Adaptive Semiblind Channel Estimation for OFDM/OQAM Systems

Tianze Su

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

September 2015

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

2015/09/16

Abstract

In this thesis, we propose and investigate novel adaptive semi-blind channel estimation algorithms for OFDM/OQAM systems. OFDM/OQAM is regarded as a promising alternative to conventional CP-OFDM for multi-carrier modulation since it can provide better spectrum efficiency, albeit at the price of increased complexity.

We first formulate a general system model of an OFDM/OQAM transceiver. Based on this model, we review a recently proposed block based semi-blind channel estimation method for OFDM/OQAM systems, known as the sign covariance matrix (SCM) method. This method mainly exploits the higher-order statistical properties of the data at the receiver side but is not well-suited for applications to time-varying channels. Subsequently, to overcome the drawbacks of this block-based technique, we propose adaptive semi-blind channel estimation algorithms for application to OFDM/OQAM.

The proposed algorithms consist of an adaptive SCM technique obtained through exponential recursive averaging, as well as several constant modulus algorithms (CMA) for recursive estimation. Although all the adaptive algorithms are designed to deal with time-varying channels, they can also be used for rapid channel acquisition in the case of static or slowly-varying channels. Furthermore, we explore the coherence bandwidth of the channel and make use of this concept to improve the estimation accuracy via a frequency averaging technique that can be combined with the adaptive SCM.

Simulation results validate the efficacy of the proposed adaptive estimation algorithms over both time-invariant and varying channels, showing their robustness in terms of convergence speed, tracking capability and residual estimation error in steady-state. In particular, the CMA with recursive least squares (CMA-RLS) updating proves to be the most preferable due to its excellent trade-off between convergence rate and residual error level. The CMA-RLS also offers the best performance in tracking a time-varying channel. In addition, simulation experiments demonstrate the effectiveness of combining the frequency averaging technique with the proposed adaptive SCM algorithm.

Sommaire

Dans ce mémoire, nous introduisons de nouveaux algorithmes d'estimation de canal semi-aveugles pour les systèmes OFDM/OQAM. La modulation à porteuses muliples OFDM/OQAM pourrait venir à remplacer la modulation plus conventionnelle de type CP-OFDM car elle offre une meilleure efficacité spectrale, malgré sa plus grande complexité.

Nous introduisons d'abord le modèle mathématique général pour un système de transmission et réception basé sur la modulation OFDM/OQAM. Nous examinons ensuite un récent algorithme d'estimation semi-aveugle, nommé méthode de matrice de covariance à signes (MCS). Cet algorithme exploite les statistiques d'ordre supérieur des données, mais n'est pas approprié pour l'estimation des canaux variables dans le temps puisqu'il repose sur l'estimation par bloc. Pour résoudre ce problème, nous proposons de nouveaux algorithmes adaptatifs semi-aveugles pour l'estimation de canal dans les système OFDM/OQAM.

Ces algorithmes incluent une réalisation adaptative de l'algorithme MCS obtenue via un moyennage exponentiel récursif, de même que plusieurs algorithmes à module constant (AMC) permettant d'effectuer l'estimation récursive. En plus de permettre le suivi des canaux variants dans le temps, comme la plupart des méthodes adaptatives, nous pouvons utiliser ces nouveaux algorithmes afin d'effectuer l'acquisition rapide de canaux statiques. Pour améliorer la performance d'estimation, nous explorons une technique de moyennage fréquentiel sur la bande de cohérence du canal, technique qui peut tre combinée avec adaptative de l'algorithme MCS.

L'efficacité des algorithmes proposés est vérifiée via des simulations sur des canaux variants et invariants dans le temps. Ces simulations démontrent la robustesse des nouveaux algorithmes en termes de vitesse de convergence, de vitesse de suivi, et d'erreur résiduelle en régime permanent. En particulier, l'AMC avec mise à jour basée sur l'approche des moindres carrés récursive (AMC-MCR) est la méthode préférentielle, grâce à un compromis optimal entre la vitesse de convergence et l'erreur résiduelle. La AMC-MCR offre aussi la meilleure performance en termes de suivi d'un canal variant dans le temps. De plus, les simulations démontrent un accroissement d'efficacité lorsque la technique de moyennage fréquentiel est couplée à l'algorithme adaptative MCS proposés.

Acknowledgements

I have been studying in McGill for two years. During this period, I have learnt both academic knowledge and a positive attitude towards life and work. With my supervisors, parents and friends' help, my master's thesis work has been completed. Here, I give my appreciation to all of you.

Firstly, I would like to thank my supervisors, Prof. Benoit Champagne and Prof. Xiaowen Chang, for their numerous help, support and guidance on both academic and personnel level over the past two years. I am deeply grateful for their advices and help. I can never forget the days we exchanged emails and discussed academic problems. The rigorous attitude and enthusiasm on research I learned from them is an invaluable fortune of my life. Working with them is an unforgettable and precious experience for me. All I can say is that I can not thank them enough.

Secondly, I gratefully thank Dr. Weikun Hou, from Huawei Corporation (Ottawa), for his constructive comments and suggestions. His previous work helped me build up solid foundation for my research. On certain occasions when I met with difficulties in basic technical issues, he was patient enough to explain them to me as clearly as possible.

Also, I would like to thank my parents who gave me endless care. It would be much harder for me to continue my research work without their selfless love. Over the past few decades, they have made their best efforts to help me grow up as an optimistic and considerate girl. It is them who made me who I am today.

Finally, I would like to express my sincere thanks to my friends and lab mates, who were companying me all the time. I appreciate all the inspiring discussions we made in the lab. The joyful time we shared together will be kept in my mind forever.

Contents

List of Figures

List of Tables

List of Acronyms

AFB	Analysis Filter Bank
AR	Auto-Regressive
BER	Bit Error Rate
CFR	Channel Frequency Response
CIR	Channel Impulse Response
C2R	Complex to Real
CM	Constant Modulus
CMA	Constant Modulus Algorithm
CP	Cyclic Prefix
DD	Decision Directed
DFT	Discrete Fourier Transform
EPA	Extended Pedestrian A
EVD	Eigenvalue Decomposition
E/D	Equalization and Detection
FFT	Fast Fourier Transform
FM	Frequency Modulation
GN	Gauss-Newton
ICI	Inter Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
ISI	Inter Symbol Interference
LMS	Least Mean Square
MCM	Multicarrier Modulation
MMA	Multi-modulus Algorithm

NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency Division Multiplex
OQAM	Offset Quadrature Amplitude Modulation
PM	Phase Modulation
P/S	Parallel to Serial
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RLS	Recursive Least Square
RRC	Root Raised Cosine
R2C	Real to Complex
SCM	Sign Covariance Matrix
SFB	Synthesis Filter Bank
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
S/P	Serial to Parallel

Chapter 1

Introduction

In this chapter, first we provide an overview of the technical problem to be studied. Then the existing literature aimed at solving this problem is surveyed. Next, we summarize the contributions made by this thesis. Finally, the thesis organization along with certain mathematical notations are explained.

1.1 Problem Overview

Nowadays, the increasing demand for high data rate transmissions continues to call for the search of improved signal processing techniques to enable wideband wireless communications with improved quality of service (QoS). Consideration of fundamental issues, including frequency selective fading due to multipath propagation, demodulation complexity and flexible multiple access, naturally leads to the use of multicarrier modulation (MCM) techniques [1]. In the MCM scheme, the available channel bandwidth is subdivided into several parallel subchannels over which different data streams (from one or multiple users) are multiplexed for parallel transmission. One of the classical MCM techniques, currently used in many standards [2], is orthogonal frequency division multiplex (OFDM) with a cyclic prefix (CP). Using the CP as a guard interval, OFDM manages to turn a frequency selective channel into a set of nearly flat parallel subchannels with independent noises. This, in turn, greatly simplifies both the channel estimation and the data recovery. CP-OFDM succeeds in reducing inter-symbol interference (ISI) compared with single carrier modulation but at the same time, the extra CP entails a waste in transmitted power as well as in spectral efficiency. Furthermore, the advantages of CP-OFDM come at cost of

an increased sensitivity to carrier frequency offset and symbol timing errors.

Alternatively, OFDM based on offset quadrature amplitude modulation, known as OFDM/OQAM, has received much attention in recent years as it can mitigate these drawbacks [3]. Nevertheless, efficient channel estimation and equalization schemes are needed for the effective implementation of data detection on the receiver side of an OFDM/OQAM system.

In this thesis, we seek to develop and investigate new adaptive algorithms for the semi-blind estimation and tracking of wireless channels in OFDM/OQAM systems.

1.2 Literature Survey

Over the past few decades, MCM has been well accepted in both wired and wireless communications due to its capability to realize reliable high data rate transmission and to cope with frequency selective channels [4, 5]. As a MCM technique, OFDM converts a single high speed data stream into multiple low speed data streams, and modulates them onto different subcarriers, with the help of an inverse discrete Fourier transform (DFT) operation. To preserve the orthogonality of the subcarriers, a cyclic prefix (CP) is added as a guard interval. However, the CP leads to a loss of spectral efficiency as it contains redundant information, which is a main drawback of CP-OFDM [6]. Other disadvantages of CP-OFDM, such as the increased sensitivity to carrier frequency and timing errors, are mentioned in [7, 8].

As a special type of MCM, OFDM based on offset quadrature amplitude modulation (OQAM), referred to as OFDM/OQAM, has been considered as an alternative to CP-OFDM [9,10]. Compared to CP-OFDM that transmits complex-valued symbols at a given symbol rate, OFDM/OQAM transmits real-valued symbols at twice this rate [11]. In OFDM/OQAM, every subcarrier is modulated with a staggered offset QAM symbol. Instead of using the cyclic prefix, OFDM/OQAM system utilizes well designed prototype filters that, in theory, satisfy perfect reconstruction conditions [12]. In order to mitigate ISI and achieve distortion free transmission, the receiver still requires knowledge of channel information to finalize equalization; therefore, channel estimation is an indispensable component of an OFDM/OQAM receiver, needed prior to equalization and data detection. However, all the attractive features of OFDM/OQAM come at the price of a relaxation of the orthogonality conditions that only hold in the real field. Hence, the simplicity of

the channel estimation in CP-OFDM is lost in OFDM/OQAM systems [13]. Consequently, OFDM channel estimation methods may not be suitable for application to OFDM/OQAM systems. A specific problem of intrinsic interference among adjacent subcarriers and symbols has to be solved, which is a main issue in OFDM/OQAM channel estimation.

Earlier works on OFDM/OQAM investigate training-based techniques as solutions to channel estimation. One of them is preamble-based where a block of pilot symbols precede the data in each burst [14–16]. Specific structures of preamble have been proposed, thereby leading to assorted channel estimation methods. In [14], two different preamble-based methods are proposed, namely the interference approximation method (IAM) and pairs of real pilots (POP). Motivated by the IAM method and awareness of its suboptimal nature, a general theoretical framework for IAM preamble design is formulated in [15] and [16]. Beside the preamble-based technique, another training-based technique introduces scattered pilots among the data symbols throughout the burst [17–19]. Authors of [17] find an appropriate combination of the data in the vicinity of each scattered pilot to cancel the interference while an iterative channel estimation method is provided in [18]. In [19], the pilots are scattered in the shape of a quincunx over the frame for a better coverage of the time-frequency space. In general, training based techniques mainly aim at cancelling the undesired interference from neighbours and improve estimation performance accordingly. However, they are not necessarily desirable in real life due to the high expense of sending pilots, which amounts to an inefficient use of precious bandwidth resources.

Based on considerations above, a semi-blind channel estimation technique is proposed to handle the fading channel [20], where statistical properties present in received signal components are explored. In particular, the real property of the transmitted symbol is fully exploited to blindly identify the channel-induced rotations. An essential part of this technique is the estimation and use of the sign covariance matrix (SCM), which leads to channel phase estimation over each subcarrier. Both channel phase (through the SCM) and amplitude are approximated by block averaging over multiple symbol times. Only one symbol per tone per block is required to resolve a trivial sign ambiguity, thereby reducing the pilot overhead for estimation purpose, as compared to previous training-based methods.

Nevertheless, block processing as in [20] is plagued by a number of limitations. Instead of being distributed over time, computational resources tend to be used unevenly, where most the stringent part of the estimator computations needs to be carried out rapidly after the data accumulation within a block. Additional memory is needed to buffer the received

data during the channel estimation phase, after which equalization and data detection for the given block can be applied. Block processing relies on the assumption that the underlying channel under estimation remains nearly constant during a block. Therefore, it is mostly useful for the case of time-invariant or slowly fading channel, but is not well suited to the estimation and tracking of fast time-varying channels, as often encountered in wireless environments with mobile users or scatterers.

1.3 Thesis Contributions

Traditionally, channel estimation techniques for OFDM/OQAM may be classified into two categories, namely: training-based methods and semi-blind or blind methods. Training-based methods send training sequences or pilots at regular intervals, thereby consuming bandwidth and decreasing spectrum efficiency. Blind channel estimation methods, on the other hand, appear attractive since they avoid the use of training symbols. However, existing blind methods use second-order and higher-order cyclostationary statistics to yield the unique channel estimation up to an ambiguity factor, as represented by one or more unknown scalars. In practice, this ambiguity could be resolved by using a short training sequence; hence these methods are often referred to as semi-blind. However, second-order or higher-order statistics based methods need a relatively long data-record length for accurate channel estimation. For example, the semi-blind method in [20] is based on data block, which, for reasons explained above leads to uneven use of computational resources and additional memory requirements. More importantly, it becomes impractical under rapidly time-varying channel conditions. The goal of this thesis is to propose and study new semi-blind adaptive algorithms to cope with the estimation of both time-invariant and time-varying channel in OFDM/OQAM systems.

Unlike the block-processing approach in [20], we propose four different adaptive algorithms to handle the OFDM/OQAM channel estimation. These algorithms can be classified into two categories. The first one, referred to as Adaptive SCM, is based on the existing SCM of [20], but uses exponential smoothing as opposed to block processing in the estimation of certain required quantities. The second category, referred to as constant modulus algorithms (CMA), is based on the constant modulus formulation for recursive channel estimation, which exploits certain symmetry in the digital data modulation. A variety of CMA can be devised depending on the cost function and search direction strategy. In this

thesis, three CMA are considered, as listed below:

- Gradient-based least mean square (CMA-LMS-Grad)
- Gauss-Newton based least mean square (CMA-LMS-GN)
- Recursive least squares (CMA-RLS)

These adaptive algorithms store the most recent channel estimate only and update it recursively by processing the current observations. The main advantages of adaptive algorithms not only lie in faster estimation of time-invariant (stationary) channels, but also in effective acquisition and tracking of time-varying (non-stationary) channels. Only a few pilots per tone are needed at an early stage of adaptation to resolve the sign ambiguity present in the initial estimate. Simulation results confirm the efficacy of the proposed adaptive algorithms for both time-invariant and time-varying channel conditions. Among the newly proposed adaptive algorithms, the CMA-RLS proves to have the best performance in terms of initial convergence speed, steady-state error and tracking capability. In addition, to further improve the estimation accuracy, a frequency averaging technique that takes advantage of the coherence bandwidth is appended as a modification to the Adaptive SCM algorithm.

1.4 Thesis Organization

The rest of this thesis is organized as follows. Chapter 2 reviews basic concepts of multicarrier modulation, gives a mathematical description of the OFDM/OQAM system model and then explains block based channel estimation technique in [20] thoroughly. On the basis of the system model derived in Chapter 2, the four adaptive semi-blind channel estimation algorithms are developed in Chapter 3 by recursively minimizing a specific objective function. In this chapter, we also introduce the so-called multi-modulus extensions of the proposed algorithms, as well the combination with a simple frequency averaging technique, which in theory gives rise to enhance estimation. Simulation results are presented and discussed in Chapter 4 to demonstrate the performance of the proposed adaptive channel estimation algorithms when applied to OFDM/OQAM systems. Finally, Chapter 5 draws conclusions of the whole thesis and indicates possible research directions for the future.

Notation. Vectors and matrices are denoted by bold lower-case and upper-case letters, respectively. Superscript ^{T} represents transposition. $\Re, \Im$ specify the real and imaginary parts of a complex quantity. The expectation operator is denoted by $E(.)$. The complex conjugate of a complex number z is denoted by z^*. $\|\cdot\|$ is used for the 2-norm of a vector.

Chapter 2

Block based Channel Estimation for OFDM/OQAM Systems

In this chapter, we first review the fundamentals of multicarrier modulation, including OFDM and OFDM/OQAM schemes. Then a recently proposed semi-blind channel estimation method, referred to as block based SCM algorithm, is reviewed for application to OFDM/OQAM. Finally, the limitations of this block based SCM algorithm are discussed briefly.

2.1 Overview of Multicarrier Modulation

2.1.1 General MCM

In the past decades, there has been much research and development work aiming at increasing data rates of wireless transmissions while improving quality of services. In turn, increased transmission rates have led to new and ever more demanding applications, spurring the need for further improvements in data rates, and it seems that this trend will continue [21].

To boost high data rates in broadband wireless transmissions, multicarrier modulation (MCM) systems have been extensively studied and used in many applications [4]. The basic idea behind MCM transmission schemes is to use multiple subcarriers to transmit data in parallel across a wideband channel. Equivalently, the wideband frequency selective channel is divided into several parallel subchannels with flat or mildly selective fading, facilitating

channel estimation and equalization. As a result, in a MCM system with the same data rate as a single carrier system, the symbol duration is longer, which translates into reduced inter symbol interference (ISI) and improved system performance with respect to probability of error. In addition, the division of the whole bandwidth into many subchannels also provides scalability and flexibility when configuring the communication link [5].

Fig. 2.1 illustrates the parallel transmission scheme taking place at the transmitter of a generic MCM system. A serial bit stream with rate R, in bit per second (bps) is firstly converted to M parallel low speed data streams with rate R/M bps. Each sub-stream then goes through the following chain of operations:

- Symbol Mapper: where groups of bits are mapped to M-ary symbols within a given constellation;
- Pulse shaping filter: which is used to convert the M-ary symbols into analog pulse code modulation waveforms.
- Modulation: where each baseband substream is shifted to the desired frequency subband by multiplication with $\cos(2\pi f_m t)$, where f_m is the corresponding carrier frequency.

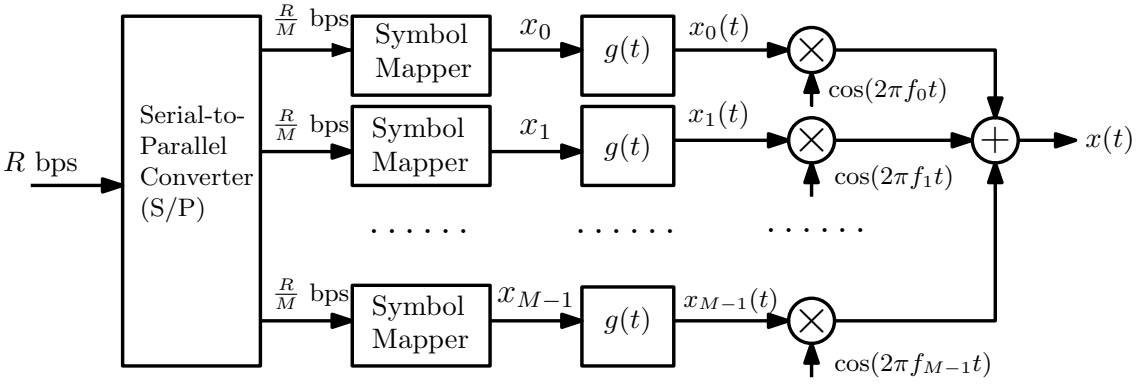


Fig. 2.1 Transmitter schematic diagram in general MCM

Following this sequence of operations, the various subband signals are added for transmission over the wideband channel. At the receiver (not shown), a corresponding sequence of operation takes place to recover the transmitted data stream from the received analog front-end signal.

2.1.2 OFDM

Orthogonal frequency division multiplexing (OFDM) is the flagship of the MCM techniques. It splits a high-rate data stream into a number of lower rate streams that are transmitted simultaneously over a number of subcarriers by exploiting properties of the discrete Fourier transform (DFT) [22]. These subcarriers are uniformly spaced in frequency and are orthogonal to each other. By choosing an appropriate number of subcarriers for transmission, the bandwidth of these parallel subchannels is smaller than the corresponding coherence bandwidth of the entire frequency selective channel. Therefore, each subchannel experiences flat fading, and low-cost (i.e. single-tap) channel estimation and equalization techniques can be applied to each of these subchannels for the purpose of data detection. To prevent the subcarriers from losing their orthogonality, however, a cyclic prefix (CP) is needed in OFDM modulation. In effect, a properly selected CP elegantly turns severely frequency selective channels into flat fading ones at the subcarrier level.

Currently, OFDM is widely applied in several standards, for example, the WiFi standard [23], which is based on the various iterations of IEEE 802.11, uses OFDM for its highest throughput profiles. Moreover, the Long Term Evolution (LTE) standard [24] uses OFDM for the downlink - that is, from the base station to the terminal to transmit the data over many narrow subbands of 180KHz each, instead of spreading one signal over the complete 5MHz bandwidth.

Fig. 2.2 shows a simplified flowchart of the CP-OFDM system. At the transmitter side (Tx), by using serial-to-parallel (S/P) converter, the transmitted bit stream is split into multiple substreams of M-ary symbols, typically taken from a quadrature amplitude modulation (QAM) constellation. After the S/P conversion, the parallel substreams of OFDM symbols across the frequency domain go through an OFDM modulator, in which the Inverse Discrete Fourier Transform (IDFT) is applied, resulting in a vector of time-domain digital samples. Following these operations, a CP is added which extends the length of this vector prior to its conversion from parallel to serial form. Finally, the resulting sequence of time-domain samples is shaped by Tx filter (rectangular filter) and sampler for transmission over the radio channel. At the receiver side (Rx), the reverse sequence of operations take place, that is: Rx filtering and sampling; conversion from serial to parallel form; removal of the CP; mapping to frequency domain via DFT; channel equalization and detection (E/D); followed by the final conversion from parallel to serial. In practice, the required DFT and

IDFT operations are implemented in the form of a fast Fourier transform (FFT) algorithm and its inverse, the IFFT, respectively. By adding and later removing the CP, wideband ISI

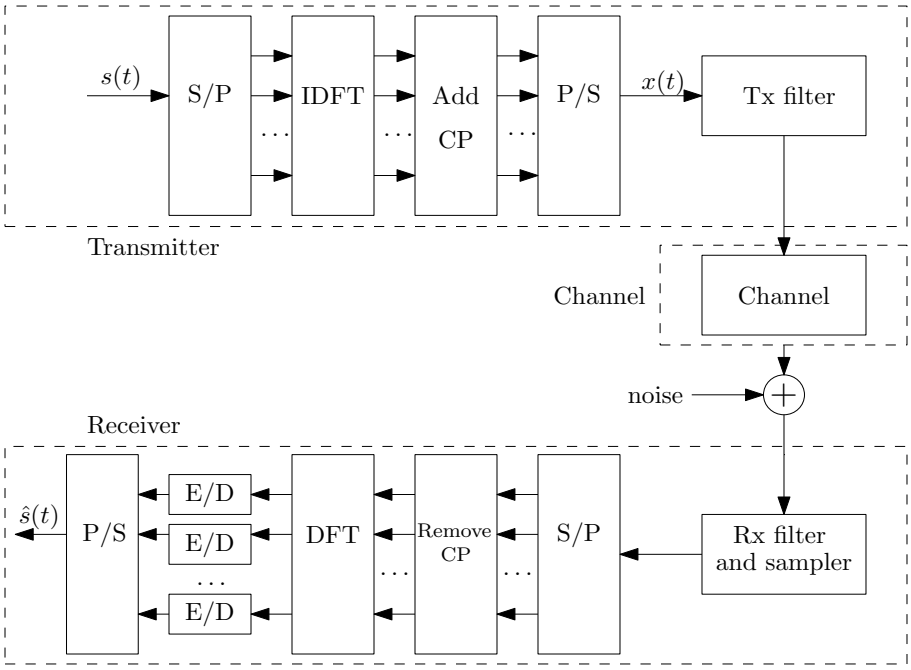


Fig. 2.2 CP-OFDM

and inter carrier interference (ICI) can be eliminated, which is one of the main motivations for using OFDM as an efficient mean to deal with dispersive channels.

However, the OFDM scheme also has its drawbacks [6]. One of its main disadvantages is related to the use of the CP. Indeed, since the CP is removed at the Rx side, the information contained in the CP is not utilized. Hence, the elimination of ISI and ICI comes at the cost of a reduction of the effective transmission rate, i.e. system capacity, which is due to adding the redundant CP message.

2.2 OFDM/OQAM

2.2.1 Pulse Shaping and Real Orthogonality

CP-OFDM is by far the most widely used of finding numerous application in both wireless and wireline communications. However, it uses rectangular pulse shaping on each subcar-

rier, leading to high out-of-band radiance, as well as ingress interference. Moreover, the OFDM system sacrifices data transmission rate because of the insertion of CP.

To overcome the aforementioned drawbacks, researchers have been looking into different MCM techniques. Recently, OFDM based on offset quadrature amplitude modulation (OQAM), referred to as OFDM/OQAM, has been considered as a promising alternative to conventional CP-OFDM for transmission over multi-path fading channels [3]. Similar to OFDM, subcarriers of the signal overlap each other in frequency to achieve a high spectral efficiency in OFDM/OQAM. However, in OFDM/OQAM, each subcarrier is modulated with a staggered OQAM symbol, i.e., the real and imaginary parts of each QAM symbol are processed separately with double the symbol rate. Another main difference comes from the pulse shaping. Instead of rectangular pulses as in OFDM, OFDM/OQAM utilizes a well designed prototype filter which has a desirable time and frequency localization property. Frequency shifted versions of the prototype filter are used to implement the filter banks needed for data modulation and demodulation, i.e., the synthesis filter bank (SFB) used for modulation of the data substreams on the individual subcarriers in the Tx, and the analysis filter bank (AFB) needed for the reverse operation at the Rx [9]. OFDM/OQAM does not use any CP and offers the possibility to use different prototype filters, enabling a more efficient use of channel resources. Furthermore, the orthogonality constraint for OFDM/OQAM is relaxed, being limited to the real field while for OFDM, it has to be satisfied in the complex field. However, while channel estimation and equalization is fairly simple in CP-OFDM, real orthogonality has a considerable impact on channel estimation in OFDM/OQAM. In particular, the presence of subcarriers interference prior to equalization has to be solved by new channel estimation techniques.

2.2.2 Transceiver System Configuration

For the purpose of comparison between CP-OFDM and OFDM/OQAM, the main processing blocks of an OFDM/OQAM transceiver system are depicted in Fig. 2.3 [10, 21]. On the Tx side, the system consists of the following components:

- S/P and P/S conversion: same as in OFDM;
- Complex to real (C2R) conversion: where each complex QAM symbol of duration T_0 is converted to a pair of real OQAM symbols with duration $T_0/2$;

- SFB: where the individual real symbol substreams are modulated on their corresponding subcarrier with proper time offset.

On the Rx side, besides S/P and P/S conversion blocks, the system also includes

- AFB: where the received substreams are individually demodulated on their corresponding subcarrier;
- Real to complex (R2C) conversion: which performs the reverse operation as the C2R block;
- Equalization and detection (E/D): where the channel effects are removed and data detection is applied.

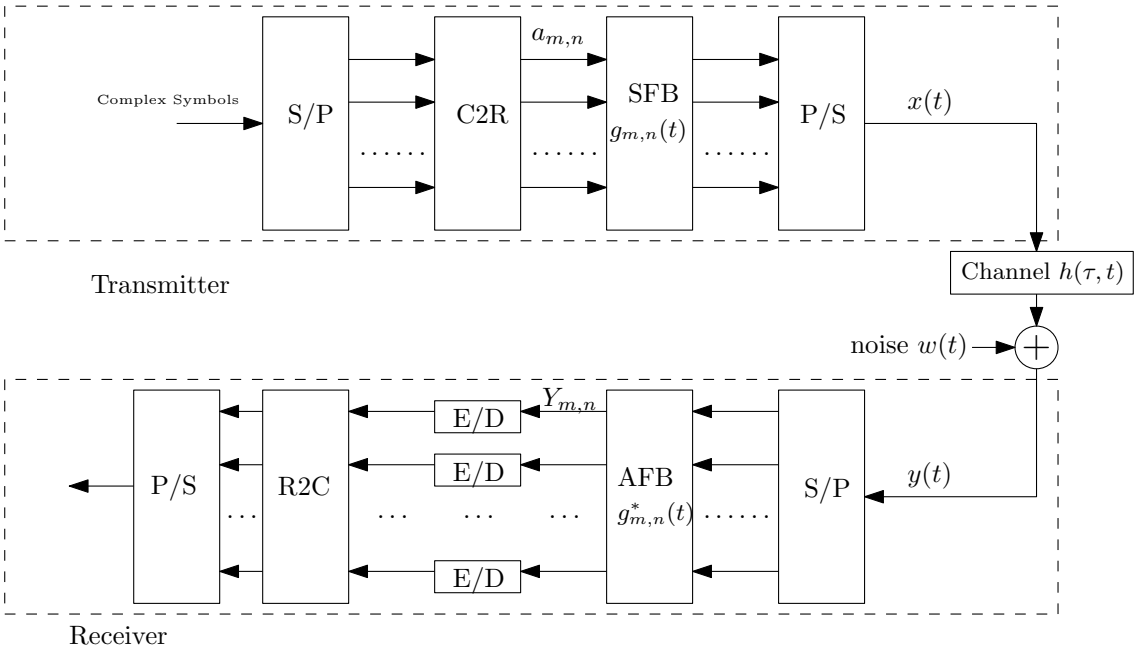


Fig. 2.3 OFDM/OQAM transceiver system model

Compared with CP-OFDM, OFDM/OQAM removes CP related blocks but add more sophisticated filter bank blocks, which represents a crucial change in system design.

2.2.3 Baseband Mathematical System Model

In this section, according to [20], we derive the mathematical system model of OFDM/OQAM to which the channel estimation algorithms can be further applied.

On the basis of transceiver system configuration in Fig. 2.3, we can write the output signal of the OFDM/OQAM synthesis filter bank (SFB) as follows [20, 25]

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{m=0}^{M-1} a_{m,n} g_{m,n}(t) \tag{2.1}$$

where

$$g_{m,n}(t) = g(t - n\tau_0) e^{j2\pi f_0 mt} e^{j\phi_{m,n}} \tag{2.2}$$

M is the total number of subcarriers, $a_{m,n}$ is the real symbol transmitted over the m-th subcarrier at the n-th time instance, $g(t)$ is a real-valued prototype filter, $\phi_{m,n} = \frac{\pi}{2}(n+m)$ stands for the phase term, $f_0 = 1/T_0$ is the subcarrier spacing and $\tau_0 = T_0/2$ denotes the half symbol delay. In general, the prototype filter is of finite duration $T_p = KT_0$, where K is a small integer, and is symmetrical in time, i.e. $g(T_p - t) = g(t)$. The sequence of real data symbols $a_{m,n}$ is modelled as a zero-mean random process, independent over time n and frequency m, with variance $\sigma_a^2 = E\{a_{m,n}^2\}$.

After transmission over the noisy radio channel, the baseband version of the received signal $y(t)$ can be written as

$$y(t) = \int_0^{\tau_{\max}} h(\tau, t) x(t - \tau) d\tau + w(t) \tag{2.3}$$

where $h(\tau, t)$ denotes the equivalent baseband impulse response of the slowly fading frequency selective channel at current time t and delay τ, with maximum delay spread τ_{max}, and $w(t)$ is an additive white Gaussian noise with zero-mean and variance σ_w^2. Upon substitution of (2.1) into (2.3), and under the assumption that the prototype function has relatively low variations in time over the interval $[0, \tau_{max}]$, that is, $g(t-\tau-n\tau_0) \approx g(t-n\tau_0)$, $y(t)$ can be equivalently written in the form

$$y(t) = \sum_{n=-\infty}^{\infty} \sum_{m=0}^{M-1} H_m(t) a_{m,n} g_{m,n}(t) + w(t) \tag{2.4}$$

where

$$H_m(t) = \int_0^{\tau_{max}} h(\tau, t) e^{-j2\pi f_0 m\tau} d\tau \tag{2.5}$$

After $y(t)$ passes through the analysis filter bank (AFB), the received signal in the

frequency domain can be further expressed as

$$\begin{aligned} Y_{m,n} &= \int_{-\infty}^{\infty} y(t) g_{m,n}^{*}(t) dt \\ &= \sum_{q=-\infty}^{\infty} \sum_{p=0}^{M-1} a_{p,q} \int_{-\infty}^{\infty} H_p(t) g_{p,q}(t) g_{m,n}^{*}(t) dt + W_{m,n} \end{aligned} \tag{2.6}$$

where $W_{m,n} = \int_{-\infty}^{\infty} w(t) g_{m,n}^{*}(t) dt$ is the corresponding white Gaussian noise in the frequency domain. Due to the orthogonality properties of $g(t)$, it follows that the noise term $W_{m,n}$ has zero-mean and variance σ_w^2.

Under distortion-free channel, perfect reconstruction of the real symbols at the Rx side is obtained thanks to the following real orthogonality condition [14]:

$$\Re\{\int_{-\infty}^{\infty} g_{m,n}(t) g_{p,q}^{*}(t) dt\} = \delta_{m,p}\delta_{n,q} \tag{2.7}$$

where $\delta_{m,p}$ denotes the Kronecker delta function. Indeed, suppose the channel impulse response (CIR) does not change appreciably within the prototype filter duration, so that $H_p(t) \approx H_p(q\tau_0) \triangleq H_{p,q}, t \in [q\tau_0, q\tau_0 + T_p]$. Using this approximation and substituting (2.7) into (2.6), then (2.6) simplifies to

$$Y_{m,n} = H_{m,n} a_{m,n} + j \sum_{(p,q)\neq(m,n)} H_{p,q} a_{p,q} G_{p,q}^{m,n} + W_{m,n} \tag{2.8}$$

where we define

$$G_{p,q}^{m,n} = \Im\{\int_{-\infty}^{\infty} g_{m,n}(t) g_{p,q}^{*}(t) dt\} \tag{2.9}$$

In (2.8), $H_{p,q}$ is the so-called channel frequency response (CFR) where the indices p and q respectively identify the the subcarrier frequency and time instance. The coefficients $G_{p,q}^{m,n}$ provides a measure of the correlation or match between $g_{m,n}(t)$ and $g_{p,q}(t)$. Equation (2.8) is the most frequently used OFDM/OQAM system model in the literature.

2.3 Block based Sign Covariance Matrix (SCM) Algorithm

In this section, we review the block based SCM algorithm to address the channel estimation in OFDM/OQAM systems, as presented in [20].

If the prototype filter $g(t)$ has a good time-frequency localization property, the level of the correlation $|G_{p,q}^{m,n}|$ decays rapidly as the distance between (p,q) and (m,n) increases. On this basis, let us define a neighbourhood $\Omega_{m,n}$ of a given center point (but excluding the latter) such that $G_{p,q}^{m,n} \approx 0$ if $(p,q) \notin \Omega_{m,n} \cup \{(m,n)\}$. Then, if we assume that the channel coefficients $H_{p,q}$ are nearly constants within the neighbourhood $\Omega_{m,n}$, we have that

$$\text{if} \quad (p,q) \in \Omega_{m,n} \quad H_{p,q} \approx H_{m,n} \tag{2.10}$$

This condition will be satisfied if the channel is slowly-varying and the frequency separation between the sub-carriers is small, i.e. large value of M for a given fixed system bandwidth. Then (2.8) becomes

$$Y_{m,n} \approx H_{m,n}(a_{m,n} + j \sum_{(p,q)\in\Omega_{m,n}} G_{p,q}^{m,n} a_{p,q}) + W_{m,n} \tag{2.11}$$

For convenience, let us define the self-interference

$$b_{m,n} = \sum_{(p,q)\in\Omega_{m,n}} G_{p,q}^{m,n} a_{p,q} \tag{2.12}$$

and *effective* transmitted data on the m-th subcarrier and the n-th time instant.

$$X_{m,n} = a_{m,n} + j b_{m,n} \tag{2.13}$$

Hence, we can rewrite (2.11) as

$$Y_{m,n} = H_{m,n} X_{m,n} + W_{m,n} \tag{2.14}$$

Separating the real part and imaginary part of each quantity in (2.14), we obtain

$$\underbrace{\begin{bmatrix} Y_{m,n}^{\Re} \\ Y_{m,n}^{\Im} \end{bmatrix}}_{\mathbf{y}_{m,n}} = \underbrace{\begin{bmatrix} H_{m,n}^{\Re} & -H_{m,n}^{\Im} \\ H_{m,n}^{\Im} & H_{m,n}^{\Re} \end{bmatrix}}_{\mathbf{H}_{m,n}} \underbrace{\begin{bmatrix} a_{m,n} \\ b_{m,n} \end{bmatrix}}_{\mathbf{x}_{m,n}} + \begin{bmatrix} W_{m,n}^{\Re} \\ W_{m,n}^{\Im} \end{bmatrix} \tag{2.15}$$

where the superscript $\Re$ and $\Im$ indicate real and imaginary parts, and where $\mathbf{y}_{m,n}$, $\mathbf{H}_{m,n}$ and $\mathbf{x}_{m,n}$ denote the (real-valued) received signal vector, 2×2 channel matrix and *effective*

transmitted signal vector.

Define the spatial sign vector as

$$\tilde{\mathbf{y}}_{m,n} = \begin{cases} \frac{\mathbf{y}_{m,n}}{\|\mathbf{y}_{m,n}\|} & \text{if} \quad \|\mathbf{y}_{m,n}\| \neq 0 \\ \mathbf{0} & \text{if} \quad \|\mathbf{y}_{m,n}\| = 0 \end{cases} \tag{2.16}$$

We can write the spatial sign covariance matrix as [26]

$$\mathbf{C}_{m,n} = E\{\tilde{\mathbf{y}}_{m,n}\tilde{\mathbf{y}}_{m,n}^T\} \tag{2.17}$$

The dominant eigenvector of $\mathbf{C}_{m,n}$, i.e., the one associated with its largest eigenvalue, is then computed and denoted as $\mathbf{u}_{m,n}$. Note that the rotation present in the received signal is induced by the channel matrix $\mathbf{H}_{m,n}$ in (2.15). Since the spatial sign covariance is rotation equivariant [27], the first column of the channel matrix may be computed as the dominant eigenvector of $\mathbf{C}_{m,n}$ [26]. Specifically, it can be shown that $\mathbf{u}_{m,n}$ is aligned to vector $[H_{m,n}^{\Re}, H_{m,n}^{\Im}]^T$, except for a scalar difference in magnitude. Therefore we have

$$\left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_{m,n}|H_{m,n}|\mathbf{u}_{m,n}, \quad \epsilon_{m,n} \in \{-1, +1\} \tag{2.18}$$

where $\epsilon_{m,n}$ is a sign ambiguity for the m-th subcarrier at time n.

So far, except for sign ambiguity, we still lack the knowledge of the magnitude $|H_{m,n}|$ to finalize the estimation of $H_{m,n}$. By further analysis of (2.12), the basic statistical properties of interference term $b_{m,n}$ can be derived as follows [20]:

$$\begin{aligned} E\{b_{m,n}\} &= \sum_{(p,q)\in\Omega_{m,n}} G_{p,q}^{m,n} E\{a_{p,q}\} = 0 \\ \sigma_b^2 = E\{b_{m,n}^2\} &= \sum_{(p,q)\in\Omega_{m,n}} |G_{p,q}^{m,n}|^2 E\{a_{p,q}^2\} \\ &\approx \sigma_a^2 \sum_{(p,q)\neq(m,n)} |G_{p,q}^{m,n}|^2 = \sigma_a^2 \end{aligned} \tag{2.19}$$

That is, the self-interference is zero-mean with variance σ_b^2 equal to that of the transmitted symbols $a_{m,n}$. In the derivation of the last line in (2.19), we include the small interference terms outside the neighbourhood $\Omega_{m,n}$ and utilized a key property of the real-valued

prototype filter, that is [14]

$$\sum_{(p,q)\neq(m,n)} |G_{p,q}^{m,n}|^2 = 1 \tag{2.20}$$

We further note that $E\{a_{m,n}b_{m,n}\} = 0$ and that

$$\mathbf{H}_{m,n}^T \mathbf{H}_{m,n} = |H_{m,n}|^2 \mathbf{I}_2 \tag{2.21}$$

where $\mathbf{I}_2$ is an identity matrix with dimension 2. Given these basic results, the power of received signal $Y_{m,n}$ in (2.14) is expressed as

$$\begin{aligned} P_{m,n} &= E\{|Y_{m,n}|^2\} = |H_{m,n}|^2(E\{a_{m,n}^2\} + E\{b_{m,n}^2\}) + \sigma_w^2 \\ &= |H_{m,n}|^2 2\sigma_a^2 + \sigma_w^2 \end{aligned} \tag{2.22}$$

where σ_w^2 is noise power. Therefore, the channel magnitude $|H_{m,n}|$ can be easily computed via

$$|H_{m,n}| = \sqrt{\frac{P_{m,n} - \sigma_w^2}{2\sigma_a^2}} \tag{2.23}$$

After $|H_{m,n}|$ and $\mathbf{u}_{m,n}$ are obtained, the subchannel estimate $[H_{m,n}^{\Re}, H_{m,n}^{\Im}]^T$ in (2.18) can be obtained, up to the sign ambiguity.

In practice, if the channel remains nearly constant over a data block with length N, $H_{m,n}$, $P_{m,n}$ and $\mathbf{C}_{m,n}$ are independent of time n and can be denoted as H_m, P_m and $\mathbf{C}_m$ respectively, so as the dominant eigenvector $\mathbf{u}_m$ and the sign ambiguity ϵ_m. In this case, the average power can be estimated by time averaging the square magnitudes of the received signal values on that same subchannel, that is

$$\hat{P}_m = \frac{1}{N} \sum_{n=1}^{N} |Y_{m,n}|^2 \tag{2.24}$$

In the same way, the spatial sign covariance matrix $\mathbf{C}_{m,n}$ in (2.17) can be estimated as

$$\hat{\mathbf{C}}_m = \frac{1}{N} \sum_{n=1}^{N} \tilde{\mathbf{y}}_{m,n} \tilde{\mathbf{y}}_{m,n}^T \tag{2.25}$$

where N is the number of symbols for time averaging.

Block based SCM is easy to implement and the estimation procedure can be summarized in 2.3.1. While it is presented only for the m-th subband, in practice, it has to be implemented separately, so that $\hat{P}_m$, $\hat{\mathbf{C}}_m$ and H_m are obtained for each frequency index $m = 1, 2, \ldots, M$. The resolution of the sign ambiguity ϵ_m requires transmitting a single

Algorithm 2.3.1 Block based SCM

1: $\hat{P}_m = \frac{1}{N}\sum_{n=1}^{N} |Y_{m,n}|^2$
2: $\hat{\mathbf{C}}_m = \frac{1}{N}\sum_{n=1}^{N} \tilde{\mathbf{y}}_{m,n}\tilde{\mathbf{y}}_{m,n}^T$
3: $|H_m| = \sqrt{\frac{\hat{P}_m - \sigma_w^2}{2\sigma_a^2}}$
4: Compute $\mathbf{u}_m$, which is the dominant eigenvector of $\hat{\mathbf{C}}_m$
5: $\left[H_m^{\Re}, H_m^{\Im}\right]^T = \epsilon_m |H_m| \mathbf{u}_m$

real pilot symbol on that subcarrier over a given period of time.

The initially obtained estimate of CFR is denoted as $\check{\mathbf{h}}_{block} = [H_1, H_2, H_3, \ldots, H_M]^T$. To further reduce the estimation noise in $\check{\mathbf{h}}_{block}$, the low rank property of the channel can be exploited. Suppose we know the length L of CIR, where $L \ll M$ is assumed. Then, the CFR $\check{\mathbf{h}}_{block}$ can be converted to a time-domain CIR via IDFT, truncated to an appropriate length L and zero-padded, and finally converted back to the frequency domain by applying a DFT. This procedure produces an enhanced estimate of the CFR, denoted as $\hat{\mathbf{h}}_{block}$, since the truncation in the time-domain amounts to the removal of estimation noise. The procedure can be expressed in matrix form as follow:

$$\hat{\mathbf{h}}_{block} = \mathbf{F}\mathbf{F}^H \check{\mathbf{h}}_{block} \tag{2.26}$$

where $\mathbf{F}$ is an $M \times L$ partial DFT matrix with entries $\mathbf{F}_{m,l} = \frac{1}{\sqrt{M}} e^{-j2\pi ml/M}$.

2.4 Limitations of Block based SCM

The block based SCM algorithm is semi-blind because it requires a pilot per tone per block to resolve the sign ambiguity. Still, compared to conventional training based techniques, SCM leads to a significant reduction in the number of required training pilots. However, block based SCM as in [20] is plagued by a number of limitations.

Small number M of subcarrier: We recall that the system model equation (2.11) is derived under the assumption (2.10). For a given (i.e. fixed) system bandwidth, as the

number M of subcarriers is reduced, the frequency separation between them increases. Eventually, when this separation reaches a value comparable to the channel coherence bandwidth, (2.10) is not satisfied and (2.11) no longer holds. In this case, the use of the above SCM estimator cannot be fully justified; in practice, it is observed that for too small values of M, its performance will deteriorate to some extent. To overcome this issue, it is necessary to replace (2.11) by a more accurate, coupled model in which received data on tone m depends, in addition to $H_{m,n}$, on surrounding (but different) values of $H_{p,q}$ in the neighbourhood of (m, n). This falls outside the scope of this thesis but remains a possible avenue for future work.

Small data block size N: From a statistical perspective, the value of block length N has a great impact on the estimation of $\hat{P}_m$ (2.24) and $\hat{\mathbf{C}}_m$ (2.25). Block processing highly relies on the assumption that the underlying channel remains nearly constant during a block. Under this assumption, the use of a larger value of N improves the quality of the above estimates, which in turn reduces the estimation errors in the CFR. In practice, this assumption limits the block length and for smaller values of N, the estimated CFR will be less accurate. This issue can be partly solved by the frequency averaging technique to be proposed in Section 3.4. Nevertheless, block processing is mostly useful for the case of time-invariant or slowly fading channels, but is not well suited to the estimation and tracking of fast time-varying channels. This behaviour is to be contrasted with the adaptive estimation algorithms proposed in the next chapter.

2.5 Chapter Summary

In this chapter, a general overview of MCM was given at first, then the classical CP-OFDM technique was briefly introduced. Given the drawbacks of CP-OFDM, OFDM/OQAM is currently considered as alternative and its general structure was described with comparison to CP-OFDM. The mathematical system model for OFDM/OQAM was subsequently derived in detail. Based on this model, we reviewed and analysed the block based SCM channel estimation algorithm proposed in [20]. Finally, the limitations of this algorithm were briefly mentioned, which motivate us to propose new adaptive estimation algorithms in the next chapter.

Chapter 3

Adaptive Semi-blind Channel Estimation

In this chapter, we first propose four adaptive algorithms to address the channel estimation problem for OFDM/OQAM systems, including Adaptive SCM and three constant modulus algorithms (CMA). The CMA are then further extended to multi-modulus algorithms (MMA) in order to accommodate higher order symbol constellations. Finally, we investigate the combination of the block based SCM and Adaptive SCM algorithm with a simple frequency averaging technique, to improve estimation performance.

3.1 Adaptive SCM

In Chapter 2, we reviewed the block based SCM algorithm for semiblind channel estimation in OFDM/OQAM systems. It can be concluded from (2.17), (2.18) and (2.23) that the unknown channel coefficient $H_{m,n}$ for the m-th subcarrier can be jointly determined by $P_{m,n}$ and $\mathbf{C}_{m,n}$ up to a sign ambiguity. In the original SCM algorithm, these quantities are assumed to remain constant over a data block, say for symbol index $n = 0, \ldots, N-1$, and they are estimated by time averaging as in (2.24) and (2.25). However, this approach is hardly applicable in a time-varying environment where the channel coefficients $H_{m,n}$ change during a data block. As explained earlier, in such a situation, a more effective approach is to us an adaptive algorithm for the estimation and tracking of $H_{m,n}$.

Hence, an idea comes to us naturally that if the required quantities, i.e., $P_{m,n}$ and $\mathbf{C}_{m,n}$ can be updated efficiently as the symbol time n increases, the corresponding channel esti-

mate of $H_{m,n}$ can also be refreshed through the main sequence of operations in the SCM algorithm, which involves computation of the dominant eigenvector of $\mathbf{C}_{m,n}$ to determine the phase rotation and final magnitude computation via (2.23). We note that when considering such an adaptive estimation scheme, subscript n cannot be ignored because the values of $H_{m,n}, P_{m,n}$ and $\mathbf{C}_{m,n}$ are intrinsically time-dependent. Now the problem comes down to designing efficient update procedures for the recursive estimation of $P_{m,n}$ and $\mathbf{C}_{m,n}$. Specifically, suppose that previous estimates of $P_{m,n-1}$, $\mathbf{C}_{m,n-1}$ are available, then given the current received signal sample $\mathbf{y}_{m,n}$, we seek simple relationships that will allow us to use this new information to update these previous estimates at time $n-1$ into new ones at time n.

We now introduce a widely used update procedure for this kind of situation, which is called exponential smoothing and mathematically characterized by the generic formula given below:

$$S_n = \alpha S_{n-1} + (1-\alpha) X_n \tag{3.1}$$

where α, with $0 < \alpha < 1$, is the smoothing factor, S_n and S_{n-1} denote the current and previous estimates, and X_n is the current observation. According to (3.1), the current estimate is expressed as a simple weighted sum of the previous estimate and the new observation. The difference equation (3.1) can be explicitly solved for S_n in terms of all the past and current values of X_k, for $k = 0, 1, \ldots, n$. In this way, it can be shown that exponential smoothing amounts to applying a weight proportional to α^{n-k} to X_k.

Based on the exponential smoothing (3.1), recursive estimates of $P_{m,n}$ and $\mathbf{C}_{m,n}$ can be obtained as

$$\begin{aligned} P_{m,n} &= \alpha P_{m,n-1} + (1-\alpha)\|\mathbf{y}_{m,n}\|^2 \\ \mathbf{C}_{m,n} &= \alpha \mathbf{C}_{m,n-1} + (1-\alpha)\tilde{\mathbf{y}}_{m,n}\tilde{\mathbf{y}}_{m,n}^T \end{aligned} \tag{3.2}$$

where $\mathbf{y}_{m,n}$ denotes the real-valued signal vector and $\tilde{\mathbf{y}}_{m,n}$ is the spatial sign vector, defined in (2.15) and (2.16), respectively.

After we have obtained $P_{m,n}$ and $\mathbf{C}_{m,n}$ as above, the instantaneous estimate of the channel coefficient $H_{m,n}$ can be generated via the main steps of the SCM algorithm exposed in Section 2.3. This includes the determination of the dominant eigenvector of $\mathbf{C}_{m,n}$, denoted as $\mathbf{u}_{m,n}$, and the determination of the channel magnitude from $P_{m,n}$, leading to

$$\left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m |H_{m,n}| \mathbf{u}_{m,n} \tag{3.3}$$

where $\epsilon_m \in \{-1, +1\}$ is the sign ambiguity for the corresponding channel coefficients $H_{m,n}$ on the m-th frequency subband.

Based on the analysis above, the Adaptive SCM algorithm can be summarized as 3.1.1. While it is presented only for the m-th subband, in practice, it has to be implemented separately, so that $P_{m,n}$, $\mathbf{C}_{m,n}$ and $H_{m,n}$ are properly updated for each frequency index $m = 1, 2, \ldots, M$.

Algorithm 3.1.1 Adaptive SCM

Initialization: $P_{m,0} = 0, \mathbf{C}_{m,0} = \mathbf{0}$

1: **for** $n = 1, 2, \ldots$
2: $\quad P_{m,n} = \alpha P_{m,n-1} + (1-\alpha)\|\mathbf{y}_{m,n}\|^2$
3: $\quad \mathbf{C}_{m,n} = \alpha \mathbf{C}_{m,n-1} + (1-\alpha)\tilde{\mathbf{y}}_{m,n}\tilde{\mathbf{y}}_{m,n}^T$
4: $\quad |H_{m,n}| = \sqrt{\frac{P_{m,n}-\sigma_w^2}{2\sigma_a^2}}$
5: $\quad$ Compute $\mathbf{u}_{m,n}$, which is the dominant eigenvector of $\mathbf{C}_{m,n}$
6: $\quad \left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m |H_{m,n}| \mathbf{u}_{m,n}$
7: **end**

If we assume that the modulated data are transmitted over a time-invariant channel, as time n increases, the channel estimate $H_{m,n}$ in the m-th subband will gradually converge to a steady state value, which we can denoted as H_m. The resulting estimate of the CFR after convergence can be represented by the vector $\check{\mathbf{h}} = [H_1, H_2, \ldots, H_M]^T$. Following this step, further processing of $\check{\mathbf{h}}$, as exposed in (2.26), can be applied to obtain an enhanced estimate, denoted as $\hat{\mathbf{h}}$. As explained earlier, this amounts to the application of lower rank IDFT and DFT matrices to the initial estimate, in order to implement truncation in the time domain. However, if the data is transmitted over a time-varying channel, Algorithm 3.1.1 can be used effectively for the initial acquisition *and* tracking of the channel. In this case, at every symbol time n, we can form an instantaneous CFR vector, which we denote as $\check{\mathbf{h}}_n = [H_{1,n}, H_{2,n}, \ldots, H_{M,n}]^T$. The latter can be further enhanced via low-rank DFT processing as above, yielding the time-dependent CFR vector $\hat{\mathbf{h}}_n$. Specifically,

$$\hat{\mathbf{h}}_n = \mathbf{F}\mathbf{F}^H \check{\mathbf{h}}_n \tag{3.4}$$

where $\mathbf{F}$ is the partial DFT matrix as defined in (2.26). The main issue regarding the application of Algorithm 3.1.1 relates to the determination of the ambiguity parameter $\epsilon_m \in \{-1, +1\}$ for the m-th subband. In effect, this sign ambiguity in the calculation

of the dominant eigenvector could be time-dependent, as represented by $\epsilon_{m,n}$. If this was the case, then the resolution of the ambiguity would require the use of one pilot per tone per time iteration, thereby rendering the algorithm useless. In practice, however, we find that by properly normalizing the eigenvector $\mathbf{u}_{m,n}$, i.e. by enforcing its first entry to be non-negative, the sign ambiguity becomes independent of the time n. In this case, we find that only a few pilot symbols per tone are needed at an early stage of the adaptation process (i.e. after matrix $\mathbf{C}_{m,n}$ has started to converge) to determine the value of ϵ_m, which then remains constant for long stretch of time. This aspect will be further investigated via simulations in Chapter 4.

3.2 Constant Modulus Algorithms (CMA)

Extracting communication signals from interference is particularly important in the design of modern wireless communication systems. The constant modulus (CM) criterion offers a bandwidth efficient approach to tackle this problem [28]. Constant Modulus Algorithms (CMA) exploit the low modulus fluctuation displayed by most communication signals, such as frequency modulation (FM), phase modulation (PM), etc. CMA are often used in the blind or semi-blind channel equalization to recover transmitted symbols from the received signals which are corrupted by interference. In practice, if we adopt a zero-forcing (ZF) equalization technique followed by data detection on the receiver, the equalizer coefficients are actually closely linked to the channel coefficients [29]; hence, the determination of the equalizer coefficients is in effect equivalent to the problem of channel estimation. From this perspective, CMA can be applied to channel estimation in a proper way.

Recall our system model in (2.15), i.e.,

$$\underbrace{\begin{bmatrix} Y^{\Re}_{m,n} \\ Y^{\Im}_{m,n} \end{bmatrix}}_{\mathbf{y}_{m,n}} = \underbrace{\begin{bmatrix} H^{\Re}_{m,n} & -H^{\Im}_{m,n} \\ H^{\Im}_{m,n} & H^{\Re}_{m,n} \end{bmatrix}}_{\mathbf{H}_{m,n}} \underbrace{\begin{bmatrix} a_{m,n} \\ b_{m,n} \end{bmatrix}}_{\mathbf{x}_{m,n}} + \underbrace{\begin{bmatrix} W^{\Re}_{m,n} \\ W^{\Im}_{m,n} \end{bmatrix}}_{\mathbf{v}_{m,n}} \tag{3.5}$$

From (3.5), we find that received signal consists of the channel information $H_{m,n}$, source symbol $a_{m,n}$, self-interference $b_{m,n}$ and noise $\mathbf{v}_{m,n}$. Let us define the channel vector for the

m-th subband at symbol time n as

$$\mathbf{h}_{m,n} = \begin{bmatrix} H_{m,n}^{\Re} \\ H_{m,n}^{\Im} \end{bmatrix} \in \mathbb{R}^2 \tag{3.6}$$

where we note that $\|\mathbf{h}_{m,n}\|^2 = |H_{m.n}|^2$. Multiplying both sides of (3.5) by $\mathbf{h}_{m,n}^T$, we obtain

$$\mathbf{h}_{m,n}^T \mathbf{y}_{m,n} = |H_{m,n}|^2 a_{m,n} + \mathbf{h}_{m,n}^T \mathbf{n}_{m,n} \tag{3.7}$$

In this way, we elegantly remove the self-interference $b_{m,n}$ thanks to the real orthogonal matrix $\mathbf{H}_{m,n}$, while the remaining term is due to the additive noise only. Equation (3.7) can be further simplified through normalization by $\|\mathbf{h}_{m,n}\|^2$, leading us to define

$$\begin{aligned} z_{m,n} &= \frac{\mathbf{h}_{m,n}^T}{\|\mathbf{h}_{m,n}\|^2} \mathbf{y}_{m,n} \\ &= a_{m,n} + \frac{\mathbf{h}_{m,n}^T}{\|\mathbf{h}_{m,n}\|^2} \mathbf{n}_{m,n} \end{aligned} \tag{3.8}$$

This procedure amounts to equalization, which removes the channel filtering imposed on transmitted symbols, so that only the transmitted symbol corrupted by filtered noise are present in the equalized signal. Equation (3.8) shows that $z_{m,n}$ can be regarded as an approximation to the transmitted symbol $a_{m,n}$, which becomes more accurate as the noise power becomes smaller, corresponding to a larger signal-to-noise ratio (SNR). In fact, $\frac{\mathbf{h}_{m,n}}{\|\mathbf{h}_{m,n}\|^2}$ is the ZF equalizer since it directly makes use of the channel information.

In the present context, a CMA aims to solve an optimization problem with respect to unknown weight vector $\mathbf{w} \in \mathbb{R}^2$, formulated as

$$\min_{\mathbf{w}} E[(\mathbf{w}^T \mathbf{y}_{m,n})^2 - \gamma]^2 \tag{3.9}$$

where γ reflects the constant modulus of the transmitted symbols $a_{m,n}$, and the optimal solution to this problem is denoted as $\mathbf{w}_{m,n}$. This criterion attempts to minimize the dispersion of the modulus of the equalized output, i.e. $\mathbf{w}^T \mathbf{y}_{m,n}$, away from a constant γ. By making use of an expectation operator, this criterion indeed attempts to minimize the effect of noise on a statistical basis. For instance, in the case of binary phase shift keying (BPSK), which corresponds to 4QAM modulation in OFDM/OQAM, symbols $a_{m,n} \in \{-1, +1\}$ and

we can set $\gamma = 1$.

With the above choice of γ, the optimal solutions $\mathbf{w}_{m,n}$ to (3.9) can be obtained, which is called the CM based equalizer. In [30], it is proved that when the channel matrix has orthogonal columns, which in our case corresponds to (2.21), the minima of CM criterion satisfy the ZF conditions. This means that the CM based equalizer is equivalent to the ZF equalizer, i.e.,

$$\mathbf{w}_{m,n} = c_{m,n} \frac{\mathbf{h}_{m,n}}{\|\mathbf{h}_{m,n}\|^2} \tag{3.10}$$

where $c_{m,n}$ is a scalar; furthermore, when the SNR is large, $c_{m,n} \approx \epsilon_{m,n} = \pm 1$. Thus we propose to use $\epsilon_{m,n} \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$ as an estimate of $\mathbf{h}_{m,n}$. For simplicity, the estimate is still denoted by $\mathbf{h}_{m,n}$, i.e.,

$$\mathbf{h}_{m,n} \approx \epsilon_{m,n} \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2} \tag{3.11}$$

Therefore, if we can solve (3.9) for the optimal equalizer vector $\mathbf{w}_{m,n}$, the corresponding channel estimate $\mathbf{h}_{m,n}$ can be obtained directly up to a sign ambiguity $\epsilon_{m,n}$. It will be illustrated in Section 4.3.1 that for slow-varying channels, after the initial transient period, the ambiguity parameter $\epsilon_{m,n}$ will display a constant behaviour and hence can be considered as time-independent. On this basis, the sign ambiguity $\epsilon_{m,n}$ is replaced by ϵ_m in the sequel for simplicity.

In the following subsections, we investigate different search strategies to solve the CM criterion (3.9) and determine the optimal $\mathbf{w}_{m,n}$ in an adaptive manner. This includes gradient-based least mean square (LMS), Guass-Newton based LMS and the recursive least squares (RLS) adaptation.

3.2.1 Gradient-based LMS

Due to its computational simplicity, the stochastic gradient descent (SGD) approach is often employed to minimize the CM cost function (3.9). In this thesis, for convenience, the resulting gradient-based LMS algorithm is referred to as CMA-LMS-Grad.

The main steps in the algorithm derivation can be summarized as follows:

- Expanding the argument of the expectation operator in (3.9);
- Replacing the statistical expectation with instantaneous estimation,
- Calculating the gradient of the error function with respect to the weight vector $\mathbf{w}$

- Using the gradient to update the previous weight estimate $\mathbf{w}_{m,n-1}$ by making a small step in the negative direction of the gradient.

These steps, which make use of matrix calculus in the derivation of the gradient, are described in more details in [31]. The resulting equation for the recursive update of $\mathbf{w}$ are obtained as follows:

$$\mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} - \mu \mathbf{y}_{m,n}(z_{m,n}^2 - 1)z_{m,n} \tag{3.12}$$

where $z_{m,n}$ is the equalizer output based on the previous weight vector $\mathbf{w}_{m,n-1}$, i.e. [1],

$$z_{m,n} = \mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n} \tag{3.13}$$

and μ is a positive step size which should be carefully selected. A small step size value will lead to a smaller residual error after convergence but a slower convergence rate, whereas a large step size may improve convergence speed but eventually result in an oscillatory or unstable behaviour in the estimated weight vector. The choice of an optimal step size is investigated in [32]. Once the weight vector $\mathbf{w}_{m,n}$ for the m-th subband has been updated at symbol time n, the corresponding channel vector can be obtained as in (3.11), where $\epsilon_m = \pm 1$ is the sign ambiguity.

The initialization and update procedure of the resulting CMA-LMS-Grad algorithm is summarized in 3.2.1. While it is presented only for the m-th subband, in practice, it has to be implemented separately for each subband $m = 1, 2, \ldots, M$.

Algorithm 3.2.1 CMA-LMS-Grad

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T$

1: **for** $n = 1, 2, \ldots$
2: $\quad z_{m,n} = \mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n}$
3: $\quad \mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} - \mu \mathbf{y}_{m,n}(z_{m,n}^2 - 1)z_{m,n}$
4: $\quad \left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$
5: **end**

In the absence of *a priori* knowledge, the algorithm can be initialized with a non-zero vector, taken here as $[1, 0]^T$. Same as for the Adaptive SCM algorithm in Section 3.1, only a few pilot symbols per tone are needed at an early stage of the adaptation

[1] The definition of $z_{m,n}$ is different with that in (3.8). Here, $z_{m,n}$ represents the equalizer output.

process to determine the value of ϵ_m, which can then be assumed to remain constant for a long stretch of time, as will be illustrated through simulations in Chapter 4. At every symbol time n, the resulting instantaneous estimate of the CFR can be represented by the vector $\check{\mathbf{h}}_n = [H_{1,n}, H_{2,n}, \ldots, H_{M,n}]^T$. Furthermore, it can be enhanced via low-rank DFT processing as in (3.4), where the enhanced estimate is denoted as $\hat{\mathbf{h}}_n$.

Some researchers proposed other gradient-based CMA to improve the convergence speed or to reduce the steady state residual level. For example, the normalized constant modulus algorithms (NCMA) in [33] prove to be useful for the reduction in steady state error.

3.2.2 Gauss-Newton based LMS

The CMA-LMS-Grad uses the stochastic gradient as the search direction and as such, it is characterized by a slow convergence rate. To address this issue, in this section, we extend the regular Gauss-Newton method to solve the stochastic non-linear least squares (NLS) problem (3.9). For convenience, since the search direction is based on the Gauss-Newton method, the resulting algorithm will be referred to as CMA-LMS-GN in this thesis.

To begin with, let us review the regular (i.e. non-stochastic) Gauss-Newton method, as introduced in, e.g. [34]. In many applications such as in the global positioning system (GPS), the observation model is non-linear and has the following form:

$$\mathbf{y} = \mathbf{f}(\mathbf{w}) + \mathbf{v} \tag{3.14}$$

where $\mathbf{y} \in \mathbb{R}^N$ is the measurement or observation vector, $\mathbf{f} : \mathbb{R}^M \to \mathbb{R}^N$ is a non-linear differentiable function, $\mathbf{w} \in \mathbb{R}^M$ is an unknown parameter vector and $\mathbf{v}$ is an additive noise vector. A typical approach to tackling the estimation of $\mathbf{w}$ is to solve the following NLS problem

$$\min_{\mathbf{w}} \|\mathbf{y} - \mathbf{f}(\mathbf{w})\|^2 \tag{3.15}$$

The solution $\hat{\mathbf{w}}$ is referred to as the NLS estimator of the true parameter vector $\mathbf{w}$.

In the Gauss-Newton method, the above problem is simplified through linearization of the objective function. Specifically, let $\mathbf{w}^{(k-1)}$ denote the estimate of $\mathbf{w}$ obtained after $k-1$ iterations of the procedure. The Taylor series expansion of $\mathbf{f}(\mathbf{w})$ about $\mathbf{w}^{(k-1)}$ gives

$$\mathbf{f}(\mathbf{w}) = \mathbf{f}(\mathbf{w}^{(k-1)}) + \mathbf{J}(\mathbf{w}^{(k-1)})(\mathbf{w} - \mathbf{w}^{(k-1)}) + O(\|\mathbf{w} - \mathbf{w}^{(k-1)}\|^2) \tag{3.16}$$

where $\mathbf{J}(\mathbf{w})$ is the $N \times M$ Jacobian matrix of $\mathbf{f}$ at a point $\mathbf{w}$. If we use the linear part of (3.16) as an approximation to $\mathbf{f}(\mathbf{w})$ in (3.15), the latter simplifies into the following linear LS problem

$$\min_{\mathbf{w}} \|\mathbf{y} - \mathbf{f}(\mathbf{w}^{(k-1)}) - \mathbf{J}(\mathbf{w}^{(k-1)})(\mathbf{w} - \mathbf{w}^{(k-1)})\|^2 \tag{3.17}$$

By defining $\mathbf{p} = \mathbf{w} - \mathbf{w}^{(k-1)}$, (3.17) can be expressed more compactly as

$$\min_{\mathbf{p}} \|\mathbf{y} - \mathbf{f}(\mathbf{w}^{(k-1)}) - \mathbf{J}(\mathbf{w}^{(k-1)})\mathbf{p}\|^2 \tag{3.18}$$

This is an ordinary linear LS problem and its solution[2] $\mathbf{p}$ can be found by QR factorization or singular value decomposition (SVD) [34]. Once the optimal solution, say $\mathbf{p}^{(k)}$ has been obtained, the updated weight vector $\mathbf{w}^{(k)}$ can be calculated as

$$\mathbf{w}^{(k)} = \mathbf{w}^{(k-1)} + \mathbf{p}^{(k)} \tag{3.19}$$

This is the general procedure for the regular Gauss-Newton method. Due to its approximate nature, convergence is not guaranteed and in some cases, the objective function may not decrease at every iteration. To overcome this difficulty, alternatively, a *damped* Gauss-Newton method is proposed which scales down the search step $\mathbf{p}^{(k)}$, i.e.,

$$\mathbf{w}^{(k)} = \mathbf{w}^{(k-1)} + \beta\mathbf{p}^{(k)} \tag{3.20}$$

where β, with $0 < \beta < 1$ represents the scale factor.

Now, we seek to apply the Gauss-Newton method to the stochastic NLS problem in (3.9). To this end, as was the case of the CMA-LMS-Grad, we first replace the statistical expectation with instantaneous estimation. The resulting deterministic NLS problem is given by

$$\min_{\mathbf{w}}(|\mathbf{w}^T\mathbf{y}_{m,n}|^2 - 1)^2 \tag{3.21}$$

Comparing (3.15) and (3.21), we can find the correspondence between quantities appearing in these two equations. Specifically, the constant value 1 in (3.21) corresponds to $\mathbf{y}$ in (3.15), while the term $(\mathbf{w}^T\mathbf{y}_{m,n})^2$, which depends on the unknown weight vector $\mathbf{w}$, corresponds to the non-linear function $\mathbf{f}(\mathbf{w})$. Hence, if we use $\mathbf{w}_{m,n-1}$ obtained at time $n-1$ as the initial

[2]Note that if $N < M$, the problem is under-determined and the solution may not be unique. In this case, we take the minimum 2-norm solution.

point for iteration, and only one iteration is performed, the updated weight vector $\mathbf{w}_{m,n}$ is calculated as

$$\mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} + \beta \mathbf{p}_{m,n} \tag{3.22}$$

where $\mathbf{p}_{m,n}$ is the optimal solution to an ordinary linear LS problem given by

$$\min_{\mathbf{p}} |1 - (\mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n})^2 - \mathbf{J}(\mathbf{w}_{m,n-1})\mathbf{p}|^2. \tag{3.23}$$

In (3.23), $\mathbf{J}(\mathbf{w}_{m,n-1})$ is the 1×2 Jacobian matrix (which in this case corresponds to the gradient vector) of $(\mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n})^2$ at a given point $\mathbf{w}_{m,n-1}$. The latter can be calculated explicitly as

$$\mathbf{J}(\mathbf{w}_{m,n-1}) = 2\mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n} \mathbf{y}_{m,n}^T \tag{3.24}$$

In practice, since $\mathbf{J}(\mathbf{w}_{m,n-1})$ is a row vector, the LS problem (3.23) is under-determined and admits a one-dimensional subspace of solutions. Among these, a unique optimal solution can be obtained by enforcing a minimum 2-norm condition. The resulting solution is given by

$$\mathbf{p}_{m,n} = \frac{1 - (\mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n})^2}{\mathbf{J}(\mathbf{w}_{m,n-1})\mathbf{J}^T(\mathbf{w}_{m,n-1})} \mathbf{J}^T(\mathbf{w}_{m,n-1}). \tag{3.25}$$

Once the weight vector $\mathbf{w}_{m,n}$ for the m-th subband has been updated at symbol time n, the corresponding channel vector can be obtained as in (3.11), where $\epsilon_m = \pm 1$ is the sign ambiguity. The resulting algorithm, referred to as CMA-LMS-GN in this thesis, is summarized in 3.2.2 for the m-th subband, where as before, $m = 1, 2, \ldots, M$.

Algorithm 3.2.2 CMA-LMS-GN

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T$

1: **for** $n = 1, 2, \ldots$
2: $\quad z_{m,n} = \mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n}$
3: $\quad \mathbf{J}_{m,n} = 2 z_{m,n} \mathbf{y}_{m,n}^T$
4: $\quad \mathbf{p}_{m,n} = \frac{1 - z_{m,n}^2}{\mathbf{J}_{m,n}\mathbf{J}_{m,n}^T} \mathbf{J}_{m,n}^T$
5: $\quad \mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} + \beta \mathbf{p}_{m,n}$
6: $\quad \left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$
7: **end**

Similar to the CMA-LMS-Grad, this algorithm is initialized with a non-zero weight

vector, taken here as $\mathbf{w}_{m,0} = [1, 0]^T$. Only a few pilot symbols per tone are needed at an early stage of the adaptation process to resolve the sign ambiguity ϵ_m, which can then be assumed to remain constant over time, as will be further discussed in Chapter 4. The resulting instantaneous estimate of the CFR at time n can be represented by the vector $\check{\mathbf{h}}_n = [H_{1,n}, H_{2,n}, \ldots, H_{M,n}]^T$, which can be enhanced via low-rank DFT processing as in (3.4), where the enhanced estimate is denoted as $\hat{\mathbf{h}}_n$.

3.2.3 Recursive Least Squares (RLS)

Due to its use of instantaneous estimates and stochastic gradient approach, the LMS algorithm has a relatively slow convergence rate, and its behaviour largely relies on the step size μ. The RLS discussed in this section offers an alternative to the LMS: its convergence rate from initialization is typically faster than LMS but at the expense of higher complexity. Here, we first provide an overview of the standard RLS algorithm and then consider its application to the CM criterion.

The underlying cost function of the RLS algorithm is given by

$$C(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i} |\mathbf{w}^T \mathbf{y}_i - d_i|^2 \tag{3.26}$$

where λ is a forgetting factor with $0 < \lambda < 1$, used to control the memory of the algorithm, $\mathbf{y}_i$ is the input vector and d_i is the desired output at time i. According to the least squares (LS) theory, the optimal weight vector $\mathbf{w}$ is found by [35]

$$\mathbf{w}_n = \mathbf{\Phi}_n^{-1} \mathbf{t}_n \tag{3.27}$$

where

$$\begin{aligned} \mathbf{\Phi}_n &= \sum_{i=1}^{n} \lambda^{n-i} \mathbf{y}_i \mathbf{y}_i^T \\ \mathbf{t}_n &= \sum_{i=1}^{n} \lambda^{n-i} \mathbf{y}_i d_i \end{aligned} \tag{3.28}$$

In (3.28), $\mathbf{\Phi}_n$ is the (scaled) correlation matrix of the input vectors $\mathbf{y}$, while $\mathbf{t}_n$ is the cross-correlation vector between the input and the desired output d_i. By expressing the defining equations for $\mathbf{\Phi}_n$ and $\mathbf{t}_n$ as first order difference equations and applying the matrix

inversion lemma [35], the standard RLS update procedure can be formulated as

$$\begin{aligned}
\mathbf{q}_n &= \mathbf{P}_{n-1}\mathbf{y}_n \\
\mathbf{g}_n &= \mathbf{q}_n/(\lambda + \mathbf{y}_n^T\mathbf{q}_n) \\
e_n &= d_n - \mathbf{w}_{n-1}^T\mathbf{y}_n \\
\mathbf{w}_n &= \mathbf{w}_{n-1} + \mathbf{g}_n e_n \\
\mathbf{P}_n &= \lambda^{-1}\mathbf{P}_{n-1} - \lambda^{-1}\mathbf{g}_n\mathbf{y}_n^T\mathbf{P}_{n-1}
\end{aligned} \tag{3.29}$$

where $\mathbf{P}_n$ is often initialized as a scaled identity matrix, i.e., $\mathbf{P}_0 = \delta^{-1}\mathbf{I}$ with $\delta > 0$.

Now we combine the CM criterion with the RLS approach and apply the update procedure above to our problem. Replacing the statistical expectation operator in (3.9) with an exponentially weighted time average, the corresponding CMA attempts to minimize the cost function

$$C_m(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i}(|\mathbf{w}^T\mathbf{y}_{m,i}|^2 - 1)^2 \tag{3.30}$$

We note that this cost function is not quadratic in $\mathbf{w}$, but instead involves 4th powers of the entries of this vector. Hence, it is not possible to directly apply the RLS algorithm (3.29) to this problem. To overcome the limitation, it is proposed in [28] to approximate the above cost function by a new one that only involves second powers in $\mathbf{w}$, and is given by

$$C'_m(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i}(\mathbf{w}^T\mathbf{y}_{m,i}\mathbf{y}_{m,i}^T\mathbf{w}_{m,i-1} - 1)^2 \tag{3.31}$$

Note that in the above cost function, the quantities $\mathbf{w}_{m,i-1}$, for $i = 1, ..., n$, are available from previous iterations.

Now, since (3.31) is quadratic in the unknown quantities, the update of $\mathbf{w}$ can proceed as in the derivation of the standard RLS algorithm in (3.29). The main differences are that the original input $\mathbf{y}_i$ in (3.26) is replaced by $\mathbf{z}_{m,i} = \mathbf{y}_{m,i}\mathbf{y}_{m,i}^T\mathbf{w}_{m,i-1}$, and the desired output

d_i is set to a constant 1. On this basis, the RLS algorithm equations (3.29) become [28]

$$\begin{aligned}
\mathbf{z}_{m,n} &= \mathbf{y}_{m,n}\mathbf{y}_{m,n}^T\mathbf{w}_{m,n-1} \\
\mathbf{q}_{m,n} &= \mathbf{P}_{m,n-1}\mathbf{z}_{m,n} \\
\mathbf{g}_{m,n} &= \mathbf{q}_{m,n}/(\lambda + \mathbf{z}_{m,n}^T\mathbf{q}_{m,n}) \\
e_{m,n} &= 1 - \mathbf{w}_{m,n-1}^T\mathbf{z}_{m,n} \\
\mathbf{w}_{m,n} &= \mathbf{w}_{m,n-1} + \mathbf{g}_{m,n}e_{m,n} \\
\mathbf{P}_{m,n} &= \lambda^{-1}\mathbf{P}_{m,n-1} - \lambda^{-1}\mathbf{g}_{m,n}\mathbf{z}_{m,n}^T\mathbf{P}_{m,n-1}
\end{aligned} \tag{3.32}$$

For the m-th subband, the corresponding channel vector $\mathbf{h}_{m,n}$ can be obtained from (3.11) once $\mathbf{w}_{m,n}$ has been updated at time n. The resulting algorithm, referred to as CMA-RLS in this thesis, is summarized in 3.2.3 for the m-th subband, $m = 1, 2, \ldots, M$.

Algorithm 3.2.3 CMA-RLS

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T, \mathbf{P}_{m,0} = \delta^{-1}\mathbf{I}, \delta = \text{small} \quad \text{positive} \quad \text{constant}$

1: **for** $n = 1, 2, \ldots$
2: $\quad \mathbf{z}_{m,n} = \mathbf{y}_{m,n}\mathbf{y}_{m,n}^T\mathbf{w}_{m,n-1}$
3: $\quad \mathbf{q}_{m,n} = \mathbf{P}_{m,n-1}\mathbf{z}_{m,n}$
4: $\quad \mathbf{g}_{m,n} = \mathbf{q}_{m,n}/(\lambda + \mathbf{z}_{m,n}^T\mathbf{q}_{m,n})$
5: $\quad \mathbf{P}_{m,n} = (\mathbf{P}_{m,n-1} - \mathbf{g}_{m,n}\mathbf{z}_{m,n}^T\mathbf{P}_{m,n-1})/\lambda$
6: $\quad e_{m,n} = 1 - \mathbf{w}_{m,n-1}^T\mathbf{z}_{m,n}$
7: $\quad \mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} + \mathbf{g}_{m,n}e_{m,n}$
8: $\quad \left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$
9: **end**

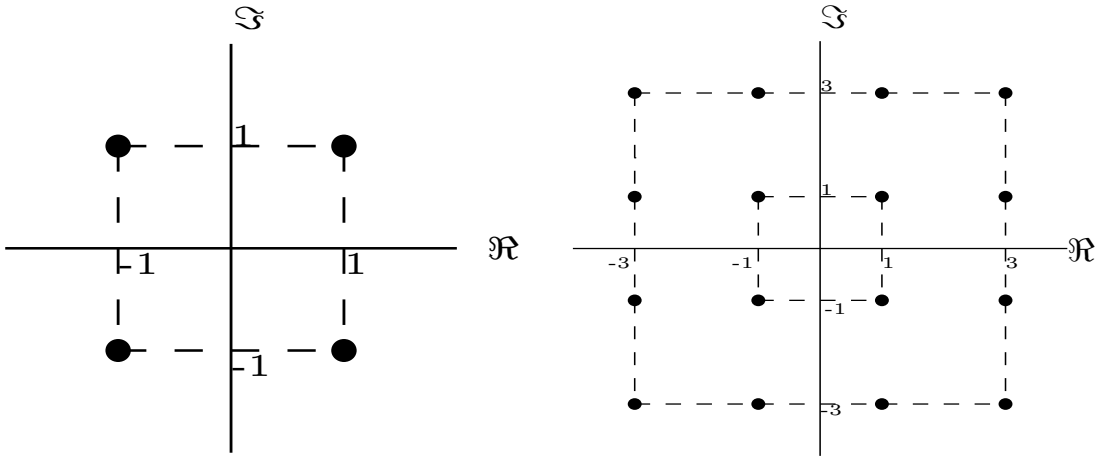
For more advanced forms of OFDM/OQAM modulations, such as M-QAM with $M \geq 16$ where $a_{m,n} \in \{\pm 1, \pm 3, \ldots, \pm(\sqrt{M} - 1)\}$, the constant γ in (3.9) can be set as [36, 37]

$$\gamma = \frac{E(|a_{m,n}|^4)}{E(|a_{m,n}|^2)} \tag{3.33}$$

However, in such situations, modifications can be made to the CM cost function (3.9) that leads to better estimation performance; this is further discussed as the multi-modulus case in Section 3.3.

3.3 Multi-modulus Algorithms (MMA)

In OFDM/OQAM systems, each subcarrier is modulated with a staggered offset QAM symbol, i.e., the real and imaginary parts of the QAM symbols are transmitted separately with half symbol duration. For illustrative purposes, Fig. 3.1 displays the 4QAM and 16QAM symbol constellations. In the case of the 4QAM constellation, after separation,



(a) 4QAM (b) 16QAM

Fig. 3.1 QAM symbol constellations

the transmitted real symbols $a_{m,n} \in \{-1, 1\}$; hence, $a_{m,n}$ satisfies the constant modulus property and γ in (3.9) is equal to 1.

However, when it comes to the 16QAM constellation, the new symbols $a_{m,n} \in \{-3, -1, 1, 3\}$ do not have a unique modulus, a situation referred to as multi-modulus (MM). For higher-order constellations, such as 64QAM, the number of permissible values for $a_{m,n}$ is even larger and there will be more variations in the modulus of $a_{m,n}$. As will be demonstrated in Chapter 4, the adaptive CMA proposed in Section 3.2 work well for the 4QAM constellation, since the recovered symbols satisfy the CM condition. However, when the transmitted symbols have MM, the estimation error obtained by applying these algorithms, which are based on the CM criterion (3.9), does not reach zero even if the channel is perfectly equalized in a noise free environment. In fact, for MM constellations, use of the CM criterion leads to an increased misadjustment and unsatisfactory performance in the corresponding adaptive CMA. This situation has been known for many years in the blind equalization literature, where the use of CMA for MM signals can only achieve a moder-

ate level of steady-state error, leading to large bit error rate (BER) [38]. Moreover, the convergence speed is another drawback of CMA for this kind of signals.

A common solution to improve the equalization performance in this case is to switch to a decision directed (DD) mode of operation [38, 39], but experiments show that this technique may not lead to a satisfactory performance in our applications. Alternatively, in [40], a simple CMA which makes full use of the property of 16QAM symbol constellation is studied where a simple transformation is applied to the demodulated signal to limit the possible values of the modulus. Here, we focus on the application of this concept for the 16QAM constellation, as it is often used in modern communication systems. Specifically, based on the theory in [40], we extend the three CMA developed in Section 3.2 into a new series, referred to as multi-modulus algorithms (MMA).

3.3.1 MMA-LMS-Grad

In [40], the authors simply adjust the coordinates of the 16QAM symbols by applying to them a surjective transformation, that elegantly transforms the corresponding MM problem with modulus values in $\{1, 3\}$ into a CM problem with unit modulus. Specifically, the following transformation is applied

$$a_{m,n} \rightarrow \bar{a}_{m,n} = a_{m,n} - 2\,\mathrm{sign}(a_{m,n}) \tag{3.34}$$

where the function $\mathrm{sign}(x) = 1$ for $x \geq 0$ and -1 otherwise. This transformation maps the values $-3, -1, +1, +3$ onto $-1, +1, -1, +1$, respectively, thereby turning the MM problem into a CM one with $|\bar{a}_{m,n}| = 1$.

In (3.9), we attempts to minimize the dispersion of the modulus of the equalizer output $\mathbf{w}^T\mathbf{y}_{m,n}$ away from $|a_{m,n}|$. Here, since $\bar{a}_{m,n}$ satisfies the CM criterion, we can minimize the dispersion of the modulus of transformed equalizer output, i.e., $\mathbf{w}^T\mathbf{y}_{m,n} - 2\,\mathrm{sign}(\mathbf{w}^T\mathbf{y}_{m,n})$ away from $|\bar{a}_{m,n}|$ ($|\bar{a}_{m,n}| = 1$ in this case). The resulting new cost function is given by

$$\min_{\mathbf{w}} E[(\mathbf{w}^T\mathbf{y}_{m,n} - 2\,\mathrm{sign}(\mathbf{w}^T\mathbf{y}_{m,n}))^2 - 1]^2 \tag{3.35}$$

In Section 3.2.1, the gradient of the instantaneous error function enclosed in (3.9) was used as the search direction in a SGD scheme. Here, we can proceed in the same way, but since a different error function is employed, the corresponding gradient formula has to be modified

as well. In this regard, we note that the sign function $\text{sign}(x)$ is not differentiable at $x = 0$ which might pose a difficulty. However, the probability that $\mathbf{w}^T\mathbf{y}_{m,n} = 0$ will be zero in practice when the optimal solution $\mathbf{w}_{m,n} \neq 0$ due to the fact that $|\mathbf{w}_{m,n}^T\mathbf{y}_{m,n}|$ has to be very close to 1 or 3 to minimize the cost function. Hence, in computing the gradient, we ignore the discontinuity of $\text{sign}(x)$ and sets its derivative to 0. The resulting algorithm, referred to as MMA-LMS-Grad, is summarized in 3.3.1 for the m-th subband, where $\mu > 0$ is the adaptation step-size.

Algorithm 3.3.1 MMA-LMS-Grad

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T$

for $n = 1, 2, \ldots$

$z_{m,n} = \mathbf{w}_{m,n-1}^T\mathbf{y}_{m,n}$

$\bar{z}_{m,n} = z_{m,n} - 2\,\text{sign}(z_{m,n})$

$\mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} - \mu\mathbf{y}_{m,n}\bar{z}_{m,n}(\bar{z}_{m,n}^2 - 1)$

$\left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$

end

3.3.2 MMA-LMS-GN

We now consider the derivation of a Gauss-Newton type LMS algorithm for the MM situation. The procedure is essentially identical to that presented in Section 3.2.2, except for some required modifications in the computation of the Jacobian matrix (i.e. gradient vector). Specifically, based on the new cost function (3.35), the non-linear term enclosed in (3.23) changes to a relatively complex one, thus the corresponding Jacobian matrix $\mathbf{J}(\mathbf{w}_{m,n-1})$ at a local point $\mathbf{w}_{m,n-1}$ should be different from that in (3.24). Algorithm 3.3.2 given below specifies the overall update procedure, referred to as the MMA-LMS-GN algorithm.

Algorithm 3.3.2 MMA-LMS-GN

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T$

1: **for** $n = 1, 2, \ldots$
2: $z_{m,n} = \mathbf{w}_{m,n-1}^T \mathbf{y}_{m,n}$
3: $\bar{z}_{m,n} = z_{m,n} - 2\text{sign}(z_{m,n})$
4: $\mathbf{J}_{m,n} = 2\bar{z}_{m,n}\mathbf{y}_{m,n}^T$
5: $\mathbf{p}_{m,n} = \frac{1-\bar{z}_{m,n}^2}{\mathbf{J}_{m,n}\mathbf{J}_{m,n}^T}\mathbf{J}_{m,n}^T$
6: $\mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} + \beta\mathbf{p}_{m,n}$
7: $\left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$
8: **end**

3.3.3 MMA-RLS

To derive an RLS-based algorithm for the MM case, we first rewrite the new cost function (3.35) as an exponentially weighted time average cost function,

$$C_m(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i}[(\mathbf{w}^T\mathbf{y}_{m,i} - 2\,\text{sign}(\mathbf{w}^T\mathbf{y}_{m,i}))^2 - 1]^2 \tag{3.36}$$

Notwithstanding the presence of the sign(), the above cost function is non-quadratic in the weight vector. To overcome this first difficulty, we can adopt a similar approximation as in (3.31) and use the modified cost function

$$C'_m(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i}[(\mathbf{w}^T\mathbf{y}_{m,i} - 2\,\text{sign}(\mathbf{w}^T\mathbf{y}_{m,i}))\bar{z}_{m,i} - 1]^2 \tag{3.37}$$

where we define

$$z_{m,i} = \mathbf{w}_{m,i-1}^T \mathbf{y}_{m,i} \tag{3.38}$$

$$\bar{z}_{m,i} = z_{m,i} - 2\,\text{sign}(z_{m,i}) \tag{3.39}$$

which are available from previous iterations. By expanding the square term in (3.37), we obtain

$$\begin{aligned} C'_m(\mathbf{w}) = &\sum_{i=1}^{n} \lambda^{n-i}(\mathbf{w}^T\mathbf{y}_{m,i}\,\bar{z}_{m,i} - 1)^2 \\ &- 4\sum_{i=1}^{n} \lambda^{n-i}(\mathbf{w}^T\mathbf{y}_{m,i}\bar{z}_{m,i} - 1)\,\bar{z}_{m,i}\,\mathrm{sign}(\mathbf{w}^T\mathbf{y}_{m,i}) \\ &+ 4\sum_{i=1}^{n} \lambda^{n-i}\bar{z}_{m,i}^2 \end{aligned} \tag{3.40}$$

Upon examination of this expression, we note that the last term does not depend on the unknown weight vector $\mathbf{w}$, while the second term averages to zero in practice due to the random nature of the product $\bar{z}_{m,i}\,\mathrm{sign}(\mathbf{w}^T\mathbf{y}_{m,i})$. For instance, near convergence, this product takes values near ± 1 with equal probability. Hence, we are left with the optimization of the first term, which corresponds exactly to the modified problem in (3.31) for which the RLS algorithm can be applied directly. The resulting algorithm, referred to as the MMA-RLS, is summarized in 3.3.3. It is for the most part identical to the CMA-RLS in 3.2.3, except that the product $z_{m,n} = \mathbf{w}_{m,n-1}^T\mathbf{y}_{m,n}$ is now modified as $\bar{z}_{m,n} = z_{m,n} - 2\mathrm{sign}(z_{m,n})$ to meet the CM criterion as per (3.34).

Algorithm 3.3.3 MMA-RLS

Initialization: $\mathbf{w}_{m,0} = [1, 0]^T, \mathbf{P}_{m,0} = \delta^{-1}\mathbf{I}, \delta = \text{small} \quad \text{positive} \quad \text{constant}$

for $n = 1, 2, \ldots$

$z_{m,n} = \mathbf{w}_{m,n-1}^T\mathbf{y}_{m,n}$

$\bar{z}_{m,n} = z_{m,n} - 2\,\mathrm{sign}(z_{m,n})$

$\mathbf{z}_{m,n} = \mathbf{y}_{m,n}\bar{z}_{m,n}$

$\mathbf{q}_{m,n} = \mathbf{P}_{m,n-1}\mathbf{z}_{m,n}$

$\mathbf{g}_{m,n} = \mathbf{q}_{m,n}/(\lambda + \mathbf{z}_{m,n}^T\mathbf{q}_{m,n})$

$\mathbf{P}_{m,n} = (\mathbf{P}_{m,n-1} - \mathbf{g}_{m,n}\mathbf{z}_{m,n}^T\mathbf{P}_{m,n-1})/\lambda$

$e_{m,n} = 1 - \bar{z}_{m,n}^2$

$\mathbf{w}_{m,n} = \mathbf{w}_{m,n-1} + \mathbf{g}_{m,n}e_{m,n}$

$\left[H_{m,n}^{\Re}, H_{m,n}^{\Im}\right]^T = \epsilon_m \frac{\mathbf{w}_{m,n}}{\|\mathbf{w}_{m,n}\|^2}$

end

3.4 Combination with Frequency Averaging Technique

In this section, we introduce a simple frequency averaging technique which takes advantage of the coherence bandwidth to improve the channel estimation accuracy. We then explain how to combine this technique with the block based SCM and Adaptive SCM algorithms presented in the previous sections.

3.4.1 Block based SCM with Frequency Averaging

In Section 2.3, under the assumption that the channel remains nearly constant over a data block, the quantities $P_{m,n}$ and $\mathbf{C}_{m,n}$ were estimated by time averaging in (2.24) and (2.25), to approximate the expectation operator in (2.17) and (2.22), respectively. Statistically, by using a larger value of N (i.e., additional realizations), the quality of the above estimates is improved, which in turn reduces the estimation errors in the CFR. In practice, however, N cannot be increased indefinitely since this will eventually violate the assumption of a constant channel over a block with size limitation. In this section, to overcome this limitation, we introduce a frequency averaging technique that makes full use of the coherence bandwidth to improve the channel estimation accuracy.

In (2.10), we assumed that the channel coefficients $H_{p,q}$ were nearly constants within the neighbourhood $\Omega_{m,n}$. If we restrict the definition of $\Omega_{m,n}$ as an immediate neighbourhood, i.e. $\Omega_{m,n} = \{(m-1,n),(m+1,n)\}$, then condition (2.10) becomes

$$H_{m\pm1,n} \approx H_{m,n} \tag{3.41}$$

This suggests that we can make use of the received signals in adjacent subbands, i.e. $m-1$ and $m+1$, to provide additional realizations for the purpose of averaging, without increasing N. Specifically, under (3.41), the received signals in the $(m-1)$-th and $(m+1)$-th subbands can also contribute to the estimates of $P_{m,n}$ and $\mathbf{C}_{m,n}$ since in theory, they will result in almost the same channel estimate as defined in (3.41). Accordingly, (2.24) and (2.25) can be replaced by:

$$\hat{P}_m = \frac{1}{3N}\sum_{n=1}^{N}(|Y_{m,n}|^2 + |Y_{m-1,n}|^2 + |Y_{m+1,n}|^2) \tag{3.42}$$

and

$$\hat{\mathbf{C}}_m = \frac{1}{3N}\sum_{n=1}^{N}(\tilde{\mathbf{y}}_{m,n}\tilde{\mathbf{y}}_{m,n}^T + \tilde{\mathbf{y}}_{m-1,n}\tilde{\mathbf{y}}_{m-1,n}^T + \tilde{\mathbf{y}}_{m+1,n}\tilde{\mathbf{y}}_{m+1,n}^T) \tag{3.43}$$

More generally, let us assume that the coherence bandwidth is approximately equal to $2kf_0$, where f_0 is the subcarrier spacing (see equation (2.2)) and k is an integer. Since $H_{m,n}$ is nearly constant within a coherence bandwidth, the neighbourhood $\Omega_{m,n}$ can be extended to cover a corresponding range, which is represented as

$$H_{m\pm k,n} \approx H_{m,n} \quad \text{for small positive integer} \quad k \tag{3.44}$$

where k is the "distance" to the center of subband m. Furthermore, while in (3.42) and (3.43), we assign equal weights to the different subbands, this needs not be the case in general. In practice, for a specific m, the central subband m might account for the largest weight while neighbouring subbands contribute less and less as we move away from this central point. On this basis, we propose following time-frequency averaging:

$$\begin{aligned}\hat{P}_m &= \frac{1}{N}\sum_{l=-k}^{k}\sum_{n=1}^{N} w_l\,|Y_{m-l,n}|^2 \\ \hat{\mathbf{C}}_m &= \frac{1}{N}\sum_{l=-k}^{k}\sum_{n=1}^{N} w_l\,\tilde{\mathbf{y}}_{m-l,n}\tilde{\mathbf{y}}_{m-l,n}^T\end{aligned} \tag{3.45}$$

where w_l is a positive weight factor used for subband $m-l$, and required to satisfy

$$\sum_{l=-k}^{k} w_l = 1. \tag{3.46}$$

Note that (3.42) and (3.43) correspond to a special case of this scheme with $k = 1$ and $w_l = \frac{1}{3}$ for all possible value of $l \in \{0, \pm 1\}$, so that in effect the weight factor w_l performs as a rectangular window. However, in practice, we can use instead any of the symmetric windows available from the signal processing literature [22], such as the triangular window, Hamming windows, etc.

3.4.2 Adaptive SCM with Frequency Averaging

The discussions above emphasize the use of time-frequency averaging by applying the weight factor w_l to the neighbouring subbands. For adaptive algorithms, time-averaging is not suitable because we seek for an instantaneous estimate at each time n; however, frequency averaging can still be useful to improve the estimation accuracy. Specifically, at each time instance n, the new sample to be processed by the adaptive algorithm, i.e., $\mathbf{y}_{m,n}$, is replaced by a weighted combination of all the samples in neighbouring subbands, i.e., $\{\mathbf{y}_{m-k,n}, \ldots, \mathbf{y}_{m+k,n}\}$.

In this case, we can easily combine frequency averaging with the Adaptive SCM proposed in Section 3.1. The resulting procedure is in essence identical to that presented in 3.1.1, except for certain required modifications in the update of quantities $P_{m,n}$ and $\mathbf{C}_{m,n}$, which are now implemented as follows:

$$\begin{aligned} P_{m,n} &= \alpha P_{m,n-1} + (1-\alpha) \sum_{l=-k}^{k} w_l |Y_{m-l,n}|^2 \\ \mathbf{C}_{m,n} &= \alpha \mathbf{C}_{m,n-1} + (1-\alpha) \sum_{l=-k}^{k} w_l \tilde{\mathbf{y}}_{m-l,n} \tilde{\mathbf{y}}_{m-l,n}^T \end{aligned} \tag{3.47}$$

where w_l is a positive weight factor used for subband $m-l$.

In Chapter 4, we will show that the Adaptive SCM developed earlier can lead to performance improvement when combined with the proposed frequency averaging technique.

3.5 Chapter Summary

In this chapter, we proposed four adaptive algorithms for application to semi-blind channel estimation in OFDM/OQAM systems. The Adaptive SCM algorithm uses exponential smoothing in the estimation of certain required quantities. Based on the CM cost function, three CMA were also proposed to search for the optimal solution, thereby leading to the desired channel estimation. These three CMA were then extended to MMA to accommodate higher order symbol constellations. Finally, we combined the block based SCM and Adaptive SCM with a frequency averaging technique which can take advantage of the coherence bandwidth.

Chapter 4

Simulation Results and Discussion

In this chapter, we first introduce the configuration of the OFDM/OQAM system under evaluation along with the methodology to be used in the simulations. Next, we present the simulation results for the block based SCM estimator and the various adaptive algorithms developed in Chapter 3. On the basis of the results so obtained, we can characterize the performance of these newly developed algorithms.

4.1 System Configurations and Methodology

In the simulated OFDM/OQAM system, the bandwidth is set to $B = 10$MHz, which leads to a sampling rate of T_s=100ns. The total number of subcarriers for data transmission is set to $M = 128$; accordingly, the subcarrier spacing is $f_0 = 78.125$kHz and the (vector) symbol duration is $T_0 = 1/f_0 = 12.8\mu$s. Referring to (2.2), a root raised cosine (RRC) filter with a roll-off factor of 1 is chosen as the prototype filter $g(t)$ to implement both the SFB and the AFB filter banks. Indeed, it is shown in [41] that this type of filters is well localized in the time-frequency domain. The filter support length is $T_p = KT_0$, where here, $K = 4$ is the overlapping factor. The latter represents the number of symbol durations in the prototype filter length. Figure 4.1 displays the shape of the prototype filter used in the simulations. Unless otherwise indicated, we assume that offset-4QAM symbols are transmitted using the above OFDM/OQAM system.

In the simulations, the underlying multipath channel is assumed to follow the Extended Pedestrian A (EPA) channel model [42]. The EPA model characterizes the radio channel as a 7-tap finite impulse response (FIR) filter to which is assigned a corresponding profile,

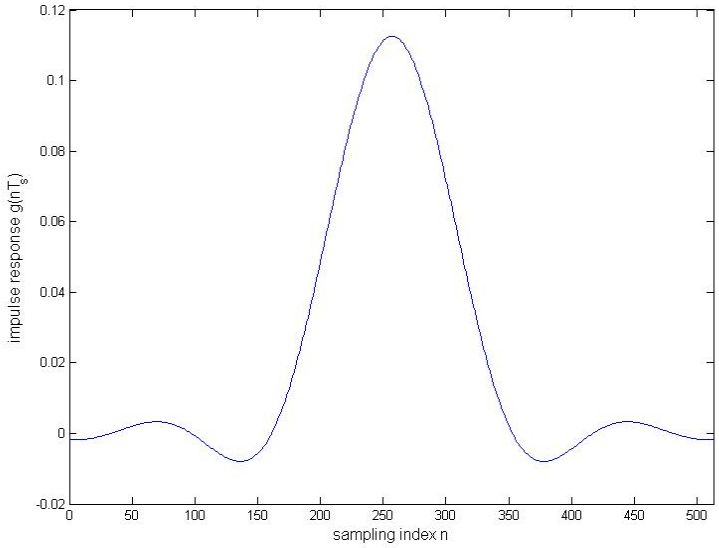


Fig. 4.1 Root raised cosine prototype filter impulse response $g(nT_s)$ versus sampling index n (here, $T_s = 100$ns)

e.g. excess tap delay and relative power. The complex gain associated with each tap is randomly generated according to a complex circular Gaussian distribution, so that the channel magnitude follows the Rayleigh distribution with uniformly distributed phase. For the simulations with time-invariant channels, the tap values generated in this way remain constant during the operation. For the simulations with time-varying channels, the tap delays remain fixed as above but the amplitude changes using a given model, as will be further explained in Section 4.3.3. As per (2.3), the transmitted signal $x(t)$ at the output of the OFDM/OQAM modulator is convolved with the channel impulse response, $h(\tau, t)$, to which white Gaussian noise $w(t)$ with power level σ_w^2 is added. In the simulations, this continuous-time model is implemented in the discrete-time domain with uniform sampling at the rate T_s.

In the simulations, the received data $y(nT_s)$ is processed by the different channel estimation algorithms for the purpose of evaluation and comparison. When applying the block based SCM algorithm from Chapter 2, the block size is chosen as $N = 80$ unless otherwise indicated. For the adaptive algorithms, i.e. the Adaptive SCM, the three adaptive CMA algorithms, and their MM extensions, different choices of parameters (i.e. α, μ, β and λ) are considered as indicated later. A selection of these algorithms is evaluated for each case

of time-invariant and time-varying channels.

Different metrics are used to evaluate and compare the performance of the different algorithms. In particular, to evaluate the accuracy of the channel estimation, we use the normalized mean square error (NMSE) between the true and estimated channels,

$$\text{NMSE} = E\big[\frac{\|\hat{\mathbf{h}} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2}\big] \tag{4.1}$$

where $\hat{\mathbf{h}}$ is the estimated CFR vector and $\mathbf{h}$ is the true CFR vector. In the Monte Carlo simulations, the expectation is evaluated as an average over multiple independent runs, on the order of 500. We also study the impact of the channel estimation on the equalization process. To this end, we evaluate the bit error rate (BER) between the transmitted and decoded bit streams after equalization, i.e., between the binary sequences $s(t)$ and $\hat{s}(t)$ in Fig. 2.2. The corresponding formula is given below:

$$\text{BER} = \frac{\text{number of bit errors}}{\text{total number of transmitted bits}} \tag{4.2}$$

where a large number of transmitted bits is considered, on the order of 10^4 or more.

Finally, the computational costs of the proposed adaptive algorithms, in terms of the required number of floating point operations (flops) per iteration per tone, will also be briefly examined.

4.2 Block based SCM

The purpose of this short section is to demonstrate the efficacy and limitations of the block based SCM [20] algorithm, which was presented in Chapter 2.

Here, we consider the time-invariant channel model, where the underlying wireless channel remains constant during a block of N consecutive symbols. We first present results for the case of $N = 80$ symbols, with the number of subcarriers set to $M = 128$ in the OFDM/OQAM system. Fig. 4.2 illustrates the NMSE of the channel estimator versus SNR [1], where the two curves represent the performance of the initial estimate $\check{\mathbf{h}}_{block}$ and the enhanced estimate $\hat{\mathbf{h}}_{block}$ respectively, as defined in (2.26). It can be seen that the

[1] In the simulations, SNR=$\frac{E|x(t)|^2}{\sigma_w^2}$, as per (2.1) and (2.3)

enhanced estimate can achieve a much lower value of the NMSE by exploiting *a priori* knowledge about the channel length. The error floor for both cases at high SNR is due to various deviations from the modelling assumptions, as discussed in Chapter 2. These results are consistent with those presented in the original paper [20].

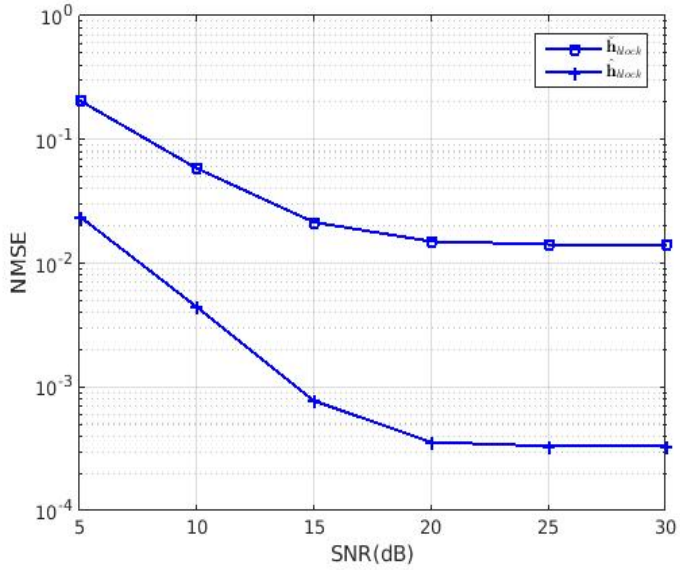


Fig. 4.2 NMSE versus SNR for block based SCM estimator with and without enhancement ($M = 128, N = 80$)

Fig. 4.3 shows the BER performance for equalized offset-4QAM modulation as a function of SNR, where the equalizer is obtained from the estimated channel coefficients. The results, also consistent with those in [20], show that the enhanced channel estimate outperforms the initial one by nearly 2dB. That is, it makes it possible to obtain the same BER while reducing the transmit power by 2dB.

To illustrate the limitations of the block based SCM algorithm, i.e., degradation of the estimation performance when M or N is too small, we tested two alternative configurations of the system model, with parameters as given in Table 4.1. The block based SCM algorithm was then applied again for these two additional configurations. Since the enhanced channel estimate $\hat{\mathbf{h}}_{block}$ in (2.26) always outperforms the initial estimate $\check{\mathbf{h}}_{block}$, we only show the result for $\hat{\mathbf{h}}_{block}$ in Fig. 4.4. It is observed from this figure that when M or N is reduced to a small value, i.e., Configurations 2 or 3, respectively, the accuracy of the SCM channel estimate deteriorates significantly.

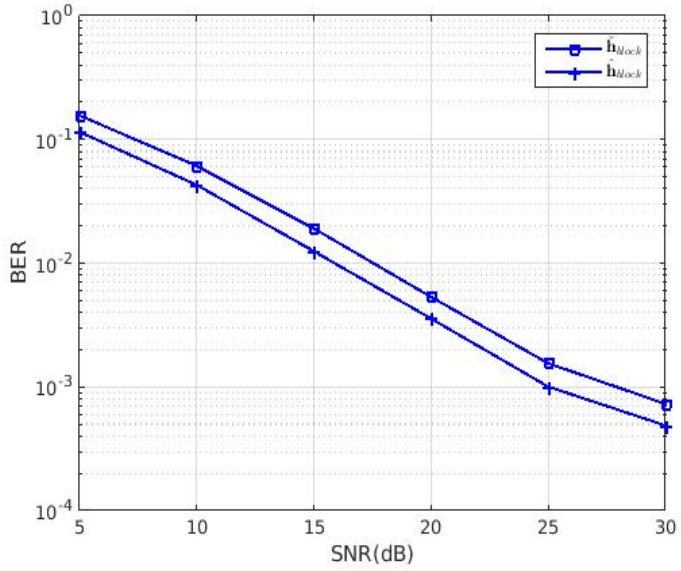


Fig. 4.3 BER versus SNR for SCM-equalized OFDM/OQAM systems ($M = 128, N = 80$)

Configuration	Parameters
1	$M = 128, N = 80$
2	$M = 32, N = 80$
3	$M = 128, N = 40$

Table 4.1 Different system configurations for evaluation of SCM estimator

In the sequel, unless otherwise indicated, all the simulation experiments are based on an OFDM/OQAM system configuration with $M = 128$ subcarriers. Also, when comparing the performance of different adaptive estimation algorithms, we only show results for the enhanced estimate $\hat{\mathbf{h}}_{block}$ in the block based SCM algorithm or the instantaneous estimate $\hat{\mathbf{h}}_n$ in the adaptive algorithms, as was defined in (3.4).

4.3 Adaptive Algorithms for Semi-Blind Channel Estimation

In this section, we investigate the performance of the four adaptive algorithms proposed in Chapter 3 for channel estimation in OFDM/OQAM systems, namely: Adaptive SCM, CMA-LMS-Grad, CMA-LMS-GN and CMA-RLS. We first discuss the resolution of the ambiguity parameter ϵ_m in these algorithms. Then we present simulation results for both

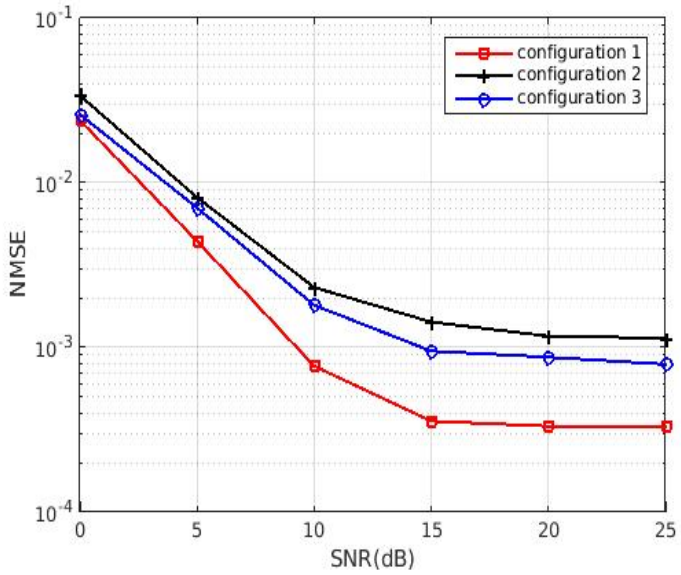


Fig. 4.4 NMSE versus SNR for block based SCM estimator in the case of Configurations 1,2 and 3

cases of time-invariant and time-varying channels. Finally, we briefly discuss the computational cost of implementation of these algorithms. In this section, due to the use of the 4QAM constellation, the parameter γ in (3.9) is set to 1.

4.3.1 Resolution of Sign Ambiguity ϵ_m

The main issue regarding the application of adaptive algorithms proposed in Chapter 3 relates to the determination of the ambiguity parameter $\epsilon_m \in \{-1, +1\}$ for the m-th subband, which appears in (3.3) for the Adaptive SCM and (3.11) for the adaptive CMA-based algorithms. In effect, this sign ambiguity could be time-dependent, as represented by $\epsilon_{m,n}$. If this was the case, then the resolution of the ambiguity would require the use of one pilot per tone per time iteration, thereby rendering the algorithms ineffective.

The purpose of this section is to demonstrate through numerical experiments that for slowly-varying channels, after a few iterations, the ambiguity parameter displays a constant behaviour and hence can be considered time-independent for all practical purposes. To this end, we ran a series of experiments and recorded the correct values of $\epsilon_{m,n}$ over time n, by comparing the estimated channel coefficients obtained with the four adaptive algorithms to the true ones. Illustrative results for a selected subband are shown in Fig. 4.5, but the

results are similar for other subbands. These results were obtained for the case of a slowly time-varying channel, which is generated by:

$$\mathbf{h}_n = 0.9995^n \mathbf{h}_0 \tag{4.3}$$

where $\mathbf{h}_0$ is an initial CFR vector obtained with the EPA model.

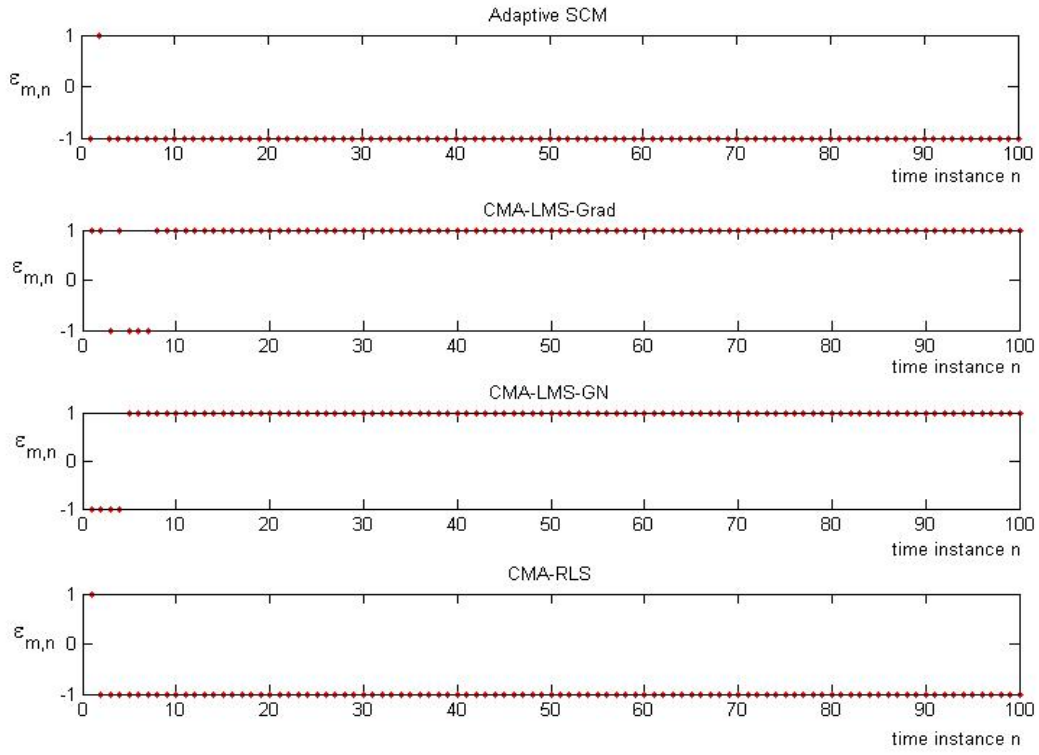


Fig. 4.5 Time evolution of sign ambiguity $\epsilon_{m,n}$ for the different adaptive algorithms (results shown for subband $m = 64$)

From these and similar results, we find that after an initial transient period characterized by oscillations between ± 1 following the onset of adaptation at $n = 0$, the sign ambiguity parameter $\epsilon_{m,n}$ becomes independent of the time n and settle to a constant value (+1 or -1 in Fig. 4.5). Hence, only a few pilot symbols per tone are needed at an early stage of the adaptation process to determine the value of $\epsilon_{m,n}$, which then remains constant for a long stretch of time. This provides the motivation to simplify $\epsilon_{m,n}$ as the time-independent quantity ϵ_m in Chapter 3.

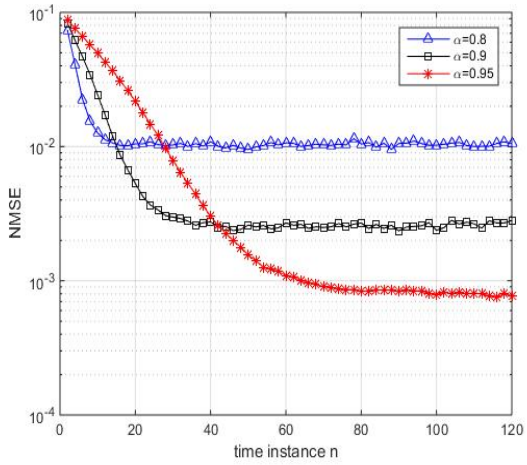
4.3.2 Time-invariant Channel Estimation

When the data is transmitted over a time-invariant channel, the channel estimate obtained with a properly working adaptive algorithm will gradually converge to a steady state value, with the residual NMSE reaching a certain floor level. The convergence speed and steady-state error are therefore two essential factors to consider in the evaluation of an adaptive algorithm.

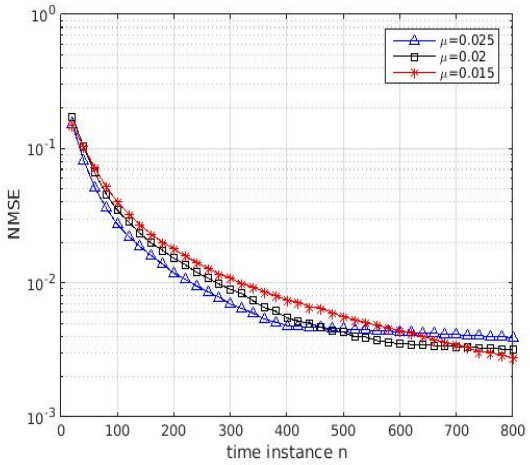
Each one of the adaptive channel estimation algorithms introduced in Chapter 3 contains one parameter, that is: smoothing factor α for Adaptive SCM; step size μ for CMA-LMS-Grad; step size β in CMA-LMS-GN method; and forgetting factor λ for CMA-RLS. It is important to understand how the choice of these parameters will affect the trade-off between convergence speed and steady-state error level in the operation of these algorithms. To this end, we show in Fig. 4.6 the time evolution of the NMSE (4.1) obtained with the above algorithms for different choices of their design parameters. The results are obtained for a time-invariant channel, as explained in Section 4.1, while the SNR is set to 25dB.

Beginning with the Adaptive SCM in Fig. 4.6(a), we note that an increase in the memory parameter α leads to a lower residual error level (i.e. error floor) in steady state by sacrificing the convergence rate. In the case of the CMA-LMS-Grad algorithm in Fig 4.6(b), the step-size μ offers a similar trade-off, i.e., decreasing the step size leads to a reduction of the residual NMSE but slows down the convergence. For the CMA-LMS-GN in Fig. 4.6(c), it appears that a smaller value of β leads to a reduction in the residual NMSE level, without significantly affecting the initial rate of convergence, at least for the considered range of β values. For the CMA-RLS in Fig. 4.6(d), increasing λ, which amounts to a longer exponential window memory, leads to a reduction of the residual NMSE but with a slow down of the initial convergence. We note in passing that among these four algorithms, the worst performance is obtained with the CMA-LMS-Grad, which cannot reach an NMSE floor as low as with the other algorithms.

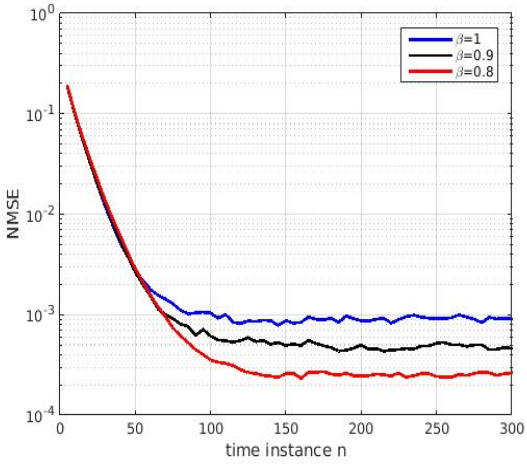
To further compare the convergence rate of the proposed adaptive algorithms, their parameters have to be adjusted to ensure that they all converge to the same residual level. For example, with an SNR=25dB, we find that the Adaptive SCM, CMA-LMS-GN and CMA-RLS algorithms will converge to the same NMSE level, i.e. $\sim 2 \times 10^{-4}$, if we make the following choice of parameters: $\alpha = 0.98$, $\lambda = 0.6$ and $\beta = 0.8$. The corresponding learning curves of these algorithms are shown in Fig. 4.7. From this figure, we observe that



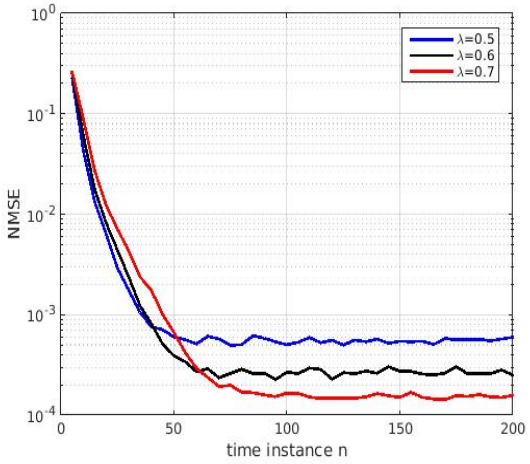
(a) Adaptive SCM



(b) CMA-LMS-Grad



(c) CMA-LMS-GN



(d) CMA-RLS

Fig. 4.6 Time evolution of NMSE for different adaptive algorithms (time-invariant channel, SNR=25dB)

the NMSE of the CMA-RLS and CMA-LMS-GN algorithms drops quickly over time to the desired error floor, while that NMSE of the Adaptive SCM goes down more gradually. Among these three algorithms, the CMA-RLS offers the best performance in convergence, i.e. fastest convergence speed under a constraint of common residual NMSE level in steady state. A representative learning curve for the CMA-LMS-Grad, obtained with $\mu = 0.05$ is also included in this figure. Unfortunately, it is not possible to adjust the step size μ of this algorithm to achieve the desired error floor of 2×10^{-4} within a reasonable time. Instead, its step size has been chosen such that it has the same initial convergence rate as the Adaptive SCM algorithm. In this case, however, it can only reach an NMSE floor of $\sim 10^{-2}$.

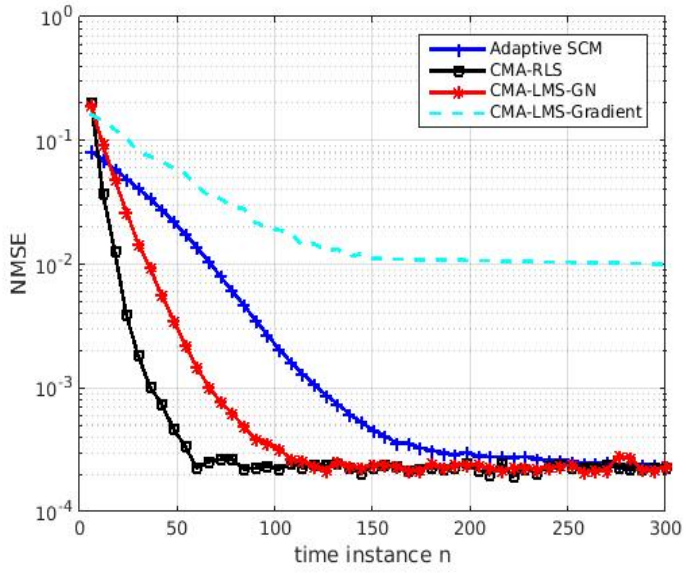


Fig. 4.7 Time evolution of NMSE for four adaptive algorithms (time-invariant channel, SNR=25dB, $\alpha = 0.98$, $\mu = 0.05$, $\beta = 0.8$ and $\lambda = 0.6$)

In Fig. 4.8, we show the learning curves of the Adaptive SCM, CMA-LMS-GN and CMA-RLS algorithms, but this time by adjusting their parameters such that their initial rate of convergence (slope of NMSE versus time n) is similar; the corresponding parameter values are $\alpha = 0.96$, $\beta = 0.8$ and $\lambda = 0.75$. Again, the CMA-RLS achieves the best performance, i.e. the lowest residual NMSE level by a significant margin.

For the above choices of parameters, Fig. 4.9 shows the experimental BER in steady-state (after time iteration $n = 100$) as a function of SNR. For each algorithm, the BER

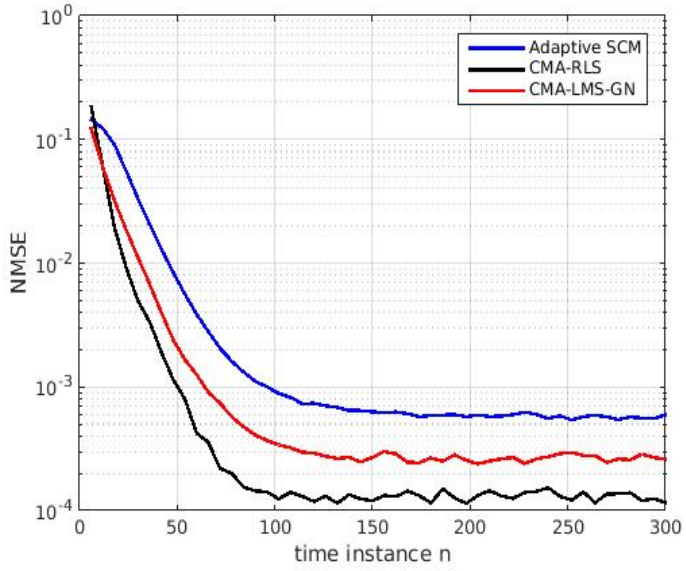


Fig. 4.8 Time evolution of NMSE for three adaptive algorithms (time-invariant channel, SNR=25dB, $\alpha = 0.96$, $\beta = 0.8$ and $\lambda = 0.75$)

was obtained after initial convergence by using the estimated CFR vector $\hat{\mathbf{h}}_n$ at each time n to perform equalization and binary data detection in the OFDM/OQAM receiver. At an SNR of 25dB, for which the above parameter values yield a similar initial convergence rate (hence tracking ability), the CMA-RLS provides the best performance, allowing a 1.25dB gain in this parameter. This shows that the more accurate channel estimation with the CMA-RLS leads to better equalization performance and lower error rate. At lower values of SNR, the Adaptive SCM seems to have a small edge, although in this case the chosen parameter values do not provide identical convergence rate.

Since block based SCM was also applied to the time-invariant channel estimation in Section 4.2, it is interesting to compare the performance of the CMA-RLS to the latter. Specifically, when the SNR = 25dB and $\lambda = 0.6$, the CMA-RLS reaches its steady-state within about 60 iterations. Accordingly, in Fig. 4.10, we show the learning curve of the CMA-RLS (i.e., NMSE versus time n) as compared to the NMSE obtained with the block based SCM when the block length is set to $N = 60$ in (2.24) and (2.25). Clearly, CMA-RLS leads to a more accurate channel estimation in steady-state than the block based SCM in this case.

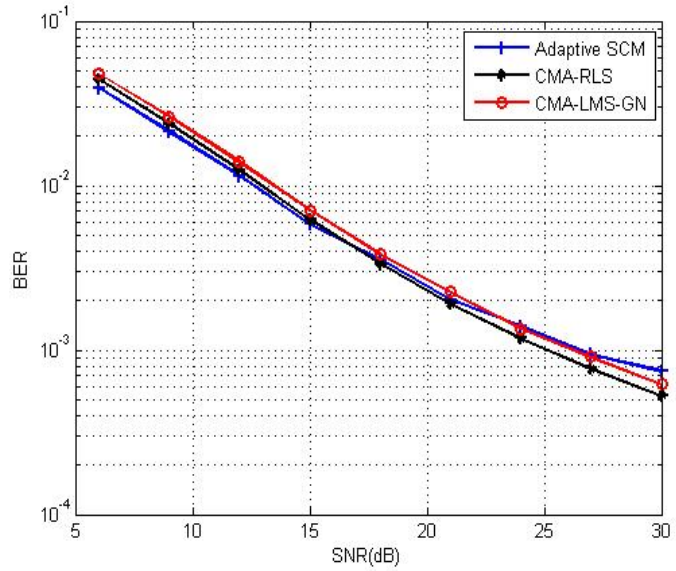


Fig. 4.9 BER versus SNR for three adaptive algorithms (time-invariant channel, $\alpha = 0.96$, $\beta = 0.8$ and $\lambda = 0.75$)

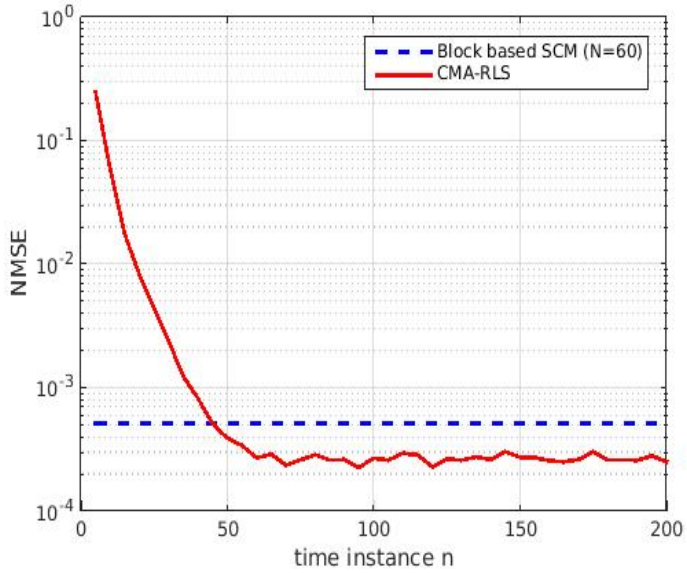


Fig. 4.10 Time evolution of NMSE for CMA-RLS and NMSE for block based SCM estimator (block size N=60, time-invariant channel, SNR=25dB)

4.3.3 Time-varying Channel Tracking

Block based algorithms are mostly useful for the case of time-invariant channels, while they are not well suited to the estimation and tracking of fast time-varying channels, as often

encountered in wireless communications. On the contrary, adaptive algorithms are capable of effective acquisition and tracking of time-varying (non-stationary) channels, as will be demonstrated in this section. We note in passing that among the four adaptive algorithms previously considered, the worst performance is obtained with the CMA-LMS-Grad. Hence, in the subsequent simulations, we only consider Adaptive SCM, CMA-LMS-GN and CMA-RLS for the purpose of performance evaluation.

In Fig. 4.11, we consider the extreme case of a single sudden change in the true channel impulse response. That is, the channel remains constant until time $n = 300$, at which time it undergoes a sudden random change (regeneration of channel coefficients with the EPA model), and then remains constant again until the end of simulation. The figure shows the time evolution of the NMSE obtained with the above three algorithms for the choices of parameters $\alpha = 0.98, \beta = 0.8, \lambda = 0.6$, with SNR=25dB. The results for the initial convergence, i.e. first phase from time $n = 0$ to 300, are similar to those in Fig. 4.7 where the CMA-RLS exhibits a much faster initial convergence rate than the other algorithms and quickly approaches a common error level, i.e., $\sim 2 \times 10^{-4}$. After the sudden change in the true channel, at time $n = 300$, the error level of all three algorithms increases dramatically, which is then followed by a second phase of convergence. It can be seen from Fig. 4.11 that after the sudden change, all three algorithms can track this change but CMA-RLS achieves the best performance, i.e., the fastest acquisition and tracking of the channels. We note that in this case, it takes a much longer time for Adaptive SCM to converge to the same NMSE level.

To further explore the tracking capability of the proposed adaptive algorithms, we carried out another set of experiments in which the underlying channel changes continuously over time by using the following model:

$$\mathbf{h}_n = (1 + 0.25\sin(0.25n))\mathbf{h}_0 \tag{4.4}$$

where $\mathbf{h}_0$ is an initial CFR vector obtained with the EPA model. Fig. 4.12 shows the imaginary part of the true and estimated channel coefficient for the first subband, i.e., $H^{\Im}_{1,n}$ and $\hat{H}^{\Im}_{1,n}$, as a function of time n, but the results for other subbands and the real parts are similar. In this case, the SNR is set to 25dB and the algorithm parameters are adjusted as in Fig. 4.11 (which would yield an identical error floor under stationary conditions). Again the best results are obtained with the CMA-RLS, followed by CMA-LMS-GN and Adaptive

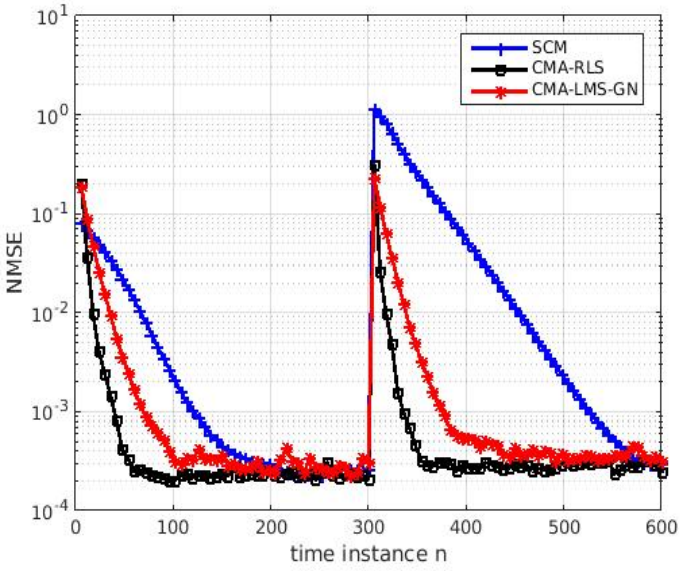


Fig. 4.11 NMSE versus time n for Adaptive SCM, CMA-LMS-GN and CMA-RLS (sudden channel change at $n = 300$, SNR=25dB)

SCM. These results are consistent with those in Fig. 4.11, since here, a slower initial convergence translates into a larger estimation lag (i.e, time delay) [35]. In particular, the CMA-RLS and CMA-LMS-GN can acquire and effectively track the channel with respect to both amplitude and phase, albeit with a small lag; However, the Adaptive SCM is too slow to properly track the true channel, which in addition to a large delay leads to significant amplitude attenuation.

For the above choices of parameters α, β and λ, Fig. 4.13 shows the experimental BER in steady-state as a function of SNR, obtained in the same way as in Fig. 4.9. At SNR=25dB, for which the algorithm parameters were adjusted, the best results are obtained with the CMA-RLS, which provides around 1.25 dB of gain in SNR. This shows that a more precise channel tracking with the CMA-RLS, as shown in Fig. 4.12, leads to better equalization performance and lower error rate.

4.3.4 Computational Cost

The above simulation results show that CMA-RLS provides the best performance, i.e., better trade-off between the convergence speed and residual NMSE level, which in turn leads to more effective tracking of time-varying channels. However, since the CMA-RLS

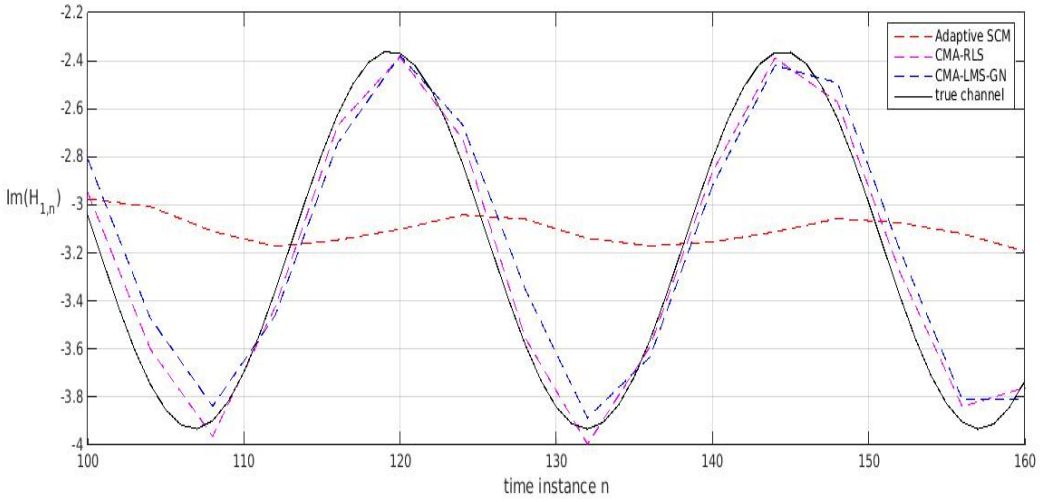


Fig. 4.12 Time evolution of channel estimate $H_{1,n}^{\Im}$ for different adaptive algorithms (fast-varying channel, SNR=25dB)

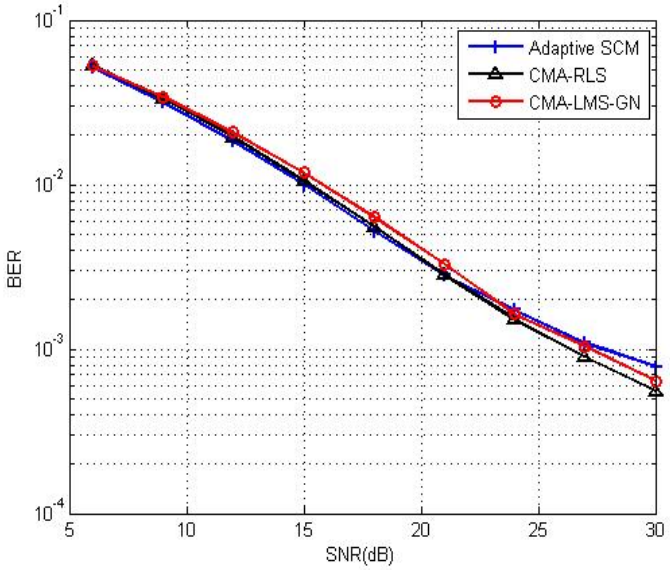


Fig. 4.13 BER versus SNR for three adaptive algorithms (time-varying channel, $\alpha = 0.98$, $\beta = 0.8$ and $\lambda = 0.6$)

is based on a matrix inversion operation [35], its computational complexity is higher than other adaptive algorithms. Table 4.2 summarizes the computational cost, i.e., the required number of floating point operations (flop) per iteration per tone of the four algorithms under study. These figures are obtained by counting the number of flops[2] inside the main loop of the corresponding algorithms in Chapter 3. Indeed, as can be seen from Table 4.2, CMA-RLS achieves those aforementioned advantages at the expense of higher computational cost.

Adaptive algorithm	Number of operations
Adaptive SCM	40
CMA-LMS-Grad	16
CMA-LMS-GN	22
CMA-RLS	48

Table 4.2 Computational cost for four adaptive algorithms per iteration per tone

4.4 MMA for Time-invariant Channel Estimation

In this section, we perform simulation experiments to demonstrate the efficacy of the adaptive MMA presented in Section 3.3. In these experiments, 16QAM symbols are transmitted using an OFDM/OQAM system with the same parameter setting as described in Section 4.1. We also present results for the Adaptive SCM, which can be applied regardless of the specific choice of discrete symbol constellation.

Fig. 4.14 illustrates the time evolution of the NMSE obtained with the Adaptive SCM (Algorithm 3.1.1), the MMA-LMS-GN (Algorithm 3.3.2) and the MMA-RLS (Algorithm 3.3.3), with the following choice of parameters: $\alpha = 0.98$, $\lambda = 0.53$, $\beta = 0.7$ and SNR=25dB. From the figure, we note that these three algorithms suffer a degradation in the NMSE performance after convergence, where the NMSE floor is increased from 2×10^{-4} to 3×10^{-3}, compared to 4QAM cases. Nevertheless, the MMA-RLS offers the fastest performance in convergence under a constraint of common residual NMSE level in steady state.

[2] 1 flop = 1 elementary floating point operation: $+, -, \times, \div$

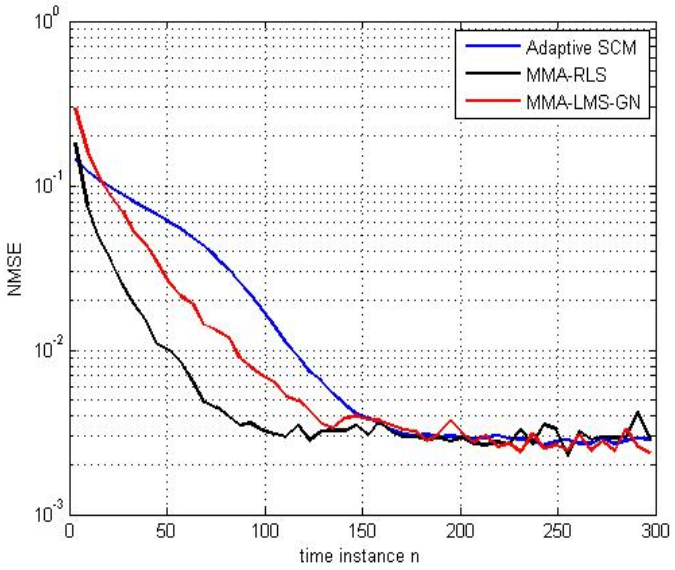


Fig. 4.14 Time evolution of NMSE for Adaptive SCM, MMA-LMS-GN and MMA-RLS (16QAM constellation, time-invariant channel, SNR=25dB)

4.5 Use of Frequency Averaging Technique

In this section, we perform simulation experiments to show that the algorithms developed earlier, including block based SCM and Adaptive SCM, can lead to performance improvements when combined with the simple frequency averaging technique.

For these experiments, we consider time-invariant channels generated randomly according to the EPA model as explained in Section 4.1, where the underlying channel remains constant during each independent run. The magnitude response of a representative channel realization is plotted in Fig. 4.15. From this figure, we can infer that a reasonable value of k in (3.44), which in essence specifies the coherence bandwidth of the channel, is on the order of 2 to 4.

We first present results for the block based SCM algorithm for two different choices of window w_l in the frequency averaging process (3.45), i.e. rectangular and triangular. Fig. 4.16 (a) illustrates the NMSE of the channel estimator versus SNR for the rectangular window, where the three curves represent the performance of $\hat{\mathbf{h}}_{block}$ with $k = 0, 4, 6$ respectively; Fig. 4.16 (b) presents the corresponding results for the case of the symmetric triangular windows.

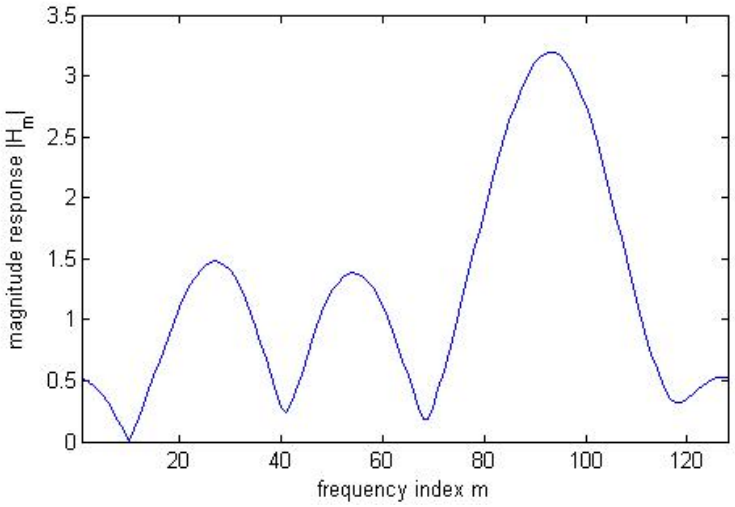


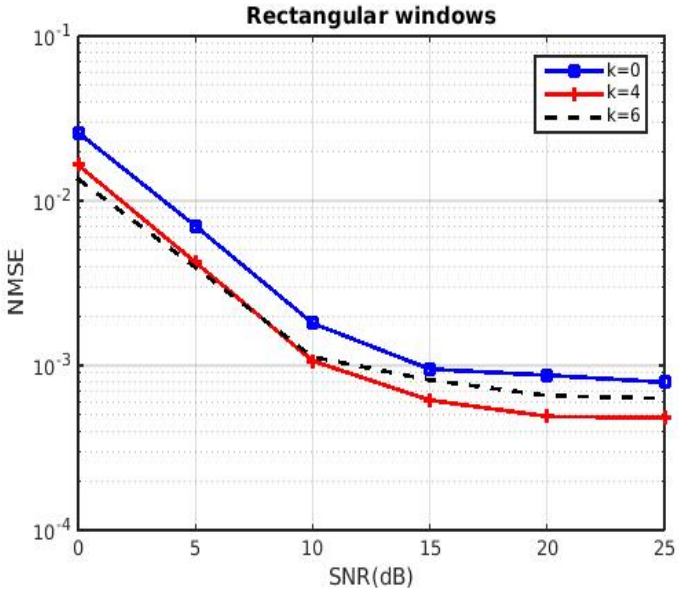
Fig. 4.15 Magnitude response of a representative EPA channel realization

Clearly, for the particular channel model under consideration, an increase in the value of parameter k from 0 to 4 leads to a notable performance enhancement, i.e. optimal reduction of about 2.2db in NMSE. However, increasing k beyond this optimal value does not provide as much improvement. The use of a triangular window, compared to the rectangular one, provides only a minor improvement in performance. In the sequel, we only consider the case of a rectangular window with $k = 4$.

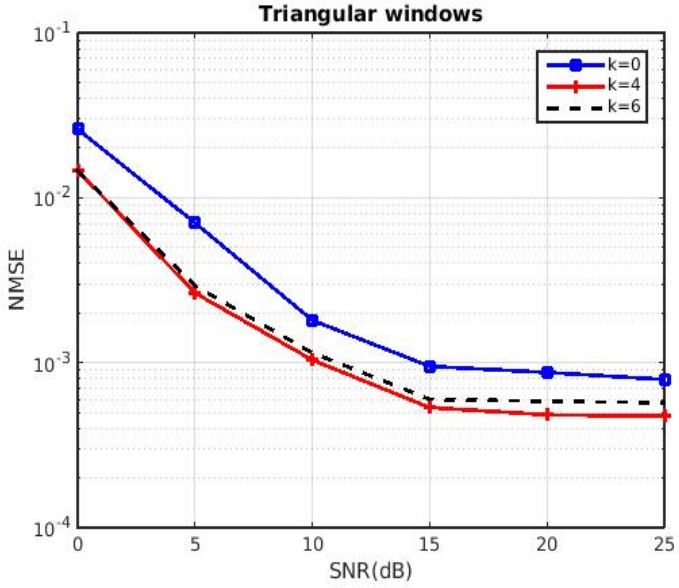
In Section 3.4.2, we briefly explained how the frequency averaging technique could be applied to the Adaptive SCM developed in the thesis. In Fig. 4.17, we show the learning curves of the Adaptive SCM algorithm with and without the use of frequency averaging. In this experiment, SNR=25dB, $k = 4$, $\alpha = 0.94$ and $w_l = \frac{1}{9}$ for all $l \in \{0, \pm 1 \ldots, \pm 4\}$. It can been seen from Fig. 4.17 that the frequency averaging technique can improve estimation performance by increasing the convergence speed initially and lowering the residual NMSE floor in steady state.

4.6 Chapter Summary

In this chapter, numerical simulation experiments have been carried out to evaluate the performance of the proposed algorithms for time-invariant channel estimation and time-varying channel tracking. CMA-RLS proves to have the best performance in terms of convergence



(a) rectangular windows



(b) triangular windows

Fig. 4.16 NMSE versus SNR of block based SCM estimator with rectangular or triangular windows ($k = 0, 4, 6$, $N = 40$, time-invariant channel)

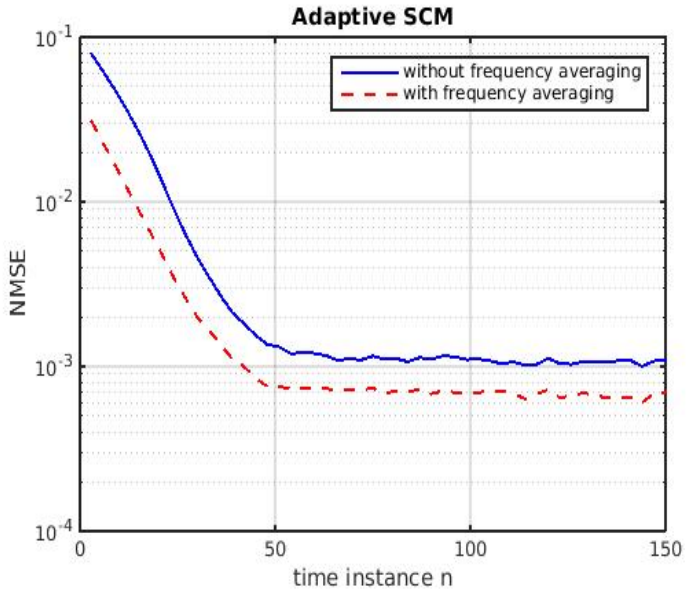


Fig. 4.17 Time evolution of NMSE for Adaptive SCM with and without frequency averaging ($\alpha = 0.94$, rectangular window with $k = 4$, time-invariant channel, SNR=25dB)

speed, steady-state error and tracking capability. However, CMA-RLS sacrifices computational cost to achieve these advantages. Also, the proposed MMA were applied to the case of a 16QAM constellation, where again the MMA-RLS showed the best performance. Finally, simulation results confirm that both the block based SCM and Adaptive SCM can lead to performance improvements under certain conditions when combined with the proposed frequency averaging technique.

Chapter 5

Conclusion and Future Work

This thesis proposed semi-blind adaptive algorithms for the estimation of both time-invariant and time-varying channels in OFDM/OQAM systems.

In Chapter 2, we introduced the fundamentals of OFDM/OQAM systems, including mathematical characterization of the transmitter and receiver sub-systems. Then the recently proposed block based SCM algorithm [20] was reviewed for application to channel estimation and its limitations were indicated.

In Chapter 3, we proposed four adaptive algorithms to overcome these limitations in addressing the semi-blind channel estimation problem for OFDM/OQAM systems. In particular, four novel adaptive algorithms were proposed, namely,

- Adaptive SCM,
- CMA-LMS-Grad,
- CMA-LMS-GN,
- CMA-RLS,

Also, we investigated the extension of these algorithms to the MM constellation as well as the combination with a simple frequency averaging technique that exploits the coherence bandwidth of the channel to improve estimation performance.

In Chapter 4, we presented the simulation results for the block based SCM estimator as well as the new adaptive algorithms developed in Chapter 3. On the basis of the results so obtained, we compared the performance of these newly developed algorithms. The

adaptive algorithms prove to be useful in faster estimation of time-invariant channels and in effective acquisition and tracking of time-varying channels. Among the newly proposed algorithms, CMA-RLS proves to be the best algorithm for our application in the case of the 4QAM constellation. Indeed, CMA-RLS achieves the best trade-off between the convergence speed and steady state error. Simulations with a higher order symbol constellation showed that proposed MMA-RLS gives the best results in the case of a 16QAM constellation. Finally, simulation results showed that the proposed frequency averaging technique can boost the estimation performance to some extent when combined with block based SCM or the Adaptive SCM algorithm.

Some methods discussed in this thesis have limitations and could be investigated further. As mentioned in Section 2.4, the block based SCM is plagued by the limitation that for a given (i.e. fixed) system bandwidth, if the number of subcarriers M is reduced, (2.10) is not satisfied and (2.11) no longer holds. To overcome this issue, it is necessary to replace (2.11) by a more accurate, coupled model in which received data on tone m depends, in addition to $H_{m,n}$, on surrounding values of $H_{p,q}$ in the neighbourhood of (m, n). In turn, the use of a more compact system model for this case could possibly lead to the development of improved estimation techniques that can better cope with such coupling of the subcarrier data.

Application of adaptive channel estimation to systems with higher order symbol constellations, i.e., 16QAM or 64QAM, calls for more effective extensions of existing CMA-based algorithms, in addition to the simple MMA developed in Section 3.3.

While this thesis focused on channel estimation in OFDM/OQAM systems, where one obvious purpose is that of equalization, the use of CMA based algorithms makes it possible to consider channel estimation and equalization jointly. Indeed, if we look at the CMA-LMS-GN algorithm for instance, we note that the quantity $z_{m,n}$ actually provides a non-quantized estimate of the transmitted data $a_{m,n}$. Within this context, it would be interesting to study how the use of a decision-directed mode of operation can help resolve the determination of the ambiguity parameter ϵ_m in the case of more rapidly varying channels.

In this thesis, the channel is estimated in the frequency domain. Each subband is estimated independently because our system model and the corresponding algorithms are single subcarrier based. Alternatively, it might be of interest to study new algorithms to address the channel estimation in the time domain for OFDM/OQAM systems. There currently exist a few approaches to do this, as in e.g. [12], but they tend to be complex and

require the use of training sequences.

References

[1] J. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[2] L. Hanzo, Y. Akhtman, L. Wang, and M. Jiang, *OFDM Standards*, pp. 37–60. Wiley-IEEE Press, 2011.

[3] H. Bölcskei, *Orthogonal frequency division multiplexing based on offset QAM*, pp. 321–352. 2003.

[4] N. Neurohr and M. Schilpp, "Comparison of transmultiplexers for multicarrier modulation," in *Proc. Fourth Inter. Conf. Signal Processing*, (Beijing, China), pp. 35–38 vol.1, Oct. 1998.

[5] S. Kai, Z. Weiwei, and W. Guangyu, "Design and implementation of MDFT filter bank based multicarrier modulation systems," in *IEEE Int. Conf. Signal Processing, Communication and Computing*, (KunMing, China), pp. 1–6, Aug. 2013.

[6] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Processing Magazine*, vol. 28, pp. 92–112, May 2011.

[7] X. Li and J. Ritcey, "Maximum-likelihood estimation of OFDM carrier frequency offset for fading channels," in *Thirty-First Asilomar Conf. Signals, Systems and Computers*, vol. 1, (Pacific Grove, CA, USA), pp. 57–61, Nov. 1997.

[8] M. Hasan, F. Nabita, A. Khandakar, I. Ahmed, and F. Ahmed, "Analytical evaluation of timing offset error in OFDM system," in *Second Int. Conf. Communication Software and Networks*, (Singapore), pp. 3–7, Feb. 2010.

[9] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Trans. Signal Processing*, vol. 50, pp. 1170–1183, May. 2002.

[10] L. Vangelista and N. Laurenti, "Efficient implementations and alternative architectures for OFDM-OQAM systems," *IEEE Trans. Communications*, vol. 49, pp. 664–675, Apr. 2001.

[11] J. Fang, Z. You, I.-T. Lu, J. Li, and R. Yang, "Comparisons of filter bank multicarrier systems," in *IEEE conf. Systems, Applications and Technology*, (Farmingdale, NY), pp. 1–6, May. 2013.

[12] D. Kong, D. Qu, and T. Jiang, "Time domain channel estimation for OQAM-OFDM systems: Algorithms and performance bounds," *IEEE Trans. Signal Processing*, vol. 62, pp. 322–330, Jan. 2014.

[13] D. Katselis, E. Kofidis, A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation for CP-OFDM and OFDM/OQAM systems: A comparative study," *IEEE Trans. Signal Processing*, vol. 58, pp. 2911–2916, May 2010.

[14] C. Lélé, J.-P. Javaudin, R. Legouable, A. Skrzypczak, and P. Siohan, "Channel estimation methods for preamble-based OFDM/OQAM modulations," *European Trans. Telecommunications*, vol. 19, no. 7, pp. 741–750, 2008.

[15] J. Du and S. Signell, "Novel preamble-based channel estimation for OFDM/OQAM systems," in *IEEE Inter. Conf. Communications*, pp. 1–6, June. 2009.

[16] E. Kofidis and D. Katselis, "Preamble-based channel estimation in MIMO-OFDM/OQAM systems," in *IEEE Inter. Conf. Signal and Image Processing Applications*, (Kuala Lumpur), pp. 579–584, Nov. 2011.

[17] C. Lélé, R. Legouable, and P. Siohan, "Channel estimation with scattered pilots in OFDM/OQAM," in *IEEE Proc. Workshop Signal Processing Advances in Wireless Communications*, (Recife), pp. 286–290, July. 2008.

[18] C. Lélé, "Iterative scattered-based channel estimation method for OFDM/OQAM," *EURASIP Journal on Advances in Signal Processing*, no. 1, 2012.

[19] J.-P. Javaudin, D. Lacroix, and A. Rouxel, "Pilot-aided channel estimation for OFDM/OQAM," in *the 57th IEEE semiannual Vehicular Technology Conf.*, vol. 3, pp. 1581–1585, April. 2003.

[20] W. Hou and B. Champagne, "Semiblind channel estimation for OFDM/OQAM systems," *IEEE Signal Processing Letters*, vol. 22, pp. 400–403, April 2015.

[21] T. H. Stitz, T. Ihalainen, A. Viholainen, and M. Renfors, "Pilot-based synchronization and equalization in filter bank multicarrier communications," *EURASIP J. Adv. Signal Process*, vol. 2010, pp. 9:1–9:11, Jan. 2010.

[22] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall Press, 3rd ed., 2009.

[23] T. Cooklev, *The IEEE Standard for WLAN: IEEE 802.11*, pp. 45–132. Wiley IEEE Standards Association, 2004.

[24] A. Molisch, *3GPP Long Term Evolution*, pp. 665–698. Wiley IEEE Press, 2011.

[25] C. Lélé, P. Siohan, R. Legouable, and J.-P. Javaudin, "Preamble-based channel estimation techniques for OFDM/OQAM over the powerline," in *IEEE International Symposium on Power Line Communications and Its Applications, 2007. ISPLC 07*, pp. 59–64, March 2007.

[26] S. Visuri, H. Oja, and V. Koivunen, "Subspace-based direction-of-arrival estimation using nonparametric statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 2060–2073, Sep. 2001.

[27] C. Croux, E. Ollila, and H. Oja, "Sign and rank covariance matrices: Statistical properties and application to principal components analysis," in *Statistical Data Analysis Based on the L1-Norm and Related Methods* (Y. Dodge, ed.), Statistics for Industry and Technology, pp. 257–269, Birkhauser Basel, 2002.

[28] Y. Chen, T. Le-Ngoc, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Trans. Signal Processing*, vol. 52, pp. 1452–1456, May. 2004.

[29] J. Mark and W. Zhuang, *Wireless Communications and Networking*, pp. 139–143. Prentice Hall, 2003.

[30] C. Xu and T. Le-Ngoc, "CMA-based channel estimation for space-time block coding transmission systems," *Wireless Personal Communications*, vol. 28, no. 4, pp. 241–258, 2004.

[31] J. Benesty and P. Duhamel, "Fast constant modulus adaptive algorithm," *IEE Proc. Radar and Signal Processing*, vol. 138, pp. 379–387, Aug. 1991.

[32] V. Zarzoso and P. Comon, "Optimal step-size constant modulus algorithm," *IEEE Trans. Communications*, vol. 56, pp. 10–13, January. 2008.

[33] C. B. Papadias and D. T. M. Slock, "On the convergence of normalized constant modulus algorithms for blind equalization," in *Proc. DSP Inter. Conf. Digital Signal Processing*, pp. 245–250, 1993.

[34] A.Bjorck, *Numerical Methods for Least Squares Problems*, pp. 339–358. SIAM, 1996.

[35] S. Haykin, *Adaptive Filter Theory, 3rd Ed.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

[36] I. Homana, M. Topa, and B. Kirei, "Semi-blind equalization using the constant modulus algorithm," in *IEEE Inter. Conf. Automation Quality and Testing Robotics*, vol. 3, pp. 1–4, May. 2010.

[37] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Communications*, vol. 28, pp. 1867–1875, Nov 1980.

[38] C. A. R. Fernandes, G. Favier, and J. C. M. Mota, "Decision directed adaptive blind equalization based on the constant modulus algorithm.," *Signal, Image and Video Processing*, vol. 1, pp. 333–346, Mar 2007.

[39] T. Endres, S. Hulyalkar, C. Strolle, T. Schaffer, and R. Casas, "A decision-directed constant modulus algorithm for higher-order source constellations," in *IEEE Inter. Conf. Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3382–3385 vol.6, 2000.

[40] W. Rao, K. min Yuan, Y.-C. Guo, and C. Yang, "A simple constant modulus algorithm for blind equalization suitable for 16-QAM signal," in *Inter. Conf. Signal Processing*, pp. 1963–1966, Oct 2008.

[41] J. Du and S. Signell, "Comparison of CP-OFDM and OFDM/OQAM in doubly dispersive channels," in *Future Generation Communication and Networking*, vol. 2, pp. 207–211, Dec. 2007.

[42] S. Sesia, I. Toufik, and M. Baker, *LTE - the UMTS long term evolution : from theory to practice*. Chichester: Wiley, 2009.