Sparse Channel Estimation in OFDM System

Hui Xie

Thèse de Doctorat

Hui XIE

Mémoire présenté en vue de l'obtention
du grade de Docteur de l'Université de Nantes, France
Sous le label de l'Université Nantes Angers Le Mans

du grade de Docteur de South China University of Technology, Guangzhou, Chine

Discipline : Electronique
Spécialité : Télécommunications
Laboratoire : IETR UMR 6164

Soutenance le 28 mai 2014

École doctorale Sciences et Technologies de l'Information et Mathématiques (STIM), France

Estimation de canal parcimonieux pour les systèmes OFDM
Sparse Channel Estimation in OFDM System

JURY

Président :	**M. Jiayin QIN**, Professeur, Sun Yat-Sen University, Chine
Rapporteurs :	**Mme Virginie DÉGARDIN,** Professeur, Université de Lille 1, France
	M. Simin YU, Professeur, Guangdong University of Technology, Chine
Directeurs de Thèse :	**M. Suili FENG**, Professeur, South China University of Technology, Guangzhou, Chine
	M. Yide WANG, Professeur, Ecole polytechnique de l'université de Nantes, France
Co-encadrant :	**M. Guillaume ANDRIEUX**, Maître de Conférences, IUT La Roche s/Yon, France

DEDICATION

.

This dissertation is a special present to my family members and all the friends who have ever helped me and supported me:

to my wife Lu lingbo for her love, patience and understanding;

to Dad and Mom for their love, support and hardworking;

to uncles, aunts for their supports and constructive suggestions;

to my friends in Cité de Casterneau (en France Nantes): Xiao xiaoting, Qi linqing and Sun Meng for the lovely time we share together.

ACKNOWLEDGMENT

This thesis represents the work far from typing some english words on the computer, it not only shows that I have grasped the experiences of finishing a whole scientific research work but also marks the first milestone of my scientific research life. Throughout the nearly past four years I have learned how to read the related papers and extract the main views and ideas of these papers; how to analyze the views and ideas that the authors have proposed; how to find the drawbacks or limitations of the methods proposed in papers, thus improve and extend them; how to propose new ideas in one specific research area and how to finish the simulations to demonstrate the ideas.

First and foremost I wish to thank my major supervisor, professor Yide Wang, director of research of Polytech Nantes. He has been supportive ever since the days I began to work in Polytech. Thanks to his guides, I have gradually learned how to make deep investigation on my research topic and proposed novel ideas, which is actually an amazing experience for me. It is also his encouragement and help, which push me out of trap and depression. Throughout four years of cooperations with Professor Wang, I deeply felt that I have learned a lot, not just the current achievements, but the way of thinking and research.

I would also like to thank Guillaume Andrieux, my assistant supervisor, who has actually made so much work for me, including providing suggestions and ideas, reading my papers and making modifications. Thanks to his wise and reasonable guidance, I have avoided a lot troubles.

I would also like to give thanks to professor Suili Feng, my Chinese supervisor. During the passed four years, he has always cared about the progress of my research work. Either in the major directions, or in some specific details of wireless communications, he has given me a lot of helps.

I would also like to thank the professor Jiayin Qing for presiding my oral PhD defense and the professors Virginie Dégardin and Simin Yu for accepting the reviewing of the thesis and participating in my oral defense jury.

Finally, I should thank doctor Ding, who has provided me so much help during the whole PhD study.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

As our modern technology-driven civilization acquires and exploits ever-increasing amounts of data, 'everyone' now knows that most of the data we acquire 'can be thrown away' with almost no perceptual loss - witness the broad success of lossy compression formats for sounds, images and specialized technical data. The phenomenon of ubiquitous compressibility raises very nature questions: why go to so much efforts to acquire all the data when most of what we get will be thrown away? Can't we just directly measure the part that won't end up being thrown away?

—- David Donoho Compressed Sensing (2006)

1.1 Background and Motivation

Information is one of the main characteristics of our modern and rapid developing world. There is no doubt that the information revolution accelerates the pace of life of ordinary people and scientific research in various fields. Each day, we use vast amount of information, including sound, images, text data and so on. Among all these data, a considerable amount is transmitted by various kinds of wireless communication systems.

It is generally accepted that a basic wireless communication system shown in Fig 1.1.1 includes five main parts: data acquisition, data compression, data transmission, differential detection or CSI && coherent detection, data decompression.

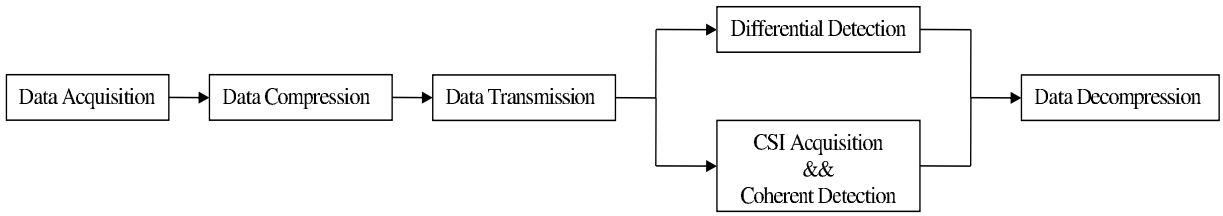

Figure 1.1.1: Basic wireless communication system

At the transmitter, data is firstly acquired for transmission. Following that, data compression is required for efficient transmission. After transmission, at the receiver, differential detection or coherent detection is employed to obtain the transmitted data. Finally, the useful data is obtained by decompression. The differential detection doesn't require the channel state information (CSI), however, it requires high SNR for the received signal. Comparatively, coherence detection requires the CSI, which can be used to reduce the impact brought by the physical channel during the transmission. Generally, coherence detection uses about 3dB less SNR than differential detection for the same BER performance [1]. In order to pursue better estimation performance, majority of communication systems adopt the coherent detection.

For coherent detection, as an effective tool to obtain CSI, channel estimation is essential at the receiver. In terms of the classification of whether to use the frequency bands or not to estimate the channel, training based channel estimation [2–9] and blind channel estimation [10–12] are the two major types of channel estimation methods. Among them, training based channel estimation is the more widely used type. Unlike the blind channel estimation, which can realize the channel estimation depending only on some specific statistics properties [10, 11], training based channel estimation can fully utilize its training sequence and pilots [13–15] to obtain effective CSI. Most training based channel estimation methods are realized

by undergoing two steps, which are channel sensing and channel reconstruction. Channel sensing process mainly focuses on how to effectively make use of limited frequency bands to fully make sense of the CSI. To realize this, optimal or suboptimal training sequence or pilot pattern arrangement should be developed in terms of different channel models. Different from the channel sensing process, channel reconstruction process concentrates on efficiently extracting CSI obtained by the channel sensing process. In this process, it is important to effectively balance the estimation performance, spectrum efficiency and the computational complexity [16], which is the guarantee for the realization of high data rate, high frequency efficiency and green communication [17, 18] in modern communications. Both of those two steps mentioned above can be and can only be effectively realized based on having good knowledge of wireless communication environment, its corresponding physical channel and the characteristics of the channel. Traditional training based channel estimation methods are effective to estimate the rich multipath channels, however, if the channels are sparse, which are demonstrated to be existed in many wireless communication environments, the traditional methods can hardly be effective [19]. The primary reason is that the sparsity of the sparse channel can hardly be explored by the current channel estimation methods. If the characteristics of sparsity of the physical channels can be fully utilized, it can actually benefit spectral efficiency, channel estimation performance and the computational complexity. As one of the major discoveries in the 21th century, compressed sensing (CS) [20, 21] theory provides an effective way to extract the CSI of the sparse channel with limited frequency bands and acceptable computational costs [19, 22]. The following three sections in this chapter mainly focus on the review of OFDM system and sparse channel propagation; the review of CS theory and sparse channel estimation in OFDM system.

1.2 Review of OFDM System and The Characteristics of Sparsity for the Wireless Channel Propagation

1.2.1 Review of OFDM System

As one of the techniques of the physical layer of wireless communication system, orthogonal frequency devision multiplexing (OFDM) [23–25] technique provides reliable solution to the future needs on high speed data transmission rate. Unlike the single subcarrier transmission system, which requires adaptive equalization to effectively reduce the intersymbol interference (ISI), OFDM system provides a much simpler way for channel equalization, moreover, the orthogonality of subcarriers can greatly improve the efficiency of transmission. Currently, OFDM technique is widely accepted for various standards, including IEEE 802.11, IEEE 802.15, IEEE 802.16, Long Term Evolution (LTE), digital audio broadcasting (DAB) and digital video broadcasting (DVB) etc. As one of the key techniques in OFDM system, channel estimation is a challenging task and also the guarantee for high date rate transmission.

1.2.2 The Wireless Channel Propagation and its Characteristics of Sparsity

In practical wireless communication environment, the radiowaves are transmitted through different paths, which leads to significant gains at different arrivals of delay time at the receiver and it is the multipath propagation. For the wireless multipath channel propagation, fading is an important characteristic. Fading reflects the power attenuation of different arriving paths, which is the main casue of high error probability of the transmitted data at the receiver [26]. To combat against fading, the diversity gains of wireless channels can be explored, including, delay, Doppler, angle of arrival (AOA) and angle of departure

(AOD) [19, 27]. In terms of the number of the effective channel degrees of freedom (DOF), wireless channels can be classified into two categories, the rich multipath channel and sparse multipath channel [19]. The wireless channel is a rich multipath channel, if the number of DOF is expected to scale linearly with the channel dimension, while if the minority of the channel DOF is expected to be effective, the channel is a sparse multipath channel [19]. In some wireless communication environments, for example, in the indoor environment, the time interval of adjacent arrival path is not very big compared to one sampling interval (sampling interval is related with the system frequency band, which should not be wide and ultra wideband (UWB) systems are not considered here.), therefore, the channel presents approximately the characteristics of rich multipath channel [28, 29]. While for the rest of majority wireless communication environments, the time interval of adjacent arrival paths is comparatively high with respect to one sampling interval. The channel tends to exhibit the characteristics of sparse multipath channel. Many practical communication environments can be considered as sparse channels, such as, the high definition television (HDTV) channels [30], the underwater acoustic channels [22, 31], some particular urban channels for long term evolution LTE system [32], ITU-R vehicular channels (channel A and channel B) [33] and some of the SUI channels (SUI4-SUI6) [34] et al. Additionally, in terms of whether the significant taps of CIR locate on the integer sampling points, wireless sparse channel can be classified into the sample spaced sparse channel [35–38] and non-sample spaced channel [22, 39, 40].

1.3 Review of Compressed Sensing Theory

In the digital information age, various analog signals are converted in to digital signals, stored, transmitted and processed. As a key technique for acquiring the digital signal, analog to digital converter (ADC) is essential. Traditional sampling technique is generally

based on the Shannon-Nyquist sampling theorem [41, 42], which requires that the sampling frequency is at least twice times of the frequency band of the signal. Indeed, Shannon-Nyquist sampling theorem provides us a sufficient condition for signal recovery. However, Is it a necessary condition? Compressed sensing (CS) [20, 21, 43] theory gives a negative answer.

In the year 2004, Donoho, Candes, Romberg and Tao firstly proposed the concept of CS and demonstrated that a signal with the characteristics of sparsity can be exactly recovered from a small number of linear, nonadaptive and incoherent measurements [20,21]. This result fully reveals that a sparse signal can possibly be reconstructed from far few linear measurements. Now, CS has an improved theory based on the combination of matrix analysis, statistics, combinatorics optimizations and operational research etc. In applications, CS covers a wide-ranging from astronomy, biology, wireless communications, pattern recognition, radar, audio, image and video processing etc.

Before considering the CS theory, it is necessary to review various transforms, such as, Fourier transform [44], discrete cosine transform [45], wavelet transform [46] and so on. In signal processing field, transforms provide effective tools for us to deal with signals with different characteristics. For example, Fourier transform is usually considered in wireless communication signals, while wavelet transform is employed for processing audio, image and video signals etc. After carefully studying all those transforms, it is shown that many signals can be expressed by sparse or approximately sparse signals in specific domains. For instance, as mentioned before, many physical wireless channels have the sparse CIR in time domain, audio, image, video signals exhibit sparse characteristics in wavelet domain [46]. Generally, the bases of those sparse transforms are redundant, which is also called redundant dictionary [47].

Measurement matrix ($M \times N \quad M < N$) design is the first but one of the most important steps in CS, which has deep impacts on the precision of the sparse signal reconstruction. Generally,

the components of the measurement matrix can subject to i.i.d Gaussian distribution [48,49] or Bernoulli distribution [49]. The measurement matrix can also be the partial Fourier matrix or its derived partial Fourier matrix [19,50]. The measurement matrix can also be the matrix with its components subject to i.i.d Gaussian distribution combined with the matrix of a specific transform [22, 51].

Once the measurement matrix is designed, the sparse signal can be recovered by different reconstruction algorithms. There are mainly three categories of methods. The first type is composed of the l_1 norm based methods, which solves the l_1 norm minimization, such as Basis pursuit (BP) [52,53], Dantzig selector (DS) [54] and Lasso [55]. The second solves the l_2 norm minimization, greedy pursuits based methods are the most classical type, such as: Matching Pursuit (MP) or Orthogonal Matching Pursuit (OMP) [49, 56] etc. Beyond the first two types, Iterative Hard Thresholding (IHT) [57, 58] method can be an alternative.

1.4 Sparse Channel Estimation in OFDM System

For sparse channel, it is essential to explore the characteristics of sparsity of the wireless channel to realize effective channel estimation. To achieve this, several steps given in Fig 1.4.1 should be carried out.

As can be seen from Fig 1.4.1, there are roughly four steps for sparse channel estimation in OFDM system. Firstly, the sparse channel model considered should be clearly defined. Generally, there are sparse channel models like, sparse frequency selective channel, sparse doubly selective channel, sample spaced and non-sample spaced channels etc. After that, analysis should be given to the selected channel model, which is actually important for the remaining two steps—- measurement matrix design and sparse channel recovery. Once the analysis on the sparse channel model has been finished, measurement matrix should be constructed properly to match with the channel model and satisfy the communication demands

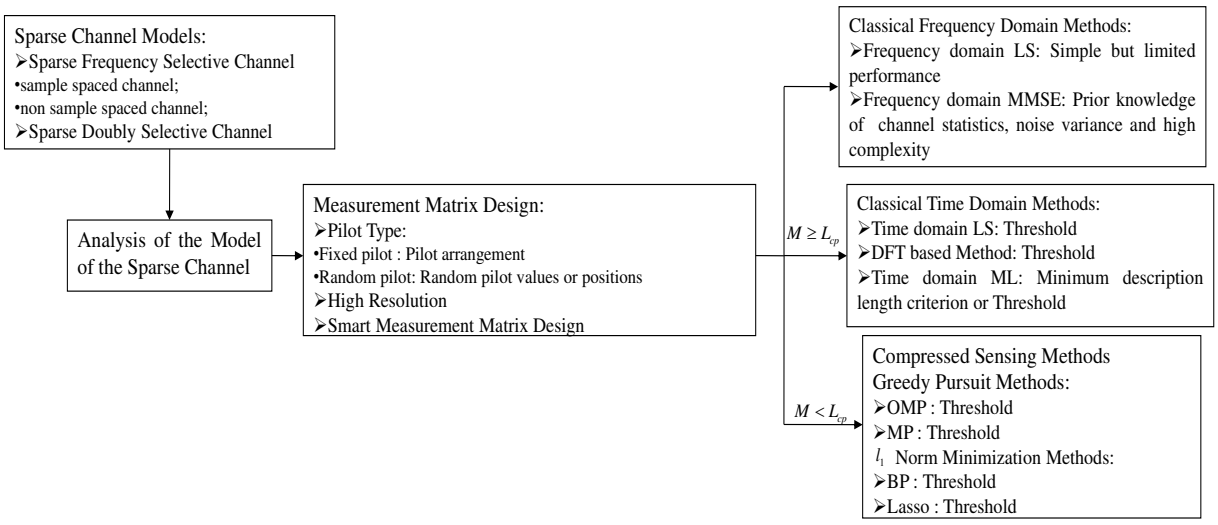

Figure 1.4.1: Framework for sparse channel estimation in OFDM system

(effectively balance among frequency band consumption, computational complexity and estimation performance). Finally, different sparse channel estimation methods are employed to recover the sparse channels in different channel communication environments. Among these four steps, measurement matrix design and sparse channel recovery are the two main focuses in research.

1.4.1 Measurement Matrix Design

The design of measurement matrix is an important work for CS based sparse channel estimation. In the process of measurement matrix design, there are several factors which should be considered. The first factor is the pilot types, generally, fixed pilot [59, 60] and random pilot [61, 62] are the most common two. Fixed pilot usually has fixed value and their positions are determined according to some specific pilot arrangement methods, comparatively, random pilot usually has random value [62] or random pilot arrangements [61]. For fixed pilot, the pilot arrangement is important in the measurement matrix design. Different pilot

arrangement methods are proposed in [50, 59, 60, 63] to realize effective sparse channel estimation. While for random pilot, there is no need to make pilot arrangement. The second factor is the measurement matrix with high resolution [22], which is essential in many cases like non-sample spaced sparse channel [22] and doubly selective channel [31, 64–66]. The third factor is the smart measurement matrix. In many practical channel estimation cases, redundant measurement matrix is effective for CS recovery, however, it can cause huge computational complexity, especially for sparse channel. In order to reduce the computational complexity, smart measurement matrix should be used.

1.4.2 Sparse Channel Recovery

Channel estimation methods are crucial to realize effective sparse channel recovery. Generally, channel estimation can be classified into two categories, frequency domain channel estimation and time domain channel estimation.

Frequency domain least squares (LS) and minimum square error (MMSE) [67, 68] are two major classical frequency domain channel estimation methods. Generally, frequency domain MMSE method can achieve better channel estimation performance than frequency LS method, but frequency domain MMSE method requires the prior knowledge of channel statistics and noise variance. Moreover, it has an increased computational complexity [67, 69]. Comparatively, the complexity of frequency domain LS is low and it is popular to combine LS method with different interpolation algorithms to realize effective channel estimation [15, 69, 70]. For example, linear interpolation, second-order interpolation, low pass interpolation etc are commonly used types. Among them, linear interpolation with lowest computational cost is the most classical one [15, 69–72], but its performance relies on comparatively high percentage of pilots and it is vulnerable to noise, especially when the channel is sparse. Frequency domain channel estimation methods have the advantages of the convenience of equalization, however, these methods estimate the channel without con-

sidering its characteristics of sparsity of the physical channel, therefore, the characteristics of channel sparsity are not explored by those methods.

Besides the frequency domain channel estimation methods, time domain channel estimation methods are also popular in channel estimation, especially in sparse channel estimation. Time domain LS method [19, 67], DFT based method [40, 73] and time domain maximum likelihood (ML) method [39, 74] are the three major classical methods. Different from frequency domain methods, time domain methods are sensitive to the maximum delay of the sparse channel, in other words, the classical time domain channel estimation methods can only be employed under the assumption that $M \geq L_{cp}$. (M is the number of pilots and L_{cp} is the length of cyclic prefix, which is used to prevent inter-symbol interference.) Additionally, time domain channel estimation methods are effective in estimating rich multipath channel, however, if the channel is sparse, threshold is essential to promote the channel estimation performance [37, 40, 73].

However, compared with classical time domain channel estimation methods, CS reconstruction methods fully explore the characteristics of the channel sparsity, which can effectively balance the channel estimation performance spectral efficiency and computational complexity. Different CS reconstruction algorithms may have different channel estimation performance. Currently, majority of channel reconstruction methods are based on two types. The first type is l_1 norm minimization methods [22, 61, 64, 65, 75], BP [22] and Lasso [65] are the two commonly used methods. The second type is greedy pursuit algorithms [22,30,32,76–82], such as MP and OMP. Comparatively, greedy pursuit algorithms are more popular due to its low complexity and convenient realization [32]. In the early stages, MP algorithm was employed to estimate sparse channel with improved estimation performance [30, 76–78, 80]. Recently, OMP method gradually replaces MP method to realize more effective sparse channel estimation [22,32,81] in OFDM system. For MP, OMP, BP and Lasso methods, threshold is necessary to guarantee the estimation performance.

1.5 Major Contributions

In order to improve the sparse channel estimation performance to satisfy the needs of the future communication system, we have developed several novel sparse channel estimation methods focusing on the goal of effective balance among the channel estimation performance, spectral efficiency, computational complexity and the burden of the whole communication system. The ideas behind all the proposed methods are motivated by fully considering the characteristics of sparse channel and exploring its special inherent quality. In the rest of this section, we highlight the contributions of our research work.

1) Based on the analysis of both the characteristics of channel sparsity and noise, an effective threshold is proposed to effectively improve the sparse channel estimation performance in the case of $M \geq L_{cp}$. Compared with other existing methods, the proposed method can realize effective sparse channel estimation without the prior knowledge of channel statistics and the noise standard deviation.

2) A novel effective threshold is proposed for CS based sparse channel estimation in the case of $M < L_{cp}$. The proposed method makes full use of the characteristics of channel sparsity to estimate the noise standard deviation in the case of $M < L_{cp}$, which is essential for a threshold estimation without the prior knowledge of channel statistics and noise standard deviation. Both theoretical derivation and simulation results demonstrate that the proposed method can significantly improve the channel estimation performance as well as the spectral efficiency for sparse channel with different channel sparsity.

3) For non-sample spaced sparse channel, we have derived the theoretical result on the relationships between oversampling factor R and leakage of the observed CIR at the receiver. Additionally, suboptimal pilot arrangement strategy is considered for the case of $M < L_{cp}$ and the concept of smart measurement is introduced to effectively balance the channel estimation performance, computational complexity and spectral efficiency.

1.6 Organization

The rest of the thesis is organized as follow:

In chapter 2, channel propagation and channel estimation in OFDM system are considered. Firstly, the characteristics of the effect of channel propagation including the path loss, shadowing and multipath fading are discussed. Following that, the structure and the mathematical model of OFDM system are analyzed. Finally, traditional LS, MMSE and ML channel estimation methods are introduced and compared in both the cases of rich multipath and sparse multipath channels.

In chapter 3, CS principles and the models based on CS for sparse channel estimation are presented. The introduction of CS, including its basic ideas, measurement matrix and recovery conditions are firstly discussed, then, different recovery algorithms are presented. In the end, based on the principles of CS, different measurement matrices are constructed to match with different channel models.

In chapter 4, effective threshold for LS based sparse channel estimation is addressed. We firstly introduce the frequency interpolated based channel estimation and analyze its performance. Comparatively, time domain LS can overcome the drawbacks of the frequency interpolated based channel estimation ($M \geq L_{cp}$), however, it is highly affected by noise when estimating the sparse channel. After carefully analyzing the noise characteristics, median absolute deviation (MAD) based noise standard deviation estimation method is full investigated. Based on the MAD method, a novel effective time domain threshold is derived according to the estimated noise standard deviation and without requiring the prior knowledge of both channel statistics.

In chapter 5, CS based sample spaced and non-sample spaced sparse channel estimation are investigated. As previous mentioned, the sparse channel can be classified into sample spaced sparse channel and non-sample spaced sparse channel. For sample spaced sparse channel, the

case where $M < L_{cp}$ cannot be solved by classical LS, however, it is an interesting research topic in CS [60]. By making fully use of the CS reconstruction in sparse channel estimation, an effective threshold estimation method is proposed without any prior knowledge of channel statistics and noise standard deviation. The estimated threshold can effectively improve the sparse channel estimation performance. For the non-sample spaced sparse channel, by introducing the smart measurement matrix, which is constructed under the guide of rough detection on the hot zones, effective channel estimation can be obtained with a small number of pilots ($M < L_{cp}$) and comparatively low computational complexity.

Chapter 2

Channel Estimation in OFDM System

To realize effective channel estimation in OFDM system, the characteristics of both the wireless channel and OFDM system are fundamental. In this chapter, we firstly focus on the effect of channel propagation, then, OFDM system model including its mathematical description is presented. After that, classical channel estimation methods are discussed. Finally, simulation comparisons are given for classical channel estimation methods for both the rich multipath and sparse multipath channels.

2.1 Effect of Channel Propagation

Wireless radio channels present huge challenge for modern high-speed and reliable communication since they are not only vulnerable to interferences and noise, but also they are highly affected by impediments, Doppler Shift (Moving Environment) and changing communication environment dynamics. The transmitted radio signals are likely to be diffracted, scattered, reflected and attenuated during their transmission, therefore, when the radio signals reach the receiver, they may have different transmission paths and random phases, furthermore their power is attenuated.

During the transmission, there are different scales of fading. Roughly, the fading of wireless channel propagation can be classified into large scale fading, which is generally used to characterize the intensity of the transmitted signal over comparatively long transmitter-receiver (T-R) separation distance with several hundreds or thousands meters and small scale fading, which is usually employed to characterize the transmitted signal intensity over several wavelengths. For large scale fading, the main contributors are path loss and shadowing. While for small scale fading, the main contributor is multipath fading. In the following, the characteristics of both large scale fading and small scale fading of wireless channel are presented.

2.1.1 Path Loss in Free Space

As one of the important large scale fadings, path loss is caused by the dissipation of the power radiated by the transmitter as well as by the effect of the natural propagation path. Path loss in free-space, which assumes that the line of sight (LOS) transmission path without any obstruction is the simplest model. In practice, satellite communication system propagates radiowaves, which typically undergo free space propagation. Considering a free space propagation with distance d, the mathematical relation between the transmitted power and the received power is given by the Friis equation [83, 84], expressed as:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \tag{2.1.1}$$

where $P_r(d)$ is the received power, which falls proportionally to the square of d, P_t is the transmitted power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, λ is the wavelength (meter), L is the system loss factor without relation to propagation ($L \geq 1$).

Consider the average signal to noise ratio (SNR), given by:

$$SNR = \frac{\bar{P}_r}{P_n} = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \frac{1}{N_0 B} \tag{2.1.2}$$

where N_0 is the unilateral noise power spectral density, B is the bandwidth of the signal, which is composed in the case of OFDM of the bandwidth of pilots (B_p) in OFDM system and useful transmitted data bandwidth (B_d). Therefore, (2.1.2) can be rewritten as:

$$SNR = \frac{\bar{P}_r}{P_n} = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \frac{1}{N_0 (B_p + B_d)} \tag{2.1.3}$$

Generally, each communication system has a required minimum SNR, which can be expressed by SNR_0 and we should have $SNR \geq SNR_0$. Therefore, the useful data bandwidth B_d has an upper bound, which can be expressed by:

$$B_d \leq \frac{P_t G_t G_r \lambda^2}{N_0 (4\pi d)^2 L} \frac{1}{SNR_0} - B_p \tag{2.1.4}$$

Therefore, both pilots (B_p) and the path loss will generate the consumption of bandwidth. The loss of bandwidth due to path loss can not be avoided. However, the consumption of bandwidth by pilots can be reduced if the sparse channel characteristics are properly explored. This actually can be realized by effective pilot arrangements [50, 59, 60].

2.1.2 Shadowing

Besides path loss, shadowing is another major large scale fading. The primary reason for the cause of shadowing is that the intensity of radiowave suffers a typical random loss caused by the electromagnetic field shadowing due to the blockage of different objects, such as mountains, buildings and so on. A typical model characterizing this random variations of

power is the log Gaussian distribution (log Gaussian shadowing), which can be expressed by the following expression [84]:

$$p(r) = \frac{1}{\sqrt{2\pi}\sigma r} exp\left(-\frac{(\ln r - \mu)^2}{2\sigma^2}\right) \tag{2.1.5}$$

where r is the received signal amplitude, which is assumed to be random. Additionally, the log-normal distribution is determined by μ and σ, which are the mean value and standard deviation of the received complex random signal respectively.

2.1.3 Multipath Fading Channel

Different from large scale fading, small scale fading reflects the changes of the received signal in several wavelengths. Multipath fading is the major cause of this phenomena. During the transmission process, the radiowave is likely to be scattered and reflected around different objects, which can produce the received signal, which is the combination of different signals coming from different paths. Because of the random phases, some of the received signals with same time arrival may be strengthened due to the same or similar direction of phases, while others may be weakened by the opposite direction of phases, it is the phenomena of fading. Due to presence of independent transmission paths, the received signal is composed of different copies of original signal with different delays and attenuated fading power. The channel impulse response is composed of the sum of different discrete multipath components, which can be expressed by:

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - \tilde{\tau}_l) \tag{2.1.6}$$

where $\alpha_l \in C$ is the amplitude of lth path of the channel and $\tilde{\tau}_l$ is the lth path delay of the channel.

In (2.1.6), let $\tilde{\tau}_l = \tau_l T_s$ and T_s be the sampling interval of the communication system, (2.1.6)

can be rewritten as:

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - \tau_l T_s) \tag{2.1.7}$$

The channel frequency response (CFR) of $h\left[\tau\right]$ can be written as:

$$\begin{aligned} g[f] &= \int_{-\infty}^{\infty} h[\tau] e^{-j2\pi f\tau} \mathrm{d}\tau \\ &= \sum_{l} \alpha_l e^{-j2\pi f \tau_l T_s} \end{aligned} \tag{2.1.8}$$

Denote $f = \frac{k}{NT_s}, k = 0, 1, 2, \ldots, N-1$, the discrete form of $g[f]$ can be expressed by:

$$g[k] = \sum_{l} \alpha_l e^{-j2\pi\tau_l k/N} \tag{2.1.9}$$

After inverse discrete Fourier transform (IDFT), the observed CIR at the receiver is given by [67]:

$$\begin{aligned} h[n] &= \frac{1}{N} \sum_{l} \alpha_l \frac{1 - e^{j2\pi(n-\tau_l)}}{1 - e^{j2\pi(n-\tau_l)/N}} \\ &= \frac{1}{N} \sum_{l} \alpha_l \frac{1 - e^{-j2\pi\tau_l}}{1 - e^{j2\pi(n-\tau_l)/N}} \\ &= \frac{1}{N} \sum_{l} \alpha_l \frac{e^{-j\pi\tau_l}\left(e^{j\pi\tau_l} - e^{-j\pi\tau_l}\right)}{e^{j\pi(n-\tau_l)/N}\left(e^{-j\pi(n-\tau_l)/N} - e^{j\pi(n-\tau_l)/N}\right)} \\ &= \frac{1}{N} \sum_{l} \alpha_l e^{-j\frac{\pi}{N}(n+(N-1)\tau_l)} \frac{sin(\pi\tau_l)}{sin(\frac{\pi}{N}(\tau_l - n))} \end{aligned} \tag{2.1.10}$$

Sample Spaced Multipath Channel

In (2.1.10), when τ_l is an integer, we have $h[\tau_l] = \alpha_l$, which means that all energy of α_l is mapped to the channel tap $h[\tau_l]$ $(n = \tau_l)$. There is no leakage of power. For sample spaced channel, its channel tap positions are located exactly in the sampling points.

Non-sample Spaced Multipath Channel

In (2.1.10), when τ_l is not an integer, we have $h[\tau_l] \neq \alpha_l$, which means that not all the energy of α_l is mapped to the channel tap $h[\tau_l]$ $(n \neq \tau_l)$. There exists leakage of power to other most adjacent channel taps. For non-sample spaced channel, its channel tap positions are not only located in the sampling points.

Channel Parameters

The power delay profile is an important parameter for characterizing multipath channel. From (2.1.6), the power delay profile (PDP) versus delay can be written as:

$$p(\tau) = \sum_l E[|\alpha_l|^2]\delta(\tau - \tilde{\tau}_l) \tag{2.1.11}$$

For different delayed paths, their powers attenuate usually with the increase of delays. Mathematically, many papers, such as [22,36,67] consider that the power decreases exponentially with the delay. Specifically, the relationship can be written by:

$$E\left[|\alpha_l|^2\right] = e^{-\frac{\tilde{\tau}_l}{\tilde{\tau}_{rms}}} \tag{2.1.12}$$

where $\tilde{\tau}_{rms}$ denotes the root mean square (RMS) delay used to describe the delay spread of the channel defined as [85]:

$$\tilde{\tau}_{rms} = \sqrt{\overline{\tilde{\tau}^2} - \bar{\tilde{\tau}}^2} \tag{2.1.13}$$

where $\bar{\tilde{\tau}}$ and $\overline{\tilde{\tau}^2}$ are the channel mean delay and the channel mean square delay respectively. They are defined by [85]:

$$\bar{\tilde{\tau}}^2 = \frac{\sum_l {\alpha_l}^2 \tilde{\tau}_l}{\sum_l {\alpha_l}^2} \tag{2.1.14}$$

$$\overline{\tilde{\tau}^2} = \frac{\sum_l {\alpha_l}^2 \tilde{\tau}_l^2}{\sum_l {\alpha_l}^2} \tag{2.1.15}$$

Figure 2.1.1. shows a typical attenuated power profile with time delay for different $\tilde{\tau}_{rms}$ ($\tilde{\tau}_{rms} = 6ms, 4ms, 3ms$ and maximum delay of the channel $25ms$). As can be seen from the figure, with the decrease of $\tilde{\tau}_{rms}$, the speed of the attenuation of channel power increases.

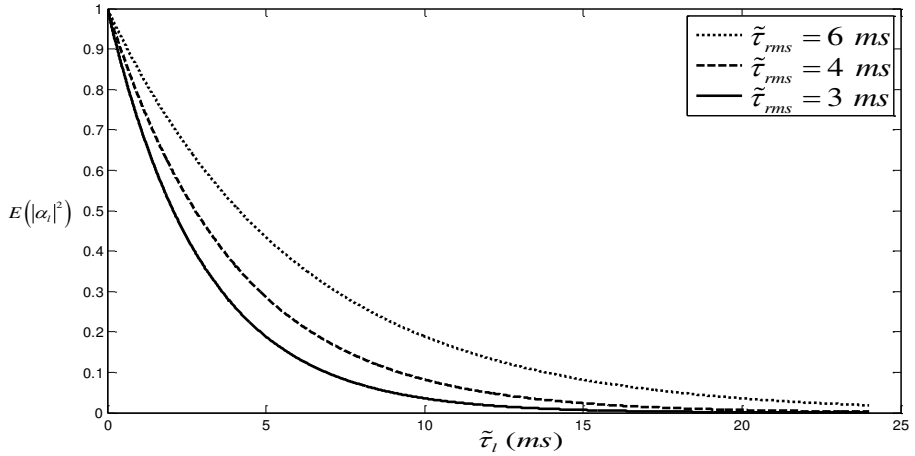

Figure 2.1.1: Attenuated power with time delays

$\tilde{\tau}_{rms}$ is an effective parameter not only for reflecting the time extent of the dispersive channel, but also for characterizing the frequency selectivity of the channel, since $\tilde{\tau}_{rms}$ is related to the average number of fades per bandwidth and to the average bandwidth of the fades [85].

Rayleigh Fading Channel

In the scattering multipath wireless communication environment, each path is the sum of large amount of incoherent random processes, in this case, the central limit theorem tells us that the real part and imaginary part independently subject to Gaussian distribution. Assume that $\alpha_l = a + bj$ and $a, b \sim N(0, \sigma^2)$. We have the joint probability density function expressed by:

$$f(a,b) = \frac{1}{2\pi\sigma^2} e^{-\frac{(a^2+b^2)}{2\sigma^2}} \tag{2.1.16}$$

In terms of polar coordinates ($r^2 = a^2 + b^2$ and $\theta = arctan(\frac{b}{a})$), the probability density function is given by [86]:

$$f(r,\theta) = \frac{r}{2\pi\sigma^2} e^{\frac{-r^2}{2\sigma^2}} \tag{2.1.17}$$

$f(r,\theta) = f(r)f(\theta)$, r and θ are two independent random variables. The phases of the scattered waves are uniformly distributed within $[0, 2\pi)$ and the distribution of the amplitude of α_l is given by:

$$f(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, r \geq 0 \tag{2.1.18}$$

where $\sigma^2 = \frac{E(|\alpha_l|^2)}{2}$. (2.1.18) is the well known Rayleigh distribution, which is one of the mostly used distributions for wireless channel.

2.2 OFDM System

2.2.1 System Model

OFDM is one of the multicarrier modulation techniques, which converts the serial high speed data stream to N independent slow data streams modulated by N orthogonal subcarriers in frequency domain, which can be shown in Fig 2.2.1, for parallel transmission. By adopting multicarrier transmission technique, the wideband frequency selective channel has been divided into different narrowband flat fading subchannels, therefore, compared with the single carrier communication systems, OFDM systems have the strong capability of anti-multipath fading and interference. Additionally, compared with the single carrier communication system, the equalization becomes much easier for OFDM system, because the bandwidth of each subchannel is only a small part of the bandwidth of original channel. Moreover, OFDM systems allow for the subcarriers aliasing in frequency domain due to the orthogonality of different subcarriers, which can promote the spectral efficiency. Fig 2.2.2 illustrates a typical baseband model for OFDM system.

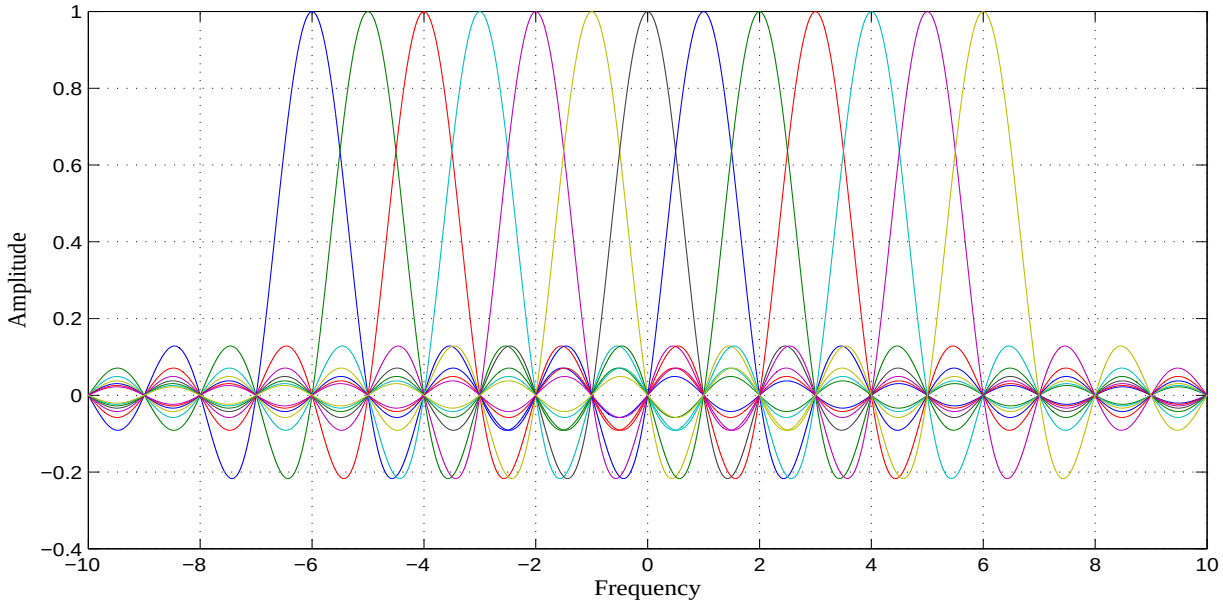

Figure 2.2.1: OFDM Subcarriers in frequency domain

In Figure 2.2.2, the high speed serial input bit stream firstly passes the QAM mapping block to realize the pre-constellation modulation. After that, serial to parallel block is used to get parallel data streams and inverse Fourier transform (IFFT) is adopted to transform the modulated data into the time domain signal. Following that, as it is shown in Fig 2.2.3, a cyclic prefix (guard interval), which a copy of a portion of the end of an OFDM symbol (l^{th} or $(l+1)^{th}$ OFDM symbol) is added in the front of its corresponding transmitted OFDM symbol to prevent the inter symbol interference (ISI). As a result, the whole OFDM symbol length (T_{sym}) is the combination of the durations of subcarriers (T_{sub}) and the duration of cyclic prefix (T_{cp})). Finally, digital to analog converter (DAC) is used to get the analog signal. During the transmission, the transmitted signal passes through the channel with additive white Gaussian noise (AWGN). At the receiver, analog to digital converter (ADC) is initially used to convert the received analog signal into digital signal, then, before removing guard interval, the serial to parallel block is employed to obtain the parallel data stream. In order to get the frequency domain signal, FFT is used, following that, channel estimation,

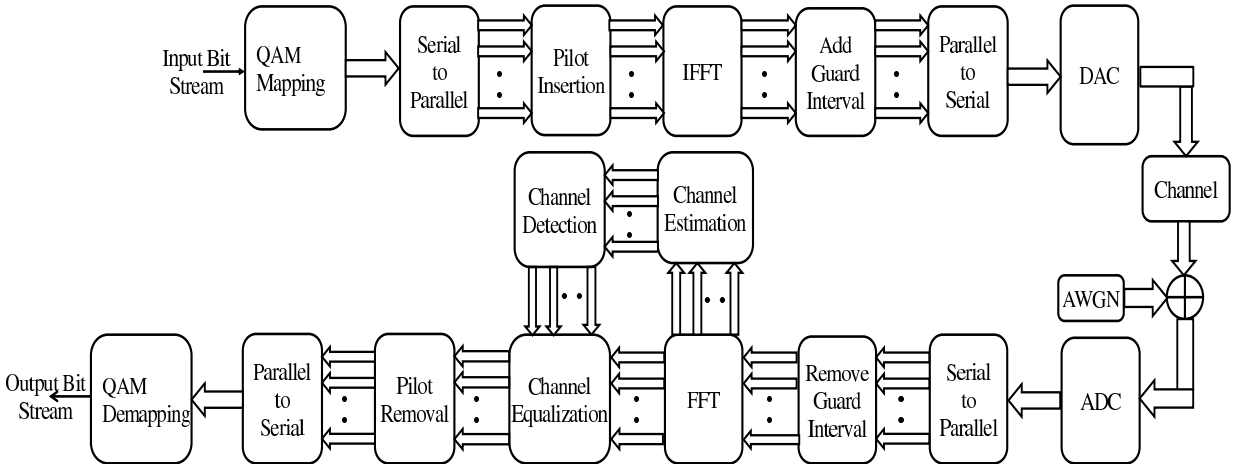

Figure 2.2.2: Baseband model for OFDM system

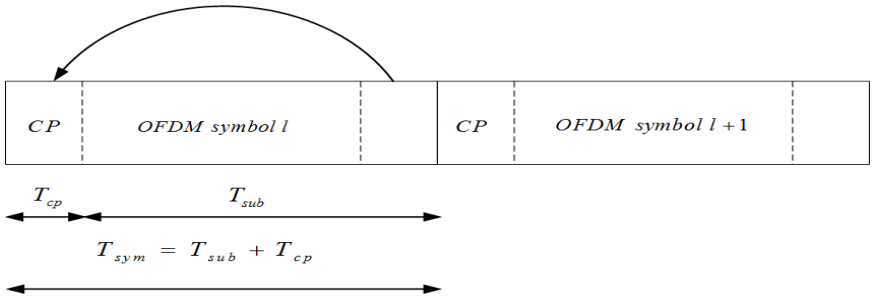

Figure 2.2.3: Cyclic prefix

detection and equalization blocks combined with the pilot removal block are employed to obtain the data symbols. Finally, the QAM demapping demodulates the data symbols to obtain the high speed serial output bit stream.

2.2.2 Mathematical Model

The basic principle behind multicarrier modulation technique is to use the multiple orthogonal sub-carriers to realize modulation and demodulation. Generally, there are two different types of subcarriers, which employ different orthogonal bases. The first one uses the Fourier

bases to realize the multi-carrier transmission [87], while the second one uses the wavelet transform [88] or wavelet packet transform [89] to realize the multi-carrier transmission. Both multi-carrier transmission systems can be realized by the equivalent filter banks at the transmitter and receiver. In this thesis, we mainly consider the first one and the realization of filter banks at both the transmitter and receiver for OFDM [87], which is illustrated in Figure 2.2.4:

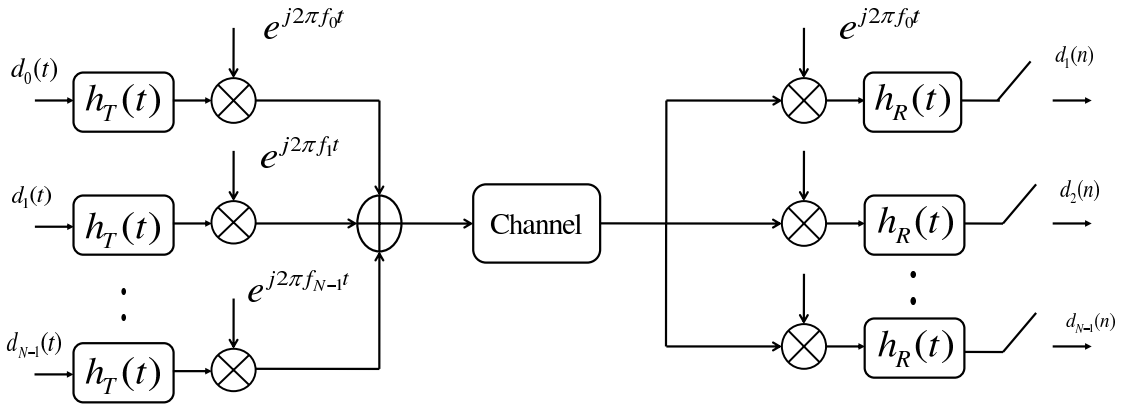

Figure 2.2.4: OFDM framework with filter banks

In Fig 2.2.4, for the transmitter, we have the input data signal:

$$d_p(t) = \sum_n d_p(n)\delta(t - nT) \tag{2.2.1}$$

where $d_p(n)$ is the data symbol with the subcarrier index p and n is the OFDM symbol index, T is one OFDM symbol duration.

With the input data signal expressed in (2.2.1) modulated by N subcarriers at the trans-

mitter, the transmitted signal is obtained, which can be expressed by [87]:

$$s(t) = \sum_{n} \sum_{p=0}^{N-1} d_p(n) h_T(t - nT) e^{j2\pi f_p t} \tag{2.2.2}$$

where $e^{j2\pi f_p t}$ is the p^{th} subcarrier with subcarrier frequency f_p. $h_T(t)$ is the rectangle signal of the transmitter without modulation (width T) given by:

$$h_T(t) = \begin{cases} 1, & 0 \le t \le T \\ 0, & else \end{cases} \tag{2.2.3}$$

In (2.2.2), assume $f_p = f_b + p\Delta f$ and $\Delta f = 1/T$. Theoretically, we have $T\Delta f = 1$ to maintain the orthogonality between the subcarriers.

Broadly, for different subcarriers and different OFDM symbols, the orthogonality is also maintained, which can be given by:

$$\begin{aligned} &\frac{1}{T}\langle h_R(t - nT)e^{j2\pi t f_q}, h_T(t - kT)e^{j2\pi t f_p}\rangle \\ &= \frac{1}{T}\int_{-\infty}^{\infty} h_R(t - nT)h_T^*(t - kT)e^{-j2\pi t(f_q - f_p)}\mathrm{d}t \\ &= \frac{1}{T}\int_{nT}^{(n+1)T} h_R(t - nT)h_T^*(t - kT)e^{-j2\pi t(q-p)\Delta f}\mathrm{d}t \\ &= \delta(n - k)\delta(p - q) \end{aligned} \tag{2.2.4}$$

where $h_R(t)$ is the rectangle signal of the receiver, which has the same width as $h_T(t)$, $\delta(\cdot)$ is the delta function defined by:

$$\delta(n) = \begin{cases} 1, & if\ n = 0 \\ 0, & else \end{cases} \tag{2.2.5}$$

(2.2.4) shows that the transmitter and receiver filter banks satisfy the biorthogonal condition.

Thus, at the receiver, by fully using the orthogonality of the subcarriers, the correct data symbols can be demodulated, which is given by:

$$\tilde{d_p}(n) = \frac{1}{T}\langle d_p(n)h_T(t-nT)e^{j2\pi t f_p}, h_R(t-kT)e^{j2\pi t f_p}\rangle \tag{2.2.6}$$

The above presentation is only for the ideal transmission situation, in which the channel impulse response is considered as a Dirac function. In practical communication, the transmission in channel can't be considered as ideal. In the following sections, we will discuss about that.

2.3 Training based Frequency Domain Channel Estimation Methods

Frequency domain least squares (LS) [67,90], minimum mean square error (MMSE) [4,67,91], time domain LS [4, 36, 67, 91] and maximum likelihood (ML) [4, 91] are the most classical channel estimation methods. In the followings, LS, MMSE and ML estimation methods are presented.

Consider a N subcarriers OFDM system. M pilots with index $k_0, k_1, \ldots, k_{M-1}$ are employed to estimate a channel. The received pilot vector can be expressed by [67, 90, 92]:

$$\boldsymbol{y}_p = \boldsymbol{X}_p \mathbf{g}_p + \boldsymbol{v}_p \tag{2.3.1}$$

where $\boldsymbol{X}_p = diag[x_{k_0}, x_{k_1}, \ldots, x_{k_{M-1}}]$ is the diagonal matrix of transmitted pilots; $\boldsymbol{y}_p = [y_{k_0}, y_{k_1}, \ldots, y_{k_{M-1}}]^T$ is the received pilot vector; $\mathbf{g}_p = [g_{k_0}, g_{k_1}, \ldots, g_{k_{M-1}}]^T$ is the partial frequency response including only the pilots locations and $\boldsymbol{v}_p = [v_{k_0}, v_{k_1}, \ldots, v_{k_{M-1}}]^T$ is the additive complex white Gaussian noise vector ($\boldsymbol{v}_p \sim CN(\mathbf{0}_M, \sigma^2 \boldsymbol{I}_M)$).

If the channel impulse response $h[\tau]$ is assumed to be static in one OFDM symbol, the system can be modeled as the transmitted pilots passing through a parallel of independent complex Gaussian channels, as shown in Fig 2.3.1:

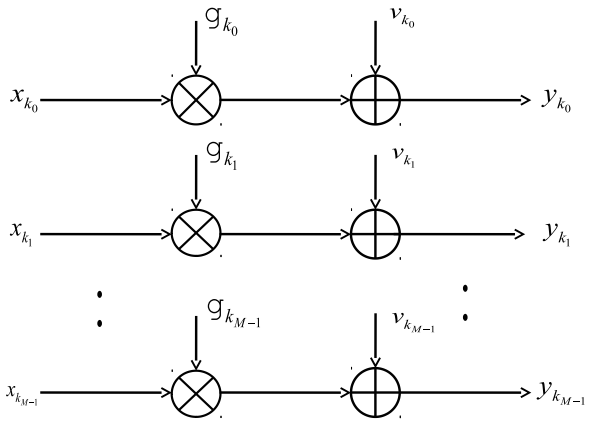

Figure 2.3.1: Pilots passing through a parallel of complex Gaussian channels

2.3.1 Frequency Domain Least Squares Estimation (LS)

Principle of Frequency Domain LS Estimator

For frequency domain LS estimator, the main goal is to realize the minimization of $\|\boldsymbol{y}_p - \boldsymbol{X}_p\mathbf{g}_p\|_2^2$, which can be expressed by:

$$\hat{\mathbf{g}}_{p,LS} = argmin\|\boldsymbol{y}_p - \boldsymbol{X}_p\mathbf{g}_p\|_2^2 \tag{2.3.2}$$

The solution is classical and given by [67, 90]:

$$\hat{\mathbf{g}}_{p,LS} = \boldsymbol{X}_p^{-1}\boldsymbol{y}_p = \mathbf{g}_p + \boldsymbol{X}_p^{-1}\boldsymbol{v}_p \tag{2.3.3}$$

From (2.3.3), the frequency domain LS estimator can be illustrated by Fig 2.3.2:

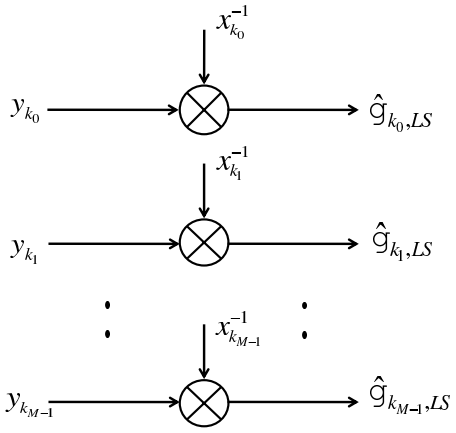

Figure 2.3.2: Frequency domain LS estimation model

Application of Frequency Domain LS Estimator in OFDM System

From (2.3.3) and Fig 2.3.2, frequency domain LS estimation has a simple structure on both mathematics and physical realization, however, its performances can be highly affected by noise. For frequency domain LS estimator, its mean square error (MSE) is give by [19]:

$$MSE_{\boldsymbol{g}_{p,LS}} = E[(\hat{\mathbf{g}}_{p,LS} - \mathbf{g})^H(\hat{\mathbf{g}}_{p,LS} - \mathbf{g})] = Trace(\sigma^2(\boldsymbol{X}_p^H \boldsymbol{X}_p)^{-1}) \tag{2.3.4}$$

where Trace(.) is the trace of a matrix. Here, only M estimated CFR samples are obtained, and the rest $N - M$ samples can generally be got by two types of methods. The first category of methods is to employ various kinds of interpolation methods (linear interpolation method is considered in chapter 4). The other category of methods is to use time domain LS based methods, which convert the $\hat{\boldsymbol{g}}_{p,LS}$ into time domain CIR. Thus, different time domain methods can be employed to improve the estimation performance. Additionally, if the channel is the non-sparse channel (rich multipath channel in chapter 3, the MSE can hardly be reduced furthermore. However, if the channel is sparse, this MSE can be significantly reduced. The details will be discussed in chapter 3.

2.3.2 Minimum Mean Square Error (MMSE) Estimation

Principle of Frequency Domain MMSE Estimator

The main goal of frequency domain MMSE estimator is to achieve the minimum mean square error of the estimated parameters, which is given by:

$$minE[(\hat{\mathbf{g}}_p - \mathbf{g})^H(\hat{\mathbf{g}}_p - \mathbf{g})] \tag{2.3.5}$$

By employing the orthogonal principle expressed by [91]:

$$E[(\mathbf{g}_p - \hat{\mathbf{g}}_p)^H \boldsymbol{y}_p] = 0 \tag{2.3.6}$$

The frequency domain MMSE is given by [91, 93]:

$$\hat{\mathbf{g}}_{p,MMSE} = \boldsymbol{R}_{\mathbf{g}_p \boldsymbol{y}_p} \boldsymbol{R}^{-1}_{\boldsymbol{y}_p \boldsymbol{y}_p} \boldsymbol{y}_p \tag{2.3.7}$$

where $\boldsymbol{R}_{\mathbf{g}_p \boldsymbol{y}_p} = \boldsymbol{R}_{\mathbf{g}_p \mathbf{g}_p} \boldsymbol{X}_p^H$ is the covariance matrix between $\boldsymbol{g}_p$ and $\boldsymbol{y}_p$; $\boldsymbol{R}_{\boldsymbol{y}_p \boldsymbol{y}_p} = \boldsymbol{X}_p \boldsymbol{R}_{\mathbf{g}_p \mathbf{g}_p} \boldsymbol{X}_p^H + \sigma^2 \boldsymbol{I}_M$ is the auto-covariance matrix of vector $\boldsymbol{y}_p$. More commonly, the frequency domain MMSE is expressed by [4, 67, 91]:

$$\hat{\mathbf{g}}_{p,MMSE} = \boldsymbol{R}_{\mathbf{g}_p \mathbf{g}_p} (\boldsymbol{R}_{\mathbf{g}_p \mathbf{g}_p} + \sigma^2 (\boldsymbol{X}_p \boldsymbol{X}_p^H)^{-1})^{-1} \hat{\mathbf{g}}_{p,LS} \tag{2.3.8}$$

From (2.3.8), we see that frequency domain MMSE algorithm is actually the further extension of frequency domain LS algorithm, which fully makes use of the prior knowledge of both the noise and channel statistics.

Application of Frequency Domain MMSE Estimator in OFDM System

Due to the presence of $(\boldsymbol{X}_p\boldsymbol{X}_p^H)^{-1}$, which is not convenient for calculations, $E[(\boldsymbol{X}_p\boldsymbol{X}_p^H)^{-1}]$ is introduced. Additionally, average power of signal to noise ratio (SNR) (Here signal is actually the pilot signal) defined as $\frac{E[|x_k|^2]}{\sigma^2}$ can be involved to make further simplification, thus, (2.3.8) can be rewritten as:

$$\hat{\mathbf{g}}_{p,Av,MMSE} = \boldsymbol{R}_{\mathbf{g}_p\mathbf{g}_p}(\boldsymbol{R}_{\mathbf{g_p}\mathbf{g_p}} + \frac{\beta}{SNR}\boldsymbol{I})^{-1}\hat{\mathbf{g}}_{p,LS} \tag{2.3.9}$$

where β is a constant related to the used constellation modulation. For example, in the case of QPSK, $\beta = 1$, while in the case of 16QAM, $\beta = 17/9$ [69]. The performance of the modified MMSE relies on both the statistics of the channel and noise.

2.3.3 Time Domain LS Estimation

Unlike the previous two estimators LS and MMSE in frequency domain, the time domain LS estimator mainly focuses on the channel estimation in time domain.

Principle of Time Domain LS Estimator

In time domain, (2.3.1) can be rewritten as [19]:

$$\boldsymbol{y}_p = \boldsymbol{X}_p\boldsymbol{F}_{M\times L_{cp}}\boldsymbol{h} + \boldsymbol{v}_p \tag{2.3.10}$$

where $\boldsymbol{F}_{M\times L_{cp}}$ (L_{cp} is the length of cyclic prefix) is the partial Fourier matrix, which can be obtained by selecting the rows of Fourier matrix with index $k_0, k_1, \ldots, k_{M-1}$ and the first L_{cp}

columns as follows:

$$\boldsymbol{F}_{M\times L_{cp}}=\begin{bmatrix} W_N^{k_0 0} & W_N^{k_0 1} & \dots & W_N^{k_0(L_{cp}-1)} \\ W_N^{k_1 0} & W_N^{k_1 1} & \dots & W_N^{k_1(L_{cp}-1)} \\ \vdots & \vdots & \vdots & \vdots \\ W_N^{k_{M-1}0} & W_N^{k_{M-1}1} & \dots & W_N^{k_{M-1}(L_{cp}-1)} \end{bmatrix} \tag{2.3.11}$$

where $W_N^{k_p l} = e^{\frac{-j2\pi k_p l}{N}}$, $0 \le p \le M-1$, $0 \le l \le L_{cp}-1$. Let $\boldsymbol{X}_p\boldsymbol{F}_{M\times L_{cp}} = \boldsymbol{A}$, which is generally known as the measurement matrix, (2.3.10) can be rewritten as:

$$\boldsymbol{y}_p = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_p \tag{2.3.12}$$

Time domain LS criterion solves the following minimization problem [19, 91]:

$$\hat{\boldsymbol{h}}_{p,LS} = \arg\min_{\boldsymbol{h}} \|\boldsymbol{y}_p - \boldsymbol{A}\boldsymbol{h}\|_2^2 \tag{2.3.13}$$

This is a classical optimization problem, which has the following closed solution [4,36,67,91]:

$$\hat{\boldsymbol{h}}_{p,LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y}_p \tag{2.3.14}$$

Application of Time Domain LS Estimator in OFDM System

For time domain LS estimator, its MSE is what we are actually interested in practical channel estimation. From (2.3.13), we know the MSE for time domain LS estimator is given by:

$$MSE_{\boldsymbol{h}_{p,LS}} = E[(\hat{\boldsymbol{h}}_{p,LS} - \boldsymbol{h})^H(\hat{\boldsymbol{h}}_{p,LS} - \boldsymbol{h})] = Trace(\sigma^2(\boldsymbol{A}^H\boldsymbol{A})^{-1}) \tag{2.3.15}$$

The question is whether $MSE_{\boldsymbol{h}_{p,LS}}$ achieves the Cramer-Rao Bound (CRB) [94]?

Let's simplify (2.3.15) by the following formula:

$$MSE_{\boldsymbol{h}_{p,LS}} = Trace(\sigma^2(\boldsymbol{A}^H\boldsymbol{A})^{-1}) = (L_{cp}/M)\sigma^2 \tag{2.3.16}$$

From the analysis and proof above, $MSE_{\boldsymbol{h}_{p,LS}}$ reaches CRB, when the considered channel is the rich multipath channel [19]. However, in the case of sparse channel, does the $MSE_{\boldsymbol{h}_{p,LS}}$ can still achieve CRB? From (2.3.14), an intuitive answer is no, since a sparse CIR can hardly have L_{cp} non-zero channel taps. For mathematical details, we will discuss it in chapter 4.

2.3.4 Time Domain Maximum Likelihood Estimation

Principle of Time Domain ML Estimator

In the case of $\boldsymbol{v}_p \sim CN(\boldsymbol{0}_M, \sigma^2\boldsymbol{I}_M)$, time domain ML estimator has the same expression with the time domain LS estimator, which is given by [4, 91, 95]:

$$\hat{\boldsymbol{h}}_{p,ML} \equiv \hat{\boldsymbol{h}}_{p,LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y}_p \tag{2.3.17}$$

Application of Time Domain ML Estimator in OFDM System

As one of the major channel estimation methods, time domain ML is popular in channel estimation in OFDM system [4, 74, 95]. [4] considers the time domain ML estimator and its derived forms in rich multipath channel environments, the estimation performance is generally effective. Comparatively, the performance of time domain ML estimation is not effective when estimating the sparse channel. In order to promote the estimation performance, a low complexity for tap selective ML estimator is proposed, the estimation results are convincing [74].

2.3.5 Simulations

We evaluate the performance of the frequency domain LS, MMSE estimators and time domain LS estimator (The time domain LS and ML estimators have the same expression, therefore, time domain LS estimator is only considered here.) in both rich multipath and sparse multipath channel environments for two different OFDM systems, the parameters of which are given as follows:

1) The first OFDM system modulated by QPSK has 1024 subcarriers and cyclic prefix of 16.

2) The second OFDM system modulated by QPSK has 1024 subcarriers and cyclic prefix of 64.

The first OFDM system is used to evaluate the estimation performance of rich multipath channel, while the second OFDM system is employed to evaluate the estimation performances of sparse multipath channel. For the rich multipath channel, 16 taps Rayleigh channel are considered while the sparse multipath channel has 6 taps with their taps position randomly generated. The channel power delay profile has an exponential distribution described by $\phi(\tau') = e^{-\frac{\tau'}{\tau_{rms}}}$, with $\tau_{rms} = \frac{L_{cp}}{4}$ [67]. The unit delay is the OFDM sample period. In the simulations, most of the three considered channel estimation methods use OFDM symbol with full pilots to estimate the channel.

In the following, we consider the NMSE (Normalized Mean Square Error) of the channel impulse response defined by [36]:

$$NMSE = \frac{E[\left\|\mathbf{h} - \hat{\mathbf{h}}\right\|_2^2]}{E[\|\mathbf{h}\|_2^2]} \tag{2.3.18}$$

where $\hat{\mathbf{h}}$ is an estimate of $\mathbf{h}$.

Fig 2.3.3 compares the NMSE performance of frequency domain LS estimator, MMSE estimator and time domain LS estimator in the case of rich multipath channel environment. As

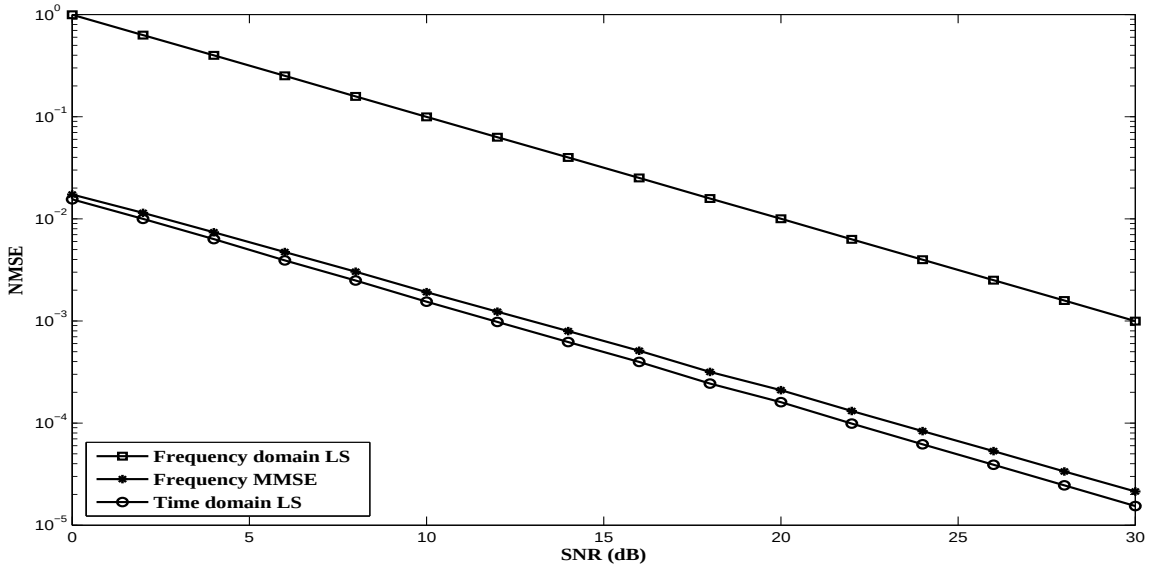

Figure 2.3.3: Performance of NMSE comparison for rich multipath channel estimation

can be seen from the figure, time domain LS estimator has the best and optimal channel estimation due to its measurement matrix with most of the considered L_{cp} non-zero channel taps and without any noise taps. Because of the utilization of full pilots and the prior knowledge of both channel statistics and noise variance, MMSE estimator achieves the good channel estimation performance throughout the whole considered SNR, however, its performance can hardly be optimal without considering the number of channel taps and tap positions. Frequency domain LS estimator has the poorest channel estimation.

The NMSE performance of the three estimators mentioned above in the sparse multipath channel environment is compared in Fig 2.3.4. By using the prior knowledge of the sparsity of the channel, time domain LS estimator with known sparsity (number of non-zero channel taps) has the best estimation performance. Frequency domain MMSE estimator achieves similar estimation performance as that of the time domain LS estimator, which actually considers the first L_{cp} taps, which includes noise taps. Frequency domain LS still has the poorest estimation performance.

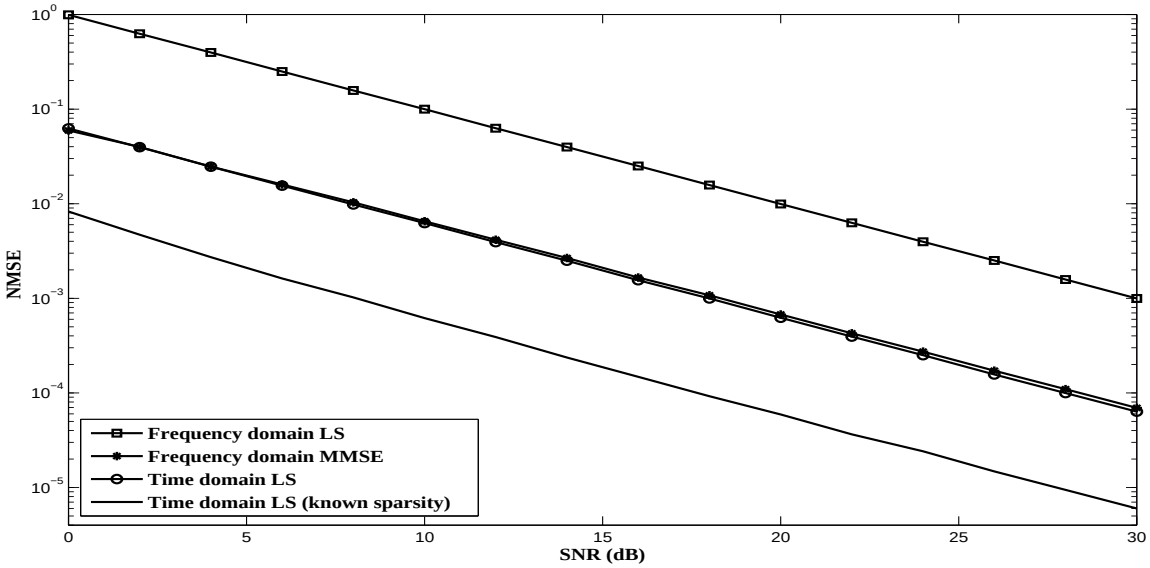

Figure 2.3.4: Performance of NMSE comparison for sparse multipath channel estimation

Table 2.1: Computational complexity comparison

Algorithm	Frequency domain LS	Time domain LS	MMSE
Complexity	$O(N)$	$O(ML_{cp})$	$O(N^2)$

Compared the three different channel estimation methods, frequency domain LS method has the lowest computational complexity shown in Table 2.1, however, its performance is the poorest in both the rich multipath and the sparse multipath channel. Frequency domain MMSE has comparatively good channel estimation performance in both rich multipath and sparse multipath channels, but it has the highest computational complexity and it requires both the channel and noise statistics, which severely burden its applications. Time domain LS estimator has optimal channel estimation in rich multipath channel, however, even its performance in sparse multipath channel is good, it is not so effective. Therefore, in chapters 4 and 5, effective channel estimation methods based on time domain LS estimator are developed.

2.4 Conclusion

In this chapter, we review the characteristics of both the wireless channel and OFDM system. Additionally, the mathematical models of classical channel estimation methods LS, MMSE and ML are presented and analyzed. At the end of this chapter, simulations are carried out to evaluate the computational complexity and estimation performance in the cases of both the rich multipath and sparse multipath channels.

Chapter 3

Compressed Sensing Theory and Models for Sparse Channel Estimation in OFDM System

The compressed sensing theory provides us a new vision and mind on how to acquire and process signal, as long as the signal is sparse in a certain transform domain. To compressed sensing based sparse channel estimation, how to design effective measurement matrix according to the specific sparse channel model is the essential pre-condition for effective sparse channel estimation in OFDM system.

3.1 Compressed Sensing Theory

3.1.1 Basic Principles of Compressed Sensing

Traditional signal processing methods are generally based on Shannon theorem, which states that to reconstruct a signal without distortion, the sampling frequency should be at least two times the bandwidth of the signal. It is a sufficient guarantee for any bandlimited signals. However, in both mathematical or signal processing fields, there exists a huge amount of signals, which are sparse or approximately sparse in some specific transform domains. In

this case, is the two times of the bandwidth for the sampling frequency still essential to reconstruct the signal? Actually, in this situation, if the high dimensional sparse signals are projected onto low dimensional spaces, it is possible to reconstruct these sparse signals by low dimensional observation vectors with almost no perceptual loss [20]. This is actually what compressed sensing (CS) is concerned about.

Basic Mathematical Model of Compressed Sensing

In classical signal processing field, a signal $\boldsymbol{z} \in C^N$ can generally be expressed by linear combination of bases $\{\boldsymbol{\psi}_k = 0, \ldots, N-1\}$ in a full rank $N \times N$ transform matrix $\boldsymbol{\Psi} = [\boldsymbol{\psi_0}, \boldsymbol{\psi_1}, \ldots, \boldsymbol{\psi_{N-1}}]$:

$$\boldsymbol{z} = \boldsymbol{\Psi x} \tag{3.1.1}$$

where $\boldsymbol{x}$ is the weighting vector with each coefficient weighting its corresponding basis.

Traditionally, as long as we get signal $\boldsymbol{z}$, it is not difficult to reconstruct the weighting vector $\boldsymbol{x}$ without regarding whether $\boldsymbol{x}$ is a sparse signal or not. Therefore, the dimension of $\boldsymbol{z}$ named observation vector is always N, which may cause huge loss in efficiency or greatly increase complexity, especially when $\boldsymbol{x}$ is a sparse signal.

Assume $\boldsymbol{x}$ is a K sparse vector with $\|\boldsymbol{x}\|_0 = |supp(\boldsymbol{x})| = |\Gamma| = K \leq K_{max} \ll N$; $\Gamma = \{i : x_i \neq 0\}$ is the support of the vector $\boldsymbol{x}$; K is also known as the value of sparsity; K_{max} is the maximum sparsity (maximum value of K). The main focus is on using just a $M (M \ll N)$ dimension observation vector $\boldsymbol{y} \in C^M$ to reconstruct vector $\boldsymbol{x}$ with almost no perceptual loss or small loss on precision.

Note: There are two different definitions of K sparse vector. The first one assumes that the sparse vectors contain exactly K non-zero coefficients if the sparse vector is K sparse. While the second one assumes that the sparse vectors

have at most K non-zero coefficients. Actually, both $\|x\|_0$ and its maximum value are important. Therefore, $\|x\|_0 = K$ and K_{max} are used.

In order to get a lower dimension observation vector $\boldsymbol{y}$, an observation matrix $\boldsymbol{\Phi} \in C^{M\times N}$ is required, therefore, we have the following formula:

$$\boldsymbol{y} = \boldsymbol{\Phi z} + \boldsymbol{w} = \boldsymbol{\Phi\Psi x} + \boldsymbol{w} \tag{3.1.2}$$

where $\boldsymbol{w} \in C^M$ is the CAWGN vector distributed as $CN(\boldsymbol{0}_M, \sigma^2\boldsymbol{I}_M)$. $\boldsymbol{A} = \boldsymbol{\Phi\Psi}$ is the measurement matrix. Then, a classical mathematical model in CS is expressed as follows:

$$\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{w} \tag{3.1.3}$$

In CS, the main task is to reconstruct the K sparse vector $\boldsymbol{x}$ with almost no perceptual loss. A key property for $\boldsymbol{A}$ in reconstructing the sparse signal named restricted isometry property (RIP) can be expressed as [96–99]:

$$(1-\delta_{K_{max}})\|\boldsymbol{x}\|_2^2 \le \|\boldsymbol{Ax}\|_2^2 \le (1+\delta_{K_{max}})\|\boldsymbol{x}\|_2^2 \tag{3.1.4}$$

where $\delta_{K_{max}} \in (0,1)$ is the constrained parameter, which guarantees the approximation performance.

Beyond the RIP property, which can guarantee the optimal CS reconstruction, mutual coherence of $\boldsymbol{A}$ is also an important parameter on effective reconstruction, which can be expressed by [56, 97, 100]:

$$\mu(\boldsymbol{A}) = \max_{0\le j,k\le N-1, j\ne k} \frac{|\langle \boldsymbol{a}_j, \boldsymbol{a}_k\rangle|}{\|\boldsymbol{a}_j\|_2\|\boldsymbol{a}_k\|_2} \tag{3.1.5}$$

where $\boldsymbol{A} = [\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{N-1}]$.

The effective reconstruction of sparse vector $\boldsymbol{x}$ from $\boldsymbol{A}$, which is composed of the bases

(atoms) can only be guaranteed by a small value of $\mu(\boldsymbol{A})$.

Sparse Solution to Underdetermined Linear Equations by l_p Norm

The sparse solutions to underdetermined linear equations can usually be realized by minimizing l_p (p=0, 1, 2) norm minimization. For sparse vector $\boldsymbol{x}$, its l_0 norm $\|\boldsymbol{x}\|_0$ is defined previously in this chapter, its l_1 norm and l_2 norm are defined by $\|\boldsymbol{x}\|_1 = \sum_{i=0}^{N-1}|x_i|$ and $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=0}^{N-1}|x_i|^2}$ respectively.

l_0 Norm Minimization

To reconstruct the sparse vector $\boldsymbol{x}$, l_0 norm minimization is the most direct solution, which can be expressed by the following formula:

$$min\|\boldsymbol{x}\|_0 \; subject \; to \; \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon \tag{3.1.6}$$

In the above formula, l_0 norm minimization finds the sparsest coefficients vector by counting the smallest number of non-zero components under the condition that the error is within the error tolerance bound ε. Generally speaking, it is a NP (non-deterministic polynomial-time) problem [43].

l_1 Norm Minimization

Since l_0 norm minimization is not computational efficient, therefore, the research focus is converted to l_1 norm minimization, which can replace l_0 norm minimization to effectively reconstruct a sparse signal. The l_1 norm minimization can be expressed in the following formula:

$$min\|\boldsymbol{x}\|_1 \; subject \; to \; \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon \tag{3.1.7}$$

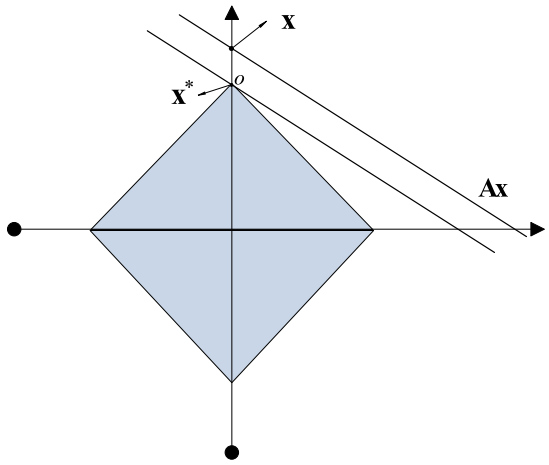

Figure 3.1.1: Optimal solution to l_1 norm

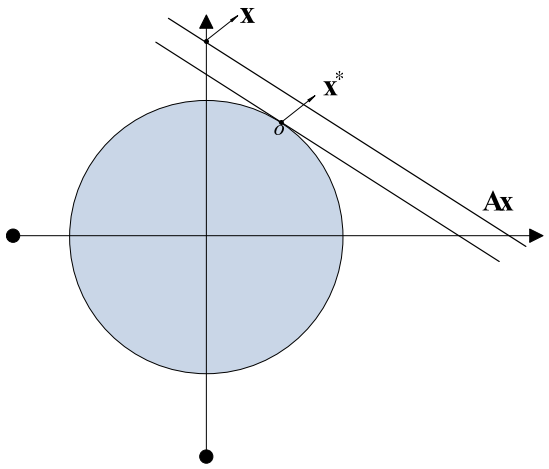

Figure 3.1.2: Optimal solution to l_2 norm

l_1 norm minimization pursues to find the sparsest solution to minimize $\|\boldsymbol{x}\|_1$, meanwhile the reconstruction error is bounded by $(0, \varepsilon]$.

Candes and Tao further limits the boundary of the reconstructed signal $\hat{\boldsymbol{x}}$ by [96, 97]:

$$\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2 \leq C_{K_{max}} \varepsilon \tag{3.1.8}$$

where $C_{K_{max}}$ is a constant value related to the maximum sparsity K_{max}.

Fig 3.1.1 illustrates the l_1 norm minimization from graphical angle. As can be seen from the figure, the goal of l_1 norm minimization method is to search the sparse vector $\boldsymbol{x}$ in terms of the minimum of $\|\boldsymbol{x}\|_1$. However, due to the existence of noise, we can only get the noisy observation vector $\boldsymbol{y}$, therefore, $\boldsymbol{x}^*$ located at the point O, which is also the top point of diamond. There exists a distance error between vector $\boldsymbol{x}$ and $\boldsymbol{x}^*$, which can hardly be reduced due to the random characteristics of noise.

l_2 Norm Minimization

Besides the l_1 norm minimization, l_2 norm minimization is also considered in sparse approximation, which has the following expression:

$$min\|\boldsymbol{x}\|_2 \; subject \; to \; \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{y}\|_2 \leq \varepsilon \tag{3.1.9}$$

The graphical model of l_2 norm minimization is shown in Fig 3.1.2. The sparsest solution is $\boldsymbol{x}^*$ the minimum of $\|\boldsymbol{x}\|_2$, which is generally realized by the least squares (LS) method. $\boldsymbol{x}^*$ is located at the point O, which is the tangent point for the tangent line $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}$ and the circle. Theoretically, l_2 norm minimization can hardly get the sparse vector $\boldsymbol{x}$ even without regarding noise vector $\boldsymbol{w}$, which is quite different from l_1 norm minimization.

Both l_1 norm minimization and l_2 norm minimization with their own characteristics are popular in CS reconstruction. For l_1 norm minimization, from Fig 3.1.1, it has the optimal solution to sparse signal reconstruction, however, its complexity reaches $O(M^2N^{\frac{3}{2}})$ [48, 101] which is comparatively high. For l_2 norm minimization, from Fig 3.1.2, its solution to sparse signal recovery can hardly be optimal, however, its complexity is comparatively low; the greedy algorithms, which are the most common l_2 norm solution, have the complexity around only $O(KMN)$ [48]. Statisticians and mathematicians prefer l_1 norm minimization algorithms due to its perfections in mathematics, while l_2 norm is more popular with engineering scientists due to its effective reconstruction performance, easy realization and low complexity.

3.1.2 Sparse Signal Reconstruction Algorithms in Compressed Sensing

Before investigating the sparse reconstruction algorithms, it is necessary to take eigenanalysis on some of the important correlation matrices. Consider the normalized measurement matrix defined by $\bar{\boldsymbol{A}} = [\frac{\boldsymbol{a}_0}{\|\boldsymbol{a}_0\|_2}, \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|_2}, \ldots, \frac{\boldsymbol{a}_{N-1}}{\|\boldsymbol{a}_{N-1}\|_2}]$ and its support Γ, the new measurement matrix

can be written by $\bar{\boldsymbol{A}}_\Gamma$. Then, the new CS model regarding to (3.1.3) can be expressed by:

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{A}}_\Gamma \boldsymbol{x}_\Gamma + \bar{\boldsymbol{w}} \tag{3.1.10}$$

where $\bar{\boldsymbol{w}}$ is the normalized AWGN with zero mean and noise variance $\sigma_w^2 = \frac{\sigma^2}{M}$. Let $\boldsymbol{c} = \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{y}} = \bar{\boldsymbol{A}}_\Gamma^H (\bar{\boldsymbol{A}}_\Gamma \boldsymbol{x}_\Gamma + \bar{\boldsymbol{w}}) = \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \boldsymbol{x}_\Gamma + \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{w}}$ whose autocorrelation is given by:

$$\boldsymbol{R}_{cc} = E[\boldsymbol{c}\boldsymbol{c}^H] = E[(\bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \boldsymbol{x}_\Gamma + \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{w}})(\boldsymbol{x}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma + \bar{\boldsymbol{w}}^H \bar{\boldsymbol{A}}_\Gamma)] = \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \Lambda \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma + \sigma_w^2 \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \tag{3.1.11}$$

where $\Lambda = E(diag[\boldsymbol{x}_\Gamma])^2$. Let $\boldsymbol{D}$ be the matrix composed of the eigenvectors of $\boldsymbol{R}_{cc}$ corresponding to the significant eigenvalues. We have:

$$\begin{aligned} \boldsymbol{R}_{cc}\boldsymbol{D} &= \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \Lambda \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \boldsymbol{D} + \sigma_w^2 \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma \boldsymbol{D} \\ &= \boldsymbol{C}\Lambda \boldsymbol{C}^H \boldsymbol{D} + \sigma_w^2 \boldsymbol{C}\boldsymbol{D} \end{aligned} \tag{3.1.12}$$

In (3.1.12), $\boldsymbol{C} = \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma$ is typically the coherence matrix in CS, which fully reflects the coherences between concerned bases. Obviously, the eigenvalues of $\boldsymbol{R}_{cc}$ are not the coefficients of Λ, mainly because $\boldsymbol{C}$ is not a normalized orthogonal matrix and the noise part $\sigma_w^2 \bar{\boldsymbol{A}}_\Gamma^H \bar{\boldsymbol{A}}_\Gamma$ exists. What we are really concerned about is the square of the l_2 norm of the error vector $\|\boldsymbol{e}\|_2^2 = \|\boldsymbol{\alpha} - \boldsymbol{\beta}\|_2^2$, where $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_{K-1}]^T$ is the vector composed of the eigenvalues of $\boldsymbol{R}_{cc}$ and $\boldsymbol{\beta} = [x_0^2, x_1^2, \ldots, x_{N-1}^2]_\Gamma^T$.

The signal reconstruction algorithms in CS are developed mainly based on l_1 and l_2 norms minimization. Of course, there still remains other reconstruction methods like iterative hard threshold (IHT) [57, 58] etc. The rest of this section will discuss about these major signal reconstruction algorithms in CS.

Greedy Pursuit Algorithm

The greedy pursuit algorithm is one of the major sparse signal approximation methods which is based on the solution on classical l_2 norm minimization. Traditionally, we minimize the l_2 norm problem (3.1.9) by least squares (LS) method expressed in the following formula:

$$\hat{\boldsymbol{x}}_{LS} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \tag{3.1.13}$$

where $\boldsymbol{A} \in M \times N$ is the measurement matrix. Actually, (3.1.13) is a classical well known minimization problem. If $M \geq N$ and $Rank(\boldsymbol{A}) = N$, $\hat{\boldsymbol{x}}_{LS}$ has the following solution:

$$\hat{\boldsymbol{x}}_{LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y} \tag{3.1.14}$$

As previously discussed in chapter 2, if $\boldsymbol{x}$ is a rich signal (most elements in $\boldsymbol{x}$ are non-zero), it is easy to argue the optimality of the LS estimator, which is equivalent to the ML estimator. However, in this work, sparse vector ($\|\boldsymbol{x}\|_0 = |supp(x)| \ll N$) is considered. In this case, LS can hardly be optimal. Indeed, in this case, an optimal or a suboptimal threshold is essential to get an optimal or suboptimal estimate of $\boldsymbol{x}$.

Although $M \geq N$ and $Rank(\boldsymbol{A}) = N$ can theoretically guarantee the reconstruction of vector $\boldsymbol{x}$ by LS, however, it can hardly guarantee the efficiency, which means that too many measurements than necessary are required especially when ($\|\boldsymbol{x}\|_0 = |supp(x)| \ll N$). Therefore, it is essential to consider the case $M < N$. In this case, classical LS method can hardly solve this problem. However, it doesn't mean that the problem in this case can't be solved. Actually, there are typical reconstruction methods of sparse signal from an overcomplete dictionary [100].

Greedy pursuit methods are effective to solve the sparse signal reconstruction from an overcomplete dictionary. Most greedy pursuit methods iteratively update the indexes of the

significant coefficients of the sparse vector $\boldsymbol{x}$ by the coherence calculation of the measurement matrix $\boldsymbol{A}$. The value of the significant coefficients in $\boldsymbol{x}$ is obtained by a constraint LS method to improve the sparse signal reconstruction precision. Finally, a suboptimal reconstruction of the sparse vector is obtained. In the following, we briefly discuss the Orthogonal Matching Pursuit (OMP) and some of its further extended and more sophisticated greedy pursuit methods.

Orthogonal Matching Pursuit (OMP) method is one of the most classical sparse signal reconstruction methods among greedy pursuit methods. Its specific realization of OMP method is expressed in [49, 102]:

Input: measurement matrix $\boldsymbol{A}$, observation vector $\boldsymbol{y}$ and stopping condition

Initialize residual vector $\boldsymbol{r}_0 = \boldsymbol{y}$, subset index for significant taps $\Gamma_0 = \emptyset$ and present iteration number $t = 0$.

While (Stopping criterion is not satisfied)

Project $\boldsymbol{r}_t$ onto the columns of $\boldsymbol{A}$: $\boldsymbol{c}_t = \boldsymbol{A}^H \boldsymbol{r}_t$;

Find the index of the highest coherence between the residual vector $\boldsymbol{r}_t$ and columns of $\boldsymbol{A}$ and update the subset index:

$\Gamma_{t+1} = \Gamma_t \bigcup \{\arg\max_i |\boldsymbol{c}_t(i)|\}$;

Update $\boldsymbol{x}_{t+1}$, $\boldsymbol{r}_{t+1}$ and t:

Obtaining suboptimal $\boldsymbol{x}_{t+1}$ by using the bases in $\boldsymbol{A}$ with the subset index Γ_{t+1}:

$\boldsymbol{x}_{t+1} = \arg\min_{\boldsymbol{u}} \|\boldsymbol{y} - \boldsymbol{A}_{\Gamma_{t+1}} \boldsymbol{u}\|_2$;

$\boldsymbol{r}_{t+1} = \boldsymbol{y} - \boldsymbol{A}_{\Gamma_{t+1}} \boldsymbol{x}_{t+1}$;

$t = t + 1$;

End

Output: estimated signal $\hat{\boldsymbol{x}}_{omp}$ by OMP.

Most of the OMP realization steps and parameters are fixed except the stopping condition, which is not determined. There are three possible stopping criteria proposed in [103]:

1) The loop of OMP algorithm stops after K_{max} iterations.

2) The residual has the constrained threshold: $\|\boldsymbol{r}_{t+1}\|_2 \leq \theta_1$.

3) when no column gives a significant amount of energy in the residual: $\|\boldsymbol{A}^H \boldsymbol{r}_{t+1}\|_\infty \leq \theta_2$.

The first stopping criteria employs K_{max}, which is the maximum possible number of non-zero coefficients in the sparse vector. However, it can hardly be optimal when $\|\boldsymbol{x}\|_0 < K_{max}$ and the signal to noise ratio (SNR) is low.

The stopping criteria 2) depends on θ_1. It's interesting to have discussions about the value of θ_1. Assume that the support Γ of the vector $\boldsymbol{x}$ is known, then, we can get its Oracle estimate ($\hat{\boldsymbol{x}}_{Oracle}$) [104] by LS:

$$\hat{\boldsymbol{x}}_{oracle} = \begin{cases} (\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)^{-1} \boldsymbol{A}_\Gamma^H \boldsymbol{y}, & on\ \Gamma \\ \boldsymbol{0}, & else \end{cases} \tag{3.1.15}$$

Note: Oracle estimator is LS method with ideal knowledge of the sparse signal on both the number of non-zero elements and their positions

From (3.1.15) we know that Oracle estimate is originated from LS method, which requires the knowledge of the support of the sparse vector $\boldsymbol{x}$. In this case, the mean square error (MSE) of the Oracle estimate can be obtained by the following formula [19, 96]:

$$E\{\|\hat{\boldsymbol{x}}_{oracle} - \boldsymbol{x}\|_2^2\} = \sigma^2 Tr((\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)^{-1}) \geq \frac{K^2\sigma^2}{Tr(\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)} = \frac{K^2\sigma^2}{\sum_{m=r_0}^{r_{K-1}} \|\boldsymbol{a}_m\|_2^2} \tag{3.1.16}$$

where $\boldsymbol{a}_m$ is the column in $\boldsymbol{A}$ with index m and $m \in \Gamma$. The equality holds, if and only if $\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma = diag[\|\boldsymbol{a}_{r_0}\|_2^2, \|\boldsymbol{a}_{r_1}\|_2^2, \ldots, \|\boldsymbol{a}_{r_{K-1}}\|_2^2]\boldsymbol{I}_K$ and $\|\boldsymbol{a}_{r_0}\|_2^2 = \|\boldsymbol{a}_{r_1}\|_2^2 = \cdots = \|\boldsymbol{a}_{r_{K-1}}\|_2^2$.

What's the relationship among the MSE or root mean square error (RMSE) of other estimators, the RMSE of Oracle estimate and CRB? Indeed, any unbiased estimator ($\hat{\boldsymbol{x}}$) of $\boldsymbol{x}$ satisfies the following relationship with CRB [105]:

$$RMSE(\hat{\boldsymbol{x}}) = \sqrt{E\{\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_2^2\}} \geq \sigma\sqrt{Tr((\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)^{-1})} = RMSE_{oracle}(\hat{\boldsymbol{x}}) = CRB \quad (3.1.17)$$

Therefore, for the $\|\boldsymbol{r}_{t+1}\|_2$ bounded noise, its error tolerance bound θ_1 should follow the following relations:

$$\theta_1 \geq \sigma\sqrt{Tr((\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)^{-1})} \quad (3.1.18)$$

For θ_2, the universal threshold originated from [106], [107], is widely accepted for the error tolerance bound of the l_∞ bounded noise [54]:

$$\theta_2 = \sqrt{2(a+1)lnN}\sigma \quad (3.1.19)$$

where $a \geq 0$ is a constant value.

OMP algorithm is one of the most original and direct greedy pursuit algorithms. It is simple but effective in many cases. However, there are still some cases, in which OMP algorithm can't work effectively or efficiently. There are two major characteristics for OMP algorithm. The first major characteristic of OMP algorithm is that once a basis in the measurement matrix is selected according to the largest coherence between the residual vector and the bases at each iteration, it will never be kicked out. The second one is that any bases can never be chosen twice. If there is one mischosen basis, it may cause considerable error in reconstruction. Additionally, choosing only one basis at one iteration results in comparatively high computational complexity for large scale problems [108]. Therefore, many extended algorithms of greedy pursuit have been developed, they mainly focuse on making changes on these two major characteristics to reduce the computational complexity, enhance the stability

of recovery or promote the recovery precision of sparse signals. There are algorithms such as: stagewise OMP (StOMP) [108], regularized OMP (ROMP) [109], compressive sampling matching pursuit (CoSaMP) [110], subspace pursuit (SP) [111] and back-tracking based adaptive orthogonal matching pursuit (BAOMP) [48]. StOMP and ROMP realize effective sparse signal reconstruction by iteratively selecting the subspaces composed of the bases (not just one basis) having the highest coherence with the residual. Comparatively, CoSaMP and SP iteratively choose the subspaces of bases with a fixed dimension ($2K_{max}$ or K_{max}) of the support vectors. Unlike the previous four algorithms, BAOMP method incorporates a simple backtracking technique to detect the reliability of previous chosen bases and then deletes the unreliable bases at each iteration. In the following of this chapter, StOMP and CoSaMP are presented respectively.

Stagewise orthogonal matching pursuit (StOMP) is a greedy pursuit algorithm extended from OMP. Different from OMP algorithm, which iteratively selects the highest coherences between the bases and the residual vector, StOMP method iteratively chose the subspaces composed of the bases with the highest coherence between the remaining bases and the residual vector according to a threshold with respect to the residual vector. Specifically, it can be realized as follows [108]:

Input: measurement matrix $\boldsymbol{A}$, observation vector $\boldsymbol{y}$ and stopping condition

Initialize residual vector $\boldsymbol{r}_0 = \boldsymbol{y}$, subset index for significant taps $\Gamma_0 = \emptyset$ and present iteration number $t = 0$.

While (Stopping criterion is not satisfied)

Project $\boldsymbol{r}_t$ onto $\boldsymbol{A}$: $\boldsymbol{c}_t = \boldsymbol{A}^H \boldsymbol{r}_t$;

Find the indices of bases whose coherence with the residual vector is higher than a threshold:

Λ_t=$\{i : |c_t(i)| > \lambda_t \sigma_t\}$

Here $\sigma_t = \|\boldsymbol{r}_t\|_2/\sqrt{M}$ is the noise level and $\lambda_t \in [2, 3]$ is a parameter. Renew the subset

index at the iteration t:

$\Gamma_{t+1} = \Gamma_t \bigcup \Lambda_t$

Obtaining suboptimal $\boldsymbol{x}_{t+1}$ by using the bases in $\boldsymbol{A}$ with the subset index Γ_t:

$\boldsymbol{x}_{t+1} = \arg\min_{\boldsymbol{u}} \|\boldsymbol{y} - \boldsymbol{A}_{\Gamma_{t+1}}\boldsymbol{u}\|_2$;

$\boldsymbol{r}_{t+1} = \boldsymbol{y} - \boldsymbol{A}_{\Gamma_{t+1}}\boldsymbol{x}_{t+1}$;

t=t+1;

End

Output: estimated signal $\hat{\boldsymbol{x}}_{StOMP}$ by StOMP.

Similar with the StOMP method, Compressive Sampling Matching Pursuit (CoSaMP) iteratively chooses the $2K_{max}$ bases index in terms of the maximum coherence between the bases and the residual vector whose realization procedure is given by [110]:

Input measurement matrix $\boldsymbol{A}$, observation vector $\boldsymbol{y}$, maximum sparsity K_{max} and stopping condition

Initialize residual vector $\boldsymbol{r}_0 = \boldsymbol{y}$, approximated sparse vector $\boldsymbol{x}_t$ at iteration t ($\boldsymbol{x}_0 = \boldsymbol{0}$), and present iteration number $t = 0$.

While (Stopping criterion is not satisfied)

Project $\boldsymbol{r}_t$ onto $\boldsymbol{A}$: $\boldsymbol{c}_t = \boldsymbol{A}^H\boldsymbol{r}_t$;

Find the indices for the $2K_{max}$ highest coherence between the residual vector and bases and update the subset index:

Λ=supp($[\boldsymbol{c}_t]_{2K_{max}}$);

Γ_{t+1}=$\Lambda \bigcup$ supp($\boldsymbol{x}_t$);

$\boldsymbol{v}|_{\Gamma_{t+1}} = \arg\min_{\boldsymbol{u}} \|\boldsymbol{y} - \boldsymbol{A}_{\Gamma_{t+1}}\boldsymbol{u}\|_2$;

$\boldsymbol{v}|_{\Gamma_{t+1}{}^c} = \boldsymbol{0}$

Obtain the approximated sparse vector at the current iteration:

$\boldsymbol{x}_{t+1}$=$[\boldsymbol{v}]_{K_{max}}$

$\boldsymbol{r}_{t+1}$=$\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{t+1}$;

$t = t + 1$

End

Output: estimated signal $\hat{\boldsymbol{x}}_{CoSaMP}$ by CoSaMP.

Instead of choosing one basis at each iteration, both StOMP and CoSaMP algorithms realize the improved recovery performances by searching the subspace of bases at each iteration, which can reduce the computational complexity and provide more effective recovery performance in many cases [103, 108, 110]. Compared with the greedy pursuit algorithms, l_1 norm minimization based algorithms have the advantages in recovery performance [22, 103], although their computational complexities are much higher. In the following, several l_1 norm minimization based algorithms are presented.

l_1 Norm Minimization Based Algorithms

Most l_1 minimization norm based algorithms try to solve the l_1 minimization problem with the constrained conditions in (3.1.7). Dantzig selector (DS) is one of the most popular l_1 norm based algorithm, which tries to solve the following optimization problem [54]:

$$min\|\hat{\boldsymbol{x}}\|_1 \; subject \; to \; \|\boldsymbol{A}^H(\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{y})\|_\infty \leq \lambda \tag{3.1.20}$$

where $\lambda = \sqrt{(2(a+1)lnN)}\sigma$ is the recommended error tolerance bound for the l_∞ norm of the residual.

Additionally, l_1 norm minimization can also be converted into Linear programming (LP) problem [52] and realized by iterative approximation process. For simplicity, details will

not be discussed. Beyond DS method and LP method, Basis pursuit (BP) method, which was early studied in [53] is also popular in solving the l_1 minimization problem. BP can be used in the noiseless case. While in noisy cases, BP employs a primal-dual interior point method [53] to realize the denoising procedure.

Generally, the l_1 norm minimization based algorithms are converted into a convex optimization [112] and solved by interior-point methods, projected gradient methods or iterative thresholding methods [22].

Iterative Hard Thresholding

The concept of threshold is widely used in CS, not only in denoising, but also in the processes of pursuit algorithms. Besides the previous StOMP and BAOMP algorithms, threshold also involves in the realization of the iterative hard thresholding (IHT) [57, 58] algorithm. IHT approach is a simple but effective method to reconstruct the sparse signal, which can be realized by iterations ($\boldsymbol{x}_0 = \mathbf{0}$, $t = 0$):

$$\boldsymbol{x}_{t+1} = H_{K_{max}}(\boldsymbol{x}_t + \boldsymbol{A}^H(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_t)) \tag{3.1.21}$$

where $H_{K_{max}}(\cdot)$ is a non-linear operator that sets all elements to zero except the K_{max} elements having largest amplitudes.

3.1.3 Simulations

In the simulation part, we consider the linear reconstruction of a sparse vector from a few random observations by CS. The general CS model without noise is employed, which can be expressed by:

$$\boldsymbol{y}_{noiseless} = \boldsymbol{A}\boldsymbol{x} \tag{3.1.22}$$

where $\boldsymbol{A} = \boldsymbol{\Phi}\boldsymbol{\Psi} \in C^{M\times N}$ is the measurement matrix; $\boldsymbol{\Phi} \in C^{M\times N}$ is a random matrix with each element following independent and identical complex Gaussian distribution $CN(0,1)$; $\boldsymbol{\Psi}$ is the Fourier matrix; $\boldsymbol{y}_{noiseless} \in C^M$ is the noiseless observation vector; $\boldsymbol{x} \in C^N$ is the sparse signal vector with values distributed as complex Gaussian $CN(0,1)$.

Simulation focuses on realizing effective reconstruction of sparse vector $\boldsymbol{x}$ with low dimension from the observation vector $\boldsymbol{y}_{noiseless}$ (low value of M). The dimension of sparse vector $\boldsymbol{x}$ is 1024 and 20 non-zero elements are randomly located. The purpose of this simulation is to show how wasteful it is to take so many measurements to reconstruct a sparse signal, therefore, only OMP algorithm is adopted for simplicity.

To evaluate the performance of sparse vector $\boldsymbol{x}$ reconstruction, mean square error (MSE) is adopted, which is given by:

$$MSE_{\boldsymbol{x}} = \frac{1}{N_{it}}(\sum_{t=0}^{N_{it}} \frac{1}{N} \sum_{i=0}^{N-1} |x_i^{(t)} - x_i|^2) \tag{3.1.23}$$

where N_{it} is the number of Monte Carlo simulations.

Figure 3.1.3 shows MSE performance of the sparse vector reconstruction by OMP algorithm in function of M (Dimension of the Observation Vector). Generally, the performance of MSE decreases slightly (MSE performance decreases less than 3.5×10^{-6}) with the significant decrease of M (From 100 to 30). More specifically, when M increases from 30 to 60 and from 60 to 100, MSE gains are less than 3×10^{-6} and about 5×10^{-7} respectively. This simulation fully shows that to recover a sparse vector, a small number of measurements is enough and there exists great redundancy in the measurements.

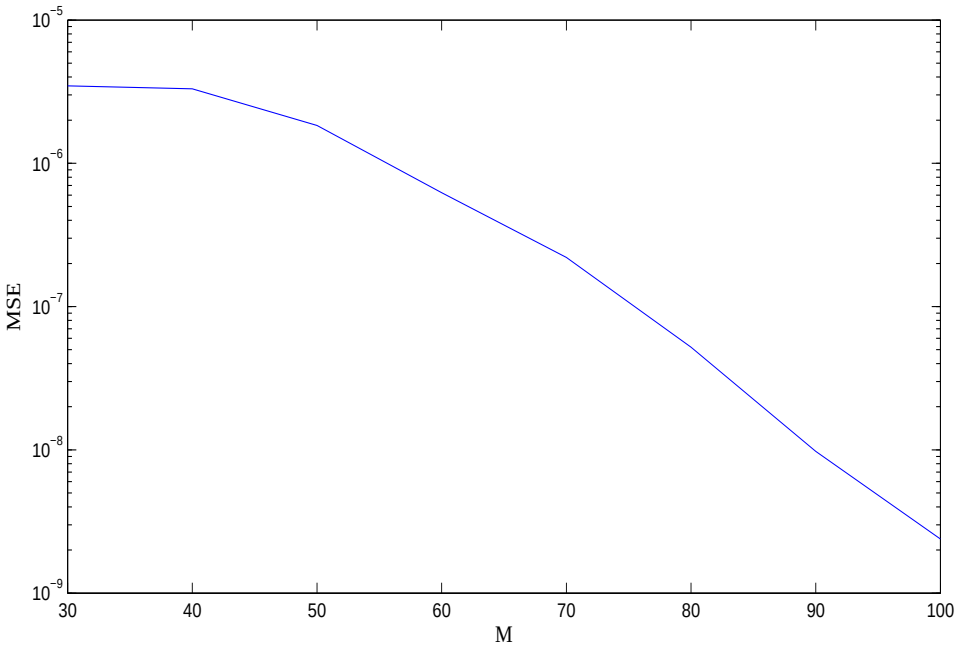

Figure 3.1.3: Performance of MSE for sparse vector recovery

3.2 Compressed Sensing based Sparse Channel Estimation in OFDM System

In the previous section, the compressed sensing theory, which provides the fundamental basis of the sparse channel estimation is discussed. In this section, we focus on constructing the mathematical model of CS based sparse channel estimation. In order to build the mathematical model of CS based sparse channel estimation we should firstly get to know the characteristics of sparse channel.

3.2.1 Characteristics of Sparse Channel

The degrees of freedom (DOF) concept is widely used in the field of statistics, physics and mechanics of structures etc. It can also be used in the wireless communication to describe

the wireless channel [19]. Roughly speaking, the number of degrees of freedom for wireless channel and system is the maximum number of all resolvable quantities, including delays, Doppler shifts, angles of arrival for the receiver and angles of departure for the transmitter etc. For simplicity, the DOF concerning the wireless channel could be considered with only delays [19]:

$$D = L_{cp} = \tau_{cp} W \tag{3.2.1}$$

where W is usable bandwidth of the transmitter signal over the channel, τ_{cp} is the duration of cyclic prefix and L_{cp} is the length of cyclic prefix, therefore, generally, we have $\tau_{cp} = L_{cp}\frac{1}{W} = L_{cp}T_s$ (In [19], $\frac{1}{W} = T_s$ is called the delay resolution bins (in Figure 3.2.1) with Nyquist sampling for delay). However, in many communication applications, the actual or

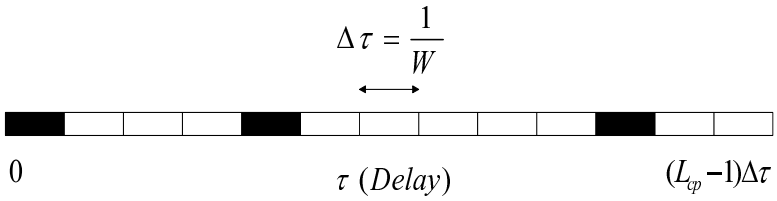

Figure 3.2.1: Sparse Channel with Delay Resolution Bins

effective number of DOF d satisfies $d \ll D$. Practically, d can be obtained by [19, 113]:

$$d = |[i] : |h[i]| > \theta| \tag{3.2.2}$$

θ is a properly chosen threshold. As can be seen from Figure 3.2.1, the number of black bins, which present the effective DOF is extremely small compared with the total number of bins. Therefore, it is a typical sparse channel.

The above analysis is mainly for sample spaced channels. In practical wireless communications, the values of delays of multipath channel taps may be non multiple of the sampling period, the channel can be a non-sample sparse spaced channel. In chapter 2, both sample spaced and non-sample spaced channels have been discussed. For the sample spaced sparse

channel, both pilot arrangement and threshold are important to effective channel estimation with high spectral efficiency. For the non-sample spaced sparse channel, due to the existence of leakage, beyond pilot arrangement and threshold, it is important to reduce the width of the resolution bins and increase the resolution. As previously stated in the first chapter, the sparse channel can be encountered in various different practical wireless communication environments, such as the underwater acoustic channels [22, 31], ITU-R vehicular channels (channel A and channel B) [33] etc. To estimate the sparse channel, its corresponding mathematical model is essential. In the following, the mathematical models of both sample spaced and non sample spaced sparse channels based on CS are presented concerning the OFDM system.

3.2.2 Mathematical Model for Compressed Sensing based Sparse Channel Estimation in OFDM System

Mathematical Model for CS Based Sample Spaced Sparse Channel Estimation

Consider a N subcarriers OFDM system with M pilots. Here we adopt the mathematical model given in (2.3.12):

$$\boldsymbol{y}_p = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_p \tag{3.2.3}$$

where $\boldsymbol{A} = \boldsymbol{X}_p \boldsymbol{F}_{M \times L_{cp}}$ is generally known as the measurement matrix.

The main task is to reconstruct the K (K non-zero channel taps, $K \ll L_{cp}$) sparse channel taps from a few number of pilots with almost no perceptual loss in performance. The above mathematical model can be used for both traditional LS ($M \geq L_{cp}$) and CS ($M < L_{cp}$) based sample spaced sparse channel estimation, which will be discussed in both chapter 4 and chapter 5.

Mathematical Model for CS based Non-sample Spaced Sparse Channel Estimation

In the section 2.1.3 and 3.2.1, the characteristics of non sample spaced sparse multipath channel are discussed. One of the main characteristics of the non-sample spaced channel is the leakage of non-zero channel taps at the receiver, which is caused by the limitations of the sampling speed of the devices in wireless communication (such as ADC). More specifically, sampling interval of communication systems may not be small enough, which can hardly guarantee most of the channel taps locating at the sampling points, just as Figure 3.2.1 shows, the channel taps may locate in the middle of the black bins (between two sampling points).

In order to reduce the leakage of the non-sample spaced sparse channel, a redundant Fourier matrix (overcomplete dictionary or redundant dictionary in compressed sensing), which has finer resolution $lT_s/R,\ l = 0, 1 \ldots (L_{cp}-1)R$ (R is the oversampling factor) [19] is introduced:

$$\boldsymbol{F}_{M\times(R(L_{cp}-1)+1)}=\begin{bmatrix} W_N^{k_0 0} & W_N^{k_0(1/R)} & \cdots & W_N^{k_0(L_{cp}-1)} \\ W_N^{k_1 0} & W_N^{k_1(1/R)} & \cdots & W_N^{k_1(L_{cp}-1)} \\ \vdots & \vdots & \vdots & \vdots \\ W_N^{k_{M-1} 0} & W_N^{k_{M-1}(1/R)} & \cdots & W_N^{k_{M-1}(L_{cp}-1)} \end{bmatrix} \tag{3.2.4}$$

With the redundant dictionary, the mathematical relationship between the transmitter and receiver for the non-sample spaced sparse channel can be similarly constructed as the sample spaced sparse channel:

$$\boldsymbol{y}_p = \boldsymbol{X}_p \boldsymbol{F}_{M\times(R(L_{cp}-1)+1)} \boldsymbol{h} + \boldsymbol{v} \tag{3.2.5}$$

Here $\boldsymbol{h}$ is channel impulse response vector with the dimension of $R(L_{cp}-1)+1$. The parameters for $\boldsymbol{y}$, $\boldsymbol{X}_p$ and $\boldsymbol{v}$ have the same definition as the sample spaced sparse channel.

By introducing the parameters $\boldsymbol{y}_p$, $\boldsymbol{A}_{fr}$ and $\boldsymbol{v}_p$, (3.2.5) can be rewritten as:

$$\boldsymbol{y}_p = \boldsymbol{A}_{fr}\boldsymbol{h} + \boldsymbol{v}_p \tag{3.2.6}$$

where $\boldsymbol{A}_{fr} = \boldsymbol{X}_p\boldsymbol{F}_{M\times(R(L_{cp}-1)+1)}$ is the measurement matrix with finer resolution, which can be rewritten as:

$$\boldsymbol{A}_{fr}= \begin{bmatrix} a_{k_0 0} & a_{k_0 1} & \dots & a_{k_0(L_{cp}-1)R} \\ a_{k_1 0} & a_{k_1 1} & \dots & a_{k_1(L_{cp}-1)R} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k_{M-1} 0} & a_{k_{M-1} 1} & \dots & a_{k_{M-1}(L_{cp}-1)R} \end{bmatrix} \tag{3.2.7}$$

where $a_{k_m n} = x_{k_m} e^{\frac{-j2\pi k_m n}{RN}}, \quad 0 \le m \le M-1,\, 0 \le n \le (L_{cp}-1)R.$

How to use the limited number of pilots M to reconstruct $\boldsymbol{h}$ with dimension of $R(L_{cp}-1)+1$ ($M \ll R(L_{cp}-1)+1$) is a typical CS reconstruction problem, which will be discussed in chapter 5.

3.3 Conclusion

In this chapter, the basic compressed sensing theory is presented, based on which, the mathematical models for CS based sample spaced and non-sample spaced sparse channels estimations are introduced. In chapters 4 and 5, the CS theory and the two mathematical models are employed to realize effective sparse channel estimation.

Chapter 4

Classical LS based Sparse Channel Estimation

Traditional frequency domain and time domain LS estimation methods are popular in channel estimation. Frequency domain LS estimator combined with different interpolation methods is the mainstream of the frequency domain channel estimation. Comparatively, the channel impulse response (CIR) of sparse channel estimated by time domain LS method in the case of $M \geq L_{cp}$ is vulnerable to noise, which can actually be resolved by effective threshold. To effectively reduce the impact of noise, different threshold estimation methods are proposed, however, these threshold estimation methods either rely on the prior knowledge of channel statistics and noise standard deviation or haven't got the desired estimation precision. This chapter fully extracts both the characteristics of noise and channel sparsity to realize effective sparse channel estimation without the prior knowledge of channel statistics and the noise standard deviation.

4.1 Frequency Interpolated based Channel Estimation

4.1.1 Linear Interpolation based Channel Estimation

Frequency domain interpolated channel estimation methods realize channel estimation in frequency domain combined with interpolation technique for OFDM system having comb type of pilots [15, 69], which are convenient for equalization and widely used in traditional channel estimation [15, 69, 72]. For comb type of pilots based channel estimation, only M pilot subcarriers are known to the receiver and can be used to estimate the CFR coefficients at the remaining $N-M$ data subcarriers. As previously mentioned, linear interpolation and second-order interpolation are all popular in estimating comb type pilot based channel in frequency domain. Among them, linear interpolation method is the most simple and classical one, which is given by [72]:

$$\hat{g}_k = (\frac{\hat{g}_{k_{p+1}} - \hat{g}_{k_p}}{k_{p+1} - k_p})(k - k_p) + \hat{g}_{k_p} \tag{4.1.1}$$

where $\hat{g}_{k_{p+1}}$ and $\hat{g}_{k_p}$ are the estimated CFR coefficients corresponding to the $p+1^{th}$ and p^{th} pilots respectively. For comb type pilot, the pilot interval PI $= k_{p+1} - k_p$ is a constant value. $\hat{g}_k$ is the estimated $k^{th}, 0 < k - k_p <$ PI coefficient of CFR.

4.1.2 Simulations and Analysis

The simulations evaluate the estimation performance of the classical frequency domain channel estimation method (frequency domain LS method combined with linear interpolation method) with different PI.

The parameters for the simulated QPSK modulated OFDM system are taken as: $N = 1024$, $L_{cp} = 64$. Sparse channels with 6 channel taps and randomly distributed tap positions

are considered. The channel power delay profile has an exponential distribution described by $\phi(\tau) = e^{-\frac{\tau}{\tau_{rms}}}$, with $\tau_{rms} = \frac{L_{cp}}{4}$. The unit delay is the OFDM sample period. In the simulations, the channel is assumed to be static during the duration of one OFDM symbol.

Figure 4.1.1 shows the NMSE (frequency domain) performance of classical frequency domain channel estimation method with different pilot intervals PI=16, 8, 4 and their corresponding pilots numbers are 64, 128 and 256 respectively. As can be seen from the figure, the NMSE performance of the classical frequency domain channel estimation method degrades with the increase of PI overall. Of course, some slight differences for classical frequency domain channel estimation method with (PI=4, 8) in the (0-8dB for Eb/N0) are observed beyond this overall trend. The primary reason is that the NMSE performance is measured not in terms of signal to noise ratio (SNR) but in terms of Eb/N0, which is more fair to measure the performances of different channel estimation methods or one channel estimation method with different parameters. In Appendix A, the relationship between SNR and Eb/N0 is presented in terms of the number of pilots and cyclic prefix. In terms of the parameters of N=1024, L_{cp}=64 and QPSK modulation considered in this simulation, SNR=Eb/N0+2.1dB in the case of PI=8, while in the case of PI=4, SNR=Eb/N0+1.4dB. Obviously, the power payload (the power for actual data transmission) of classical frequency domain channel estimation method with PI=4 is 0.7dB higher than that of PI=8. That can explain why the PI=8 outperforms PI=4 for Eb/N0 0-8dB.

Fig 4.1.2 shows that the variation of BER performance of classical frequency channel estimation method with different pilot intervals bears great similarities with that of the NMSE performance which means that if PI decreases, M will increase, the performance will be improved.

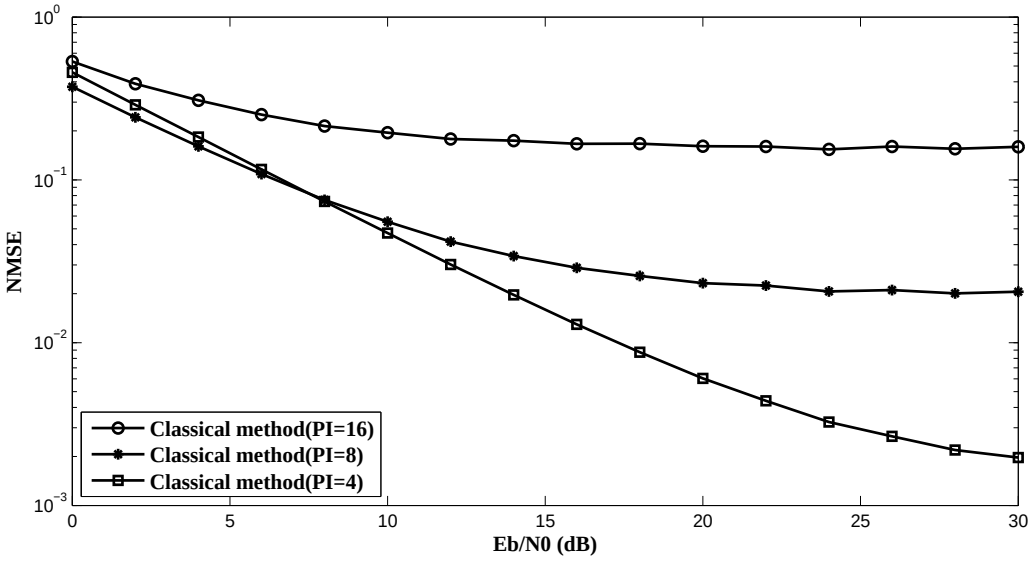

Figure 4.1.1: NMSE performance of classical frequency channel estimation method with different pilot intervals

4.2 LS based Time Domain Channel Estimation

From the above analysis, we observe that for interpolation methods, especially linear interpolation method, the estimation performance highly depends on the pilot interval (PI). Comparatively, time domain least squares (LS) estimator does not so closely relate to PI, but it depends on the maximum delay of the channel ($M \geq L_{cp}$). LS method is expressed by the following formula, initially discussed in subsections 2.3.3 and 3.1.2.

$$\hat{\boldsymbol{x}}_{LS} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{y} \tag{4.2.1}$$

For LS estimator, it requires that the measurement matrix $\boldsymbol{A}$ ($M \geq L_{cp}$) is of full column rank. The mean square error of LS estimator can be expressed by:

$$E\{\|\hat{\boldsymbol{x}}_{LS} - \boldsymbol{x}\|_2^2\} = \sigma^2 Tr((\boldsymbol{A}^H \boldsymbol{A})^{-1}) \tag{4.2.2}$$

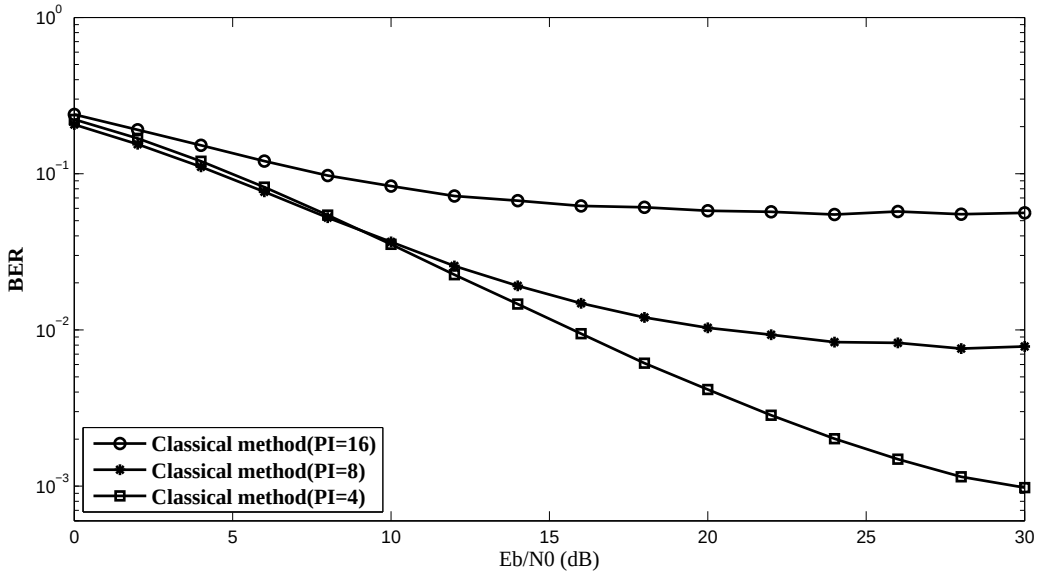

Figure 4.1.2: BER performance of classical frequency channel estimation method with different pilot intervals

According to the analysis in subsection 2.3.3, $E\{\|\hat{\boldsymbol{x}}_{LS} - \boldsymbol{x}\|_2^2\}$ equals to its Cramer Rao Bound (CRB), however, it is based on the fact that $\boldsymbol{x}$ is a rich signal vector. For sparse vector $\boldsymbol{x}$, does (4.2.2) still equal to CRB? Assume that the support of $\boldsymbol{A}$ is Γ, we may have the Oracle estimator. The relationship between Oracle estimator and CRB in the case of known support of sparse vector $\boldsymbol{x}$, which has already been presented in chapter 3, is recalled by:

$$E\{\|\hat{\boldsymbol{x}}_{oracle} - \boldsymbol{x}\|_2^2\} = \sigma^2 Tr((\boldsymbol{A}_\Gamma^H \boldsymbol{A}_\Gamma)^{-1}) = CRB \tag{4.2.3}$$

Obviously, for sparse signal vector $\boldsymbol{x}$, $E\{\|\hat{\boldsymbol{x}}_{oracle} - \boldsymbol{x}\|_2^2\}$ is the theoretical achievable smallest value of MSE instead of the MSE expressed by (4.2.2). Consider Γ^c (the complement of Γ), $\boldsymbol{x}_{\Gamma^c}$ is obviously the set of indices of noise coefficients, which have no useful information and should be eliminated. To effectively eliminate the noise coefficients, noise standard deviation is the basic and essential information. With the knowledge of noise standard deviation, threshold can be employed to effectively eliminate noise components. In the

following section, the characteristics of noise and its standard deviation estimation are firstly considered. Then, two different thresholds based on noise standard deviation are analyzed in terms of false alarm rate (FAR) and probability of detection (POD).

4.3 Noise Characteristics and Threshold Analysis

4.3.1 Noise Standard Deviation Estimation

A stationary noisy signal is usually described as a stationary stochastic process, which follows a certain type of distributions (Gaussian noise, Poisson noise and so on). Noise variance or noise standard deviation is one of the main characteristics to describe the random noise. Assume $Z(t)$ is a random noise and its samples are $z_i, 0 \leq i \leq N-1$. There are different categories of methods to estimate the standard deviation of $Z(t)$. There are methods, such as the empirical sample standard deviation, the Bayesian method [114] and the median absolute deviation (MAD) based method [106, 115]. We will discuss them in details.

Note: Median absolute deviation based method is popular in standard deviation estimation of noise in presence of sparse signal. [72, 106]

The sample standard deviation (biased estimator) of the noise random process $Z(t)$ is given by:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(z_i - \bar{z})^2} \tag{4.3.1}$$

where $\bar{z} = \frac{1}{N}\sum_{i=0}^{N-1} z_i$ is the sample mean value of noise samples $z_i, 0 \leq i \leq N-1$.

For simplicity the Bayesian estimation method is not discussed in this paper. [114] gives the details of the realization of the algorithm for Rayleigh distribution.

Besides the above two different standard deviation estimation methods, the median absolute deviation (MAD) based method is also widely used in the noise standard deviation estima-

tion, particularly a sparse signal is presented in the noise [106, 107, 115]. The fundamental idea behind the MAD based method is the close relationship between MAD of noise and the noise standard deviation, which can be expressed by:

$$\sigma = S \cdot MAD \tag{4.3.2}$$

where MAD is the median absolute deviation, which is given by:

$$MAD = \text{median}(|Z|) \approx \text{median}(|z_i, 0 \leq i \leq N-1|) \tag{4.3.3}$$

In (4.3.2), S is a constant factor, which depends on the specific distribution of random variables. For example, if real Gaussian random variable is considered, we have [106]:

$$\sigma = \frac{1}{0.6745} \cdot MAD \tag{4.3.4}$$

where the constant factor $S = \frac{1}{0.6745}$.

In wireless communications, zero mean complex additive white Gaussian noise (CWGN) is often considered. If random variable Z subjects to CWGN distribution, the amplitudes of Z subjects to Rayleigh distribution. The cumulative distribution function of Rayleigh distribution is expressed as:

$$F(r) = 1 - e^{-r^2/2\sigma_1^2} \tag{4.3.5}$$

where σ_1 ($\sigma = \sqrt{2}\sigma_1$) is the standard deviation of either the real or imaginary parts of Z.

When $F(r) = 0.5$, the corresponding value of r is the median of $|Z|$, therefore the relation between the standard deviation of $|Z|$ and its median value can be written as [115]:

$$\sigma = \sqrt{2}\sigma_1 = \sqrt{2}\frac{\text{median}(|Z|)}{\sqrt{\ln 4}} \tag{4.3.6}$$

Here the constant factor is $S = \sqrt{\frac{1}{\ln 2}}$.

If the samples of a signal are all noise coefficients, most of the three noise standard deviation methods are all effective. However, if the coefficients are the samples of a sparse signal contaminated by noise, the estimated standard deviations based on the empirical sample standard deviation and Bayesian method are obviously not the noise standard deviation due to the presence of signal components, while for the MAD method, the estimated standard deviation can be approximately considered as the estimate of noise standard deviation. In the section 4.4, the performance of these three methods will be compared and discussed.

4.3.2 Threshold Analysis

Threshold is an important parameter for effective detection, which is essential for effective sparse channel estimation in OFDM systems. Probability of detection (POD) and false alarm rate (FAR) are two important criteria for determining the threshold. The details for POD and FAR will be discussed in the following.

Assume that $\boldsymbol{h} \in C^L$ is a K sparse channel vector whose non-zero coefficients follow the distribution of complex Gaussian with zero means and indices set Ω; $\boldsymbol{z} \in C^L$ is a noise vector with each element following the distribution $CN(0, \sigma_z^2 \boldsymbol{I}_L)$, therefore, the coefficients of observed noisy sparse channel vectors $\boldsymbol{r} \in C^L$ and its distribution are given by:

$$\boldsymbol{r} = \boldsymbol{h} + \boldsymbol{z} \tag{4.3.7}$$

$$r_i \sim \begin{cases} CN(0, \sigma_{h_i}^2 + \sigma_z^2) & i \in \Omega \\ CN(0, \sigma_z^2) & i \notin \Omega \end{cases} \quad 0 \le i \le L-1 \tag{4.3.8}$$

where $\sigma_{h_i}^2$ is the variance of the non-zero channel coefficient with index i.

POD reflects the probability of a signal being detected. Each element of $\boldsymbol{h}$ is statistically independent and subjects to complex Gaussian distribution, but has different variance. Meanwhile, the elements of $\boldsymbol{z}$ are also statistically independent, subject to complex Gaussian distribution and have the same variance. Therefore, the POD of observed channel vector $\boldsymbol{r}$ is the mean value of POD of its each element, which is given by:

$$P_d = \frac{1}{K} \sum_{i \in \Omega} P_{d_i} \tag{4.3.9}$$

where P_{d_i} is the POD of the $i^{th}, i \in \Omega$ element, which is given mathematically by:

$$P_{d_i} = \int_{Th}^{\infty} f_{r_i}(t) \mathrm{d}t \tag{4.3.10}$$

where $f_{r_i}(t)$ is the probability density function of $|r_i|$ (the amplitude of the i^{th} element of $\boldsymbol{r}$); Th is the threshold value. r_i subjects to $CN(0, \sigma_{h_i}^2 + \sigma_z^2) \quad i \in \Omega$, its amplitude subjects to Rayleigh distribution, therefore, we have:

$$P_{d_i} = \int_{Th}^{\infty} \frac{2t}{\sigma_{r_i}^2} e^{\frac{-t^2}{\sigma_{r_i}^2}} \mathrm{d}t = e^{-\frac{Th^2}{\sigma_{r_i}^2}} \tag{4.3.11}$$

where $\sigma_{r_i}^2 = \sigma_{h_i}^2 + \sigma_z^2$ is the variance of the i^{th} element of the vector $\boldsymbol{r}$.

FAR reflects the probability of false detection of noise coefficients as the signal coefficients. All elements of vector $\boldsymbol{z}$ are statistically independent, subject to Rayleigh distribution and have the same variance, therefore, the FAR of $\boldsymbol{z}$ is mathematically given by:

$$P_{FAR} = P_{FAR_m} = \int_{Th}^{\infty} \frac{2t}{\sigma_z^2} e^{\frac{-t^2}{\sigma_z^2}} \mathrm{d}t = e^{-\frac{Th^2}{\sigma_z^2}} \tag{4.3.12}$$

where P_{FAR_m} is the FAR of the m^{th} $m \in \Omega^c$ element of $\boldsymbol{z}$, which is equal to the P_{FAR} (the FAR of $\boldsymbol{z}$), Th is the threshold value.

From (4.3.11) and (4.3.12), either POD and FAR are related to the chosen threshold, therefore, optimal threshold should optimally balance the POD and FAR to obtain optimal MSE. Moreover, optimal threshold requires the prior knowledge of both the channel statistics (channel power delay profile (PDP), which is also the variance of none-zero channel taps and the exact channel sparsity) and the noise variance [40]. Consequently, the threshold estimation methods can be classified into two categories in terms of using or not using the prior knowledge of the channel statistics. The first category requires the prior knowledge of channel statistics [37, 116], which significantly increases the complexity and the burden of the whole communication system. Comparatively, the second category of thresholds, which doesn't require the prior knowledge of channel statistics and only relies on the noise power or noise variance [73, 117] is preferred.

The second category of thresholds generally has something to do with the noise variance for the detection of powers of the channel taps or the noise standard deviation for the detection of amplitudes of channel taps. Here, the amplitudes of channel taps are considered. A typical example is the threshold proposed by Kang [73], which has the following formula:

$$Th_1 = \sqrt{2}\sigma_z \tag{4.3.13}$$

where σ_z is the noise standard deviation.

Consider the CWGN vector $\boldsymbol{z} \sim \mathbb{CN}(0, \sigma_z^2 \boldsymbol{I}_L)$. The false alarm rate (FAR) of the above threshold can be calculated by the following formula:

$$P\{|z_i| \geq \sqrt{2}\sigma_z\} = e^{-2} \approx 0.1353 \tag{4.3.14}$$

Obviously, the above threshold can only eliminate the noise coefficient with probability of 0.8647, there is still a probability of 0.1353 that the noise coefficients still exist.

Beyond the above threshold, the universal threshold firstly proposed in [106], is widely used in compressed channel sensing [113], [27], [19]:

$$Th_2 = \sqrt{2\ln L}\sigma_z \tag{4.3.15}$$

The threshold given in (4.3.15) depends only on σ_z and L, which is the number of elements. If $L \to \infty$, the threshold tends to positive infinity. However, practically, we can hardly deal with the signal with length of infinity. We do care about is the speed of the probability of $|z_i| \geq \sqrt{2lnL}\sigma_z$ approaching to zero. It has important impact on the FAR. The FAR of the universal threshold is derived in the Appendix B.

From the aspect of FAR, it is not difficult to show that the universal threshold is obviously better than the threshold proposed by Kang by comparing the false alarm rate of threshold of Kang and that of the universal threshold with a specific value of L. As to the POD, if the channel power delay profile of non-zero elements is known, it can be calculated by (4.3.11). The following section will mainly focus on developing an effective threshold for sparse channel estimation in OFDM system.

4.4 Proposed Time Domain Threshold for LS based Channel Estimation in OFDM System

In this section, an effective threshold is proposed for LS based channel estimation in OFDM system. Firstly, we consider the time domain LS estimator $\boldsymbol{h}_{LS}$ in (2.3.14):

$$\hat{\boldsymbol{h}}_{LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y}_p \tag{4.4.1}$$

Combining (3.2.3) and (4.4.1), the following formula is obtained:

$$\hat{\boldsymbol{h}}_{LS} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{A}\boldsymbol{h} + (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{v}_p \tag{4.4.2}$$

Consider the case where $L_{cp} \leq M < N$ (N is an integer multiple of M) and the pilots are uniformly distributed, in this case $\boldsymbol{A}^H \boldsymbol{A} = M\boldsymbol{I}_{L_{cp}}$, (4.4.2) can be rewritten as:

$$\hat{\boldsymbol{h}}_{LS} = \boldsymbol{h} + \boldsymbol{n} \tag{4.4.3}$$

where $\boldsymbol{n} = \frac{1}{M}\boldsymbol{A}^H \boldsymbol{v}_p$, it is a linear combination of independent Gaussian noise, therefore, it is still an AWGN vector, with covariance matrix $\boldsymbol{C}$ expressed as:

$$\boldsymbol{C} = E(\boldsymbol{n}\boldsymbol{n}^H) = \frac{1}{M}\sigma^2 \boldsymbol{I}_{L_{cp}} \tag{4.4.4}$$

From (4.4.3) and (4.4.4), the estimated CIR by LS is highly affected by the noise when the channel is sparse. Therefore, it is important to denoise the estimated CIR by an appropriate threshold. To do this, the universal threshold introduced in (4.3.15) is employed:

$$\lambda = \sqrt{2\ln L_{cp}}\sigma_n \tag{4.4.5}$$

An accurate estimation of standard deviation σ_n of each element in the noise vector $\boldsymbol{n}$, is necessary in practical communication. However, it is difficult to obtain an effective estimate of standard deviation of each element in $\boldsymbol{n}$ when the noise vector $\boldsymbol{n}$ and the channel vector $\boldsymbol{h}$ are present together. In the following, we propose an efficient threshold based on the estimated CIR and the characteristics of the sparse channel.

For sparse channel, the majority of coefficients in CIR are noise, therefore it is possible to obtain an approximated estimation of noise standard deviation $\hat{\sigma}'_n$ by using (4.3.6) and the

initial estimated CIR $\boldsymbol{h}_{LS}$:

$$\hat{\sigma}'_n = \sqrt{2}\hat{\sigma}' = \sqrt{2}\frac{\text{median}\left|\hat{\boldsymbol{h}}_{LS}\right|}{\sqrt{\ln 4}} \tag{4.4.6}$$

However, the presence of channel taps results in bias in the estimated noise standard deviation, especially when the SNR is high and the channel is not sufficiently sparse. In general, we have $\hat{\sigma}'_n > \sigma_n$. If the majority number of channel taps in $\hat{\boldsymbol{h}}_{LS}$ are removed, the remaining coefficients can be regarded approximately as noise coefficients, which can be used to get a better estimate of the noise standard deviation $\hat{\sigma}''_n$. With this estimated noise standard deviation $\hat{\sigma}''_n$, an effective threshold can be obtained.

4.4.1 Proposed Method

The main framework of the proposed threshold for sparse channel estimation is shown in Fig 4.4.1. The different steps are described as follows.

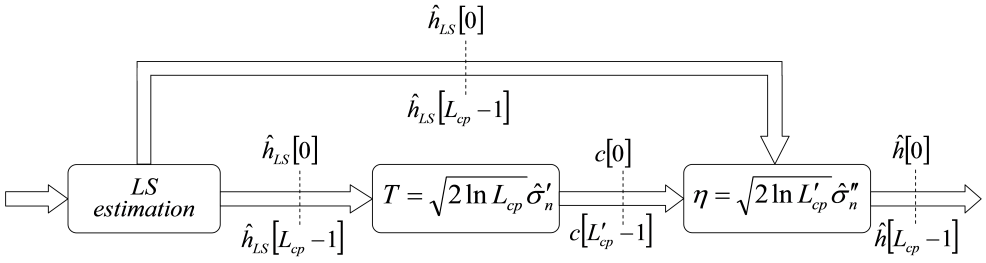

Figure 4.4.1: Proposed sparse channel estimation scheme

Step.1. LS is used to get an initial CIR estimate with length L_{cp}.

Step.2. To get a good estimate of the noise standard deviation, a threshold is needed to eliminate the majority of channel taps in the initial estimated CIR. An initial rough estimate of noise standard deviation can be obtained from the coefficients of sparse channel vector

$\hat{\boldsymbol{h}}_{LS}$ by (4.4.6). Then, $T = \sqrt{2\ln L_{cp}}\hat{\sigma}'_n$ is used as a threshold to eliminate the majority of channel taps present in the estimated CIR. By comparing with T, the vector of noise coefficients (the coefficients $\hat{h}_{LS}[j]$ with amplitude equal or smaller than T) denoted by $\boldsymbol{c}$ ($\boldsymbol{c} = [c[0], c[1], \ldots, c[L_{cp}{}' - 1]]$, $L_{cp}{}' < L_{cp}$) is extracted.

Step.3. With the vector of noise coefficients $\boldsymbol{c}$ (with no or much fewer channel taps than in $\hat{\boldsymbol{h}}_{LS}$), $\hat{\sigma}''_n$ is estimated by $\hat{\sigma}''_n = \sqrt{2}\frac{median(|\boldsymbol{c}|)}{\sqrt{\ln 4}}$. Then, an effective threshold is obtained by $\eta = \sqrt{2\ln L_{cp}{}'}\hat{\sigma}''_n$. The final estimated CIR is given by:

$$\hat{h}[n] = \begin{cases} \hat{h}_{LS}[n], & \left|\hat{h}_{LS}[n]\right| > \eta \\ 0, & \left|\hat{h}_{LS}[n]\right| \leq \eta \end{cases} \quad 0 \leq n \leq L_{cp} - 1 \tag{4.4.7}$$

4.4.2 Analysis and Performances Comparison

Simulations are carried out to evaluate the estimation performance of the proposed method and compare it with that of other existing methods. We consider two QPSK modulated OFDM systems for two different channel models:

1) An OFDM system with 1024 subcarriers, among which 64 subcarriers are pilots. The length of cyclic prefix is $L_{cp} = 64$.

2) An OFDM system with 1024 subcarriers, among which 256 subcarriers are pilots. The length of cyclic prefix is $L_{cp} = 256$.

For the first OFDM system, the channel model is a simplified version of DVB-T channel model whose channel impulse response is given in Table 1 [35].

For the second OFDM system, we use the ATTC (Advanced Television Technology Center) and the Grand Alliance DTV laboratory's ensemble E model whose CIR is given by [35]:

$$\begin{aligned} h[n] &= \delta[n] + 0.3162\delta[n-2] + 0.1995\delta[n-17] + 0.1296\delta[n-36] \\ &+ 0.1\delta[n-75] + 0.1\delta[n-137]. \end{aligned} \tag{4.4.8}$$

Table 4.1: CIR for the first OFDM system

Delay (OFDM samples)	Gain	Phase(rad)
0	0.2478	-2.5694
1	0.1287	-2.1208
3	0.3088	0.3548
4	0.4252	0.4187
5	0.4900	2.7201
7	0.0365	-1.4375
8	0.1197	1.1302
12	0.1948	-0.8092
17	0.4187	-0.1545
24	0.3170	-2.2159
29	0.2055	2.8372
49	0.1846	2.8641

The coefficients in (4.4.8) and the gains in Table 4.1 represent the standard deviation of the corresponding zero mean complex Gaussian random variable. In the simulations, one OFDM sample period is assumed to be the same as the unit delay of the channel, the CIR is static for each OFDM symbol duration and each OFDM symbol has a newly generated Rayleigh channel. Additionally, the channel tap gains are obtained by multiplying the CIR coefficients with zero mean unit variance complex Gaussian random variables. Moreover, 10 OFDM symbols are considered for each iteration; and there are totally 800 iterations. Therefore, 8000 independent channel realizations have been considered in each simulation. In the simulations, the same independent and identically distributed (i.i.d) complex Gaussian channels and noise are used for the algorithms with different pilot percentages and different SNR, while for the different OFDM symbols in each iteration and different iterations, different independent and identically distributed (i.i.d) complex Gaussian channels and noise are employed.

The simulations focus on the performance of bit error rate (BER) and normalized mean square error (NMSE) comparison between the proposed method, the classical method (frequency domain LS method with linear interpolation), Oracle estimator, LS method with

MST proposed by Minn et al [35] (MST method uses a fixed number of MST, which is double of the designed number of channel taps as recommended in [35]) and LS method with the threshold proposed by Kang et al [73] (for convenient comparisons, the exact noise standard deviation is used for the threshold proposed by Kang et al).

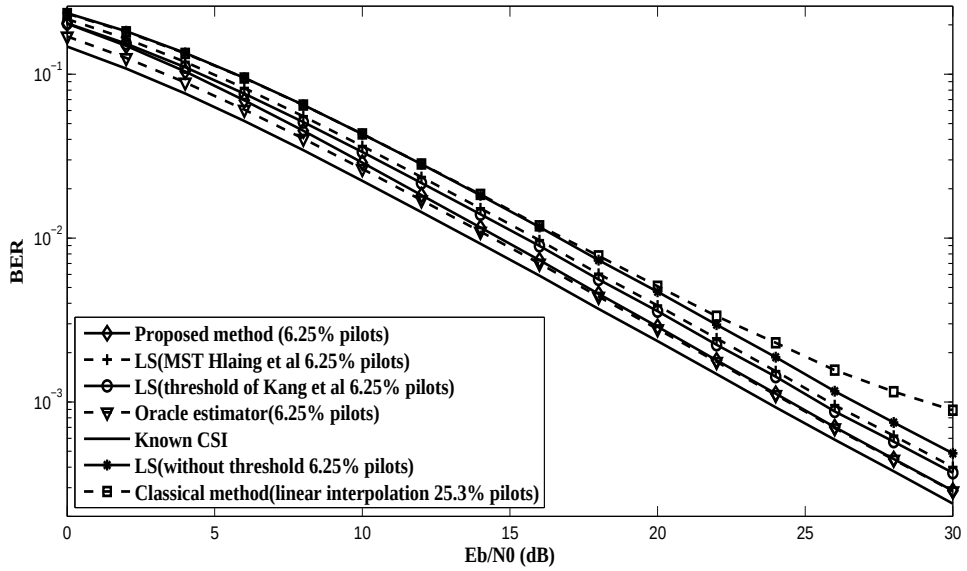

Figure 4.4.2: BER performance comparison of the first OFDM system

Fig 4.4.2 illustrates the BER performance of different algorithms of the first OFDM system. With only 6.25% of pilots, the proposed method outperforms the classical frequency domain method with 25.3% of pilots in the overall considered Eb/N0. Meanwhile, with the same percentage (6.25%) of pilots, compared with LS method with the threshold proposed by Kang et al and LS method with MST proposed by Minn et al, the proposed method achieves the same BER performance with at least 1dB gain in high Eb/N0 (13dB-30dB). Moreover, the proposed method has almost the same performance as Oracle estimator in the majority of considered Eb/N0 (8dB-30dB). Additionally, LS method without threshold has the poorest performance, there is at least 2dB gap in Eb/N0 between LS method without threshold and Oracle estimator for a same BER. Comparatively, the known CSI (instantaneous channel

frequency response is known) has the best BER performance, however, for the majority of considered Eb/N0 (8dB-30dB), the performance gap between the proposed method and known CSI is less than 1dB in Eb/N0.

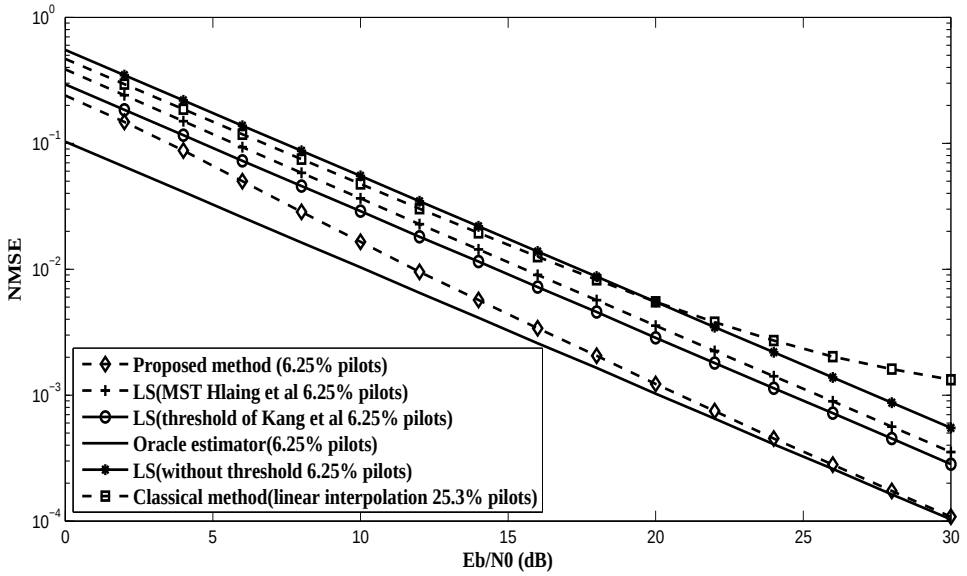

Figure 4.4.3: NMSE performance comparison of the first OFDM system

Fig 4.4.3 shows that the NMSE performance of the first OFDM system has similar trends as that of the BER performance except that the differences between algorithms are much more obvious. For example, when the NMSE reaches 10^{-3}, there is about 4dB gains in Eb/N0 for the proposed method compared with LS method with threshold proposed by Kang et al.

In Fig 4.4.4, the NMSE performance of different algorithms of the second OFDM system is compared. The proposed method maintains at least 4dB performance advantage in the all considered Eb/N0 for a same NMSE compared with LS method with MST proposed by Minn et al and LS method with the threshold proposed by Kang et al, which is bigger than in the first OFDM system. Furthermore, in the second OFDM system, the NMSE performance gap between the proposed method and Oracle estimator is slightly smaller than in the first OFDM system due to more accurate estimation on noise standard deviation and changes on

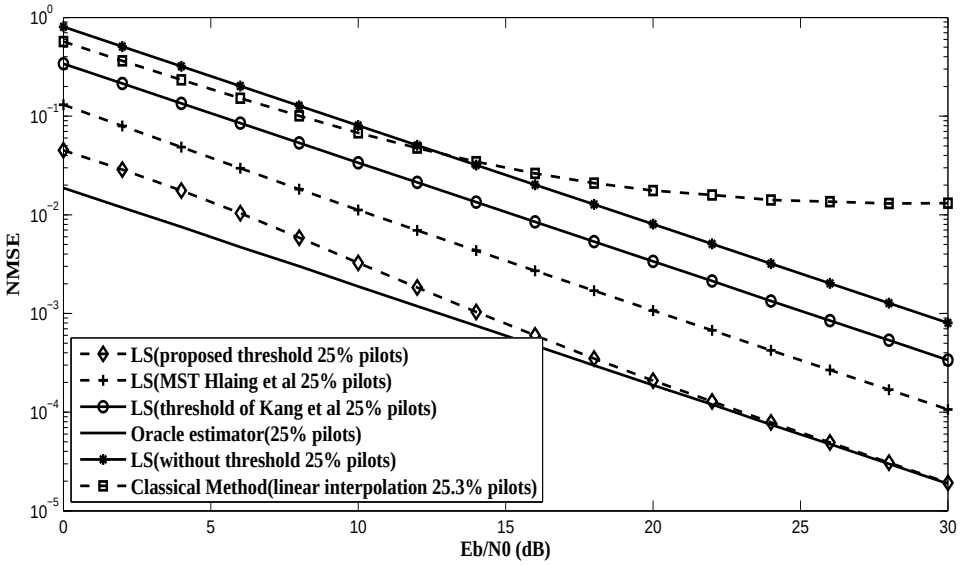

Figure 4.4.4: NMSE performance comparison of the second OFDM system

σ_n (See (4.4.4), σ_n will be reduced with the increase of the number of pilots).

The noise standard deviation estimation plays a central role in the proposed method. In order to show the performance of the proposed noise standard deviation estimation method, we consider the absolute relative error on the estimated standard deviation $\varepsilon = \frac{\sum_{i=0}^{N-1} |\hat{\sigma}_n[i]-\sigma_n|}{N|\sigma_n|}$ (σ_n is the actual noise standard deviation for each element of vector $\boldsymbol{n}$ and $\hat{\sigma}_n$ is the estimated one and N is the number of Monte Carlo simulations) of different estimation methods for both DVB and DTV channels with 8000 independent channel realizations. As can be seen from Table 4.2, for both channel models, when Eb/N0 increases, there is slight changes on the relative estimation error of standard deviation estimation for both the MAD method $\hat{\sigma}'_n = \sqrt{2}\frac{\text{median}|\hat{\boldsymbol{h}}_{LS}|}{\sqrt{\ln 4}}$ in (4.4.6) (used in the first threshold to eliminate the majority of channel taps) and the proposed method (the improved standard deviation estimation after eliminating the majority of channel taps). We can also see that the relative estimation error of the proposed method is smaller than the MAD based estimation method. Furthermore, the sample standard deviation estimation $\hat{\sigma}'''_n = \sqrt{\frac{1}{L_{cp}} \sum_{i=0}^{L_{cp}-1} |\hat{h}_{LS}[i] - mean(\hat{\boldsymbol{h}}_{LS})|^2}$ and Bayesian

estimation method [114] provide biased estimation, especially for high Eb/N0, due to the presence of non-zero channel taps.

Table 4.2: Absolute relative error of different noise standard deviation estimation methods

Channel Model	DVB Model			DTV Model		
Eb/N0(dB)	10	20	30	10	20	30
Proposed method	0.0860	0.0808	0.0778	0.0368	0.0363	0.0368
MAD method	0.1650	0.1785	0.1778	0.0391	0.0392	0.0394
Definition of sample standard deviation	3.3607	12.4768	41.5017	2.6978	10.3039	34.6210
Bayesian method	3.5083	12.9357	42.9288	2.9407	11.0520	36.9536

From the above analysis of simulation results, we observe that even though the sparsity rate [118] $\frac{K}{L_{cp}}$ has changed significantly (2.34% for the second channel and 18.75% for the first channel), the proposed method still maintains good performance of both the noise standard deviation estimation and channel estimation. Therefore, we can draw the conclusion that even without prior knowledge of channel statistics and noise standard deviation, the proposed method can still work efficiently within a wide range of sparsity rate.

Table 4.3: (Complex) Computational complexity comparison

	Proposed Method		Classical Method
Alg	LS+Thr	FFT	LS+Lin Inter
Comp	$O(\frac{M}{2}log_2M + M)$	$O(\frac{N}{2}log_2N)$	$O(N)$

In Table 4.3, the computational complexity (for simplicity only complex multiplications are considered) comparison for the proposed method and the classical frequency domain method is presented. The proposed method is composed of three algorithms, which are LS algorithm (time domain), threshold estimation algorithm and FFT algorithm. The LS and threshold estimation have a total complexity of $O(\frac{M}{2}log_2M + M)$ (in the case of this thesis, LS algorithm can mainly be realized by a M size IFFT computation), while the FFT method has complexity of $O(\frac{N}{2}log_2N)$. Therefore, the total computational complexity of the proposed method is around $O(Nlog_2N)$. Obviously, the proposed method has higher

computational complexity than that of the classical method, which is the combination of LS (frequency domain) and linear interpolation algorithms and has computational complexity of $O(N)$. However, for the same performance of BER and NMSE, the proposed method allows to use fewer pilots than that of the classical frequency domain method, thus it has better spectral efficiency.

4.5 Conclusion

In this chapter, we firstly review the frequency interpolated based channel estimation. Then, the noise characteristics, noise standard deviation estimation and threshold are presented. Finally, an effective threshold is proposed in the case of $M \geq L_{cp}$ with good spectral efficiency, effective channel estimation performance, low computational complexity and no requirement for the prior knowledge of both channel statistics and noise standard deviation. The proposed method is actually effective for sparse channels. In the next chapter, the case of $M < L_{cp}$ is considered to realize high spectral efficiency transmission.

Chapter 5

Compressed Sensing Based Sparse Channel Estimation

In chapter 4, traditional time domain LS method is adopted to realize effective sparse channel estimation in OFDM system in the case of $M \geq L_{cp}$. In order to obtain higher spectral efficiency, the case of $M < L_{cp}$ should be considered. In the case of $M < L_{cp}$, traditional time domain LS method cannot be used due to the rank deficiency of the measurement matrix, however, it can be resolved by compressed sensing (CS) effectively. This chapter addresses both the sample spaced and non-sample spaced sparse channel estimation in the case of $M < L_{cp}$. Firstly, for sample spaced sparse channel, we propose an effective threshold for CS based sparse channel estimation without the prior knowledge of channel statistics and noise standard deviation. Secondly, smart measurement matrix is adopted for non-sample spaced sparse channel estimation, which can effectively reduce the computational complexity, promote the spectral efficiency and achieve effective channel estimation performances.

5.1 Compressed Sensing Based Sample Spaced Sparse Channel Estimation

In the previous chapter, the LS based sample spaced sparse channel estimation is discussed for $M \geq L_{cp}$. In this chapter, the case $M < L_{cp}$, an interesting topic in CS, will be discussed. The mathematical model in (3.2.3) is considered, which is given by:

$$\boldsymbol{y}_p = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_p \tag{5.1.1}$$

In the case of $M \geq L_{cp}$, the measurement matrix $\boldsymbol{A}$ is a full column rank matrix, which effectively guarantees the channel estimation by LS method. However, if $M < L_{cp}$ is considered, the measurement matrix $\boldsymbol{A}$ is a rank deficient matrix. If $\boldsymbol{h}$ is a rich multipath channel, it can hardly be reconstructed, however, if $\boldsymbol{h}$ is a sparse channel, it is possible to reconstruct $\boldsymbol{h}$ under some specific conditions. Because for sparse channel reconstruction, $\boldsymbol{A}_\Gamma$ mentioned in chapter 3, (Γ is the support of $\boldsymbol{A}$), not $\boldsymbol{A}$ is really important. If $\boldsymbol{A}_\Gamma$ is a full column rank matrix, the $\boldsymbol{h}_\Gamma$ can still be reconstructed by a constrained LS, furthermore, sparse channel vector $\boldsymbol{h}$ can be obtained. To obtain the support Γ for both $\boldsymbol{A}$ and $\boldsymbol{h}$, different greedy pursuit methods are derived, which are presented in chapter 3.

5.1.1 A Novel Robust Threshold for Compressed Sensing Based Sample Spaced Sparse Channel Estimation

Pilot Arrangement for CS based Sparse Channel Estimation

The coherence of measurement matrix is an important parameter in CS, which guarantees the performance of the sparse signal's reconstruction. It is defined by (3.1.5). In general, the smaller the coherence of measurement matrix, the better the precision will be for the

sparse channel estimation. Taking the coherence of measurement matrix into account, the pilot arrangement can be realized by:

$$\arg\min_{\Gamma} \max_{0\leq i,j\leq L_{cp}-1, i\neq j} \frac{|\langle \boldsymbol{a}_i, \boldsymbol{a}_j\rangle|}{\|\boldsymbol{a}_i\|_2\|\boldsymbol{a}_j\|_2} \tag{5.1.2}$$

where Γ is the optimal pilot arrangement subset. There are $\binom{N}{M}$ possible pilot arrangements and it is actually computationally exhausted task. Therefore, the technique proposed in [59] is adopted, which proposes to randomly generate a limited number of pilot arrangements and search the suboptimal pilot arrangement.

m Selected Bases for the Partial CIR Reconstruction for Sparse Channel

In the case of $M < L_{cp}$, $Rank(A) < L_{cp}$. To obtain an effective threshold, the noise standard deviation should be estimated. However, due to the limitation of $Rank(A) = M$, reconstructing the whole CIR with length L_{cp} and estimating the noise standard deviation are impossible. In this case, partial CIR with $m \quad (m \leq M)$ can be extracted to estimate the noise standard deviation. There are two ways to realize this:

(1) Extract the first m $(m < M)$ coefficients [119];

(2) Extract the m coefficients by considering its corresponding bases, which satisfy that the mutual coherence of any two chosen bases should be smallest possible:

$$\arg\min_{i,j\in\Lambda, 0\leq i,j\leq L_{cp}-1, i\neq j} \frac{|\langle \boldsymbol{a}_i, \boldsymbol{a}_j\rangle|}{\|\boldsymbol{a}_i\|_2\|\boldsymbol{a}_j\|_2}, |\Lambda| = m \quad (m \leq M) \tag{5.1.3}$$

where Λ is the subset of m optimally selected bases.

m Selected Bases of the Measurement Matrix

To obtain Λ, there are $\binom{L_{cp}}{m}$ possible choices. It is a computational exhausted task. Therefore, threshold of coherence between bases is introduced to select m bases. In the proposed algorithm, the threshold of coherence can iteratively find the indices of the eliminated bases from the subset indices for candidate bases meanwhile renew the subset indices for eliminated bases, subset indices for candidate bases and the subset indices for selected bases. The specific realization of the algorithm is given by:

Input: Normalized Measurement matrix: $\bar{\boldsymbol{A}} = \frac{\boldsymbol{A}}{\sqrt{M}}$ (Defined in 3.1.2);

Subset indices for eliminated bases: $\Omega = \emptyset$;

Subset indices for candidate bases: $\Theta = \{1, 2, \ldots, L_{cp}\}$;

Subset indices for eliminated bases at each iteration: $\boldsymbol{aaa} = \emptyset$;

Subset indices for selected bases: $\Lambda = 1$

Threshold for coherence of bases: Th;

Current index: $ci = 1$.

$\boldsymbol{C} = abs(\bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}})$; Obtain the coherency matrix

$\boldsymbol{C}(i, i) = 0, i = 0, 1, \ldots, L_{cp} - 1$; Diagonal elements of coherency matrix are set to zero;

AlgorithmRealization

while $(max(max(\boldsymbol{C}(:, \Theta))) > Th$ & $max(max(\boldsymbol{C}(:, \Theta))) < 1)$ Exist the coherence indices for candidate bases within $(Th, 1)$?

$\boldsymbol{aaa} = find(\boldsymbol{C}(:, ci) > Th$ & $\boldsymbol{C}(:, ci) \neq 1)$; Find the indices of bases whose coherences with the basis of ci are bigger than Th and not equal to 1;

$\Omega = \Omega \bigcup \boldsymbol{aaa}$; Update the subset indices for eliminated bases;

$\boldsymbol{bbb} = \Omega^c$; Obtain the complement of Ω;

$\boldsymbol{ccc} = \boldsymbol{bbb} - \boldsymbol{\Lambda}$; Obtain the initial subset for candidate indices without the indices for selected bases;

$\boldsymbol{order} = randperm(length(\boldsymbol{ccc}))$; Obtain the random order for candidate indices without the indices for selected bases;

$ci = \boldsymbol{ccc}(\boldsymbol{order}(\mathbf{1}))$; Obtain the newly index for selected basis at the current iteration;

$\Lambda = [\Lambda\ ci]$; Update the subset indices for selected bases;

$\Theta = \boldsymbol{bbb} - \Lambda$; Update the subset of indices for candidate bases;

$\boldsymbol{C}(:, \Omega) = 1$; Set one to the columns of the coherent matrix with eliminated indices.

end

The elements of subset Λ are just the indices of the m selected bases with constrained coherence.

[119] uses the first m channel coefficients to estimate the noise standard deviation. However, $Rank(A) = m$ doesn't necessarily mean that the first m columns of $\boldsymbol{A}$ are linearly independent, therefore, the inverse computation in LS may not exist. Different from [119], if the m selected bases are optimized by minimizing the coherence of any two bases, it is highly likely that the m selected bases of $\boldsymbol{A}$ are linearly independent.

Reconstruction of the m Selected Coefficients of the Partial CIR and Error Vector

(5.1.1) can be rewritten as:

$$\boldsymbol{y}_p = \boldsymbol{A}_m \boldsymbol{h}_m + \boldsymbol{A}_b \boldsymbol{h}_b + \boldsymbol{v}_p \tag{5.1.4}$$

where $\boldsymbol{A}_m = [\boldsymbol{a}[p_0], \boldsymbol{a}[p_1], \ldots, \boldsymbol{a}[p_{m-1}]$ and $\boldsymbol{h}_m = [h[p_0], h[p_1], \ldots, h[p_{m-1}]]^T$ are the matrix with the m selected columns of $\boldsymbol{A}$ and vector with the m selected channel taps of $\boldsymbol{h}$ respectively; $\boldsymbol{A}_b$ and $\boldsymbol{h}_b = [h[p_m], h[p_{m+1}], \ldots, h[p_{L_{cp}-1}]]^T$ are the matrix with the columns

different from the m selected columns of $\boldsymbol{A}$ and vector with corresponding channel taps of $\boldsymbol{h}$ respectively.

Using LS method to estimate the m selected channel taps, we have:

$$\hat{\boldsymbol{h}}_{ls-m} = (\boldsymbol{A}_m^H \boldsymbol{A}_{\boldsymbol{m}})^{-1} \boldsymbol{A}_m^H \boldsymbol{y}_p \tag{5.1.5}$$

Combine (5.1.4) and (5.1.5), we have:

$$\hat{\boldsymbol{h}}_{ls-m} = (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H (\boldsymbol{A}_m \boldsymbol{h}_m + \boldsymbol{A}_b \boldsymbol{h}_b + \boldsymbol{v}_p) \tag{5.1.6}$$

If $\boldsymbol{A}_m^H \boldsymbol{A}_m$ is not deficient, (5.1.6) can be rewritten as:

$$\hat{\boldsymbol{h}}_{ls-m} = \boldsymbol{h}_m + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b \boldsymbol{h}_b + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{5.1.7}$$

which is composed of three parts, the first part is the m selected channel taps, the second one is the interference due to the significant taps in $\boldsymbol{h}_b$, which should be cared about, and the last part is the noise part.

It is necessary to reduce the impact of the significant taps $h[r_0], h[r_1], \ldots, h[r_{s-1}]$ of $\boldsymbol{h}_b$ in (5.1.7) with s the number of most significant taps in $\boldsymbol{h}_b$. Therefore, the significant taps estimated by OMP expressed by $\hat{h}_{omp}[d_0], \hat{h}_{omp}[d_1], \ldots, \hat{h}_{omp}[d_{K_{max}-1}],\ 0 \leq d_0, d_1, ..., d_{K_{max}-1} \leq L_{cp} - 1$ are introduced. The significant taps of $\hat{\boldsymbol{h}}_{omp}$ in subset Λ^c (the complement of Λ) are extracted to get a new sparse vector $\hat{\boldsymbol{h}}_{omp-b}$. Then, a new equation is obtained with regard to the second part of (5.1.7):

$$\hat{\boldsymbol{h}}_m = \boldsymbol{h}_m + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b (\boldsymbol{h}_b - \hat{\boldsymbol{h}}_{omp-b}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{5.1.8}$$

In order to reduce the impact of the significant taps in vector $\boldsymbol{h}_m$, the significant taps of

$\hat{\boldsymbol{h}}_{omp}$ in subset Λ are extracted to get a new sparse vector $\hat{\boldsymbol{h}}_{omp-m}$. Then, an estimated error vector $\hat{\boldsymbol{e}}_m$ can be obtained by eliminating the significant taps in $\hat{\boldsymbol{h}}_m$ in the first part of (5.1.8):

$$\hat{\boldsymbol{e}}_m = (\boldsymbol{h}_m - \hat{\boldsymbol{h}}_{omp-m}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b (\boldsymbol{h}_b - \hat{\boldsymbol{h}}_{omp-b}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{5.1.9}$$

5.1.2 Noise Standard Deviation Estimation and Threshold Estimation

If the initial estimated channel by OMP is accurate, $\hat{\boldsymbol{e}}_m$ approximates an error vector with length of m contaminated only by noise, which follows complex Gaussian distribution and its amplitude subjects to Rayleigh distribution. The noise standard deviation of the real part $\hat{\sigma}_r$ and imaginary part $\hat{\sigma}_i$ can be estimated by the median function in (4.3.6):

$$\hat{\sigma}_r = \hat{\sigma}_i = \hat{\sigma}_1 = \frac{median(|\hat{\boldsymbol{e}}_m|)}{\sqrt{ln4}} \tag{5.1.10}$$

Therefore, the noise standard deviation of $(\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p$ can be estimated by $\hat{\sigma} = \sqrt{2}\hat{\sigma}_1$. With the estimated $\hat{\sigma}$, the universal threshold can be employed for the detection of sparse channel [19, 106]:

$$T = R\hat{\sigma};\ (R = \sqrt{2ln(m)}) \tag{5.1.11}$$

5.1.3 Main Framework of the Proposed Method

Fig 5.1.1 gives the main framework of the proposed method, the specific steps of which are summarized as follows.

First step: The m bases are selected from the measurement matrix based on the threshold of coherence between bases.

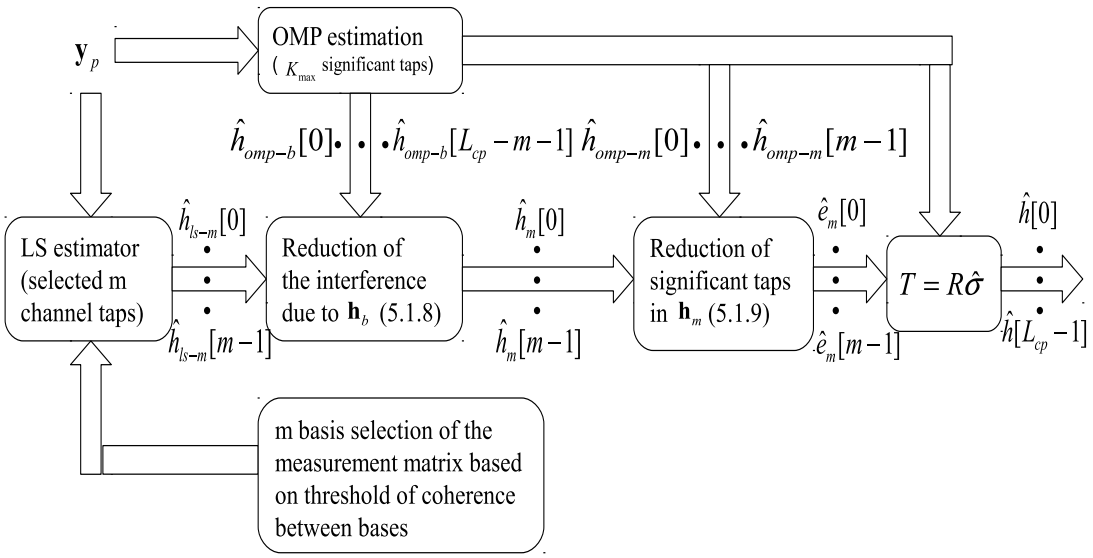

Figure 5.1.1: Proposed Threshold for OMP Algorithm

Second step: OMP algorithm is used to get an initial CIR with K_{max} significant taps and the rest of the channel taps are setting to zero. Meanwhile, LS estimator is employed to estimate the partial CIR (the m selected channel taps) with the received pilot vector $\boldsymbol{y}_p$ (M pilots).

Third step: In order to improve the estimation performance of CIR with the m selected channel taps, the interference caused by the significant taps in $\boldsymbol{h}_b$ is reduced by the estimated significant taps in $\hat{\boldsymbol{h}}_{omp-b}$.

Fourth step: In order to reduce the impact of the significant taps in $\boldsymbol{h}_m$, the estimated significant taps in $\hat{\boldsymbol{h}}_{omp-m}$ are introduced, thus an estimated error vector $\hat{\boldsymbol{e}}_m$ is obtained.

Fifth step: Get the noise standard deviation estimate $\hat{\sigma}$ by using the median of absolute value of $\hat{\boldsymbol{e}}_m$.

Sixth step: With $\hat{\sigma}$, the threshold $T = R\hat{\sigma}$ is used for the MST detection, which is given by:

$$\hat{h}[n] = \begin{cases} \hat{h}_{omp}[n], & |\hat{h}_{omp}[n]| \geq T \\ 0, & |\hat{h}_{omp}[n]| < T \end{cases}, \ 0 \leq n \leq L_{cp} - 1 \tag{5.1.12}$$

Simulations and Comparisons

In simulations, we consider a QPSK modulated OFDM system. The system has total bandwidth of 10MHz [120] and 1024 subcarriers, of which 128 subcarriers are pilots. The durations of the whole OFDM symbol and cyclic prefix are $128\mu s$ and $25.6\mu s$ respectively. The pilot pattern arrangement is obtained by the method given in [59] and the specific pilot positions in the simulations are [0,12,21,25,37,41,51,52,78,87,88,91,105,108,126,128,129,134,137,145,152,163, 175,176,178,180,193,205,210,216,223,236,249,252,269,272,280,293,295,299,303,316,323,334,340, 343,348,375,376,381,386,387,391,404,419,442,446,454,463,466,468,470,484,485,486,501,514,518 520,534,554,566,571,575,584,586,588,599,604,611,613,620,623,633,634,639,640,646,654,670,675 719,731,733,742,759,762,767,791,801,812,828,838,840,847,851,855,857,878,883,886,889,893,926 931,941,948,950,952,955,959,962,972,986,988,991,1013,1023]. Additionally, we consider $Th = 0.14$ and $K_{max} = 24$ for the proposed method.

Two different channel models are considered for the performance evaluation. The 12 tap Hilly Terrain model is adopted as the first channel model whose power delay profile is given in Table 5.1 [120]:

The second channel model used for performance evaluation is the ATTC (Advanced Television Technology Center) and the Grand Alliance DTV laboratory's ensemble E model whose CIR is given by [35]:

$$h[n] = \delta[n] + 0.3162\delta[n-2] + 0.1995\delta[n-17] + 0.1296\delta[n-36] + 0.1\delta[n-75] + 0.1\delta[n-137] \tag{5.1.13}$$

Table 5.1: Power delay profile for 12 tap Hilly Terrain channel

Delay in time [μs]	Relative power [dB]
0	-10
0.2	-8
0.4	-6
0.6	-4
0.8	0
2.0	0
2.4	-4
15	-8
15.2	-9
15.8	-10
17.2	-12
20	-14

The performance of both BER and MSE are compared for different channel estimation methods, such as the proposed method, the DFT based method with threshold proposed by Kang et al [73] (Firstly, frequency LS is combined with linear interpolation to get CFR, then, DFT transform is adopted to get CIR and threshold of Kang is used for significant tap detection, Finally, final CFR is obtained), method proposed in [119] and Oracle estimator etc.

Fig 5.1.2 illustrates the performance of BER comparison of different channel estimation methods for the first channel model. Even though with only 12.5% of pilots, the proposed method outperforms significantly the DFT based method with threshold proposed by Kang [73] with 25.3% of pilots throughout Eb/N0. Additionally, without the prior knowledge of either channel statistics and noise standard deviation, the proposed method achieves almost the same BER performance compared with OMP method with 12 taps (exact number of taps) and Oracle estimator in the overall considered Eb/N0. Moreover, if compared with OMP method with 24 tap, the performance gains range of the proposed method is about 2dB throughout Eb/N0. In the case $m = 80$, the performance of the method proposed in [119] is significantly degraded due to the bad condition of $\boldsymbol{A}_m^H \boldsymbol{A}_m$ in (5.1.6), comparatively, the

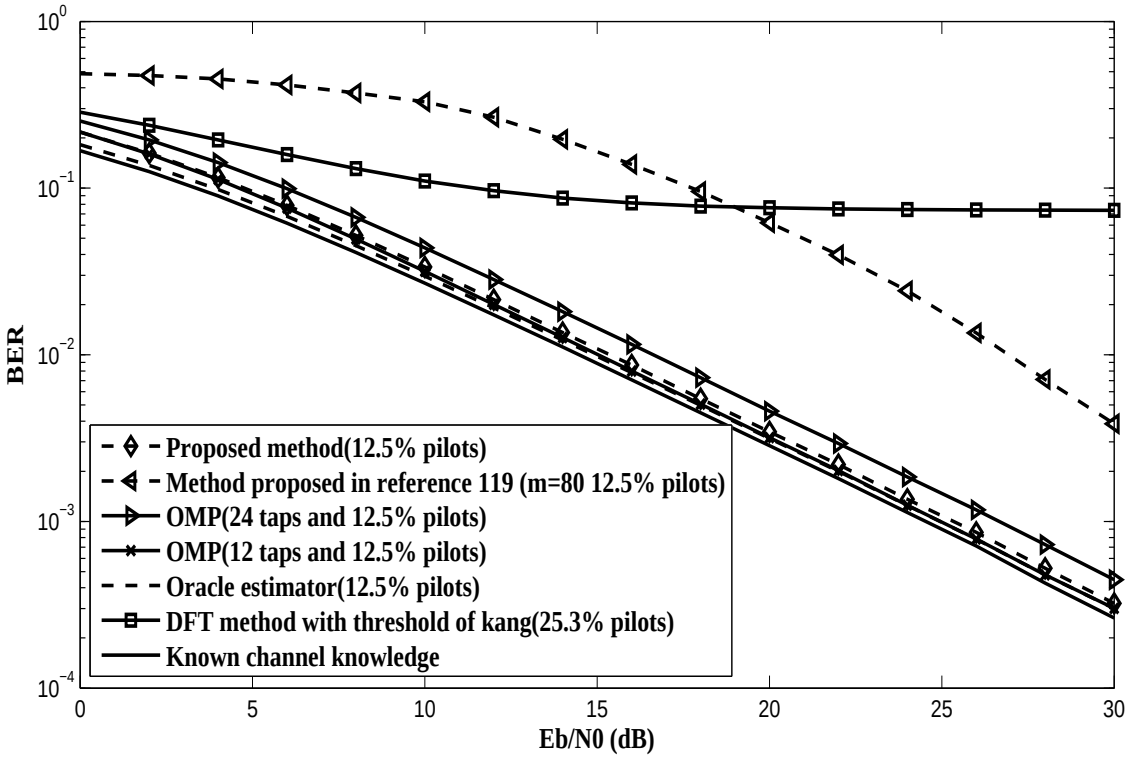

Figure 5.1.2: Performance of BER for the first channel model

proposed method maintains effective performance due to the better choice of the m selected bases.

Generally, the MSE performance of different algorithms for the first channel model has better distinction than the BER performance. Two facts can be observed from Fig 5.1.3. The first observation is that when compared with OMP method with 12 taps, the proposed method has at most 2.5dB performance degradation on MSE in the overall Eb/N0 mainly due to the sub-optimal m selected channel coefficients and the reduced interference in the first and second parts of (5.1.9). The second observation is that if compared with OMP method with 24 taps, the performance gains range of the proposed method is 4-6dB.

As shown in Fig 5.1.4, the performance of BER has similar trend for different algorithms for the second channel model compared with the first channel model.

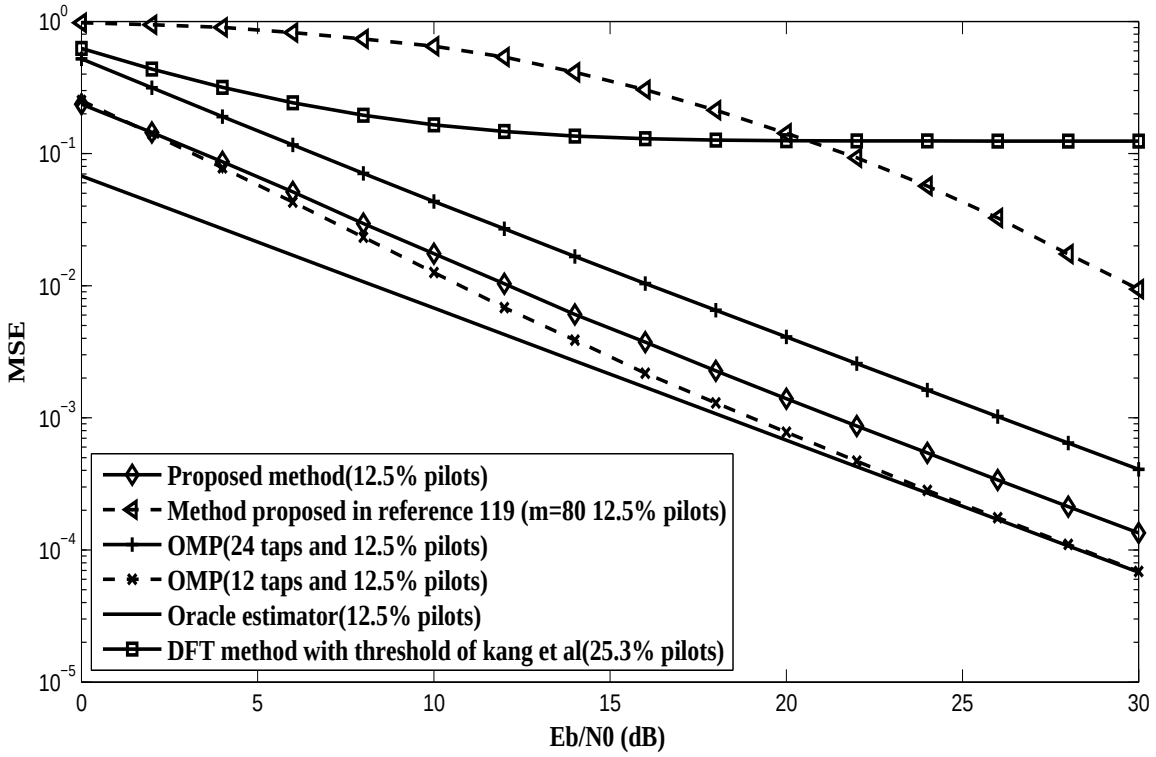

Figure 5.1.3: Performance of MSE for the first channel model

Table 5.2: Computational complexity comparison

	Proposed Method		DFT based method
Alg	OMP+Thr	FFT	Lin+FFT
Comp	$O(ML_{cp}K_{max})$	$O(Nlog_2N)$	$O(Nlog_2N)$

Table 5.2 evaluates the complexity of the proposed method and the traditional DFT based method. As can be seen from the table, the computational complexity of the proposed method is composed of that of OMP method $O(ML_{cp}K_{max})$, threshold estimation method $O(M^2m)$ and FFT method $O(Nlog_2N)$. In the proposed algorithm, $O(ML_{cp}K_{max}) \approx O(M^2m)$ and its total complexity is about $O(ML_{cp}K_{max})$. The DFT based method has computational complexity which is composed of linear interpolation $O(N)$ and FFT $O(Nlog_2N)$, its total computational complexity is around $O(Nlog_2N)$. The proposed method has a higher

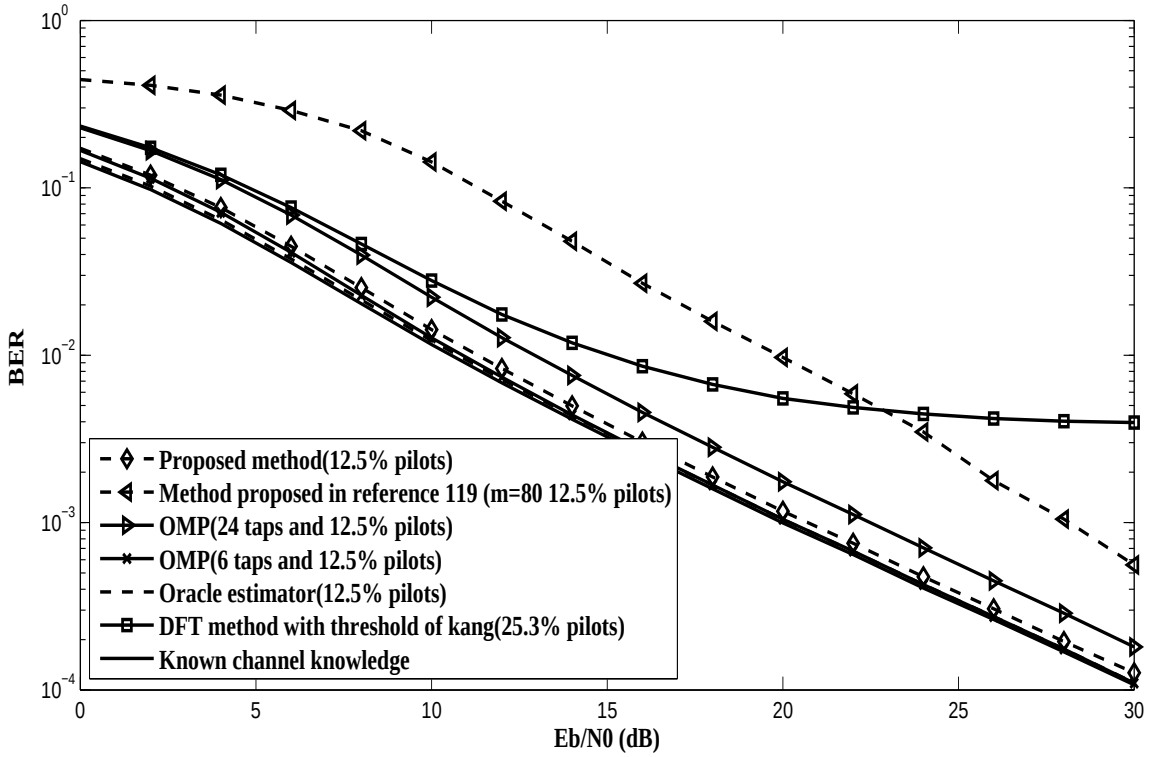

Figure 5.1.4: Performance of BER for the second channel model

computational complexity than that of the DFT based method, however, the performance of the proposed method is much better than the DFT based method.

Note: The complexity of the pilots arrangement finding and m bases selection are not included in any of the above algorithms, because both of these two algorithms can be realized previously to the actual communications.

5.2 Non-Sample Spaced Sparse Channel Estimation based on Compressed Sensing

The model of multipath channel and the observed CIR concerning the sampling interval T_s are given by equations (2.1.7) and (2.1.10) respectively. From the analysis, there is no leakage for the sample spaced channel since for a particular delay, all the power is mapped into the corresponding channel tap. However, if the non-sample spaced channel is considered, the path power corresponding to a delay is mapped to other channel taps, this is the phenomenon of leakage. Therefore, how to reduce leakage becomes a critical issue. For a wireless communication system, T_s is a constant parameter and cannot be changed. However, the physical channel reveals the characteristics of sparsity, therefore, finer resolution $T_s{}' = T_s/R$ (R is the oversampling factor) of CIR can be considered. (2.1.7) can be rewritten as [19]:

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - R\tau_l T_s{}') \tag{5.2.1}$$

The channel frequency response of $h\left[\tau\right]$ can be written as:

$$\begin{aligned} g[f] &= \int_{-\infty}^{\infty} h[\tau] e^{-j2\pi f\tau} \mathrm{d}\tau \\ &= \sum_{l} \alpha_l e^{-j2\pi f R\tau_l T_s{}'} \end{aligned} \tag{5.2.2}$$

Denote $f_k = \frac{k}{RNT_s{}'},\ k = 0, 1, 2, \ldots, RN-1$, taking the FFT of $h[lT_s{}']$, $l = 0, 1, 2, \ldots, RN-1$, we obtain the discrete channel frequency response:

$$g[k] = \sum_{l} \alpha_l e^{-j2\pi R\tau_l k/RN} \tag{5.2.3}$$

Its IFFT with size of RN is used to obtain the channel impulse response:

$$\begin{aligned} h[n] &= \frac{1}{RN}\sum_{l}\alpha_l \sum_{k=0}^{RN-1} e^{j2\pi(n-R\tau_l)k/RN} \\ &= \frac{1}{RN}\sum_{l}\alpha_l \frac{1-e^{j2\pi(n-R\tau_l)}}{1-e^{j2\pi(n-R\tau_l)/(RN)}} \end{aligned} \tag{5.2.4}$$

Let $n - R\tau_l = x$, (5.2.4) becomes:

$$\frac{1}{RN}\sum_{l}\alpha_l \frac{1-e^{j2\pi x}}{1-e^{j2\pi x/(RN)}} \tag{5.2.5}$$

Consider the case where $x \to 0$, we have:

$$\begin{aligned} &\lim_{x\to 0} \frac{1}{RN}\frac{1-e^{j2\pi x}}{1-e^{j2\pi x/(RN)}} \\ &= \lim_{x\to 0} \frac{1}{RN}\frac{-j2\pi x + o(x^2)}{-j2\pi x/(RN)} \\ &= 1 \end{aligned} \tag{5.2.6}$$

For the case where $R \to \infty$ and $x \not\to 0$, we have:

$$\begin{aligned} &\lim_{R\to\infty}\lim_{x\not\to 0} \frac{1}{RN}\frac{1-e^{j2\pi x}}{1-e^{j2\pi x/(RN)}} \\ &= \lim_{R\to\infty}\lim_{x\not\to 0} \frac{1}{RN}\frac{1-e^{j2\pi x}}{1-e^{-j2\pi\tau_l/N}e^{j2\pi n/(RN)}} \\ &= 0 \end{aligned} \tag{5.2.7}$$

Combining the derivation of the two cases, in the case of $R \to \infty$, we can theoretically obtain:

$$h[n] = \begin{cases} \alpha_l, & n = R\tau_l \\ 0, & n \neq R\tau_l \end{cases} \quad R \to \infty \tag{5.2.8}$$

From (5.2.8), when $R \to \infty$, $h[n]$ becomes a continuous channel, therefore, the power of path $\delta(t-\tau_l)$ will be completely mapped into $h[R\tau_l]$ when $R \to \infty$. It is obviously not realistic for digital processors in wireless communication, however, R can be properly chosen to make a tradeoff between the computational complexity and channel estimation accuracy. In the following, a novel smart redundant dictionary design method is proposed to reduce the computational complexity, meanwhile, maintaining effective channel estimation performance.

5.2.1 Smart Measurement Matrix Design for Non-sample Spaced Sparse Channel Estimation

In chapter 3, finer resolution measurement matrix $\boldsymbol{A}_{fr}$ with oversampling factor R, which has the total dimensions of $M \times (R(L_{cp}-1)+1)$ is given in (3.2.7):

$$\boldsymbol{A}_{fr}= \begin{bmatrix} a_{k_0 0} & a_{k_0 1} & \dots & a_{k_0 (L_{cp}-1)R} \\ a_{k_1 0} & a_{k_1 1} & \dots & a_{k_1 (L_{cp}-1)R} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k_{M-1} 0} & a_{k_{M-1} 1} & \dots & a_{k_{M-1} (L_{cp}-1)R} \end{bmatrix} \tag{5.2.9}$$

where $a_{k_m n} = x_{k_m} e^{\frac{-j2\pi k_m n}{RN}}, \quad 0 \le m \le M-1,\, 0 \le n \le (L_{cp}-1)R.$

If it is employed to reconstruct the non-sample spaced sparse channel, it will cause considerable computational cost due to the characteristics of the sparse channel. [121] proposes a simple CS based channel estimation algorithm for time varying sparse channel in OFDM system. Under the assumption of $M > L_{cp}$ and equispaced pilots arrangement, the author fully utilizes the monotonicity property of the delay matching function within $[l-1/2, l+1/2]$ ($l, 0 \le l \le L_{cp}-1$ is an integer sampling) to significantly reduce the computational cost. According to [121], specifically, assume that $\boldsymbol{g}_p$ is the partial frequency response including only the pilots locations (pilots are equispaced and $M > L_{cp}$), $\boldsymbol{h}$ is the channel impulse response

containing only one channel tap (sparsity $K = 1$), the delay τ'' is uniformly distributed over $[0, L_{cp} - 1]$, we have:

$$|< \boldsymbol{f}_l, \boldsymbol{g}_p >| = |\bar{g}_0 \sum_{m=0}^{M-1} e^{\frac{-j2\pi k_m \Delta l}{N}}| = |\bar{g}_0| \varphi(\Delta l) \tag{5.2.10}$$

where $\boldsymbol{f}_l$ is the l^{th} column of $\boldsymbol{F}_{M \times L_{cp}}$, which is Fourier matrix with original resolution (defined in section 3.2.2); $\Delta l = \tau'' - l,\ \Delta l \in [-L_{cp} + 1, L_{cp} - 1]$ is the deviation between the delay of the only one channel tap τ'' and the integer sampling position l; $|\bar{g}_0|$ is the amplitude of the channel tap; $\varphi(\Delta l) = |\frac{sin(\pi \Delta l)}{sin(\pi \Delta l / M)}|$ is function of Δl; We have $\varphi(\Delta l) = \varphi(-\Delta l)$, therefore, $\Delta l \in [0, L_{cp} - 1]$ is considered. According to the result of [121], two properties are given:

(1) $\Delta l \in [0, 1/2]$, $\varphi(\Delta l)$ monotonously decreases as Δl increases;

(2) $\Delta l \in [1/2, L_{cp} - 1]$, $\varphi(\Delta l) < \varphi(1/2)$. For simplicity, this thesis doesn't gives specific proof, for details see Appendix of [121]. From the above analysis, the algorithm proposed in [121] has an important assumption, which requires $M > L_{cp}$. However, $M > L_{cp}$ can not satisfy the high spectral efficiency transmission in many cases, therefore, $M < L_{cp}$ should be considered. In the case of $M < L_{cp}$, the equispaced pilots arrangement can hardly be effective [59, 63]. In order to obtain optimal pilot arrangement, we should find solution to (5.1.3), which requires that pilot arrangement should minimize the coherence of the measurement matrix, however, from the previous analysis, it is a computational exhausted task. (Actually, equispaced pilot arrangement can be an optimal solution to (5.1.3) in the special case of $M > L_{cp}$). In order to realize effective sparse channel estimation in the case of $M < L_{cp}$, this work adopts the suboptimal pilot arrangement proposed in [59], which is already considered in the sample spaced sparse channel estimation ($M < L_{cp}$) in the previous section in this chapter. The question is that in the case of $M < L_{cp}$ and non-equispaced pilot arrangement, are the monotonicity properties of $\varphi(\Delta l)$ in $\Delta l \in [0, 1/2]$ and $\varphi(\Delta l) < \varphi(1/2)$ when $\Delta l \in [1/2, L_{cp} - 1]$ still satisfied? Intuitively, the answers are no. From (5.2.10), in the case of

$M < L_{cp}$, $|\sum_{m=0}^{M-1} e^{\frac{-j2\pi k_m \Delta l}{N}}|$ is related to the pilot arrangement and it can hardly be expressed by a certain formula. However, the fact that $\tau \in [-1/2+l, l+1/2]$ with comparatively high probability still exists. This fact can be full utilized to reduce computational complexity.

In order to reduce the computational complexity, we introduce the concept of smart measurement matrix. Smart measurement matrix is a much more efficient measurement matrix compared with (5.2.9), which contains oversampling bases only in the "hot zones" (black square) as shown in Fig 5.2.1. "hot zones" can be described as the zones, which may contain the significant taps with high probability. Practically, the smart measurement matrix can be obtained by employing the measurement matrix with original resolution to detect those "hot zones" and setting finer resolution only in those "hot zones". The specific realization is given by:

In the first step, the indices in the measurement matrix with original resolution related with "hot zones" can be initially detected by the following procedures:

$\boldsymbol{c} = \boldsymbol{A}^H \boldsymbol{y}_p$; (Project $\boldsymbol{y}_p$ onto $\boldsymbol{A}$.);

$\Omega = supp(\boldsymbol{c}_S)$; (Find the S highest coherences in coherence vector $\boldsymbol{c}$.)

where $\boldsymbol{c}_S$ is a vector composed by S coefficients in vector $\boldsymbol{c}$; S is the number of "hot zones"; $\boldsymbol{A} = \boldsymbol{X}_p \boldsymbol{F}_{M \times L_{cp}}$ is the measurement matrix.

In the second step, a smart measurement matrix can be obtained by setting finer resolution only in the nearest surrounding zones to the integer samples $[l-1/2, l+1/2]$ (If l is detected among the highest coherences in the first step). In the cases of $l = 0$ and $l = L_{cp} - 1/2$, $[0, l/2]$ and $[l-1/2, L_{cp}-1/2]$ are considered respectively. The indices in detected subset Ω is regarded as the "hot zones", while the indices beyond the Ω retains the original resolution (See Figure 5.2.1).

Correspondingly, the smart measurement matrix $\boldsymbol{A}_{sm}$, which is the partial matrix of finer

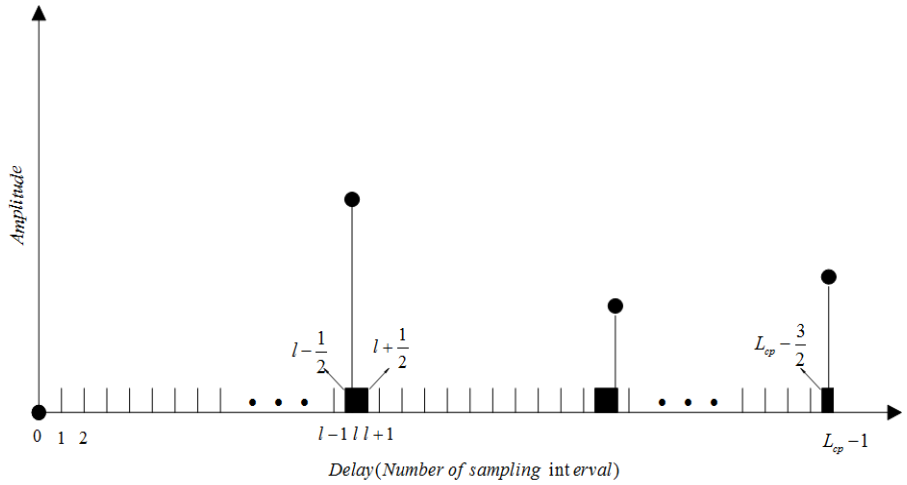

Figure 5.2.1: Hot zones and their corresponding non-sample spaced channel taps

resolution measurement matrix $\boldsymbol{A}_{fr}$, can be expressed by:

$$\boldsymbol{A}_{sm}=\begin{bmatrix} a_{k_0 0} & \cdots & a_{k_0(lR-\frac{R}{2})} & \cdots & a_{k_0 lR} & \cdots & a_{k_0(lR+\frac{R}{2})} & \cdots & a_{k_0((L_{cp}-1)R)} \\ a_{k_1 0} & \cdots & a_{k_1(lR-\frac{R}{2})} & \cdots & a_{k_1 lR} & \cdots & a_{k_1(lR-\frac{R}{2})} & \cdots & a_{k_1((L_{cp}-1)R)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k_{M-1}0} & \cdots & a_{k_{M-1}(lR-\frac{R}{2})} & \cdots & a_{k_{M-1}lR} & \cdots & a_{k_{M-1}(lR-\frac{R}{2})} & \cdots & a_{k_{M-1}((L_{cp}-1)R)} \end{bmatrix} \tag{5.2.11}$$

where $l = 0, 1, 2, \ldots, L_{cp} - 1$.

5.2.2 Simulations

In the simulations, an OFDM system with 1024 subcarriers is considered. In order to estimate a non-sample spaced multipath Rayleigh sparse channel having 6 taps with delays uniformly distributed over $[0, L_{cp}-1]$, 128 pilots (The pilot positions are the same with the previous CS based sample spaced method.) and cyclic prefix with $L_{cp} = 256$ are employed. Additionally,

the channel has power delay profile with an exponential distribution ($\phi(\tau') = e^{-\frac{\tau'}{\tau_{rms}}}$ and $\tau_{rms} = \frac{L_{cp}}{4}$). In the simulations, the channel is assumed to be static during the duration of one OFDM symbol.

Simulation comparisons mainly focus on the performance of BER and NMSE combined with the computational cost of different algorithms.

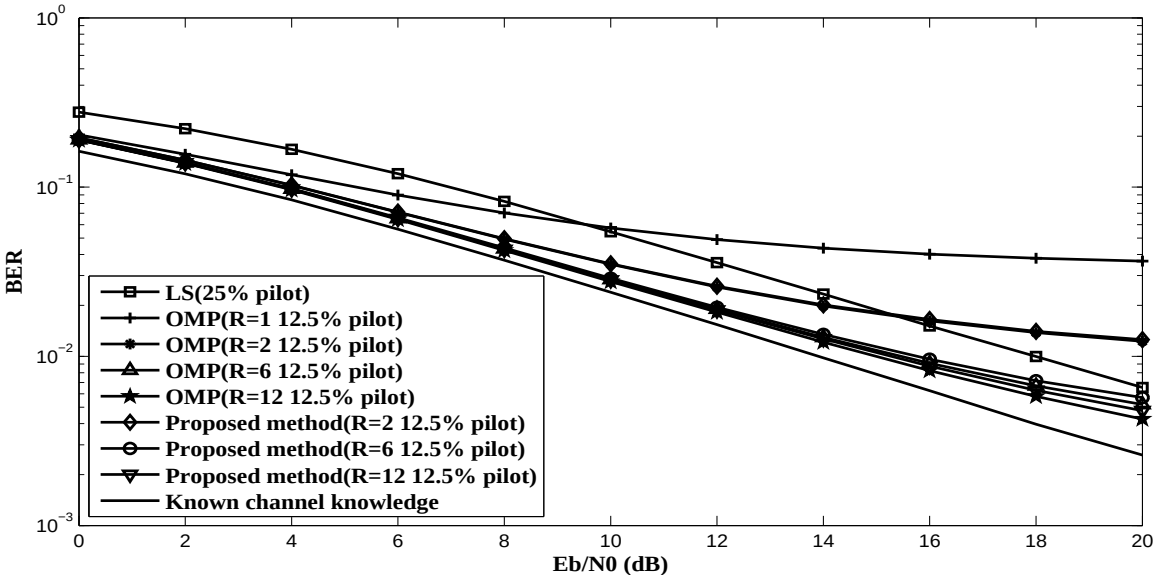

Figure 5.2.2: Performance of BER comparison for the non-sample spaced sparse channel

Fig.5.2.2 illustrates the performance of BER comparison for the non-sample spaced sparse channel with different algorithms. Generally, under different time resolutions ($R = 2, 6, 12$), the proposed method achieves the approaching BER performance with OMP method with traditional finer resolution measurement matrix. Additionally, with only 12.5% of pilots, the proposed method with resolution $R = 6, 12$ outperforms the LS method with 25% of pilots, since the leakage of channel powers of different taps is significantly reduced, meanwhile the noise taps are eliminately greatly. (For simplicity, we assume that the number of non-zero channel taps is known for both the proposed method and the OMP method with different resolutions.) OMP method with the original resolution ($R = 1$) has poorest performance,

because the leakage powers of the channel taps to the noise taps are ignored and they are combined with noise taps when considering the sparsity of 6. Known channel knowledge has the best performances, however, within the considered range of Eb/N0, its performance advantage is less than 3dB compared with the proposed method with $R = 12$.

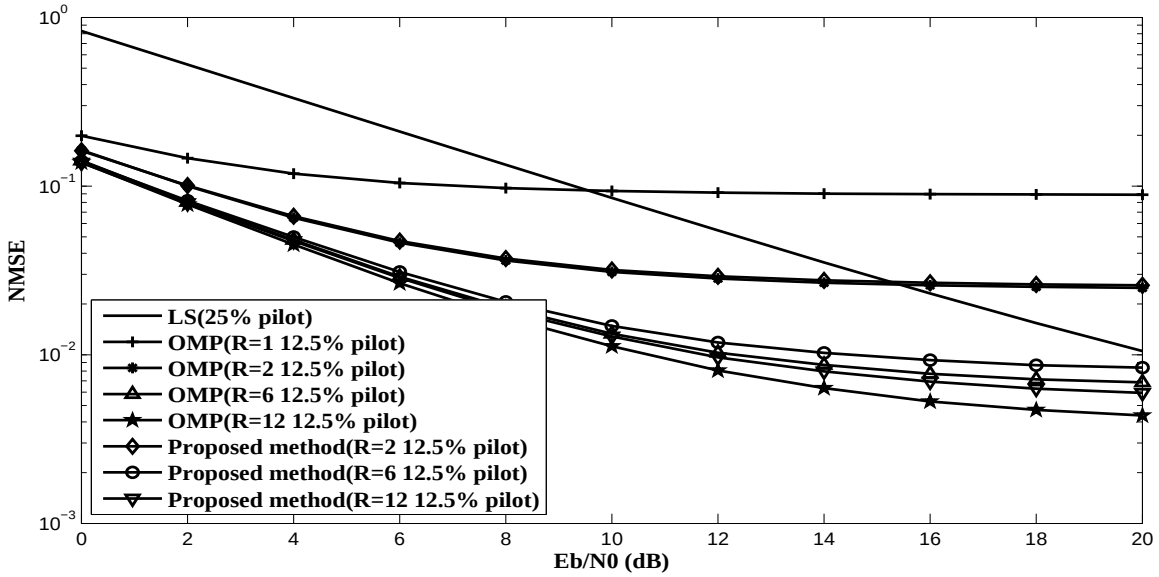

Figure 5.2.3: Performance of NMSE comparison for the non-sample spaced sparse channel

The NMSE performance of different algorithms has the same trend as that with the performance of BER, although the performance differences between algorithms are much more obvious.

Table 5.3: Computational complexity comparison

Algorithm	OMP $(R = 6, 12)$	proposed $(R = 6, 12)$
Complexity	$O(RM(L_{cp} - 1)K)$	$O(RMSK + ML_{cp}K)$

Table 5.3 gives the computational complexity of OMP method with different resolutions and the proposed method. It is well known that OMP method has the complexity of $O(RM(L_{cp}-1)K)$ (R is the resolution) [48]. Due to the presence of the smart measurement matrix, the computational complexity reduces to $O(RMSK + ML_{cp}K)$ (Maximum possible calculation)

in the cases of $RS \ll L_{cp}$.

In order to make the complexity more intuitive, computational complexity reduction percentage (CCRP) is defined to compare the proposed method and OMP method with different resolutions. The CCRP is defined as:

$$CCRP = (1 - \frac{computational\ complexity\ of\ the\ proposed\ scheme}{computational\ complexity\ of\ the\ OMP\ method\ with\ finer\ resolution}) \times 100\% \tag{5.2.12}$$

Consider the parameters ($R = 2, 6, 12$ and $S = 48$), we can have the following observations:

1) if $R = 6$ is considered for the proposed method and OMP, its CCRP is 64.4%.

2) if $R = 12$ is considered for the proposed method, and $R = 6, 12$ for OMP, their CCRP are 45.6% and 72.8% respectively.

With $R = 6$ or $R = 12$, the complexity of the proposed method is significantly reduced with respect to that of OMP. However, according to the above two observations, the estimation performance is close, especially in terms of BER. Actually, in practical communications, it can be neglected. Additionally, the proposed method with $R = 12$ has CCRP of 45.6% with respect to OMP method with $R = 6$, however, its performance advantage in both NMSE and BER is obvious.

5.3 Conclusion

In this chapter, both sample spaced and non-sample spaced sparse channel estimations based on CS are addressed. For the sample spaced sparse channel, an effective threshold is proposed for CS without the prior knowledge of channel statistics and the noise standard deviation to improve the channel estimation performance. For non-sample spaced sparse channel, the smart measurement matrix is introduced to effectively balance the spectral efficiency, computational complexity and channel estimation performance.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

Over the past two decades, the world has witnessed the arrival and experienced the conveniences of two generations of wireless communications: 3th and 4th generations (3G and 4G). As the rapid increasing standard of living, the demands to wireless transmission performances will increase accordingly, which conversely impose high requirements to the whole communication system. Additionally, the number of users of wireless communications is growing continually, which leads to the tension of spectral resources. Moreover, green communications are becoming a trend for future wireless communications, which require the high level of simplicity of communication system. As one of the key challenges in wireless communications, channel estimation has always been a research focus. In many communication environments, traditional channel estimation methods can hardly meet those challenges mentioned above. Thanks to the exploration of the characteristics of the sparsity of channel and the development of compressed sensing theory, overcoming all those challenges becomes possible. The research work mainly focuses on the sparse channel estimation and a series of novel channel estimation methods have been proposed. Specifically, we have made some contributions in the following points.

(1) Based on the exploration and analysis of the characteristics of wireless channels and their sparsity, the mathematical models and the frameworks of sparse channel estimation are constructed.

(2) An effective sparse channel estimation method based on LS is proposed in OFDM system in the case of $M \geq L_{cp}$. In this method, a novel effective time domain threshold depending only on the effective noise standard deviation estimated from the noise coefficients obtained by eliminating the channel coefficients with an initial estimated threshold is proposed. Both theoretical analysis and simulation results show that the proposed method can achieve better performance in both BER and NMSE than the compared methods within a wide range of sparsity rate, has good spectral efficiency and moderate computational complexity.

(3) An novel effective threshold is proposed for CS based sparse channel estimation in OFDM system in the case of $M < L_{cp}$. The proposed method can realize effective channel estimation with low consumption of spectral resources without the prior knowledge of channel statistics. Simulations show that the proposed method can be used to estimate the sparse channel with different sparsity in various kinds of wireless communication environments.

(4) We have also addressed the CS based non-sample spaced sparse channel estimation with effective channel estimation performance, high spectral efficiency and low computational complexity. By fully exploring the characteristics of the non-sample spaced sparse channel and the measurement matrix, smart measurement matrix is proposed to realize our goal. Simulations show that the proposed method can effectively solve the challenge of non-sample spaced sparse channel estimation.

6.2 Future Works

Due to the time constraints, the research work is not deep and comprehensive. Future works will mainly concentrate on the following aspects:

(1) Beyond the static or quasi-static channel, they are still many dynamic channels, such as, time selective channels and doubly selective channels, which are varying in one OFDM symbol. For time selective channels or doubly selective channels, the sparsity still exists in many cases. How to fully utilize the sparsity to realize effective channel estimation with low consumption of frequency band and low computational complexity is the main challenge.

(2) For fixed constellation pilots, pilot arrangement is a key issue. Some of the methods focusing on pilot arrangements are initially developed, however, there are still enough space to explore, especially, focusing on some specific channel patterns.

(3) Majority of current sparse channel estimation methods are realized by adopting the fix constellation pilots, which greatly depends on the pilots arrangement, however, pilot arrangement is a challenging and time consuming task, comparatively, random pilots are obvious more convenient and don't require any arrangements. Therefore, how to make full use of random pilot arrangement to effectively extract CSI is an obvious good direction.

(4) Up until now, CS reconstruction methods are gradually developed. However, for different channel models, different reconstruction algorithms can be explored to promote the system overall performance.

Appendix A

Derivation of the Relationship Between SNR and Eb/N0 in OFDM System

Assume that P_s is the power of signal; P_n is the noise power; E_b is the bit energy. Consider the definition of SNR, we have:

$$SNR = \frac{P_s}{P_n} = \frac{E_s/T_s}{P_n} = \frac{E_s}{N_0 W T_s} = \frac{E_s}{N_0} \ (W = \frac{1}{T_s}) \tag{A.0.1}$$

where E_s is the symbol engergy; N_0 is the noise power spectral density; W and T_s are the bandwidth and sampling interval respectively. If the case where no pilot is considered, the energy bit E_b is given by:

$$E_b = \frac{E_{OFDM}}{N n_b} \tag{A.0.2}$$

where n_b is the bit number per symbol; E_{OFDM} is the enegy per OFDM symbol; N is the size of FFT. Comparatively, if M $M < N$ in one OFDM symbol is considered, we have:

$$E_b^{'} = \frac{E_{OFDM}}{(N - M) n_b} \tag{A.0.3}$$

where $E_b^{'}$ is the energy bit considering M pilots. Therefore, we have the relationship between SNR and $\frac{E_b^{'}}{N_0}$ given by:

$$SNR = \frac{n_b E_b}{N_0} = \frac{n_b(N-M)}{N}\frac{E_b^{'}}{N_0} = \alpha_1 \frac{E_b^{'}}{N_0} \tag{A.0.4}$$

where $\alpha_1 = \frac{n_b(N-M)}{N}$ is a constant value. If both pilots and cyclic prefix are considered, we have:

$$E_b^{'} = \frac{N}{N+L_{cp}} E_b^{''} = \alpha_2 E_b^{''} \tag{A.0.5}$$

where L_{cp} is the length of cyclic prefix, $\alpha_2 = \frac{N}{N+L_{cp}}$ is a constant value. Finally, the relationship between SNR and $\frac{E_b^{''}}{N_0}$ is obtained [122]:

$$SNR = \alpha_1\alpha_2 \frac{E_b^{''}}{N_0} = \frac{n_b E_b}{N_0} = \frac{n_b(N-M)}{N}\frac{N}{N+L_{cp}}\frac{E_b^{''}}{N_0} \tag{A.0.6}$$

Appendix B

Derivation of FAR of the Universal Threshold

Assume complex vector $\boldsymbol{z} \sim \mathbb{CN}(0, \sigma_z^2 \boldsymbol{I}_L)$, with real part $\Re\{\boldsymbol{z}\} \sim \mathbb{N}(0, \frac{\sigma_z^2}{2} \boldsymbol{I}_L)$, and imaginary part $\Im\{\boldsymbol{z}\} \sim \mathbb{N}(0, \frac{\sigma_z^2}{2} \boldsymbol{I}_L)$. The amplitude of the i^{th} element z_i of vector $\boldsymbol{z}$ follows Rayleigh distribution. Consider formula (4.3.5) we have:

$$P\{|z_i| \geq t\frac{\sigma_z}{\sqrt{2}}\} = 1 - F(t\frac{\sigma_z}{\sqrt{2}}) = e^{-\frac{t^2}{2}} \tag{B.0.1}$$

if $t = \sqrt{2}\sqrt{2\ln L}$, (the threshold is $\sqrt{2\ln L}\sigma_z$), then we have:

$$P\{|z_i| \geq \sqrt{2\ln L}\sigma_z\} = \frac{1}{L^2} \tag{B.0.2}$$

$$P\{|z_i| \leq \sqrt{2\ln L}\sigma_z\} \to 1 \quad s.t. \quad L \to \infty \tag{B.0.3}$$

From (D.0.2) when $L \to \infty$, the probability of $|z_i| \geq \sqrt{2\ln L}\sigma_z$ approaches to zero at the speed of L^2.

Appendix C

Résumé étendu (French Extended Abstract)

C.1 Introduction

L'ère de l'information est l'une des principales caractéristiques de l'évolution rapide de notre monde moderne. Il ne fait aucun doute que la révolution de l'information accélère le rythme de vie de l'ensemble de la population ainsi que la recherche scientifique dans divers domaines. Chaque jour, nous échangeons une vaste quantité d'information, comprenant des sons, des images, du texte et bien d'autres choses. Parmi l'ensemble de ces données, une quantité considérable est transmise par différents types de systèmes de communication sans fil.

Les communications sans fil sont très souvent dégradées à cause des perturbations (interférences, bruit, ...) et de la propagation non maîtrisée des ondes (atténuation, déphasage, trajets multiples, ...). Si dans certains cas, un fort rapport signal à bruit est suffisant pour une bonne réception des données, dans la majorité des cas, la connaissance à priori du canal de propagation est nécessaire. Le récepteur cohérent utilise la connaissance des informations du canal pour réduire les effets indésirables amenés par le canal physique durant la transmission. La majorité des systèmes de communication utilisent un récepteur cohérent afin d'obtenir de meilleures performances lors des communications.

Pour utiliser ce genre de récepteur, la connaissance des informations concernant le canal est nécessaire. C'est pour cela que l'estimation du canal est essentielle à la réception. Les méthodes d'estimation de canal peuvent être classées en deux catégories : celles utilisant une séquence d'apprentissage [2–9] et les méthodes dites aveugles [10–12]. Parmi ces méthodes, celles utilisant une séquence d'apprentissage sont les plus courantes. Elles sont généralement efficaces dans l'estimation de canaux à trajets multiples. Cependant, dans le cas de canaux parcimonieux ("sparse channel"), qui sont présents dans de nombreux environnements de communication sans fil, les méthodes traditionnelles d'estimation de canal peuvent s'avérer peu efficaces [19]. Si les caractéristiques de la parcimonie des canaux physiques sont pleinement exploitées, cela peut être bénéfique pour l'efficacité spectrale, pour les performances de l'estimation de canal et pour la complexité de calcul. De plus, l'une des découvertes scientifiques récentes, qui est la théorie de l'acquisition comprimée ("compressed sensing" ou "compressive sensing"), peut fournir un moyen efficace pour extraire les informations des canaux parcimonieux en utilisant une bande de fréquence limitée avec des coûts de calcul réduits.

La majorité des systèmes de communication sans fil haut-débit actuels sont basés sur la modulation OFDM (Orthogonal Frequency Division Multiplexing). C'est le cas des standards de communication sans fil Wi-Fi (Wireless Fidelity), DVB-T (Digital Video Broadcasting - Terrestrial), LTE (Long Term Evolution), WiMAX (Worldwide Interoperability for Microwave Access) et bien d'autres encore.

Les travaux de cette thèse concernent donc l'estimation de canal pour les systèmes OFDM, et plus particulièrement l'estimation des canaux parcimonieux. Les parties suivantes de cette introduction s'intéressent aux canaux de propagation parcimonieux, à la théorie de CS (Compressed Sensing) et à l'estimation des canaux parcimonieux dans les systèmes OFDM.

C.1.1 Les canaux de propagation et leurs caractéristiques de parcimonie

Dans un environnement réel de communication sans fil, les ondes radio sont transmises à travers différents chemins, ce qui conduit à l'arrivée de signaux de puissance significative à des instants différents au récepteur. Dans le cas d'une propagation à travers un canal à trajets multiples, le fading (ou évanouissement) est un paramètre important. Il reflète l'atténuation de puissance due aux différents chemins d'arrivée, ce qui est la source principale de perturbation et donc d'augmentation du taux d'erreurs des données à la réception [26]. Pour combattre ce phénomène, le gain de diversité peut être exploité [19, 27]. Les canaux de propagation sans fil à trajets multiples peuvent être classés en deux catégories : les canaux à trajets multiples dits riches et les canaux à trajets multiples parcimonieux [19].

Dans certains environnements "indoor" et dans des espaces confinés, le nombre de trajets multiples est très important et l'écart temporel entre l'arrivée de deux trajets est comparable à l'intervalle d'échantillonnage. Dans ce cas, le canal présente les caractéristiques d'un canal à trajets multiples dit riche [28, 29]. Par contre, dans certains environnements "outdoor" et dans de vastes espaces, le nombre de trajets multiples peut être relativement faible et l'écart temporel entre l'arrivée de deux trajets peut être très grand comparé à l'intervalle d'échantillonnage. Dans ce cas, le canal présente les caractéristiques d'un canal à trajets multiples parcimonieux. Dans les standards de communication actuels, il existe plusieurs modèles de canaux parcimonieux. C'est le cas pour le DVB-T [30], les ondes acoustiques sous-marines [22, 31], et le standard LTE [32–34].

C.1.2 La théorie de l'acquisition comprimée ("Compressed Sensing Theory")

A l'ère du numérique, divers signaux analogiques sont convertis en signaux numériques, stockés, tranmis et traités. Dans ce contexte, le convertisseur analogique-numérique (CAN) est un élément essentiel. Les techniques d'échantillonnage traditionnelles sont généralement basées sur le théorème d'échantillonnage de Shannon-Nyquist [41, 42], qui exige que la fréquence d'échantillonnage soit au moins deux fois supérieure à la bande de fréquence du signal. Le théorème de Shannon-Nyquist nous donne une condition suffisante pour la récupération du signal, mais cette condition est-elle nécessaire? La théorie de l'acquisition comprimée ou CS (Compressed Sensing ou Compressive Sensing) [20, 21, 43] donne une réponse négative à cette question.

Durant l'année 2004, Donoho, Candes, Romberg et Tao furent les premiers à proposer le concept de CS et démontrèrent qu'un signal avec les caractéristiques de parcimonie peut être exactement récupéré à partir d'un faible nombre de mesures [20, 21]. Aujourd'hui,la théorie de CS a été améliorée et est utilisée dans de nombreuses applications (l'astronomie, la biologie, les communications sans fil, la reconnaissance de formes, le radar, le traitement vidéo, etc).

La conception de la matrice de mesure est la première et certainement la plus importante étape dans la théorie de CS. Cette étape a des répercussions profondes sur la précision de la reconstruction du signal parcimonieux. Il existe des possibilités très variées pour construire cette matrice [19, 22, 48–51]. Une fois la matrice construite, le signal parcimonieux peut être reconstruit à l'aide de différents algorithmes de reconstruction. Ces algorithmes sont basés sur des méthodes de minimistion de la norme l_1 [52–55], sur des méthodes de minimistion de la norme l_2 [49, 56] ou encore sur des méthodes alternatives [57, 58].

C.1.3 Estimation de canal parcimonieux dans les systèmes OFDM

Pour réaliser une estimation de canal efficace, il est essentiel d'analyser et d'explorer les caractéristiques de parcimonie de ces canaux. C'est une phase importante pour la conception de la matrice de mesure et pour la reconstruction du canal.

Conception de la matrice de mesure

La construction de la matrice de mesure est une étape importante dans l'estimation de canal parcimonieux basée sur la théorie de CS. Il y a plusieurs facteurs à prendre en compte. Le premier facteur est l'arrangement des pilotes OFDM qui peut être fixé [59, 60] ou aléatoire [61, 62]. Dans le cas d'un arrangement de pilotes fixe, des méthodes d'arrangement ont été proposées [50, 59, 60, 63] pour améliorer l'efficacité de l'estimation du canal. Le deuxième facteur est la qualité de résolution de la matrice [22]. C'est important dans le cas de canaux doublement sélectifs [31, 64–66]. Un autre facteur est l'utilisation d'une matrice de mesure dite intelligente qui permet de réduire considérablement la complexité de calcul lorsque l'étalement temporel du canal est très grand.

Estimation et reconstruction de canaux parcimonieux

L'estimation de canal peut être réalisée soit dans le domaine fréquentiel, soit dans le domaine temporel. Les méthodes LS (Least Squares) et MMSE (Minimum Mean Square Error) [67,68] sont deux méthodes d'estimation de canal majeures dans le domaine fréquentiel. La méthode MMSE fournit les meilleures performances mais nécessite, notamment, la connaissance à priori de la variance du bruit. La méthode LS à une complexité moindre et est associée à des algorithmes d'interpolation [15, 69, 70] pour réaliser une estimation de canal efficace. Cependant, cette méthode requiert un fort pourcentage de pilotes et est sensible au bruit, particulièrement dans le cas de canaux parcimonieux.

Les méthodes d'estimation dans le domaine temporel sont plus adaptées aux canaux parcimonieux [19, 39, 40, 67, 73, 74]. Contrairement aux méthodes fréquentielles, les méthodes temporelles sont sensibles à l'étalement temporel du canal. Elles ne sont utilisables que dans le cas où $M \geq L_{cp}$. (M est le nombre de pilotes et L_{cp} est la longueur du préfixe cyclique). Dans le cas d'un large étalement temporel (comparativement à la valeur de L_{cp}), les méthodes basées sur la théorie de CS doivent être considérées. Différents algorithmes de reconstruction fournissant des performances variées existent [22, 30, 32, 61, 64, 65, 75–82].

C.1.4 Organisation

Le deuxième chapitre, traite de l'estimation de canal dans les systèmes OFDM. Les différents effets du canal de propagation sont d'abord introduits puis, le principe d'un système OFDM est présenté avant de détailler et comparer les méthodes d'estimation de canal.

Le troisième chapitre présente la théorie de l'acquisition comprimée (CS pour "Compressed Sensing"). La construction des matrices de mesure et les méthodes de reconstruction y sont développées.

Dans le quatrième chapitre, les méthodes d'estimation de canal dans le domaine temporel sont évoquées dans le cas ($M \geq L_{cp}$). Pour séparer les trajets du bruit, la détermination d'un seuil est nécessaire. Après l'étude de seuils existants et des caractéristiques du bruit, une nouvelle méthode d'estimation de canal dans le domaine temporel est présentée. Elle repose sur l'utilisation d'un seuil efficace sans connaissance à priori des statistiques du canal et de la variance du bruit.

Dans le cinquième chapitre, le cas où $M < L_{cp}$ est considéré. L'estimation de canal parcimonieux basée sur la théorie de CS est présentée. Une nouvelle méthode d'estimation de canal utilisant un seuil efficace est proposée. Pour terminer ce chapitre, le cas où les trajets sont présents entre deux instants d'échantillonnage (non-sample spaced channel) est abordé.

C.2 Estimation de canal dans les systèmes OFDM

Dans cette partie, les effets du canal de propagation sont d'abord détaillés, puis, le modèle du systéme OFDM est présenté. Enfin, les méthodes d'estimation de canal courantes sont abordées.

C.2.1 Effets du canal de propagation

Les signaux radioélectriques sont susceptibles d'être diffractés, dispersés, réfléchis et atténués au cours de leur transmission, par conséquent, lorsque les signaux radio atteignent le récepteur, ils peuvent avoir différents chemins de transmission et des phases aléatoires, outre l'atténuation de leur puissance. Par la suite, les caractéristiques d'évanouissement (fading) des canaux sans fil sont présentées.

Atténuation en espace libre

L'atténuation en espace libre est provoquée par la dissipation de la puissance rayonnée par l'émetteur, ainsi que par l'effet de la propagation naturelle. C'est le modèle le plus simple qui considère une transmission en vue directe sans aucun obstacle. Pour une distance d entre l'émetteur et le récepteur, la relation entre la puissance émise et la puissance reçue est donnée par l'équation de Friis [83, 84], dont l'expression est:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2 L} \tag{C.2.1}$$

où $P_r(d)$ est la puissance reçue, P_t est la puissance émise, G_t est le gain de l'antenne d'émission, G_r celui de l'antenne de réception, λ est la longueur d'onde, L est le facteur de perte du système ($L \geq 1$).

Shadowing

Le shadowing, parfois appelé ombrage est un autre phénomène important. La principale raison de ce phénomène est que l'intensité des ondes radioélectriques subit une atténuation aléatoire causée par l'occultation du champ électromagnétique dû à la présence de différents objets, tels que les montagnes, les bâtiments, etc. Un modèle simplifié pour ce phénomène est le suivant :

$$p(r) = \frac{1}{\sqrt{2\pi}\sigma r} exp\left(-\frac{(\ln r - \mu)^2}{2\sigma^2}\right) \tag{C.2.2}$$

où r représente l'amplitude du signal reçu, considérée comme aléatoire. La distribution log-normale est déterminée par μ et σ, qui sont respectivement la valeur moyenne et l'écart type du signal aléatoire complexe reçu.

Canal à trajets multiples

Pendant la transmission, l'onde radioélectrique est soumise à la diffusion et aux réflexions. Le signal reçu est, dans ce cas, la combinaison de différents signaux provenant de différents trajets. La présence de chemins de transmission indépendants fait que le signal reçu est composé de différentes copies du signal original avec des temps d'arrivée différents et des atténuations de puissance différentes. La réponse impulsionnelle du canal est donc donnée par :

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - \tilde{\tau}_l) \tag{C.2.3}$$

où $\alpha_l \in C$ est l'amplitude du l^{eme} trajet du canal et $\tilde{\tau}_l$ est le retard du l^{eme} trajet du canal.

Si $\tilde{\tau}_l = \tau_l T_s$ avec T_s l'instant d'échantillonnage, l'équation s'écrit :

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - \tau_l T_s) \tag{C.2.4}$$

C.2.2 Le système OFDM

Modèle du système OFDM

L'OFDM est une des techniques de modulation multi-porteuses qui convertit un flux série de données haut-débit en N flux indépendants de données faible-débit modulés par N sous-porteuses orthogonales. En adoptant cette technique de transmission, le canal large bande sélectif en fréquence a été divisé en N sous-canaux bande étroite non sélectifs en fréquence. Ainsi, le système OFDM améliore l'efficacité de transmission en étant robuste au multi-trajets et aux interférences sur chaque sous-canal.

La figure E.2.1 présente un modèle bande de base pour le systéme OFDM. Tout d'abord, pour l'émission, le flux série de données haut-débit subit une modulation QAM. Puis, les symboles obtenus sont mis en parallèle avant l'insertion de pilotes qui serviront de séquence d'apprentissage pour l'estimation du canal. Ensuite, une IFFT (Inverse Fast Fourier Transform) est appliquée. Un préfixe cyclique (ou intervalle de guarde), qui est la copie de la fin d'un symbole OFDM, est ajouté devant le symbole OFDM correspondant pour lutter contre les interférences inter-symboles (IIS). Le symbole OFDM ainsi obtenu traverse un CNA (Convertisseur Numérique-Analogique) pour fournir le signal analogique qui sera transmis. Pour la réception, les opérations inverses sont effectuées.

Modèle mathématique

Le principe de base de la modulation multi-porteuses est d'utiliser des sous-porteuses orthogonales pour la modulation et la démodulation. Ce type de système de transmission peut être réalisé à partir de bancs de filtres (réseaux de filtres passe-bandes qui sépare le signal d'entrée en plusieurs composantes) équivalents à l'émetteur et au récepteur, comme le montre la figure E.2.2.

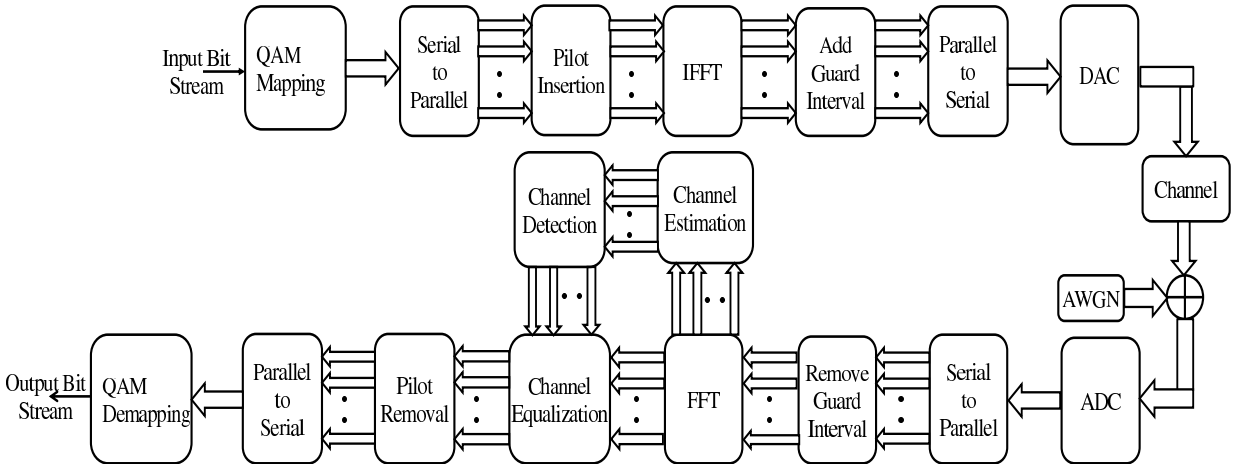

Figure C.2.1: Modèle bande de base pour le systéme OFDM

A la figure E.2.2, le signal d'entrée est donnée par l'équation suivante:

$$d_p(t) = \sum_n d_p(n)\delta(t - nT) \tag{C.2.5}$$

où $d_p(n)$ est le symbole de données avec l'indice de sous-porteuse p et n est l'indice du symbole OFDM, T est la durée d'un symbole OFDM.

Si ce signal est modulé par N sous-porteuses, le signal total émis s'exprime [87]:

$$s(t) = \sum_n \sum_{p=0}^{N-1} d_p(n) h_T(t - nT) e^{j2\pi f_p t} \tag{C.2.6}$$

où $e^{j2\pi f_p t}$ est la p^{eme} sous-porteuse de fréquence f_p. $h_T(t)$ est le filtre d'émission rectangulaire de largeur T.

Considérons que $f_p = f_b + p\Delta f$ et $\Delta f = 1/T$. Pour maintenir l'orthogonalité des sous-porteuses, il faut que $T\Delta f = 1$. De façon générale, pour différentes sous-porteuses et

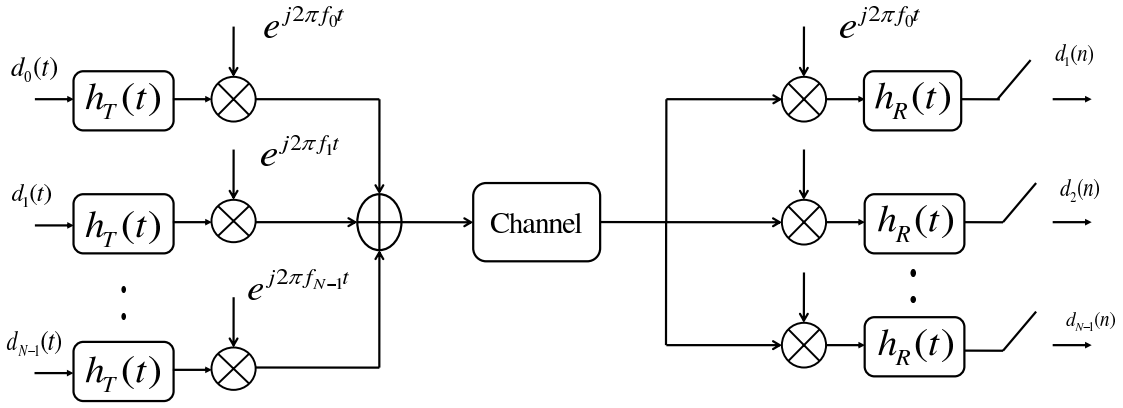

Figure C.2.2: Structure OFDM avec les bancs de filtres

différents symboles OFDM, l'orthogonalité est également maintenue.

$$\frac{1}{T}\langle h_R(t-nT)e^{j2\pi t f_p}, h_T(t-kT)e^{j2\pi t f_q}\rangle = \delta(n-k)\delta(p-q) \tag{C.2.7}$$

où $h_R(t)$ est le filtre de réception rectangulaire de largeur T.

Cette équation montre que la condition de biorthogonalité est satisfaite par les bancs de filtres à l'émetteur et au récepteur. Ainsi, en utilisant l'orthogonalité des sous-porteuses, les symboles sont correctement démodulés à la réception :

$$\tilde{d}_p(n) = \frac{1}{T}\langle d_p(n)h_T(t-nT)e^{j2\pi t f_p}, h_R(t-kT)e^{j2\pi t f_p}\rangle \tag{C.2.8}$$

Cette présention est vraie dans le cas idéal où la réponse impulsionnelle du canal est un dirac. Dans un environnement réel de communications sans fil, ce n'est jamais le cas. Nous en rediscuterons dans les sections suivantes.

C.2.3 Méthodes d'estimation de canal

Les méthodes LS (Least Squares) [67,90] et MMSE (Minimum Mean Square Error) [4,67,91] dans le domaine fréquentiel, LS [4, 36, 67, 91] dans le domaine temporel et ML (Maximum Likelihood) [4, 91] sont les méthodes d'estimation de canal les plus courantes. Elles sont présentées dans ce qui suit.

Considérons un système OFDM avec N sous-porteuses. M pilotes d'indices $k_0, k_1, \ldots, k_{M-1}$ sont utilisés pour estimer le canal. Le vecteur des pilotes reçus s'écrit [67, 90, 92]:

$$\boldsymbol{y}_p = \boldsymbol{X}_p \mathbf{g}_p + \boldsymbol{v}_p \tag{C.2.9}$$

où $\boldsymbol{X}_p = diag[x_{k_0}, x_{k_1}, \ldots, x_{k_{M-1}}]$ est la matrice diagonale des pilotes émis; $\boldsymbol{y}_p = [y_{k_0}, y_{k_1}, \ldots, y_{k_{M-1}}]^T$ est le vecteur des pilotes reçus; $\mathbf{g}_p = [g_{k_0}, g_{k_1}, \ldots, g_{k_{M-1}}]^T$ est le vecteur comprenant la position des pilotes et $\boldsymbol{v}_p = [v_{k_0}, v_{k_1}, \ldots, v_{k_{M-1}}]^T$ est le vecteur du bruit blanc additif Gaussien ($\boldsymbol{v}_p \sim CN(\mathbf{0}_M, \sigma^2 \boldsymbol{I}_M)$).

La méthode LS dans le domaine fréquentiel

L'objectif de l'estimateur LS est de minimiser $\|\boldsymbol{y}_p - \boldsymbol{X}_p \mathbf{g}_p\|_2^2$, ce qui s'écrit:

$$\hat{\mathbf{g}}_{p,LS} = argmin \|\boldsymbol{y}_p - \boldsymbol{X}_p \mathbf{g}_p\|_2^2 \tag{C.2.10}$$

La solution est classique et est donnée par [67, 90]:

$$\hat{\mathbf{g}}_{p,LS} = \boldsymbol{X}_p^{-1} \boldsymbol{y}_p = \mathbf{g}_p + \boldsymbol{X}_p^{-1} \boldsymbol{v}_p \tag{C.2.11}$$

Dans ce cas, seulement M échantillons de la réponse fréquentielle du canal sont estimés et les $N - M$ échantillons restants sont généralement obtenus par des méthodes d'interpolation.

La méthode MMSE dans le domaine fréquentiel

L'objectif de l'estimateur MMSE est de minimiser l'erreur quadratique moyenne des paramètres estimés, ce qui s'écrit:

$$minE[(\hat{\mathbf{g}}_p - \mathbf{g})^H(\hat{\mathbf{g}}_p - \mathbf{g})] \tag{C.2.12}$$

La solution est donnée par [91, 93]:

$$\hat{\boldsymbol{g}}_{p,MMSE} = \boldsymbol{R}_{\boldsymbol{g}_p\boldsymbol{y}_p}\boldsymbol{R}^{-1}_{\boldsymbol{y}_p\boldsymbol{y}_p}\boldsymbol{y}_p \tag{C.2.13}$$

où $\boldsymbol{R}_{\boldsymbol{g}_p\boldsymbol{y}_p} = \boldsymbol{R}_{\boldsymbol{g}_p\boldsymbol{g}_p}\boldsymbol{X}_p^H$ est la matrice de covariance entre $\boldsymbol{g}_p$ et $\boldsymbol{y}_p$; $\boldsymbol{R}_{\boldsymbol{y}_p\boldsymbol{y}_p} = \boldsymbol{X}_p\boldsymbol{R}_{\boldsymbol{g}_p\boldsymbol{g}_p}\boldsymbol{X}_p^H + \sigma^2\boldsymbol{I}_M$ est la matrice d'auto-covariance du vecteur $\boldsymbol{y}_p$. Plus communément, la solution s'écrit [4, 67, 91]:

$$\hat{\mathbf{g}}_{p,MMSE} = \boldsymbol{R}_{\mathbf{g}_p\mathbf{g}_p}(\boldsymbol{R}_{\mathbf{g}_p\mathbf{g}_p} + \sigma^2(\boldsymbol{X}_p\boldsymbol{X}_p^H)^{-1})^{-1}\hat{\mathbf{g}}_{p,LS} \tag{C.2.14}$$

La méthode MMSE est une extension de la méthode LS qui nécessite la connaissance à priori des statistiques du canal et celles du bruit.

La méthode LS dans le domaine temporel

Dans le domaine temporel, l'équation (E.2.9) peut être réécrite [19]:

$$\boldsymbol{y}_p = \boldsymbol{X}_p\boldsymbol{F}_{M\times L_{cp}}\boldsymbol{h} + \boldsymbol{v}_p \tag{C.2.15}$$

où $\boldsymbol{F}_{M\times L_{cp}}$ (L_{cp} est la longueur du préfixe cyclique) est la matrice de Fourier partielle, obtenu en sélectionnant les lignes de la matrice de Fourier avec les indices $k_0, k_1, \ldots, k_{M-1}$ et les L_{cp}

premières colonnes comme suit:

$$\boldsymbol{F}_{M\times L_{cp}}=\begin{bmatrix} W_N^{k_0 0} & W_N^{k_0 1} & \dots & W_N^{k_0(L_{cp}-1)} \\ W_N^{k_1 0} & W_N^{k_1 1} & \dots & W_N^{k_1(L_{cp}-1)} \\ \vdots & \vdots & \vdots & \vdots \\ W_N^{k_{M-1} 0} & W_N^{k_{M-1} 1} & \dots & W_N^{k_{M-1}(L_{cp}-1)} \end{bmatrix} \tag{C.2.16}$$

où $W_N^{k_p q} = e^{\frac{-j2\pi k_p l}{N}}$, $0 \leq p \leq M-1$, $0 \leq l \leq L_{cp}-1$.

Soit $\boldsymbol{X}_p\boldsymbol{F}_{M\times L_{cp}} = \boldsymbol{A}$, généralement connue comme matrice de mesures, l'équation (E.2.15) devient:

$$\boldsymbol{y}_p = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_p \tag{C.2.17}$$

L'objectif est de résoudre le problème de minimisation suivant [19, 91]:

$$\hat{\boldsymbol{h}}_{p,LS} = \arg\min_{\boldsymbol{h}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{h}\|_2^2 \tag{C.2.18}$$

La solution est donnée par [4, 36, 67, 91] (Appendice A):

$$\hat{\boldsymbol{h}}_{p,LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y}_p \tag{C.2.19}$$

Time Domain Maximum Likelihood Estimation

Dans le cas où $\boldsymbol{v}_p \sim CN(\boldsymbol{0}_M, \sigma^2\boldsymbol{I}_M)$, l'estimateur ML a la même expression que l'estimateur LS (Appendice B). La solution est donnée par [4, 91, 95]:

$$\hat{\boldsymbol{h}}_{p,ML} \equiv \hat{\boldsymbol{h}}_{p,LS} = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{y}_p \tag{C.2.20}$$

Simulations

Dans cette partie, les performances des méthodes d'estimation précédemment présentées sont évaluées (exceptée la méthode ML qui a les mêmes performances que celles de LS dans le domaine temporel). Deux cas différents sont considérés: le cas d'un canal à trajets multiples riche (le nombre de trajets est égal à la longueur du préfixe cyclique) et celui d'un canal à trajets multiples parcimonieux. Pour les simulations, le symbole OFDM est constitué pleinement de pilotes pour estimer le canal.

Pour évaluer les performances, l'erreur quadratique moyenne normalisée (NMSE) de la réponse impulsionnelle du canal est considérée. Elle est définie par [36]:

$$NMSE = \frac{E[\left\|\mathbf{h} - \hat{\mathbf{h}}\right\|_2^2]}{E[\|\mathbf{h}\|_2^2]} \tag{C.2.21}$$

où $\hat{\mathbf{h}}$ est l'estimation de $\mathbf{h}$.

Les simulations comparent d'abord les performances de l'erreur quadratique moyenne normalisée des estimateurs considérés dans le cas d'un canal à trajets multiples riche. La méthode LS dans le domaine temporel donne les performances optimales grâce à la matrice de mesures dont les L_{cp} colonnes considérées correspondent aux L_{cp} trajets du canal. La méthode MMSE aurait été optimale si le nombre de trajets du canal ainsi que leur position été connus. La méthode LS dans le domaine fréquentiel donne les plus mauvais résultats.

Dans le cas d'un canal à trajets multiples parcimonieux, en utilisant la connaissance à priori de la parcimonie du canal (nombre de trajets du canal non nuls), la méthode LS dans le domaine temporel donne les meilleures performances. La méthode LS dans le domaine temporel considère les L_{cp} premiers trajets dont certains sont du bruit dans ce cas. Les performances se rapprochent de celles obtenues par la méthode MMSE. La méthode LS dans le domaine fréquentiel donne toujours les plus mauvais résultats.

Table C.1: Comparaison de la complexité de calcul

Algorithme	LS domaine fréquentiel	LS domaine temporel	MMSE
Complexité	$O(N)$	$O(ML_{cp})$	$O(N^2)$

Le tableau E.1 compare la complexité de calcul des trois méthodes. La méthode LS dans le domaine fréquentiel a la complexité la plus faible mais donne les performances les plus mauvaises. La méthode MMSE qui a de bonnes performances présente la complexité la plus élevée et requiert la connaissance des statistiques du canal et du bruit. La méthode LS dans le domaine temporel présente le meilleur compromis avec les meilleures performances et une compexité relative.

C.2.4 Conclusion

Dans ce chapitre, après une présentation des caractéristiques du canal de propagation et du système OFDM, des méthodes d'estimation de canal ont été décrites. Des simulations ont permis d'évaluer leurs performances avant de comparer leur complexité de calcul.

C.3 Théorie de l'acquisition comprimée (Compressed Sensing Theory)

C.3.1 Les principes de base de l'acquisition comprimée

Les méthodes de traitement du signal traditionnelles sont basées sur le théorème de Shannon. Dans le domaine des mathématiques ou du traitement du signal, il existe des signaux parcimonieux. Si les signaux parcimonieux de grande dimension sont projetés sur des espaces de faible dimension, il est possible de reconstruire ces signaux à l'aide de vecteurs d'observation de faible dimension sans perte perceptible [20]. C'est le principe de l'acquisition comprimée (compressed sensing ou compressive sensing).

Le modèle mathématique

Un signal $\boldsymbol{z} \in C^N$ peut être exprimé par la combinaison linéaire de bases $\{\boldsymbol{\psi}_k = 0, \ldots, N-1\}$ avec une matrice $N \times N$ de rang plein $\boldsymbol{\Psi} = [\boldsymbol{\psi_0}, \boldsymbol{\psi_1}, \ldots, \boldsymbol{\psi_{N-1}}]$:

$$\boldsymbol{z} = \boldsymbol{\Psi x} \tag{C.3.1}$$

où $\boldsymbol{x}$ est le vecteur des pondérations.

Si le vecteur $\boldsymbol{x}$ est parcimonieux, la grande dimension de $\boldsymbol{z}$ peut causer une perte d'efficacité ou une augmentation de la complexité.

Supposons que $\boldsymbol{x}$ est un vecteur de parcimonie K (K éléments non nul) avec $\|\boldsymbol{x}\|_0 = |supp(\boldsymbol{x})| = |\Gamma| = K \leq K_{max} \ll N$; $\Gamma = \{i : x_i \neq 0\}$ est le support du vecteur $\boldsymbol{x}$; K_{max} est la valeur maximale de K. L'objectif est d'utiliser un vecteur d'observation $\boldsymbol{y} \in C^M$ de dimension $M (M \ll N)$ pour reconstruire le vecteur $\boldsymbol{x}$ sans perte perceptible.

Afin d'obtenir un vecteur d'observation $\boldsymbol{y}$ de plus faible dimension, une matrice d'observation $\boldsymbol{\Phi} \in C^{M \times N}$ est introduite. Nous avons donc :

$$\boldsymbol{y} = \boldsymbol{\Phi z} + \boldsymbol{w} = \boldsymbol{\Phi \Psi x} + \boldsymbol{w} \tag{C.3.2}$$

où $\boldsymbol{w} \in C^M$ est le vecteur complexe du bruit blanc additif gaussien tel que $CN(\boldsymbol{0}_M, \sigma^2 \boldsymbol{I}_M)$. $\boldsymbol{A} = \boldsymbol{\Phi \Psi}$ est la matrice de mesures. Un modèle mathématique classique est donné par :

$$\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{w} \tag{C.3.3}$$

L'objectif principal est de reconstruire le vecteur $\boldsymbol{x}$ de parcimonie K sans perte perceptible. Une propriété importante pour la reconstruction est la propriété RIP (Restricted Isometry Property) définie par [96–99]:

$$(1 - \delta_{K_{max}}) \|\boldsymbol{x}\|_2^2 \leq \|\boldsymbol{Ax}\|_2^2 \leq (1 + \delta_{K_{max}}) \|\boldsymbol{x}\|_2^2 \tag{C.3.4}$$

où $\delta_{K_{max}} \in (0, 1)$ est un paramètre qui garantit la qualité de l'approximation.

Outre la propriété RIP, la cohérence mutuelle de $\boldsymbol{A}$ est aussi un facteur important pour l'efficacité de la reconstruction. Elle s'écrit [56, 97, 100]:

$$\mu(\boldsymbol{A}) = \max_{0 \leq j,k \leq N-1, j \neq k} \frac{|\langle \boldsymbol{a}_j, \boldsymbol{a}_k \rangle|}{\|\boldsymbol{a}_j\|_2 \|\boldsymbol{a}_k\|_2} \tag{C.3.5}$$

où $\boldsymbol{A} = [\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{N-1}]$.

La reconstruction efficace du vecteur parcimonieux $\boldsymbol{x}$ à partir de $\boldsymbol{A}$ n'est garantie que pour une petite valeur de $\mu(\boldsymbol{A})$.

La solution parcimonieuse

La solution parcimonieuse optimale au système d'équations linéaires sous-déterminé est obtenue par la minimisation de la norme l_p (p=0, 1, 2). Pour le vecteur $\boldsymbol{x}$, ses normes l_1 et l_2 sont définies par $\|\boldsymbol{x}\|_1 = \sum_{i=0}^{N-1}|x_i|$ et $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=0}^{N-1}|x_i|^2}$ respectivement.

Pour reconstruire le vecteur parcimonieux $\boldsymbol{x}$, la minimisation de la norme l_0 est la solution la plus directe. Son expression est donnée par la formule suivante:

$$min\|\boldsymbol{x}\|_0 \textit{ sous contrainte de } \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon \tag{C.3.6}$$

La minimisation de la norme l_0 n'est pas mathématiquement efficace. C'est pourquoi la minimisation de la norme l_1 a été proposée. Elle est exprimée par :

$$min\|\boldsymbol{x}\|_1 \textit{ sous contrainte de } \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon \tag{C.3.7}$$

Outre la minimisation de la norme l_1, celle de la norme l_2 est aussi considérée. Elle est décrite par l'expression suivante :

$$min\|\boldsymbol{x}\|_2 \textit{ sous contrainte de } \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2 \leq \varepsilon \tag{C.3.8}$$

La minimisation des normes l_1 et l_2 sont toutes les deux des méthodes très utilisées. La minimisation de la norme l_1 donne la solution optimale pour la reconstruction du signal parcimonieux, mais a une complexité de calcul élevée [48,101]. La minimisation de la norme l_2 fournit une solution difficilement optimale mais a une complexité plus faible. Les mathématiciens préfèrent la norme l_1 alors que la norme l_2 est plus utilisée pour l'ingénierie.

C.3.2 Les algorithmes de reconstruction

Les algorithmes de reconstruction sont essentiellement basés sur la minimisation des normes l_1 et l_2. D'autres algorithmes alternatifs existent tel que l'algorithme IHT (Iterative Hard Thresholding) [57, 58].

L'algorithme GP (Greedy Pursuit)

L'algorithme GP (Greedy Pursuit) est l'une des méthodes majeures pour la reconstruction des signaux parcimonieux. Cette méthode est basée sur la minimisation de la norme l_2. Traditionnellement, la solution au problème de minimisation de la norme l_2 est obtenue en utilisant la méthode LS (Least Squares):

$$\hat{\boldsymbol{x}}_{LS} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \tag{C.3.9}$$

où $\boldsymbol{A} \in M \times N$ est la matrice de mesures.

Si $M \geq N$ et $Rank(\boldsymbol{A}) = N$ une solution optimale peut être obtenue. Par contre, l'efficacité n'est pas garantie dans le cas d'un vecteur parcimonieux ($\|\boldsymbol{x}\|_0 = |supp(x)| \ll N$). En effet, plus de mesures que nécessaires sont exploitées. Si nous considérons le cas où $M < N$, la méthode LS peut difficilement résoudre le problème. Actuellement, il existe des méthodes de reconstruction de signaux parcimonieux basées sur l'utilisation d'un dictionnaire [100]. Ce sont les méthodes GP (Greedy pursuit).

Les algorithmes de poursuite sont des algorithmes itératifs. A chaque itération, ils effectuent deux opérations : sélectionner le meilleur atome dans le dictionnaire (le plus souvent, c'est l'atome le plus corrélé au résidu précédent) puis, mettre à jour le résidu courant (information restante) en retirant une approximation n'utilisant que des atomes déjà sélectionnés. Les algorithmes se différencient par leurs règles de mise à jour et leurs critères d'arrêt.

L'algorithme OMP (Orthogonal Matching Pursuit) [49, 102] est l'un des plus simples et des plus efficaces parmi les méthodes GP. Cependant, dans certains cas, il pourrait être plus efficace [108].

Ainsi, de nombreux autres algorithmes de poursuite, souvent basés sur OMP, ont été développés dans le but de réduire la complexité et d'améliorer la précision dans la reconstruction du signal parcimonieux. Parmi ces algorithmes, il y a : stagewise OMP (StOMP) [108], regularized OMP (ROMP) [109], compressive sampling matching pursuit (CoSaMP) [110], subspace pursuit (SP) [111] and back-tracking based adaptive orthogonal matching pursuit (BAOMP) [48].

Les algorithmes StOMP et CoSaMP améliorent les performances de reconstruction du signal tout en réduisant la complexité de calcul [103, 108, 110]. Cependant, même si les algorithmes basés sur la minimisation de la norme l_1 sont plus complexes, ils fournissent les meilleures performances de reconstruction [22, 103].

Les algorithmes basés sur la minimisation de la norme l_1

La plupart des algorithmes basés sur la minimisation de la norme l_1 tentent de résoudre le problème de l'équation (E.3.7). La méthode DS (Dantzig selector), l'une des plus connues, a pour principe la résolution du problème suivant [54]:

$$min\|\hat{\boldsymbol{x}}\|_1 \textit{ sous contrainte de } \|\boldsymbol{A}^H(\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{y})\|_\infty \leq \lambda \tag{C.3.10}$$

où $\lambda = \sqrt{(2(a+1)lnN)}\sigma$ est la marge d'erreur tolérée pour la norme l_∞ du résidu.

La méthode BP (Basis Pursuit) [53] est une méthode souvent rencontrée qui résout le problème de minimisation de la norme l_1 dans le cas d'un environnement sans bruit. En présence de bruit, l'algorithme BP fait appel à une méthode de points intérieurs pour réaliser le processus de débruitage.

La méthode IHT (Iterative Hard Thresholding)

Le concept de seuil ("threshold") est largement utilisé dans la théorie de l'acquisition comprimée, et ce, pas uniquement pour le débruitage, mais aussi dans les processus des algorithmes de poursuite. Hormis pour les algorithmes StOMP et BAOMP précédemment cités, l'utilisation d'un seuil est également nécessaire pour la réalisation de l'algorithme IHT [57,58]. Son principe est simple mais efficace pour la reconstruction de signaux parcimonieux. Cette méthode consiste à réaliser les itérations suivantes ($\boldsymbol{x}_0 = \mathbf{0}$, $t = 0$):

$$\boldsymbol{x}_{t+1} = H_{K_{max}}(\boldsymbol{x}_t + \boldsymbol{A}^H(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_t)) \tag{C.3.11}$$

où $H_{K_{max}}(\cdot)$ est un opérateur non-linéaire qui fixe tous les éléments à zéro exceptés les K_{max} éléments ayant les amplitudes les plus grandes.

C.3.3 Conclusion

Dans ce chapitre, la théorie de l'acquisition comprimée (Compressed Sensing) a été introduite. Son modèle mathématique et les algorithmes de reconstruction du signal parcimonieux ont été présentés. Dans le cinquième chapitre, cette théorie est utilisée pour réaliser une estimation de canal parcimonieux efficace.

C.4 Estimation de canal parcimonieux basée sur la méthode LS

Que ce soit dans le domaine fréquentiel ou temporel, la méthode LS est une méthode d'estimation de canal très utilisée. Dans le domaine fréquentiel, elle est associée à des méthodes d'interpolation. Dans le domaine temporel, l'estimation de la réponse impulsionnelle d'un canal parcimonieux contient de nombreux échantillons qui ne sont que du bruit, en particulier pour le cas où $M \geq L_{cp}$ (le nombre de pilote est supérieur à la longueur du préfixe cyclique). L'utilisation d'un seuil efficace est nécessaire pour extraire du bruit les échantillons estimés du canal parcimonieux.

C.4.1 Les caractéristiques du bruit et analyse du seuil

Estimation de l'écart-type du bruit

La variance ou l'écart-type du bruit est un des paramètres numériques important pour décrire le phénomène aléatoire du bruit. Considérons que $Z(t)$ est un bruit aléatoire et que ses échantillons sont $z_i, 0 \leq i \leq N-1$. Il existe différentes méthodes pour estimer l'écart-type de $Z(t)$ telles que l'estimation biaisée, l'estimation Bayesienne [114] ou encore les méthodes basées sur la déviation absolue de la médiane [106, 115].

L'estimateur biaisé du processus aléatoire $Z(t)$ est donné par :

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(z_i - \bar{z})^2} \tag{C.4.1}$$

où $\bar{z} = \frac{1}{N}\sum_{i=0}^{N-1} z_i$ est la valeur moyenne des échantillons de bruit $z_i, 0 \leq i \leq N-1$.

Outre l'estimation biaisée et l'estimation Bayesienne [114], la déviation absolue de la médiane ou MAD (Median Absolute Deviation) est largement utilisée pour l'estimation de l'écart-type du bruit, particulièrement dans le cas de signaux parcimonieux en présence du bruit [106, 107, 115]. L'idée repose sur la relation étroite entre l'écart-type du bruit et la déviation absolue de sa médiane $MAD = \text{median}(|Z|) \approx \text{median}(|z_i, 0 \leq i \leq N-1|)$. Cette relation est donnée par :

$$\sigma = S \cdot MAD \tag{C.4.2}$$

où S est une constante qui dépend de la distribution spécifique des variables aléatoires. Par exemple, pour une variable aléatoire Gaussienne réelle, nous avons [106]:

$$\sigma = \frac{1}{0.6745} \cdot MAD \tag{C.4.3}$$

Pour une distribution de Rayleigh, qui est souvent utilisée dans le cas des communications sans fil, nous avons [115]:

$$\sigma = \sqrt{2}\sigma_1 = \sqrt{2}\frac{\text{median}(|Z|)}{\sqrt{\ln 4}} \tag{C.4.4}$$

où la constante vaut $S = \sqrt{\frac{1}{\ln 2}}$.

Si tous les échantillons ne sont que du bruit, les trois méthodes d'estimation citées sont efficaces. Par contre, dans le cas d'un signal parcimonieux, quelques composantes du signal sont présentes dans le bruit. L'estimation de l'écart-type obtenue par la méthode basée sur la déviation absolue de la médiane peut être considérée comme l'estimation de l'écart-type du bruit.

Analyse du seuil

Le seuil est un paramètre important pour une détection efficace, ce qui est essentiel pour isoler les trajets du canal du bruit dans une estimation de canal parcimonieux. Deux paramètres sont importants pour déterminer le seuil : la probabilité de détection ou POD (Probability Of Detection) et le taux de fausse alarme ou FAR (False Alarm Rate) qui consiste à considérer des échantillons du bruit comme étant des échantillons du signal.

Un seuil efficace doit équilibrer le POD et le FAR pour minimiser l'erreur quadratique moyenne du signal estimé. La détermination d'un seuil optimal requiert la connaissance à priori des statistiques du canal ainsi que la variance du bruit [40]. Par conséquence, il existe deux catégories de seuils. La première requiert la connaissance des statistiques du canal [37, 116], ce qui augmente la complexité du système. La seconde, qui est préférable, nécessite seulement une information sur la puissance ou la variance du bruit [73, 117].

Cette dernière catégorie de seuils a besoin de la variance du bruit pour la détection des puissances des trajets du canal ou de l'écart-type du bruit pour la détection des amplitudes des trajets. Dans notre cas, les amplitudes des trajets du canal sont considérées. Un exemple de seuil typique est le seuil proposé par Kang [73], qui s'écrit :

$$Th_1 = \sqrt{2}\sigma_z \tag{C.4.5}$$

où σ_z est l'écart-type du bruit.

Un autre seuil, appelé seuil universel [106], est couramment utilisé dans l'acquisition comprimée [113], [27], [19]:

$$Th_2 = \sqrt{2\ln L}\sigma_z \tag{C.4.6}$$

Ce seuil dépend de σ_z et de L, qui est le nombre d'éléments. En terme de FAR, le seuil universel présente de meilleures performances.

C.4.2 Proposition d'une méthode d'estimation de canal parcimonieux pour les systèmes OFDM

Nous considérons d'abord l'estimateur LS $\boldsymbol{h}_{LS}$ dans le domaine temporel qui est donné par (E.2.19):

$$\hat{\boldsymbol{h}}_{LS} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{y}_p \tag{C.4.7}$$

En combinant avec (E.2.17), la formule suivante est obtenue:

$$\hat{\boldsymbol{h}}_{LS} = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{A}\boldsymbol{h} + (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H \boldsymbol{v}_p \tag{C.4.8}$$

Considérons le cas où $L_{cp} \leq M < N$ (N est un entier multiple de M) et les pilotes sont uniformément distributés, dans ce cas $\boldsymbol{A}^H \boldsymbol{A} = M\boldsymbol{I}_{L_{cp}}$, l'équation précédente peut être réécrite comme suit :

$$\hat{\boldsymbol{h}}_{LS} = \boldsymbol{h} + \boldsymbol{n} \tag{C.4.9}$$

où $\boldsymbol{n} = \frac{1}{M}\boldsymbol{A}^H \boldsymbol{v}_p$ est une combinaison linéaire de bruits Gaussiens indépendants, donc c'est un vecteur AWGN, dont la matrice de covariance $\boldsymbol{C}$ est:

$$\boldsymbol{C} = E(\boldsymbol{n}\boldsymbol{n}^H) = \frac{1}{M}\sigma^2 \boldsymbol{I}_{L_{cp}} \tag{C.4.10}$$

L'estimation de la réponse impulsionnelle du canal ainsi obtenue par la méthode LS est fortement affectée par le bruit quand le canal est parcimonieux. Pour débruiter cette estimation, le seuil universel introduit précédemment est utilisé :

$$\lambda = \sqrt{2\ln L_{cp}}\sigma_n \tag{C.4.11}$$

Une estimation de l'écart-type σ_n du bruit $\boldsymbol{n}$ est donc nécessaire. Il est difficile d'obtenir une estimation efficace quand les échantillons du bruit $\boldsymbol{n}$ et les trajets du canal $\boldsymbol{h}$ sont présents ensemble dans le même signal. Pour un canal parcimonieux, la majorité des coefficients de l'estimation de la réponse impulsionnelle sont du bruit. Il est donc possible d'obtenir une estimation approximative de l'écart-type du bruit $\hat{\sigma}'_n$ en utilisant (E.4.4) et l'estimation du canal $\boldsymbol{h}_{LS}$:

$$\hat{\sigma}'_n = \sqrt{2}\hat{\sigma}' = \sqrt{2}\frac{\text{median}\left|\hat{\boldsymbol{h}}_{LS}\right|}{\sqrt{\ln 4}} \tag{C.4.12}$$

Cependant, la présence des trajets du canal dans le signal crée un biais dans l'estimation de l'écart-type du bruit ($\hat{\sigma}'_n > \sigma_n$). Si la majorité des trajets sont retirés de $\hat{\boldsymbol{h}}_{LS}$, les coefficients restants ne seront, approximativement, que du bruit qui pourra être utilisé pour obtenir une meilleure estimation de l'écart-type du bruit $\hat{\sigma}''_n$. C'est le principe de la méthode proposée.

La méthode proposée

La structure principale de la méthode proposée pour l'estimation de canal parcimonieux est décrite à la figure E.4.1. Les différentes étapes sont décrites ci-dessous.

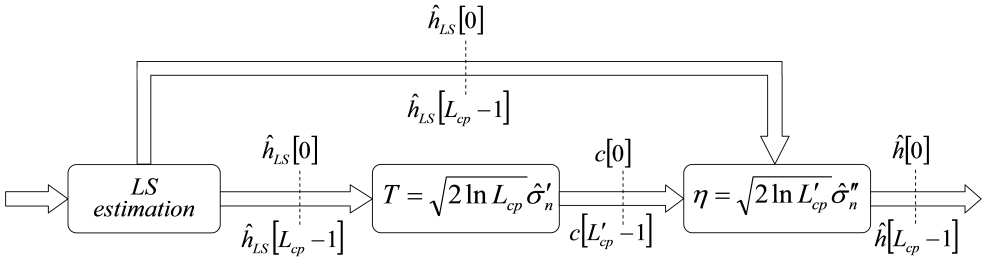

Figure C.4.1: Schéma de l'estimation de canal parcimonieux proposée

Etape.1. La méthode LS est utilisée pour obtenir une estimation initiale de la réponse impulsionnelle du canal de longueur L_{cp}.

Etape.2. Pour avoir une bonne estimation de l'écart-type du bruit, un seuil est nécessaire pour éliminer la majorité des trajets du canal dans l'estimation de la réponse impulsionnelle. Une première estimation approximative de l'écart-type du bruit est obtenu à partir des coefficients de l'estimation du canal parcimonieux $\hat{\boldsymbol{h}}_{LS}$ en utilisant (E.4.12). Puis, $T = \sqrt{2\ln L_{cp}}\hat{\sigma}'_n$ est utilisé comme seuil pour éliminer la majorité des trajets du canal présents dans l'estimation de la réponse impulsionnelle. En comparant avec T, le vecteur des coefficients du bruit (les coefficients $\hat{h}_{LS}[j]$ dont l'amplitude est inférieure ou égale à T) noté $\boldsymbol{c}$ ($\boldsymbol{c} = [c[0], c[1], \ldots, c[L_{cp}{}' - 1]]$, $L_{cp}{}' < L_{cp}$) est extrait.

Step.3. Avec le vecteur des coefficients du bruit $\boldsymbol{c}$ (contenant aucun ou beaucoup moins de trajets du canal que $\hat{\boldsymbol{h}}_{LS}$), $\hat{\sigma}''_n$ est estimé par $\hat{\sigma}''_n = \sqrt{2}\frac{median(|\boldsymbol{c}|)}{\sqrt{\ln 4}}$. Puis, un seuil efficace $\eta = \sqrt{2\ln L_{cp}{}'}\hat{\sigma}''_n$ est obtenu. L'estimation finale de la réponse impulsionnelle du canal est donnée par :

$$\hat{h}[n] = \begin{cases} \hat{h}_{LS}[n], & \left|\hat{h}_{LS}[n]\right| > \eta \\ 0, & \left|\hat{h}_{LS}[n]\right| \leq \eta \end{cases} \quad 0 \leq n \leq L_{cp} - 1 \tag{C.4.13}$$

Analyse et comparaison des performances

Des simulations sont proposées pour évaluer les performances d'estimation de la méthode proposée et la comparer avec des méthodes existantes. Pour les simulations, nous considérons un système OFDM simplifié avec 1024 sous-porteuses dont 256 sous-porteuses sont des pilotes. Les données sont modulées en QPSK et la longueur du préfixe cyclique est $L_{cp} = 256$.

Le modèle de canal considéré a une réponse impulsionnelle définie par [35]:

$$\begin{aligned} h[n] &= \delta[n] + 0.3162\delta[n-2] + 0.1995\delta[n-17] + 0.1296\delta[n-36] \\ &+ 0.1\delta[n-75] + 0.1\delta[n-137]. \end{aligned} \tag{C.4.14}$$

Les simulations présentent les performances, en terme de NMSE (Normalized Mean Squared

Error), de la méthode proposée, la méthode classique (LS dans le domaine temporel avec interpolation linéaire), l'estimateur Oracle (LS avec la connaissance idéale du canal), la méthode LS avec MST proposée par Minn et al [35] (MST signifie Most Significant Taps)) et la méthode LS avec le seuil proposé par Kang et al [73].

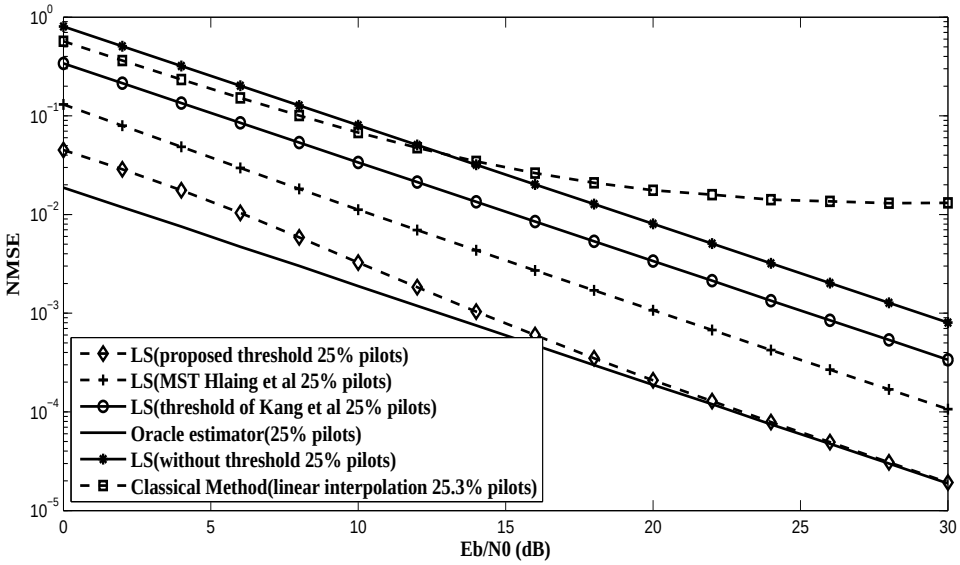

Figure C.4.2: Comparaison des performances des méthodes d'estimation de canal

La figure E.4.2, compare les performances des différents algorithmes. La méthode proposée donne de bonnes performances et tend vers l'estimateur Oracle lorsque l'impact du bruit diminue. Les performances de la méthode proposée maintiennent un avantage d'au moins 4dB sur l'ensemble des rapports signal à bruit considérés comparées aux performances de la méthode LS avec MST proposée par Minn et al et celles de la méthode LS avec le seuil proposé par Kang et al. En conclusion, même sans connaissance à priori des statistiques du canal et de l'écart-type du bruit, la méthode proposée fournit de bonnes performances.

C.4.3 Conclusion

Dans ce chapitre, les caractéristiques du bruit et l'analyse du seuil ont été présentées. Une méthode d'estimation utilisant un seuil efficace dans le cas où $M \geq L_{cp}$ a été proposée. Elle présente une grande efficacité spectrale, des performances d'estimation de canal efficaces, une faible complexité de calcul le tout sans connaissance à priori des statistiques du canal et de l'écart-type du bruit. La méthode proposée est efficace pour des canaux parcimonieux avec un faible étalement temporel. Dans le cas de canaux avec un large étalement temporel, le cas $M < L_{cp}$ doit étre considéré.

C.5 Estimation de canal parcimonieux basée sur l'acquisition comprimée

Contrairement à l'étude précédente, nous considérons dans ce chapitre le cas $M < L_{cp}$. Le mauvais conditionnement de la matrice de mesures dans le cas où $M < L_{cp}$ rend les méthodes d'estimation classiques non utilisables. La théorie de l'acquisition comprimée ou CS (Compressed Sensing) peut résoudre ce problème.

C.5.1 Estimation de canal parcimonieux lorsque $M < L_{cp}$

Le modèle mathématique de l'équation (E.2.17) est considéré :

$$\boldsymbol{y}_p = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_p \tag{C.5.1}$$

Si $M < L_{cp}$, la matrice de mesures $\boldsymbol{A}$ est mal conditionnée. Si $\boldsymbol{h}$ est un canal parcimonieux, il est possible de reconstruire $\boldsymbol{h}$ sous certaines conditions.

Nouveau seuil robuste pour une méthode d'estimation de canal parcimonieux basée sur l'acquisition comprimée

La cohérence de la matrice de mesures, définie par l'équation (E.3.5) est un paramètre important pour la performance de reconstruction du signal. En général, plus ce paramètre est petit, meilleure est la précision. En prenant en compte cela, l'arrangement des pilotes peut être réalisé par :

$$\arg\min_{\Gamma} \max_{0\leq i,j\leq L_{cp}-1, i\neq j} \frac{|\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle|}{\|\boldsymbol{a}_i\|_2 \|\boldsymbol{a}_j\|_2} \tag{C.5.2}$$

où Γ est le sous-ensemble d'arrangement des pilotes optimal.

Dans le cas où $M < L_{cp}$, $Rank(A) < L_{cp}$. Pour obtenir un seuil efficace, l'écart-type du bruit doit être estimé. Comme $Rank(A) = M$, reconstruire la réponse impulsionnelle du canal de longueur L_{cp} est impossible. Une réponse impulsionnelle partielle avec $m \quad (m \leq M)$ peut être extraite pour estimer l'écart-type du bruit. Il y a deux possibilités :

(1) Extraire les m $(m < M)$ premiers coefficients [119];

(2) Extraire les m coefficients en considérant :

$$\arg \min_{i,j \in \Lambda, 0 \leq i,j \leq L_{cp}-1, i \neq j} \frac{|\langle \boldsymbol{a}_i, \boldsymbol{a}_j \rangle|}{\|\boldsymbol{a}_i\|_2 \|\boldsymbol{a}_j\|_2}, |\Lambda| = m \quad (m \leq M) \tag{C.5.3}$$

où Λ est le sous-ensemble des m bases sélectionnées de façon optimale.

Pour obtenir Λ, il y a $C_{L_{cp}}^{m}$ choix possibles. Un algorithme efficace a été proposé pour obtenir les éléments de Λ avec une contrainte de cohérence.

L'équation (E.5.1) peut être ainsi réécrite :

$$\boldsymbol{y}_p = \boldsymbol{A}_m \boldsymbol{h}_m + \boldsymbol{A}_b \boldsymbol{h}_b + \boldsymbol{v}_p \tag{C.5.4}$$

où $\boldsymbol{A}_m = [\boldsymbol{a}[p_0], \boldsymbol{a}[p_1], \ldots, \boldsymbol{a}[p_{m-1}]$ et $\boldsymbol{h}_m = [h[p_0], h[p_1], \ldots, h[p_{m-1}]]^T$ sont la matrice avec les m colonnes de $\boldsymbol{A}$ sélectionnées et le vecteur avec les m trajets du canal $\boldsymbol{h}$ sélectionnés; $\boldsymbol{A}_b$ et $\boldsymbol{h}_b = [h[p_m], h[p_{m+1}], \ldots, h[p_{L_{cp}-1}]]^T$ sont la matrice avec les autres colonnes de $\boldsymbol{A}$ et le vecteur avec les autres trajets de $\boldsymbol{h}$.

En utilisant la méthode LS pour estimer les m trajets du canal, nous obtenons :

$$\hat{\boldsymbol{h}}_{ls-m} = (\boldsymbol{A}_m^H \boldsymbol{A}_{\boldsymbol{m}})^{-1} \boldsymbol{A}_m^H \boldsymbol{y}_p \tag{C.5.5}$$

En combinant les deux équations précédentes, nous avons :

$$\hat{\boldsymbol{h}}_{ls-m} = (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H (\boldsymbol{A}_m \boldsymbol{h}_m + \boldsymbol{A}_b \boldsymbol{h}_b + \boldsymbol{v}_p) \tag{C.5.6}$$

Si $\boldsymbol{A}_m^H \boldsymbol{A}_m$ est inversible, l'équation s'écrit :

$$\hat{\boldsymbol{h}}_{ls-m} = \boldsymbol{h}_m + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b \boldsymbol{h}_b + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{C.5.7}$$

Pour réduire l'impact des trajets significatifs présents dans $\boldsymbol{h}_b$, ils sont estimés par la méthode OMP. Alors, une nouvelle équation est donnée par :

$$\hat{\boldsymbol{h}}_m = \boldsymbol{h}_m + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b (\boldsymbol{h}_b - \hat{\boldsymbol{h}}_{omp-b}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{C.5.8}$$

Les trajets significatifs du vecteur $\boldsymbol{h}_m$ sont estimés par la méthode OMP pour donner le vecteur $\hat{\boldsymbol{h}}_{omp-m}$. Ainsi, l'erreur $\hat{\boldsymbol{e}}_m$ peut être estimé par :

$$\hat{\boldsymbol{e}}_m = (\boldsymbol{h}_m - \hat{\boldsymbol{h}}_{omp-m}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{A}_b (\boldsymbol{h}_b - \hat{\boldsymbol{h}}_{omp-b}) + (\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p \tag{C.5.9}$$

Estimation de l'écart-type du bruit

Si l'estimation est précise, l'erreur $\hat{\boldsymbol{e}}_m$ ne contient que du bruit, qui suit une distribution Gaussienne et dont l'amplitude suit une loi de Rayleigh. L'écart-type de la partie réelle $\hat{\sigma}_r$ et de la partie imaginaire $\hat{\sigma}_i$ du bruit est estimé par :

$$\hat{\sigma}_r = \hat{\sigma}_i = \hat{\sigma}_1 = \frac{median(|\hat{\boldsymbol{e}}_m|)}{\sqrt{ln4}} \tag{C.5.10}$$

Donc, l'écart-type du bruit $(\boldsymbol{A}_m^H \boldsymbol{A}_m)^{-1} \boldsymbol{A}_m^H \boldsymbol{v}_p$ est estimé par $\hat{\sigma} = \sqrt{2}\hat{\sigma}_1$. Le seuil universel,

qui peut être utilisé pour la détection du canal parcimonieux est donné par [19, 106]:

$$T = R\hat{\sigma};\ (R = \sqrt{2ln(m)}) \tag{C.5.11}$$

Structure de la méthode proposée

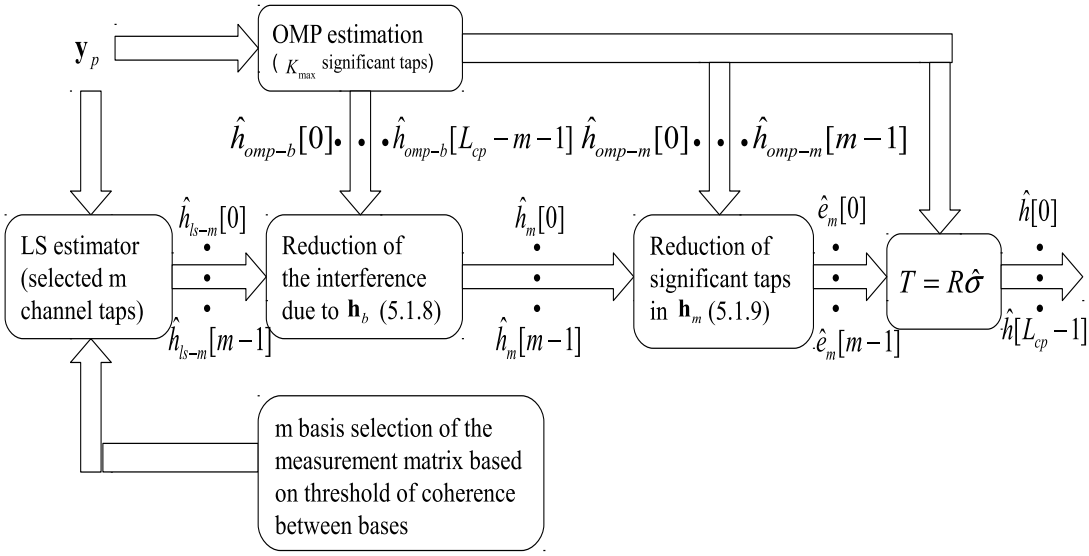

Figure C.5.1: Schéma de l'estimation de canal parcimonieux proposée

La figure E.5.1 présente la structure de la méthode proposée. Les étapes spécifiques sont décrites comme suit.

Etape.1. Les m bases sont sélectionnées à partir de la matrice de mesures.

Etape.2. L'algorithme OMP fournit une estimation du canal avec K_{max} trajets significatifs alors que l'estimateur LS donne une estimation de la réponse impulsionnelle partielle du canal (les m trajets du canal sélectionnés) avec le vecteur des pilotes reçus $\boldsymbol{y}_p$ (M pilotes).

Etape.3. Pour améliorer l'estimation, l'interference causée par les trajets de $\boldsymbol{h}_b$ est réduite par l'utilisation de $\hat{\boldsymbol{h}}_{omp-b}$.

Etape.4. Pour réduire l'impact des trajets significatifs de $\boldsymbol{h}_m$, $\hat{\boldsymbol{h}}_{omp-m}$ est introduit. Une estimation de l'erreur $\hat{\boldsymbol{e}}_m$ est obtenue.

Etape.5. Obtenir l'estimation de l'écart-type du bruit $\hat{\sigma}$ à partir des valeurs de $\hat{\boldsymbol{e}}_m$.

Etape.6. Avec $\hat{\sigma}$, le seuil $T = R\hat{\sigma}$ est utilisé pour la détection des trajets les plus significatifs :

$$\hat{h}[n] = \begin{cases} \hat{h}_{omp}[n], & |\hat{h}_{omp}[n]| \geq T \\ 0, & |\hat{h}_{omp}[n]| < T \end{cases}, \ 0 \leq n \leq L_{cp} - 1 \tag{C.5.12}$$

Simulations et comparaisons

Pour les simulations, un système OFDM avec une modulation QPSK est considéré. La bande passante est de 10MHz [120] et il y a 1024 sous-porteuses dont 128 sont des pilotes. Les durées du symbole OFDM et du préfixe cyclique sont $128\mu s$ et $25.6\mu s$. L'arrangement des pilotes est donné par [59]. De plus $Th = 0.14$ et $K_{max} = 24$ sont considérés pour la méthode proposée.

Le modèle de canal suivant est considéré [35]:

$$h[n] = \delta[n] + 0.3162\delta[n-2] + 0.1995\delta[n-17] + 0.1296\delta[n-36] + 0.1\delta[n-75] + 0.1\delta[n-137] \tag{C.5.13}$$

Les performances sont comparées pour différentes méthodes d'estimation telles que la méthode proposée, la méthode avec seuil proposée par Kang et al [73], la méthode proposée dans [119] et l'estimateur Oracle.

La figure E.5.2 compare les performances en termes de BER. La méthode proposée donne les meilleures performances avec seulement 12.5% de pilotes. De plus, elle ne requiert pas la connaissance à priori des statistiques du canal et de l'écart-type du bruit.

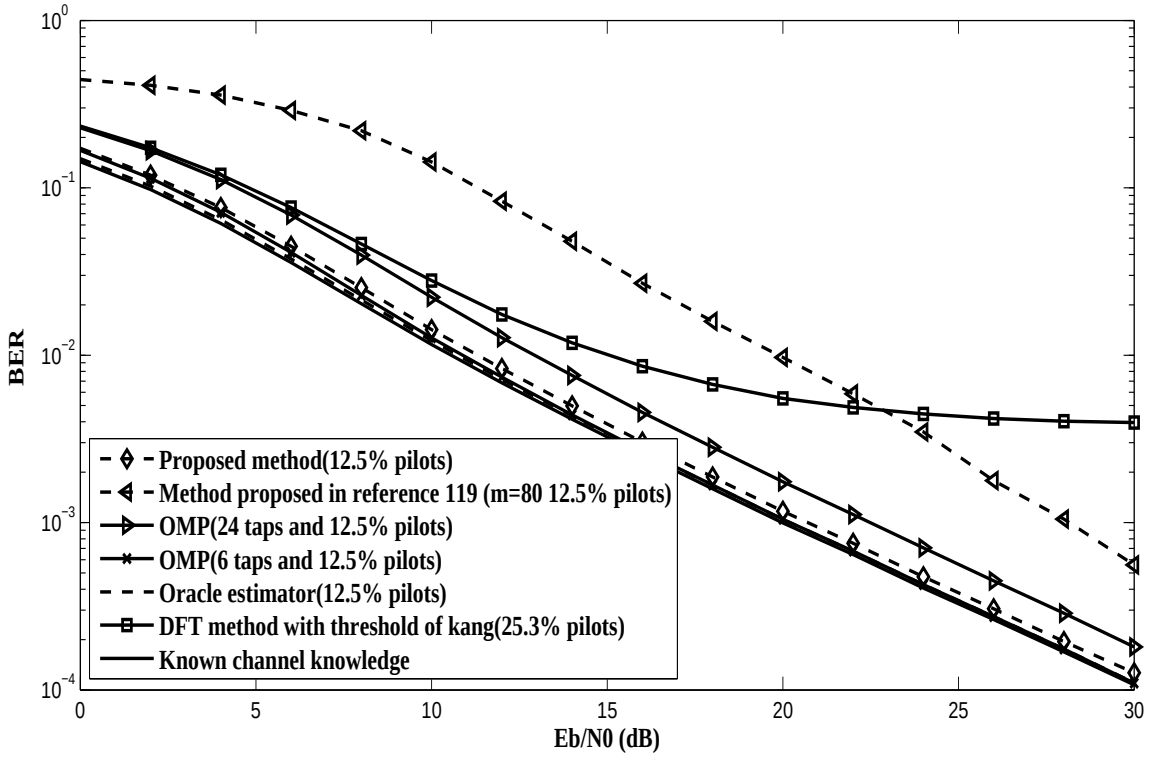

Figure C.5.2: Performance en termes de taux d'erreur binaire (BER)

C.5.2 Estimation de canaux parcimonieux dont les trajets sont situés en dehors des instants d'échantillonnage

Si les trajets du canal se situent en dehors des instants d'échantillonnage, il y a un phénomène de perte d'information. La période d'échantillonnage T_s est un paramètre constant. Une résolution plus fine $T_s' = T_s/R$ (R est le facteur de suréchantillonnage) peut être considérée. L'équation (E.2.4) se réécrit [19]:

$$h\left[\tau\right] = \sum_{l} \alpha_l \delta(\tau - R\tau_l {T_s}') \tag{C.5.14}$$

La réponse fréquentielle du canal est donnée par :

$$\begin{aligned} g[f] &= \int_{-\infty}^{\infty} h[\tau] e^{-j2\pi f\tau} \mathrm{d}\tau \\ &= \sum_{l} \alpha_l e^{-j2\pi f R \tau_l T_s{}'} \end{aligned} \tag{C.5.15}$$

Posons $f_k = \frac{k}{RNT_s'},\ k = 0, 1, 2, \ldots, RN-1$, prenons la FFT de $h[lT_s'],\ l = 0, 1, 2, \ldots, RN-1$, nous obtenons la réponse fréquentielle discrète du canal :

$$g[k] = \sum_{l} \alpha_l e^{-j2\pi R \tau_l k/RN} \tag{C.5.16}$$

Conception d'une matrice de mesures intelligente

Une matrice de mesures avec une résolution plus fine $\boldsymbol{A}_{fr}$ de dimensions $M \times (R(L_{cp}-1)+1)$ peut être obtenue :

$$\boldsymbol{A}_{fr} = \begin{bmatrix} a_{k_0 0} & a_{k_0 1} & \ldots & a_{k_0 (L_{cp}-1)R} \\ a_{k_1 0} & a_{k_1 1} & \ldots & a_{k_1 (L_{cp}-1)R} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k_{M-1} 0} & a_{k_{M-1} 1} & \ldots & a_{k_{M-1} (L_{cp}-1)R} \end{bmatrix} \tag{C.5.17}$$

où $a_{k_m n} = x_{k_m} e^{\frac{-j2\pi k_m n}{RN}},\quad 0 \le m \le M-1,\ 0 \le n \le (L_{cp}-1)R.$

Si cette matrice est utilisée pour la reconstruction, cela causera une augmentation du coût de calcul considérable. Comme le canal est parcimonieux, [121] propose un simple algoritme d'estimation de canal basé sur l'acquisition comprimée pour les canaux parcimonieux variant dans le temps.

Une matrice de mesures intelligente $\boldsymbol{A}_{sm}$, qui est la matrice partielle de la matrice de mesures

$\boldsymbol{A}_{fr}$, est alors obtenue :

$$\boldsymbol{A}_{sm}=\begin{bmatrix} a_{k_0 0} & \cdots & a_{k_0(lR-\frac{R}{2})} & \cdots & a_{k_0 lR} & \cdots & a_{k_0(lR+\frac{R}{2})} & \cdots & a_{k_0((L_{cp}-1)R)} \\ a_{k_1 0} & \cdots & a_{k_1(lR-\frac{R}{2})} & \cdots & a_{k_1 lR} & \cdots & a_{k_1(lR-\frac{R}{2})} & \cdots & a_{k_1((L_{cp}-1)R)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k_{M-1} 0} & \cdots & a_{k_{M-1}(lR-\frac{R}{2})} & \cdots & a_{k_{M-1} lR} & \cdots & a_{k_{M-1}(lR-\frac{R}{2})} & \cdots & a_{k_{M-1}((L_{cp}-1)R)} \end{bmatrix} \tag{C.5.18}$$

où $l = 0, 1, 2, \ldots, L_{cp} - 1$.

Simulations

Pour ces simulations, un système OFDM avec 1024 sous-porteuses est consideré. 128 pilotes et un préfixe cyclique de longueur $L_{cp} = 256$ sont utilisés. Le canal est considéré comme statique sur toute la durée d'un symbole OFDM.

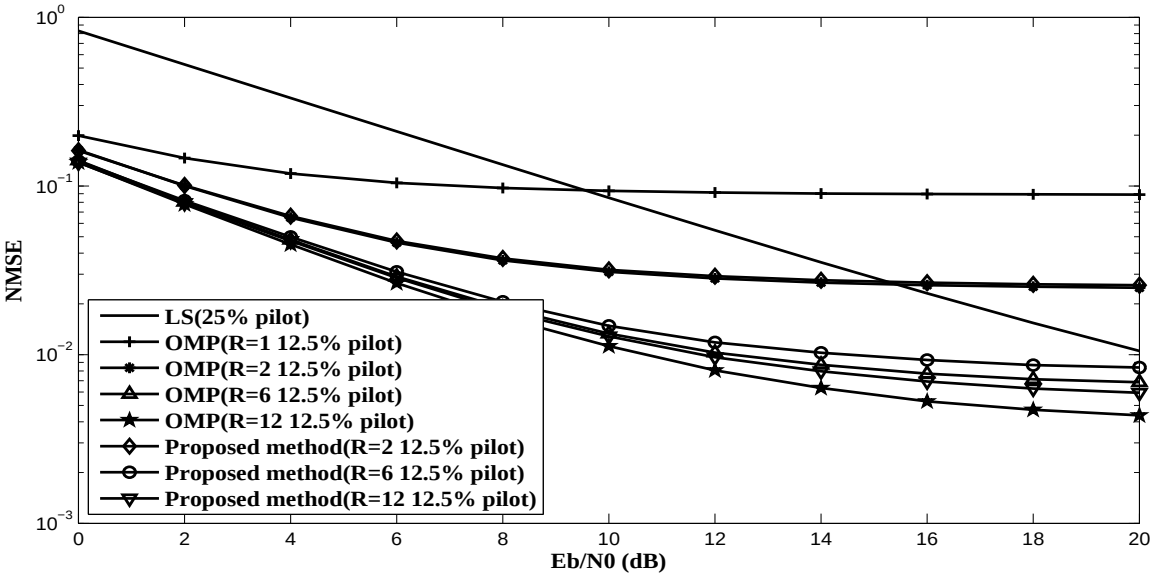

Figure C.5.3: Performances en termes de NMSE

La figure E.5.3 présente les performances en termes de NMSE des différents algorithmes

d'estimation pour le cas de canaux parcimonieux dont les trajets sont situés en dehors des instants d'échantillonnage. Pour les différentes résolutions temporelles, la méthode proposée approche les performances de l'OMP utilisant la matrice complète de résolution fine. De plus, avec seulement 12.5% de pilotes, les performances dépassent celles de la méthode LS avec 25% de pilotes.

C.5.3 Conclusion

Dans ce chapitre, une méthode d'estimation de canal parcimonieux basée sur l'acquisition comprimée est proposée. Cette méthode, utilisant un seuil efficace, ne requiert aucune connaissance à priori des statistiques du canal et de l'écart-type du bruit. Dans le cas des canaux dont les trajets sont situés en dehors des instants d'échantillonnage, l'utilisation d'une matrice de mesure intelligente permet d'améliorer l'efficacité des méthodes d'estimation.

C.6 Conclusion et perspectives

Dans cette thèse, après une large introduction bibliographique, les caractéristiques des canaux de propagation ont été présentées suivies du système OFDM et des méthodes classiques d'estimation de canal. Dans le cas où le nombre de pilotes est plus grand que la longueur du préfixe cyclique, une méthode d'estimation de canal parcimonieux basée sur l'utilisation d'un seuil temporel original est proposée. Cette méthode avec une haute efficacité spectrale, de bonnes performances d'estimation, une faible complexité de calcul, requiert aucune connaissance à priori des statistiques du canal et de l'écart-type du bruit. Si le canal est parcimonieux avec un grand étalement temporel (le nombre de pilotes est plus faible que la longueur du préfixe cyclique), une technique d'estimation de canal basée sur la théorie de l'acquisition comprimée est proposée. Cette méthode requiert un plus petit nombre de pilotes que les techniques classiques. Ce travail se termine avec l'étude des canaux dont les trajets sont situés en dehors des instants d'échantillonnage. L'utilisation d'une matrice de mesures intelligente permet d'améliorer l'efficacité des méthodes d'estimation.

Parmi les perspectives possibles, la première est l'étude de l'optimisation de l'arrangement des pilotes. A partir des statistiques du canal, il est possible de déterminer un arrangement de pilotes qui optimise les performances des estimations de canal. Une autre perspective possible est la recherche d'un compromis (nombre de pilotes/qualité d'estimation) qui permettrait d'obtenir le nombre de pilotes minimal pour conserver une bonne efficacité d'estimation. L'efficacité spectrale pourrait ainsi être optimisée.

Bibliography

[1] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels [J]," *IEEE Trans. Commun*, vol. 46, pp. 902–915, 1998.

[2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels [J]," *IEEE Trans. Veh. Technol*, vol. 40, pp. 686–693, 1991.

[3] A. J. Coulson, "Maximum likelihood synchronization for OFDM using a pilot symbol: algorithms [J]," *IEEE J. Select. Areas Commun*, vol. 19, pp. 2486–2494, 2001.

[4] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems [J]," *IEEE Trans. Signal Process*, vol. 49, pp. 3065–3073, 2001.

[5] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels [J]," *IEEE Trans. Signal Process*, vol. 5, pp. 1615–1624, 2003.

[6] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels [J]," *IEEE Trans. Signal Process*, vol. 51, pp. 1351–1366, 2003.

[7] B. M. S. L. Tong and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing [J]," *IEEE Signal Process. Mag*, vol. 21, pp. 12–25, 2004.

[8] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems [J]," *IEEE Trans. Wireless Commun*, vol. 4, pp. 765–774, 2005.

[9] Z. J. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems [J]," *IEEE Trans. Signal Process*, vol. 55, pp. 2226–2238, 2007.

[10] G. X. L. Tong and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach [J]," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, 1994.

[11] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods [J]," *Proc. IEEE*, vol. 86, pp. 1951–1968, 1998.

[12] S. Roy and C. Y. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix [J]," *IEEE Trans. Wireless Commun*, vol. 1, pp. 572–579, 2002.

[13] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas [J]," *IEEE Trans. Wireless Commun*, vol. 1, pp. 67–75, 2002.

[14] P. Stoica and O. Besson, "Training design for frequency offset and frequency-selective channel estimation [J]," *IEEE Trans. Commun*, vol. 51, pp. 1910–1917, 2003.

[15] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems [J]," *IEEE Trans. Broadcast*, vol. 48, pp. 223–229, 2002.

[16] P. Hammarberg, F. Rusek, and O. Edfors, "Iterative receivers with channel estimation for multi-user MIMO-OFDM: complexity and performance [J]," *EURASIP Journal on Wireless Commun and Network*, vol. 75, pp. 1–17, 2012.

[17] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: The ecological and economic perspective [J]," *IEEE Commun. Mag*, vol. 49, pp. 55–62, 2011.

[18] C. Despins, F. Labeau, T. L. Ngoc, R. Labelle, M. Cheriet, F. G. C. Thibeault, A. Leon-Garcia, O. Cherkaoui, B. S. Arnaud, J. Mcneill, Y. Lemieux, and M. Lemay, "Leveraging green communications for carbon emission reductions: Techniques, testbeds, and emerging carbon footprint standards [J]," *IEEE Commun. Mag*, vol. 49, pp. 101–109, 2011.

[19] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels [J]," *Proc of IEEE*, vol. 98, pp. 1058–1076, 2010.

[20] D. Donoho, "Compressed sensing [J]," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, 2006.

[21] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information [J]," *IEEE Trans. Inform. Theory*, vol. 52, pp. 489–509, 2006.

[22] C. R. Berger, Z. H. Wang, J. Z. Huang, and S. L. Zhou, "Application of compressive sensing to sparse channel estimation [J]," *IEEE Commun Mag*, vol. 48, pp. 164–174, 2010.

[23] J. L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency-devision multiplexing [J]," *IEEE Trans. Commun*, vol. 33, pp. 665–675, 1985.

[24] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come [J]," *IEEE Commun Mag*, vol. 28, pp. 5–14, 1990.

[25] Z. Wang and G. G. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon [J]," *IEEE Signal Process. Mag*, vol. 17, pp. 29–48, 2000.

[26] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels [J]," *IEEE Trans. Inf. Theory*, pp. 1073–1096, 2003.

[27] W. U. Bajwa, A. M. Sayeed, and R. Nowak, "Sparse multipath channels: Modeling and estimation [A]," *in Proc. 13th IEEE Digital Signal Process Workshop [C]*, pp. 320–325, 2009.

[28] A. A. M. Saleh, A. J. Rustako, and R. S. Roman, "Distributed antennas for indoor radio communications [J]," *IEEE Trans. Commun*, vol. 35, pp. 1245–1251, 1987.

[29] H. Nishimoto, Y. Ogawa, T. Nishimura, and T. Ohgane, "Measurement-based performance evaluation of MIMO spatial multiplexing in a multipath-rich indoor environment [J]," *IEEE Trans. Antennas Propagat*, vol. 55, pp. 3677–3689, 2007.

[30] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization [J]," *IEEE Trans. Commun*, vol. 50, pp. 374–377, 2002.

[31] C. R. Berger, S. Zhou, J. C. Preisig, and P.Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing [J]," *IEEE Trans. Signal Process*, vol. 58, pp. 1708–1721, 2010.

[32] P. Maechler, P. Greisen, B. Sporrer, S. Steiner, N. Felber, and A. Burg, "Implementation of greedy algorithms for LTE sparse channel estimation [A]," *in Proc. 44th Asilomar Conf. Signals, Syst and Comp [C]*, pp. 400–405, 2010.

[33] . Recommendation ITU-R M.1225, International Telecommunication Union, *Guidelines for evaluation of radio transmission technologies for IMT-2000 [M]*. 1997.

[34] V. E. et al, "Channel models for fixed wireless applications [S]," *IEEE 802.16.3c-01/29r4*, pp. 1–36, 2001.

[35] H. Minn and V. K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation [J]," *IEEE Trans. Broadcast*, vol. 46, pp. 240–248, 2000.

[36] M. R. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels [J]," *IEEE Signal Precess. Lett*, vol. 12, pp. 52–55, 2005.

[37] J. Oliver, R. Aravind, and K. M. M. Prabhu, "Sparse channel estimation in OFDM system by threshold-based pruning [J]," *Electron. Lett*, vol. 44, pp. 830–832, 2008.

[38] W. Z. F. Wan and M. Swamy, "Semi-blind most significant tap detection for sparse channel estimation of OFDM systems [J]," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, pp. 703–713, 2010.

[39] M. R. Raghavendra, E. Lior, S. Bhashyam, and K. Giridhar, "Parametric channel estimation for pseudo-random tile-allocation in uplink OFDMA [J]," *IEEE Trans. Signal Process*, vol. 55, pp. 5370–5381, 2007.

[40] S. Rosati, G. E. Corazza, and A. Venelli-Coralli, "OFDM channel estimation based on impulse response decimation: analysis and novel algorithms [J]," *IEEE Trans. Commun*, vol. 60, pp. 1996–2008, 2012.

[41] C. E. Shannon, "Communications in the presence of noise [J]," *Proc. IRE*, vol. 37, pp. 10–21, 1949.

[42] H. Nyquist, "Certain topics in telegraph transmission theory [J]," *AIEE Trans*, vol. 47, pp. 617–644, 1928.

[43] R. G. Baraniuk, "Compressive sensing [J]," *IEEE Signal Process. Mag*, vol. 24, pp. 118–120, 124, 2007.

[44] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform [J]," *Proc. IEEE*, vol. 66, pp. 51–83, 1978.

[45] N. Ahmed, T. Natarajan, and K. R. Rao, "On image processing and a discrete cosine transform [J]," *IEEE Trans. Comput*, vol. C-23, pp. 90–93, 1974.

[46] S. Mallat, *A wavelet tour of signal processing [M].* Academic Press, 2009.

[47] E. J. Candès, Y. C. Elad, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries [J]," *Appl and Comput. Harmo Anal*, vol. 31, pp. 59–73, 2011.

[48] H. Huang and A. Makur, "Backtracking-based matching pursuit method for sparse signal reconstruction [J]," *IEEE Signal Process. Lett*, vol. 18, pp. 391–394, 2011.

[49] J. A. Tropp, "Signal recovery from random measurements via orthogonal matching pursuit [J]," *IEEE Trans. Inf. Theory*, vol. 53, pp. 4655–4666, 2007.

[50] C. Qi and L. Wu, "Tree-based backward pilot generation for sparse channel estimation [J]," *Electron. Lett*, vol. 48, pp. 501–503, 2012.

[51] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling [J]," *IEEE Signal Process. Mag*, vol. 25, pp. 21–30, 2008.

[52] E. J. Candès and T. Tao, "Decoding by linear programming [J]," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, 2005.

[53] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit [J]," *SIAM J. Scientific Comput*, vol. 20, pp. 33–61, 1999.

[54] E. J. Candès and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n [J]," *Ann. Stat.*, vol. 35, pp. 2313–2351, 2007.

[55] R. Tibshirani, "Regression shrinkage and selection via the lasso [J]," *J. R. Stat. Soc. B*, vol. 58, pp. 267–288, 1996.

[56] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries [J]," *IEEE Trans. Inform. Theory*, vol. 54, pp. 2210–2219, 2008.

[57] T. Blumensath and M. Davies, "Iterative hard thresholding for sparse approximations [J]," *J. Fourier Anal. Appl*, vol. 14, pp. 629–654, 2008.

[58] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing [J]," *Appl. Comput. Harmonic Anal*, vol. 27, pp. 265–274, 2009.

[59] X. He and R. Song, "Pilot pattern optimization for compressed sensing based sparse channel estimation in OFDM systems [A]," *in Proc. Conf. Wireless Commun and Signal Process [C]*, 2010.

[60] C. Qi and L. Wu, "Optimized pilot placement for sparse channel estimation in OFDM systems [J]," *IEEE Signal Process. Lett*, vol. 18, pp. 749–752, 2011.

[61] G. Tauböck and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots [A]," *in Proc. IEEE Int. Conf. Acoust, Speech and Signal Process (ICASSP) [C]*, pp. 2885–2888, 2008.

[62] J. Meng, Y. Li, N. Nguyen, W. Yin, and Z. Han, "Compressive sensing based high resolution channel estimation for OFDM system [J]," *IEEE J. Sel. Topics Signal Process*, vol. 6, pp. 15–25, 2012.

[63] J. C. Chen, C. K. Wen, and P. Ting, "An efficient pilot design scheme for sparse channel estimation in OFDM systems [J]," *IEEE Commun. Lett*, vol. 17, pp. 1352–1355, 2013.

[64] G. Tauböck and F. Hlawatsch, "Compressed sensing based estimation of doubly selective channels using a sparsity-optimized basis expansion [A]," *in Proc. Eur. Signal Process. Conf (EUSIPCO) [C]*, 2008.

[65] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing [J]," *IEEE J. Sel. Topics Signal Process*, vol. 4, pp. 255–271, 2010.

[66] F. H. D. Eiwen, G. Taubock and H. G. Feichtinger, "Compressive tracking of doubly selective channels in multicarrier systems based on sequential delay-doppler sparsity [A]," *in Proc. IEEE Int. Conf. Acoust, Speech and Signal Process (ICASSP) [C]*, pp. 2928–2931, 2011.

[67] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems [A]," *in Proc. IEEE Vehicular Technology Conf.(VTC) [C]*, vol. 2, pp. 815–819, 1995.

[68] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system [J]," *IEEE Trans. Consum Electron*, vol. 44, pp. 1122–1128, 1998.

[69] M. Hsieh and C. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels [J]," *IEEE Trans. Consumer Electron*, vol. 44, pp. 217–225, 1998.

[70] S. G. Kang, Y. M. Ha, and E. K. Joo, "A comparative investigation on channel estimation algorithms for OFDM in mobile communications [J]," *IEEE Trans. Broadcast*, vol. 49, pp. 142–149, 2003.

[71] X. Dong, W. S. Lu, and A. Soong, "Linear interpolation in pilot symbol assisted channel estimation for OFDM [J]," *IEEE Trans. Wireless Commun*, vol. 6, pp. 1910–1920, 2007.

[72] Y. S. Lee, H. C. Shin, and H. N. Kim, "Channel estimation based on time-domain threshold for OFDM systems [J]," *IEEE Trans. Broadcast*, vol. 55, pp. 656–662, 2009.

[73] Y. Kang, K. Kim, and H. Park, "Efficient DFT-based channel estimation for OFDM system on multipath channels [J]," *IET Commun*, vol. 1, pp. 197–202, 2007.

[74] J. K. Hwang and R. L. Chung, "Low-complexity algorithm for tap-selective maximum likelihood estimation over sparse multipath channels [A]," *in Proc. IEEE Global Telecommunications Conference (GLOBECOM) [C]*, pp. 2857–2862, 2007.

[75] G. Gui, Q. Wan, S. Qin, and A. Huang, "Sparse multipath channel estimation using ds algorithm in wideband communication systems [A]," *in Proc. Cong. Image and Signal Process (CISP) [C]*, pp. 4450–4453, 2010.

[76] S. F. Cotter and B. D. Rao, "Matching pursuit based decision-feedback equalizers [A]," *in Proc. IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP) [C]*, pp. 2713–2716, 2000.

[77] S. F. Cotter and B. D. Rao, "The adaptive matching pursuit algorithm for estimation and equalization of sparse time-varying channels [A]," *in Proc. Asilomar Conf. Signals Syst. Comput., Pacific Grove [C]*, 2000.

[78] W. Li and J. C. Preisig, "Estimation of rapidly time-varying sparse channels [J]," *IEEE J. Ocean. Eng*, vol. 32, pp. 927–939, 2007.

[79] C.-J. Wu and D. W. Lin, "A group matching pursuit algorithm for sparse channel estimation for OFDM transmission [A]," *in Proc. IEEE Int. Conf. Acoust, Speech and Signal Process (ICASSP) [C]*, pp. 429–432, 2006.

[80] T. Kang and R. A. Iltis, "Matching pursuit channel estimation for an underwater acoustic OFDM modem [A]," *in Proc. IEEE Int. Conf. Acoust, speech and signal process (ICASSP) [C]*, pp. 5296–5299, 2008.

[81] M. A. Khojastepour, K. Gomadam, and X. D. Wang, "Pilot-assisted channel estimation for MIMO OFDM systems using theory of sparse signal recovery [A]," *in Proc. IEEE Intern. Conf. Acoust, speech and signal process (ICASSP) [C]*, pp. 2693–2696, 2009.

[82] S. Kim, "Angle-domain frequency-selective sparse channel estimation for underwater MIMO-OFDM systems [J]," *IEEE Commun. Lett*, vol. 16, pp. 685–687, 2012.

[83] W. C. Jakes, *Microwave communications [M].* John Wiley, 1947.

[84] A. Goldsmith, *Wireless communications [M].* Cambridge University Press, 2005.

[85] R. Prasad, *OFDM for wireless communications systems [M].* Artech House Press, 2004.

[86] J. G. Proakis, *Digital communications 5th edition [M].* McGraw-Hill Book Co., 2007.

[87] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier [J]," *IEEE Signal Process. Mag*, vol. 28, pp. 92–112, 2011.

[88] B. G. Negash and H. Nikookar, "Wavelet-based multicarrier transmission over multipath wireless channels [J]," *Electron. Lett*, vol. 36, pp. 1778–1788, 2000.

[89] K. M. Wong, J. F. Wu, T. N. Davidson, Q. Jin, and P. C. Ching, "Performance of wavelet packet-division multiplexing in impulsive and gaussian noise [J]," *IEEE Trans. Commun*, vol. 48, pp. 1083–1086, 2000.

[90] J.-J. v. d. B. S. K. W. O. Edfors, M. Sandell and P. O. Brjesson, "OFDM channel estimation by singular value decomposition [J]," *IEEE Trans. Commun*, vol. 46, pp. 931–939, 1998.

[91] S. M. Kay, *Fundamentals of statistical signal processing: Estimation Theory [M]*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[92] G. Auer and E. Karipidis, "Pilot aided channel estimation for ofdm: A separated approach for smoothing and interpolation [A]," *in Proc. IEEE Int. Conf. Commun (ICC) [C]*, vol. 4, pp. 2173–2178, 2005.

[93] L. L. Scharf, *Statistical signal processing: detection, estimation, and time series analysis [M]*. Addison-Wesley, 1991.

[94] P. Stoica and A. Nehorai, "Music, maximum likelihood, and Cramér-Rao bound [J]," *IEEE Trans. Acoust, Speech and Signal Process*, vol. 37, pp. 720–741, 1989.

[95] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, and M. Engels, "A low-complexity ML channel estimator for OFDM [J]," *IEEE Trans. Commun*, vol. 51, pp. 135–140, 2003.

[96] E. J. Candès and T. Tao, "Stable signal recovery from incomplete and inaccurate information [J]," *Commun. Pure Appl. Math*, vol. 59, pp. 1207–1233, 2005.

[97] E. J. Candès, "The restricted isometry property and its implications for compressed sensing [J]," *C. R. Acad. Sci. Paris, Ser. I*, vol. 346, pp. 589–592, 2008.

[98] R. G. Baraniuk, M. Davenport, R. A. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices [J]," *Constr. Approx*, vol. 28, pp. 253–263, 2008.

[99] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit approach via restricted isometry property [J]," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4395–4401, 2010.

[100] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise [J]," *IEEE Trans. Inf. Theory*, vol. 52, pp. 6–18, 2006.

[101] I. E. Nesterov, A. Nemirovskii, and Y. Nesterov, "Interior-point polynominal algorithms in convex programming [J]," *Philadelphia, PA: SIAM*, 1994.

[102] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation [J]," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, 2004.

[103] J. A. Tropp, "Computational methods for sparse solution of linear inverse problems [J]," *Proc of IEEE*, vol. 98, pp. 948–958, 2010.

[104] Z. Ben-Haim and Y. C. Eldar, "The Cramér-Rao bound for estimating a sparse parameter vector [J]," *IEEE Trans. Signal Process*, vol. 58, pp. 3384–3389, 2010.

[105] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise [J]," *IEEE Trans. Signal Process*, vol. 58, pp. 5030–5043, 2010.

[106] D. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage [J]," *Biometrika*, vol. 81, pp. 425–455, 1994.

[107] D. Donoho, "De-noising by soft thresholding [J]," *IEEE Trans. Inf. Theory*, vol. 41, pp. 613–627, 1995.

[108] D. Donoho, Y. Tsaig, I. Drori, and J. C. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit [J]," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1094–1121, 2012.

[109] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit [J]," *Found. Comput. Math*, vol. 9, pp. 317–334, 2009.

[110] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples [J]," *Appl and comput. harmo anal*, vol. 26, pp. 301–321, 2009.

[111] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction [J]," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2230–2249, 2009.

[112] E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies? [J]," *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406–5425, 2006.

[113] W. U. Bajwa, G. R. J. Haupt, and R. Nowak, "Compressed channel sensing [A]," *in Proc. 42nd Conf. Information Science and Systems [C]*, pp. 5–10, 2008.

[114] A. M. Abd-Elfattah, S. H. Amal, and D. M. Ziedan, "Efficiency of Bayes estimator for rayleigh distribution," *Statistics on the Internet*, 2006.

[115] K. Peter, "Phase preserving denoising of images [A]," *in Proc. Conf. Australian Pattern Recognition Society [C]*, pp. 212–217, 1999.

[116] L. Najjar, "Sparsity level-aware threshold-based channel structure detection in OFDM systems [J]," *Electron. Lett*, vol. 48, pp. 495–496, 2012.

[117] Z. P. W. Yi, L. Lihua and L. Zemin, "Optimal threshold for channel estimation in MIMO-OFDM system [A]," *in Proc. IEEE Int. Conf. Commun (ICC) [C]*, pp. 4376–4380, 2008.

[118] S. Sarvotham, D. Bron, and R. G. Baraniuk, "Sudocodes - fast measurement and reconstruction of sparse signals [A]," *in Proc. IEEE Int. Symp. Inf. Theory [C]*, 2006.

[119] H. Xie, G. Adrieux, Y. Wang, J. F. Diouris, and S. Feng, "A novel effective compressed sensing based sparse channel estimation in OFDM system [A]," *in Proc. IEEE Int. Conf. Signal Process, Commun and Comput (ICSPCC) [C]*, pp. 1–6, 2013.

[120] K. M. Nasr, J. P. Cosmas, M. Bard, and J. Gledhill, "Performance of an echo canceller and channel estimator for on-channel repeaters in DVB-T/H networks [J]," *IEEE Trans. Broadcast*, vol. 53, pp. 609–618, 2007.

[121] D. Hu, X. Wang, and L. He, "A new sparse channel estimation and tracking method for time-varying ofdm systems [J]," *IEEE Trans. Veh. Tech*, vol. 62, pp. 4648–4653, 2013.

[122] C. Heegard and S. B. Wicker, *Turbo coding [M].* Kluwer Academic Publishers, 1999.

Research and Published Papers During the PhD

Journal Paper

1. Hui Xie, Guillaume Andrieux, Yide Wang, Jean-François Diouris and Suili Feng, *Efficient time domain threshold for sparse channel estimation in OFDM system*, International Journal of Electronics and Communications (Elsevier), vol. 68, april 2014, pp. 277-281 (SCI).

Conference Papers

1. Hui Xie, Guillaume Andrieux, Yide Wang, Jean-François Diouris and Suili Feng, *Threshold based most significant taps detection for sparse channel estimation in OFDM system*, IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC'2013), 5-8 august 2013, Kunming, Yunnan, China (EI).

2. Hui Xie, Guillaume Andrieux, Yide Wang and Suili Feng, *A novel effective compressed sensing based sparse channel estimation in OFDM system*, IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC'2013), 5-8 august 2013, Kunming, Yunnan, China (EI).

3. Hui Xie, Guillaume Andrieux, Yide Wang, Jean-François Diouris and Suili Feng, *Compressed sensing based non-sample spaced sparse channel estimation in OFDM system*,

Second Sino-French Workshop on Education and Research collaborations in Information and Communication Technologies (SIFWICT 2013), 3-4 june 2013, Guangzhou, China.

l'université nantes angers le mans
PÔLE DE RECHERCHE ET D'ENSEIGNEMENT SUPÉRIEUR

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Hui XIE

Estimation de canal parcimonieux pour les systèmes OFDM

Sparse Channel Estimation in OFDM System

Résumé

L'OFDM est très présent dans les communications sans-fil dû à sa capacité de transmission haut-débit sur des canaux sélectifs en fréquence. Pour une détection cohérente des symboles OFDM, les réponses fréquentielles du canal sont estimées à partir de sous-porteuses pilotes.

Les méthodes d'estimation fréquentielles sont souvent employées avec des méthodes d'interpolation rarement efficaces. La solution simple d'accroitre le nombre de pilotes diminue l'efficacité spectrale du système. Une autre solution est de travailler dans le domaine temporel. En général, la réponse impulsionnelle du canal contient un nombre limité de valeurs significatives. Dans le cas d'un canal parcimonieux, ce nombre est beaucoup plus petit que celui des pilotes. Pour améliorer l'estimation du canal, l'utilisation d'un seuil est nécessaire.

Dans cette thèse, si le nombre de pilotes est plus grand que la longueur du préfixe cyclique, une méthode d'estimation de canal parcimonieux basée sur l'utilisation d'un seuil temporel original est proposée. Cette méthode avec une haute efficacité spectrale, de bonnes performances d'estimation, une faible complexité de calcul, requiert aucune connaissance à priori des statistiques du canal et du bruit.

Si le canal est parcimonieux avec un grand étalement temporel, une technique d'estimation de canal basée sur la théorie de l'acquisition comprimée est proposée. Cette méthode requiert un plus petit nombre de pilotes que les techniques classiques.

Ce travail se termine avec l'étude des canaux dont les trajets sont situés en dehors des instants d'échantillonnage. L'utilisation d'une matrice de mesure intelligente permet d'améliorer l'efficacité des méthodes d'estimation.

Mots clés

Canal parcimonieux, échantillonnage, estimation de canal, méthode LS, estimation de l'écart-type du bruit, acquisition comprimée, seuil, système OFDM.

Abstract

OFDM is widely used in wireless communications due to its capacity of high data rate transmission over frequency selective channel. For coherent detection of OFDM symbols, channel frequency responses must be estimated, which is usually done with the help of pilot tones.

Frequency domain estimation methods are often employed with interpolation methods. Usually, interpolation methods introduce an error floor. The easy solution of increasing the number of pilots decreases the spectral efficiency of the system. Another solution is to work in the time domain. In general, the channel impulse response contains a limited number of significant values having more energy than the noise. In the case of sparse channel, this number is much smaller than that of pilot subcarriers. To improve the sparse channel estimation, some kind of threshold is needed.

In this thesis, in the case where the number of pilots is larger than the length of cyclic prefix, a time domain sparse channel estimation method based on an original threshold is proposed. This method with high spectral efficiency, good channel estimation performance, low computational complexity, requires no prior knowledge of both the channel statistics and noise standard deviation.

In the case where the channel is sparse with large delay spread, we propose an original channel estimation technique based on compressed sensing theory. The proposed method requires smaller number of pilots than that of classical frequency domain techniques.

This work ends with the study of non-sample spaced sparse channel; the idea of smart measurement matrix is proposed to improve the efficiency of the classical CS based estimation methods.

Key Words

Sparse channel, non-sample spaced sparse channel, channel estimation, LS method, noise standard deviation estimation, compressed sensing, threshold, OFDM system.