The simulation parameters, which are listed in Table 3, were purposely chosen to help ascertain the performance of the method in terms of PAPR reduction and BER degradation and to allow for comparison with other methods. The problem in (12)

Table 2 Computational complexity comparison

Method	Training complexity	Runtime complexity
Proposed method	None	$I \times O(N \log_2 N)$
ELM-TR	$O(N_i N_1 + N_s N_1 N_o)$	$O(N_i N_1 + N_1 N_o)$
SSCR-TR	None	$I_{sscr} \times O(N + N \log_2(N))$
IVO-TR	$N_s \times O(N_i N_1 + N_1 N_o + M N_o)$	$O(N_i N_1 + N_1 N_o)$
CF-TR	None	$I_{cf} \times O(N \log_2 N)$
LSA-TR	None	$2 I_{ls} \times O(N \log_2 N)$

Table 3 Simulation parameters

Total subcarriers, N	64
Subcarrier modulation	QPSK
Number of OFDM symbols	10^4
Oversampling factor	4
Number of reserved subcarriers, L	4, 8
Power amplifier model	Rapp model, $p = 2$, IBO $= 8$ dB

was first re-formulated in MATLAB to take into account the real and imaginary parts of the inequality constrains. The interior-point method was then employed to solve the linear program for the peak-reduction coefficients.

To have a good estimate of the continuous-time PAPR, all the discrete-time signals were oversampled by a factor of 4. In addition, the Rapp's model of HPA [25] was used with the smoothness parameter p set at 2 and the IBO at 8 dB, which is approximately 1 dB above the PAPR of peak-reduced signals at the CCDF $= 10^{-3}$. This IBO setting ensures that the percentage of signal amplitudes clipped by HPA is less than 1%. Additionally, since the type of subcarrier modulation does not affect the level of PAPR reduction, only the QPSK modulation was used during the simulations.

In Fig. 2, the PAPR reduction performance of the proposed method is shown for two cases of 4 and 8 reserved subcarriers out of the 64 subcarriers. The two cases give a data-rate loss of 6.25% and 12.5%, respectively. The PAPR reduction at CCDF $= 10^{-3}$ is 4.06 dB and 5.75 dB for 4 and 8 reserved subcarriers, respectively. This shows that the proposed method can achieve high PAPR reductions with only a small percentage of the total number of subcarriers reserved for peak-reduction coefficients. It can also be observed that the reduction in PAPR increases with the number of reserved subcarriers. However, since the PAPR reduction is at the expense of a data-rate loss due to the reserved subcarriers, a compromise between the two is necessary depending on the requirements of the communication system.

The PAPR reduction for the case of 4 reserved subcarriers was used to compare the performance of the proposed method to ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR methods. For the IVO-TR and ELM-TR methods, $N_O = 400$, $N_1 = 1000$, $N_S = 10^5$ and the size of the test-data set is 10^4. In addition, the training target is 100 iterations of CC-TR method. The number of iterations for the CF-TR, LSA-TR and SSCR-TR methods is 2, 3 and 5, respectively.

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol* (2022) 9:12

Page 10 of 15
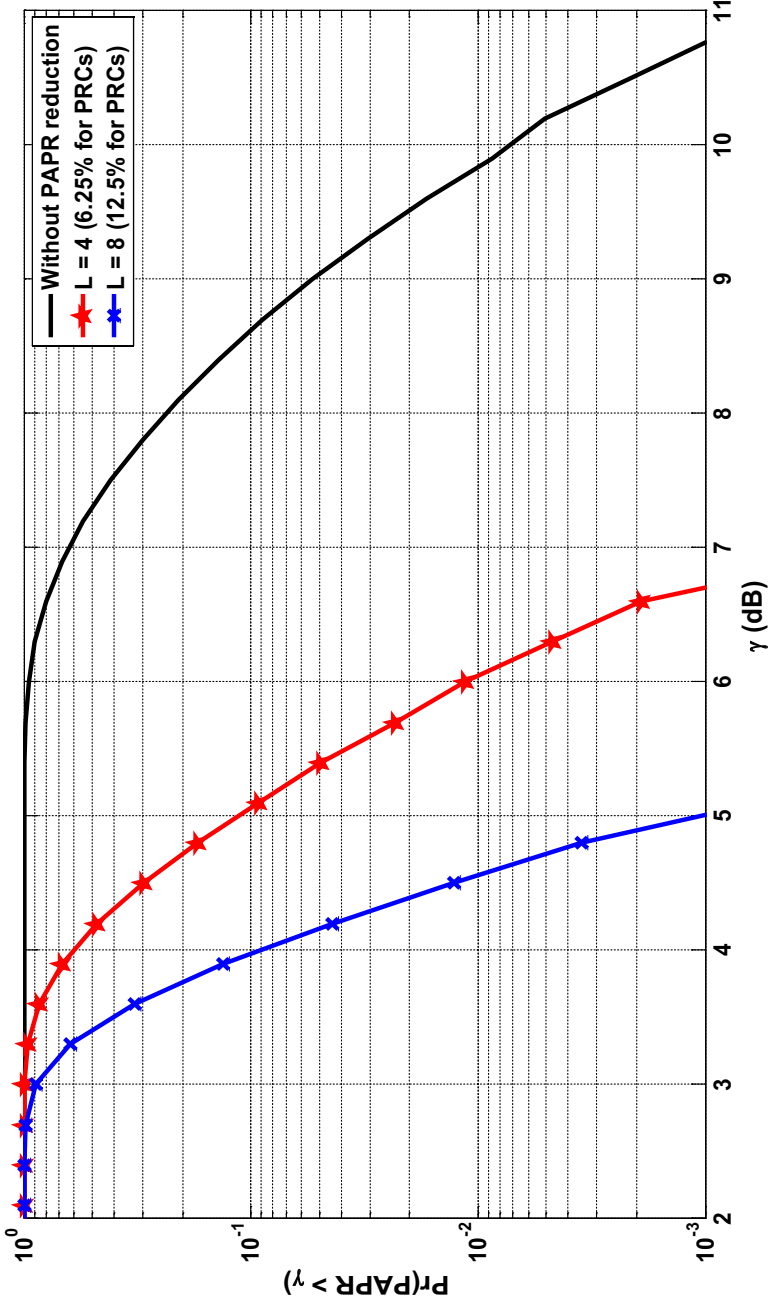
Fig. 2 PAPR reduction with 4 and 8 reserved subcarriers
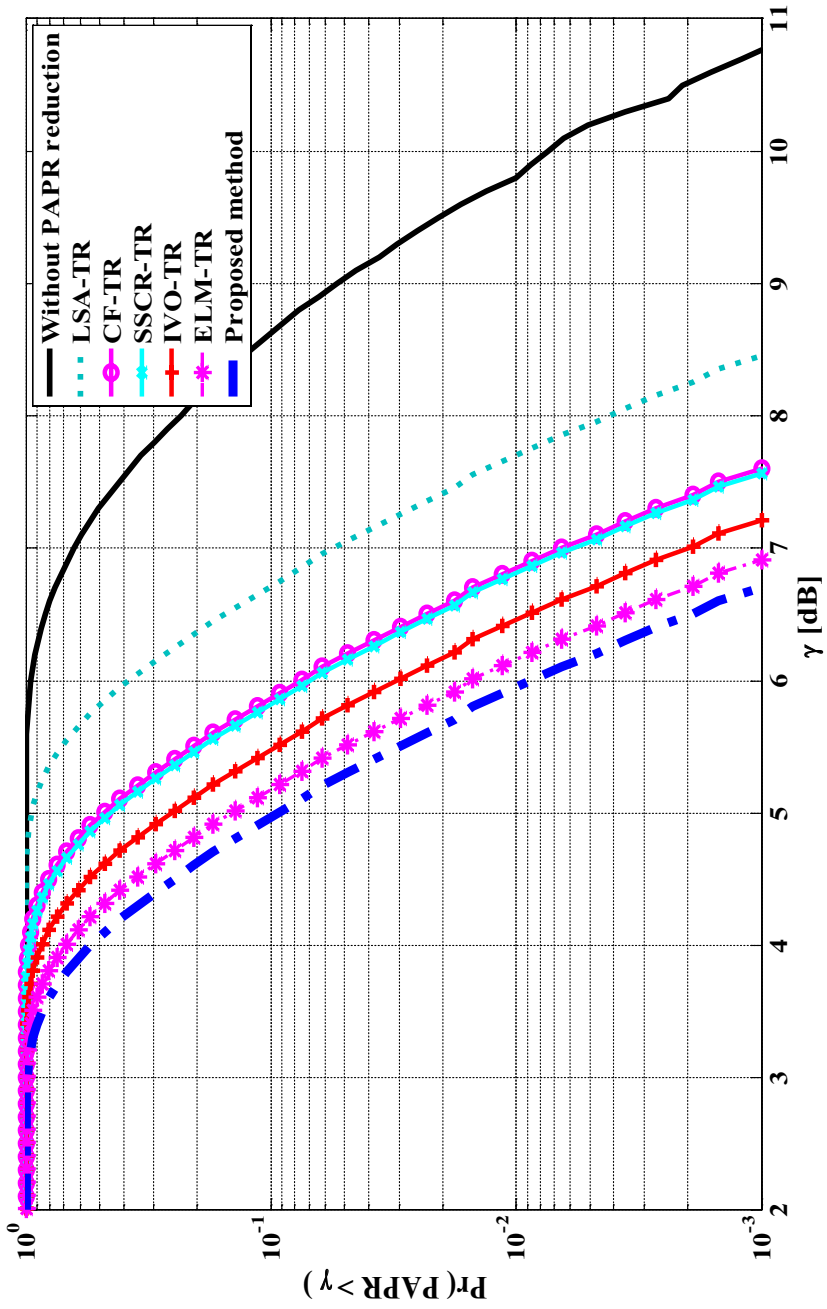
Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*　　*(2022) 9:12*　　　　　　　　　　Page 11 of 15

Fig. 3 PAPR reduction performance comparison

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   *(2022) 9:12*   Page 12 of 15

Fig. 4 BER performances over AWGN channel

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 13 of 15
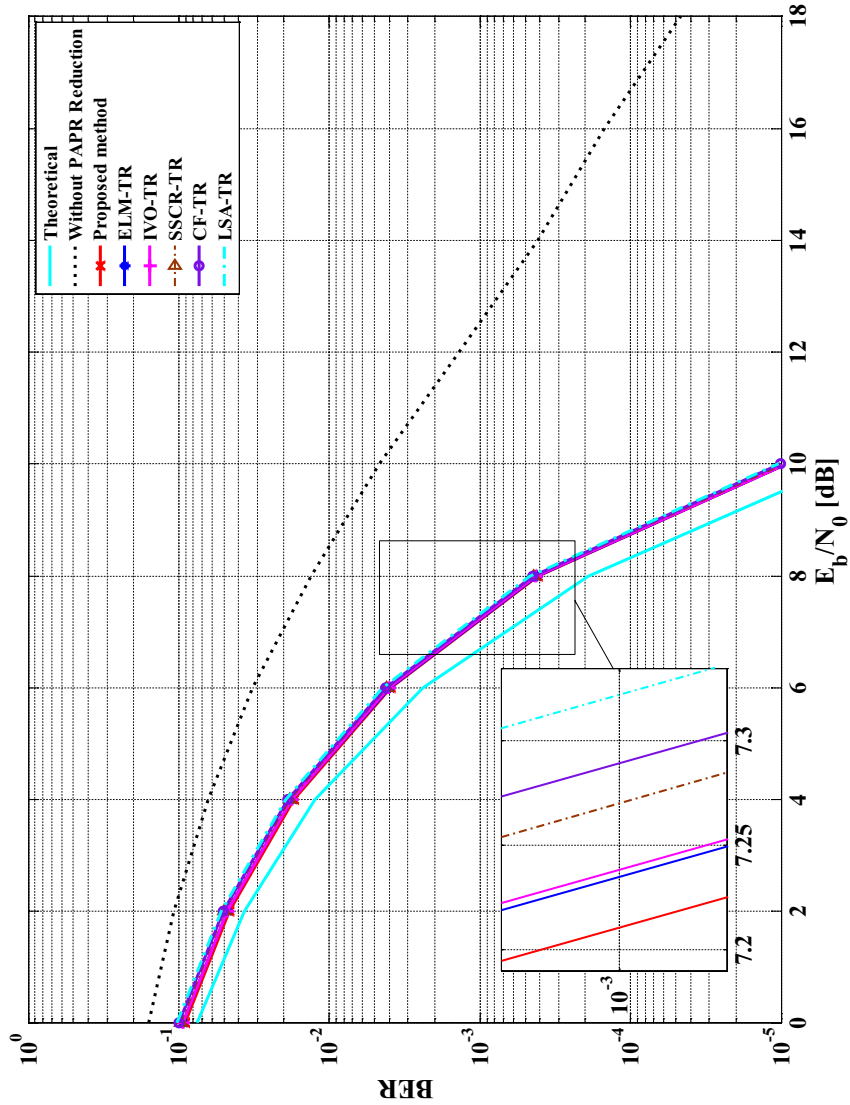
Table 4 PAPR reduction and average power increase

Method	PAPR reduction [dB]	Power increase [dB]
Proposed LP-TR	4.06	0.46
ELM-TR	3.85	0.59
SSCR-TR	3.19	0.25
IVO-TR	3.55	0.57
CF-TR	3.16	0.84
LSA-TR	2.31	0.30

Table 5 Required E_b/N_o at BER$=10^{-3}$

Method	E_b/N_o (dB)
Proposed method	7.21
ELM-TR	7.23
SSCR-TR	7.27
IVO-TR	7.24
CF-TR	7.29
LSA-TR	7.32

The CCDF curves for the six methods are depicted in Fig. 3, and the results for PAPR reductions at CCDF$=10^{-3}$ and the average power increase are summarized in Table 4. From these results, it can be observed that the proposed method has better PAPR reduction performance than the rest. At the CCDF$=10^{-3}$, the proposed method exhibits a higher PAPR reduction than the ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR method by 0.21, 0.87, 0.51, 0.90 and 1.75 dB, respectively.

For the BER performance, the results for the six methods are given in Fig. 4. The BER performances are for the cases of transmission of amplified peak-reduced signals over additive white Gaussian noise (AWGN) channels. The curve labelled *Theoretical* gives the lower limit or the best-expected performance as it corresponds to the performance given by the BER formula of the QPSK modulation. The curve labelled *Without PAPR Reduction* is for the case when the OFDM signals were amplified through the HPA with the IBO$=0$ dB and therefore is the worst expected BER performance.

The required SNR per bit, i.e. E_b/N_o, at BER$=10^{-3}$ for the methods is presented in Table 5. As it is expected of methods based on tone-reservation concept, all the six methods have approximately the same BER performance. However, due to the setting of the IBO and the level of PAPR reduction, the proposed LP-TR method has a slightly better BER performance than ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR method by 0.02, 0.06, 0.03, 0.08 and 0.11 dB, respectively.

Conclusions

In this work, we have proposed a new optimal tone reservation method for reducing PAPR of OFDM signals. The method first generates a desired peak-reduction signal, and then, using linear programming of the Chebyshev approximation problem, it designs the actual peak-reduction signal, while utilising only a small number of reserved subcarriers for peak-reduction coefficients.

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   (2022) 9:12   Page 14 of 15

With a small number of reserved subcarriers, the proposed method achieves significant PAPR reductions, e.g. with 4 and 8 reserved subcarriers out of a total of 64, 4.06 and 5.75 dB of PAPR reductions are attained, respectively. In addition, the method only causes only a small increase in transmit power, e.g. for the case of 4 reserved subcarriers, the power increase is 0.46 dB. Additionally, the method does not affect the BER of the underlying OFDM system.

In comparison with five other methods, namely ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR method, the proposed method has better PAPR reduction performance. At $CCDF = 10^{-3}$ for the case of 4 reserved subcarriers out of 64, the proposed method achieves 0.21, 0.87, 0.51, 0.90 and 1.75 dB of PAPR reduction above the ELM-TR, SSCR-TR, IVO-TR, CF-TR and LSA-TR method, respectively.

In future work, the proposed method can be employed to reduce PAPR in an OFDM system employing adaptive modulation and coding during one symbol duration. Additionally, the peak-reduction signals generated by the proposed method can be used as training targets for a PAPR reduction method based on machine learning. Another future research is to develop a faster algorithm than the interior-point algorithm to solve the formulated Chebyshev approximation problem in this paper and thereby reduce convergence time and computational complexity.

Abbreviations

AWGN	Additive white Gaussian noise
BER	Bit-error rate
SNR	Signal-to-noise ratio
CCDF	Complementary cumulative distribution function (CCDF)
CF	Curve fitting
HPA	High power amplifier
IBO	Input back-off
IVO	Initial value optimization
OFDM	Orthogonal frequency division multiplexing
LSA	Least squares approximation
PAPR	Peak-to-average power ratio
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase-shift keying
TR	Tone reservation

Declarations

Appendix G: Sixth Publication

Kiambi *et al. Journal of Electrical Systems and Inf Technol*   *(2022) 9:12*   Page 15 of 15

References

1. Prasad R (2004) OFDM for wireless communications systems. Artech House, London
2. Bahai R, Saltzberg R, Ergen M (2004) Multi-carrier digital communications: theory and applications of OFDM. Springer, New York
3. ETSI EN 302 755 V.1.3.1 Digital video broadcasting (DVB); frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). https://www.dvb.org/standards
4. NTT DOCOMO Inc. (2020) White paper 5G evolution and 6G, January 2020
5. Juwono F, Reine R (2021) Future OFDM-based communication systems towards 6G and beyond: machine learning approaches. Green Intell Syst Appl 1(1):19–25
6. Louët Y, Palicot J (2008) A classification of methods for efficient power amplification of signals. Ann Telecom 63(7–8):351–368
7. Takizawa S, Ochiai H (2019) PAPR reduction of OFDM with trellis shaping based on p-norm minimization. IEEE Wirel Commun Lett 8(4):988–991
8. Tran V-N, Dang T-H (2021) New clipping-and-filtering method for peak-to-average power ratio reduction in OFDM. In: 2021 international conference engineering and telecommunication (En &T), 2021, pp 1–5
9. Liu K, Cui X, Xing Z, Liu Y (2022) Generalized Continuous Piecewise Linear Companding Transform Design for PAPR Reduction in OFDM Systems. IEEE Trans Broadcast. https://doi.org/10.1109/TBC.2022.3171134
10. Valluri SP, Kishore V, Vakamulla VM (2020) A new selective mapping scheme for visible light systems. IEEE Access 8:18087–18096
11. Musafer H, Faezipour M (2021) A novel partitioning scheme for partial transmit sequence method. In 2021 IEEE 12th annual ubiquitous computing, electronics & mobile communication conference (UEMCON), 2021, pp 121–125
12. Liu Z, Liu W, Hu L, Zhang L, Yang F (2021) A low complexity improved tone reservation method based on ADMM for OFDM systems' PAPR reduction. In 2021 13th international conference on wireless communications and signal processing (WCSP), 2021, pp 1–5
13. Jiang T, Ni C, Xu C, Qi Q (2014) Curve fitting based tone reservation method with low complexity for PAPR reduction in OFDM systems. IEEE Commun Lett 18(5):805–808
14. Li H, Jiang T, Zhou Y (2011) An improved tone reservation scheme with fast convergence for PAPR reduction in OFDM systems. IEEE Trans Broadcast 57(4):902–906
15. Li H, Wei J, Jin N (2019) Low-complexity tone reservation scheme using pre-generated peak-canceling signals. IEEE Commun Lett 23(9):1586–1589
16. Gatherer A, Polley M (1997) Controlling clipping probability in DMT transmission. In: Proceedings of 31st Asilomar conference on signals, systems and computers, vol 1, November 1997, pp 578–584
17. Wang J, Lv X, Wu W (2019) SCR-based tone reservation schemes with fast convergence for PAPR reduction in OFDM system. IEEE Wirel Commun Lett 8(2):624–627
18. Li Z, Jin N, Wang X, Wei J (2021) extreme learning machine-based tone reservation scheme for OFDM systems. IEEE Wirel Commun Lett 10(1):30–33
19. Sharif M, Gharavi-Alkhansari M, Khalaj BH (2002) New results on the peak power of OFDM signals based on oversampling. In: Proceeding of IEEE ICC, vol 2, 2002, pp 866–871
20. Cho Y, Kim J, Yang W, Kang C (2010) MIMO-OFDM wireless communications with MATLAB. Wiley, Singapore
21. Tellado J (2000) Peak to average power reduction for multicarrier modulation, Ph.D. dissertation, Department of Electrical Engineering, Stanford University, Stanford, CA, USA, 2000
22. Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, New York
23. Messac A (2015) Optimization in practice with MATLAB for engineering students and professionals. Cambridge University Press, New York
24. Boyd S, Vandenberghe L, Grant M (1994) Efficient convex optimization for engineering design. In proceedings of IFAC symposium on robust control design, Rio De Janeiro, Brazil, September 1994
25. Rapp (1991) Effects of HPA nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal. In: Proceeding of 2nd European conference on satellite communications, Liège, Belgium, October 1991, pp 179–184

Publisher's Note

Appendix H: MATLAB Programs

This appendix presents the programs that were developed in MATLAB for generating some of the figures in this thesis and for the simulation of the proposed methods. The following are the main functions of the programs:

i) The *OFDMSignalDistribution.m* function that generates OFDM signal, computes and plots the probability density function of the amplitudes.

ii) The *plotCCDF_Diff_N.m* function that generates OFDM signals with different number of subcarriers and for each signal computes and plots the CCDF of PAPR.

iii) The *trSOCP.m* function that implements the proposed SCOP-TR method for reduction of PAPR of OFDM signals.

iv) The *trLpCheby.m* function that implements the proposed LP-TR method for reduction of PAPR of OFDM signals.

v) The *trIRLS.m* function that implements the proposed IRLS-TR method for reduction of PAPR of OFDM signals.

vi) The *lcsaMethod().m* function that realizes the proposed LCSA method for reduction of PAPR of OFDM signals.

vii) The *lcasmMIMO.m* function that implements the proposed LCASM method for reducing PAPR of MIMO-OFDM signals.

H.1 Program to Illustrate Distribution of OFDM Signal Amplitudes

```matlab
% ----------------------------------------------------------------------%
% Purpose: To plot the distribution of OFDM signal amplitudes; real and  %
%          imaginary part, and the combination                           %
% Author:  Stephen Kiambi                                                %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)      %
%-----------------------------------------------------------------------%


function OFDMSignalDistribution_new()
    clc; clear all; close all; clf;
    %% Parameters
    Nfft=16; %FFT size
    nBit=2; %No. of bits per symbol
    Fs=16; %oversampling factor
    NF=Nfft*Fs; %Length of oversampled signal
    T=1/NF; %Plot time interval
```

```matlab
t = 0:T:1-T; %time-variable
nHists=1e3;%No. of histograms
nBins=30; %No. of bins


%% Generate OFDM time-signal
[X,modTyp] = mapperSymb(nBit,Nfft); %Generate 16 QPSK modulation symbols
X(1) = 0; %Set DC subcarrier to zero
for kk = 1 : Nfft % Generate 16 subcarrier time-signals
    if kk <= Nfft/2,
        x_sig = ifft([zeros(1, kk-1) X(kk) zeros(1, NF-kk+1)],NF);
    else
        x_sig = ifft([zeros(1,NF-Nfft + kk-1) X(kk) zeros(1,Nfft-kk)],NF);
    end
    xRe(kk, :) = real(x_sig); % Real parts of the 16 signals
    xIm(kk,:) = imag(x_sig); % Imaginary part of the 16 signals
end
sumxRe = sum(xRe); %Combine the real parts of the 16 signals
sumxIm = sum(xIm); %Combine the imaginary parts of the 16 signals


%% Plotting the real, imaginary and their combination
figure(1)
subplot(3,1,1)
plot(t,xRe,'k:')


hold on;
fig_1=plot(t,sumxRe,'b');
ylabel('x_{R}(t)','FontSize',14,'FontWeight','bold');
set(fig_1,'LineWidth',2.5);
title([modTyp ', N=' num2str(Nfft)]);
set(gca,'FontSize',14,'FontWeight','bold');


subplot(3,1,2)
plot(t,xIm,'k:');
hold on,
fig_2=plot(t,sumxIm,'b');
ylabel('x_{I}(t)','FontSize',14,'FontWeight','bold');
set(fig_2,'LineWidth',2.5);
set(gca,'FontSize',14,'FontWeight','bold');


subplot(3,1,3)
fig_3 = plot(t, abs(sumxRe+1i*sumxIm),'b');
set(fig_3,'LineWidth',2.5);
hold on;
ylabel('|x(t)|','FontSize',14,'FontWeight','bold');
xlabel('t','FontSize',14,'FontWeight','bold');
```

225

```matlab
fig_d = plot (t, ones(1,length(t))*mean(abs(sumxRe+1i*sumxIm)),'r-');
set(fig_d,'LineWidth',2.5);
hold off
legend('|x(t)|','Average of |x(t)|')
set(gca,'FontSize',14,'FontWeight','bold');


clear('xI'), clear('xQ')


%% Generate the distribution of the real, imaginary and combined amplitudes
for m = 1:nHists
    [X,modTyp] = mapperSymb(nBit,Nfft); %Generate 16 QPSK symbols
    X(1)=0; %Set DC subcarrier to zero
    for kk = 1:Nfft
        if (kk<= Nfft/2)
            x_sig=ifft([zeros(1,kk-1) X(kk) zeros(1,NF-kk+1)],NF);
        else
            x_sig=ifft([zeros(1,NF-Nfft/2+kk-Nfft/2-1) X(kk) zeros(1,Nfft-kk)],NF);
        end
        xRe(kk,:) = real(x_sig);
        xIm(kk,:) = imag(x_sig);
    end


    histRe(NF*(m-1)+1:NF*m)=sum(xRe); %Histogram for real part
    histIm(NF*(m-1)+1:NF*m)=sum(xIm); %Histogram for imaginary part
end


%% Plot the distributions of the real, imaginary and combined amplitudes
figure(2)
subplot(3,1,1)
[xRd,bins]=hist(histRe,nBins);
bar(bins,xRd/sum(xRd),'k');
title([modTyp ', N=' num2str(Nfft)]);
ylabel('pdf of x_{R}(t)','FontSize',14,'FontWeight','bold');
set(gca,'FontSize',14,'FontWeight','bold');


subplot(3,1,2)
[xId,bins]=hist(histIm,nBins);
bar(bins,xId/sum(xId),'k');
ylabel('pdf of x_{I}(t)','FontSize',14,'FontWeight','bold');
set(gca,'FontSize',14,'FontWeight','bold');


subplot(3,1,3)
[x_combd,bins]=hist(abs(histRe+1i*histIm),nBins);
bar(bins,x_combd/sum(x_combd),'k');
```

```matlab
        ylabel('pdf of |x(t)|','FontSize',14,'FontWeight','bold');
        xlabel('x_{0}','FontSize',14,'FontWeight','bold');
        set(gca,'FontSize',14,'FontWeight','bold');
end


%Generate Modulation Symbols
function [modSymbols,modTyp] = mapperSymb(nBit,N)
    modOrder=2^nBit; % Number Symbols digital Modulation


    if nBit == 1
        modTyp='BPSK';


        normEs=1;
        modObject=modem.pskmod('M',modOrder);


    elseif nBit == 2
        modTyp='QPSK';


        normEs=1;


        %Create QPSK modulator
        modObject=modem.pskmod('M',modOrder, 'PhaseOffset',pi/4,'SymbolOrder','gray');


    else
        modTyp=[num2str(2^nBit) 'QAM'];


        %Es is the average energy of modulation symbols
        Esym=1;


         %factor to normalize symbols
        normEs=sqrt(3*Esym/(2*(modOrder-1)));


        %Create M-ary QAM Gray encoded modulator
        modObject=modem.qammod('M',modOrder,'SymbolOrder','gray');
    end


    %Generate N random symbols
    mod_int= randi([0,modOrder-1],1,N);
    modSymbols = normEs*modulate(modObject,mod_int);
end
```

H.2 Program to Illustrate Increase of PAPR with Number of Subcarriers

```matlab
% ----------------------------------------------------------------------%
% Purpose: To plot the PAPR CCDF of OFDM signal for different number     %
%          of subcarriers                                                %
% Author: Stephen Kiambi                                                 %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)      %
%-----------------------------------------------------------------------%

function plotCCDF_Diff_N()
    clc; clear all; close all; clf;
    %% Parameters
    Nsub=2.^(6:10); %No. of subcarriers
    nBits=2; %No. of bits per symbol
    M=2^nBits; %Modulation size
    Nsym=1e4; %Frame size or number of OFDM symbols
    ydB=4:0.1:12; % PAPR threshold variable
    n_ydB=length(ydB);%No. of points on the horizontal axis
    %
    %% CCDF equation (3.43)
    formulaCCDF=inline('1-((1-exp(-y.^2/(2*sigma))).^N)','N','sigma','y');


    %% Generation of CCDFs
    for k = 1:length(Nsub)
        Ns = Nsub(k);
        xSig = zeros(Nsym,Ns);
        for i=1:Nsym
            X=mapperSymb(nBits,Ns);
            xSig(i,:)=ifft(X,Ns)*sqrt(Ns);
            PAR(i)=paprSig(xSig(i,:)); %Store PAPR values
        end
        sigma = mean(mean(abs(xSig)))^2/(pi/2); %Standard deviation
        formulatedCCDF=formulaCCDF(Ns,sigma,10.^(ydB/20));


        for n=1:n_ydB,
            simulatedCCDF(n)=sum(PAR>ydB(n))/Nsym;
        end


        % Plot CCDFs from the formula and simulation
        fig_a=semilogy(ydB,formulatedCCDF,'r-'); hold on; grid on;
        set(fig_a,'LineWidth',2.0)
        fig_b=semilogy(ydB(1:3:end),simulatedCCDF(1:3:end),'b:*');
        set(fig_b,'LineWidth',2.0)
    end
    axis([ydB([1 length(ydB)]) 1e-3 1]);
    title('Simulation of CCDF for OFDM Signals with Different Numbers of Subs');
    xlabel('\gamma [dB]','FontSize',14,'FontWeight','bold');
    ylabel('CCDF','FontSize',14,'FontWeight','bold');
```

228

```matlab
        legend('By CCDF formula','By simulation');
        set(gca,'FontSize',14,'FontWeight','bold');
end


%Calculate and return the PAPR of signal x in dB
function paprdB = paprSig(xSig)
lenSig=length(xSig);
xRe=real(xSig);
xIm=imag(xSig);
pwrSig = xRe.*xRe + xIm.*xIm;
avgPwr = sum(pwrSig)/lenSig;
peakPwr = max(pwrSig);
paprdB = 10*log10(peakPwr/avgPwr);
```

H.3 Program for SOCP-TR Method

```matlab
% ------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of OFDM signals using the proposed SOCP-TR method%
% Author:  Stephen Kiambi                                                 %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)       %
%-------------------------------------------------------------------------%

% The proposed method employs second-order cone programming (SOCP) to minimize
% the infinity norm, i.e., min max |Q_hat*C_hat+x|                              %
% The trSOCP.m function generates the peak-reduction coefficient vector C
% for OFDM signal (x) by calling the cOptSocp.m function.
% The function cOptSocp.m is formulated in CVX (a MATLAB software for
% disciplined convex programming) syntax.
% C is converted to peak-reduction time-domain signal c.
% The data vetor, X, and peak-reduction coefficient vector, C, occupy
% disjoint frequency subspaces.
% X has N-L nonzero values while C has L nonzero values.
% The PAPR-reduced signal is s=x+c.


%% main function
function [PAR_sdB, Avg_p_inc]=trSOCP()
clear all; clc; close all; clf;


startDate=date; %Store date
tic;        %start stopwatch timer


%% Parameters
N=64;%FFT size
nBits=2; %Number of bits per symbol
Nsym=1e0; %frame size
Fs=4; %oversampling factor
```

```matlab
NF=N*Fs; %length of oversampled signal
%
ydB=0:0.1:16; % PAPR threshold variable
N_ydB=length(ydB);


%% Initialize container objects
PAR_sdB={zeros(1,Nsym); zeros(1,Nsym);zeros(1,Nsym)}; %PAPR container object
CCDFbySimulation={zeros(1,N_ydB); zeros(1,N_ydB);zeros(1,N_ydB)}; %CCDF container
object


%% Set the number of reserved subcarriers
Lprt=4;%ceil([0.05*N 0.2*N]); %5% and 20% reservation


for n=1:length(Lprt)
    if mod(Lprt(n),2)==1
        Lprt(n)=Lprt(n)-1; %number of reserved subcarriers
    end
end


%% File for storing PAPR
file_name=['OFDM_PAPR_SOCP_TR', num2str(N), '_L' num2str(Lprt), '_' ...
    datestr(now,'dd-mmm-yy_HH_MM') '.dat'];
fid=fopen(file_name, 'w+');
%
%%
Avg_p_inc=zeros(length(Lprt),1); %initialise placeholder for average power increase


for p=1:Nsym
    count=p %#ok<NASGU,NOPRT>
    %
    [X,Mod] = mapperSymb(nBits,N);    %Generate modulation symbols


    % Frequency domain oversampling
    X=[X(1 : N/2) zeros(1, N*(Fs-1)) X(N/2+1 : end)];


    x=ifft(X,NF)*sqrt(NF)/sqrt(N);          %Convert the output of IFFT to OFDM signal
    %x=resample(x,Fs,1);         % oversampling


    [PeakPx,AvgPx] = paprSig(x); %Return peak and average power of x
    PAR_sdB{1}(p)=10*log10(PeakPx/AvgPx); %Calculate PAPR
    pwr_x_dB=10*log10(mean(abs(x).^2)); %peak power of x


    for n=1%:length(Lprt)
        X_os=X;
%
```

230

```matlab
        L=Lprt(n);
        out=randperm(N); m=sort(out(1:L));    %Generate L random PRT locations


        %positions of PRT in oversampled sig.
        m(m>N/2)=m(m>N/2)+N*(Fs-1);
        %
        X_os(m)=0;                            %reserve L subcarriers
        %
        C=zeros(1,NF);                        %initialize PRC vector
        C(m)=cOptSocp(m,NF,X_os);              %generate peak-reduction coefficients


        s=ifft(X_os+C,NF)*sqrt(NF)/sqrt(N);   %transmit signal s=x+c


        [PeakPs] = paprSocpTr(s);             %peak power of transmit signal


        PAR_sdB{n+1}(p)=10*log10(PeakPs/AvgPx);  %PAPR of the transmit signal


        % Power increase calculation
        pwr_s_dB=10*log10(mean(abs(s).^2));
        p_inc=pwr_s_dB-pwr_x_dB;
        Avg_p_inc(n)=(Avg_p_inc(n)*(p-1)+p_inc)/p;
    end;
end;


%% CCDF data and plotting
linColr={('r-'), ('b:'),('g--')}; %colour for the plots
%
for k=1:length(Lprt)+1
    for i=1:N_ydB,
        CCDFbySimulation{k}(i)=sum(PAR_sdB{k}>ydB(i))/Nsym; %generate CCDF
    end
    semilogy(ydB(1:3:end),CCDFbySimulation{k}(1:3:end),linColr{k});%plot the CCDFs
    grid on;
    hold on;
end;
%
toc %end stopwatch timer
execTime=round(100*toc/3600)/100;    %display the duration of execution on the
screen.
%
%CCDF plot labels
axis([ydB([1 end]) 1e-5 1]);
title(['PAPR of Signal for N = ',  num2str(N),' subcarriers, ','Mod. = ',
num2str(Mod), ', NBlk =', num2str(Nsym),...
    ', Date: ', num2str(startDate), ', Exec. time=', num2str(execTime), ' hours']);
```

231

```matlab
xlabel('\gamma(dB)'); ylabel('Pr(PAPR > \gamma)');legend('Original PAPR','PRC=6.25%');
hold off


%% Scatter/bubble plot
r=max(max(abs(real(x))),max(abs(imag(x))))+0.5;
figure
scatter(real(x),imag(x)), hold on,
scatter(real(s),imag(s),[],'filled'), hold off,
axis([-r, r, -r, r]),
axis square,
xlabel('Re'); ylabel('Im');legend('x','s=x+c');
title(['Scatter plots of time-domain signals x and s, ','N = ',  num2str(N),', Mod. =
', num2str(Mod)]);
end


%%Calculate PAPR of signal
function [peakPwr,avgPwr] = paprSocpTr(xSig)
lenSig=length(xSig);
xRe=real(xSig);
xIm=imag(xSig);
pwrSig = xRe.*xRe + xIm.*xIm;
avgPwr = sum(pwrSig)/lenSig;
peakPwr = max(pwrSig);
paprdB = 10*log10(peakPwr/avgPwr);


%Returns peak-reduction coefficients
function C_hat=cOptSocp(m,N,X)
%This function generates peak-reduction coefficients in the frequency     %
%domain by minimizing max |Q_hat*C_hat+x|                                 %
%Utilizes the complex number multiplication (a+jb)(y+jz)=(ay-bz)+j(az+by) %
%for the matrix multiplication Q_hat*C_hat.                               %


%% inputs
L=length(m); %Number of peak-reduction coefficients
x=(ifft(X,N)*sqrt(N)).'; %time-domain signal
Q_N=ifftmatSocp(N); %generate IFFT matrix, Q_hat, of size N x N
Q_hat=Q_N(:,m);    %generate sub-ifft matrix, Q_hat, of size N x L


%% Matrix A=[A1;A2; ... Ap]; p=N
Ar=real(Q_hat);
Ai=imag(Q_hat);
A = Ar + 1i*Ai;


%% Matrix b=[b1; b2; ... bp]; p=N
br=real(x);
```

```matlab
bi=imag(x);
b = br + 1i*bi;


%% minimize     max(abs(Q_hat*C_hat)+x)
cvx_begin quiet
    variable y(L) complex
    minimize    (norm(A*y+b,inf))
cvx_end
%% Optimal y
C_hat = y;      %Peak-reduction coefficient vector of size L x 1 of complex numbers
echo off
```

H.4 Program for LP-based TR Method

```matlab
%-----------------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of OFDM signals using the proposed LP-based TR method%
% Author:  Stephen Kiambi                                                          %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)               %
%-----------------------------------------------------------------------------------%


% The proposed method employs linear programming (LP) to minimize the infinity norm,
% i.e. min max |Q_hut*C_hut - d|
% The trLpCheby.m function generates the peak-reduction coefficient vector C
% for OFDM signal x by calling the linprogChebyNorm.m function with the
% desired peak-reduction signal.
% C is converted to a peak-reduction time-domain signal c.
% The data vetor, X, and peak-reduction coefficient vector, C, occupy
% disjoint frequency subspaces.
% X has N-L nonzero values, while C has L nonzero values.
% The peak-reduced signal is s=x-c.


%main function
%-----------------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of OFDM signals using the proposed LP-based TR method%
% Author:  Stephen Kiambi                                                          %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)               %
%-----------------------------------------------------------------------------------%


% The proposed method employs linear programming (LP) to minimize the infinity norm,
% i.e. min max |Q_hut*C_hut - d|
% The trLpCheby.m function generates the peak-reduction coefficient vector C
% for OFDM signal x by calling the linprogChebyNorm.m function with the
% desired peak reduction signal.
% C is converted to a peak-reduction time-domain signal c.
% The data vetor, X, and peak-reduction coefficient vector, C, occupy
```

```matlab
% disjoint frequency subspaces.
% X has N-L nonzero values, while C has L nonzero values.
% The PAPR-reduced signal is s=x-c.


%%
function [PAR_sdB, Avg_p_inc]=trLpCheby()
    clear all; clc; close all; clf;
    disp('Program running, please wait ...')
    disp('  ')
    %
    startDate=date; %store date
    tic %start stopwatch


%% Global parameters and initialization
    N=64;    %FFT window size
    b=2;   %bits per symbol
    Nsym=1e3; %frame size
    Fs=4; %oversampling factor
    NF=N*Fs; %length of oversampled signal
    Lprt=[4, 8];    %number of subcarriers reserved for PRT
    %
    ydB=0:0.1:16; % PAPR threshold variable
    N_ydB=length(ydB);


    for k=1:length(Lprt)+1,
        PAR_sdB{k}=zeros(1,Nsym); %initialize PAPR container object
    end;


    PAR_sdB=PAR_sdB.';
    CCDFbySimulation={zeros(1,N_ydB); zeros(1,N_ydB)};%CCDF container object


%% File for storing PAPR
    fileName=['OFDM_PAPR_LpTR_oversampled', num2str(N), '_L' num2str(Lprt), '_'
datestr(now,'dd-mmm-yy_HH_MM') '.dat'];
    fileId=fopen(fileName, 'w+');
    %
    %%


    Avg_p_inc=zeros(length(Lprt),1); %initialize placeholder for increases in average
power
    %%
    for p=1:Nsym
        count=p %#ok<NASGU,NOPRT>
        %
        [X,Mod] = mapperLpTR(b,N);    % generate data vector X
```

```matlab
x=ifft(X.',N)*sqrt(N);      % converts ifft sig. to OFDM sig.
x=resample(x,Fs,1);         % oversampling


par_dB=paprSig(x); %PAPR of x
PAR_sdB{1}(p)=par_dB;
pwr_x_dB=10*log10(mean(abs(x).^2)); %power of x


%if par_dB>=targetPapr;
for n=1:length(Lprt)
    Xr=X.';


    L=Lprt(n); %number of reserved subcarriers


    out=randperm(N); m=sort(out(1:L));  %random L positions of PRT
    Xr(m)=0;         %reserve positions of PRT in the freq. domain
    xr=ifft(Xr,N)*sqrt(N);
    xr=resample(xr,Fs,1);


    %% clipping level
    x_clip=mean(abs(xr));


    %% Generation of desired peak-reducing signal
    d=zeros(NF,1);
    for i=1:NF
        if abs(xr(i))>x_clip
            d(i)=xr(i)/abs(xr(i))*(abs(xr(i))-x_clip);
        end
    end
    %% forming IFFT submatrix
    Q_N=conj(dftmtx(NF))/sqrt(N);


    m(m>N/2)=m(m>N/2)+N*(Fs-1);%positions of PRT in oversampled sig.


    Q_hat=Q_N(:,m); %generate sub-ifft matrix, Q_hat of size NF x L


    %% Generating Peak-reducing signal coefficients
    C_hat=linprogChebyNorm(m,NF,Q_hat,d); %peak-reduction coefficients


    c=Q_hat*C_hat; %peak-reducing signal


    %% Peak-reduced signal
    s=xr-c;
```

```matlab
        %% PAPR calculation
         PAR_sdB{n+1}(p) = max(abs(s).^2)/mean(abs(x).^2);


        % Power increase calculation
         pwr_s_dB=10*log10(mean(abs(s).^2));
         p_inc=pwr_s_dB-pwr_x_dB;
         Avg_p_inc(n)=(Avg_p_inc(n)*(p-1)+p_inc)/p;
     end;
 end;
 %% CCDF data and plot
 linColr={('k-'), ('m-o'),('b-x')};
 %


 for k=1:length(Lprt)+1
     for i=1:N_ydB,
         CCDFbySimulation{k}(i)=sum(PAR_sdB{k}>ydB(i))/Nsym;%CCDF values
     end;
     %plotting of CCDF
     fig_a=semilogy(ydB(1:3:end),CCDFbySimulation{k}(1:3:end),linColr{k});


     grid on;
     hold on;
 end;


set(gca,'FontSize',14,'FontWeight','bold');
set(fig_a,'LineWidth',2.5,'MarkerSize',10);


 %% print CCDF values to file
 for kk=1:N_ydB,
     fprintf(fileId, '%d\t %11.3e %11.3e\n', ydB(kk).',
(CCDFbySimulation{1}(kk)).', (CCDFbySimulation{2}(kk)).');
 end
 %  %% Plotting CCDF from file
 % if (fileId~=0),
 %     fclose(fileId);
 % end
 % plot_CCDF(fileName);
 %grid on
 %
 toc %stop timer
 execTime=round(100*toc/3600)/100;      %displays the execution time.
 %
 %CCDF plot labels
 axis([ydB([2 length(ydB)]) 1e-3 1]);
```

236

```matlab
    title(['PAPR of OFDM with N = ',num2str(N),' subcarriers, ','L = ',...
        num2str(Lprt),', Mod. = ', num2str(Mod), ', Nsym =', num2str(Nsym),...
        ', PwrInc =', num2str(Avg_p_inc.'), ', Date: ', num2str(startDate), ...
        ', Exec. time=', num2str(execTime), ' hours']);


    xlabel('\gamma(dB)'); ylabel('Pr(PAPR > \gamma)');legend('Original
PAPR','PRC=6.25%','PRC=12.50%');


    hold off
    %
    disp('  ')
    disp('End of program')
end




%% Function for finding peak-reduction coefficients
function [C_hat]=linprogChebyNorm(m,N,Q_hat,d)
%------------------------------------------------------------------------
% Aim:find the value of x that minimizes the Chebyshev norm ||Ax - b||_inf,
% i.e % min max |Ax - b| by% employing the linprog command with default
% options (see: optimoptions(@linprog) in the Matlab Optimization Toolbox:
% f = [ zeros(n,1); 1 ];
% Ane = [ +A, -ones(m,1) ; ...
%         -A, -ones(m,1) ];
% bne = [ +b; -b ];
% xt = linprog(f,Ane,bne);
% x_cheb = xt(1:n,:);
%
% With CVX, the same problem can be specified as follows:
% cvx_begin
%  variable x(n)
%  minimize( norm(A*x-b,Inf) )
% cvx_end
%%
L=length(m);


%Matrix A=[a1;a2; ... ap]; p=N vectors
Ar=real(Q_hat);
Ai=imag(Q_hat);
A = [Ar -Ai;Ai Ar];


b=[real(d);imag(d)];
```

```
n=2*L; m=2*N;
%%
f = [ zeros(n,1); 1 ];
Ane = [ +A, -ones(m,1) ; ...
        -A, -ones(m,1) ];
bne = [ +b; -b ];
xt = linprog(f,Ane,bne);
x_cheb = xt(1:n,:);


C_hat=x_cheb(1:L)+1i*x_cheb(L+1:2*L);%peak-reduction coefficients


echo off



% Calculates and plots the BER when the LP-TR Method is used to reduce PAPR
function trLpCheby_BER()
    % Generates OFDM signal x and peak-reduction c using TR method based
    % on linear programming to solve min max |Q_hut*C_hut - d|.
    % Transmits the peak-reduced signal s=x-c over AWGN channel
    % Then calculates the BER at the receiver

    %%
    clear all; clc; close all; clf;
    disp('Program running, please wait ...')
    disp(' ')
    %
    tic  % start stopwatch timer

    %% Global parameters and initialization
    N=64;    %FFT window size
    b=2;     % no. of bits per symbol
    Nsym=1e4; % number of OFDM symbols
    Fs=4;     % Oversampling factor
    NF=N*Fs;   % FFT size for the oversampled signal
    EbN0dB=0:2:18; % SNR per bit (Eb/N0) in dB
    L =4; %no.of reserved subcarriers

    %% HPA parameters
    p_HPA=2; %smoothness parameter
    IBO = 8; %IBO sets the saturation level: Psat = Pav x 10^(IBO/10)


    %% initialize average symbol power
    sigPow=0;


%%
```

```matlab
%% File for storing BER
fileName=['OFDM_BER_LpCheby_AWGN_N', num2str(N), '_L' num2str(L),'_b' ...
    num2str(b), '_IBO',num2str(IBO), '_' datestr(now,'dd-mmm-yy_HH_MM') '.dat'];
fileId=fopen(fileName, 'w+');



%% Monte Carlo Simulations: Outer loop
for k=0:length(EbN0dB)
    counter_EbN0=k   %#ok<NASGU,NOPRT>


    % Initialize the number of bit errors, and total bits
    nBErrors=0; nBits=0;


    %Inner loop
    for p=1:Nsym
        symb=p       %#ok<NASGU,NOPRT>


        [X,~] = mapperSymb(b,N); %generate data symbols and corresponding integers


        %% Reserve subcarriers
        Xr=X.';
        out=randperm(N);
        m=sort(out(1:L));       % random L positions of PRT
        Xr(m)=0;                % positions of PRT in the freq. domain
        xr=ifft(Xr,N)*sqrt(N);
        xr=resample(xr,Fs,1);


        % clipping level
        x_clip=mean(abs(xr));


        % Generation of desired peak-reducing signal
        d=zeros(NF,1);
        for i=1:NF
            if abs(xr(i))>x_clip
                d(i)=xr(i)/abs(xr(i))*(abs(xr(i))-x_clip);
            end
        end
        %% forming IFFT submatrix
        Q_N=conj(dftmtx(NF))/sqrt(N);
        m_ext=m;
        m_ext(m_ext>N/2)=m_ext(m_ext>N/2)+N*(Fs-1);%PRT in the oversampled signal


        Q_hat=Q_N(:,m_ext); %generate sub-ifft matrix, Q_hat, of size N x L
```

```matlab
% Generating Peak-reducing coefficients
C_hat=linprogChebyNorm(m_ext,NF,Q_hat,d);
c=Q_hat*C_hat;
%
% Peak-reduced signal
s=xr-c;


%% Channel: add AWGN noise_____
if k==0 % measure power for use to add AWGN
    sigPow = sigPow + mean(abs(s).^2); % average sample(symbol) power
    continue;% jump to the next iteration in the current for loop
end


%% pass signal to HPA: p=2, and a given IBO in dB
% Create memoryless nonlinear HPA using Rapp model and apply signal
Asat=sqrt(sigPow*10^(IBO/10));%setting the saturation
y = s./(1+(abs(s)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));%Rapp HPA model


%% calculate noise and pass to AWGN
EsN0dB = EbN0dB(k)+10*log10(b); %  Eb/N0 conversion to Es/N0
%c1:EsN0_num = 10^(EsN0dB/10);% EsN0 in numeral
%c2:var_noise = sigPow/(2*EsN0_num);% noise variance
%c3:sigma_noise = sqrt(var_noise); % noise standard deviation;
%c4:noise=sigma_noise*(randn(size(y))+1i*randn(size(y)));% AWGN noise
%c5:y_noise = y + noise; %noise plus signal
y_noise = awgn(y,EsN0dB,10*log10(sigPow));%c1-5 above replaced by this.


%% Receiver_____
r = downsample(y_noise,Fs); %downsampling to recover input time samples


%% demodulation
%find data indices
a = 1:N; %all indices
inA = 1-ismember(a,m);
data_ind = find(inA); %data subcarrier positions


%Tx symbols
X_data = Xr(data_ind);
[X_int,~] = demapperLS(b,X_data); % demodulated tx integers
txBits = reshape(de2bi(X_int,b,'left-msb'),1,[]);%transmitted bits


%Rx symbols;
R = fft(r)/sqrt(N);
R_data = R(data_ind);
```

```matlab
            [R_int,~] = demapperLS(b,R_data); % demodulated rx integers
            rxBits = reshape(de2bi(R_int,b,'left-msb'),1,[]);%received bits


        %BER computation
        nBerrors=nBerrors+biterr(txBits,rxBits);%total bit-errors
        tBits=tBerror+length(R_int)*b; %total number of bits in received symbols;
    end;


    %% Storage of BER results in a file
    if k==0
        sigPow = sigPow/Nsym; %average symbol power over total OFDM symbols
    else
        BER = nBerrors/tBits;


        %print on screen
        fprintf('EbN0=%3d[dB],BER=%6d/%8d =%11.3e\n',EbN0dB(k),nBerror,tBits,BER)


         %print to file
        fprintf(fileId, '%d\t%11.3e\n', EbN0dB(k), BER);
    end
end


%% Plotting BER
if (fileId~=0)
    fclose(fileId);
end
plotBER(fileName,b,EbN0dB);


%% display the execution time.
toc
execTime=round(100*toc/3600)/100;


disp('  ')
disp('End of program')


end
```

H.5 Program for IRLS-TR Method

```matlab
% ------------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of OFDM signals using the proposed IRLS-TR method %
% Author:  Stephen Kiambi                                                       %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)            %
%-------------------------------------------------------------------------------%

% The proposed method utilises iterative re-weighted least-squares (IRLS) algorithm
% to minimize the infinity norm, i.e., min max |Q_hut*C_hut - d|
```

```matlab
% The trIRLS.m function generates the peak-reduction coefficients vector C
% for OFDM signal x by calling the irlsOpt.m function with the desired peak-
% reduction signal.
% C is converted to a peak-reduction time-domain signal c.
% The data vector, X, and peak-reduction coefficient vector, C, occupy
% disjoint frequency subspaces.
% X has N-L nonzero values, while C has L nonzero values.
% The peak-reduced signal is s=x-c.


%% Main function
function [PAR_sdB, Avg_p_inc]=trIRLS()
clear all; clc; close all; clf;


disp('Program running, please wait ...')
disp('  ')
%
startDate=date;
tic %start stopwatch timer
%% Global parameters and initialization
N=64; %FFT window size
b=2; %number of bits per symbol
Nsym=1e4; %frame size
Fs=4; %oversampling factor
NF=N*Fs; %length of oversampled signal


Lprt=4; %number of reserved subcarriers


ydB=0:0.1:16; % PAPR threshold variable
N_ydB=length(ydB);
for k=1:length(Lprt)+1,
    PAR_sdB{k}=zeros(1,Nsym); %initialize PAPR container object
end;
PAR_sdB=PAR_sdB.'; %transpose


CCDFbySimulation={zeros(1,N_ydB);zeros(1,N_ydB);zeros(1,N_ydB)};%CCDF container object
%
%% File for storing PAPR
file_name=['OFDM_PAPR_IRLS-TR_N', num2str(N), '_L' num2str(Lprt), '_' datestr(now,'dd-
mmm-yy_HH_MM') '.dat'];
fid=fopen(file_name, 'w+');


%%
Avg_p_inc=zeros(length(Lprt),1); %initialize placeholder for changes in transmit power


for p=1:Nsym
    symb=p  %#ok<NASGU,NOPRT>
    %
    [X,Mod] = mapperSymb(b,N);  % generate data vector X
```

242

```matlab
x = ifft(X)*sqrt(N);        %OFDM signal
x=resample(x,Fs,1);         % oversampling by 4


PAR_sdB{1}(p) = paprSig(x); %PAPR of x


pwr_x=10*log10(mean(abs(x).^2)); %power of x


for n=1:length(Lprt)
    Xr=X;


    L=Lprt(n); %number of reserved subcarriers
    out=randperm(N); m=sort(out(1:L)); % random positions of PRT


    Xr(m) = 0;  % Reservation of PRT locations


    xr = ifft(Xr,N)*sqrt(N);  % convert IFFT signal to OFDM signal
    xr = resample(xr,Fs,1);


    %% clipping level
     x_clip = mean(abs(xr));


    %%Generation of desired peak-reducing signal
    d=zeros(NF,1);
    for i=1:NF
        if abs(xr(i))>x_clip
            d(i)=xr(i)/abs(xr(i))*(abs(xr(i))-x_clip);
        end
    end


    %% Generation of peak-reduction signal
    Q_N=conj(dftmtx(NF))/sqrt(N);%IFFT matrix
    m(m>N/2)=m(m>N/2)+N*(Fs-1);%reserved locations in the oversampled signal
    Q_hut=Q_N(:,m); %sub-IFFT matrix, Q_hut, of size NF x L


    C_hut = irlsOpt(Q_hut,d); %generate peak-reduction coefficients


    c=Q_hut*C_hut; %peak-reduction signal


    %% Generate peak-reduced signal
    s=xr.'-c;


    %% PAPR of the transmit signal
```

243

```matlab
        PAR_sdB{n+1}(p) = max(abs(s).^2)/mean(abs(x).^2);


        %% Calculate increase in the transmit power
        pwr_s_dB=10*log10(mean(abs(s).^2));
        p_inc=pwr_s_dB - pwr_x;


        Avg_p_inc(n)=(Avg_p_inc(n)*(p-1)+p_inc)/p;
    end;
end;
%
%% display the duration of execution
toc %end stopwatch timer
execTime=round(100*toc/3600)/100;


%% CCDF data and plots
linColr={('r-'), ('b:'),('g-'),('r-.'), ('b-'),('r-*'),('g:'),('r--'),('m-.')};
%
for k=1:length(Lprt)+1
    for i=1:N_ydB,
        CCDFbySimulation{k}(i)=sum(PAR_sdB{k}>ydB(i))/Nsym; %CCDFs
    end;

    fig_a = semilogy(ydB(1:1:end),CCDFbySimulation{k}(1:1:end),linColr{k}); %plotting
CCDFs
    set(fig_a,'LineWidth',2.5,'MarkerSize',10);
    grid on;
    hold on;
end;


set(gca,'FontSize',14,'FontWeight','bold');


%% print CDDF values to file
for kk=1:N_ydB,
    fprintf(fid, '%d\t %11.3e %11.3e\n', ydB(kk).', (CCDFbySimulation{1}(kk)).',...
            (CCDFbySimulation{2}(kk)).');
end


%% labels for CCDF plots
axis([ydB([2 end]) 1e-3 1]);


title(['PAPR Red: N = ',  num2str(N),', L = ', num2str(Lprt),', Mod. = ', ...
      num2str(Mod), ', Nsym =', num2str(Nsym),', PwrInc =', num2str(Avg_p_inc.') ...
      ,',Date: ', num2str(startDate), ', Exec. time=', num2str(execTime), ' hours']);
```

244

```matlab
xlabel('\gamma(dB)'); ylabel('Pr(PAPR > \gamma)');


legend('Without PAPR reduction', 'L = 13 (5%) PRCs');


hold off


disp(' ')
disp('End of program')




%------------------------------------------------------------------------
% Function irlsOpt.m finds the solution to Ax=b where
% A=Q_hut, x=C_hut and b=d
% It minimizes the L_p norm ||Ax-b||_p, with large p to approximate the
% infinity norm, using IRLS with Newtonian iterative update of the solution x
% For 2<p<infinity, use the homotopy parameter, 1.01<a<2
% For 0<p<2, 0.7 <a<0.9
%------------------------------------------------------------------------
function C_hut = irlsOpt(Q_hut,d)
%Parameters and initialization
norm_p = 20;                    %target p-norm to approximate infinity-norm
K = 15;                         %maximum number of iterations
a =10^(log10(norm_p)/K);        %homotopy variation parameter
pk = 2;                         %Initial homotopy value


C_hut = Q_hut\d;                %Initial least-squares solution


%%
k=1;
e = Q_hut*C_hut - d;
%
while k<K+1 && max(abs(e))>1e-4
    pk = min(norm_p,a*pk);        % Homotopic change of p


    e = Q_hut*C_hut - d;          % error vector


    w = abs(e).^((pk-2)/2);       % weights for the errors
    W = diag(w/sum(w));           % normalize the weights for proportionate weighting


    WQ = W*Q_hut;                 % apply the weights


    C_h = WQ\(W*d);               % weighted least-squares solution
```

245

```matlab
    u = 1/(pk-1);                % Newton parameter
    C_hut = C_hut + u*(C_h - C_hut);  % partial update of solution
%   ep = norm(e,norm_p); E = [E ep];  % p-norm error at each iteration
 k=k+1;
end
```

H.6 Program for LCSA Method

```matlab
% --------------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of OFDM signals using the proposed LCSA method   %
% Author:  Stephen Kiambi                                                       %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)            %
%--------------------------------------------------------------------------------%


% The proposed method generates a peak-reduction signal, c, that is similar to
% the desired signal, d.
% The peak-reduced signal is x-c.
% Nonzero samples of PRC are appended to one end of the peak-reduced signal
% to form what is called a composite transmit signal

%% Main function
function [PAR_sdB, Avg_p_inc]=lcsaMethod()
    clear all; clc; close all; clf;


    disp('Program running, please wait ...')
    disp('  ')
    %
    startDate=date;
    tic  % start stopwatch timer
    %% Global parameters and initialization
    N=64; %FFT window size
    b=2;   % number of bits per symbol
    Nsym=1e4; % frame size
    Fs=4; % Oversampling factor
    NF=N*Fs;  % size of oversampled signal
    Lsamp=4; % number of nonzero PRS samples


    ydB=0:0.1:16;    % PAPR threshold variable
    N_ydB=length(ydB);


    for k=1:length(Lsamp)+1,
        PAR_sdB{k}=zeros(1,Nsym); %initialize PAPR container object
    end;
    PAR_sdB=PAR_sdB.';
    CCDFbySimulation={zeros(1,N_ydB); zeros(1,N_ydB)};%initialize CCDF container


    %% File for storing PAPR
    fileName=['OFDM_PAPR_FLCTR', num2str(N), '_L' num2str(Lsamp), '_' ...
        datestr(now,'dd-mmm-yy_HH_MM') '.dat'];
```

```matlab
fileId=fopen(fileName, 'w+');


%%
Avg_p_inc=zeros(length(Lsamp),1); %initialise average power increase


for p=1:Nsym
    symb=p   %#ok<NASGU,NOPRT>


    [X,Mod] = mapperSymb(b,N); % generate data vector X


    x = ifft(X.'); %original time signal with no reduction


    x = resample(x,Fs,1);   %oversampled signal
    PAR_sdB{1}(p) = paprSig(x);  %PAPR calculation
    pwr_x_dB=10*log10(mean(abs(x).^2)); %peak power of x


    for n=1:length(Lsamp)
        M=Lsamp(n);


        %% clipping level
        x_dsc=sort(abs(x),'descend');
        x_clip=x_dsc(M+1);


        %% Generation of desired peak-reducing signal
        d=zeros(NF,1);
        for i=1:NF
            if abs(x(i))>x_clip
                d(i)=x(i)/abs(x(i))*(abs(x(i)) - x_clip);
            end
        end
        %
        d_hut=d(d~=0);%returns nonzero elements in d
        %
        c=d;        % peak-reducing signal
        c_hut=d_hut;  % nonzero samples vector


        %% Generation of peak-reduced signal and its PAPR
        s=[x-c; c_hut]; %composite transmit signal


        %% PAPR calculation
        PAR_sdB{n+1}(p) = max(abs(s).^2)/mean(abs(x).^2);


        % Power increase calculation
```

```matlab
        pwr_s_dB=10*log10(mean(abs(s).^2));
        p_inc=pwr_s_dB-pwr_x_dB;
        Avg_p_inc(n)=(Avg_p_inc(n)*(p-1)+p_inc)/p;
    end;
end;


%
%% display execution time.
toc %stop the timer
execTime=round(100*toc/3600)/100;


%% CCDF data and plots
linColr={('k-'), ('m:+'),('b->'), ('g-o'),('y-x'),('r-*')};
%


% Generate CCDF
for k=1:length(Lsamp)+1
    for i=1:N_ydB,
        CCDFbySimulation{k}(i)=sum(PAR_sdB{k}>ydB(i))/Nsym;
    end;
    fig=semilogy(ydB(1:1:end),CCDFbySimulation{k}(1:1:end),linColr{k});
    hold on;
    set(fig,'LineWidth',1.5);
end;
set(gca,'FontSize',14,'FontWeight','bold');


%% print CCDF values to file
for kk=1:N_ydB,
    fprintf(fileId, '%d\t %11.3e %11.3e\n', ydB(kk).',(CCDFbySimulation{1}(kk)).'…
            , (CCDFbySimulation{2}(kk)).');
end
%
%  %% Plotting CCDF from file
% if (fileId~=0),
%     fclose(fileId);
% end
% plot_CCDF(fileName);
grid on


%% CCDF plot labels
axis([ydB([1 end]) 1e-4 1]);
title(['PAPR of OFDM with N = ',num2str(N),' subcarriers, ','Lsamp = '…
    , num2str(Lsamp),', Mod. = ', num2str(Mod), ', Nsym =', num2str(Nsym),…
    ', PwrInc =', num2str(Avg_p_inc.'), ', Date: ', num2str(startDate), …
    ', Exec. time=', num2str(execTime), ' hours']);
```

```matlab
    xlabel('\gamma_0(dB)','FontSize',14,'FontWeight','bold');
    ylabel('Pr(PAPR > \gamma_0)','FontSize',14,'FontWeight','bold');
    legend({'Original','5.88% loss'},'FontSize',14,'FontWeight','bold');
end



% Calculate and plot BER performance when LCSA Method is used to reduce PAPR
function lcsaBER()
    % Generates OFDM signal and peak-reduction using LCSA method
    % Transmits the composite transmit signal s=[x-c; c_hut] over AWGN channel
    % Then calculates the BER at the receiver
    %
    clear all; clc; close all; clf;
    disp('Program running, please wait ...')
    disp('  ')
    %
    tic % start stopwatch timer
    %% Global parameters and initialization
    N=64; %FFT window
    b=2;  % no. of bits per symbol
    Nsym=1e4; % frame size
    Fs=4; % Oversampling factor
    NF=N*Fs;  % size of oversampled signal
    Lsamp=4; % number of nonzero PRS samples
    EbN0dB=0:2:18; % SNR per bit (Eb/N0) in dB


    %% HPA parameters
    p_HPA=2; %smoothness parameter
    IBO = 7; %IBO to set saturation level: Psat = Pav x 10^(IBO/10)


    %% initialise average symbol power
    sigPow=0;


    %% File for storing BER
    fileNname=['OFDM_BER_TR_AWGN_N', num2str(N), '_L' num2str(Lsamp),'_b'…
            num2str(b),'_IBO',num2str(IBO), '_' datestr(now,'dd-mmm-yy_HH_MM')
            '.dat'];
    fileId=fopen(fileName, 'w+');
    %
    %% Monte Carlo Simulations: Outer loop
    for k=0:length(EbN0dB)
        counter_EbN0=k  %#ok<NASGU,NOPRT>


        % Initialize the number of bit errors, and total bits
```

```matlab
nBErrors=0; tBits=0;


%Inner loop
for p=1:Nsym
    %symb=p   %#ok<NASGU,NOPRT>


    [X_int,X] = mapperSymb(b,N); %data symbols and corresponding integers


    %Generate transmit bits
    transmittedBits = reshape(de2bi(X_int,b,'left-msb'),1,[]);


    x = ifft(X.')*sqrt(N);   %original time signal


    x = resample(x,Fs,1);   %oversampling


    %% signal clipping level
    x_dsc=sort(abs(x),'descend');
    x_clip=x_dsc(Lsamp+1);


    %% Generation of desired peak-reducing signal
    d=zeros(NF,1);
    for i=1:NF
        if abs(x(i))>x_clip
            d(i)=x(i)/abs(x(i))*(abs(x(i))-x_clip);
        end
    end
    %nz_indices=find(d); %indices nonzero elments in d
    d_hut=d(d~=0);%returns nonzero elements in d
    nz_ind=find(d);
    c=d;  % peak-reduction signal
    c_hut=d_hut; %nonzero samples of peak-reduction signal


    %% Generation of peak-reduced signal and its PAPR
    s=[x-c; c_hut];%composite transmit signal


    %% Channel: add AWGN noise_____
    if k==0 % Find signal power to help adding AWGN
        sigPow = sigPow + mean(abs(s).^2); %mean(abs(s).^2)= the average power
        continue;% jump the code lines below to the next iteration of for loop;
    end


    %% pass signal to HPA: p=2, and a given IBO in dB
    % Create memoryless nonlinear HPA using Rapp model and apply signal
```

```matlab
    Asat=sqrt(sigPow*10^(IBO/10));%setting the saturation point
    y = s./(1+(abs(s)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));%Rapp HPA model


    %% Boosted signal.Because of adding samples of signal c_hut to s,
    % SNR is reduced, therefore the signal from the P/S converter
    % has to be boosted before adding noise
    y=sqrt((NF+Lsamp)/(NF))*y;


    %% calculate noise and pass to AWGN
    %Eb/N0 conversion to Es/N0 to include the effect of extension by L samples
    EsN0dB = EbN0dB(k)+10*log10(b)+10*log10(NF/(NF+Lsamp));


    %c1:EsN0_num = 10^(EsN0dB/10);% EsN0 in numeral
    %c2:var_noise = sigPow/(2*EsN0_num);% noise variance
    %c3:sigma_noise = sqrt(var_noise); % noise standard deviation;
    %c4:noise=sigma_noise*(randn(size(y))+1i*randn(size(y)));% AWGN noise
    %c5:y_noise = y + noise; %noise plus signal
    y_noise = awgn(y,EsN0dB,10*log10(sigPow));%c1-5 above replaced by this.


    %% Receiver
    %form peak-reducing signal
    c_noise=zeros(NF,1);
    c_noise(nz_ind)=y_noise(NF+1:end); %#ok<*FNDSB>


    %Remove L peak-reducing samples from the received signal
    y_noise = y_noise(1:NF);


    %Recovery of transmitted data samples by adding the formed
    %peak-reducing signal to the received signal
    r = y_noise+c_noise; %reconstruct data samples
    r = downsample(r,Fs); %downsampling to recover input time samples


    %% demodulation
    Y = fft(r)/sqrt(N);%Received symbols;
    [Y_int,~] = demapperLS(b,Y); % demodulated integers
    receivedBits = reshape(de2bi(Y_int,b,'left-msb'),1,[]);%received bits


    %BER calculations
    nBErrors =nBErrors +biterr(transmittedBits,receivedBits);%total bit-errors
    tBits=tBits+length(Y)*b; %total number of bits in received symbols;
end;


%% Storage of BER results
if k==0
```

```matlab
            sigPow = sigPow/Nsym; %average symbol power for the OFDM frame
        else
            BER = nBErrors/tBits;

             %print on screen
            fprintf('EbN0=%3d[dB], BER=%6d/%8d =%11.3e\n', EbN0dB(k), …
                    nBErrors, tBits, BER)

            %print to file
            fprintf(fileId, '%d\t%11.3e\n', EbN0dB(k), BER);
        end
    end


    %% Plotting BER
    if (fileId~=0)
        fclose(fileId);
    end
    plotBER(fileName,b,EbN0dB);


    %
    %% display the execution time.
    toc
    execTime=round(100*toc/3600)/100


    disp('  ')
    disp('End of program')


end


%% function for Plotting BER
function plotBER(fileName,b,EbN0dB)
M=2^b;
sqM=sqrt(M);
EbN0_num = 10.^(EbN0dB/10);


%% Theoretical BER calculations
if b == 1 %BPSK
    BER = qfunc(sqrt(2*EbN0_num));
elseif b==2 %QPSK
    SER = 2*qfunc(sqrt(2*EbN0_num)).*(1-1/2*qfunc(sqrt(2*EbN0_num)));%SER by formula
for QPSK
    BER = SER/b;%BER=SER/b
else %M-QAM, M>4
    BER_sqM = 2*(1-1/sqM)*qfunc(sqrt(6*log2(sqM)/(M-1).*EbN0_num));%SER by formula for
sqM-PAM
```

```matlab
        SER = (1-(1-BER_sqM).^2);%SER by formula for M-QAM

        BER = SER/b; %BER=SER/b

end


% load BER file

a= load(fileName);


%plotting

fig=semilogy(EbN0dB,BER,'g-', a(:,1),a(:,2),'r-');

axis([EbN0dB(1) EbN0dB(end) 1e-7 1]);

xlabel('E_b/N_0 [dB]');

ylabel('BER');

set(fig,'LineWidth',2.0);

legend('BER by formula', 'BER by simulation');

title(['SNR per bit for M = ', num2str(M), ' Constellation']);

grid on,

%

end
```

H.7 Program for LCASM Method

```matlab
% ----------------------------------------------------------------------------------%
% Purpose: To reduce the PAPR of MIMO-OFDM signals using the proposed LCASM method %
% Author:   Stephen Kiambi                                                          %
% Research Work: PAPR reduction in OFDM Systems (for PhD at EIE-UoN)                %
%-----------------------------------------------------------------------------------%


% The proposed method generates a peak-reduction signal, c, that is similar to
% the desired signal, d.
% The peak-reduced signal is x-c.
% Nonzero samples of PRC are appended to one end of the peak-reduced signal
% to form what is called a composite transmit signal

function lcasmMIMO()

    clear all; clc; close all; clf;

    %

    disp('Program running, please wait ...')

    disp(' ')

    %

    startDate=date;

    tic   % start stopwatch timer


    %% Global parameters

    N = 64;                    %FFT window size

    Nt = 2;                    %Number of transmit antennas

    b = 4;                     %Number of bits per symbol

    Nsym = 1e4;                %Number of OFDM symbols per frame

    Fs = 4;                    %Oversampling factor

    NF = N*Fs;                 %FFT size for the oversampled signal
```

253

```matlab
Lsamp = 4;                    %no of extension samples
linColr={('k-'), ('m:+'),('b->'), ('g-o'),('y-x'),('r-*')};%colours for BER plots


%% plotting parameters and initialization
ydB=0:0.1:16;               %PAPR thresholds
N_ydB=length(ydB);
for k=1:length(Lsamp)+1,
    PAR_sdB{k}=zeros(1,Nsym);
end;
PAR_sdB=PAR_sdB.';
CCDFbySimulation={zeros(1,N_ydB); zeros(1,N_ydB)};


%% File for storing CCDF
fileName=['OFDM_PAPR_MIMO_LCSA', num2str(N), '_L' num2str(Lsamp), '_' ...
    datestr(now,'dd-mmm-yy_HH_MM') '.dat'];
fileId=fopen(fileName, 'w+');
%
%%
Avg_p_inc = zeros(length(Lsamp),1); %initialize average power increase


%% File for storing BER
    fileName = ['MIMO_OFDM_BER_LCSA_FLAT-FADE_N', num2str(N), '_L' …
                num2str(L), '_IBO',num2str(IBO), '_' datestr(now, …
                'dd-mmm-yy_HH_MM')'.dat'];
    fileId = fopen(fileName, 'w+');


%% Outer loop
for p = 1:Nsym
    symb=p %#ok<NASGU,NOPRT>


    %--------------------------TRANSMITTER-------------------------------
    % generate transmit symbols and corresponding integers
    [X,Mod] = mapper_mimo(b,N,Nt);


    %For antenna 1 and 2
    X1 = X(:,1);
    X2 = X(:,2);


    %% ----Pass X1 and X2 through the Space-Time Block Encoder:
    % We will only consider transmission during the 1st symbol duration
    X11 = X1; %1st antenna
    X21 = X2; %2nd antenna

  %% ----Convert to time domain via IFFT for antenna 1 and 2
```

```matlab
        % for the 1st symbol duration
        x11 = ifft(X11);
        x21 = ifft(X21);


        %%Digital-to-analogue conversion (DAC) => oversampled signals for antenna 1&2
        % for the 1st symbol duration
        x11 = resample(x11,Fs,1);
        x21 = resample(x21,Fs,1);


%           figure(1)
%           plot_PSD(x11,Fs,linColr{1})
%           hold on;


        PAR_sdB{1}(p) = max(paprSig(x11), paprSig(x21)); %effective PAPR
        pwr_x_dB = 10*log10(max(mean(abs(x11).^2), mean(abs(x21).^2)));%peak power


        %% inner loop
        for n = 1:length(Lsamp)
            M = Lsamp(n);
            %% ----signal clipping level
            %for the 1st symbol duration
            x11_dsc = sort(abs(x11),'descend'); %sort in descending order
            x11_CL = x11_dsc(M+1);% clipping level


            x21_dsc = sort(abs(x21),'descend'); %sort in descending order
            x21_CL = x21_dsc(M+1);% clipping level


            % Effective clipping level for the 1st symbol duration
            x_CL1 = max(x11_CL, x21_CL);


            %% ----Generation of desired peak-reducing signal for antenna 1 and 2
            %for the 1st symbol duration
            d11 = zeros(NF,1);
            for i = 1:NF
                if abs(x11(i)) > x_CL1
                    d11(i) = x11(i)/abs(x11(i))*(abs(x11(i))-x_CL1);
                end
            end
            %
            d21 = zeros(NF,1);
            for i = 1:NF
                if abs(x21(i)) > x_CL1
                    d21(i)= x21(i)/abs(x21(i))*(abs(x21(i))-x_CL1);
                end
```

255

```
          end


%% ----Signals to be appended to the transmit signal on antenna 1 and 2
%for the 1st symbol duration
c11_hut = d11(d11~=0);%returns nonzero elements in d
c11_hut = [c11_hut; zeros(M-length(c11_hut),1)]; %zero-pad to length L


c21_hut = d21(d21~=0);
c21_hut = [c21_hut; zeros(M-length(c21_hut),1)];


%      %subplot(2,1,1)
%      stem(abs(d11))
%      hold on
%      stem(abs(d21),'r-.x')
%      axis([0 length(d11)-1 0 max(max(abs(d11)),max(abs(d21)))+0.01]);
%      xlabel('n','FontSize',14,'FontWeight','bold');
%      ylabel('Magnitude','FontSize',14,'FontWeight','bold');
%      legend({'Desired PRS-Antenna 1','Desired PRS-Antenna 2'},…
%              'FontSize',14,'FontWeight','bold');
%      set(gca,'FontSize',14,'FontWeight','bold');
%      hold off
%      figure


%% ----Reduce PAPR and generate peak-reduced signal for antenna 1 and 2
%for the 1st symbol duration
s11 = [x11-d11; c11_hut];%length of txd signal increases by L PRS samples
s21 = [x21-d21; c21_hut];


%      plot_PSD(x11-d11,Fs,linColr{n+1})


% plotting transmitted signal
%      subplot(2,1,1)
%      plot(abs(x11-d11))
%      hold on
%      plot((NF+1:NF+M),abs(c11_hut),'r-')
%      axis([0 length(s11) 0 max(abs(x11-d11))+0.01]);
%      xlabel('n','FontSize',14,'FontWeight','bold');
%      ylabel('Transmit signal-antenna 1','FontSize',14,'FontWeight','bold');
%      legend({'Peak-reduced signal part ','Peak-reduction signal part'},…
%              'FontSize',14,'FontWeight','bold');
%      set(gca,'FontSize',14,'FontWeight','bold');
%      hold off
%
%      subplot(2,1,2)
          end
```

256

```matlab
%            plot(abs(x21-d21))
%            hold on
%            plot((NF+1:NF+M),abs(c21_hut),'r-')
%            axis([0 length(s11) 0 max(abs(x21-d21))+0.01])
%            %title('Transmitted signal');
%            xlabel('n','FontSize',14,'FontWeight','bold');
%            ylabel('Transmit signal-antenna 2','FontSize',14,'FontWeight','bold');
%            legend({'Peak-reduced signal part ','Peak-reduction signal part'},…
%                    'FontSize',14,'FontWeight','bold');
%            set(gca,'FontSize',14,'FontWeight','bold');
%            hold off
%            figure


            %% PAPR calculation
            %for the 1st symbol duration
            PAR_sdB{n+1}(p) = max(max(abs(s11).^2)/mean(abs(x11).^2),...
                max(abs(s21).^2)/mean(abs(x21).^2));


            %% Power increase calculation
            %for the 1st symbol duration
            pwr_s_dB = 10*log10(max(mean(abs(s11).^2), mean(abs(s21).^2)));
            p_inc = pwr_s_dB - pwr_x_dB;
            Avg_p_inc(n) = (Avg_p_inc(n)*(p-1)+p_inc)/p;


        end %End of FOR loop for L


    end; %end of FOR loop for Nysm
    legend({'Original','1% loss','2.3% loss','4.8% loss','6.9% loss','9.2% loss'},...
        'FontSize',14,'FontWeight','bold');


    %% display the execution time.
    toc
    execTime=round(100*toc/3600)/100;


    fprintf('Execution time =%2.2g\n', round(100*toc/3600)/100)%print on screen


    %% CCDF data and plots
    %figure(2)
    for k=1:length(Lsamp)+1
        for i=1:N_ydB,
            CCDFbySimulation{k}(i)=sum(PAR_sdB{k}>ydB(i))/Nsym;
        end;
        fig=semilogy(ydB(1:1:end),CCDFbySimulation{k}(1:1:end),linColr{k});
        hold on;
```

```matlab
            set(fig,'LineWidth',1.5);
    end;
    set(gca,'FontSize',14,'FontWeight','bold');
    grid on


    %% print CDDF values to file
    for kk=1:N_ydB,
        fprintf(fileId, '%d\t %11.3e %11.3e\n', ydB(kk).',(CCDFbySimulation{1}(kk)).',…
            (CCDFbySimulation{2}(kk)).');
    end


     %% Plotting CCDF
    if (fileId~=0)
        fclose(fileId);
    end
    plot_CCDF(fileName);
    grid on


    %% CCDF plot labels
    axis([ydB([1 end]) 1e-4 1]);
    title(['CCDF for N = ',num2str(N),' scs, ','Res. scs = ', num2str(M),...
        ', Mod. = ', num2str(Mod), ', Nsym =', num2str(Nsym),...
        ', PwrInc =', num2str(Avg_p_inc.'), ', Date: ', num2str(startDate), ...
        ', Exec. time=', num2str(execTime), ' hours']);
    xlabel('\gamma_0(dB)','FontSize',14,'FontWeight','bold');
    ylabel('Pr(PAPR > \gamma_0)','FontSize',14,'FontWeight','bold');
    legend({'Original','1% loss','2.3% loss','4.8% loss','6.9% loss',...
        '9.2% loss'},'FontSize',14,'FontWeight','bold');


    disp('  ')
    disp('End of program')
end %function end


%generation of modulation symbols for the Nt transmit antennas
function [modSymbols,modTyp] = mapper_mimo(nBit,N,Nt)
    modOrder=2^nBit; % Number Symbols digital Modulation


    if nBit == 1
        modTyp='BPSK';


        normEs=1;
        modObject=modem.pskmod('M',modOrder);


    elseif nBit == 2
```

```matlab
        modTyp='QPSK';


        normEs=1;


        %Create QPSK modulator
        modObject=modem.pskmod('M',modOrder, 'PhaseOffset',pi/4,'SymbolOrder','gray');


    else
        modTyp=[num2str(2^nBit) 'QAM'];


        %Es is the average energy of modulation symbols
        Esym=1;


         %factor to normalize symbols
        normEs=sqrt(3*Esym/(2*(modOrder-1)));


        %Create M-ary QAM Gray encoded modulator
        modObject=modem.qammod('M',modOrder,'SymbolOrder','gray');
    end



    % Generate a block of N random 2^b-PSK/QAM modulation symbols
     mod_int= randi([0,M-1],N,Nt);
     mod_symbols = normEs*modulate(mod_object,mod_int);
end



% Calculate and plot BER performance when LCASM method is used to reduce PAPR
function lcasm_MIMO_OFDM_BER()
    % Generates OFDM signals and peak-reductions using LCASM method
    % Transmits the composite transmit signals over Rayleigh-fading channels
    % Then calculates the BER at the receivers
    %%
    clear all; clc; close all; clf;
    disp('Program running, please wait ...')
    disp('  ')
    %
    tic  % start stopwatch timer
    %% Global parameters and initialization
    N = 256;                  %FFT window size
    Nt = 2;                   %Number of transmit antennas
    b = 2;                    %No. of bits per symbol
    Nsym = 1e4/2;             %Number of OFDM symbols
    Fs = 4;                   %Oversampling factor
```

```matlab
    NF = N*Fs;                %FFT size for the oversampled signal
    EbN0dB = 0:2:20;          %SNR per bit (Eb/N0) in dB
    L = 13;                    %no of added samples in the transmit signal
    chan_len = 1;             %Number of channel taps: 1 for flat-fading Rayleigh
Channel
    cp_len = chan_len+1;      %length of cyclic prefix: must be > that chan_len to avoid
ISI
    fade_var_1D = 0.5;        %1-dimensional fade variance


    %% HPA parameters
    p_HPA = 2;                %Smoothness parameter
    IBO = 7;                  % Sets saturation level, i.e., Psat = Pav x 10^(IBO/10)


    %% initialise average symbol power
    sigPow = 0;


    %% File for storing BER
    fileName = ['MIMO_OFDM_BER_LCSA_Flat-fading_RaleighCh_N', num2str(N),...
        '_L' num2str(L), '_IBO_',num2str(IBO), '_' datestr(now,'dd-mmm-yy_HH_MM')
'.dat'];
    fileId = fopen(fileName, 'w+');
    %
    %% Monte Carlo simulations: Outer loop
    for k = 0:length(EbN0dB)
        counter_EbN0=k   %#ok<NASGU,NOPRT>


        % Initialize the number of bit errors and total bits
        nBErrors = 0;
        tBits = 0;


        %Inner loop1
        for p = 1:Nsym
            symb=p   %#ok<NASGU,NOPRT>


            %-------------------------TRANSMITTER----------------------------------
-
            % generate transmit symbols and corresponding integers
            [X_int,X] = mapper_mimo(b,N,Nt);


            %Generate coresponding transmit bits
            transmittedBits = reshape(de2bi(X_int,b,'left-msb'),1,[]);


            %For antenna 1 and 2
            X1 = X(:,1);
            X2 = X(:,2);
```

260

```matlab
%% ----Pss X1 and X2 through the Space-Time Block Encoder
% for the 1st symbol duration
X11 = X1; %1st antenna
X21 = X2; %2nd antenna


% for the 2nd symbol duration
X12 = -conj(X2); %1st antenna
X22 =  conj(X1); %2nd antenna


%% ----Convert to time domain via IFFT for antenna 1 and 2
% for the 1st symbol duration
x11 = ifft(X11);
x21 = ifft(X21);


% for the 2nd symbol duration
x12 = ifft(X12);
x22 = ifft(X22);


%% ----Digital-to-Analogue Converstion (DAC) = oversampling of signals
% for the 1st symbol duration
x11 = resample(x11,Fs,1);
x21 = resample(x21,Fs,1);


% for the 2nd symbol duration
x12 = resample(x12,Fs,1);
x22 = resample(x22,Fs,1);


%% ----signal clipping level
%for the 1st symbol duration
x11_dsc = sort(abs(x11),'descend'); %sort in discending order
x11_CL = x11_dsc(L+1);% clipping level


x21_dsc = sort(abs(x21),'descend'); %sort in discending order
x21_CL = x21_dsc(L+1);% clipping level


% Effective clipping level for the 1st symbol duration
x_CL1 = max(x11_CL, x21_CL);


%% for the 2nd symbol duration
x12_dsc = sort(abs(x12),'descend'); %sort in discending order
x12_CL = x12_dsc(L+1);% clipping level
```

261

```matlab
x22_dsc = sort(abs(x22),'descend'); %sort in discending order
x22_CL = x22_dsc(L+1);% clipping level


% Effective clipping level for the 2nd symbol duration
x_CL2 = max(x12_CL, x22_CL);


%% ----Generation of desired peak-reducing signal for antenna 1 and 2
%for the 1st symbol duration
d11 = zeros(NF,1);
for i = 1:NF
    if abs(x11(i))>x_CL1
        d11(i) = x11(i)/abs(x11(i))*(abs(x11(i))-x_CL1);
    end
end
%
d21 = zeros(NF,1);
for i = 1:NF
    if abs(x21(i)) > x_CL1
        d21(i)= x21(i)/abs(x21(i))*(abs(x21(i))-x_CL1);
    end
end


%for the 2nd symbol duration
 d12 = zeros(NF,1);
for i = 1:NF
    if abs(x12(i))>x_CL2
        d12(i) = x12(i)/abs(x12(i))*(abs(x12(i))-x_CL2);
    end
end
%
d22 = zeros(NF,1);
for i = 1:NF
    if abs(x22(i)) > x_CL2
        d22(i)= x22(i)/abs(x22(i))*(abs(x22(i))-x_CL2);
    end
end



%% ----Signals to be appended to the transmit signal on antenna 1 and 2
%for the 1st symbol duration
c11_hut = d11(d11~=0);%returns nonzero elments in d
c11_hut = [c11_hut; zeros(L-length(c11_hut),1)]; %pad with zeros to L


c21_hut = d21(d21~=0);
```

```matlab
c21_hut = [c21_hut; zeros(L-length(c21_hut),1)];


%for the 2nd symbol duration
c12_hut = d12(d12~=0);
c12_hut = [c12_hut; zeros(L-length(c12_hut),1)];


c22_hut = d22(d22~=0);
c22_hut = [c22_hut; zeros(L-length(c22_hut),1)];


%% ----Indices of nonzero elments in PRS d11, d21, d12, and d22
%for the 1st symbol duration
nz11 = find(d11);
nz21 = find(d21);


%for the 2nd symbol duration
nz12 = find(d12);
nz22 = find(d22);


%% ----Reduce PAPR and generate peak-reduced signal for antenna 1 and 2
%for the 1st symbol duration
s11 = [x11-d11; c11_hut];
s21 = [x21-d21; c21_hut];


%for the 2nd symbol duration
s12 = [x12-d12;c12_hut];
s22 = [x22-d22;c22_hut];


%Insert Cyclic Prefix of length cp_len
%for the 1st symbol duration
s11 = [s11(length(s11)-cp_len+1:end); s11];
s21 = [s21(length(s21)-cp_len+1:end); s21];


%for the 2nd symbol duration
s12 = [s12(length(s12)-cp_len+1:end); s12];
s22 = [s22(length(s22)-cp_len+1:end); s22];


%% ----Calculate average signal power
if k==0 % Only to measure the signal power for adding AWGN noise
   sigPow = sigPow + mean(abs(s11).^2); %mean(.) is the av symbol power
   continue;% continue to the next iteration (k+1) in current FOR loop;
end


%% ----pass signal to HPA: p=2, and a given IBO in dB for antenna 1 and 2
```

263

```matlab
% Create memoryless nonlinear HPA using Rapp model and apply signal
Asat=sqrt(sigPow*10^(IBO/10));%setting the saturation level


%for the 1st symbol duration
t11 = s11./(1+(abs(s11)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));
t21 = s21./(1+(abs(s21)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));


%for the 2nd symbol duration
t12 = s12./(1+(abs(s12)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));
t22 = s22./(1+(abs(s22)/Asat).^(2*p_HPA)).^(1/(2*p_HPA));




%% ----Boost signal power: because of addition of L and CP samples,
%SNR is reduced, therefore the signal from the P/S converter has
%to be boosted before being passed to the channel
%for antenna 1 and 2 for symbol duration 1
t11 = sqrt((NF+L+cp_len)/NF)*t11;
t21 = sqrt((NF+L+cp_len)/NF)*t21;


%for antenna 1 and 2 for symbol duration 2
t12 = sqrt((NF+L+cp_len)/NF)*t12;
t22 = sqrt((NF+L+cp_len)/NF)*t22;


%% -------------------- RAYLEIGH CHANNEL IMPULSE RESPONSE-----------------
%Channel form h=1/sqrt(n)*[a1+jb1 a2+jb2 ... an+jbn)]: each of a and
%b are of mean = 0 and variance = 1/2 i.e
%h=1/sqrt(n)*1/sqrt(2)*[randn(1,n)+1i*randn(1,n)] or use normrnd as below.
% fade channel for transmitter 1 and receiver 1
h11 = normrnd(0,sqrt(fade_var_1D),1,chan_len)+1i*...
    normrnd(0,sqrt(fade_var_1D),1,chan_len);
%H11 = fft(h11,N); %H11=h11 since the latter is a constant


% fade channel for transmitter 2 and receiver 1
h12 = normrnd(0,sqrt(fade_var_1D),1,chan_len)+1i*...
    normrnd(0,sqrt(fade_var_1D),1,chan_len);
%H12 = fft(h12,N);%same comment as for H11


% fade channel for transmitter 1 and receiver 2
h21 = normrnd(0,sqrt(fade_var_1D),1,chan_len)+1i*...
    normrnd(0,sqrt(fade_var_1D),1,chan_len);
%H21 = fft(h21,N);%same comment as for H11


% fade channel for transmitter 2 and receiver 2
h22 = normrnd(0,sqrt(fade_var_1D),1,chan_len)+1i*...
```

```matlab
    normrnd(0,sqrt(fade_var_1D),1,chan_len);
%H22 = fft(h22,N);%same comment as for H11


%% Eb/N0 conversion to Es/N0 to include effect of extension by L and CP
EsN0dB = EbN0dB(k)+10*log10(b)+10*log10(NF/(NF+L+cp_len));


%% ----Channel output: assume quasi-static channel i.e. channel
%is constant during symbol duration 1 and 2
%First the transit signal is pass through Rayleigh channel
%(convolution with the impulse response) the added noise
%through passing it to awgn.
%Get channel output for first receiver (diversity arm) for symbol duration
1
ch_op11 = awgn(conv(h11,t11)+conv(h12,t21),EsN0dB,10*log10(sigPow));


% Get channel output for second receiver (diversity arm) symbol duration 1
ch_op21 = awgn(conv(h21,t11)+conv(h22,t21),EsN0dB,10*log10(sigPow));


%Get channel output for first receiver (diversity arm) for symbol duration
2
ch_op12 = awgn(conv(h11,t12)+conv(h12,t22),EsN0dB,10*log10(sigPow));


% Get channel output for second receiver (diversity arm) symbol duration 2
ch_op22 = awgn(conv(h21,t12)+conv(h22,t22),EsN0dB,10*log10(sigPow));


%% --------------RECEIVER------------------------------------
%Remove CP
start_inx = cp_len + 1; % starting index


%for the 1st symbol duration
r11 = ch_op11(start_inx:end);
r21 = ch_op21(start_inx:end);


%for the 2nd symbol duration
r12 = ch_op12(start_inx:end);
r22 = ch_op22(start_inx:end);


%% ----forming peak-reducing signal
%for the 1st symbol duration
c11 = zeros(NF,1);
c11(nz11) = r11(NF+1:NF+length(nz11));


c21 = zeros(NF,1);
c21(nz21) = r21(NF+1:NF+length(nz21));
```

```matlab
%for the 2nd symbol duration
c12 = zeros(NF,1);
c12(nz12) = r12(NF+1:NF+length(nz12));


c22 = zeros(NF,1);
c22(nz22) = r22(NF+1:NF+length(nz22));


%% ----Remove L PRS samples from received signals on antenna 1 and 2
%for the 1st symbol duration
rx11 = r11(1:NF);
rx21 = r21(1:NF);


%for the 2nd symbol duration
rx12 = r12(1:NF);
rx22 = r22(1:NF);


%% ----Recovery of transmitted data samples by add the formed
%peak-reducing signal to the received signal
%for the 1st symbol duration
y11 = rx11 + c11;
y21 = rx21 + c21;


%for the 2nd symbol duration
y12 = rx12 + c12;
y22 = rx22 + c22;


%% ----ADC: downsampling to recover input time samples
%for the 1st symbol duration
y11 = downsample(y11,Fs);
y21 = downsample(y21,Fs);


%for the 2nd symbol duration
y12 = downsample(y12,Fs);
y22 = downsample(y22,Fs);


%% ----Equalization
%forming matrix Y = [Y11; Y12*; Y21; Y22*]
Y11 = fft(y11,N);
Y12 = fft(y12,N); Y12_c = conj(Y12);
Y21 = fft(y21,N);
Y22 = fft(y22,N); Y22_c = conj(Y22);
```

266

c22 = zeros(NF,1);

```matlab
        % forming channel matrix and its conjugate transpose
        %H=[h11 h12; h12* -h11*; h21 h22; h22* -h21*]
        %Hc=[h11* h12 h21* h22; h12* -h11 h22* -h21];
        h11_c = conj(h11);
        h12_c = conj(h12);
        h21_c = conj(h21);
        h22_c = conj(h22);


        sum_h = h11*h11_c + h12*h12_c + h21*h21_c + h22*h22_c;


        %estimtaes of transmitted symbols X1 and X2
        %X_hut = Hc * Y/sum_h
        X1_hut = (h11_c*Y11 + h12*Y12_c + h21_c*Y21 + h22*Y22_c)/sum_h;
        X2_hut = (h12_c*Y11 - h11*Y12_c + h22_c*Y21 - h21*Y22_c)/sum_h;
        X_hut = [X1_hut X2_hut];


        %% demodulation: demapping from symbols to binary stream
        [X_hut_int,~] = demapper_mimo(b,X_hut); % demodulated integers
        receivedBits = reshape(de2bi(X_hut_int,b,'left-msb'),1,[]);%received bits
%
        %% ----BER calculations: Counting bit errors-----------------------------
----------
        %Use biterr, a Matlab function, that computes number of bit errors and BER
        nBErrors = nBErrors + biterr(transmittedBits,receivedBits);
        %
        tBits = tBits + numel(X_hut)*b; %total number of bits in received symbols;
    end; %end of FOR loop for Nysm


    %% Storage of BER results
    if k==0 %to find the average power of the transmit signal
        sigPow = sigPow/Nsym;
    else
        BER = nBErrors/tBits;
        fprintf('EbN0 =%3d[dB], BER =%6d/%8d =%11.3e\n', EbN0dB(k), nBErrors,...
            tBits,BER)%print on screen
        fprintf(fileId, '%d\t%11.3e\n', EbN0dB(k), BER);%print to file
    end
  end %End of FOR loop for EbN0


  %% Plotting BER
  if (fileId~=0), fclose(fileId); end
plotBERRayleigh(fileName,b,EbN0dB);


  %
  %% display the execution time.
```

267

```matlab
    toc

    fprintf('Execution time =%2.2g\n', round(100*toc/3600)/100)%print on screen

    disp('  ')
    disp('End of program')


end %fn end



% Function for plotting BER
function plotBERRayleigh(file_name,b,EbN0dB)
    M=2^b;
    sqM=sqrt(M);

    %Formula for BER in Rayleigh Fading Channel
    fBER= 2*(1-power(sqM,-1))/log2(sqM);
    b= 6*log2(sqM)/(M-1);
    r = b*10.^(EbN0dB/10)/2;
    RayleighBER = 0.5*fBER*(1-sqrt(r./(r+1)));


    %Rayleigh theory 2
    % EbN0 = 10.^(EbN0dB/10);
    % formulaBER = 0.5.*(1-sqrt(EbN0./(EbN0+1)));%same as RayleighBER.
    % semilogy(EbN0dB,formulaBER,'g-','LineWidth',2);


    %Rayleigh theory 1
    semilogy(EbN0dB,RayleighBER,'r-','LineWidth',2)
    hold on

    fBER= load(file_name);

    %plotting
    fig = semilogy(fBER(:,1),fBER(:,2),'b-','LineWidth',2);

    axis([EbN0dB(1) EbN0dB(end) 1e-7 1]);

    xlabel('E_{b}/N_{0} [dB]');

    ylabel('BER');

    set(fig,'LineWidth',2.0);
```

268

```matlab
        legend('Rayleigh Theory', 'Rayleigh Simulation');
        title(['SNR per bit for M = ', num2str(M), ' Constellation']);


        grid on
        hold off


end
```