2017

Joint CFO Estimation and Data Detection in OFDM systems

Chunyu Mao
University of Windsor

Joint CFO Estimation and Data Detection in OFDM systems

by

Chunyu Mao

A Thesis

Submitted to the Faculty of Graduate Studies

through Electrical and Computer Engineering

in Partial Fulfilment of the Requirements for

the Degree of Master of Applied Science at the

University of Windsor

Windsor, Ontario, Canada

2016

Joint CFO Estimation and Data Detection in OFDM systems

by

Chunyu Mao

APPROVED BY:

Dr. Robin Gras

School of Computer Science

Dr. Esam Adbel-Raheem

Department of Electrical and Computer Engineering

Dr. Behnam Shahrrava, Advisor

Department of Electrical and Computer Engineering

October 4, 2016

AUTHOR'S DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyones copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Orthogonal frequency division multiplexing (OFDM) is a multicarrier modulation technique that is widely used in wireless broadband communication systems. The spectral efficiency of OFDM is very high since the subcarriers are spaced as closely as possible while maintaining orthogonality. However, one of the major problems with OFDM that can cause performance degradation is carrier frequency offset (CFO) which impairs the orthogonality among OFDM subcarriers, as a consequence, results in inter-subcarrier interference.

In this thesis, an iterative algorithm for joint CFO estimation and data detection in OFDM systems over frequency selective channels is proposed. The proposed algorithm is performing both CFO estimation and data detection in the frequency domain based on the Expectation-Maximization (EM) algorithm. The proposed algorithm can achieve the same bit-error-rate (BER) performance as that of its time-domain counterpart with much lower complexity.

Simulation results show that the proposed algorithm can converge after three iterations and an estimate of CFO can be obtained with high accuracy.

DEDICATION

To my parents, Huibin Mao and Yukun Wang, for their selfless contribution to my life.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Behnam Shahrrava, professor of the department of Electrical and Computer Engineering at University of Windsor first. The door to Prof. Shahrrava office was always open whenever I ran into a trouble spot or had a question about my research. He consistently allowed this thesis to be my own work, but steered me in the right direction whenever he thought I needed it.

I would also like to appreciate my thesis committee members Dr. Robin Gras and Dr. Esam Adbel-Raheem. Without their passionate participation and input, the thesis could not have been successfully conducted. I also want to thank Dr. Shervin Erfani for his guidance of my research.

I would also like to acknowledge my uncle Dr. Yingming Zhao and my aunt Li Wang for their support. They encourage me to chase my dream without any heisitate, and I am gratefully indebted to their very valuable comments on my life.

Finally, I must express my very profound gratitude to my parents for providing me with financial support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. I would also express my appreciate to my colleagues and friends, Siyu Zhang, Huyuan Li, Lin Li, Pengzhao Song, Peibo Zhao, Akilan Thangarajah for their support and help. This accomplishment would not have been possible without them. At last, I would like to extend my gratitude to University of Windsor. Thank you.

Chunyu Mao

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AWGN Additive White Gaussian Noise

BER Bit Error Rate

BLUE Best Linear Unbiased Estimator

CDMA Code Division Multiple Access

CFO Carrier Frequency Offset

CIR Channel Impulse Response

CP Cyclic Prefix

CRLB Cramer-Rao Lower Bound

DFT Discrete Fourier Transform

E-step Expectation Step

EM Expectation Maximization

ICI Intercarrier Interference

IDFT Inverse Discrete Fourier Transform

ISI Intersymbol Interference

LMMSE Linear Minimum Mean Squared Error

LS Least Squares

LTI Linear Time Invariant

LTV Linear Time-Varying

M-step Maximization Step

MAP Maximum A Posteriori

MC Multicarrier

ML Maximmum Likelihood

MVU Minimum Variance Unbiased

OFDM Orthogonal Frequency Division Multiplexing

PAPR Peak to Average Power

PDF Probability Density Function

PSK Phase shift Keying

RBLS Rao-Blackwell-Lehmann-Scheffe

TDMA Time Division Multiple Access

WCDMA Wideband Code Division Multiple Access

ZF Zero Forcing

1 Introduction

For the past centuries, people desire to communicate with each other in a convenience and economical way. Due to this demand, communications has been a vibrant field of research for more than hundred years. Thanks to the development of Very Large Scale Integration (VLSI) and Signal Processing technology since the 1960s, wireless communications has become one of the most significant research areas in the field of modern communications, [1]. A tremendous advantage of wireless communication compared with wired communication is that it does not require any physical cables. This superiority not only enables devices to work anywhere without considering the limitation of the wires but also saves the money.

1.1 Wireless Communications and OFDM Technique

1.1.1 The Development of OFDM Technique

In modern society, wireless communication has various excellent applications, such as smart buildings, telecommunication, process industry, etc. As the technology development, several techniques have been developed for wireless communication, such as Time Division Multiple Access (TDMA), Code Division Multiple Access (CDMA), Wideband Code Division Multiple Access (WCDMA), etc. Currently, Orthogonal Frequency Division Multiplexing (OFDM) is the most popular multicarrier band-efficient physical layer technique for wireless communication [2]. Multicarrier (MC) modulation technique was first developed by *Chang* [3] in 1966. Todays OFDM was introduced by *Weistein* and *Ebert* [4] in 1971, which is a special form of MC transmission. In 1985, *Crmini* first applied OFDM in mobile wireless communication [5]. After these years development, OFDM is the most recognized technique in wireless communication. It has various popular applications, such as digital audio broadcasting, digital television, wireless local area networks and the 4th generation of mobile telecommunication.

1.1.2 The Advantages of OFDM Technique

OFDM has several outstanding advantages in compared with other techniques [6]. First of all, frequency domain modulation is easy to deal with the physical channel without complex time domain equalization. Secondly, OFDM can be implemented efficiently by using Fast Fourier Transform (FFT). Moreover, OFDM has not only low sensitivity to time synchronization errors but also easy to integrate with multi-antenna systems. Furthermore, intersymbol interference (ISI) can be simply handled by adding the cyclic prefix (CP). An essential feature of OFDM [7] is that it employs multiple narrowband orthogonal subcarriers. Due to this feature, not only the available spectrum can be used efficiently but also intercarrier interference (ICI) can be completely eliminated. At last, this kind of subcarriers can preserve a high transmission speed. Because of the above advantages, OFDM has become an excellent technique for wireless communication.

The essential characteristic of OFDM system is that it employs orthogonal subcarriers. In a conventional frequency division multiplexing (FDM), to eliminate ICI, the subcarriers are spaced apart by inserting guard bands between them in the frequency domain, as shown in Fig. 1.1a. However, this kind of arraignment for the subcarriers wastes bandwidth. An alternative method is to overlap the subcarriers to save the bandwidth as shown in Fig. 1.1b. But overlapping the subcarriers could introduce ICI. Thus, to save bandwidth without having ICI, orthogonal subcarriers are the best choice. OFDM is the technique that employs orthogonal subcarriers to transmit the data. Figure 1.2 shows that the OFDM subcarriers exhibit orthogonality on a symbol interval if they are spaced in the frequency domain exactly at the reciprocal of the symbol interval, which can be accomplished by utilizing the discrete Fourier transform (DFT).

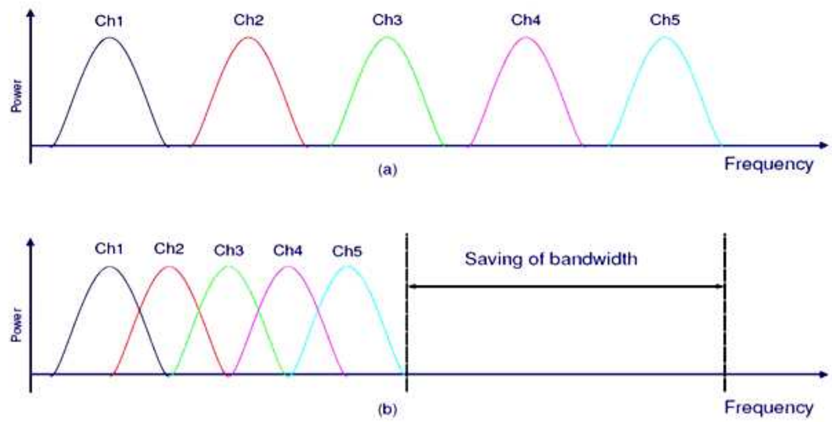

Fig. 1.1: Multicarriers

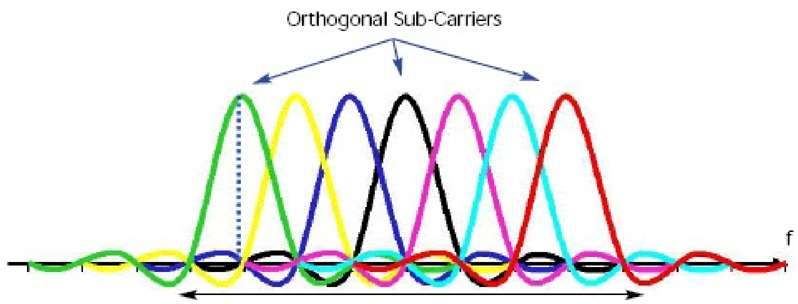

Fig. 1.2: Orthogonal Subcarriers

1.2 Challenges of OFDM

Nonetheless, due to the imperfect channel condition, OFDM still has several limitations that would affect the performance of the system. First, in practice, a wireless channel cannot be considered as an ideal channel [1]. As a result, the received signal is not the same as the transmitted signal. Thus the data need to be recovered according to the actual channel. Also, an OFDM signal can have a large peak-to-average-power ratio (PAPR) value with high probability [8]. Moreover, strict frequency synchronization is impossible to OFDM system [9] which is related to carrier frequency offset (CFO). Furthermore, both CP and null guard tones at the edge of the spectrum can decrease the spectral efficiency of the system. However, this thesis will focus only on reducing the effect of CFO on the performance of a practical OFDM system.

When considering the performance of a wireless communication system, the bit error rate (BER) plays an essential role. The probability of bit error, also called, the BER is a measure of deterioration in digital communications. Lower BER means higher performance. In OFDM systems, due to the characteristics of the physical channel, the BER is highly related to both parameters estimation and data detection. Generally speaking, the parameters contain timing error, channel information and carrier frequency offset (CFO). These parameters are independent of each other. Since the OFDM system is not sensitive to timing error [10], and the channel can be estimated accurately by inserting a set of known symbols, called pilots, [11]. Hence, the performance of the system is slightly affected by these two parameters. However, the CFO cannot be easily estimated by using the pilots. The CFO is the normalized frequency error between the transmitter and the receiver. In practical OFDM systems, the CFO can be caused by the mismatch between the oscillator in the transmitter and the receiver or the Doppler shift in frequency selective channels, [12]. In practical mobile communication systems, the channel is a frequency selective channel. The details of communication channels will be given in Chapter 3. Fig. 1.3 shows how the Doppler shift leads to a frequency error. As seen in, in OFDM systems, the existence of CFO destroys the orthogonality among OFDM subcarriers, as a consequence, results in inter-subcarrier interference [13]. And this inter-subcarrier interference can cause BER performance degradation. As seen from Fig. 1.4, OFDM is very sensitive to CFO [14], even a small CFO in an OFDM system can lead to a huge error.

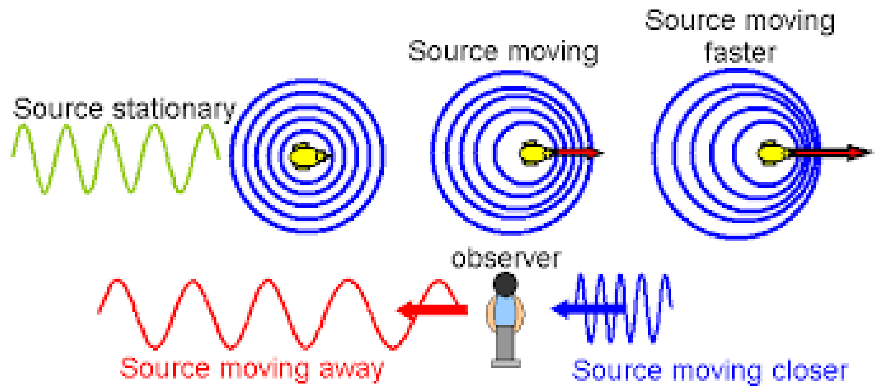

Fig. 1.3: Doppler Effect

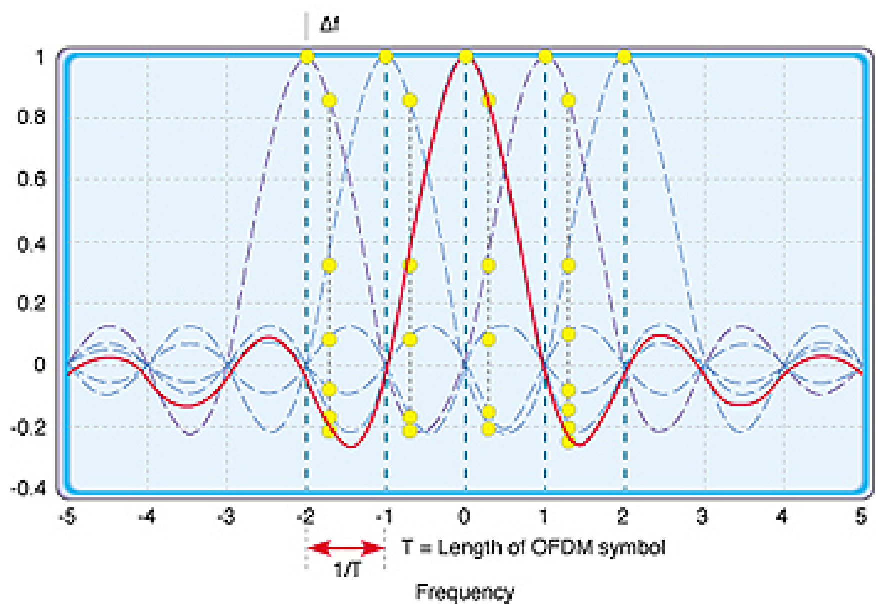

Fig. 1.4: CFO Effect

1.3 Research Background on CFO effect

In the literature, the CFO in an OFDM signal is modeled as a multiplicative exponential function of time in the received signal in the time domain. Thus, if the CFO can be estimated accurately, its impact can be simply eliminated by multiplying the received symbol in the time domain with the following factor

$$e^{\frac{j2\pi(\Delta f)n}{f_s}},$$

which will be explained in Chapter 2. As a result, the problem is leading to how to accurately estimate the CFO. In [12], the CFO was divided into two parts: integer part and fractional part. The integer part does not cause inter-subcarrier interference and only introduces a cyclic shift of data subcarriers. The integer part of CFO can be estimated accurately by using the pilots. However, the fractional part can cause interference among subcarriers, which results in destroying the orthogonality of the subcarriers. The OFDM system is extremely sensitive to the fractional part of the CFO. Since the fractional part of the CFO cannot be estimated with high accuracy by using pilots.

Since CFO is a key issue of OFDM systems, thus there has been numerous researches on how to estimate the CFO and how to compensate it. The current approach in the literature is to estimate the CFO first, and then correct the frequency of the received signal based on the estimated CFO. At last, the transmitted data is being detected based on the assumption that there is no CFO in the system. The drawback of this approach is that an accurate estimate of the CFO cannot be obtained and any estimation error in the CFO estimate can affect the accuracy of data detection. As a result, if the data is detected with the assumption that the frequency synchronization is perfect, then the would be larger than the real one, which means decrease the performance of the system. From [15], current techniques prefer to estimate the CFO by exploiting the pilots, which is called pilots-aided CFO estimation. And then detect the data without considering the impact of CFO. In this chapter, an overview of both current and a novel approach for CFO estimation and data detection is being presented. Figure 1.5 shows the differences between these two approaches.

1.3.1 Current Approach

Most of current CFO estimation algorithms are based on the pilots inserted either in the time domain or the frequency domain. In the time domain, the pilots are the preamble blocks in compared to the known symbols inserted in the selected subcarrier of the OFDM system in the frequency domain. These pilots need to be designed carefully to increase the accuracy of the CFO estimation.

In [16], with pilots are in the time domain, a Maximum Likelihood (ML) algorithm was developed for CFO estimation. The algorithm exploited repeated data symbols to track the CFO by exploring the relation between the repeated blocks. The author in [17] extends this idea by using one specified training sequence to estimate the CFO. Furthermore, an improved CFO estimation algorithm was proposed in [18], by redesigning the training symbols, to increase the accuracy at the expense of increasing computational complexity.

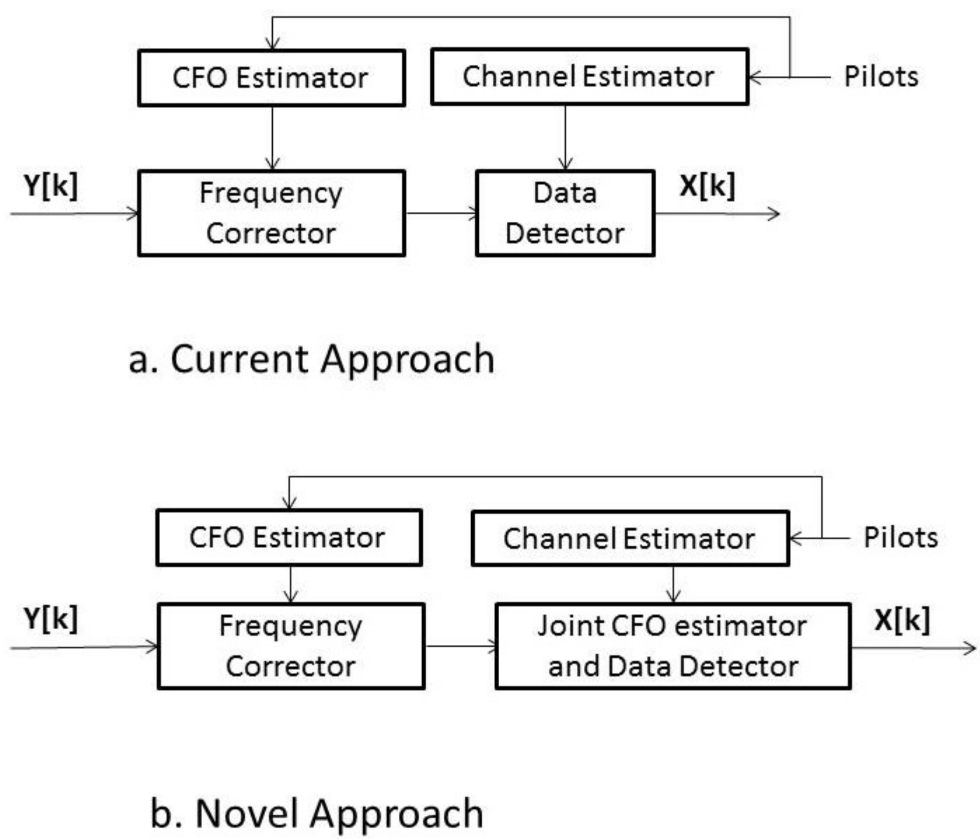

Fig. 1.5: Different Approachs

Alternatively, some others prefer to estimate the frequency offset with the added data in the frequency domain. An estimation algorithm based on this approach, which had two stages, was presented in [19]. Considering the performance of the CFO estimation algorithm, pilots needed to be designed carefully. The author in [20] proposed two novel methods to design the pilots. In [21], another method to design pilots was presented.

The accuracy of the CFO estimation can be improved by exploiting the statistical information about the wireless channel. A pilot-aided CFO estimator based on the ML algorithm has been proposed in [22], and the estimation range can be improved without decreasing accuracy [23]. By taking the channel error into consideration, two simple methods were proposed in [24]. A joint ML-based CFO and channel estimation algorithm was developed in [25] that has excellent performance on accuracy but with a relatively high computational complexity.

Figure 1.5a shows that after frequency correction, the data would be detected without considering the estimation error associated with the CFO estimate. There are four different detection algorithms. Specifically, zero forcing (ZF) equalizer, least squares (LS) equalizer, maximum likelihood (ML) detector, and expectation maximization (EM) detector.

ZF and LS are the algorithms based only on the transmitted symbols, the received symbols, and the estimated channel impulse response (CIR), [26]- [27]. The ZF equalizer is a single-tap frequency-domain linear equalizer [26]. This type of equalizer, proposed by *Robert Lucky*, is simply set to the inverse of the channel frequency response, and it can be applied to the received symbols to recover the transmitted symbols. The term *zero-forcing* means when there is no noise in the system, the ISI will be equal to zero. Based on that condition, the estimated symbols with perfect channel information can be equal to the transmitted symbols. The idea of ZF is simple, but it works well when the ISI is significant compared to the noise. Meanwhile, an LS equalizer is a detector that minimizes the squared error between the modulated transmitted symbols and the received symbols. From this aspect, LS cannot eliminate the ISI. However, the total power of noise can be minimized by this algorithm.

In addition to the ZF and LS equalizers, the ML and EM detectors take advantage of the statistical information of the channel estimation. The ML detector is an algorithm that estimates the data symbols by maximizing the likelihood function [28]. The ML algorithm can be easily implemented compared with the EM algorithm. However, one of the major problems with the ML detector is that it requires that both the pilot symbols and the data symbols are equipowered. The simulation results show that the ML detector performs better than the other three detectors, [28]. However, the ML algorithm as a sequence detector in fading channels typically has prohibitive computational complexity. The EM algorithm can be used an iterative method for finding ML or maximum a posterior (MAP), which can be applied for data detection in OFDM systems[29]. There are two steps for EM algorithms which are being called Expectation and Maximization.

1.3.2 Novel Approach

Due to the limitation of the accuracy of pilot-aided CFO estimation, the novel approach, shown in Fig. 1.5b is employed for improving the accuracy of CFO estimate. Even after the frequency corrector, the estimation error associated with CFO estimate must be taken into consideration while performing data detection. There are a few researches on the idea of joint CFO estimation and data detection.

In [30], an algorithm was proposed for joint CFO estimation and data detection based on pilots in OFDM systems. The algorithm employs linear approximation to present the system function and then obtains the approximated function by going through a linear search method, such as those in [31] and [32]. Then, to reduce the computational complexity, the joint estimator is divided into a CFO estimator and a data detector. That algorithm requires linear approximation and pilots information. However, this technique highly depends not only on the pilots, which reduce the spectral efficiency of the system but also the accuracy of the pilot estimator. In [33, 34], another algorithm for performing joint estimation that was based on a linear search method was proposed. In each iteration, the parameters were updated based on the minimum mean-squared error (MMSE) criterion. Moreover, in [35], an iterative EM-based algorithm that still separately performed CFO estimation and data detection was designed. In [36], the EM algorithm was employed for joint estimate CFO and detect data in the time domain for OFDM systems over frequency selective fast fading channel. However, there were two major problems with this algorithm. First, the computational complexity of the algorithm was extremely high due to the matrix inversion and partial eigendecomposition. Second, the frequency shift in fast fading channel cannot be separated as a constant from the channel impulse response. In addition, the frequency selective fast fading channel can only happen in high mobility, such as avionic communications. In most cases, frequency selectivity and fast fading cannot happen at the same time as it will be explained in Chapter 2. Thus, in most mobile communication systems, channels are frequency selective. Hence, the idea of jointly CFO estimation and data detection by using an iterative EM algorithm for OFDM systems over frequency selective channels is stimulated by this research.

1.4 Objectives of the Thesis

Motivated by the issues mentioned in the section above, the objectives of this thesis are stated as follows:

- Considering the novel approach for OFDM systems to solve the problem caused by CFO.
- Proposing an algorithm to overcome the CFO effects in the novel approach.

- Comparing the performance of the proposed algorithm with those of the conventional algorithms.

1.5 Organization of the Thesis

This thesis is mainly focused on reducing the effect of CFO on the performance of the OFDM systems over frequency selective channels. The rest of this thesis is organized as follows. Chapter 2 gives a background of the basic structure of OFDM systems. Chapter 3 presents an introduction of the wireless channel. Chapter 4 gives an overview of the algorithms related to estimation theory. Also, the expectation-maximization (EM) algorithm, which is related to the proposed algorithm, will be introduced in chapter 4. In chapter 5, a joint CFO estimation and data detection algorithm in the time domain is presented. In chapter 6, a joint CFO estimation and data detection algorithm in the frequency domain is proposed. Simulation results and analysis are presented in chapter 7. A summary of the results abstained in this research and suggestions for further studies are given in Chapter 8. The relevant mathematical derivation is given in Appendix.

2 OFDM Structure

In this chapter, first, the basic structure of an OFDM system is introduced. Then, the signal transmission procedure is stated from the mathematical point of view. Beyond this, the effect of CFO on the transmitted signal is presented in the second part of this chapter.

2.1 The Basis of OFDM system

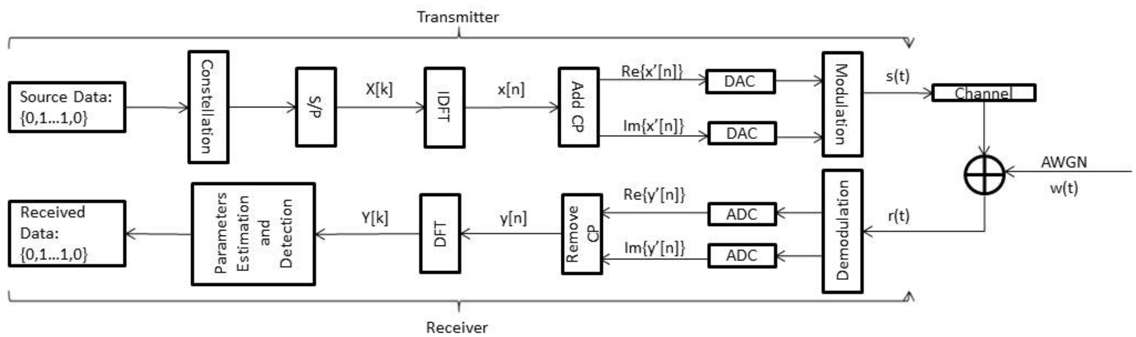

Fig. 2.1: Block Diagram of OFDM system

In an OFDM system, there are mainly two components, transmitter and receiver. Each component has several steps to process the signal in order to properly transmit the data from the transmitter to the receiver. Fig. 2.1 illustrates is a block diagram of a typical OFDM communication system.

The source data are a sequence of binary bits, for example, 11100100. The incoming bit stream of 1s and 0s is multiplexed into N parallel bit streams. Then, the N parallel bit stream are independently mapped in the frequency domain into complex symbols, denoted by X[k], in a given constellations such as phase shift Keying (PSK) or quadrature amplitude modulation (QAM), [37].

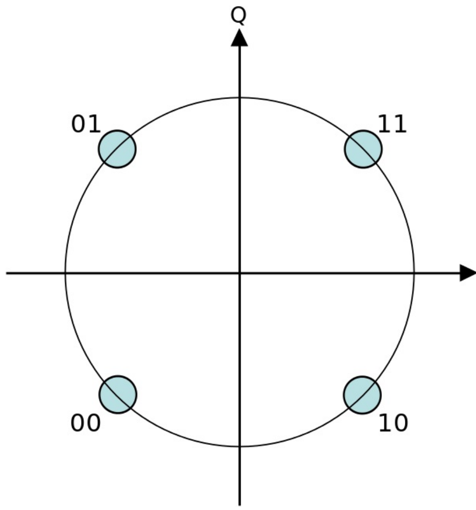

Fig. 2.2: 4-PSK

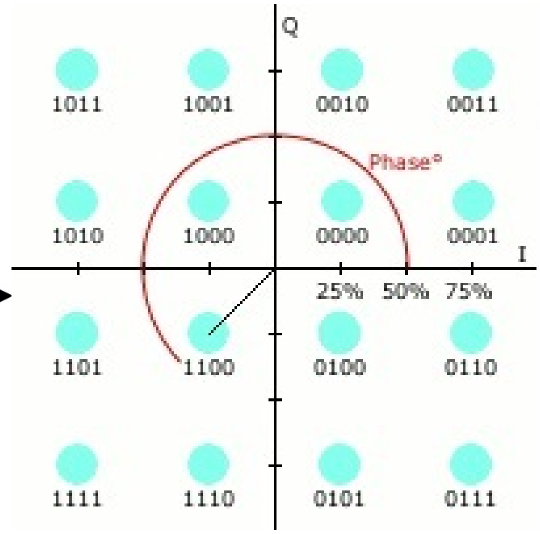

Fig. 2.3: 16-QAM

PSK is a digital modulation scheme that transfers the source data by modulating the phase of the carrier (or subcarriers), which uses a finite number of phases, each assigned to a unique pattern of data bits. Usually, each phase encodes an equal number of bits. Each pattern of bits forms the symbol that is represented by the particular phase. Fig. 2.2 shows the constellation diagram for 4-PSK modulation.

QAM is another digital modulation scheme that transfers source data by modulating the phase and amplitude at the same time. QAM arranges the source data in a square grid with equal vertical and horizontal spacing, each constellating point within the grid represent a set of bits. The constellation diagram for 16-QAM is shown in Fig. 2.3.

Moreover, the Discrete Fourier Transform (DFT) is an essential tool for a DFT-based OFDM system since the orthogonality of its subcarriers is preserved by the DFT. The DFT and Inverse Discrete Fourier Transform (IDFT) are being used to transfer the data between the time domain and the frequency domain. In an OFDM system, ISI caused by a time-dispersive channel can be eliminated by adding a cyclic prefix (CP) as a guard interval to a block of OFDM signal. The length of the CP must be greater than the delay spread of the channel. A CP is a set of symbols that is copied from the last part of each OFDM symbol block after the IDFT operation and then it is appended to the beginning of the block. Finally, the receiver can completely eliminate ISI by discarding the samples of the CP part before the DFT operation, as shown in Fig. 2.1.

A mathematical expression for an OFDM signal is derived in this section. The carrier modulated signal for an OFDM signal with a block of N symbols can be expressed in the following complex envelope form

$$s(t) = \mathrm{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\} \quad \text{for} \quad 0 \leq t \leq NT_s \tag{1}$$

where the complex envelope $\tilde{s}(t)$ can be written in the following form

$$\tilde{s}(t) = \sum_{n=0}^{N-1} \mathrm{IDFT}\{\mathrm{X}[k]\}p(t - nT_s), \tag{2}$$

and $p(t)$ which is known as the baseband pulse shape, may, for example, be a simple unit-energy rectangular pulse of duration T_s, or it could be a raised-cosine pulse, a bandlimited pulse, etc. Computing the N-point IDFT of $\{X[k]\}$ gives a complex-valued sequence $\{x[n]\}$ as

$$x[n] = \mathrm{IDFT}\{X[k]\} = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j2\pi nk/N}, \quad n = 0, 1, \ldots, N-1. \tag{3}$$

For convenience the following notation is used:

$$x_n(t) = x[n]p(t - nT_s). \tag{4}$$

Now the OFDM signal given in (1) can be written as follows:

$$s(t) = \mathrm{Re}\{\tilde{s}(t)e^{j2\pi f_c t}\} = \mathrm{Re}\{\sum_{n=0}^{N-1} x_n(t)e^{j2\pi f_c t}\} \tag{5}$$

or

$$s(t) = \mathrm{Re}\{\tilde{s}(t)\}\cos(2\pi f_c t) - \mathrm{Im}\{\tilde{s}(t)\}\sin(2\pi f_c t) \tag{6}$$

$$= \sum_{n=0}^{N-1} \mathrm{Re}\{x_n(t)\}\cos(2\pi f_c t) - \sum_{n=0}^{N-1} \mathrm{Im}\{x_n(t)\}\sin(2\pi f_c t) \tag{7}$$

At the receiver part, in an ideal communication system, which means a system with perfect time and frequency synchronization, an ideal channel, and with no additive noise, the received signal $r(t)$ and the transmitted signal $s(t)$ are identical. After all the operations are shown in Fig. 2.1, it can be easily shown that the detected symbols $y[n]$ and the transmitted symbols $x[n]$ are equal.

2.2 The Basis of CFO Effect in OFDM system

In this section, the effect of CFO on the received signal under perfect channel condition is investigated from a mathematical point of view. The demodulation operation at the receiver is performed by multiplying the received signal $r(t)$ by the signal of the local oscillator as follows:

$$z(t) = \tilde{s}(t)e^{j2\pi f_c t} \cdot e^{j2\pi f_r t} \tag{8}$$

where f_r is the frequency of the local oscillator. In order to show the CFO caused by the mismatch between the oscillator in the transmitter and that in the receiver, or the Doppler shift in the channel, the frequency of the local oscillator is written as follows:

$$f_r = f_c + \Delta f \tag{9}$$

where Δf represents the CFO in the system. Then, the output signal of the demodulator after lowpass filtering is given by

$$y(t) = \tilde{s}(t)e^{j2\pi \Delta f t}, \tag{10}$$

or

$$y(t) = \sum_{n=0}^{N-1} x_n(t)e^{j2\pi\Delta ft}, \tag{11}$$

Finally, after sampling $y(t)$ at sampling rate T_s, the received symbol in the time domain can be obtained as follows:

$$y[n] = x[n] \cdot e^{\dfrac{j2\pi\Delta fn}{f_s}} \tag{12}$$

where f_s is the sampling rate and $\dfrac{\Delta f}{f_s}$ is called normalized CFO (or just CFO for simplicity).

As seen in Equation (12), the effect of CFO in the system appears as a phase error between the transmitted symbol and the received symbol. Also, if the CFO can be accurately estimated, the impact of CFO on the received symbol in the time domain can be completely eliminated. This can be simply done by multiplying $y[n]$ by the correction factor $e^{-j2\pi\Delta fn/f_s}$, which is the reciprocal of the exponential term in Eq. (12), as mentioned in Chapter 1.

3 The Wireless Channel

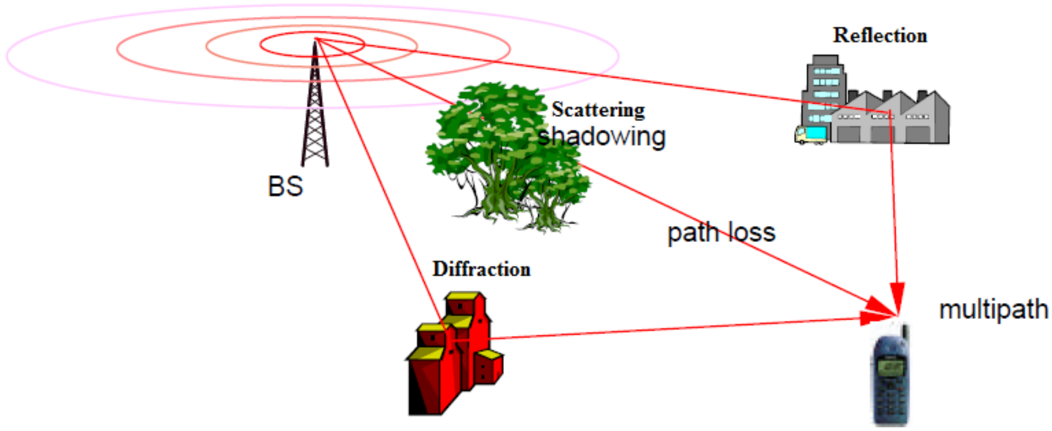

Fig. 3.1: The Wireless Channel

As mentioned in the 1st chapter, the key advantage of wireless communication compared with wireline communication is that the former one has no limitation on the communication channel. For wireline communication, the channel is a fixed environment which means that the channel would stay constant for a relatively long time. However, in wireless communication, the signal is transmitted via electromagnetic waves. Due to the variation of the wave paths and the mobility of both the transmitter and the receiver, the wireless channel becomes a significant factor to the communication system. Figure 3.1 is a typical diagram of the wireless channel. In this chapter, the fundamentals of the wireless channel are presented first, and then focusing on the channel model in OFDM system. Current channel estimation techniques are stated in the last section of this chapter.

3.1 Fundamentals of the Wireless Channel

Since the power of the radio waves is affected by the physical phenomenon of the environment. Thus the environment lays the foundation of the wireless channel. Generally speaking, there are two components of it: large scale fading and small scale fading [38].

Large scale fading represents the average signal power attenuation due to path loss and shadowing. More specially, path loss means the strength of the radio wave decay with respect to a function of distance. As the distance goes longer, the power of the wave becomes lower. In addition, even if two receivers have the same distance with the transmitter, they may still receive a different signal due to the shadowing. Shadowing is related to the environment between the transmitter and the receiver. Specifically, the shadowing effect caused by obstacles in the signal paths. In most cases, this fading effect can be accurately modeled as a random variable [39]. Fortunately, both path loss and shadowing are typically frequency independent and could be modeled properly. However, small scale fading does depend on the frequency which would severely decrease the performance the communication system.

Small scale fading is caused not only by the constructive and destructive interference of the different signal paths from the transmitter to the receiver which is being called multipath propagation but also by the mobility. Considering multipath propagation first, because of the different lengths of individual paths, each path would suffer different delays, which also termed as time shift or time spread. If the maximum time spread is greater than the symbol time, then ISI would be introduced by the channel. Under this condition, the received symbol combines two or more transmitted symbols which leading to difficulty in data detection. Otherwise, if all the spreads are within the symbol time, then the channel induced ISI would not occur. Since time spread in the time domain is equal to frequency selective in the frequency domain. Thus the wireless channel under multipath propagation effect could be characterized as frequency selective fading channel and frequency non-selective (flat) fading channel. Flat means all the subcarriers would suffer almost the same distortion. The frequency selectivity is related to the symbol rate and channel coherence bandwidth which is the bandwidth shows how rapidly a fading channel change with frequency.

If the symbol rate is greater than the channel coherence bandwidth, the channel would be frequency selective fading channel. Otherwise, the channel would be a flat fading channel. Furthermore, there is a relation between the multipath delay and the channel coherence bandwidth. In most wireless communication, such as mobile communication, the channel is frequency selective channel which is being interested in this thesis.

Moreover, if both the transmitter and the receiver are moving, then the Doppler Effect and time variation of the channel need to be taken into consideration. The author in [38] proves that the Doppler Effect caused by the motion leads to a spectral compression or dilation, which is also called frequency shift. The frequency shift is related to both the moving speed of the terminal and the carrier frequency. Furthermore, due to the motion of the terminals, the propagation paths would change all the time. This effect makes the channel as a time-varying system from nature. Under both Doppler effect and time variation conditions, the wireless channel can be characterized as slow fading channel and fast fading channel with respect to the symbol time and the channel coherence time. The channel coherence time shows how rapidly a fading channel changes with time. The fast fading channel is corresponding to the channel with channel coherence time less than the symbol interval. For fast fading channel, the channel is modeled as a linear time varying system which both contain the Doppler shift and time-varying channel impulse response. The slow fading channel is used to described the channel with the channel coherence time greater than the symbol duration. The slow fading channel can be approximately modeled as a linear time-invariant system with frequency shift within each symbol interval. Under this model, the frequency shift can be characterized as CFO. In most wireless communication, the channel is slow fading channel.

From above analysis, it has been figured out that multipath propagation leading to frequency selectivity of the signal, the motion of the terminal leading to time-varying channel impulse response. It is desired to transmitted signal in flat slow fading channel in order to make the transmission efficient. Thus the symbol time should be greater than the maximum multipath time spread regarding the frequency selectivity, and smaller than the channel coherence time with respect to the time variation. In mobile communication, the channel coherence time is much larger than the multipath time spread (milliseconds compared to microsecond or nanosecond). As a result, frequency selective fading and fast fading could not happen at the same time in relative low mobility wireless communication. But for avionic communication, both frequency selectivity and fast fading can appear at the same time. Both of the model of frequency selective fast fading channel and frequency selective slow fading channel would be derived in this thesis. Figure 3.2 is the block diagram of characterization of the wireless channel in relative low mobility.

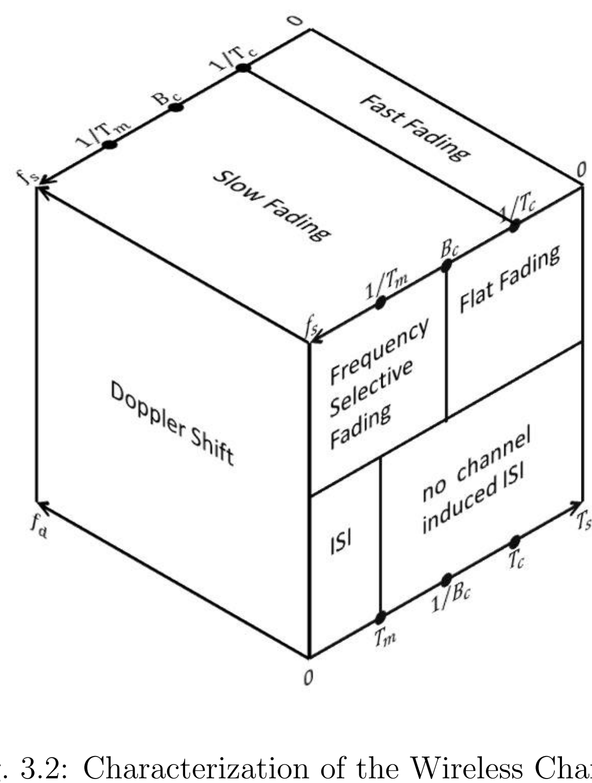

Fig. 3.2: Characterization of the Wireless Channel

3.2 The Wireless Channel and System Model in OFDM

The last section shows that the wireless channel is an LTV system with frequency selectivity due to the multipath propagation and moving transmitter and receiver. One important advantage of OFDM system is that it converts a frequency selective fading channel into flat fading channel by transmitting high-speed data to low speed data stream via multicarrier technique. Thus the channel induced ISI is avoided, but the signal distortion still there.

Since the effect of frequency selectivity is overcome by the multicarrier technique in OFDM system, then only fast fading and slow fading should be considered. For fast fading channel, the channel is modeled as an LTV system which both contain the Doppler effect and the time-varying path propagation. One important assumption for most of the communication system is that the major noise is at the receiver and independent of the channel. Under this assumption, the noise is modeled as an Additive White Gaussian Noise (AWGN). Then the relation between the transmitted signal and received signal can be presented as

$$\begin{aligned} \mathrm{r}(t) &= \mathrm{h}(t) * \mathrm{s}(t) + \mathrm{w}(t) \\ &= \int_{-\infty}^{\infty} \mathrm{h}(t,\tau)\,\mathrm{s}(\tau)\,d\tau + \mathrm{w}(t) \end{aligned} \tag{13}$$

where $\mathrm{s}(t)$ is the transmitted signal, $\mathrm{r}(t)$ is the received signal, $\mathrm{h}(t)$ is the channel impulse response, $\mathrm{w}(t)$ is the AWGN.

If this equation is properly sampled, then the system model in discrete form would be

$$\begin{aligned} \mathrm{r}[n] &= \mathrm{h}[n] * \mathrm{s}[n] + \mathrm{w}[n] \\ &= \sum_{m=-\infty}^{\infty} \mathrm{h}[n,m]\,\mathrm{s}[m] + \mathrm{w}[n] \end{aligned} \tag{14}$$

After removing the CP, equation (14) can be expressed in matrix form as

$$\mathbf{y} = \mathbf{CIR}\left(\mathbf{h}\right)\mathbf{x} + \mathbf{w} \tag{15}$$

where $\mathbf{y}\left[n\right]$ is the received symbols, $\mathbf{x}\left[n\right]$ is the transmitted symbols, and the channel coefficients matrix is as follow

$$\mathbf{CIR}\left(\mathbf{h}\right) = \begin{bmatrix} h\left(0,0\right) & 0 & \dots & h\left(0,2\right) & h\left(0,1\right) \\ h\left(1,1\right) & h\left(1,0\right) & \dots & h\left(1,3\right) & h\left(1,2\right) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & h\left(N-2,0\right) & 0 \\ 0 & 0 & \dots & h\left(N-1,1\right) & h\left(N-1,0\right) \end{bmatrix} \tag{16}$$

This channel is a frequency selective fast fading channel. However, this kind of channel is only for high mobility, like avionic communication. In most cases, the transmitter and receiver are not moving that fast, under this situation, the channel is modeled as a Linear Time Invariant (LTI) system with frequency shift within each symbol interval.

Alternatively, for an LTI system, the relation would change to

$$\begin{aligned} \mathrm{r}\left(t\right) &= \mathrm{h}\left(t\right) * \mathrm{s}\left(t\right) + \mathrm{w}\left(t\right) \\ &= \int_{-\infty}^{\infty} \mathrm{h}\left(t-\tau\right)\mathrm{s}\left(\tau\right) d\tau + \mathrm{w}\left(t\right) \end{aligned} \tag{17}$$

After sampling equation (17), the system model would be

$$\begin{aligned} \mathrm{r}\left[n\right] &= \mathrm{h}\left[n\right] * \mathrm{s}\left[n\right] + \mathrm{w}\left[n\right] \\ &= \sum_{m=-\infty}^{\infty} \mathrm{h}\left[n-m\right]\mathrm{s}\left[m\right] + \mathrm{w}\left[n\right] \end{aligned} \tag{18}$$

After removing the CP, equation (18) can be expressed as

$$\mathbf{y} = \mathbf{H}(\mathbf{h})\,\mathbf{x} + \mathbf{w} \tag{19}$$

where

$$\mathbf{H}(\mathbf{h}) = \begin{bmatrix} h(0) & 0 & \dots & h(2) & h(1) \\ h(1) & h(0) & \dots & h(3) & h(2) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & h(0) & 0 \\ 0 & 0 & \dots & h(1) & h(0) \end{bmatrix} \tag{20}$$

Now, considering frequency shift to the system, then system in time domain can be modeled as

$$\mathbf{y} = \mathbf{T}(\epsilon)\,\mathbf{H}(\mathbf{h})\,\mathbf{x} + \mathbf{w} \tag{21}$$

where $\mathbf{T} = diag\left\{\left[1, e^{-\frac{j2\pi\epsilon}{N}}, \dots, e^{-\frac{j2\pi\epsilon(N-1)}{N}}\right]^T\right\}$ is the CFO effect matrix.

The channel model in the above equation is called frequency selective slow fading channel, short for frequency selective channel. Note that, for OFDM system, the frequency shift is not only coming from Doppler effect, but also from the difference of the oscillator between the transmitter and the receiver. The effect of frequency shift which is termed as CFO which has already been proved in the 2nd chapter.

The corresponding frequency domain system model would be

$$\begin{aligned} \mathbf{Y} &= \mathbf{FT}(\epsilon)\,\mathbf{H}(\mathbf{h})\,\mathbf{x} + \mathbf{W} \\ &= \mathbf{FT}(\varepsilon)\,\mathbf{F}^{-1}\mathbf{HX} + \mathbf{W} \end{aligned} \tag{22}$$

where $\mathbf{Y} = [Y_0, \dots, Y_{N-1}]^T$ is the frequency domain received symbols and $\mathbf{X} = [X_0, \dots, X_{N-1}]^T$ is the frequency domain transmitted symbols, and $\mathbf{F}$ is the $N \times N$ DFT matrix, and $\mathbf{F}^{-1} = \frac{\mathbf{F}^H}{N}$. Moreover, $\mathbf{H} = diag\left\{[H(0), \dots, H(N-1)]^T\right\}$ is the channel frequency response diagonal matrix and also $\mathbf{H} = diag\{\mathbf{F}_L\mathbf{h}\}$, where $\mathbf{h} = [\mathrm{h}(0), \ldots, \mathrm{h}(L-1)]^T$ stands for CIR vector while L is the length of channel.

Since in most applications, the system is not under high mobility, thus the frequency selective channel is more adaptive. This thesis would only consider the CFO effect in OFDM system over frequency selective channel. There is two statistic model of frequency selective channel: Rician fading channel and Rayleigh fading channel. If the channel has a fixed light of sight, then the CIR have a Rician distribution. Otherwise, the CIR follows Rayleigh distribution. Since in most cases, the light of sight is not possible, thus Rayleigh fading channel model is being used in this thesis.

3.3 Techniques of Channel Estimation in OFDM

Before properly detecting the data, the channel should be estimated first. In an OFDM system, pilot-aided channel estimation is the most popular used technique. The pilots can be either inserted in time domain as a preamble or in frequency domain [40, 41, 42]. Two key factors that affect the performance of the pilot-aided channel estimation are pilot design and interpolation. There have many researches on pilots design related problems. The author in [43] proposes an algorithm to estimate the channel by using pilot tones. The optimal design of pilots has been researched for frequency selective channel in [44] and [45]. An overview of pilot-aided channel estimation for wireless communication has been presented in [46]. The accuracy of channel estimation can be improved by involving channel statistic information [42].

This thesis would focus on joint CFO estimation and data detection for OFDM system over frequency selective channel. Thus perfect channel information is assumed to be achieved as a prior. In addition, in order to perform joint estimation, the estimation algorithm should be considered. In next chapter, a review of estimation algorithms would be given. And the joint estimation algorithm would be picked based on the analysis of several different algorithms.

4 Estimation Algorithms

As a result from previous chapters, it is desired to design a joint estimation algorithm to deal with the CFO effect for an OFDM system with frequency selective channel. From the previous chapters, we can see that the system model which both considering the CFO and the transmitted data is a linear complex matrix equation with nonlinear parameters. Although it is extremely difficult to jointly solve this kind of problem, there are still several algorithms that can be considered to deal with it. In the following paragraphs, the estimation theory and associated algorithms would be reviewed first. Then based on the comparison of the presented algorithms, the EM algorithm would be picked to deal with the joint estimation problem and explained in the second part of this chapter.

4.1 Review of Estimation Algorithms

Estimation can be found in many signal processing systems which designed to extract the desired information, such as Speech, Communication, Control, etc. In order to properly estimate the parameters, various algorithms had been developed to handle the estimation problem. Based on the characteristic of the unknown parameters, there are two approaches: classical approach and Bayesian approach [47]. Specifically, in classical approach, the unknown parameter is assumed to be a deterministic constant which means the probability density function (PDF) of the parameter is not available, and the observed data information is summarized by the PDF that depends on the parameter. In contrast, in Bayesian approach, the unknown parameter is modeled as a stochastic process, and the PDF of it is known as a prior. According to this classification, the associated algorithms have been stated in 4.1.

Table 4.1: Classification Of Estimation Algorithm

Classical Approach	Bayesian Approach
CRLB	MMSE
RBLS	MAP
BLUE	LMMSE
ML	
LS	
MOM	

Since the problem that being interested is to estimate parameter vector based on an observed data vector, which means a discrete estimation problem, then we start with a general discrete estimation problem. After that, the algorithm is picked by both considering the conditions and comparing different algorithms. From the mathematical point of view, the problem is modeled as follow:

$$\mathbf{Y} = S(\mathbf{X}) + \mathbf{W} \tag{23}$$

where $\mathbf{Y}$ is the observed data vector, $\mathbf{X}$ is the unknown parameter vector that need to be estimated, $\mathbf{W}$ is the AWGN vector in most cases, but in general, it does not have to be white Gaussian noise, $S(\cdot)$ represents a general transformation on the unknown parameter.

Based on above information, the estimation problem is to derive a solution as

$$\hat{\mathbf{X}} = f(\mathbf{Y}) \tag{24}$$

where $\hat{\mathbf{X}}$ is the estimated parameters, $f(\cdot)$ is a specified operation on the observed data. Moreover, since the noise is modeled as AWGN, then the conditional PDF of the observed data $p(\mathbf{Y}|\mathbf{X})$ is available.

Now, considering Bayesian Approach first, MMSE, MAP, and LMMSE represent

Minimum Mean Squared Error, Maximum A Posteriori, and Linear Minimum Mean Squared Error, respectively. Since all of these algorithms require statistic information of the unknown parameter $\mathbf{X}$ as a prior, nonetheless the CFO and data are actually deterministic without the required prior information. As a result, Bayesian approach is not suitable for the joint estimation problem in this thesis. The algorithm for joint estimation should be selected from the classical approach.

For classical approach, a good estimator of unknown parameters is to see whether the average of the estimated value yields the true value and whether the variance has the minimum value. If the average is equal to the true value, $E\left\{\hat{\mathbf{X}}\right\} = \mathbf{X}$, the estimator is called unbias estimator. And if the variance is minimized, $Var\left\{\hat{\mathbf{X}}\right\} = Var_{min}\left\{\hat{\mathbf{X}}\right\}$, the estimator is called minimum variance estimator. Thus the goal in classical approach is to find a minimum variance unbiased (MVU) estimator. Based on this criteria, several algorithms are reviewed as follows.

4.1.1 Cramer-Rao Lower Bound

First of all, Cramer-Rao Lower Bound (CRLB) is the easiest method to determine the lower bound of the variance of any unbiased estimator. Then the estimator can be found according to the first derivative of the ln conditional PDF.

It first assumes that

$$E\left\{\frac{\partial \ln p\left(\mathbf{Y}|\mathbf{X}\right)}{\partial \mathbf{X}}\right\} = \mathbf{0} \tag{25}$$

Then the varince of the estimator would satisfy

$$Var\left\{\frac{\partial \ln p\left(\mathbf{Y}|\mathbf{X}\right)}{\partial \mathbf{X}}\right\} \geq \frac{1}{-E\left[\frac{\partial^2 \ln p(\mathbf{Y}|\mathbf{X})}{\partial^2 \mathbf{X}}\right]} \tag{26}$$

Then if the first derivative of the ln conditional PDF can be expressed as

$$\frac{\partial \ln p\left(\mathbf{Y}|\mathbf{X}\right)}{\partial \mathbf{X}} = \mathbf{I}\left(\mathbf{X}\right) \cdot \left(\mathbf{g}\left(\mathbf{Y}\right) - \mathbf{X}\right) \tag{27}$$

As a result, $\mathbf{X}$ would be estimated as

$$\hat{\mathbf{X}} = \mathbf{g}\left(\mathbf{Y}\right) \tag{28}$$

From equation (28), CRLB is a straightforward method to estimate the parameters under MVU condition. However, for nonlinear transformation as CFO, the derivative can not be achieved easily. And even if the derivative can be achieved, the specified expression would be a problem. Thus CRLB is not appropriate.

4.1.2 Rao-Blackwell-Lehmann-Scheffe

Moreover, Rao-Blackwell-Lehmann-Scheffe (RBLS) is another MVU estimator. By factorizing the conditional PDF in the following form

$$p\left(\mathbf{Y}|\mathbf{X}\right) = g\left(T\left(\mathbf{Y}\right), \mathbf{X}\right) h\left(\mathbf{Y}\right) \tag{29}$$

The estimator can be found as

$$\hat{\mathbf{X}} = \mathbf{k}\left(T\left(\mathbf{Y}\right)\right) \tag{30}$$

where $\mathbf{k}\left(T\left(\mathbf{Y}\right)\right)$ is a function that satisfy $E\left\{\mathbf{k}\left(T\left(\mathbf{Y}\right)\right)\right\} = \mathbf{Y}$, and $T\left(\mathbf{Y}\right)$ is called sufficient statistic.

The process of finding the estimator is simple. Nevertheless, it is hard to find the sufficient statistic and factor the PDF in the required form. Thus RBLS is not suitable either.

4.1.3 Best Linear Unbiased Estimatore

In addition, Best Linear Unbiased Estimator (BLUE) is a suboptimal MVU estimator which only requires the first and the second moments of the noise. The condition is that if

$$S(\mathbf{X}) = \mathbf{H}\mathbf{X} \tag{31}$$

where $\mathbf{H}$ is a known matrix, which means $\mathbf{Y}$ is linearly related to $\mathbf{X}$. Moreover, the noise has zero mean and known covariance matrix as $\mathbf{C}$. Then the estimator would be

$$\hat{\mathbf{X}} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{Y} \tag{32}$$

Since complete PDF of the observed data $\mathbf{Y}$ is not necessary, BLUE is more suitable for practical implementation. However, a key issue of BLUE is that it requires the system has a linear relation which can not be satisfied because of the CFO as equation (22). Thus BLUE is also not appropriate.

4.1.4 Least Squared

Furthermore, LS estimator is a basic algorithm that only minimizes the squared error. It does not need any statistic information. The estimated parameter is the value that minimizes the squared error function, where the squared error function is

$$J(\mathbf{X}) = (\mathbf{Y} - S(\mathbf{X}))^T (\mathbf{Y} - S(\mathbf{X})) \tag{33}$$

But the drawbacks are that it is neither an MVU estimator nor optimal in general. Moreover, the direct solution is not available. Thus this algorithm is not proper to be chosen to perform the joint estimation.

4.1.5 Method of Moments

Besides that, Method of Moments is another algorthm that do not need entire PDF of the observed data. It only requires the N moments of the observed data $\mathbf{Y}$, where N is the size of $\mathbf{Y}$. Suppose the ith moment is $\mu_i = h_i(\mathbf{X})$. The parameter is estimated as

$$\hat{\mathbf{X}} = h^{-1}\left(\hat{\mu}\right) \tag{34}$$

where $\hat{\mu}_i = \frac{1}{N}\sum_{n=0}^{N} X^i\left[n\right]$

The advantage of this method is that it is easy to implement. However, due to the none optimality and the requirement of multiple moments, it can not be selected as the joint estimation algorithm.

4.1.6 Maximum Likelihood

At last, ML estimator would be the one that being chosen to solve the problem that mentioned in the 1st chapter because of the advantages specified in the following paragraph.

Compared with above algorithms, ML has several advantages. First of all, ML is an optimal algorithm if the MVU estimator exists. If not, it is the approximate MVU estimator. Moreover, it can be implemented for complicated estimation problem. Lastly, the only requirement of it is the PDF of the observed data. Due to these reasons, almost all practical estimators are based on the ML principle. Thus ML is being selected at the beginning. The idea of ML is simple, the unknown parameter $\mathbf{X}$ is merely the value that maximize the likelihood function $p\left(\mathbf{Y}|\mathbf{X}\right)$. However, due to the nonlinearity of the parameters in our system model, the direct form of ML estimator can not be solved easily. Then alternative algorithms are being considered to solve the ML problem.

Firstly, since the ML can always be found for a given data set numerically, thus grid search is the safest method. However, considering the huge computation problem, grid search is not a suitable solution. Under this condition, we are forced to consider the iterative procedure. The Newton-Raphson method and the scoring approach are two iterative methods that had been considered. Nonetheless, the convergence of these two methods can not be guaranteed. At last, another iterative method, EM algorithm

[48], is being considered as the algorithm to maximize the likelihood function.

In comparing with the mentioned algorithms in the last paragraph, EM algorithm is a better choice since it can maximize the likelihood function in an iterative way with guaranteed convergence. Then the problem can be solved iteratively by treating the observations as incomplete data. In addition, the EM algorithm does not require any approximation and pilots information in solving the CFO problem. At last, the EM algorithm is excellent also because of its computation simplicity. Due to above advantages, EM algorithm is being picked to jointly estimate the CFO and detect the data for OFDM system with frequency selective channel in this thesis. In the following paragraphs, more details would be given to help understand the EM algorithm.

4.2 Expectation Maximization Algorithm

As stated in the last chapter, EM algorithm is a better method to solve ML detection problem. Let's start from ML algorithm.

Suppose that the PDF of a random known vector $\mathbf{Y}$ is given as $f\left(\mathbf{Y}|\mathbf{X}\right)$, where $\mathbf{X}$ is the vector of interest. Then the given conditional PDF of $\mathbf{Y}$ is the likelihood function for ML algorithm. As a result, the solution of the ML algorithm in this condition would be the $\mathbf{X}$ that maximize the likelihood function or log likelihood function. However, if the likelihood function $f\left(\mathbf{Y}|\mathbf{X}\right)$ highly nonlinear depends on any parameters or variables, then analytical solution would not be easy to derive. Alternatively, if there exist another random vector $\mathbf{H}$ which can not be observed directly but has a relation with $\mathbf{X}$, then according to Leibiz Integral Rule, we have:

$$f\left(\mathbf{Y}|\mathbf{X}\right) = \int f\left(\mathbf{Y},\mathbf{H}|\mathbf{X}\right) d\mathbf{H} \tag{35}$$

Under this condition, the ML problem convert to maximize the integral of the conditional marginal PDF in equation (35). In this case, $\mathbf{Y}$ is called incomplete data,

$\mathbf{H}$ is called missing data, $\{\mathbf{Y},\mathbf{H}\}$ is called complete data.

By using Bayes' rule, we have

$$f(\mathbf{Y},\mathbf{H}|\mathbf{X}) = f(\mathbf{H}|\mathbf{Y},\mathbf{X})\, f(\mathbf{Y}|\mathbf{X}) \tag{36}$$

To proceed, take the log to equation (36) and rearrange the equation, then

$$\log f(\mathbf{Y}|\mathbf{X}) = \log f(\mathbf{Y},\mathbf{H}|\mathbf{X}) - \log f(\mathbf{H}|\mathbf{Y},\mathbf{X}) \tag{37}$$

Assume we have the conditional PDF of $\mathbf{H}$ conditioned on $\mathbf{Y}$ and $\mathbf{X}^i$, where $\mathbf{X}^i$ represents the estimated $\mathbf{X}$ in the ith iteration. Then take expectation with respect to $\mathbf{H}$ in equation (37) on both sides, we have

$$\begin{aligned} E_{\mathbf{H}}\{\log f(\mathbf{Y}|\mathbf{X})\} &= E_{\mathbf{H}}\{\log f(\mathbf{Y},\mathbf{H}|\mathbf{X})\} - E_{\mathbf{H}}\{\log f(\mathbf{H}|\mathbf{Y},\mathbf{X})\} \\ &= \int \log f(\mathbf{Y},\mathbf{H}|\mathbf{X})\, f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^i\right) d\mathbf{H} \\ &\quad - \int \log f(\mathbf{H}|\mathbf{Y},\mathbf{X})\, f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^i\right) d\mathbf{H} \end{aligned} \tag{38}$$

Since $f(\mathbf{Y}|\mathbf{X})$ is nothing do with $\mathbf{H}$, then

$$E_{\mathbf{H}}\{\log f(\mathbf{Y}|\mathbf{X})\} = \log f(\mathbf{Y}|\mathbf{X}) \tag{39}$$

Combining equation (38) and (39), we have

$$\log f(\mathbf{Y}|\mathbf{X}) = \int \log f(\mathbf{Y},\mathbf{H}|\mathbf{X}) \cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^i\right) d\mathbf{H} - \int \log f(\mathbf{H}|\mathbf{Y},\mathbf{X}) \cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^i\right) d\mathbf{H} \tag{40}$$

At this point, the ML problem reduce to maximize the right side of equation (40).

For simplicity, define

$$Q\left(\mathbf{X}^{|}\mathbf{X}^{i}\right)=\int\log f\left(\mathbf{Y},\mathbf{H}|\mathbf{X}\right)\cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)d\mathbf{H} \tag{41}$$

$$P\left(\mathbf{X}|\mathbf{X}^{i}\right)=\int\log f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}\right)\cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)d\mathbf{H} \tag{42}$$

In order to increase the value of $\log f\left(\mathbf{Y}|\mathbf{X}\right)$ in each iteration, we should ensure that

$$\log f\left(\mathbf{Y}|\mathbf{X}^{i+1}\right)\geq\log f\left(\mathbf{Y}|\mathbf{X}^{i}\right) \tag{43}$$

which means

$$Q\left(\mathbf{X}^{i+1}|\mathbf{X}^{i}\right)-P\left(\mathbf{X}^{i+1}|\mathbf{X}^{1}\right)\geq Q\left(\mathbf{X}^{i}|\mathbf{X}^{i}\right)-P\left(\mathbf{X}^{i}|\mathbf{X}^{i}\right) \tag{44}$$

Notice that

$$\begin{aligned}P\left(\mathbf{X}^{i}|\mathbf{X}^{i}\right)-P\left(\mathbf{X}^{i+1}|\mathbf{X}^{i}\right) &= \int\log f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)\cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)d\mathbf{H}\\ &\quad-\int\log f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i+1}\right)\cdot f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)d\mathbf{H}\\ &= -\int f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)\cdot\log\frac{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i+1}\right)}{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)}d\mathbf{H}\end{aligned} \tag{45}$$

Since for natural logarithm inequality, $\log x\leq x-1$when $x\geq 0$, then

$$\begin{aligned}\int f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)\cdot\log\frac{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i+1}\right)}{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)}d\mathbf{H} &\leq \int f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)\cdot\left(\frac{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i+1}\right)}{f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)}-1\right)d\mathbf{H}\\ &= \int f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i+1}\right)d\mathbf{H}-\int f\left(\mathbf{H}|\mathbf{Y},\mathbf{X}^{i}\right)d\mathbf{H}\\ &= 1-1=0\end{aligned} \tag{46}$$

Thus

$$P\left(\mathbf{X}^{i}|\mathbf{X}^{i}\right)-P\left(\mathbf{X}^{i+1}|\mathbf{X}^{i}\right)\geq 0 \tag{47}$$

for any $\mathbf{X}$.

Then according to equation (44), what we need to do is to find $\mathbf{X}^{i+1}$ that satisfies

$$Q\left(\mathbf{X}^{i+1}|\mathbf{X}^{i}\right) \geq Q\left(\mathbf{X}^{i}|\mathbf{X}^{i}\right) \tag{48}$$

By performing this operation, the Q function increases iteration by iteration, which means iteratively maximize the likelihood function.

The method introduced above is called EM algorithm, since it constructs an expectation function first, and then maximizing the expectation function iteration by iteration. There are two steps which are called Expectation Step (E-step) and Maximization Step (M-step). The E-step is used to handle the randomness of some parameters. Then the solution can be searched through maximizing the function derived in E-step. In the following paragraphs, in order to give a clear idea of how the EM algorithm works, the algorithm would be introduced via solving a simple linear model.

Considering a general linear model as:

$$\mathbf{Y} = \mathbf{H}\left(\mathbf{h}\right)\mathbf{X} + \mathbf{W} \tag{49}$$

where $\mathbf{Y}$ is a known column vector, $\mathbf{h}$ is a column vector that can not be observed, but independent of any other parameters and follow a specified distribution, such as Gaussian Distribution. Moreover, $\mathbf{H}$ is nonlinear depend on $\mathbf{h}$. And $\mathbf{X}$ stands for a column vector need to be estimated. $\mathbf{W}$ is the AWGN column vector.

$\mathbf{X}$ should be estimated by using the above knowledge. The solution of the ML estimate of $\mathbf{X}$ is

$$\mathbf{X} = \mathbf{argmax} f\left(\mathbf{Y}|\mathbf{X}\right) \tag{50}$$

Nonetheless, since the system nonlinearly depends on $\mathbf{h}$, the direct solution of

ML estimate would not be solved easily. At this point, the EM algorithm is being introduced to solve this problem in a relatively easy way. Within EM algorithm, $\mathbf{Y}$ is called incomplete data, $\mathbf{h}$ is called the missing data, and $\{\mathbf{Y},\mathbf{h}\}$ is called complete data. Then the likelihood function can be expressed as the integration of the marginal likelihood function as

$$f\left(\mathbf{Y}|\mathbf{X}\right) = \int f\left(\mathbf{Y},\mathbf{h}|\mathbf{X}\right) d\mathbf{h} \tag{51}$$

Then ML estimation can be simplified by using the marginal PDF. The solution of the EM algorithm is the variables that maximize the expectation of the log marginal likelihood function as follow

$$\mathbf{X} = \mathbf{argmax} E_h \left\{ log \left\{ f\left(\mathbf{Y},\mathbf{h}|\mathbf{X}^i\right)\right\} |\mathbf{Y}, \mathbf{X}^{i-1}\right\} \tag{52}$$

There are two iterative steps to solve this problem:

E-step:

Forming the Q function as

$$Q\left(\mathbf{X}^i|\mathbf{X}^{i-1}\right) = E_h \left\{ log \left\{ f\left(\mathbf{Y},\mathbf{h}|\mathbf{X}^i\right)\right\} |\mathbf{Y}, \mathbf{X}^{i-1}\right\} \tag{53}$$

M-step:

Solving the Q function to find the solution at current iteration as

$$\mathbf{X} = \mathbf{argmax} Q(\left(\mathbf{X}^i|\mathbf{X}^{i-1}\right) \tag{54}$$

Figure 4.1 is the summarized procedure of the EM algorithm.

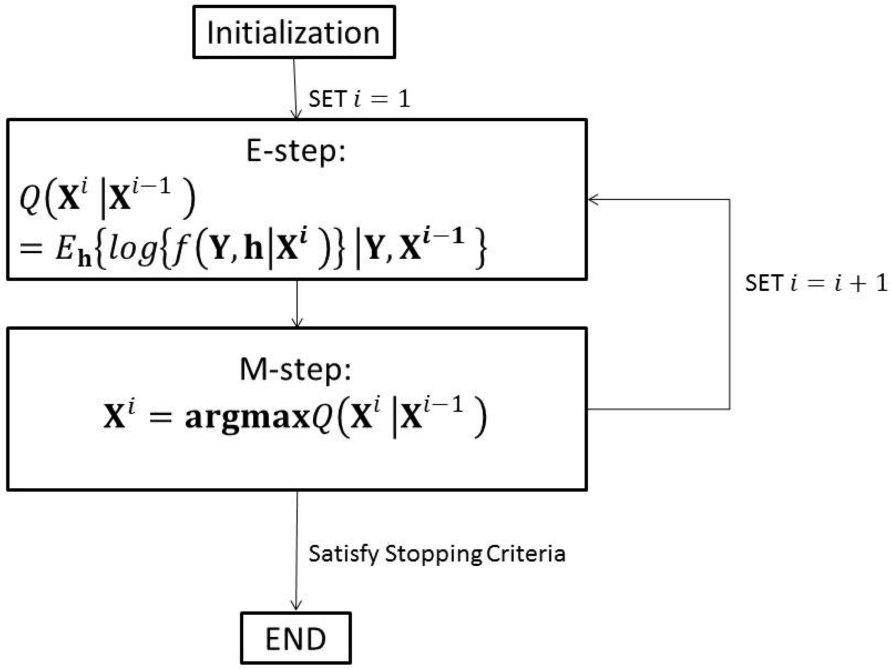

Fig. 4.1: EM Algorithm Procedure

From this chapter, we would have a clear idea about how the EM algorithm derived and the procedure of the algorithm. It is obvious that this algorithm is a better choice for those systems with nonlinear parameters or variables that can not be solved directly via ML algorithm. However, although the algorithm is simple which only involves two steps in each iteration, the mathematical derivation for an individual case is not as simple as the procedure. In next chapter, the EM algorithm would be applied to an OFDM system with frequency selective channel to jointly estimate the CFO and detect the data in time domain.

5 Joint CFO Estimation and Data Detection Algorithm for Time Domain

In this chapter, a joint CFO estimation and data detection algorithm for OFDM system over frequency selective channel is presented based on EM algorithm. As I mentioned before, this joint algorithm is proposed in [36] which deal with frequency selective fast fading channel. However, for fast fading channel, the Doppler shift can not be characterized as CFO as described in [36]. Thus the method is not correct in that paper. But in this thesis, the EM algorithm is being applied to frequency selective channel which is appropriate as explained in chapter 3, and the requirement of the pilot information is eliminated for computation simplicity. The time domain system model is derived in chapter 3 as equation (21). The details would be given in the following paragraphs.

The time domain OFDM system can be modeled as follow:

$$\mathbf{y} = \mathbf{T}\mathbf{H}(\mathbf{h})\,\mathbf{x} + \mathbf{w} \tag{55}$$

where $\mathbf{T} = diag\left\{\left[1, e^{-\frac{j2\pi\epsilon}{N}}, \ldots, e^{-\frac{j2\pi\epsilon(N-1)}{N}}\right]^T\right\}$ represents the matrix of CFO, $\mathbf{y} = [y_0, \ldots, y_{N-1}]^T$ is the time domain received symbols and $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T$ is the time domain transmitted symbols with the size of N, $\mathbf{w}$ is the AWGN matrix with the covariance $\sigma_n^2 I_N$. Thus the received symbol vector $\mathbf{y}$ follows Complex Gaussian distribution with mean $\mathbf{T}\mathbf{H}(\mathbf{h})\,\mathbf{x}$ and covariance matrix $\sigma_n^2 I_N$ as

$$f(\mathbf{y}|\mathbf{x}, \mathbf{h}, \epsilon) = \frac{1}{\pi^N \sigma_n^2} \exp\left\{-\frac{(\mathbf{y} - \mathbf{T}\mathbf{H}(\mathbf{h})\,\mathbf{x})^H (\mathbf{y} - \mathbf{T}\mathbf{H}(\mathbf{h})\,\mathbf{x})}{\sigma_n^2}\right\} \tag{56}$$

$\mathbf{H}(\mathbf{h})$ is the $N \times N$ convolution matrix, $\mathbf{h} = [\mathrm{h}(0), \ldots, \mathrm{h}(L-1)]^T$ is the CIR and

$$\mathbf{H}(\mathbf{h}) = \begin{bmatrix} h(0) & 0 & \ldots & h(2) & h(1) \\ h(1) & h(0) & \ldots & h(3) & h(2) \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & h(0) & 0 \\ 0 & 0 & \ldots & h(1) & h(0) \end{bmatrix} \tag{57}$$

Since $\mathbf{H}(\mathbf{h})\mathbf{x}$ represents the convolution, then the following property hold:

$$\mathbf{H}(\mathbf{h})\mathbf{x} = \mathbf{D}(\mathbf{x})\mathbf{h} \tag{58}$$

where

$$\mathbf{D}(\mathbf{x}) = \begin{bmatrix} x_0 & x_{N-1} & \ldots & x_{N-L+2} & x_{N-L+1} \\ x_1 & x_0 & \ldots & x_{N-L+3} & x_{N-L+2} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ x_{N-2} & x_{N-3} & \ldots & x_{N-L} & x_{N-L-1} \\ x_{N-1} & x_{N-2} & \ldots & x_{N-L+1} & x_{N-L} \end{bmatrix} \tag{59}$$

For the ML algorithm,the joint marginal likelihood function would be maximized as

$$(\mathbf{x}, \epsilon) = \mathbf{argmax} f(\mathbf{y}|\mathbf{x}, \epsilon) \tag{60}$$

However the direct solution for the marginal likelihood cannot be found. Thus it is desired to use EM algorithm to iteratively maximize the likelihood function.

The Rayleigh fading channel with covariance matrix $\mathbf{R}$ is being used as

$$p(\mathbf{h}) = \frac{1}{\pi^L det\mathbf{R}} \exp\left\{-\mathbf{h}^H \mathbf{R}^{-1} \mathbf{h}\right\} \tag{61}$$

By using Bayes' rule

$$
\begin{aligned}
f(\mathbf{y},\mathbf{h}|\mathbf{x},\epsilon) = & \quad f(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon)\, p(\mathbf{h}|\mathbf{x},\epsilon) \\
= & \quad f(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon)\, p(\mathbf{h})
\end{aligned} \tag{62}
$$

Now the solution of equation (60) can be found iteratively through EM algorithm as follows.

E-step:

$$
\log\{f(\mathbf{y},\mathbf{h}|\mathbf{x},\epsilon)\} = \log\{f(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon)\} + \log p\{(\mathbf{h})\} \tag{63}
$$

Putting equation (56) and (61) into equation (63), and dropping the terms independent of $\mathbf{x}$ and ϵ

$$
\begin{aligned}
&\log\{f(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon)\} \propto \log\{f(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon)\} \\
&= \log\left\{\frac{1}{\pi^N\sigma_n^2}\exp\left\{-\frac{(\mathbf{y}-\mathbf{TH}(\mathbf{h})\mathbf{x})^H(\mathbf{y}-\mathbf{TH}(\mathbf{h})\mathbf{x})}{\sigma_n^2}\right\}\right\} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{(\mathbf{y}-\mathbf{TH}(\mathbf{h})\mathbf{x})^H(\mathbf{y}-\mathbf{TH}(\mathbf{h})\mathbf{x})}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{y}\mathbf{y}^H - 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TH}(\mathbf{h})\mathbf{x}\right\} + \mathbf{x}^H\mathbf{H}(\mathbf{h})^H\mathbf{T}^H\mathbf{TH}(\mathbf{h})\mathbf{x}}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{y}\mathbf{y}^H}{\sigma_n^2} + \frac{2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TH}(\mathbf{h})\mathbf{x}\right\} - \mathbf{x}^H\mathbf{H}(\mathbf{h})^H\mathbf{H}(\mathbf{h})\mathbf{x}}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{y}\mathbf{y}^H}{\sigma_n^2} + \frac{2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TD}(\mathbf{x})\mathbf{h}\right\} - \mathbf{h}^H\mathbf{D}(\mathbf{x})^H\mathbf{D}(\mathbf{x})\mathbf{h}}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{y}\mathbf{y}^H}{\sigma_n^2} + \frac{2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TD}(\mathbf{x})\mathbf{h}\right\} - \mathrm{Tr}\left\{\mathbf{D}(\mathbf{x})^H\mathbf{D}(\mathbf{x})\mathbf{h}\mathbf{h}^H\right\}}{\sigma_n^2} \\
&\propto 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TD}(\mathbf{x})\mathbf{h}\right\} - \mathrm{Tr}\left\{\mathbf{D}(\mathbf{x})^H\mathbf{D}(\mathbf{x})\mathbf{h}\mathbf{h}^H\right\}
\end{aligned} \tag{64}
$$

Then the expectation of the joint log likelihood function is

$$\begin{aligned} Q\left(\mathbf{x}^i, \epsilon^i\right) &= E_{\mathbf{h}}\left\{\log\left\{f\left(\mathbf{y}, \mathbf{h}|\mathbf{x}^i, \epsilon^i\right)\right\}|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\} \\ &= 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)E\left\{\mathbf{h}|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\}\right\} \\ &\quad -\mathrm{Tr}\left\{\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)E\left\{\mathbf{h}\mathbf{h}^H|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\}\right\} \end{aligned} \tag{65}$$

In order to simplify the above euqation, make

$$\boldsymbol{\mu} = E\left\{\mathbf{h}|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\} \tag{66}$$

$$\mathbf{cov} = E\left\{(\mathbf{h}-\boldsymbol{\mu})(\mathbf{h}-\boldsymbol{\mu})^H|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\} \tag{67}$$

Then

$$E\left\{\mathbf{h}^H\mathbf{h}|\mathbf{y}, \mathbf{x}^{i-1}, \epsilon^{i-1}\right\} = \mathbf{cov} + \boldsymbol{\mu}\boldsymbol{\mu}^H \tag{68}$$

Now, equation (65) can be expressed as

$$\begin{aligned} Q\left(\mathbf{x}^i, \epsilon^i\right) &= 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\} - \mathrm{Tr}\left\{\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\left(\mathbf{cov} + \boldsymbol{\mu}\boldsymbol{\mu}^H\right)\right\} \\ &= 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\} - \boldsymbol{\mu}^H\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu} - \mathrm{Tr}\left\{\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\mathbf{cov}\right\} \end{aligned} \tag{69}$$

Based on eigen-decomposition

$$\mathbf{cov} = \sum \beta\mathbf{f}\mathbf{f}^H \tag{70}$$

where β is the mth eigenvalue of $\mathbf{cov}$, $\mathbf{f}$ is the corresponding eigenvector, then

$$\begin{aligned} \mathrm{Tr}\left\{\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\mathbf{cov}\right\} &= \mathrm{Tr}\left\{\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\sum\beta\mathbf{f}\mathbf{f}^H\right\} \\ &= \sum\beta\mathbf{f}^H\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\mathbf{f} \end{aligned} \tag{71}$$

Then equation (69) would be changed to

$$\begin{aligned}
Q\left(\mathbf{x}^i,\epsilon^i|\mathbf{x}^{i-1},\epsilon^{i-1}\right) &= 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\}-\boldsymbol{\mu}^H\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\\
&\quad -\sum\beta\mathbf{f}^H\mathbf{D}\left(\mathbf{x}^i\right)^H\mathbf{D}\left(\mathbf{x}^i\right)\mathbf{f}\\
&= 2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\}\\
&\quad -\mathbf{x}^{i^H}\mathbf{H}\left(\boldsymbol{\mu}\right)^H\mathbf{H}\left(\boldsymbol{\mu}\right)\mathbf{x}^i-\sum\beta\mathbf{x}^{i^H}\mathbf{H}\left(\mathbf{f}\right)^H\mathbf{H}\left(\mathbf{f}\right)\mathbf{x}^i
\end{aligned} \tag{72}$$

M-step:

In M-step, $Q\left(\mathbf{x}^i,\epsilon^i|\mathbf{x}^{i-1},\epsilon^{i-1}\right)$ in equation (69) is maximized with respect to $\mathbf{x}$ and ϵ. Dropping those terms independent of $\mathbf{x}$ and ϵ.

$$\begin{aligned}
\left(\mathbf{x}^i,\epsilon^i\right) &= \mathbf{argmax}Q\left(\mathbf{x}^i,\epsilon^i|\mathbf{x}^{i-1},\epsilon^{i-1}\right)\\
&= \mathbf{argmax}\left\{2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{T}\mathbf{D}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\}\right.\\
&\quad \left.-\mathbf{x}^{i^H}\mathbf{H}\left(\boldsymbol{\mu}\right)^H\mathbf{H}\left(\boldsymbol{\mu}\right)\mathbf{x}^i-\sum\beta\mathbf{x}^{i^H}\mathbf{H}\left(\mathbf{f}\right)^H\mathbf{H}\left(\mathbf{f}\right)\mathbf{x}^i\right\}
\end{aligned} \tag{73}$$

By setting the first derivative of $\mathbf{argmax}Q\left(\mathbf{x}^i,\epsilon^{i-1}|\mathbf{x}^{i-1},\epsilon^{i-1}\right)$ in equation (72) with respect to $\mathbf{x}$ to zero, $\mathbf{x}$ maximizing the function in the ith iteration as

$$\begin{aligned}
\hat{\mathbf{x}^i} &= \mathbf{argmax}Q\left(\mathbf{x}^i,\epsilon^{i-1}|\mathbf{x}^{i-1},\epsilon^{i-1}\right)\\
&= \left[\mathbf{H}\left(\boldsymbol{\mu}\right)^H\mathbf{H}\left(\boldsymbol{\mu}\right)+\sum\beta\mathbf{H}\left(\mathbf{f}\right)^H\mathbf{H}\left(\mathbf{f}\right)\right]^{-1}\mathbf{H}\left(\boldsymbol{\mu}\right)^H\mathbf{T}^H\mathbf{y}
\end{aligned} \tag{74}$$

Then demapping this continues data to discrete data according to the constellation algorithm, then the estimated symbols in the ith iteration are

$$\mathbf{x}^i=\mathbf{F}^{-1}\cdot \mathit{Demapping}\left\{\mathbf{F}\hat{\mathbf{x}^i}\right\} \tag{75}$$

At this point, the problem reduce to solve equation (66) and (67) which are the conditional mean and conditional correlation matrix. Both of them are derived in Appendix A as follows

$$\boldsymbol{\mu} = \left(\mathbf{R}^{-1} + \mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y} \tag{76}$$

$$\mathbf{cov} = \left(\mathbf{R}^{-1} + \mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1} \tag{77}$$

For the ith CFO, it can be estimated by using the $\mathbf{x}^i$ as

$$\begin{aligned}\epsilon^i &= \mathbf{argmax} Q\left(\mathbf{x}^i, \epsilon^i|\mathbf{x}^{i-1}, \epsilon^{i-1}\right)\\ &= \mathbf{argmax}\left\{2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TD}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\} - \mathbf{x}^{i^H}\mathbf{H}\left(\boldsymbol{\mu}\right)^H\mathbf{H}\left(\boldsymbol{\mu}\right)\mathbf{x}^i\right.\\ &\quad \left. -\sum\beta\mathbf{x}^{i^H}\mathbf{H}\left(\mathbf{f}\right)^H\mathbf{H}\left(\mathbf{f}\right)\mathbf{x}^i\right\}\end{aligned} \tag{78}$$

Since ϵ is only a real number, then only one dimension numerical search is required to get ϵ^i that maximize the Q function. Then repeat the E-step and M-step until satisfy the stopping criteria.

Convergence is a key issue for an iterative algorithm. For the above algorithm, the convergence can be proof.

Since

$$Q\left(\mathbf{x}^i, \epsilon^{i-1}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \geq Q\left(\mathbf{x}^{i-1}, \epsilon^{i-1}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \tag{79}$$

$$Q\left(\mathbf{x}^i, \epsilon^{i}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \geq Q\left(\mathbf{x}^{i}, \epsilon^{i-1}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \tag{80}$$

Then

$$Q\left(\mathbf{x}^i, \epsilon^{i}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \geq Q\left(\mathbf{x}^{i-1}, \epsilon^{i-1}|\mathbf{x}^{i-1}, \epsilon^{i-1}\right) \tag{81}$$

The Q function is decreased iteration by iteration. Thus the convergence of this algorithm is proved.

Summary of the joint algorithm:

Initialization:

At the beginning, simply setting $\epsilon^0 = 0$. Then the system model reduce to normal linear model without CFO as

$$\mathbf{y} = \mathbf{H}(\mathbf{h})\mathbf{x} + \mathbf{w} \tag{82}$$

Using LS data detector with the help of the estimated channel to have the initial $\mathbf{x}^0$ as

$$\mathbf{x}^0 = \mathbf{F}^{-1}\left(|\mathbf{H}|^2\right)^{-1}\mathbf{H}^H\mathbf{Y} \tag{83}$$

$\mathbf{F}^{-1}$ represents the IDFT. $\mathbf{H}$ and $\mathbf{Y}$ is the corresponding frequency domain CIR and received symbols, respectively.

Then compute in each iteration:

$$\boldsymbol{\mu} = \left(\mathbf{R}^{-1} + \mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y} \tag{84}$$

$$\mathbf{cov} = \left(\mathbf{R}^{-1} + \mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1} \tag{85}$$

Based on eigen-decomposition, we have

$$\mathbf{cov} = \sum \beta\mathbf{f}\mathbf{f}^H \tag{86}$$

Then

$$\hat{\mathbf{x}^i} = \left[\mathbf{H}(\boldsymbol{\mu})^H\mathbf{H}(\boldsymbol{\mu}) + \sum \beta\mathbf{H}(\mathbf{f})^H\mathbf{H}(\mathbf{f})\right]^{-1}\mathbf{H}(\boldsymbol{\mu})^H\mathbf{T}^H\mathbf{y} \tag{87}$$

$$\mathbf{x}^i = \mathbf{F}^{-1}\cdot \textit{Demapping}\left\{\mathbf{F}\hat{\mathbf{x}^i}\right\} \tag{88}$$

$$\begin{aligned}\epsilon^i = \mathbf{argmax}\Big\{&2\mathrm{Re}\left\{\mathbf{y}^H\mathbf{TD}\left(\mathbf{x}^i\right)\boldsymbol{\mu}\right\} - \mathbf{x}^{i^H}\mathbf{H}(\boldsymbol{\mu})^H\mathbf{H}(\boldsymbol{\mu})\mathbf{x}^i \\ &- \sum \beta\mathbf{x}^{i^H}\mathbf{H}(\mathbf{f})^H\mathbf{H}(\mathbf{f})\mathbf{x}^i\Big\}\end{aligned} \tag{89}$$

Finally, transferring the estimated time domain symbols to frequency domain symbols by taking DFT.

From above results, it can be seen that this algorithm can jointly detect the data and estimate the CFO in each iteration, and increase the accuracy iteration by iteration. The drawback of this algorithm is the computation complexity due to matrix inversion and eigendecomposition. Because of the eigendecomposition, a large number of matrix multiplication and addition is being performed in each iteration which dominates the complexity. Notice that this algorithm works in time domain which involve convolution, and convolution, is the reason of eigendecomposition. Thus an alternative algorithm is proposed for the same purpose but in the frequency domain in next chapter, which would not only decrease the computation complexity but also would maintain the accuracy.

6 Proposed Algorithm

In this chapter, a joint CFO estimation and data detection algorithm is proposed based on EM algorithm for OFDM system over frequency selective channel in the frequency domain. Unlike the algorithm in the last chapter which is dealing with time domain model, it is preferred to derive over frequency domain model in order to decrease the computation complexity. The frequency domain system model which is derived in chapter 3 have the following form

$$\mathbf{Y} = \mathbf{FTF}^{-1}\mathbf{HX} + \mathbf{W} \tag{90}$$

where $\mathbf{T} = diag\left\{\left[1, e^{-\frac{j2\pi\epsilon}{N}}, \ldots, e^{-\frac{j2\pi\epsilon(N-1)}{N}}\right]^T\right\}$ represents the CFO effect matrix and ϵ is the CFO. $\mathbf{Y} = [Y_0, \ldots, Y_{N-1}]^T$ is the frequency domain received symbols,$\mathbf{X} = [X_0, \ldots, X_{N-1}]^T$ is the frequency domain transmitted symbols, and $\mathbf{F}$ is the $N \times N$ DFT matrix, $\mathbf{F}^{-1} = \frac{\mathbf{F}^H}{N}$. Moreover, $\mathbf{H} = diag\left\{[H(0), \ldots, H(N-1)]^T\right\}$ is the channel frequency response matrix and also $\mathbf{H} = diag\{\mathbf{F}_L\mathbf{h}\}$,and $\mathbf{h} = [h(0), \ldots, h(L-1)]^T$ stands for CIR vector where L is the length of channel. $\mathbf{W}$ is the additive white Gaussian noise (AWGN) vector with zero means and covariance matrix $\sigma_n^2\mathbf{I}_N$. In order to simplify the model, due to $\mathbf{F}$ is a constant matrix, making $\mathbf{C} = \mathbf{FTF}^{-1}$, then the system model would be

$$\mathbf{Y} = \mathbf{CHX} + \mathbf{W} \tag{91}$$

Or equivalently as

$$\mathbf{Y} = \mathbf{CMF}_L\mathbf{h} + \mathbf{W} \tag{92}$$

where $\mathbf{M} = diag\{\mathbf{X}\}$

Since the system nonlinearly depends on CFO, thus the exact solution of the joint likelihood function is too complex to compute. As a result, EM algorithm algorithm is being employed to solve the ML problem. The idea of ML algorithm is as follows

$$(\mathbf{X}, \epsilon) = \mathbf{argmax} f\left(\mathbf{Y}|\mathbf{X}, \epsilon\right) \tag{93}$$

where $\mathbf{X}$ and ϵ are the parameters of interest.

Since the analytical solution of equation (93) cannot be derived easily, then considering EM algorithm. For the system expressed as equation (91), the joint likelihood function is denoted as $f\left(\mathbf{Y}, \mathbf{h}|\mathbf{X}, \epsilon\right)$. Denoting the received symbols $\mathbf{Y}$ as incomplete data, $\mathbf{h}$ as the missing data. Then $\{\mathbf{Y}, \mathbf{h}\}$ is being called complete data.

In E-step, forming the Q function by taking the expectation with respect to $\mathbf{h}$ for the conditional log joint likelihood function. In the M-step, updating the data by maximizing the Q function based on the symbols from the last iteration.

Since the noise in general is AWGN, then the conditional probability density function (pdf) of $\mathbf{Y}$ is

$$f\left(\mathbf{Y}|\mathbf{X}, \mathbf{h}, \epsilon\right) = \frac{1}{\pi^N \sigma_n^2} \exp\left\{-\frac{\left(\mathbf{Y} - \mathbf{CHX}\right)^H \left(\mathbf{Y} - \mathbf{CHX}\right)}{\sigma_n^2}\right\} \tag{94}$$

Moreover the pdf of Rayleigh fading channel is

$$p\left(\mathbf{h}\right) = \frac{1}{\pi^L det\mathbf{R}} \exp\left\{-\mathbf{h}^H \mathbf{R}^{-1} \mathbf{h}\right\} \tag{95}$$

where $\mathbf{R}$ denote the covariance matrix of channel which has already assumed to know in prior. The channel is independent of the CFO and transmitted symbols.

The details of the proposed algorithm would be given in the following paragraphs.

E-step:

By using Bayes' rule

$$f\left(\mathbf{Y}, \mathbf{h}|\mathbf{X}, \epsilon\right) = f\left(\mathbf{Y}|\mathbf{X}, \mathbf{h}, \epsilon\right) p\left(\mathbf{h}\right) \tag{96}$$

Then the conditional log joint likelihood function would be

$$\log\{f(\mathbf{Y},\mathbf{h}|\mathbf{X},\epsilon)\} = \log\{f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\} + \log p\{(\mathbf{h})\} \tag{97}$$

Putting equation (94) and (95) into equation (97) and dropping those terms independent of $\mathbf{X}$ and ϵ

$$\begin{aligned}
\log\{f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\} &\propto \log\{f(\mathbf{Y},\mathbf{h}|\mathbf{X},\epsilon)\} \\
&= \log\left\{\frac{1}{\pi^N\sigma_n^2}\exp\left\{-\frac{(\mathbf{Y}-\mathbf{CHX})^H(\mathbf{Y}-\mathbf{CHX})}{\sigma_n^2}\right\}\right\} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{(\mathbf{Y}-\mathbf{CHX})^H(\mathbf{Y}-\mathbf{CHX})}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{Y}^H\mathbf{Y} - 2Re\left\{\mathbf{Y}^H\mathbf{CHX}\right\} - \mathbf{X}^H\mathbf{H}^H\mathbf{HX}}{\sigma_n^2} \\
&= \log\left(\frac{1}{\pi^N\sigma_n^2}\right) - \frac{\mathbf{Y}^H\mathbf{Y}}{\sigma_n^2} + \frac{2Re\left\{\mathbf{Y}^H\mathbf{CHX}\right\} - \mathbf{X}^H\mathbf{H}^H\mathbf{HX}}{\sigma_n^2} \\
&\propto 2Re\left\{\mathbf{Y}^H\mathbf{CHX}\right\} - \mathbf{X}^H\mathbf{H}^H\mathbf{HX}
\end{aligned} \tag{98}$$

Then the expectation of the joint log likelihood function is

$$\begin{aligned}
&Q\left(\mathbf{X}^i,\epsilon^i|\mathbf{X}^{i-1},\epsilon^{i-1}\right) = E_{\mathbf{h}}\left\{\log\left\{f\left(\mathbf{Y},\mathbf{h}|\mathbf{X}^i,\epsilon^i\right)\right\}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\} \\
&= 2Re\left\{\mathbf{Y}^H\mathbf{C}\left(\epsilon^i\right)E\left\{\mathbf{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\}\mathbf{X}^i\right\} - \mathbf{X}^{i^H}E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\}\mathbf{X}^i
\end{aligned} \tag{99}$$

M-step:

In M-step, maximizing $Q\left(\mathbf{X}^i,\epsilon^i|\mathbf{X}^{i-1},\epsilon^{i-1}\right)$ in equation (99) with respect to $\mathbf{X}$

and ϵ.

$$\begin{aligned}\left(\mathbf{X}^i, \epsilon^i\right) &= \mathbf{argmax}\left\{Q\left(\mathbf{X}^i, \epsilon^i|\mathbf{X}^{i-1}, \epsilon^{i-1}\right)\right\} \\ &= \mathbf{argmax}\left\{2Re\left\{\mathbf{Y}^H\mathbf{C}\left(\epsilon^i\right)E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\mathbf{X}^i\right\}\right. \\ &\quad \left.-\mathbf{X}^{i^H}E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\mathbf{X}^i\right\} \end{aligned} \tag{100}$$

By setting the first derivative of $\mathbf{argmax}\left\{Q\left(\mathbf{X}^i, \epsilon^i|\mathbf{X}^{i-1}, \epsilon^{i-1}\right)\right\}$ in equation (98) with respect to $\mathbf{X}$ to zero, then $\mathbf{X}$ maximizing the function as

$$\hat{\mathbf{X}}^i = E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}^{-1}E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}^H\mathbf{C}\left(\epsilon^{i-1}\right)^H\mathbf{Y} \tag{101}$$

Notice that

$$\begin{aligned}E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} &= diag\left\{E\left\{\mathbf{F}_L\mathbf{h}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\right\} \\ &= diag\left\{\mathbf{F}_LE\left\{\mathbf{h}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\right\} \qquad (102) \\ E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} &= diag\left\{diag\left\{E\left\{\mathbf{F}_L\mathbf{h}\mathbf{h}^H\mathbf{F}_L^H|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\right\}\right\} \\ &= diag\left\{diag\left\{\mathbf{F}_LE\left\{\mathbf{h}\mathbf{h}^H|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}\mathbf{F}_L^H\right\}\right\} \end{aligned} \tag{103}$$

And from Appendix B

$$\begin{aligned}\mathbf{m} &= E\left\{\mathbf{h}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} \\ &= \left(\mathbf{R} - \mathbf{R}\mathbf{A}^H\left(\mathbf{B} + \mathbf{A}\mathbf{R}\mathbf{A}^H\right)^{-1}\mathbf{A}\mathbf{R}\right)\mathbf{A}^H\mathbf{B}^{-1}\mathbf{Y} \end{aligned} \tag{104}$$

$$\begin{aligned}\mathbf{COV} &= E\left\{(\mathbf{h} - \mathbf{m})(\mathbf{h} - \mathbf{m})^H|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} \\ &= \mathbf{R} - \mathbf{R}\mathbf{A}^H\left(\mathbf{B} + \mathbf{A}\mathbf{R}\mathbf{A}^H\right)^{-1}\mathbf{A}\mathbf{R} \end{aligned} \tag{105}$$

$$E\left\{\mathbf{h}^H\mathbf{h}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} = \mathbf{COV} + \mathbf{m}\mathbf{m}^H \tag{106}$$

Equation (101) can be solved now. Then demapping these continues data to discrete data according to predefined constellation scheme.

$$\mathbf{X}^i = Demapping\left\{\hat{\mathbf{X}^i}\right\} \tag{107}$$

Putting $\mathbf{X}^i$ back to equation (100), the estimated CFO of current iteration could be solved via

$$\begin{aligned}\epsilon^i = \mathbf{argmax} \quad & \left\{2Re\left\{\mathbf{Y}^H\mathbf{C}\left(\epsilon^i\right)E\left\{\mathbf{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\}\mathbf{X}^i\right\}\right. \\ & \left. -\mathbf{X}^{i^H}E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\}\mathbf{X}^i\right\} \end{aligned} \tag{108}$$

Notice that the CFO is only a real number, thus only one dimensional numerical search is required to find the ϵ^i that can result in the maximum value of fucntion $Q\left(\mathbf{X}^i,\epsilon^i|\mathbf{X}^{i-1},\epsilon^{i-1}\right)$.

Now considering initialization, in order to reduce the complexity of the proposed algorithm, simply setting the initial CFO equal to zero, and then using LS data detector to estimate the initial data. LS can be applied with any constellation algorithms. The idea is simple, but the simulation results prove that this initialization technique is feasible. Remark that, the initialization algorithm is not limited to LS, any existing data detection algorithm can be applied when CFO is equal to zero. The initialization step can be summarized as follows:

$$\epsilon^0 = 0 \tag{109}$$

$$\mathbf{X}^0 = \left(|\mathbf{H}|^2\right)^{-1}\mathbf{H}^H\mathbf{Y} \tag{110}$$

where E_s is the average power of the constellation technique.

In summary of the proposed algorithm:

Initialization;

Then compute in each iteration:

$$\mathbf{m} = \left(\mathbf{R} - \mathbf{R}\mathbf{A}^H \left(\mathbf{B} + \mathbf{A}\mathbf{R}\mathbf{A}^H\right)^{-1} \mathbf{A}\mathbf{R}\right) \mathbf{A}^H \mathbf{B}^{-1} \mathbf{Y} \tag{111}$$

$$\mathbf{COV} = \mathbf{R} - \mathbf{R}\mathbf{A}^H \left(\mathbf{B} + \mathbf{A}\mathbf{R}\mathbf{A}^H\right)^{-1} \mathbf{A}\mathbf{R} \tag{112}$$

$$\tag{113}$$

Compute the parameters of Q function as

$$E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} = diag\left\{\mathbf{F}_L \mathbf{m}\right\} \tag{114}$$

$$E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} = diag\left\{diag\left\{\mathbf{F}_L \left(\mathbf{COV} + \mathbf{m}\mathbf{m}^H\right) \mathbf{F}_L^H\right\}\right\} \tag{115}$$

Solve the Q function as

$$\mathbf{X}^i = Demapping\left\{E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}^{-1} E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\}^H \mathbf{C}\left(\epsilon^{i-1}\right)^H \mathbf{Y}\right\} \tag{116}$$

$$\begin{aligned} \epsilon^i = \; & \mathbf{argmax}\left\{2Re\left\{\mathbf{Y}^H \mathbf{C}\left(\epsilon^i\right) E\left\{\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} \mathbf{X}^i\right\}\right. \\ & \left. -\mathbf{X}^{i^H} E\left\{\mathbf{H}^H\mathbf{H}|\mathbf{Y}, \mathbf{X}^{i-1}, \epsilon^{i-1}\right\} \mathbf{X}^i\right\} \end{aligned} \tag{117}$$

The convergence of the proposed algorithm is the same as the algorithm in the last chapter. However, comparing the solution with the algorithm in last chapter, both of them have the same form of the Expectation function. But, in maximizing the Expectation function, due to the convolution in time domain, the solution of M-step in time domain is significantly more complex than the proposed algorithm in the frequency domain. The computation complexity is reduced mainly by eliminating the partial eigendecomposition which consumes a large scale computation. Specifically, the eigendecomposition in equation (86)is disappeared and replaced by simple matrix multiplication as equation (114) and (115) . In addition, the number of matrix multiplication in the decision matrix, equation (87,88 and 89) compared with equation (116 and 117), is increasingly reduced in every iteration. Furthermore, since the OFDM is a frequency domain modulation technique, then all the operations could stay in the frequency domain without transferring back and forth. Besides the computation complexity, the proposed algorithm suppose to maintain the same performance as the time domain algorithm. The simulation results in next chapter would proof this statement.

7 Simulation Results and Analysis

The simulation results present in this chapter. There are three components of it. First of all, the convergence of the proposed algorithm is being proved by calculating the BER in each iteration. Then, the BER is being investigated under different CFO in order to compare the performance of several different algorithms. The accuracy of the estimated CFO is being presented in the last part.

In this simulation, the results are averaged over 2000 runs. The simulated OFDM system has 64 subcarriers, and the source data is modulated by 16 QAM with average power equal to 1. Thus 256 binary bits can be transmitted at each symbol interval. The length of the simulated Rayleigh channel is 8, which is the same as the length of the added CP. The optimal CFO is searched within the range $[-0.1, 0.1]$. Since the proposed algorithm does not require pilot information, thus the simulation is performed without pilots. In the practical environment, if some of the transmitted data are pilots, then the real BER would be lower than this simulation results.

7.1 Convergence of the Proposed Algorithm

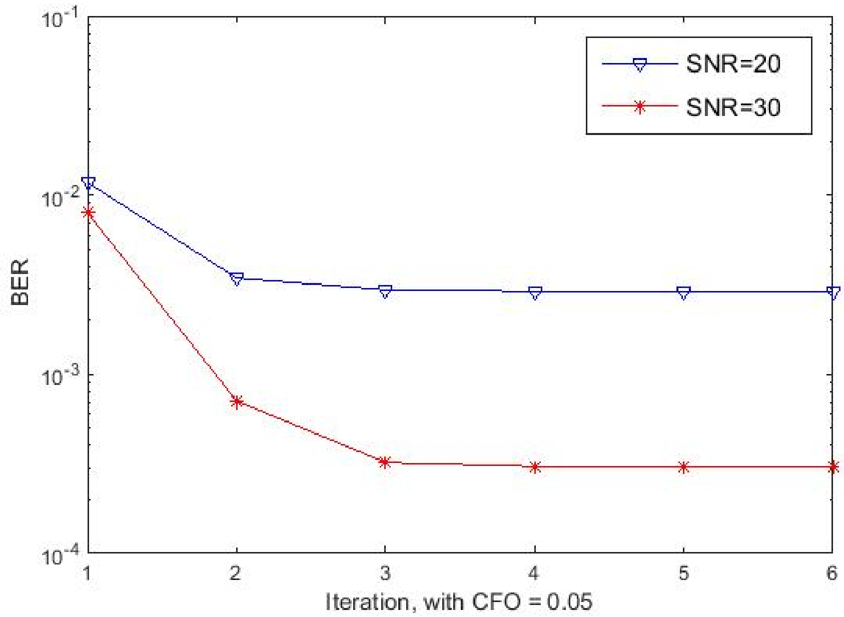

Fig. 7.1: Convergence of the proposed algorithm

BER in each iteration is being used to illustrate the convergence performance as figure 7.1. The convergence is being compared at SNR equal to 20db and 30db, respectively. The CFO is equal to 0.05 in this part. There are totally 6 iterations. The figure shows that the proposed algorithm converges within three iterations, it is the same as the time domain algorithm and the result in [36]. This result shows that in real time implementation, only three iterations are needed to stop the algorithm. Moreover, the BER under SNR equal to 30db is much lower that under 20db which means higher SNR leads to better performance.

7.2 Performance of the Proposed Algorithm

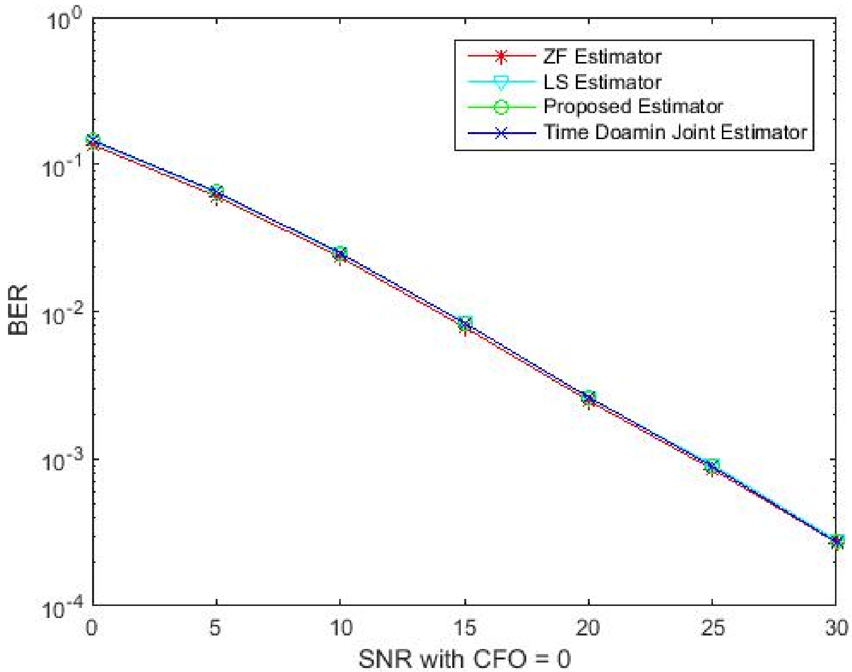

Fig. 7.2: BER with CFO=0

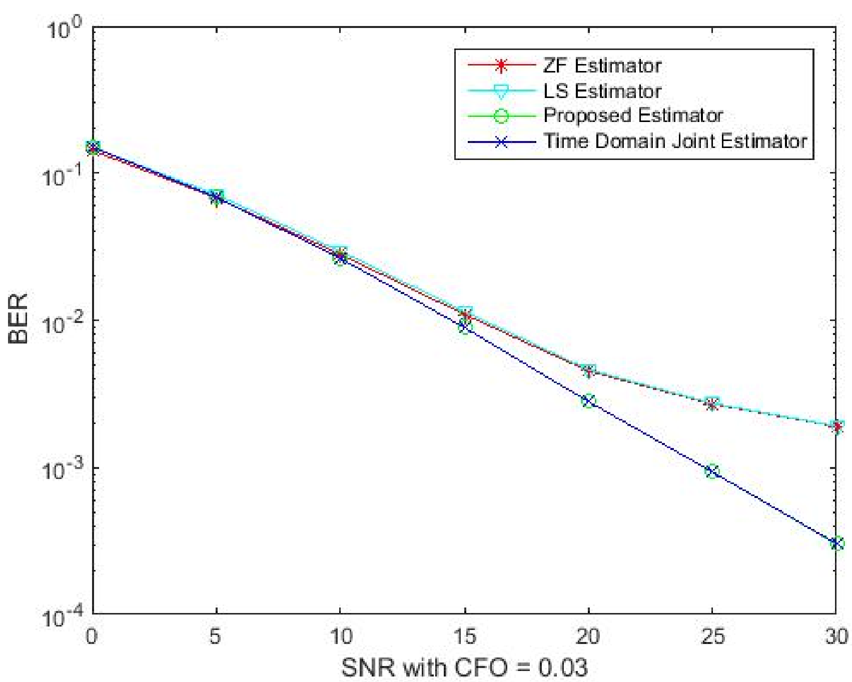

Fig. 7.3: BER with CFO=0.03

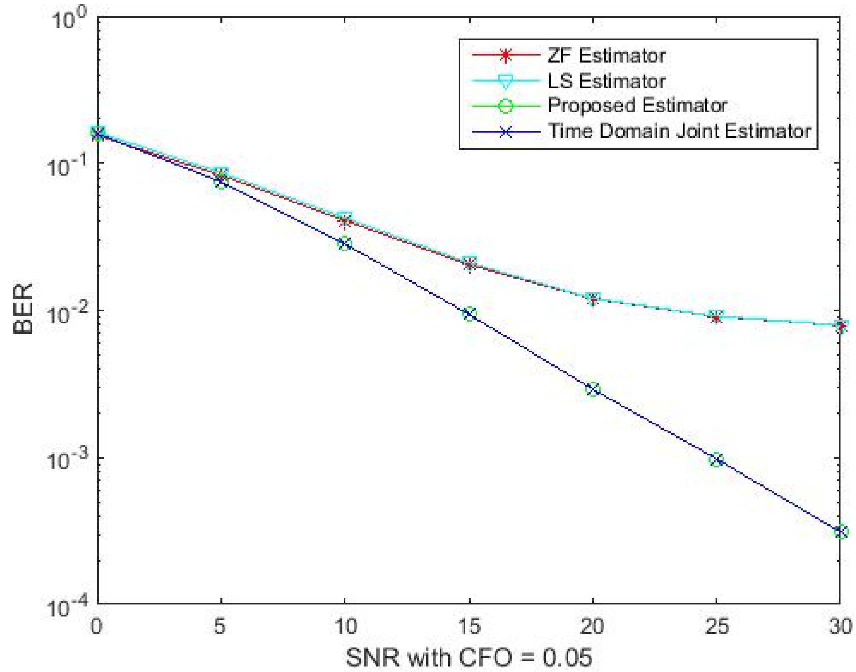

Fig. 7.4: BER with CFO=0.05

Figure 7.2, 7.3 and 7.4 compare the BER among the proposed algorithm, time domain algorithm in chapter 5, ZF equalizer and LS equalizer under different CFO. In figure 7.2, 7.3 and 7.4, the CFO is 0, 0.03, 0.05,respectively. When CFO is equal to 0, all the simulated algorithm have almost the same performance. This result shows that the proposed algorithm can also be applied to current approach as mentioned in figure 1.5. As CFO goes greater, the proposed algorithm and the algorithm in chapter 5 can significantly increase the performance of the system in compare to the other two algorithms. They can accurately detect the data when the CFO exist. The result states that current approach in figure 1.5 are sensitive to CFO while the proposed approach is robust to CFO. The gain goes larger and larger as the SNR increase. The maximum gain in this simulation is near 15 dB, which is an exciting result in wireless communication system. When comparing the proposed algorithm with the algorithm in chapter 5, the former one has lower computation complexity but preserve the same performance.

7.3 Estimated CFO

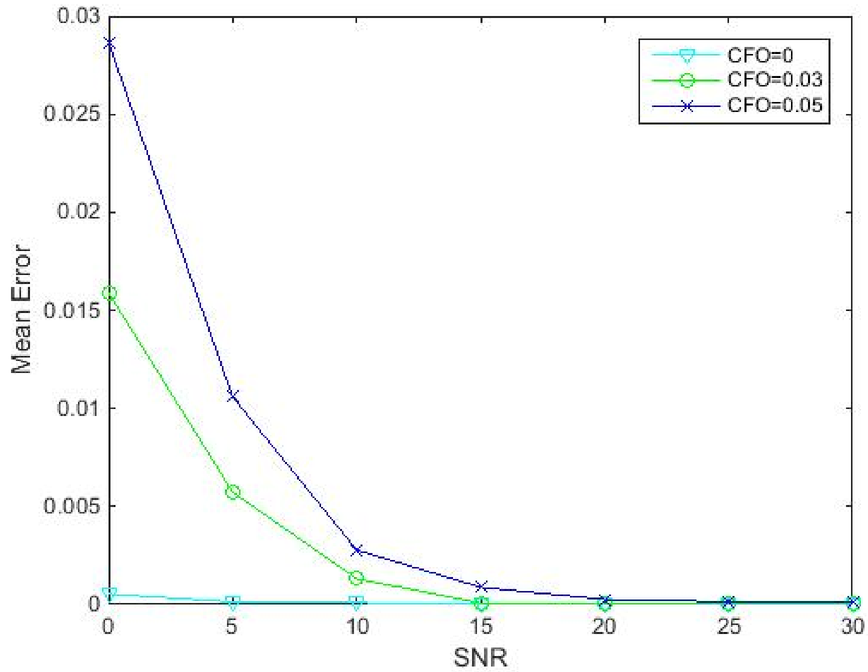

Fig. 7.5: Mean Error of Estimated CFO

Figure 7.5 shows the mean error between the estimated CFO and the real CFO. In this simulation, the real CFO is equal to 0, 0.03, 0.05, respectively. From the simulation results, it can be seen that the proposed algorithm achieves a more accurate estimated CFO when SNR goes higher than 20db. And this estimated CFO can be used in some other parts of the communication system or as a feedback to previous frequency corrector.

8 Conslusion and Further Research

8.1 Overview

In this thesis, the problem of carrier frequency offset (CFO) in orthogonal frequency division multiplexing (OFDM) systems over frequency selective channels has been addressed. To tackle this problem, first, an algorithm for joint CFO estimation and data detection in the time domain was developed. Then a new algorithm as the frequency-domain counterpart of the first one was proposed to reduce the computational complexity of the former one. Both algorithms are based on the expectation maximization (EM) algorithm which is employed as an iterative algorithm. By exploiting the channel estimate, the proposed algorithms can accurately perform CFO estimation and data detection in OFDM systems over frequency selective channels.

In Chapter 1, the background of OFDM technique was introduced. Many advantages, such as easy to deal with the fading channel, robust to intercarrier interference (ICI) and intersymbol interference (ISI), and high transmission speed, make OFDM technique being widely applied. However, due to the orthogonality of the subcarriers, synchronization is a major issue in OFDM systems. In order to overcome this issue, in most existing approaches the CFO estimation and data detection are performed in two separate parts. After frequency correction by an estimate of CFO, the transmitted data are detected based on the assumption that frequency synchronization is perfect. The detector in these approaches does not consider the estimation error associated with the CFO estimate. However, considering the limitation of the accuracy of CFO estimation, a novel approach which can jointly estimate the CFO and detect the data after frequency.

In chapter 2, a comprehensive introduction to OFDM technique and CFO effect were presented from the mathematical point of view. It was shown that the effect of CFO in the system appeared as a phase error between the transmitted symbol and the received symbol. Also, it was shown that with the assumption based on availability an accurate estimate of the CFO, the impact of CFO on the received symbol in the time domain could be completely eliminated. This was simply done by multiplying the received symbol by a correction factor.

In chapter 3, an introduction of the wireless channel was given. Based on the conditions of the wireless environment and the status of the terminals, the wireless channel was classified as frequency selective fading or flat fading and slowing fading or fast fading. With respect to popularity, the focus was on the frequency selective channel in this thesis.

In chapter 4, the estimation algorithms were introduced. By considering a general model, classical approach and Bayesian approach were compared. Then, according to the problem tackled in this thesis, the classical approach which was more appropriate to solve the problem was chosen. By comparing several classical estimation algorithms, the ML algorithm was selected. However, due to the prohibitive computational complexity of the ML sequence algorithm, the EM algorithm was considered to solve the ML problem in this thesis. Finally, the EM algorithm was introduced with the help of a comprehensive math derivation.

In chapter 5, a joint CFO estimation and data detection algorithm for OFDM in the time domain was proposed based on the EM algorithm. The idea is originally proposed in [36] which was dealing with an OFDM system in high mobility. In that case, the channel is a fast time-varying system with the channel impulse response varied sample by sample. However, the problem with that paper is that the Doppler shift cannot be classified as CFO and separate from the channel coefficients which leads to a mistake in the system model. Alternatively, if the system is under relatively slow mobility which means frequency selective channel that has more applications, then the algorithm could also solve the problem due to the CFO effect. At this point, the existing algorithm has been applied in an OFDM system with frequency selective channel. From the derivation, the transmitted data can be detected with an analytical solution while the CFO can be searched by numerical method. Nevertheless, the major problem of the algorithm is its computational complexity, which is dominated by the matrix inversion and eigendecomposition. Under this condition, in chapter 6, the frequency domain algorithm was proposed in order to reduce computational complexity.

In chapter 6, a frequency-domain algorithm for joint CFO estimation and data detection a was proposed. The proposed algorithm is also based on the EM algorithm but derives in the frequency domain in order to convert the complex convolution in the time domain to simple multiplication in the frequency domain. Under this modification, the eigendecomposition is eliminated without adding any extra operation in comparing the solution with the time domain algorithm, which means reducing computational complexity. In addition, since OFDM is a frequency domain modulation technique, then the transformation from the time domain to the frequency domain can be avoided.

In chapter 7, all the simulation results have been presented. First of all, the convergence of the proposed algorithm has been proved through the bit error rate (BER) in each iteration. After that, the performance of both the algorithms in chapter 5 and 6 have been shown in compared with the conventional approach. The simulation results show that the joint estimation algorithms are superior over the conventional algorithms up to 15 dB gain, while the proposed algorithm has the same performance as the time domain algorithm. The last part of this chapter is the accuracy of the estimated CFO. It is shown that the proposed algorithm can accurately estimate the CFO which indicates the good performance of the algorithm. As a result, The proposed algorithm can achieve the same bit-error-rate (BER) performance as that of its time-domain counterpart with much lower complexity.

8.2 Further Research

The results of this thesis can serve as the basis for the following further research:

- *Computational complexity*: It would be of theoretical and practical importance to reduce the computational complexity of the proposed algorithm.

- *Implementation issues*: In order to implement the proposed algorithm in practice, there are some key issues that were not discussed here, such as time synchronization. The start of the OFDM signal cannot be identified with sufficient accuracy, then the orthogonality of the subcarriers are destroyed.

- *Performance improvement*: The CFO estimate obtained from the proposed algorithm can be used as a feedback to increase the accuracy of the frequency alignment between the transmitter and receiver oscillators. It seems that this kind of feedback can improve the BER performance of the system.

- *CFO issue in OFDMA*: Since OFDM has multiple access capability, it always conjunct with multiple access techniques in most cellular networks. Thus, it could be very interesting research problem to develop an algorithm for joint CFO estimation and data detection in orthogonal frequency-division multiple access (OFDMA) networks.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications.* Cambridge University Press, 2005.

[2] T. Hwang, C. Y. Yang, G. WU, S. Q. Li, and G. Y. Li, "OFDM and its Wireless Applications: A Survey," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[3] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Syst. Tech. J.*, vol. 45, pp. 1775–1797, Dec. 1966.

[4] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. COM-19, no. 5, pp. 628–634, Oct.1971.

[5] L. J. Cimini and Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, no. 7, pp. 665–675, Jul. 1985.

[6] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM? An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 80–93, Nov. 2014.

[7] X. Wang and H. V. Poor, *Wireless Communication Systems: Advanced Techniques for Signal Reception.* Prentice-Hall, 2003.

[8] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 282–289, Feb. 2001.

[9] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296–306, Feb. 2004.

[10] C. R. N. Athaudage, "Ber sensitivity of ofdm systems to time synchronization error," in *Proc. 8th IEEE Int. Conf. Communication Systems, Singapore, Nov 2002*. IEEE, pp. 42–46.

[11] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[12] W.Aziz, 1E.Ahmed, G.Abbas, S.Saleem, and Q.Islam, "Performance Analysis of Carrier Frequency Offset (CFO) in OFDM using MATLAB," *Journal of Engineering (JOE)5*, vol. 1, no. 1, 2012.

[13] G. Matz and F. Hlawatsch, *Fundamentals of Time-Varying Communication Channels*. New York: Academic Press, 2011.

[14] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun*, vol. 43, no. 2, pp. 191–193, Feb. 1995.

[15] M. Morelli, C.-C. J. Kuo, and M.-O. Pun, "Synchronization Techniques for Orthogonal Frequency Division Multiple Access (OFDMA): A Tutorial Review," *Proceedings of the IEEE*, vol. 95, no. 7, July 2007.

[16] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct 1994.

[17] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 11, pp. 1613–1621, Dec 1997.

[18] M. Morelli and U. Mengali, “An improved frequency offset estimator for OFDM applications,” *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 75–77, Mar. 1999.

[19] F. Classen and H. Meyr, “Frequency synchronization algorithm for OFDM systems suitable for communication over frequency selective fading channels,” *in Proc. IEEE Veh. Technol. Conf.*, vol. 3, pp. 1655–1659, Jun. 1994.

[20] W. Zhang, X.-G. Xia, and P. C. Ching, “Clustered pilot tones for carrier frequency offset estimation in OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 101–109, Jan. 2007.

[21] Y. Li, H. Minn, N. Al-Dhahir, and A. R. Calderbank, “Pilot designs for consistent frequency-offset estimation in OFDM systems,” *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 864–877, May 2007.

[22] W. C. Kuo and M. P. Fitz, “Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading,” *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 1412–1416, Nov. 1997.

[23] M. Morelli, U. Mengali, and G. M. Vitetta, “Further results in carrier frequency estimation for transmissions over flat fading channels,” *IEEE Commun. Lett.*, vol. 2, no. 12, pp. 327–330, Dec. 1998.

[24] O. Besson and P. Stoica, “On frequency offset estimation for flat-fading channels,” *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 402–404, Oct. 2001.

[25] M. Morelli and U. Mengali, “Carrier-frequency estimation for transmission over selective channels,” *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.

[26] D. Kongkham, P.Malarvezhi, and R.Kumar, "A novel detection algorithm for performance analysis of mimo-ofdm systems using equalizer over a rayleigh fading channel."

[27] A. U. Ahmed, S. C. Thompson, and J. R. Zeidler, "Channel estimation and equalization for ce-ofdm in multipath fading channels," in *Military Communications Conference, 2008. MILCOM 2008.* IEEE.

[28] T. Cui and C. Tellambura, "Semiblind Channel Estimation and Data Detection for OFDM Systems With Optimal Pilot Design," *IEEE Trans. Commun.*, vol. 55, no. 5, MAY 2007.

[29] J. Choi, *Adaptive and Iterative Signal Processing in Communications.* Cambridge University Press, 2006.

[30] K. Lee, S. Moon, and I. Lee, "A Low-Complexity Semi-Blind Joint CFO and Data Estimation Algorithm for OFDM systems," *VTC Spring, 2012 IEEE 75th*, pp. 1–5, May 2012.

[31] Y. Na and H. Minn, "Line Search Based Iterative Joint Estimation of Channels and Frequency Offsets for Uplink OFDMA Systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 4374–4382, December 2007.

[32] X. N. Zeng, A. Ghrayeb, and D. Mao, "CFO Estimation for Uplink OFDM systems: An application of the variable projection method," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 2306–2311, May 2009.

[33] Y. Na and H. Minn, "Line Search Based Iterative Joint Estimation of Channels and Frequency Offsets for Uplink OFDMA Systems," *IEEE Transactions on Wirelesss Communications*, vol. 6, pp. 4374–4382, 2007.

[34] K. Lee, S. Moon, and I. Lee, "A Low-Complexity Semi-Blind Joint CFO and Data Estimation Algorithm for OFDM Systems," *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, vol. 6, pp. 1–5, May 2012.

[35] Y. Sun, Z. Xiong, and X. Wang, "EM-based iterative receiver design with carrier frequency offset estimation for MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 581–586, Apr. 2005.

[36] L. He, S. Ma, Y.-C. Wu, and T.-S. Ng, "Semiblind Iterative Data Detection for OFDM Systems with CFO and Doubly Selective Channels," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3491–3499, Dec.2010.

[37] T. Ha, *Theory and Design of Digital Communication Systems*. Cambridge Univ., 2010.

[38] G. Matz and F. Hlawatsch, *Fundamentals of Time-Varying Communication Channels*. New York: Academic Press, 2011.

[39] A. F. Molisch, *Wireless Communications*. Wiley, Chichester, UK, 2005.

[40] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[41] P. Hoeher, S. Kaiser, and P. Robertson, "Two-dimensional pilot symbol-aided channel estimation by wiener filtering," in *in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Apr. 1997*. IEEE, pp. 1845–1848.

[42] Y. G. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[43] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[44] S. Adireddy and L. Tong, "Optimal placement of training for frequency selective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 44, pp. 2338–2353, Aug. 2002.

[45] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138–2145, Sep. 2004.

[46] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12–25, Nov. 2004.

[47] S. M. Kay, *Fundamentals of Statistical Signal Processsing: Estimation Theory.* Prentice Hall PTR, 1993.

[48] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. Roy. Stat. Soc.*, vol. 39, pp. 1–38, 1997.

APPENDIX

A Conditional Channel Statistics in Time Domain

The conditional mean and covariance matrix, which is required by the joint CFO estimation and data detection algorithm in time domain, are derived in this chapter. The details are as follows.

By using Bayes' rules

$$\begin{aligned} f\left(\mathbf{h}|\mathbf{y},\mathbf{x},\epsilon\right) &= \frac{f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}|\mathbf{x},\epsilon\right)}{p\left(\mathbf{y}|\mathbf{x},\epsilon\right)} \\ &= \frac{f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)}{\int f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)d\mathbf{h}} \end{aligned} \tag{118}$$

Considering the time domain model as equation (55) and the property as equation (58), the time domain system model can be represented in an alternate form as

$$\mathbf{y} = \mathbf{TD}\left(\mathbf{x}\right)\mathbf{h} + \mathbf{w} \tag{119}$$

chapter 5 states that

$$\begin{aligned} f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right) &= \frac{1}{\pi^N\sigma_n^2}\exp\left\{-\frac{\left(\mathbf{y}-\mathbf{TD}\left(\mathbf{x}\right)\mathbf{h}\right)^H\left(\mathbf{y}-\mathbf{TD}\left(\mathbf{x}\right)\mathbf{h}\right)}{\sigma_n^2}\right\} \\ &= \frac{1}{\pi^N\sigma_n^2}\exp\left\{-\frac{\left(\mathbf{y}-\mathbf{TH}\left(\mathbf{h}\right)\mathbf{x}\right)^H\left(\mathbf{y}-\mathbf{TH}\left(\mathbf{h}\right)\mathbf{x}\right)}{\sigma_n^2}\right\} \end{aligned} \tag{120}$$

$$p\left(\mathbf{h}\right) = \frac{1}{\pi^L det\mathbf{R}}\exp\left\{-\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h}\right\} \tag{121}$$

In order to simplify the equation, making $\mathbf{a} = \mathbf{TD}\left(\mathbf{x}\right)$, $\mathbf{b} = \sigma_n^2\mathbf{I}_N$, then combing equation (120) and (121), then

$$f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right) = \frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\left\{-\left(\mathbf{y}-\mathbf{ah}\right)^H\mathbf{b}^{-1}\left(\mathbf{y}-\mathbf{ah}\right)-\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h}\right\} \tag{122}$$

Notice that

$$
\begin{aligned}
& \left(\mathbf{y}-\mathbf{a}\mathbf{h}\right)^H \mathbf{b}^{-1}\left(\mathbf{y}-\mathbf{a}\mathbf{h}+\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h}\right) \\
= \; & \mathbf{y}^H\mathbf{b}^{-1}\mathbf{y}-\mathbf{y}^H\mathbf{b}^{-1}\mathbf{a}\mathbf{h}-\mathbf{h}^H\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}+\mathbf{h}^H\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\mathbf{h}+\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h} \\
= \; & \mathbf{h}^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)\mathbf{h}-2Re\left\{\mathbf{h}^H\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right\}+\mathbf{y}^H\mathbf{b}^{-1}\mathbf{y} \\
= \; & \mathbf{h}^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)\mathbf{h}+\mathbf{y}^H\mathbf{b}^{-1}\mathbf{y} \\
& -2Re\left\{\mathbf{h}^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right\} \\
= \; & \mathbf{q}+\mathbf{p}
\end{aligned} \tag{123}
$$

where

$$
\begin{aligned}
\mathbf{q} = \; & \mathbf{h}^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)\mathbf{h} \\
& -2Re\left\{\mathbf{h}^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right\} \\
& +\left[\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right) \\
& \times\left[\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right] \\
= \; & \left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right) \\
& \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]
\end{aligned} \tag{124}
$$

and

$$
\begin{aligned}
\mathbf{p} = \; & \mathbf{y}^H\mathbf{b}^{-1}\mathbf{y}-\left[\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right) \\
& \times\left[\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]
\end{aligned} \tag{125}
$$

Then equation (122) can be expressed as

$$f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)=\frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\left\{-\mathbf{q}-\mathbf{p}\right\} \tag{126}$$

Equation (124) and (125) show that $\mathbf{q}$ is depending on $\mathbf{h}$, and $\mathbf{p}$ is independent of $\mathbf{h}$, thus combining equation (118) and (126), we have

$$\begin{aligned} f\left(\mathbf{h}|\mathbf{y},\mathbf{x},\epsilon\right) &= \frac{f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)}{\int f\left(\mathbf{y}|\mathbf{x},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)d\mathbf{h}} \\ &= \frac{\frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\left\{-\mathbf{q}-\mathbf{p}\right\}}{\int \frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\left\{-\mathbf{q}-\mathbf{p}\right\}d\mathbf{h}} \\ &= \frac{\exp\left\{-\mathbf{q}\right\}}{\int \exp\left\{-\mathbf{q}\right\}d\mathbf{h}} \end{aligned} \tag{127}$$

Moreover, notice that

$$\int \exp\left\{-\mathbf{q}\right\}d\mathbf{h}=\pi^L det\left(\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\right) \tag{128}$$

Combining equation (124), (127) and (128) we have

$$\begin{aligned} f\left(\mathbf{h}|\mathbf{y},\mathbf{x},\epsilon\right) &= \frac{\exp\left\{-\mathbf{q}\right\}}{\int \exp\left\{-\mathbf{q}\right\}d\mathbf{h}} \\ &= \frac{1}{\pi^L det\left(\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\right)} \\ &\quad \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right]^H\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right) \\ &\quad \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{a}^H\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^H\mathbf{b}^{-1}\mathbf{y}\right] \end{aligned} \tag{129}$$

Equation (129) presents that the conditional CIR follows a complex Gaussian

Random Distribution, then

$$\begin{aligned}
\boldsymbol{\mu} &= E\left\{\mathbf{h}|\mathbf{y},\mathbf{x}^{i-1},\epsilon^{i-1}\right\} \\
&= \left(\mathbf{R}^{-1}+\mathbf{a}^{H}\mathbf{b}^{-1}\mathbf{a}\right)^{-1}\mathbf{a}^{H}\mathbf{b}^{-1}\mathbf{y} \qquad (130)\\
\mathbf{cov} &= E\left\{\left(\mathbf{h}-\boldsymbol{\mu}\right)\left(\mathbf{h}-\boldsymbol{\mu}\right)^{H}|\mathbf{y},\mathbf{x}^{i-1},\epsilon^{i-1}\right\} \\
&= \left(\mathbf{R}^{-1}+\mathbf{a}^{H}\mathbf{b}^{-1}\mathbf{a}\right)^{-1} \qquad (131)
\end{aligned}$$

where $\mathbf{a}=\mathbf{TD}\left(\mathbf{x}\right)$, $\mathbf{b}=\sigma_n^2\mathbf{I}_N$

Applying matrix inverse lemma

$$\left(\mathbf{R}^{-1}+\mathbf{a}^{H}\mathbf{b}^{-1}\mathbf{a}\right)^{-1}=\mathbf{R}-\mathbf{R}\mathbf{a}^{H}\left(\mathbf{b}+\mathbf{a}\mathbf{R}\mathbf{a}^{H}\right)^{-1}\mathbf{a}\mathbf{R} \qquad (132)$$

B Conditional Channel Statistics in Frequency Domain

The conditional mean and correlation matrix, which is required by the proposed algorithm in chapter 6, are derived in the following paragraphs.

By using Bayes' rules

$$\begin{aligned} f(\mathbf{h}|\mathbf{Y},\mathbf{X},\epsilon) &= \frac{f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h}|\mathbf{X},\epsilon)}{p(\mathbf{h}|\mathbf{x},\epsilon)} \\ &= \frac{f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h})}{\int f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h})\,d\mathbf{h}} \end{aligned} \tag{133}$$

Considering the frequency domain system model as equation (92)

$$\mathbf{Y} = \mathbf{CMF}_L\mathbf{h} + \mathbf{W} \tag{134}$$

From chapter 6, it is known that

$$f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon) = \frac{1}{\pi^N\sigma_n^2}\exp\left\{-\frac{(\mathbf{Y}-\mathbf{CMF}_L)^H(\mathbf{Y}-\mathbf{CMF}_L)}{\sigma_n^2}\right\} \tag{135}$$

$$p(\mathbf{h}) = \frac{1}{\pi^L det\mathbf{R}}\exp\left\{-\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h}\right\} \tag{136}$$

To simplify the expression, making $\mathbf{A} = \mathbf{CMF}_L$,and $\mathbf{B} = \sigma_n^2\mathbf{I}_N$, then combining equation (135) and (136), then

$$\begin{aligned} f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h}) &= \frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}} \\ &\quad\times\exp\left\{-(\mathbf{Y}-\mathbf{A}\mathbf{h})^H\mathbf{B}^{-1}(\mathbf{Y}-\mathbf{A}\mathbf{h})-\mathbf{h}^H\mathbf{R}^{-1}\mathbf{h}\right\} \end{aligned} \tag{137}$$

Since

$$
\begin{aligned}
& \left(\mathbf{Y}-\mathbf{A}\mathbf{h}\right)^{H}\mathbf{B}^{-1}\left(\mathbf{Y}-\mathbf{A}\mathbf{h}+\mathbf{h}^{H}\mathbf{R}^{-1}\mathbf{h}\right) \\
= \;& \mathbf{Y}^{H}\mathbf{B}^{-1}\mathbf{Y}-\mathbf{Y}^{H}\mathbf{B}^{-1}\mathbf{A}\mathbf{h}-\mathbf{h}^{H}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}+\mathbf{h}^{H}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\mathbf{h}+\mathbf{h}^{H}\mathbf{R}^{-1}\mathbf{h} \\
= \;& \mathbf{h}^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)\mathbf{h}-2Re\left\{\mathbf{h}^{H}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right\}+\mathbf{Y}^{H}\mathbf{B}^{-1}\mathbf{Y} \\
= \;& \mathbf{h}^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)\mathbf{h}+\mathbf{Y}^{H}\mathbf{B}^{-1}\mathbf{Y} \\
& -2Re\left\{\mathbf{h}^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right\} \\
= \;& \mathbf{Q}+\mathbf{P}
\end{aligned}
\tag{138}
$$

where

$$
\begin{aligned}
\mathbf{Q} = \;& \mathbf{h}^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)\mathbf{h} \\
& -2Re\left\{\mathbf{h}^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right\} \\
& +\left[\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right]^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right) \\
& \times\left[\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right] \\
= \;& \left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right]^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right) \\
& \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right]
\end{aligned}
\tag{139}
$$

and

$$
\begin{aligned}
\mathbf{P} = \;& \mathbf{Y}^{H}\mathbf{B}^{-1}\mathbf{Y}-\left[\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right]^{H}\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right) \\
& \times\left[\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y}\right]
\end{aligned}
\tag{140}
$$

Then equation (137) can be expressed as

$$
f\left(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon\right)p\left(\mathbf{h}\right)=\frac{1}{\pi^{N+L}\sigma_{n}^{2}det\mathbf{R}}\exp\left\{-\mathbf{Q}-\mathbf{P}\right\}
\tag{141}
$$

Equation (139) and (140) show that $\mathbf{Q}$ is depending on $\mathbf{h}$, and $\mathbf{P}$ is independent of $\mathbf{h}$, thus combining equation (133) and (141), we have

$$\begin{aligned} f(\mathbf{h}|\mathbf{Y},\mathbf{X},\epsilon) &= \frac{f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h})}{\int f(\mathbf{Y}|\mathbf{X},\mathbf{h},\epsilon)\,p(\mathbf{h})\,d\mathbf{h}} \\ &= \frac{\frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\{-\mathbf{Q}-\mathbf{P}\}}{\int \frac{1}{\pi^{N+L}\sigma_n^2 det\mathbf{R}}\exp\{-\mathbf{Q}-\mathbf{P}\}\,d\mathbf{h}} \\ &= \frac{\exp\{-\mathbf{Q}\}}{\int \exp\{-\mathbf{Q}\}\,d\mathbf{h}} \end{aligned} \tag{142}$$

Notice that

$$\int \exp\{-\mathbf{Q}\}\,d\mathbf{h} = \pi^L det\left(\left(\mathbf{R}^{-1}+\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\right) \tag{143}$$

Putting equation (139) and (143) into equation (142)

$$\begin{aligned} f(\mathbf{h}|\mathbf{Y},\mathbf{X},\epsilon) &= \frac{\exp\{-\mathbf{Q}\}}{\int \exp\{-\mathbf{Q}\}\,d\mathbf{h}} \\ &= \frac{1}{\pi^L det\left(\left(\mathbf{R}^{-1}+\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\right)} \\ &\quad \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{Y}\right]^H\left(\mathbf{R}^{-1}+\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\right) \\ &\quad \times\left[\mathbf{h}-\left(\mathbf{R}^{-1}+\mathbf{A}^H\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{B}^{-1}\mathbf{Y}\right] \end{aligned} \tag{144}$$

From the above equation we can see that the conditional pdf of CIR follows

Complex Gaussian Distribution, then

$$\begin{aligned}
\mathbf{m} &= E\left\{\mathbf{h}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\} \\
&= \left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{Y} \qquad (145) \\
\mathbf{COV} &= E\left\{(\mathbf{h}-\mathbf{m})(\mathbf{h}-\mathbf{m})^{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\} \\
&= \left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1} \qquad (146) \\
E\left\{\mathbf{h}\mathbf{h}^{H}|\mathbf{Y},\mathbf{X}^{i-1},\epsilon^{i-1}\right\} &= \mathbf{COV}+\mathbf{m}\mathbf{m}^{H} \qquad (147)
\end{aligned}$$

where $\mathbf{A} = \mathbf{CMF}_L$,and $\mathbf{B} = \sigma_n^2\mathbf{I}_N$

Applying matrix inverse lemma

$$\left(\mathbf{R}^{-1}+\mathbf{A}^{H}\mathbf{B}^{-1}\mathbf{A}\right)^{-1} = \mathbf{R}-\mathbf{R}\mathbf{A}^{H}\left(\mathbf{B}+\mathbf{A}\mathbf{R}\mathbf{A}^{H}\right)^{-1}\mathbf{A}\mathbf{R} \qquad (148)$$

VITA AUCTORIS

NAME:	Chunyu Mao
PLACE OF BIRTH:	Anshan, Liaoning, China
YEAR OF BIRTH:	1987
EDUCATION:	Northeastern University, Shenyang, Liaoning, China Bachelor of Engineering, Measurement Technology and Instrument 2006-2010 University of Windsor, Windsor ON, Canada Master of Applied Science, Electrical and Computer Engineering 2014-2016