

**University of Stuttgart**  
Germany



Institute of  
Robust Power  
Semiconductor Systems

Student Research Project

## **Mmwave Transceiver System Performance Analysis using Post-Processing Analysis**

Yang Liu

Supervisor: Prof. Dr.-Ing. Ingmar Kallfass  
M. Sc. Seyyid Muhammed Dilek

Period: 26.10.2017 – 31.07.2018

Stuttgart, 31.07.2018



# Declaration

I hereby declare that this thesis is my own work and effort and follows the regulations related to good scientific practice of the University of Stuttgart in its latest form. All sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references.

Stuttgart, 31.07.2018

Yang Liu



# Executive Abstract

Abstract



# Zusammenfassung

abstract





# Contents

<b>List of Abbreviations</b>	<b>iii</b>
<b>List of Symbols</b>	<b>v</b>
<b>List of Figures</b>	<b>vii</b>
<b>List of Tables</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Organization of this thesis . . . . .	2
<b>2 Fundamentals</b>	<b>3</b>
2.1 Receiver Architecture . . . . .	4
2.2 Figures of Merit . . . . .	5
2.2.1 Signal-to-noise Ratio(SNR) . . . . .	5
2.2.2 Adjacent-Channel Power Ratio(ACPR) . . . . .	6
2.2.3 EVM . . . . .	7
2.2.4 Waveform Quality Factor ( $\rho$ ) . . . . .	8
2.2.5 Bit Error Ratio (BER) . . . . .	9
2.2.6 Modulation Error Ratio (MER) . . . . .	9
2.2.7 Eye Pattern . . . . .	11
2.2.8 Noise Power Ratio (NPR) . . . . .	11
2.3 Classification of Systems . . . . .	13
2.3.1 Deterministic System . . . . .	13
2.3.2 Linear and Nonlinear System . . . . .	13
2.3.3 Time-Invariant and Time-Varying System . . . . .	13
2.3.4 Linear Time-Invariant (LTI) System and Linear Time-Varying (LTV) System .	14
2.3.5 Linear Time-Varying (LTV) System . . . . .	14
2.3.6 Stochastic System . . . . .	14
2.4 Deterministic Impairments . . . . .	16
2.4.1 IQ Imbalance . . . . .	16
2.4.2 Carrier Feedthrough/Leakage, IQ Offset, DC Offset . . . . .	17
2.4.3 Amplitude/Phase Response depending on frequency . . . . .	18
2.4.4 Multipath Propagation . . . . .	18

2.4.5	Intersymbol Interference (LTV) . . . . .	19
2.4.6	Intermodulation distortion (det. nonlinear) . . . . .	20
2.5	Stochastic Impairments . . . . .	21
2.5.1	Phase Noise(random and multiplicative) . . . . .	21
2.5.2	Thermal Noise(random and additive) . . . . .	22
<b>3</b>	<b>Theory of Post Processing Methods</b>	<b>23</b>
3.1	Adaptive Filter Theory . . . . .	23
3.1.1	Linear Adaptive Filters . . . . .	24
3.1.2	Supervised Adaptive Filter Algorithm . . . . .	25
3.1.3	Blind Adaptive Filtering . . . . .	28
3.1.4	Circularity-Based Approach . . . . .	28
3.2	Machine Learning Method . . . . .	29
3.2.1	$k$ -nearest neighbors (KNN) Method . . . . .	30
3.2.2	SVM Method . . . . .	31
<b>4</b>	<b>Filters Design and Simulation Results</b>	<b>35</b>
4.1	Received Signal Generator . . . . .	35
4.2	Convergence Analysis . . . . .	37
4.2.1	Supervised Adaptive Compensator: LMS and RMS . . . . .	37
4.2.2	Blind Adaptive Compensator: CMA and modified CMA . . . . .	38
4.2.3	Blind Adaptive Compensator: Circularity-Based Approach . . . . .	41
4.3	Impact of the Hyperparameters . . . . .	45
4.4	Computational Load and Impact of time-varying channels . . . . .	46
4.5	KNN and SVM Detectors . . . . .	46
4.6	Simulation of the Whole System . . . . .	49
<b>5</b>	<b>Conclusion and Perspectives</b>	<b>51</b>
	<b>Bibliography</b>	<b>51</b>
	<b>Appendices</b>	<b>55</b>
	<b>Appendix A Acknowledgments</b>	<b>57</b>

## List of Abbreviations

ADS	Advanced Design System
AWGN	Additive white Gaussian noise
BER	Bit Error Ratio
CMA	Constant Modulus Algorithm
EVM	Error Vector Magnitude
IF	Intermediate Frequency
ILH	Institute of Robust Power Semiconductor Systems ( <i>Institut für Robuste Leistungshalbleitersysteme</i> )
IQ	In-phase and Quadrature
KNN	$k$ -nearest neighbors
LMS	Least Mean Square
ML	Machine Learning
MSE	mean squared error
OFDM	Orthogonal Frequency Division Multiplexing
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying
RF	Radio Frequency
RMS	Recursive Least Squares
SER	Symbol Error Rate
SNR	Signal Noise Ratio
SVM	Support Vector Machine



## List of Symbols

$a$  incident power wave



## List of Figures

2.1	Heterodyne and homodyne Architectures . . . . .	4
2.2	The Definition of ACPR . . . . .	6
2.3	Illustration of EVM . . . . .	7
2.4	EVM requirements for various M-ary QAMs . . . . .	8
2.5	Approximation of bit error probabilities of M-QAM over AWGN channel (for no channel coding, i.e., $R_c = 1$ ) . . . . .	10
2.6	Low MER and high MER . . . . .	10
2.7	Interpretation of eye pattern . . . . .	11
2.8	Definition of NPR . . . . .	12
2.9	Classes of EVM Error Causes . . . . .	13
3.1	FIR Transversal Structure . . . . .	24
3.2	Basic SVM Concept . . . . .	31
3.3	Mapping from 2-D to higher dimensions (3-D) . . . . .	32
4.1	ADS Circuit Design . . . . .	36
4.2	<b>Replace by self made</b> Supervised Adaptive Compensator . . . . .	37
4.3	A frame structure that contains reference training signal and transmitted data . . . . .	37
4.4	Equalizer with LMS Adaptive Filter for 4-QAM . . . . .	39
4.5	<b>change legend, add RMS, use 64 QAM</b> $E_b/N_0$ w/o RMS Adaptive Filter for 4-QAM . . . . .	40
4.6	<b>change PN unit to degree</b> RMS Expression in different IQ Imbalance . . . . .	40
4.7	Equalizer with CMA Adaptive Filter for 4-QAM . . . . .	42
4.8	Equalizer with Modified CMA Adaptive Filter for 4-QAM . . . . .	43
4.9	Equalizer with Circularity-Based Approach Filter for 4-QAM . . . . .	44
4.10	MSE of CMA filter in different filter length $N$ . . . . .	45
4.11	Convergence result for weight vector . . . . .	47
4.12	Convergence process for one of weight . . . . .	48
4.13	RX Constellation Points with Phase Noise Impairment . . . . .	49
4.14	Detection of SVM against Phase Noise Impairment . . . . .	50





List of Tables

4.1 Simulation Condition . . . . . 35

4.2 Time Cost for Different Adaptive Filters (with optimal hyperparameters), 4-QAM . . 49



# 1. Introduction

Since the first iPhone in 2007 published, mobile technology affects everyday life more than any moment in the history. People are extremely delighted by new interaction way with other part of world and seeking more robust but smaller and lighter devices. As the foundation of mobile life, telecommunication technology play a key role in the new century. Mobile data traffic is growing in a extraordinary rate. According to Cisco[1], with growth of 63 % annually, global mobile data traffic reached 7.2EB per month at the end of 2016. Almost half a billion (429 million) mobile devices and connections were added in 2016. To meet this increase in the bandwidth demand and solve the saturation problem in networks, carriers are pushed to 70 GHz channels.

## 1.1. Motivation

The modern communication involves on electrical and electromagnetic technologies. The first revolution in wireless communication happened in early 1900s. Guglielmo Marconi made an effective contribution to radio communications and won Nobel Prize in Physics in 1909. The industry develops rapidly from wire transmissions to radio broadcast, mobile network and satellite communications. Life and Production are highly dependent on telecommunication technology, especially on cellular networks, blue tooth and Wireless Local Area Networks (WLAN).

The next generation telecommunication standard 5G is expected to provide peak data rate as high as  $20 \text{ Gbits}^{-1}$  [2]. In order to achieve this high transmission speed, there are basically two ways: increase the bandwidth and increase bandwidth use ratio.

For higher bandwidth, we choose E-band as carrier frequency. E-band is located at frequency spectrum at 71-76 GHz and 81-86 GHz. It has the advantages as large bandwidth capacity and suitable to Point-to-Point communication. Therefore, recently Institute of Robust Power Semiconductor Systems (ILH) in University Stuttgart established e-band in the plane to ground data transmission.

As for better use efficient of bandwidth, Orthogonal Frequency Division Multiplexing (OFDM) and higher modulation are used at the same time. However, high modulation is much more sensitive than lower modulation. Hence the compensation of different modulations and precisely recognition play a critical role.

The aim of this thesis is to study different impairments in Radio Frequency (RF) and kinds of post-method for a practical impairment compensator. The method should be preferably be software based only, and should require as little extra hardware as possible.

## **1.2. Organization of this thesis**

Chapter 2 of this thesis explains the structure of receiver architecture, classifications and characters of impairments. The focus in this chapter is causes and mathematical expression of various impairments, in order to choose corresponding equalization method in Chapter 3.

Mathematical theories and algorithms are proposed in Chapter 3. For deterministic distortion we choose adaptive filters and for stochastic impairments we apply some basic machine learning ideas. Performance evaluation of the proposed algorithms is carried out in Chapter 4. We evaluate the equalizer performance according the constellation, Error Vector Magnitude (EVM) and Bit Error Ratio (BER). We also propose the computational complexity and give advise for parameter choose in given algorithms.

Finally, in Chapter 5, conclusions are drawn and recommendations are given.

## 2. Fundamentals

The millimeter-wave (mm-wave) spectrum offers a wide bandwidth and can enable high-data-rate wireless communications.

One of the challenges in the circuit implementation of mm-wave high speed transmission is the linearity/efficiency trade-off. Traditional RF processing techniques such as heterodyne receivers when applied to modern wireless systems impose heavy complexity and lack flexibility. This causes performance limitations in terms of cost, gain and noise power trade-off. To solve this disadvantage, recent wireless systems switch from traditional heterodyne receivers to low intermediate frequency (Intermediate Frequency (IF)) and homodyne (also called as direct conversion or zero-IF) receivers.

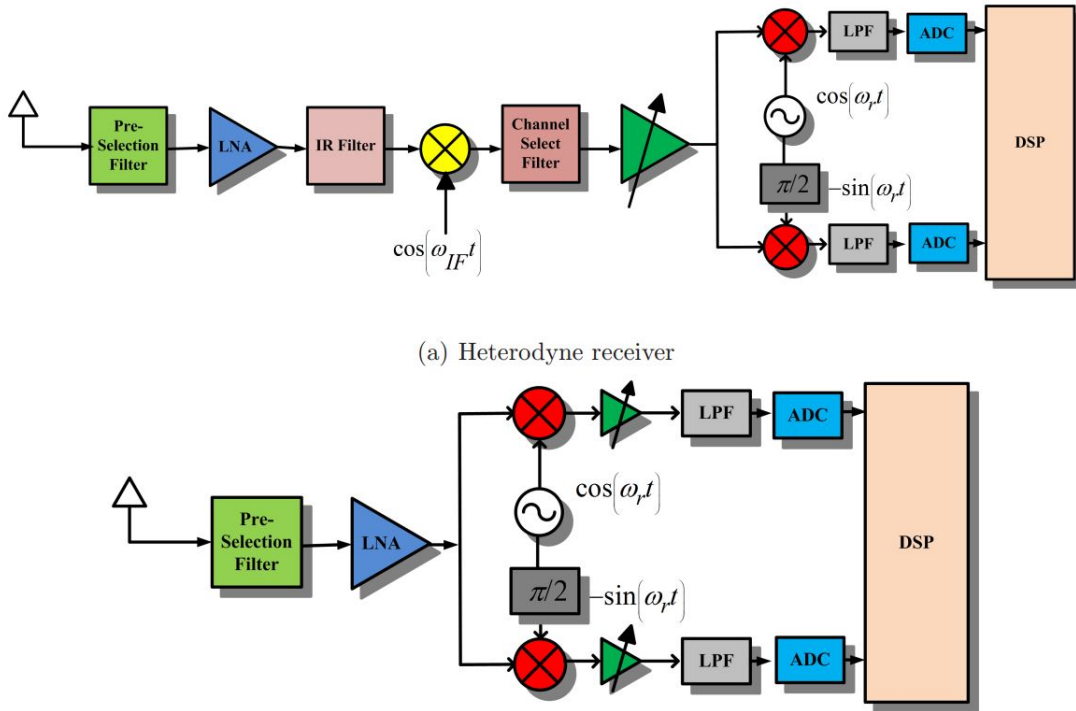
To reach high data rate, the quality of system including transmitters, signal paths, receivers should be quantitatively measured. So many figures of merits are discussed to indicate the quality of transmission system. SNR and BER shows the signal quality of the received signal. ACPR, EVM, MER are used to indicate the quality of the whole transmission system in various aspects by comparing received signal and target signal in constellation plane. Sometimes two-tone intermodulation distortion (IMD) may not be indicative for multitone signals. By creating a notched signal, NPR provide a accurately prediction for multitone IMD. Eye pattern introduces a straightforward way to observe the information of received signal. It is an effective tool to provide a visual examination of the severity of the ISI, sensitivity to timing errors, and the noise margin[3]. Anyway, different indicators together tell a more complete story than anyone by itself.

Furthermore different impairments are briefly introduced. The attractiveness of the homodyne architecture comes with various implementation issues that mainly arise from the impairments in analog hardware in the receiver front ends. In this report, several impairments associated with those homodyne receivers as well as impairments for general receiver architectures. According to different criterion, the impairments are classified to linear and nonlinear, deterministic or stochastic, time-invariant or time-varying errors. This classification will be helpful for the further impairments handling.

## 2.1. Receiver Architecture

Modern receivers serve multiple users that operate in narrow bandwidth and require ability to select the appropriate channel with effective interference suppression from other channels. To realize this, receivers typically convert the received signal frequency  $\omega_r$  into intermediate frequency (IF),  $\omega_{IF}$ , by mixing the incoming signal with a sinusoid of frequency  $\omega_r - \omega_{IF}$ . This mixing process generates frequency components at  $\omega_{IF}$  and  $2\omega_r - \omega_{IF}$ . Hence, the major issue arising in this type of down conversion is the so called image band problem, that is caused by the interference of the required signal with its mirror signal. To overcome this problem, image-rejection filters are applied before mixing the incoming signals with the local oscillator signal (LO).

Among various receiver architectures, two widely-used and well-known receiver architectures, i.e., heterodyne and homodyne, are discussed in this section. Heterodyne receivers have been traditionally used in many applications while homodyne receivers are gaining much more popularity in recent years due to the advantages they offer to wireless systems in terms of complexity.



**Figure 2.1.:** Heterodyne and homodyne Architectures

In heterodyne receivers, the incoming signal is amplified using low noise amplifier (LNA) followed by an image reject (IR) filter. The resulting signal is then mixed with local oscillator signal which performs frequency conversion. A channel selection filter is then applied to select the desired channel after which complex mixing is performed to demodulate the desired signal. The major design concern is choosing the right IF such that it is high enough to successfully

reject the image band and low enough such that channel selection filters can be designed feasibly. Although this receiver architecture provides high channel selectivity, it suffers from several disadvantages such as severe trade-off between gain, noise levels, stability and power requirements.

On the other hand, instead of translating the frequency of the incoming signal to IF, homodyne receivers translate the incoming bandpass signal directly to baseband signal after amplifying through LNA. After IQ mixing – which is required for amplitude/phase modulated signals such as frequency modulated (FM) or quadrature amplitude modulated (QAM) signals – low pass filters can be utilized to achieve channel selectivity. This has various advantages over heterodyne receivers. It does not suffer from image band problem and does not require IR filters thereby hugely reducing the system complexity.

Although homodyne architecture is very attractive for future high rate wireless applications as compared to heterodyne architecture, it introduces various RF-impairments into the received baseband signal due to the non-ideal components.

## 2.2. Figures of Merit

Wireless communication system performance can be characterized by different figures of merit. Besides the often used SNR, other system metrics are also used to quantify system performance and can be more effective and also related to SNR.

### 2.2.1. Signal-to-noise Ratio(SNR)

Signal-to-noise ratio is a basic indication of impairments and defined by a ratio value between the power of signal and of noise as

$$SNR = \frac{P_{signal}}{P_{noise}}. \quad (2.1)$$

$P_{signal}$  and  $P_{noise}$  are the power of signals and of noise.

The definition above is assuming a linear AWGN channel. However, nonlinear distortion increases the system BER for a fixed AWGN power. Hence an alternative definition of SNR is needed to develop the relationship between nonlinear distortion and system performance. The effective system SNR can be expressed in terms of the PSDs as

$$SNR = \frac{\sum_{-\frac{B}{2}}^{\frac{B}{2}} S_{correlated} df}{\sum_{-\frac{B}{2}}^{\frac{B}{2}} S_{uncorrelated} df + N_0 B}, \quad (2.2)$$

where the PSD of output of a nonlinearity is composed of two spectral components and  $S_{correlated}$  represents the correlated output and  $S_{uncorrelated}$  the uncorrelated distortion. The PSD of AWGN

represented by its PSD,  $N_0$ .

### 2.2.2. Adjacent-Channel Power Ratio(ACPR)

In communication system the signal should be transmitted in a limit frequency (channel bandwidth). However, because of the intermodulation (IM) distortion a series of harmonic signals in similar frequencies cause a lot of IM signals inside and out (adjacent-channel) of the channel. Resulting from nonlinear amplification in wireless communication systems there is spectral broadening called Adjacent-Channel Interference (ACI). Both an in-band distortion and an out-of-band distortion exist.  $P_{1dB}$  which is measured by using a single-tone signal, and IP3, which is acquired by using a two-tone signal, are not adequate for metrics of PA nonlinearities. Because modern RF waveforms are digitally- modulated signals that use complex modulations and have different signal statistics from single-tone or two-tone signals. Since other frequencies can also be used by other communication systems, ACI causes noise in other neighbor channel. So it's necessary to establish a lower limit on the ratio of the power in the main channel to the amount of power induced in the adjacent channel.

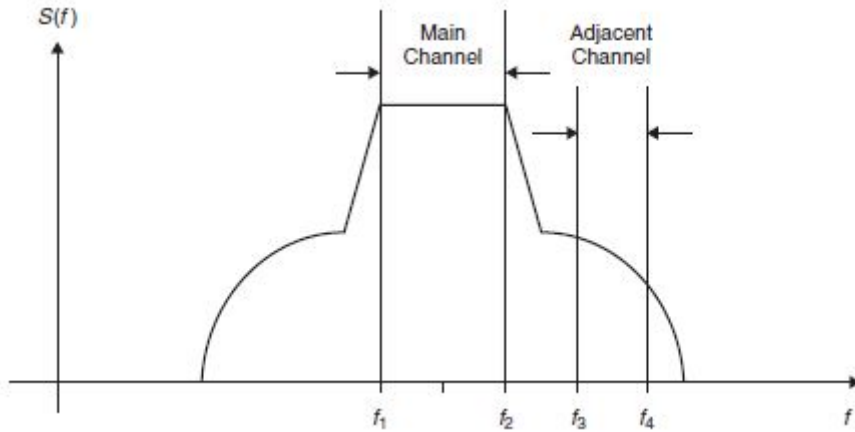


Figure 2.2.: The Definition of ACPR

The power ratio of main channel to the adjacent-channel channel is called Adjacent-Channel Power Ratio (ACLR), which is defined as

$$ACPR = \frac{\text{power spectral density in the main channel 1}}{\text{power spectral density in the offset channel 2 or 3}} = \frac{\int_{f_1}^{f_2} S_{yy}(f) df}{\int_{f_3}^{f_4} S_{yy}(f) df}. \quad (2.3)$$

The frequencies  $f_1$  and  $f_2$  are the frequency limits of the main channel, while  $f_3$  and  $f_4$  are the limits of the upper adjacent channel.



ACPR describes linearity of the communication system with the influence of third order intermodulation distortion. The poor CPR means that some energy that is supposed to be in main channels leaked o adjacent channels. It results in reducing the efficiency of signal transmissions.

### 2.2.3. EVM

There are two methods to characterize the level I-Q vector error: error vector magnitude (EVM) adn a quality factor called  $\rho$ -factor[4].

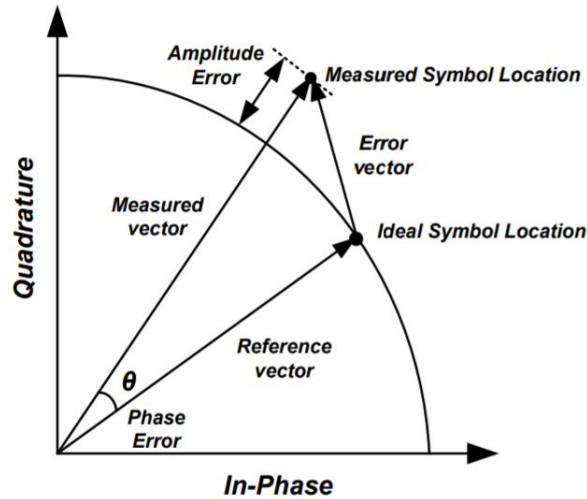


Figure 2.3.: Illustration of EVM

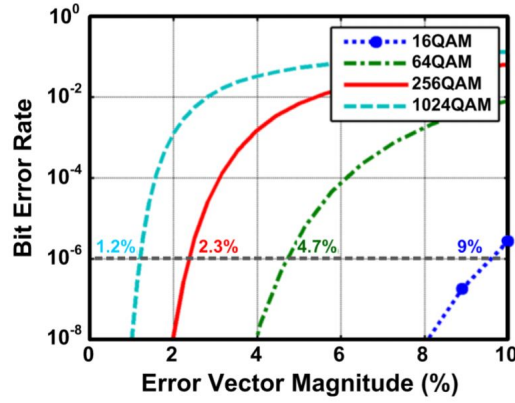
The error vector magnitude (EVM) is a measure to quantify the performance of a digital radio transmitter or receiver. Digital modulation requires a constellation diagram to identify each data point. Figure 2.3 shows that the ideal symbol locatoin is often displaced by amplitude and phase error through a transmitting chain, and the measured symbol location is found in a different location by an amount of an error vector. EVM expresses the difference between the expected complex voltage value of a demodulated symbol and the value of the actual received symbol.

The received symbols are plotted in a coordinate system. This coordinate system is called signal constellation, that allow to visualize modulation schemes and the effect of nonidealities on them. EVM is a quantitative measure of the impairments that corrupt the signal.

Through combination and normalization of measured symbols and ideal constellation diagram using root-mean-square, EVM can be expressed mathematically as

$$EVM = \frac{1}{P_{avg}} \frac{1}{N} \sum_{j=1}^N e_j^2 \quad (2.4)$$

where  $e_j$  denotes the magnitude of each error vector and  $V_{rms}$  the root mean square power of the signal. Note that to express EVM in decibels, we compute  $10\log EVM$ .



**Figure 2.4.:** EVM requirements for various M-ary QAMs

EVM measurement is sensitive to amplitude and phase imbalance at the output of the complete link. It form a powerful tool for analyzing the effect of various nonidealities in the transceiver and the propagation channel. Effects such as noise, nonlinearity and ISI readily manifest themselves in both.

EVM can be related to SNR and  $\rho$  as follows

$$\begin{aligned} EVM &= \sqrt{\frac{1}{SNR}} \\ &= \sqrt{\frac{1}{\rho} - 1} \end{aligned} \quad (2.5)$$

The following calculated EVM Equations can be formulation on either in a dB scale or a percent scale.

$$EVM(dB) = 10 \log\left(\frac{ErrorVector}{ReferenceVector}\right) \quad (2.6)$$

$$EVM(\%) = \sqrt{\frac{ErrorVector}{ReferenceVector}} \times 100\% \quad (2.7)$$

#### 2.2.4. Waveform Quality Factor ( $\rho$ )

In the CDMA system, the waveform quality factor  $\rho$  is used to represent the modulation accuracy instead of EVM[5]. The waveform quality factor is a measure of the correlation between a scaled version of the input and the total in-channel output waveforms. The waveform quality factor  $\rho$

[6] is defined as a correlation coefficient between the actual waveform  $Z(t)$  and the ideal waveform  $R(t)$ , and it is expressed as

$$\rho = \left( \frac{\sum_{k=1}^M R_k Z_k^*}{\sum_{k=1}^M R_k \sum_{k=1}^M Z_k} \right)^2, \quad (2.8)$$

where  $R_k = R_k(t_k)$  is the  $k$ th sample of the ideal signal in the measurement interval;  $Z_k = Z_k(t_k)$  is the  $k$ th sample of the actual signal in the measurement interval; and  $M$  is the measurement interval in half-chip intervals. There is an approximate relationship between  $\rho$  and the EVM when the cross-correlation between the ideal signal and the error signal is negligible,

$$\begin{aligned} \rho &= \frac{1}{1 - \text{EVM}^2} \\ &= \frac{\text{SNR}}{\text{SNR} + 1} \end{aligned} \quad (2.9)$$

### 2.2.5. Bit Error Ratio (BER)

Bit error ratio is the main performance parameter of a wireless channel and is defined as the number of error bits found in a given bit sequence at the receiver end. Bit error rate depends on the value of Signal Noise Ratio (SNR). It decreases with the increase in SNR[7].

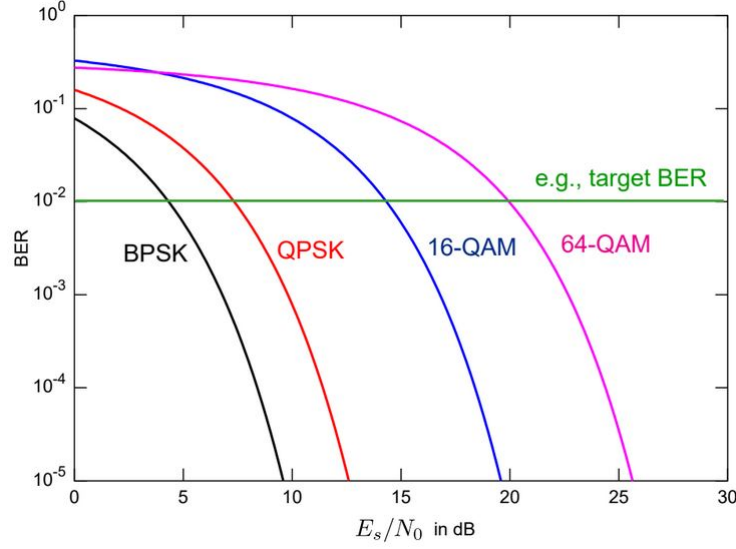
$$\text{BER} = \frac{N_{\text{error}}}{N_{\text{total}}} \quad (2.10)$$

Communication service quality can be described by the average BER. Typically, a BER of  $10^{-2}$  or better is required for voice telephony, and  $10^{-5}$  or better is required for data[4]. Different modulation schemes require more or less power to achieve the same bit error rate. As the requirement of transmission speed, higher modulation increases bit error rate and higher signal noise ratio is demanded satisfying the requirement of BER, as Figure 2.5 shows.

### 2.2.6. Modulation Error Ratio (MER)

Modulation Error Ratio (MER) is an important qualitative parameter for transmission quality of RF modulation. Similar to EVM, MER expresses the difference between the ideal complex voltage value and the value of the actual received symbol.

QAM can be expressed in an analog way by constellation diagram and also in digital way by a two digital bit streams. The two carrier waves are with phase difference by  $90^\circ$  and are called quadrature components of I component and Q component. All constellation points, which are I and Q components, will precisely locate at ideal location, if the transmitter, signal path and receiver are all ideal. However various imperfections in the implementation such as noise, phase noise, carrier suppression, etc. cause the shift between the actual and ideal constellation points.

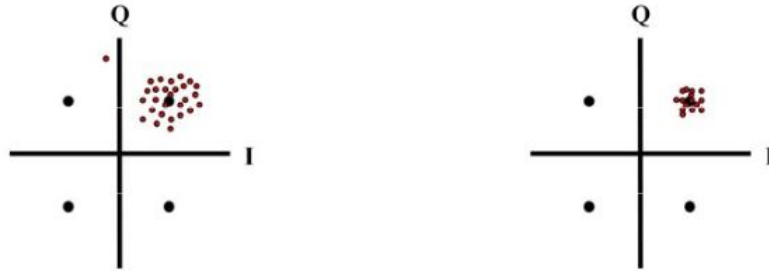


**Figure 2.5.:** Approximation of bit error probabilities of M-QAM over AWGN channel (for no channel coding, i.e.,  $R_c = 1$ )

As shown in Figure 2.3, modulation error is expressed as the error vector in the IQ plane. MER formular can be presented as

$$\begin{aligned} \text{MER}(\text{dB}) &= 10 \log \frac{P_{\text{average signal}}}{P_{\text{average error}}} \\ &= 10 \log \left( \frac{\sum_{j=1}^N (\tilde{I}_j^2 + \tilde{Q}_j^2)}{\sum_{j=1}^N [(I_j - \tilde{I}_j)^2 + (Q_j - \tilde{Q}_j)^2]} \right) \end{aligned} \quad (2.11)$$

where  $N$  is the number of measurements,  $\tilde{I}_j$  and  $\tilde{Q}_j$  are the ideal components of  $j$ -th symbol;  $I_j$  and  $Q_j$  are the measured components of  $j$ -th symbol. This definition assumes that a long enough sample is taken so that all the constellation symbols are equally likely to occur. So the higher MER value means the better RF modulation performs. Figure 2.6 expresses a large "cloud" of symbol points equates to low MER value, which is not good. The right part is what is wanted, small "cloud" and high MER value.



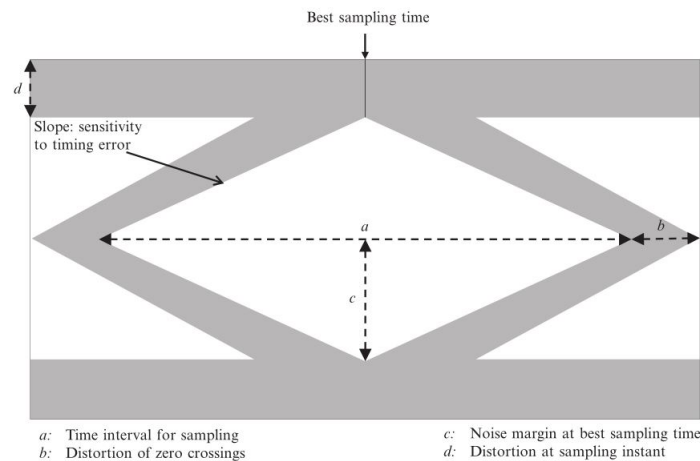
**Figure 2.6.:** Low MER and high MER

### 2.2.7. Eye Pattern

Eye pattern generation is a straightforward and visual figure of merit. It provides an effective tool for the severity of ISI, sensitivity to timing errors and noise margin.

The eye pattern is produced by displaying the received signal on an oscilloscope. The time base of the scope is triggered at a fraction of the bit rate, and it yields a sweep lasting several bit intervals. The superposition of many traces of bit intervals is showed by oscilloscope, which is the eye pattern.

Monitoring of an eye pattern can provide a qualitative measure of performance based on the signal quality. The figure 2.7 shows the important observations of the eye pattern[3]. The width of the eye opening represents the time interval during which the received signal can be sampled without error from ISI. The wider the eye is opened, the better the time to sample is. When there is no ISI, an eye is opened unitively; when there is a significant amount of ISI, the eye is completely closed. In a linear system with truly random data, all the eye openings would be identical, meanwhile, asymmetries in the eye opening generally indicate nonlinearities in the transmission channel.



**Figure 2.7.:** Interpretation of eye pattern

### 2.2.8. Noise Power Ratio (NPR)

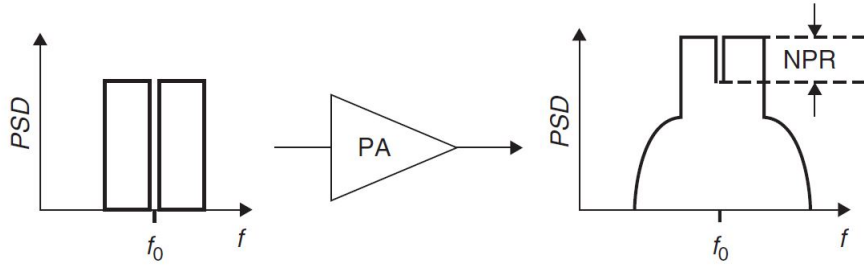
To get an accurate indication of IMD performance, Noise power ratio is developed as an indirect indication of characterizing cochannel distortion. It have been used to measure in-band distortion of nonlinear PAs. As Figure 2.8shows, in this method, white noise is used to simulate the presence of many carriers of random amplitude and phase. For many systems, including those that process multicarrier signals, white Gaussian noise is a close approximation for the real-world signals. This is a result of the Central Limit Theorem, which states that the probability distribution of a sum of

a large number of random variables will approach the Gaussian distribution. The white Gaussian noise is first passed through a bandpass filter (BPF) and then a narrow band reject filter to produce a deep notch at the center of the noise pedestal as the input of the DUT. If the notch bandwidth is sufficiently narrow, power spectral density function, observed at the output within the notch position, constitutes spectral regrowth, which is the desired cochannel distortion.

Noise power ratio is therefore defined as

$$NPR(\omega_T) = \frac{S_o(\omega_T)}{S_{wd}(\omega_T)}, \quad (2.12)$$

where  $S_o(\omega_T)$  and  $S_{wd}(\omega_T)$  are the output power spectral density functions measured in the neighbor of the test window position and within that window.



**Figure 2.8.:** Definition of NPR

Therefore, NPR represents the ratio of the output power to the effective in-band distortion power within the notch bandwidth. Despite having assumed a continuous spectrum, the NPR can also be implemented with a very large number of uncorrelated multitone signals.

## 2.3. Classification of Systems

To measure the quality of signal quantity, the constellation error of a I-Q modulation signal can be expressed by EVM measurement. EVM has more information than only SNR or BER. In this report some causes of errors are introduced and classified by different categories as Figure 2.9 describes.

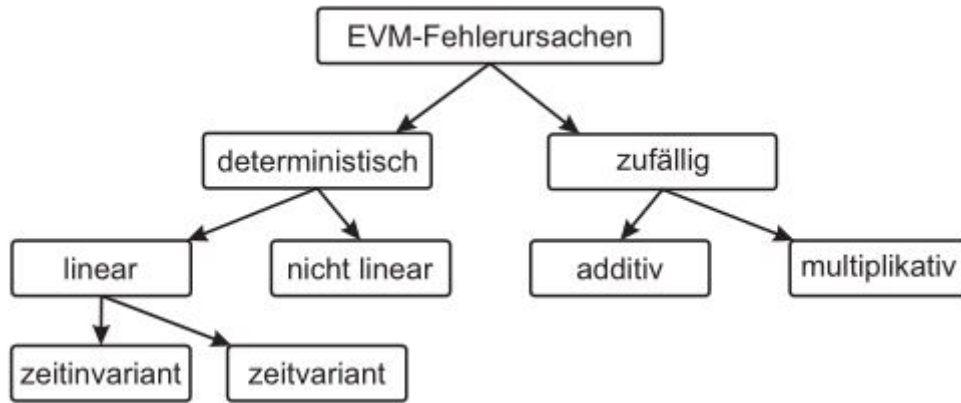


Figure 2.9.: Classes of EVM Error Causes

### 2.3.1. Deterministic System

A deterministic system is a fully-defined function of the variable time. A deterministic system is a system which there is no uncertainty with respect to its value, and every output value can be exactly determined according to the given input time. **description**

### 2.3.2. Linear and Nonlinear System

In a linear system the superposition principle is satisfied. The superposition principle is based on the additivity and homogeneity properties. A system is linear if and only if the following is satisfied:

$$\text{If } x_1(t) \rightarrow y_1(t) \text{ and } x_2(t) \rightarrow y_2(t), \text{ then } \alpha x_1(t) + \beta x_2(t) \rightarrow \alpha y_1(t) + \beta y_2(t), \quad (2.13)$$

in which  $\alpha$  and  $\beta$  are some nonzero constants.

If the superposition law is not satisfied, the system is nonlinear system.

### 2.3.3. Time-Invariant and Time-Varying System

In a time-invariant system, the input-output relationship does not change with time. This means that a time shift in the input results in a corresponding time shift in the output. A system is time

invariant if and only if the following is satisfied:

$$\text{If } x(t) \rightarrow y(t), \text{ then } x(t - \tau) \rightarrow \alpha y(t - \tau), \quad (2.14)$$

where  $\tau$  is a real constant. A system that is not time-invariant is called a time-varying system.

#### 2.3.4. Linear Time-Invariant (LTI) System and Linear Time-Varying (LTV) System

In a linear time-invariant (LTI) system, both linearity and time-invariance conditions must be satisfied. For LTI system, the system response of the impulse input can completely characterize the system and provide all relevant information to describe the system behavior for any input.

#### 2.3.5. Linear Time-Varying (LTV) System

If the Parameters of a linear system depend on time, this system is a linear time-varying system. That means a LTV system can be described as

$$y(t) = f[x(t), t] = a(t)x(t). \quad (2.15)$$

#### 2.3.6. Stochastic System

A digital communication system is influenced by random noise like Gaussian noise. Besides the transfer of time-varying signals, such as voice signals, multimedia information are unpredictable. Hence the study of random system is meaningful.

A stochastic process is assigning to every outcome  $\xi$  a function  $x(t, \xi)$ . Thus a stochastic process is a function of  $t$  and  $\xi$ . The domain of  $\xi$  is the set of all experimental outcomes and the domain of  $t$  is a set  $R$  of real numbers.

Given a stochastic process  $x(t)$ , if another process

$$y(t) = T[x(t)] \quad (2.16)$$

is created, whose samples are the function  $y(t, \xi_i)$ , the process  $y(t)$  can be considered as the output of a system with input  $x(t)$ . The system is completely specified in terms of the operate  $T$  as the rule of correspondence between the samples of the input  $x(t)$  and the output  $y(t)$ .

The system is *deterministic* if it operates only on the variable  $t$  and treats  $\xi$  as a parameter. This means if two samples  $x(t, \xi_1)$  and  $x(t, \xi_2)$  of the input are identical in  $t$ , then the corresponding samples  $y(t, \xi_1)$  and  $y(t, \xi_2)$  of the output are also identical in  $t$ . Otherwise, the system is called



*stochastic* if  $T$  operates on both variables  $t$  and  $\xi$ . This means if two samples  $x(t, \xi_1)$  and  $x(t, \xi_2)$  of the input are identical in  $t$ , but  $y(t, \xi_1) \neq y(t, \xi_2)$ .

**Memoryless Systems** A system is called memoryless system if its output

$$y(t) = T[x(t)] \quad (2.17)$$

at a given time  $t = t_1$ , the output  $y(t_1)$  depends only on  $x(t_1)$  and not on any other past or future values of  $x(t)$ .

If a signal is known only in terms of statistical averages and probabilistic description, such as its mean value, mean square value, and distribution. The collection of  $k^{th}$  – order joint cdfs or pdfs are

$$\begin{aligned} F_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k) &= P[X(t_1) \leq x_1, \dots, X(t_k) \leq x_k] \\ f_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k) &= \frac{\partial^k F_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k)}{\partial x_1 \dots \partial x_k} \end{aligned} \quad (2.18)$$

Still don't know what should be mentioned in Stochastic System, there is no books talking about classification in stochastic systems, except memoryless, linear system

The **autocorrelation and autocovariance functions of the random process  $X(t)$**  are as follows:

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x_1 x_2 f_{X(t_1)X(t_2)}(x_1, x_2) dx_1 dx_2 = R_X(\tau) \quad (2.19)$$

with  $\tau = t_2 - t_1$ , and

$$\begin{aligned} C_X(t_1, t_2) &= E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))] \\ &= R_X(t_1, t_2) - m_X(t_1)m_X(t_2) \end{aligned} \quad (2.20)$$

To analyze deterministic time-domain signals in the frequency domain, using the Fourier transform, the **power spectral density** is

$$S_X(f) = F[R_X(\tau)] = \sum_{-\infty}^{\infty} R_X(\tau) e^{2\pi f \tau} d\tau \quad (2.21)$$

## 2.4. Deterministic Impairments

### 2.4.1. IQ Imbalance

For down-conversion progress in receiver, RF signal should be multiplied by  $\cos \omega_c t$  and  $\sin \omega_c t$  signal, in order to get I part and Q part separately. Error in the  $90^\circ$  shift circuit and mismatches between the quadrature mixers result in imbalances in the amplitudes and phases of the baseband I and Q outputs. The baseband stages themselves may also contribute significant gain and phase mismatches.

The quadrature mismatches occurs because (1) the propagation through quadrature mixers experiences mismatches, a delay mismatch of 10 ps between the two mixers translates to a phase mismatch of  $18^\circ$  at 5 GHz; (2) the quadrature phases of the LO itself are introduced.

To gain insight into the effect of I/Q imbalance, consider a QPSK signal, the input is

$$x_{in} = a \cos \omega_c t + b \sin \omega_c t \quad (2.22)$$

Because of the gain and phase mismatches shown above in the LO path,

$$\begin{aligned} x_{IO,I}(t) &= 2(1 + \frac{\epsilon}{2}) \cos(\omega_c t + \frac{\theta}{2}) \\ x_{LO,Q}(t) &= 2(1 - \frac{\epsilon}{2}) \sin(\omega_c t - \frac{\theta}{2}) \end{aligned} \quad (2.23)$$

where  $\epsilon$  and  $\theta$  represent the amplitude and phase mismatches respectively. The following baseband signals after low-pass filters are:

$$\begin{aligned} x_I(t) &= a(1 + \frac{\epsilon}{2}) \cos \frac{\theta}{2} - b(1 + \frac{\epsilon}{2}) \sin \frac{\theta}{2} \\ x_Q(t) &= -a(1 - \frac{\epsilon}{2}) \sin \frac{\theta}{2} + b(1 - \frac{\epsilon}{2}) \cos \frac{\theta}{2} \end{aligned} \quad (2.24)$$

The output  $x_I$  and  $x_Q$  contain both gain impairments and phase impairments. This impairment relies not on time and the system characters don't depend on the input, which shows it's a deterministic and time-invariant system.

If  $x_{in,1} = a_1 \cos \omega_c t + b_1 \sin \omega_c t$  and  $x_{in,2} = a_2 \cos \omega_c t + b_2 \sin \omega_c t$ , then the output of  $x_{in,1} + x_{in,2}$  are

$$\begin{aligned} & a_1 \cos \frac{\theta}{2} - b_1(1 + \frac{\epsilon}{2}) \sin \frac{\theta}{2} + a_2(1 + \frac{\epsilon}{2}) \cos \frac{\theta}{2} - b_2(1 + \frac{\epsilon}{2}) \sin \frac{\theta}{2} \\ &= (a_1 + a_2)(1 + \frac{\epsilon}{2}) \cos \frac{\theta}{2} - (b_1 + b_2)(1 + \frac{\epsilon}{2}) \sin \frac{\theta}{2} \end{aligned} \quad (2.25)$$

which is the output of the signal  $x_{in,1} + x_{in,2}$ .

Hence, IQ Imbalance is a det. LTI impairment.

### 2.4.2. Carrier Feedthrough/Leakage, IQ Offset, DC Offset

Since in a homodyne topology the downconverted band extends to zero frequency, extraneous offset voltages can corrupt the signal and saturate the following stages. Consequently, the output signal is

$$V_{out}(t) = [A(t) \cos \phi + V_{OS1}] \cos \omega_c t - [A(t) \sin \phi + V_{OS2}] \sin \omega_c t \quad (2.26)$$

where  $V_{OS1}$  and  $V_{OS2}$  denote the dc offsets from the input port of the mixers. The unconverter output therefore contains the term of the unmodulated carrier:

$$V_{out}(t) = A(t) \cos(\omega_c t + \phi) + V_{OS1} \cos \omega_c t - V_{OS2} \sin \omega_c t \quad (2.27)$$

The **carrier leakage** is quantified as

$$\text{Relative Carrier Leakage} = \frac{\sqrt{V_{OS1}^2 + V_{OS2}^2}}{\sqrt{A^2(t)}} \quad (2.28)$$

It appears to the horizontal and vertical shifts in the constellation.

If one carrier leakage pair is  $V_{OS1,1}$  and  $V_{OS2,1}$ , the other carrier leakage pair is  $V_{OS1,2}$  and  $V_{OS2,2}$ , then the sum of the outputs is

$$\begin{aligned} V_{out}(t) &= A(t) \cos(\omega_c t + \phi) + V_{OS1,1} \cos \omega_c t - V_{OS2,1} \sin \omega_c t + \\ &\quad A(t) \cos(\omega_c t + \phi) + V_{OS1,2} \cos \omega_c t - V_{OS2,2} \sin \omega_c t \\ &= 2A(t) \cos(\omega_c t + \phi) + (V_{OS1,1} + V_{OS1,2}) \cos \omega_c t - (V_{OS2,1} + V_{OS2,2}) \sin \omega_c t \end{aligned} \quad (2.29)$$

which is the output of carrier leakage pair  $V_{OS1,1} + V_{OS1,2}$  and  $V_{OS2,1} + V_{OS2,2}$  and independent of the time. Hence, Carrier Feedthrough/Leakage is a LTI impairment.

### 2.4.3. Amplitude/Phase Response depending on frequency

A LTI filter for the in-band frequencies can be simplified as its low-pass equivalent (LPE). The transfer function in the frequency domain is

$$H(\omega) = |H(\omega)| e^{-j[\theta_0(\omega) + \delta\theta]}. \quad (2.30)$$

Where  $|H(\omega)|$  is the filter attenuation,  $\theta_0(\omega)$  the linear part of the phase and  $\delta\theta$  a phase offset or a phase shift in the in-band frequencies. Assuming the original I-Q signal is  $|H_s(\omega)| e^{-j\theta_s(\omega)}$ , the I-Q signal with LTI distortion is

$$X(\omega) = |H_s(\omega)H(\omega)| e^{-j[\theta_s(\omega) + \theta_0(\omega) + \delta\theta]}. \quad (2.31)$$

Above shows  $|H(\omega)|$  and  $\delta\theta$  will increase the EVM, bringing a compression and a rotation in the constellation diagram respectively. Since these effects are independent of time and the signal form, this impairment is deterministic LTI impairment.

### 2.4.4. Multipath Propagation

One of the major challenges in wireless communications is the frequency-selectivity of the channel caused by multipath propagation, with signal echoes arriving at the receiver antenna with delays similar to, or greater than the symbol duration  $T = 1/R_s$  of the transmitted signal.

It is very likely to happen if we increase the symbol rate  $R_s = 1/T$  (to obtain higher data rates) and then, the channel is no longer frequency-flat, but frequency-selective.

The multipath channel can be modelled as an FIR filter (tapped delay line) that causes intersymbol interference (ISI). The discrete time channel impulse response with  $L + 1$  taps can be described by its z-transform as

$$H(z) = \sum_{k=0}^L h_k \cdot z^{-k} \quad (2.32)$$

or in time domain

$$y(t) = \sum_{k=0}^{\infty} a_k x(n-k) \quad (2.33)$$

The input-output relationship of  $x_1$  and  $x_2$  are

$$\begin{aligned} x_1(n) &\rightarrow \sum_{k=0}^{\infty} a_k x_1(n-k) \\ x_2(n) &\rightarrow \sum_{k=0}^{\infty} a_k x_2(n-k) \end{aligned} \quad (2.34)$$

The sum of them is

$$\sum_{k=0}^{\infty} a_k x_1(n-k) + \sum_{k=0}^{\infty} a_k x_2(n-k) = \sum_{k=0}^{\infty} a_k [x_1(n-k) + x_2(n-k)] \quad (2.35)$$

which is the output of  $x_1 + x_2$ . Besides the parameters depend not on time and input.

Hence multipath propagation is a det. LTI impairment.

#### 2.4.5. Intersymbol Interference (LTV)

A digital system consists of random sequence of ONES and ZEROS. Each ONE represented by an ideal rectangular pulse and each ZERO by the absence of such a pulse. If this sequence is applied to a low-pass filter, the output can be obtained as the superposition of the responses to each input bit. Each bit level is corrupted by decaying tails created by previous bits. This phenomenon leads to higher error rate in the detection of random wave forms that are transmitted through band-limited channels.

The input bits applied to a pulse generator produce the following signal:

$$x(t) = \sum_{k=-\infty}^{\infty} a_k g(t - kT_b) \quad (2.36)$$

where  $g(t)$  is the basic shaping pulse with  $g(0) = 1$ , and the pulse amplitude  $a_k$  depends on the input bit.

The signal  $x(t)$  is transferred through transmitting filter  $H_T(f)$ , Channel  $H_C(f)$  added by Additive white Gaussian noise (AWGN), Receiving filter  $H_R(f)$ . The receiver signal in frequency domain is

$$R(f) = a_k G(f) H_T(f) H_C(f) H_R(f) + N(f) = A_k H(f), \quad (2.37)$$

where  $G(f)$  is the Fourier transform of the pulse  $g(t)$ . The received signal  $r(t)$  is sampled synchronously at  $t_m = mT_b$ , with  $m$  as integer values. The signal in time domain is

$$r(t_m) \sum_{k=-\infty}^{\infty} A_k h(mT_b - kT_b) + n(t_m) = A_m + \sum_{\substack{k=-\infty \\ k \neq m}}^{\infty} A_k h((m-k)T_b) + n(t_m) \quad (2.38)$$

The first term on the right-hand side of the expression above is the desired transmitted bit, and the second term represents the residual effect of all other transmitted bits on the  $m^{th}$  bit. This residual effect is called **intersymbol interference (ISI)**. The last term represents the noise sample.

The ISI term  $\sum_{\substack{k=-\infty \\ k \neq m}}^{\infty} A_k h((m-k)T_b)$  satisfied the form  $y(t) = a(t)x(t)$  mentioned above and therefore is LTV system.

#### 2.4.6. Intermodulation distortion (det. nonlinear)

In order to study the phenomena from nonlinearity, input/output characteristic can be approximated by

$$y(t) \approx \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t). \quad (2.39)$$

Therefore if two interferers at  $\omega_1$  and  $\omega_2$  are the input of a nonlinear system, the output is not harmonics of these frequencies. This phenomenon is called "intermodulation (IM)". Assuming input  $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$ , the output according the equation above is

$$y(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t). \quad (2.40)$$

Expanding the right-hand side and discarding the direct current terms, the results are the terms with the frequencies in

$$\left\{ \begin{array}{l} \omega_1 - \omega_2 \\ 2\omega_1 - \omega_2, \quad \omega_1, \quad \omega_2, \quad 2\omega_2 - \omega_1 \\ \omega_1 + \omega_2 \\ 2\omega_1 + \omega_2, \quad 2\omega_2 + \omega_1. \end{array} \right. \quad (2.41)$$

It is very likely that the small desired signal at  $\omega_0$  with  $2\omega_1 - \omega_2 \approx \omega_0$ . Consequently, the

intermodulation product at  $2\omega_1 - \omega_2$  falls onto the desired channel, corrupting the signal.

The term in the frequency  $2\omega_1 - \omega_2$  is  $\frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 - \omega_2)t$ . Therefore IM is nonlinear distortion.

## 2.5. Stochastic Impairments

### 2.5.1. Phase Noise(random and multiplicative)

An ideal oscillator produces a perfectly-periodic output of the form  $x(t) = A \cos \omega_c t$ . The zero crossings occur at exact integer multiples of  $T_s = 2\pi/\omega_c$ . In reality, however, the noise of the oscillator devices randomly perturbs the zero crossings, like  $x(t) = A \cos[\omega_c t + \phi_n(t)]$ , where  $\phi_n(t)$  is a small random phase quantity that deviates the zero crossings from integer multiples of  $T_c$ . The term  $\phi_n(t)$  is called the "phase noise". The frequency departs from  $\omega_c$  occasionally. As a consequence, the impulse is "broadened" to represent this random departure. Since  $\phi_n(t) \ll 1 \text{ rad}$

$$LO_{\cos}(t) = A \cos[\omega_c t + \phi_n(t)] \quad (2.42)$$

After  $90^\circ$  delay the sin signal is

$$LO_{\sin}(t) = A \sin[\omega_c t + \phi_n(t)] \quad (2.43)$$

In the receiver front-end for down-conversion situation, referring the ideal case, the desired channel is convolved with the impulse at  $\omega_{LO}$ , yielding an IF signal at  $\omega_{IF} = \omega_{RF} - \omega_{LO}$ . However, with consideration of phase noise from LO, the convolution of the desired signal and the interferer with the noisy LO results in a broadened down-converted interferer whose noise skirt corrupts the desired IF signal. This phenomenon is called "reciprocal mixing".

The IQ signal after superposition is

$$\begin{aligned} x(t) &= I \cdot LO_{\cos} - Q \cdot LO_{\sin} \\ &= I \cdot A \cos[\omega_c t + \phi_n(t)] - Q \cdot A \sin[\omega_c t + \phi_n(t)] \end{aligned} \quad (2.44)$$

After the multiplication in receiver and low pass filter, the signal with I part as real part and Q part as imaginary part are

$$\begin{aligned} y(t) &= x(t) \cos(\omega_c t) + i \cdot x(t) \sin(\omega_c t) \\ &= \frac{1}{2} I e^{-i\phi} + \frac{1}{2} Q e^{-i\phi} \cdot i + \frac{1}{2} I e^{i(2\omega_c t + \phi(t))} - \frac{1}{2} Q e^{i(2\omega_c t + \phi(t))} \cdot i \\ &= \frac{I}{2} e^{-i\phi} + \frac{Q}{2} e^{-i\phi} \cdot i + \text{high frequency signal} \end{aligned} \quad (2.45)$$

The part  $e^{-i\phi}$  shows this is a phase noise and have influence on phase instead of amplitude. It is common to treat  $\phi(t)$  as a zero-mean stationary random process[8], whose variance is defined as the LO mean square phase error[5]. The description in frequency domain about phase noise is the power spectral density  $S_{\phi}(\omega_m)$  of the phase  $\phi(t)$ . Hence phase noise is a stochastic system.

### 2.5.2. Thermal Noise(random and additive)

Due to the Brown motion of electrons in a conductor, electrical equipment is always affected by thermal noise. Thermal noise has a flat power spectral density over a very wide frequency range and is said to be white since all frequencies are equally represented. Thermal noise depends only on the temperature.

Amount of thermal noise to be found in a bandwidth of 1Hz in any device or conductor is

$$N_0 = kT(W/Hz), \quad (2.46)$$

where  $N_0$  is noise power density in watts per 1 Hz of bandwidth. Parameter  $k = 1.38 \times 10^{-23} J/K$  is Boltzmann's constant. T is temperature in klevin.



### 3. Theory of Post Processing Methods

At this point it is clear that channel estimation and equalization are necessary at the receiver port. This can be realized with or without knowing prior knowledge, such as channel parameters or some fixed training signals. The former is called *supervised learning* while the latter is *blind learning*. As mentioned above, since the deterministic and stochastic impairments exist together, the corresponding post methods are introduced by both supervised and blind ways. In this chapter, adaptive filters are introduced basically for deterministic impairments and some basic machine learning methods are aiming to stochastic impairments.

#### 3.1. Adaptive Filter Theory

The term filter is described for a hardware or software, whose goal is to extract information of interest from a set of noisy data. The data turns to be noisy by transmission through a communication channel and detected by means of noisy sensors. These noises and impairments can be both linear and nonlinear. They may be modeled by a common statistical model such as Gaussian statistical model, or random noise with unknown statistics, or may be correlated with the desired signal itself. Filtering means during the extraction of information at time  $t$  only measured data up to and including time  $t$  can be used.

The filters can be classified into linear and nonlinear. A linear filter means the output of filter device is a linear function of the filter input. Otherwise, the filter is nonlinear.

Adaptive filters are based on an approach include priori knowledge and some general models. When a priori knowledge of a dynamic process and its statistics is limited then the use of adaptive filters can offer performance improvements.

If the knowledge of the signal and noise isn't available, by using a recursive algorithm the adaptive filter can also help. If the signal and noise are stationary, the adaptive filter would converge to the optimum Wiener solution. If they are nonstationary, the adaptation rate must be faster than the rate of change.

Important features of AF: rate of convergence, misadjustment, tracking, robustness, compu-

tational requirements, structure, numerical properties.

The baseband model of QAM system is expressed by

$$x(k) = \sum_{i=0}^{M-1} h(i)a(k-1)e^{j\phi(k)} + n(k), \quad (3.1)$$

where  $h(k)$  is the overall complex base band impulse response of the transmitter filter, channel and receiver filter. The input data sequence  $a(k)$  is assumed to be independent, identically distributed (i.i.d.) random variable with zero mean.

### 3.1.1. Linear Adaptive Filters

For linear impairments, if the  $n$ th symbol  $\sum b_n u(t - nT)$  is to be sent, the received waveform in  $z$ -domain is a linear combination of

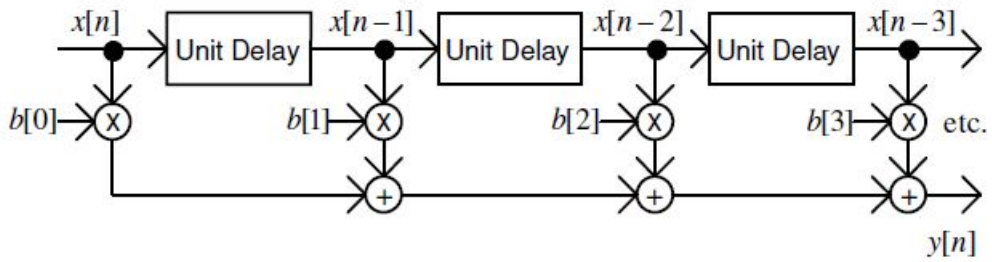
$$X(z) = \sum_{k=-\infty}^{\infty} b_k z^{-k} U(z). \quad (3.2)$$

If the received waveform is filtered with a FIR filter

$$H(z) = \frac{1}{\sum_{k=-\infty}^{\infty} b_k z^{-k}}, \quad (3.3)$$

this linear impairment is fully removed. In order to complete the channel coefficient  $b_n$ , a transversal structure is used with recursive algorithm.

Transversal structure is very successfully employed for practical systems. It is similar to the linear FIR filter structure as Figure 3.1.



**Figure 3.1.:** FIR Transversal Structure

The model can be described as

$$y[n] = \sum_{k=0}^M b[k]x[n-k] \quad (3.4)$$

In this model,  $b[n]$  is the weights of different delay paths as the filter coefficients. This filter can be made into an adaptive filter by calculating the coefficients  $b[n]$  progressively at each time step.

### 3.1.2. Supervised Adaptive Filter Algorithm

Supervised adaptive filter means in order to get channel coefficient  $a_n$ , a series of ideal data  $d_n$  should first be given as priori knowledge of reference.

One of the mostly well-used objective function adaptive filter is the **mean-square error (MSE)**, defined as

$$\xi = E[e^2(k)] = E[d^2(k) - 2d(k)y(k) + y^2(k)], \quad (3.5)$$

where  $d(n)$  is the reference signal.

For the FIR transversal structure above, the output signal is composed by a linear combination of inputs as

$$y(k) = \sum_{i=0}^N w_i(k)x_i(k) = \mathbf{w}^T(k)\mathbf{x}(k), \quad (3.6)$$

where  $\mathbf{x}(k) = [x(k)x(k-1)\dots x(k-N)]^T$  is the input vector with a tapped-delay line, and  $\mathbf{w}(k) = [w_0(k)w_1(k)\dots w_N(k)]^T$  as weight factor of each tap.

Hence the MSE function in a stationary environment is given by

$$\xi = E[d^2(k)] - 2\mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R} \mathbf{w}, \quad (3.7)$$

where  $\mathbf{p} = E(s[n]\mathbf{x}[n])$  is the cross-correlation vector between the desired and input signals, and  $\mathbf{R} = E(\mathbf{x}[n]\mathbf{x}[n]^T)$  is the input signal correlation matrix.

The gradient vector of MSE function is given by

$$\mathbf{g}_w = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}. \quad (3.8)$$

By equating the gradient vector  $\mathbf{g}_w$  to zero, the solution is

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}. \quad (3.9)$$

This solution is called Wiener solution. However, in practice, accurate estimations of correlation matrix  $\mathbf{R}$  and cross-correlation vector  $\mathbf{p}$  are not available. If the input  $\mathbf{x}(n)$  and the reference signal  $d(n)$  are ergodic, it is popular performed by most adaptive algorithms to use time averages to estimate  $\mathbf{R}$  and  $\mathbf{p}$ .

The estimation is

$$\begin{aligned}\hat{s}(n) &= \mathbf{h}^T \mathbf{x} \\ &= \mathbf{p}^T \mathbf{R}^{-1} \mathbf{x} \\ &= E[s(n)\mathbf{x}^T(n)]E[\mathbf{x}(n) \cdot \mathbf{x}^T(n)]^{-1} \mathbf{x}(n)\end{aligned}\tag{3.10}$$

### 3.1.2.1. Least-Mean-Square (LMS) Algorithm

The least-mean-square (LMS) is a search algorithm by using the gradient vector computation in order to approach the target point iteratively. The LMS algorithm is widely used in various adaptive filters. The advantages of LMS in Wiener filter are low computational complexity, unbiased convergence in the mean to the Wiener solution and stable behavior.

As mentioned above, the optimal Wiener solution is  $\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}$ . If good estimates of  $\mathbf{R}$  and  $\mathbf{p}$  denoted by  $\hat{\mathbf{R}}(k)$  and  $\hat{\mathbf{p}}(k)$  are available, a steepest-descent-based algorithm can be used as

$$\begin{aligned}\mathbf{w}(k+1) &= \mathbf{w}(k) - \mu \hat{\mathbf{g}}_{\mathbf{w}}(k) \\ &= \mathbf{w}(k) + 2\mu(\hat{\mathbf{p}}(k) - \hat{\mathbf{R}}(k)\mathbf{w}(k))\end{aligned}\tag{3.11}$$

for  $k = 0, 1, 2, \dots$ , where  $\hat{\mathbf{g}}_{\mathbf{w}}(k)$  is an estimate of the gradient vector (equation 3.8) of the objective function with respect to  $\mathbf{w}$ .

One widely used estimations of  $\mathbf{R}$  and  $\mathbf{p}$  are as follows:

$$\begin{aligned}\hat{\mathbf{R}}(k) &= \mathbf{x}(k)\mathbf{x}^T(k) \\ \hat{\mathbf{p}}(k) &= d(k)\mathbf{x}(k)\end{aligned}\tag{3.12}$$

The resulting gradient estimate  $\hat{\mathbf{g}}_{\mathbf{w}}(k)$  is expressed by

$$\begin{aligned}\hat{\mathbf{g}}_{\mathbf{w}}(k) &= 2\mathbf{x}(k)(-d(k) + \mathbf{x}^T(k)\mathbf{w}(k)) \\ &= -2e(k)\mathbf{x}(k)\end{aligned}\tag{3.13}$$

Then we get **LMS algorithm** :

First as initialization set  $\mathbf{x}(0) = \mathbf{w}(0) = [0 \ 0 \ \dots \ 0]^T$ .

Then for  $k = 0, 1, 2, \dots$

$$\begin{aligned} e(k) &= d(k) - \mathbf{x}^T(k) \mathbf{w}(k) \\ \mathbf{w}(k+1) &= \mathbf{w}(k) + 2\mu e(k) \mathbf{x}(k) \end{aligned} \quad (3.14)$$

### 3.1.2.2. Recursive Least-Squares (RLS) Algorithm

In LMS algorithm, for each iteration only the information of time interval  $n = k$  is used. However, if not only last point, but also several previous points are used together, the convergence speed may be much faster. This idea is recursive least-squares (RLS) algorithm.

The RLS algorithms performs well with its fast convergence even when the eigenvalue spread of the input signal correlation matrix is large. Hence it has excellent performance for estimating the time-varying parameters. On the other hand, all these advantages take the cost of an increased computational complexity and larger possibility of unstable than LMS-based algorithm.

Simillar to LMS, assume the input signal information vector at a given instant  $k$  is given by

$$\mathbf{x}(k) = [x(k) \ x(k-1) \dots x(k-N)]^T \quad (3.15)$$

where  $N$  is the order of the filter. The weight coefficients  $w_j(k)$ , for  $j = 0, 1, \dots, N$ , the objective function is given by

$$\begin{aligned} \xi^d(k) &= \sum_{i=0}^k \lambda^{k-i} \epsilon^2(i) \\ &= \sum_{i=0}^k \lambda^{k-i} [d(i) - \mathbf{x}^T(i) \mathbf{w}(k)]^2 \end{aligned} \quad (3.16)$$

where  $\mathbf{w}(k) = [w_0(k) \ w_1(k) \dots w_N(k)]^T$  is the adaptive-filter coefficient vector and  $\epsilon(i)$  is the posteriori output error at instant  $i$ . The parameter  $\lambda$  is an exponential weighting factor called forgetting factor that satisfies  $0 \ll \lambda \leq 1$ . The information far in the past has lower weight on the coefficient updating.

As we can observe, each error is the difference between the desired signal and the filter estimation using the most recent coefficients  $\mathbf{w}(k)$ . Because in each iteration the coefficient  $\mathbf{w}(k)$  is better than previous  $\mathbf{w}$ , the new error uses posteriori information and the algorithm consists of faster convergence.

By differentiating  $\xi^d(k)$  with respect to  $\mathbf{w}(k)$ , it is

$$\frac{\partial \xi^d(k)}{\partial \mathbf{w}(k)} = -2 \sum_{i=0}^k \lambda^{k-i} \mathbf{x}(i) [d(i) - \mathbf{x}^T(i) \mathbf{w}(k)] \quad (3.17)$$

By equating  $\frac{\partial \xi^d(k)}{\partial \mathbf{w}(k)}$  to zero, the optimal coefficient vector  $\mathbf{w}(k)$  is given by

$$\begin{aligned} \mathbf{w}(k) &= \left[ \sum_{i=0}^k \lambda^{k-i} \mathbf{x}(i) \mathbf{x}^T(i) \right]^{-1} \sum_{i=0}^k \lambda^{k-i} \mathbf{x}(i) d(i) \\ &= \mathbf{R}_D^{-1}(k) \mathbf{p}_D(k) \end{aligned} \quad (3.18)$$

where  $\mathbf{R}_D(k)$  and  $\mathbf{p}_D(k)$  are called the deterministic correlation matrix of the input signal and the deterministic cross-correlation vector between the input and desired signals.

Then we get **Conventional RLS algorithm** :

First as initialization set  $\mathbf{S}_D(-1) = \mathbf{R}_D^{-1}(-1) = \delta \mathbf{I}$  where  $\delta$  can be the inverse of the input signal power estimate

$$\mathbf{p}_D(-1) = \mathbf{x}(-1) = [0 \quad 0 \quad \dots \quad 0]^T$$

Then for  $k = 0, 1, 2, \dots$

$$\begin{aligned} \mathbf{S}_D(k) &= \mathbf{R}_D^{-1}(k) = \frac{1}{\lambda} \left[ \mathbf{S}_D(k-1) - \frac{\mathbf{S}_D(k-1) \mathbf{x}(k) \mathbf{x}^T \mathbf{S}_D(k-1)}{\lambda + \mathbf{x}^T(k) \mathbf{S}_D(k-1) \mathbf{x}(k)} \right] \\ \mathbf{p}_D(k) &= \lambda \mathbf{p}_D(k-1) + d(k) \mathbf{x}(k) \\ \mathbf{w}(k) &= \mathbf{S}_D(k) \mathbf{p}_D(k) \end{aligned} \quad (3.19)$$

### 3.1.3. Blind Adaptive Filtering

For supervised adaptive filter mentioned above, a series of ideal or required data is required to generate the error signal. However, there are lots of applications where the reference signals are not available. Furthermore, if the training data is not required, the pressure of transformation could be reduced and higher transformation speed can be reached. This kind of adaptive filter, which doesn't utilize reference signal is called unsupervised adaptive filter or blind adaptive filter. Since blind adaptive filters don't demand training data during learning, the algorithms are specific suitable for an unknown environment.

### 3.1.4. Circularity-Based Approach

For IQ Imbalance and immunity to channel noise and the fading channel [9]

### 3.1.4.1. $p$ -Constant-Modulus Algorithm

The CMA, proposed by Godard in 1980[10], is the most popular algorithm for blind adaptive filter of QAM communication.

The basic idea is to minimize the distance between the modulus of the equalizer output and some constant values[11]. In this way a reference signal is not required. These constant values, normally denoted by  $R_p$ , express the modulus of constellation symbols. The earlier blind algorithm was addressed the case of Pulse Amplitude Modulation (PAM). This method was later developed with objective function to apply to the case of Quadrature Amplitude Modulation (QAM)[12].

The cost function of CMA is expressed by

$$J(k) = E[|y(k)|^p - R_p]^2, \quad (3.20)$$

where  $E[\bullet]$  indicates statistical expectation. Employing a stochastic gradient of cost function  $J(k)$  with respect to the weights vector  $\mathbf{w}(k)$ , the iteration function is adapted by

$$\begin{aligned} w(k+1) &= w(k) - \mu \cdot \nabla J(k) \\ &= w(k) - \mu y(k) |y(k)|^{p-2} (|y(k)|^p - R_p) \mathbf{x}^*(k), \end{aligned} \quad (3.21)$$

where  $\mu$  is the step-size parameter and  $*$  denotes complex conjugation.  $R_p$  is the constant modulus depending only on the input data as[10]

$$R_p = \frac{E[|a(k)|^{2p}]}{E[|a(k)|^p]}. \quad (3.22)$$

Godard proved that this method is also suitable for rectangular QAM constellation although it's non-constant modulus.

## 3.2. Machine Learning Method

The demodulation process is basically a pattern classification problem, which is already extensively discussed in Machine Learning (ML) area. ML plays a role in various areas (exp. data mining, pattern recognition and classification, image processing) as a powerful tool with interdisciplinary. Lately, several classical ML methods are also amplified in demodulation process, such as  $k$ -NN Method [13], SVM Method [14], MLE Method.

Since all these Methods are supervised ML, a prefixed training data is necessary for further processing. Specifically for 16-QAM, various constellation points are regarded as 16 classes of data, where each received data point can be represented as a two-dimensional vector  $(I(\mathbf{x}), Q(\mathbf{x}))$ . The correct demodulation results are already known as a reference. The whole data points are

divided to two parts: a small amount with label  $y(\mathbf{x})$  as training data set and the rest as test data set. The basic idea is: With training data and corresponding labels as known information, the test set is classified to modulation points. Then compare the classify results with correct reference and calculate BER.

All ML Methods mentioned above will be discussed and applied to support demodulation process in phase noise distortion.

### 3.2.1. KNN Method

As a classification algorithm, KNN is used to classify the received signals to different constellations. It is straightforward and effective, especially for large data with low dimensions [13]. Hence, KNN is particular suitable for demodulation if it improves the appearance.

To begin with, the distances of each point in test set ( $\mathbf{x}_i$ ) to every points in train set  $\mathbf{x}_j$  need to be calculated as

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{(I(\mathbf{x}_i) - I(\mathbf{x}_j))^2 + (Q(\mathbf{x}_i) - Q(\mathbf{x}_j))^2}. \quad (3.23)$$

After the distances sorted in ascending order,  $k$  nearest training data points are determined as  $N_k(\mathbf{x}_i)$ . The majority of these  $k$  labels is regards as the predicted label for  $\mathbf{x}_i$ :

$$f_{KNN}(\mathbf{x}_i) = \arg \max_{l \in \mathbf{L}} \sum_{i \in N_k(\mathbf{x}_i)} \delta(l, l_i), \quad (3.24)$$

where  $\delta(l, l_i) = 1$  if  $l = l_i$  and  $\delta(l, l_i) = 0$  if  $l \neq l_i$ ,  $\mathbf{L} = \{1, 2, \dots, 16\}$  is the label set.

Furthermore, in order to get higher accuracy, the closer points in train set  $\mathbf{x}_j$  should have greater weights. With a weight vector  $\omega_j = 1/d(\mathbf{x}_i, \mathbf{x}_j)^2$  based on distance  $d(\mathbf{x}_i, \mathbf{x}_j)$ , a distance weighted KNN method is proposed as:

$$f_{\omega KNN}(\mathbf{x}_i) = \arg \max_{l \in \mathbf{L}} \sum_{i \in N_k(\mathbf{x}_i)} \omega_i \delta(l, l_i). \quad (3.25)$$

KNN is one of the most intuitive ML method. In next chapter the simulation results will be discussed and compared with other methods.



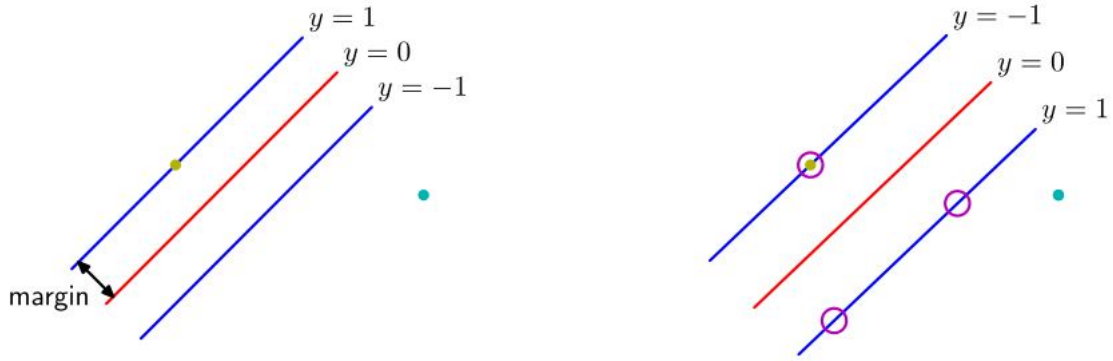


Figure 3.2.: Basic SVM Concept

### 3.2.2. SVM Method

SVM is another straightforward and well implicated decision machine. Basic SVM is a two-class classification with linear models of the form

$$y(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b, \quad (3.26)$$

where  $\phi(\mathbf{x})$  denotes a fixed feature-space transformation, and  $b$  is the bias parameter. Assume that train data set as  $N$  input vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  with corresponding labels  $t_1, t_2, \dots, t_N$  is linearly separable in feature space, where  $t_n \in \{-1, 1\}$ , and new data points  $\mathbf{x}$  are classified based on the sign of  $y(\mathbf{x})$ . According to the definition (linear separable) there exists at least one weight parameters vector  $\mathbf{w}$  and one bias parameter vector  $b$  such that  $y(\mathbf{x}) > 0$  for all  $\{\mathbf{x}_i | t_i = 1\}$  and  $y(\mathbf{x}) < 0$  for all  $\{\mathbf{x}_i | t_i = -1\}$ .

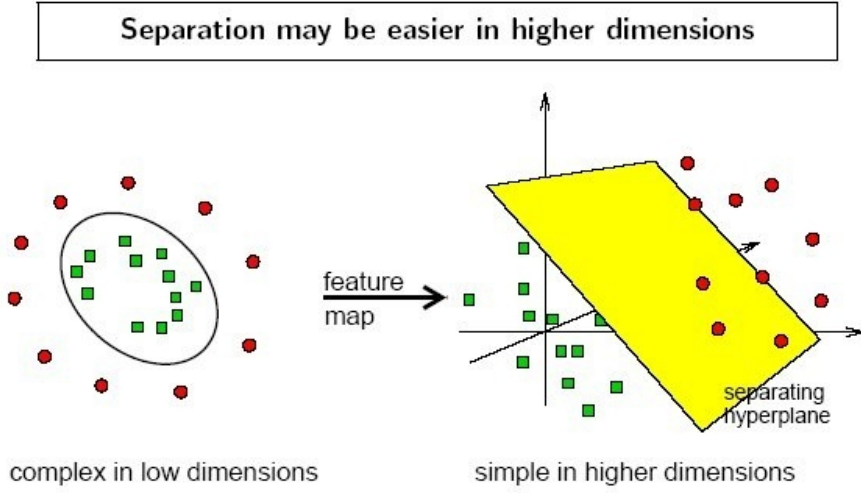
The Support Vector Machine (SVM) approaches this two-class classification problem with the concept *margin*, which is defined as the smallest distance between decision boundary and any samples like Figure 3.2.

Normally there exist several solutions for  $\mathbf{w}$  and  $b$ . By the definition of margin, the best solution is decided. According to [15], in SVM the decision boundary is chosen to maximize the margin. In a hyperplane the distance of a point  $\mathbf{x}$  from a hyperplane  $y(\mathbf{x}) = 0$  is given by

$$distance(\mathbf{x}, y) = \frac{|y(\mathbf{x})|}{\|\mathbf{w}\|}. \quad (3.27)$$

Thus the maximum margin solution is found by solving

$$\arg \max_{\mathbf{w}, b} \left\{ \frac{1}{\|\mathbf{w}\|} \min_n [t_n (\mathbf{w}^T \phi(\mathbf{x}_n) + b)] \right\} \quad (3.28)$$



**Figure 3.3.:** Mapping from 2-D to higher dimensions (3-D)

Since scaling makes no influence on solution[15], in order to simply the problem,  $(\mathbf{w}, b)$  is scaled to insure

$$\min_{\mathbf{w}, b} |\mathbf{w}^T \mathbf{x} + b| = 1. \quad (3.29)$$

Then the original problem 3.28 is transformed as

$$\begin{aligned} \operatorname{argmin}_{\mathbf{w}, b} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{s.t.} \quad & y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, \quad i = 1, 2, \dots, m. \end{aligned} \quad (3.30)$$

After that Lagrange Function, Dual Problem and Karush-Kuhn-Tucker(KKT) conditions are introduced[16], this problem becomes a convex linear optimization problem.

Until here, it is assumed that there exist a hyperplane, which can separate the training points with different labels. However, various tasks don't satisfy this condition. SVM solved the limitation of linear separation through so called *kernel trick*.

Since original vector space  $\mathbb{R}^d$  is linear inseparable, SVM uses a mapping relation  $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^{\hat{d}}$  to insure the training points linear separable in new vector space  $\mathbb{R}^{\hat{d}}$  like Figure 3.3. According to Vapnik [17] there must exists a  $\hat{d}$  if  $d$  is a limited number.

Since in SVM inner product is repeatedly used(dual problem), kernel trick converges mapping and inner product to one step, in order to reduce the compute complexity from  $O(\hat{d})$  to  $O(d)$  at the same time.

Finally, algorithm of SVM is as following:

- **Given**

- training samples  $\mathbf{x}_i$  and labels  $y_i$
- **Parameter**
  - kernel function  $K(\mathbf{x}_m, \mathbf{x}_n)$
- **Training**
  - compute  $\mathbf{Q} = [Q_{mn}]$  with  $Q_{mn} = y_m y_n K(\mathbf{x}_m, \mathbf{x}_n)$
  - compute  $\alpha$  by solving the convex quadratic problem  $\min_{\alpha} \frac{1}{2} \alpha^T \mathbf{Q} \alpha - \mathbf{1}^T \alpha$
  - determine the index set  $SV = \{n | \alpha_n > 0\}$  of support vectors
  - compute  $w_0 = \frac{1}{|SV|} (\sum_{n \in SV} y_n - \sum_{m \in SV} \sum_{n \in SV} \alpha_m y_n Q_{mn})$
- **Classification**
  - calculate the discriminant function  $f(\mathbf{x}) = \sum_{n \in SV} \alpha_n y_n K(\mathbf{x}_n, \mathbf{x}) + w_0$
  - estimate the class  $\hat{y}(\mathbf{x}) = \text{sgn}(f(\mathbf{x}))$



## 4. Filters Design and Simulation Results

In this chapter, assume that the received signal suffers from medium deterministic and stochastic impairments. The corresponding parameters are introduced in first section. The algorithms for channel equalizers including adaptive filters and ML methods are implemented in MATLAB. Original received signals are generated in Keysight Advanced Design System (ADS). The performances are compared and discussed based on EVM and BER. At last, through discussion of compute complexity and progress speed the possibility of applying is demonstrated.

In higher order modulation schemes, like 64-QAM, the distance between the constellation points is less so for the same amount of distortion. Therefore, higher order modulation schemes demand better EVM in comparison to lower order modulation, like QPSK. At this medium impairments, three signal constellations are considered: 4-QAM, 16-QAM, 64-QAM.

### 4.1. Received Signal Generator

ADS is the world's leading electronic design automation software for RF, microwave, high speed digital, and power electronics applications[18]. With easy-used interface, received signal w/o various impairments are generated in ADS.

In Figure 4.1 the relative parameters are in Table 4.1.

Transmitter	Modulation type	QPSK, 16-QAM, 64-QAM
	Modulation type for the reference signal	QPSK
	No. of frames per run	50
	No. of symbols per frame	1000
Channel	SNR	20 to 30
	IQ Imbalance	3db, 3 degree
	Multipath	5 dB
	Phase Noise	5 dB
Receiver	Adaptation step size	0.08
	No. of filter taps	6
	Adaptive methods	LMS, RMS, CMA, kNN, SVM,

**Table 4.1.:** Simulation Condition



### Figure 4.1.: ADS Circuit Design

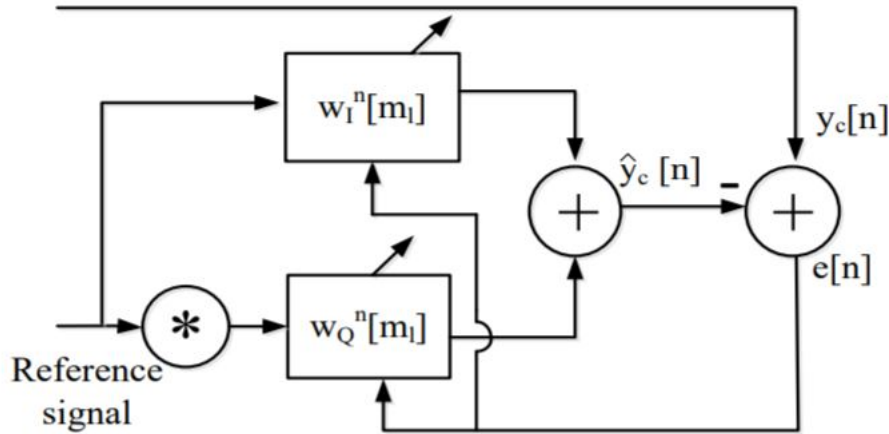


Figure 4.2.: **Replace by self made** Supervised Adaptive Compensator

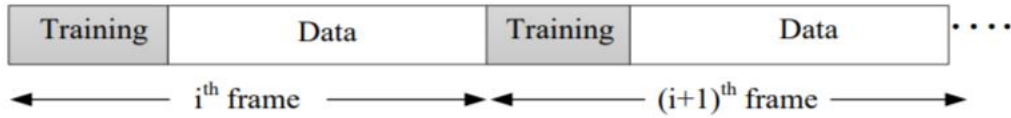


Figure 4.3.: A frame structure that contains reference training signal and transmitted data

## 4.2. Convergence Analysis

Convergence speed and results are the main part of adaptive filter performance. Hence, in this section we will discuss it in detail. Before using various adaptive filter, we remove kinds of offset in the first place. through simple data processing. We remove the constellation center of received data and proceed normalization through formula

$$\mathbf{x}_{\text{train}} = \frac{\mathbf{x}_{\text{train}} - x_{\text{mean}} \cdot \mathbf{1}}{x_{\text{mean}}}, \quad (4.1)$$

where  $\mathbf{x}_{\text{train}}$  is received training data set (column vector),  $x_{\text{mean}}$  is mean value of training data and  $\mathbf{1}$  is a column vector with  $\mathbf{x}_{\text{train}}$  size.

### 4.2.1. Supervised Adaptive Compensator: LMS and RMS

This section presents the adaptive method for IQ Imbalance and multipath propagation impairment. In this method, parameters of both impairments are estimated according to the functional block diagram of Figure 4.2.

Since a *priori* knowledge is required, the transmit data is arranged in a frame structure as shown in Figure 4.3. Each frame contains training symbols followed by data symbols.

Figure 4.4 (a) shows the constellation results scattered from the original location (the red circles)

due to the IQ imbalance and channel distortion. The black points are the constellation of received signal without compensation, while the blue points are the constellation after Compensator with LMS algorithms. From Figure 4.4 (b) shows the convergence speed in this compensation. After about 50 times iterations the learning line stays stable, which means the weight vector  $\mathbf{w}(x)$  is convergence and the algorithms can't do better. The distortion observed in constellations of the adaptive LMS compensation is much less than that of untreated data, which has been reflected later in the EVM and SER results.

The EVM structures for supervised adaptive filters LMS and RMS have been shown in Figure 4.5. It should be noted that the EVM and SER performances have been shown for a 64-QAM modulation scheme only. The legend 'Uncompensated' refers to the received data that has not been treated with any correction or detection method. The term 'Ideal' refers to the ideal case where the received signal  $x$  doesn't have any IQ or channel distortion. The only distortion for the ideal case is the AWGN.

The results show that supervised adaptive filter including LMS and RMS work really satisfying. With limited taps and iterations, the distortion for IQ imbalance and multipath channel distortions are almost faded away. In  $E_b/N_0 = 30$  dB, the EVM decreased from about 18% to 2%, which is already very close to the ideal signal with 2%.

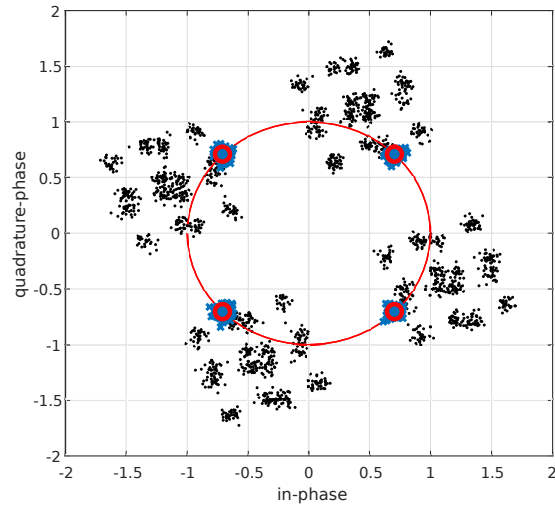
Specifically, RMS algorithms are tested more detailed as Figure 4.6 in various Gain Imbalances (in dB) and Phase Imbalance (in degree). Signal is generated in ADS software with given parameters Gain Imbalance from 0 dB to 6 dB and Phase Imbalance from 0 degree to 5 degree. The EVM value shows more sensitive in Gain Imbalance than Phase Imbalance. In this figure RMS makes a satisfying performance in different mesh points of Gain Imbalance and Phase Imbalance. IQ imbalance and channel distortion are compensated almost completely.

#### 4.2.2. Blind Adaptive Compensator: CMA and modified CMA

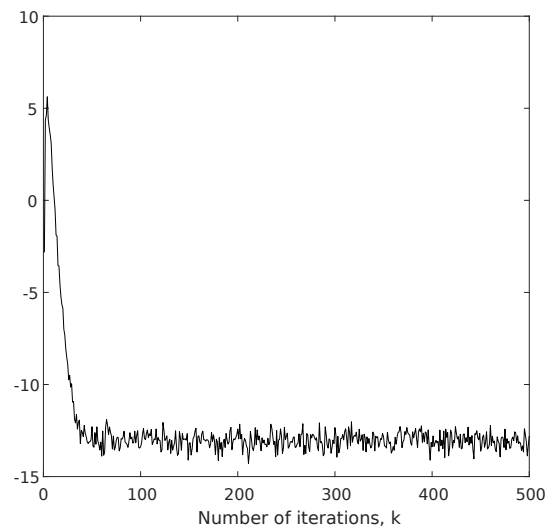
As described in Chapter 3, the CMA algorithm relies neither on the symbol decisions nor on a training sequence. This algorithm operates blindly based on the statistical property that the average power of received signal  $P(x)$  should be constant. Therefore, the CMA is not suitable for M-QAM if  $M > 4$ , since such modulations do not satisfy a constant power property any more. Savory [19] showed that the error function has a non zero optimum minimum. Since the error tends to be non zero, the stochastic gradient descent method works not as we hope and therefore degrades the performance.

Figure 4.7 (a) shows the constellations before and after compensator with CMA algorithm. The black points are received signal with IQ imbalance and channel distortion, while the blue points are the compensated signals. Figure 4.7 (b) shows the convergence line through MSE change by iteration. The CMA algorithm reaches convergence in about 2700 times iteration, while LMS algorithm requires only about 75 times iteration, since latter algorithms are supervised



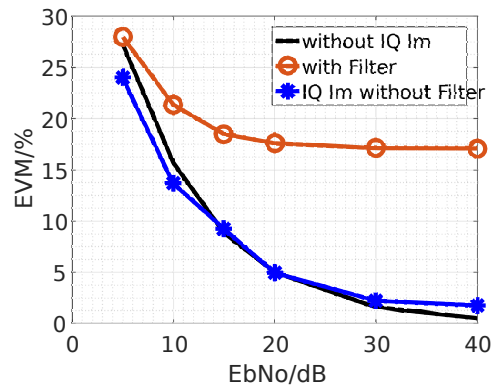


(a) Constellation Graph

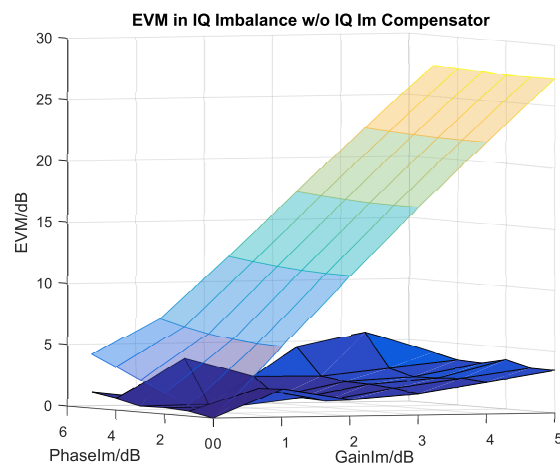


(b) Y Axis: MSE Learning / Convergence Line

**Figure 4.4.:** Equalizer with LMS Adaptive Filter for 4-QAM



**Figure 4.5.:** change legend, add RMS, use 64 QAM  $E_b/N_o$  w/o RMS Adaptive Filter for 4-QAM



**Figure 4.6.:** change PN unit to degree RMS Expression in different IQ Imbalance

method masters much more specific priori information.

Another disadvantage for CMA algorithm is that the constant power information contains only amplitude restriction and lacks phase restriction. Therefore a phase shift is introduced as Figure 4.7(a) shows.

This phase shift can be removed by a improved algorithm called Modified CMA [20]. As showed in Figure 4.8, while the traditional CMA converges independently of carrier recovery and reaches convergence point with a phase error, modified CMA solves this problem. Further more, since in modified CMA the real part and imaginary part are computed separately, the computational difficulty is actually decreased. Comparing the demand of around 2500 iterations before convergence in Constant Modulus Algorithm (CMA), modified CMA convergences at the fifth hundred iteration as Figure 4.8 (b) shows. Figure 4.8 expresses the change of different weights in  $\mathbf{w}$  in each iteration. The final weight vector  $\mathbf{w} = [-0.0070 - 0.0092i, 1.2703 - 0.5877i, 0.1753 + 0.3159i, 0.1293 - 0.0460i]^T$ . Comparing to the channel distortion  $\mathbf{H} = [1.1 + j*0.5, 0.1 - j*0.3, -0.2 - j*0.1]$ , the weight vector  $\mathbf{w}$  is closing to  $\mathbf{H}$  in someway and the rest weight  $(-0.0070 - 0.0092i)$  convergences to zero.

#### 4.2.3. Blind Adaptive Compensator: Circularity-Based Approach

As mentioned in Chapter 3, Lauri Anttila proposed in 2017 a new compensation algorithm against IQ imbalance called *Circularity-based Approach* [9]. Based on widely linear processing, this blind filter presents impressive performance and compensate IQ imbalance effective.

Figure 4.9 shows the compensation result with 5dB gain imbalance and  $5^\circ$ . The convergence as showed in Figure 4.9 (c) is  $\mathbf{w} = -0.2763 + 0.0462i$ . This makes sense because with chosen IQ imbalance of 5 dB and  $5^\circ$ , the original points  $\mathbf{x}$  is actually scaled and rotated by matrix

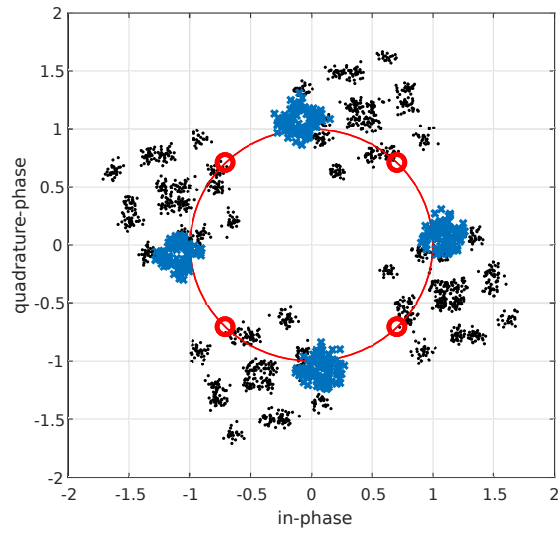
$$M_{\text{IQImbalance}} = \begin{bmatrix} 10^{0.5 \cdot \frac{I_a}{20}} \cdot \exp(-i0.5I_p \frac{\pi}{180}) & 0 \\ 0 & 10^{-0.5 \cdot \frac{I_a}{20}} \cdot \exp(i0.5I_p \frac{\pi}{180}) \end{bmatrix} = \begin{bmatrix} 1.2995 & 0 \\ 0 & 0.6910i \end{bmatrix}, \quad (4.2)$$

where  $I_a = 5$  dB represents amplitude imbalance and  $I_p = 5^\circ$  represents phase imbalance.

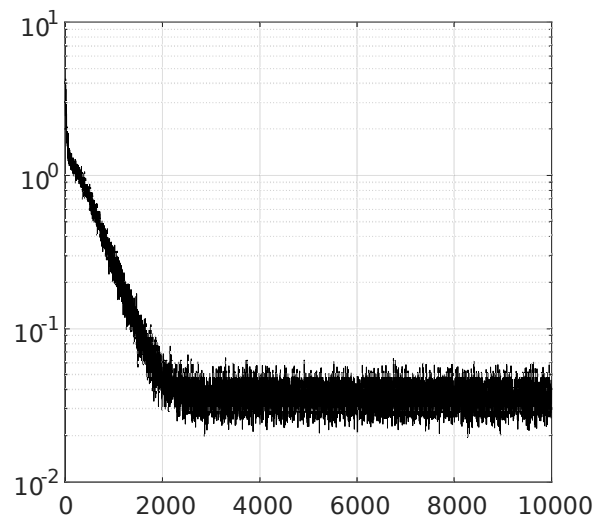
The circularity-based approach adaptive filter generate a weight matrix  $M_{\text{circularity}} = \begin{bmatrix} -0.2763 & 0 \\ 0 & 0.0462i \end{bmatrix}$  to ensure that

$$\left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + M_{\text{circularity}} \right) \cdot M_{\text{IQImbalance}} = \begin{bmatrix} 0.9391 & 0 \\ 0 & 0.0321 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (4.3)$$

which means the points are recovered as they were.

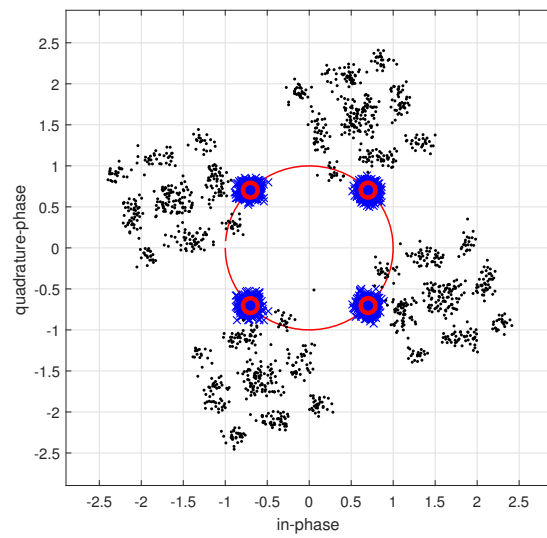


(a) Constellation Graph

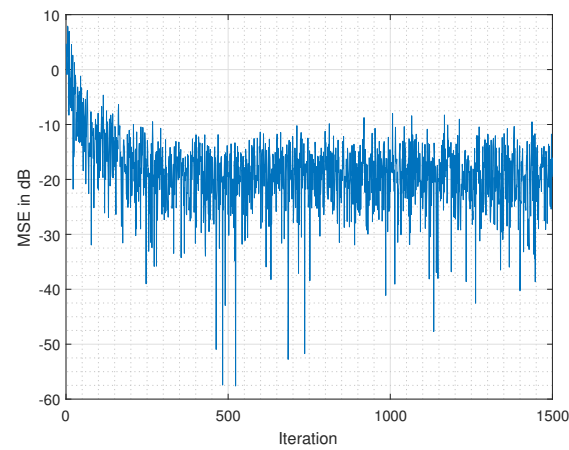


(b) Learning / Convergence Line

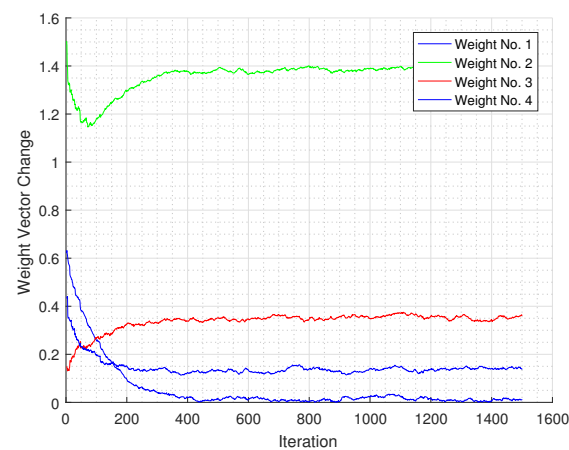
**Figure 4.7.:** Equalizer with CMA Adaptive Filter for 4-QAM



(a) Constellation Graph

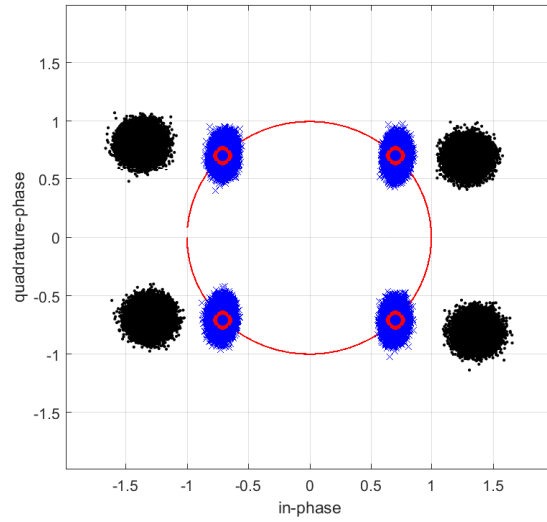


(b) Learning / Convergence Line

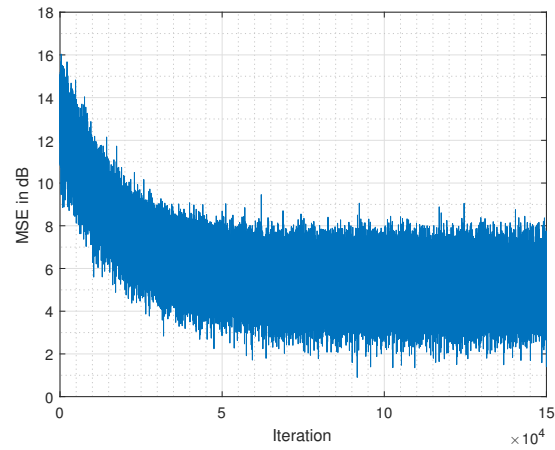


(c) Weight Vector changes during iterations

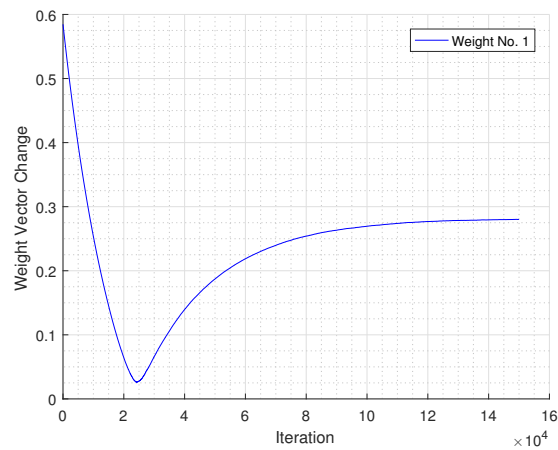
**Figure 4.8.:** Equalizer with Modified CMA Adaptive Filter for 4-QAM



(a) Constellation Graph



(b) Learning / Convergence Line



(c) Convergence Process of Weight

**Figure 4.9.:** Equalizer with Circularity-Based Approach Filter for 4-QAM

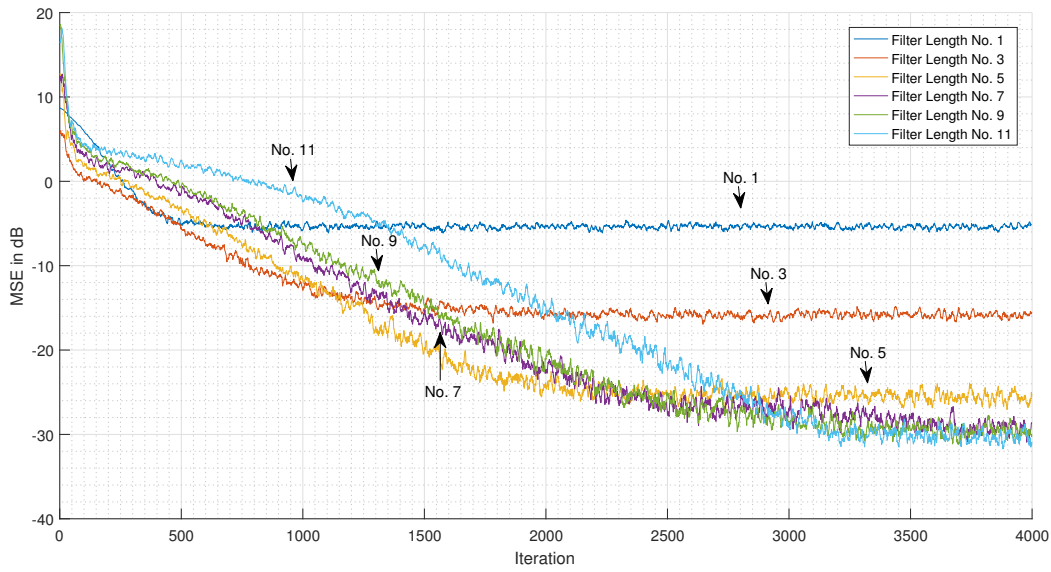


Figure 4.10.: MSE of CMA filter in different filter length  $N$

### 4.3. Impact of the Hyperparameters

Hyperparameters implies the setting parameters before adaptive filters start. Basic hyperparameters in filters contains normally equalizer length  $N$ , iteration step length  $\mu$  and iteration times. By iteration times we choose a large number to ensure reaching the convergence point. However, chooses of equalizer length and iteration step length are following experience and experiments.

The influence of filter length  $N$  is showed in Figure ???. MSE moves are processed by moving average in the consideration of clearness.

On the one hand, dealing with channel distortion  $\mathbf{H} = [1.1 + j*0.5, 0.1 - j*0.3, -0.2 - j*0.1]$ , since channel contains three taps, the LMS filter with  $N > 3$  performs obviously better than  $N \leq 3$ . MSE is reduced to about  $-28$  dB by 5 ordered LMS filter, while by lower length MSE reaches only to  $-5$  dB ( $N = 1$ ) and to  $-16$  dB.

On the other hand, Figure 4.10 also shows filter with lower length has higher convergence speed. LMS filter with  $N = 5$  convergences at second thousand times iteration, while the one with  $N = 11$  reaches convergence point at third thousand times iteration, although its MSE is controlled 4 dB better (from  $-26$  dB to  $-30$  dB). Therefore a trade-off of filter length  $N$  should be noticed.

Another hyperparameter iteration step length  $\mu$  controls the step value of each movement. Large step length will increase the convergence speed. As Figure 4.11a and 4.11b show, by step length  $\mu = 0.001$  it takes 800 iterations for the convergence, while under 100 iterations are demanded by step length  $\mu = 0.015$ .

However, some disadvantages are introduced by large step length at the same time. Large step length causes not only a rough convergence with a always changing weight vector  $\mathbf{w}$ , but the filter becomes also unstable and easy divergences. When the filter divergences, weight vector  $\mathbf{w}$  becomes infinite large and totally unusable. Figure 4.12b show this situation. Although the weight enter the desired area, comparing Figure 4.12a it walks around the optimal point in a large circus. Optimal weight vector can not be precise estimated because of large step chose.

To the contrary, 4.11c and 4.12c shows what happened if step length  $\mu$  is too small. After 1500 iterations the weight vector  $\mathbf{w}$  still doesn't reach the convergence point. This will waste compute resource and increase computational cost.

#### 4.4. Computational Load and Impact of time-varying channels

The study of tracking ability is also important for proposed equalizers. Compare to fixed filters, adaptive filters are able to adjust weight parameters and adapt time varying channel, if the impairments change relatively gently.

For the sake of simplicity, only infinite polarization rotation by Jones matrix is considered. The channel impulse response at time  $t_0$  can be written as

$$\mathbf{C}_{t_0}(t) = \begin{bmatrix} \cos(\Omega t_0) & \sin(\Omega t_0) \\ -\sin(\Omega t_0) & \cos(\Omega t_0) \end{bmatrix} \delta(t), \quad (4.4)$$

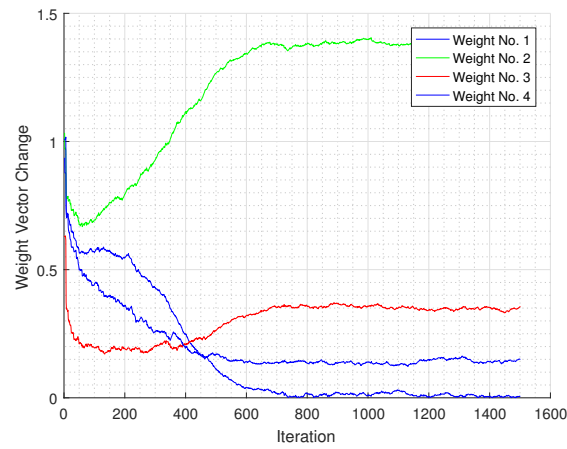
where  $\Omega$  is the rotation speed in  $\text{rad s}^{-1}$ .

In consideration of limited computational power in reality, the complexity of various algorithms is also a critical part for implementation of real time communication. The compute time cost is showed in Table 4.2. All adaptive filters are trained (for supervised filters) with 300 symbols and tested with  $10^5$  symbols. The convergence process takes time, especially for circularity-based approach because of its huge iteration demand. After convergence things becomes easier. With consideration of big test set, a processing time about 1 ms is not expensive. Of course all compute works are done in general PC. As to practice such a compute ability might be out of possession. Some more practical measurement should be proposed.

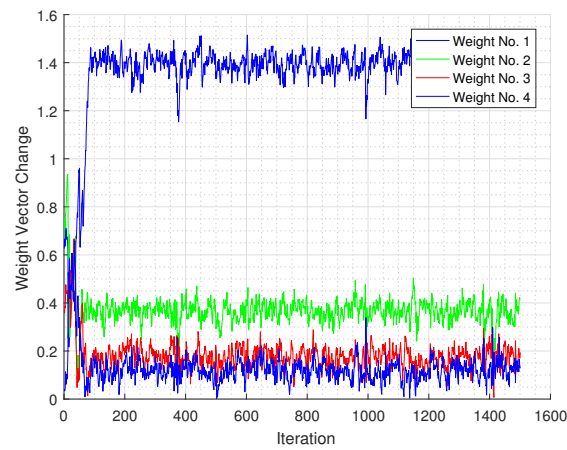
#### 4.5. KNN and SVM Detectors

As mentioned in Chapter 3, SVM is a data-aided detector. The modulation is 16 QAM and received signal as Figure 4.13 shows is distorted by phase noise with frequency offset  $f_{\text{offset}} = -40\text{Hz}$  and phase noise level  $L = 20\text{dBc/Hz}$ . The number of training symbols is kept as 100. Figure 4.14

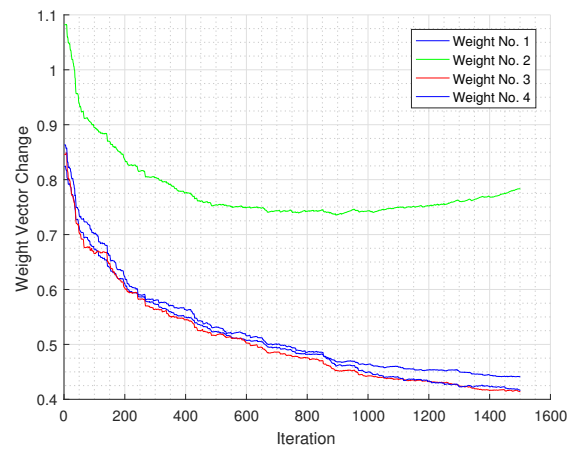




(a) LMS Filter with  $\mu = 0.001$  and  $N = 11$

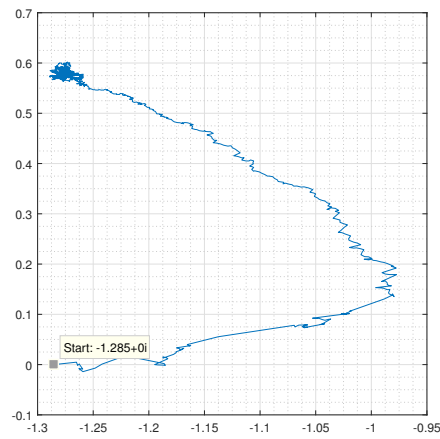


(b) LMS Filter with  $\mu = 0.015$  (too large) and  $N = 11$

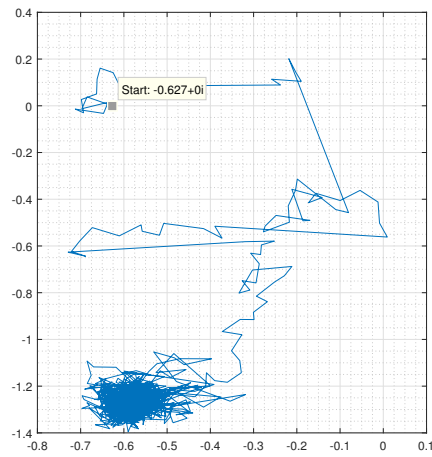


(c) LMS Filter with  $\mu = 0.0001$  (too small) and  $N = 11$

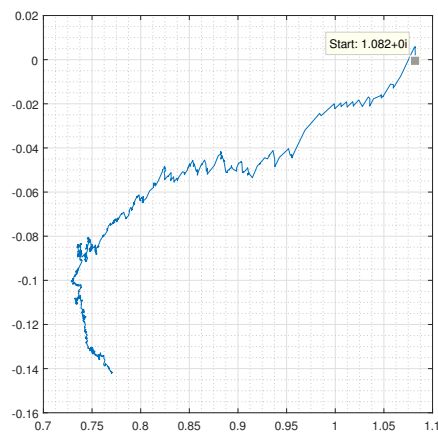
Figure 4.11.: Convergence result for weight vector



(a) LMS Filter with  $\mu = 0.001$  and  $N = 11$



(b) LMS Filter with  $\mu = 0.015$  (too large) and  $N = 11$

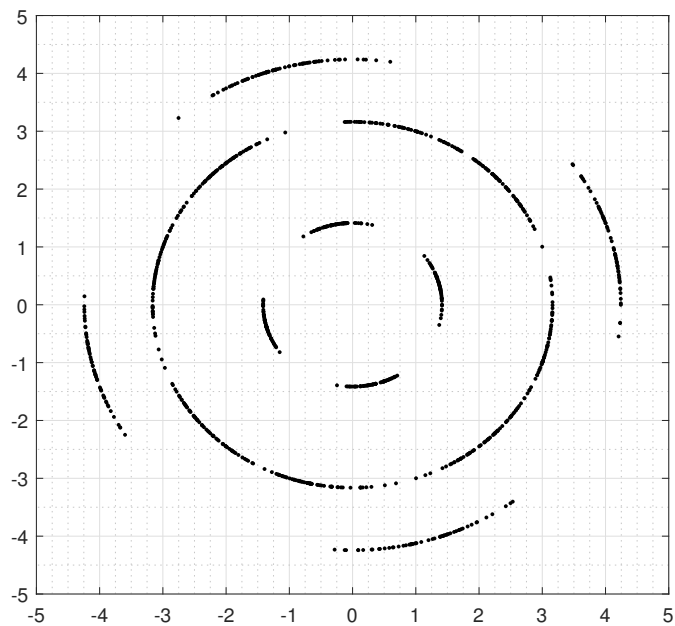


(c) LMS Filter with  $\mu = 0.0001$  (too small) and  $N = 11$

**Figure 4.12.:** Convergence process for one of weight

	LMS	RMS	CMA	Modified CMA	Circularity-based Approach
Iteration demand	100	70	3000	500	$10^5$
Each iteration ( $\mu$ s)	8	9	8	8	8
Initial Time (ms)	0.8	0.63	24	4	800
Use Time for $10^5$ symbols (ms)	1.013	1.012	0.997	1.005	0.472
Computational Complexity	$2N + 1$	$4N^2$	$2(4N + 1)$		

**Table 4.2.:** Time Cost for Different Adaptive Filters (with optimal hyperparameters), 4-QAM

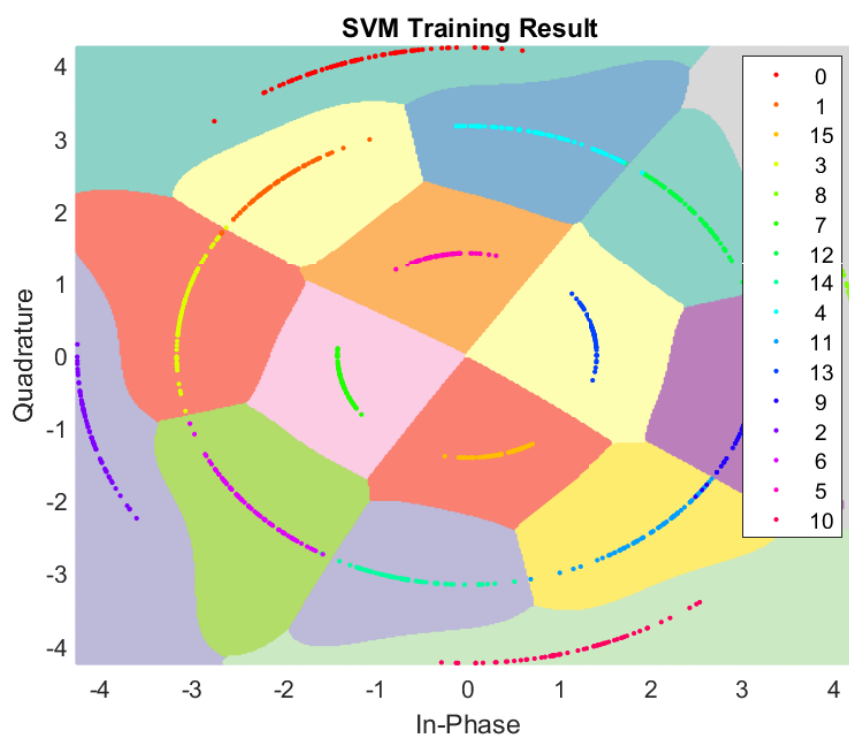


**Figure 4.13.:** RX Constellation Points with Phase Noise Impairment

shows classified areas with various labels according to training symbols. Different colored block represents different decision. It takes 0.787 s to train SVM. Once training finished, SVM can detect the symbol immediately.

4.6. Simulation of the Whole System

=====**Following is from Laura**=====



**Figure 4.14.:** Detection of SVM against Phase Noise Impairment

## 5. Conclusion and Perspectives

Newer and more demanding wireless applications are pushing the telecommunication technology forward. Increasing speed need ask for high signal quality, which though is hard to satisfy because of crowded radio spectrum and limited spectrum resource. Besides hardware design and coding technology (check bit or self-correcting code), some post-end technology makes also contribution.

The work carried out in the thesis deals with both deterministic impairments (channel distortion, IQ imbalance) and stochastic impairment (phase noise). Channel distortion merge old symbols and new symbol by deterministic weights. In addition , fixed or slow changing scaling and phase shifting are also convoluted in transmitted signal through IQ imbalance. Deterministic impairments are relative easy to equalize, since once we get weight vector, compensation can be done with satisfied result. By contrast, stochastic impairment phase noise is more difficult to eliminate. Even if the parameters is known, we can still not equalize each point precisely.

Futhermore, we introduced several algorithms and methods for improving the whole system. Higher and more powerful chips are also making potential algorithms more possible than any other moment in history. For deterministic impairments, we use basically gradient-descent algorithm through constant step-size version to approach weight vector. Different cost functions bring huge difference in convergence speed, final MSE and computational complexity. We discuss in detail the chose and its effect of step size  $\mu$ , filter length  $L$  and iteration number  $N$ .

For stochastic impairment phase noise we try a new method SVM. Pilot method is used for phase noise compensation [21]. In this article we mainly realize SVM as recognizer [14]. After trained by initial symbols, SVM separate the constellation area properly. New received symbols are demodulated to different labels according to this separation.

Finally,



## Bibliography

- [1] "Cisco visual networking index: Global mobile data traffic forecast update, 2016–2021 white paper."
- [2] "IMT vision – framework and overall objectives of the future development of IMT for 2020 and beyond," p. 21.
- [3] A. Grami, *Introduction to Digital Communications*. Academic Press. Google-Books-ID: 6nww-BAAAQBAJ.
- [4] "The RF and microwave handbook."
- [5] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Springer US.
- [6] K. M. Gharaibeh, *Nonlinear Distortion in Wireless Systems: Modeling and Simulation with MATLAB*. John Wiley & Sons. Google-Books-ID: EMI2xW9gqLQC.
- [7] S. Sharma, J. Nagar, and P. Singh, "Modeling and performance analysis of wireless channel," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 3052–3057.
- [8] D. B. Leeson, "A simple model of feedback oscillator noise spectrum," vol. 54, no. 2, pp. 329–330.
- [9] L. Anttila, M. Valkama, and M. Renfors, "Blind compensation of frequency-selective i/q imbalances in quadrature radio receivers: Circularity -based approach," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, vol. 3, pp. III–245–III–248.
- [10] D. GODARD, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," vol. 28, pp. 1867–1875.
- [11] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: a review," vol. 86, no. 10, pp. 1927–1950.
- [12] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Springer US, 4 ed.
- [13] D. Wang, M. Zhang, M. Fu, Z. Cai, Z. Li, H. Han, Y. Cui, and B. Luo, "Nonlinearity mitigation using a machine learning detector based on  $k$ -nearest neighbors," vol. 28, no. 19, pp. 2102–2105.
- [14] D. Wang, M. Zhang, Z. Cai, Y. Cui, Z. Li, H. Han, M. Fu, and B. Luo, "Combatting nonlinear phase noise in coherent optical systems with an optimized decision processor based on machine learning," vol. 369, pp. 199–208.
- [15] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," vol. 2, pp. 45–66.

- [16] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley-Interscience, 2. ed.
- [17] V. Vapnik, *The Nature of Statistical Learning Theory*. Information Science and Statistics, Springer-Verlag, 2 ed.
- [18] "Advanced design system (ADS) | keysight (formerly agilent's electronic measurement)."
- [19] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," vol. 16, no. 5, pp. 1164–1179.
- [20] K. N. Oh and Y. O. Chin, "Modified constant modulus algorithm: blind equalization and carrier phase recovery algorithm," in *ICC '95 Seattle, 'Gateway to Globalization', 1995 IEEE International Conference on Communications, 1995*, vol. 1, pp. 498–502 vol.1.
- [21] S. Randel, S. Adhikari, and S. L. Jansen, "Analysis of RF-pilot-based phase noise compensation for coherent optical OFDM systems," vol. 22, no. 17, pp. 1288–1290.



# Appendices



## A. Acknowledgments

From Laura

I would like to gratefully acknowledge the people who supported me and who have contributed directly or indirectly, to my work and my personal development.

Firstly, I would like to thank Prof. Dr.-Ing. Ingmar Kallfass, for giving me the opportunity to work as a HiWi and to write my Master Thesis in his research group and take advantage from a wealth of knowledge and experience. I am also grateful to my thesis adviser M.Sc. Christopher Grötsch for sharing valuable know-how, immense support during the lengthy way of achieving the desired results and teaching me the first steps in millimeter waves circuit design.

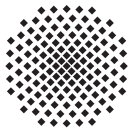
From Institute of Robust Power Semiconductor Systems (*Institut für Robuste Leistungshalbleitersysteme*, ILH) also, a special thank goes to M.Sc. Seyyid M. Dilek, for his friendship, guidance and providing me with powerful references. In particular, I would also like to thank to M.Sc. Iulia Dan for her useful and honest advices, open discussions, encouragements and also for being a source of inspiration. I am also thankful to M.Sc. Parisa Harati and M.Sc. Eswara Rao Bammidi for sharing meaningful comments and suggestions. I am also appreciative to M.Sc. Benjamin Schoch.

The passion for microwave engineering would have never been so deep without Prof. Dr.-Ing. Iulia Mocanu, from "Politehnica" University of Bucharest, who carefully shaped this career path for me. Also, a tight friendship was developed with Prof. Dr.-Ing. Ruxandra Costea, who sincerely advised me in critical decisions and who always encouraged and supported me in my entire university life.

I especially appreciate all my colleagues, tutors and professors who made this master programme an unforgettable experience. I am particularly thankful to my dear friends, Slađan, Mariangela, Jón, Baho, Eliza, Ionuț, Ahmed and Daniel with whom I have understood the value of a true friendship during difficult times. I am grateful to my colleague, Xiaohui Nie, who provided me with useful Latex tips and shortcuts during the thesis writing time. You all should know that I truly enjoyed the open and inspiring atmosphere you all contributed to.

I am thankful to Deutscher Akademischer Austausch Dienst (DAAD) for financing my studies during the entire master programme and giving me the opportunity to focus on research without any concern. In particular, I am thankful to Mrs. Gabrijela Nikolic.

An enormous thank goes to my parents and my family for their unconditional support and for their inexhaustible trust.



**University of Stuttgart**  
Germany



Institute of  
Robust Power  
Semiconductor Systems

**Yang Liu**

**Student Research Project: Mmwave Transceiver System Performance Analysis using Post-Processing Analysis**

**Period: 26.10.2017 – 31.07.2018**

**Supervisor: Prof. Dr.-Ing. Ingmar Kallfass , M. Sc. Seyyid Muhammed Dilek**

**Executive Abstract: [Abstract](#)**