國立交通大學

電信工程學系

碩 士 論 文

正交分頻多工系統下運用壓縮取樣技術執行通道估測

Channel Estimation in OFDM Systems Using Compressive Sampling Technique

研 究 生：闕瑞慶

指導教授：吳文榕 博士

中 華 民 國 九 十 九 年 七 月

正交分頻多工系統下運用壓縮取樣技術執行通道估測

Channel Estimation in OFDM Systems Using Compressive Sampling Technique

研 究 生：闕瑞慶　　Student: Ruey-Ching Chiueh

指導教授：吳文榕 博士　　Advisor: Dr. Wen-Rong Wu

國 立 交 通 大 學

電 信 工 程 學 系

碩 士 論 文

A Thesis
Submitted to Department of Communication Engineering
College of Electrical and Computer Engineering
National Chiao-Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Science
In
Communication Engineering
July 2010
Hsinchu, Taiwan, Republic of China

中 華 民 國 九 十 九 年 七 月

正交分頻多工系統下運用壓縮取樣技術執行通道估測

學生：闕瑞慶　　　　　　　　　　指導教授：吳文榕 教授

國立交通大學電信工程學系碩士班

摘要

在正交分頻多工(OFDM)系統中，通道估測(channel estimation)經常是藉由安插在 OFDM 符元(symbol)間的領航訊號(pilot)來完成。但是領航訊號的使用卻會影響到系統的效能，越多的領航訊號被安插在 OFDM 符號間則能傳送的資料量就越少，系統的傳送速率(transmission rate)便會下降；此外，在某些系統中領航訊號的數量是有所限制的，因此如何利用少量的領航訊號來達到準確的通道估測便成了一個值得探討的問題。近年來，有研究提出一項名為 Compressive Sampling (CS)的新技術，宣稱只需要運用少許的取樣值便能還原原始的訊號，只要該訊號本身擁有稀疏(sparse)的特性即可。而在時域上的通道響應其非零的位置通常不多，符合 CS 技術的要求，因此我們可以將此技術應用在通道估測的問題上。在本篇論文中，我們提出使用一個 Subspace Pursuit (SP)方法，證明其在通道估測方面比現存應用 CS 技術的眾多方法有更佳的效能表現，並透過回授(feedback)的機制使此方法能在領航訊號密度很低的時候仍保有好的準確度，最後我們延伸此估測法至時變通道。透過模擬結果，可以看出我們所提出的通道估測法在高速移動的環境中依然擁有很好的效能。

Channel Estimation in OFDM Systems Using Compressive Sampling Technique

Student: Ruey-Ching Chiueh　　Advisor: Dr. Wen-Rong Wu

Department of Communication Engineering

National Chiao-Tung University

Abstract

In pilot-assisted OFDM systems, the channel estimation problem is usually solved by the using the pilot subcarriers inserted in OFDM symbols. However, more pilots used will lead to lower transmission rate, and the number of pilots is sometimes limited due to the systems. So we are facing a problem to accurately estimate the channel response while using a small number of pilots. Recently, a novel technique called compressive sampling (CS) has emerged, asserting to recover the sparse signals with a few measurements. Since the number of non-zero taps in time-domain channel response is small, we can then apply the CS methods to the channel estimation problem in OFDM systems. In this thesis, we propose using a subspace pursuit (SP) algorithm which is shown to be superior to the existing CS methods in channel estimation. The performance of proposed method is also shown to be good when pilot density is very low by adding a decision-feedback mechanism. Then, our problem is extended to the time-variant case. And simulation results show the proposed method performs well even when the speed of mobility is high.

誌謝

本篇論文得以順利完成，首先必須要感謝我的指導教授 吳文榕老師，不論在學業或研究問題上，老師總是扮演著明燈的角色，在我遭遇困境時適時的提供專業的意見與指導，對於日常生活與生涯規劃，老師也總是給予相當的關心與建議。這兩年來，在教授身上獲得的不僅僅是更深一層的專業知識，更重要的是學習到從事研究的方法、態度與精神。

此外，我要感謝實驗室的學長姐與同學們，在這期間帶給我許多研究上的幫助、生活中的樂趣以及寶貴的人生經驗談，使我的碩士生活充實了起來。我也要感謝 97 的同學們，尤其是室友們，從大學時期以來同甘共苦經歷的這些時光，都將成為我珍惜的記憶，希望未來大家能夠繼續保持連絡。

最後，我要感謝我的家人以及朋友們，因為有你們的支持與陪伴，我才有力量能夠順利的完成碩士學位，僅以此文獻給你們。

Catalog

List of Tables

List of Figures

Chapter 1 Introduction

Wireless communication technique has attracted more and more attention in recent years since it can overcome the mobility problem. And the demand for high data rate transmission also emerges along with the popularity of high quality video and audio service. Orthogonal Frequency Division Multiplexing (OFDM) offers high spectrum efficiency and strong immunity to multipath fading channel and has become an important modulation technique in wideband wireless communications. It has also been widely used in many applications such as digital audio broadcasting (DAB), digital video broadcasting (DVB), wireless local area network (LAN), and WiMAX.

One of the most important tasks in OFDM receivers is to accurately estimate the channel response in order to recover the transmitted signals. To do that, a common practice is to insert pilot subcarriers in OFDM symbols. Since the data in pilot subcarriers are known, the related channel responses can be estimated and the response of other subcarriers can be interpolated [1],[2]. Pilot subcarriers cannot be used to transmit data and this approach affects the actual data rate. The more pilots we use, the lower the data rate will be. On the other hand, if the density of the pilot subcarriers is not high enough, the channel responses in data subcarriers cannot be accurately estimated and the data rate is affected also. In many applications, the time-domain channel response is sparse. In other words, the delay spread is large but the nonzero taps is a few. In these cases, a large number of pilots are still used. The sparsity of the channel is not explored.

In recent years, the compressive sampling (CS) technique has been developed to recover the sparse signals [3],[4]. Using the CS method, the number of measurements can be reduced dramatically since it exploits the sparse property of the signal. However, the complexity of the existing methods is high, despite the fact that some of them can be solved by the standard linear programming (LP) [5]. Another way to reconstruct the sparse signal is using the greedy algorithm, which retrieves the desired signals from a large redundant set of vectors in an iterative fashion. The matching pursuit (MP) algorithm was developed and proved to be superior to the least squares (LS) algorithm [6],[7]. Later, the orthogonal matching pursuit (OMP) algorithm was introduced on purpose to overcome the re-selection problem occurred in the MP algorithm, and it also showed better performance than MP [8],[9]. Recently, a new method, called subspace pursuit (SP), was developed for the sparse signal reconstruction [10]. It has been shown that its computational complexity is lower than that of OMP and it can have the accuracy of LP.

The CS technique has found many applications in wireless communication, including the time-domain channel estimation in OFDM systems [11]. In the channel estimation of OFDM system, the received signals in pilot subcarriers serve as the measurements. As stated in [12],[13], sparse signals can be exactly recovered under the limited number of measurements when the sensing matrix satisfies the restricted isometry property (RIP). Thus, we can use a small number of pilots to recover the sparse channel response as long as RIP is held.

In this thesis, we study the spare channel estimation problem in OFDM systems. We first apply the SP algorithm and compare it with the LP, MP, and OMP algorithms. From simulation results, we show that SP indeed outperforms other methods. Next, we reduce the pilot density in order to raise the transmission rate. However, this causes aliasing in the time-domain channel response, and the response in the aliasing region cannot be recovered. To overcome the problem, we propose a decision-feedback method. The main idea is first to conduct an initial symbol detection with the partial aliasing-free channel response, and then use some decisions as pseudo pilots. With the original and pseudo pilots, the SP algorithm can then be conducted to estimate the whole channel response. Simulations results show that the proposed SP method works well even when the pilot density is low. Finally, we discuss the channel estimation problem in the high-mobility wireless environments, where the channel becomes time-variant. In time-variant channel, the orthogonality of the subcarriers in one OFDM symbol is no longer held, causing the inter-carrier interference (ICI) effect. For ICI mitigation, accurate channel estimation is needed. We then extend the proposed decision-feedback SP algorithm to the channel estimation in time-variant channels. Simulations also show that the performance of the proposed method is satisfactory.

This thesis is organized as follows. First, we give an introduction to the OFDM system and show the main concepts and terminology of CS reconstruction technique in Chapter 2. In Chapter 3, various channel estimation methods are reviewed. In Chapter 4, decision-feedback SP algorithms for linear time-invariant and time-variant OFDM systems with uniformly distributed pilot subcarriers are proposed. In Chapter 5, we evaluate the performance of the proposed method and with simulations demonstrating its superior performance. Finally, the conclusions and future works are drawn in Chapter 6.

Chapter 2 Introduction to OFDM System and Compressive Sampling

2.1 OFDM system

OFDM is a frequency division multiplexing (FDM) scheme and can be view as a digital multi-carrier modulation technique. In FDM, the high rate stream is divided into several parallel lower rate sub-streams, and this is equivalently to divide the available wideband channel into narrowband sub-channels, and each data stream is transmitted with a subcarrier in a sub-channel. While the data to be transmitted need not to be divided equally nor do they have to originate from the same information source.

The primary advantage of OFDM over single-carrier modulation is the resistance to the frequency selective fading effect. As the bandwidth of each OFDM sub-channel is sufficiently narrow, the effect of frequency selective fading for each transmitted signal in each sub-channel can be considered as flat. Thus, the equalizer at the receiver can be simplified to an one-tap frequency-domain equalizer. Furthermore, since the symbol duration increases for lower rate subcarriers, OFDM provides additional immunity to impulse noise and other impairments and the system stability is raised.

In OFDM systems, the subcarriers are designed to be orthogonal to each other, allowing the spectrum of individual subcarrier overlapping with minimum frequency spacing, achieving high spectral efficiency. Due to the orthogonality, the signal transmitted on each subcarrier can be recovered despite the overlapped spectrum. Figure 2- 1 shows the overlapped spectrum of OFDM modulated signals. Since the sinc-shaped spectrum of one subcarrier is required to be nulled at other subcarriers' frequencies. The subcarrier spacing between two neighbor subcarriers can be calculated as $\Delta f = \frac{W}{N} = \frac{1}{T}$, where W is the bandwidth, N is the number of subcarriers, and T is the symbol period.

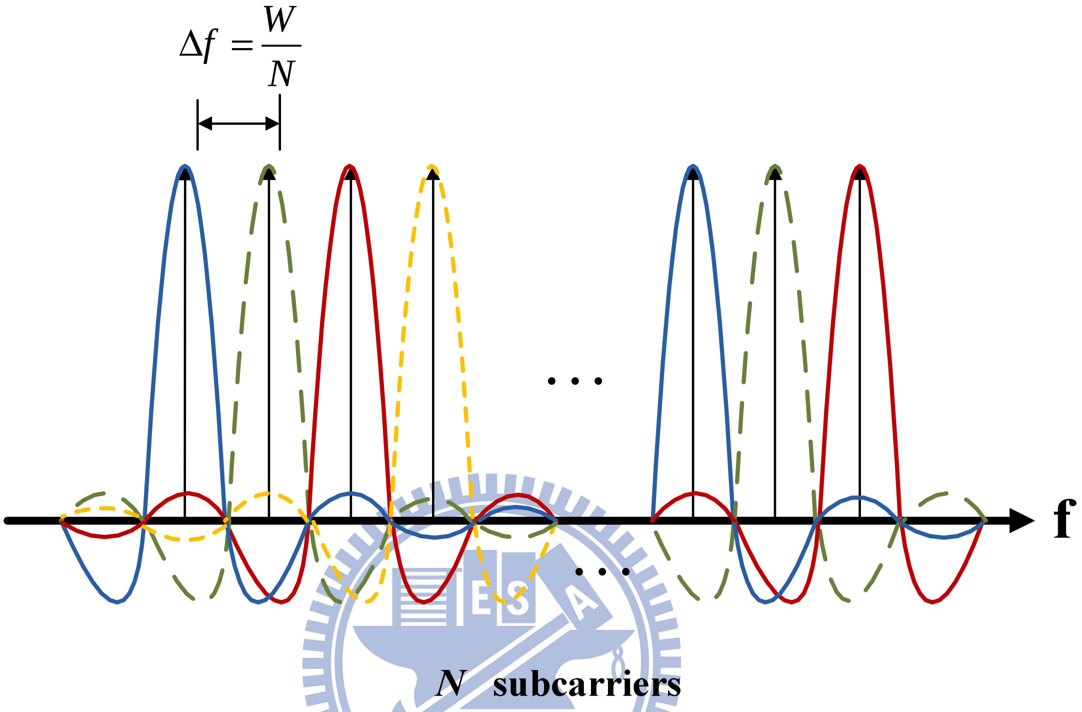

Figure 2- 1 Amplitude spectrum of an OFDM signal with N subcarriers.

In most wireless systems, signal usually travels through different paths causing the multipath effect, which results previous symbols to interfere with the latter symbols and the phenomenon is known as inter-symbol interference (ISI). By adding a cyclic prefix (CP) in front of each symbol, the OFDM scheme offers an effective solution for ISI mitigation. The size of CP is designed to be larger than the maximum channel delay spread, so that the effect of ISI is eliminated. Since CP is a copy of the end portion of an OFDM symbol, the transmitted signal becomes partially periodic, and the effect of the linear convolution with a multipath channel can be translated to a circular convolution. As is known, conducting a circular convolution in the time-domain is equal to conducting a multiplication in the frequency-domain. Thus, the received data in frequency-domain is simplified to a point-to-point multiplication of the data symbol and channel frequency response. Moreover, if the CP length is long enough, the inter-carrier interference (ICI) can also be eliminated to maintain the orthogonality of subcarriers in the multipath fading environments. Figure 2- 2 shows the generation of the CP. In the figure, T denotes the symbol duration excluding CP, T_{CP} the length of CP, and T_s the total symbol duration.

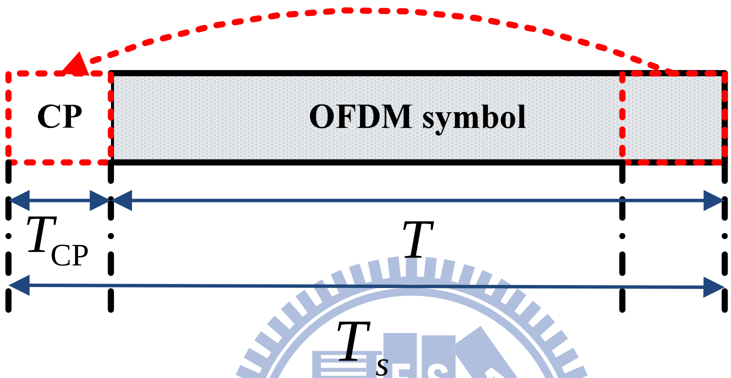

Figure 2- 2 An OFDM symbol with cyclic prefix.

2.1.1 Continuous-time OFDM signal model

Figure 2- 3 shows a typical continuous-time OFDM baseband modulator. The operation of the modulation can be described as below. The transmitted symbol stream is first split into parallel sub-streams using a serial-to-parallel converter and each sub-stream modulates a subcarrier. The modulated signals are then transmitted simultaneously.

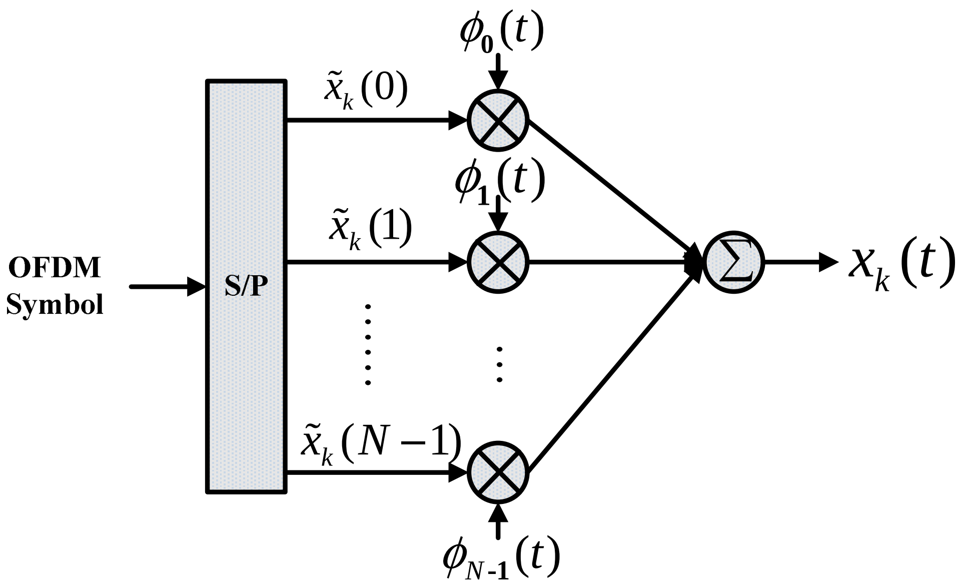

Figure 2- 3 Continuous-time OFDM baseband modulator.

The i -th modulated subcarrier $\phi_i(t)$ can be represented as

$$\phi_i(t) = \begin{cases} \frac{1}{\sqrt{T}} e^{j2\pi k \frac{(t-T_{CP})}{T}} & , t \in [0, T_s = T + T_{CP}) \\ 0 & , otherwise \end{cases} \quad (2.\ 1)$$

In Figure 2- 3, $\tilde{x}_k(i)$ denotes the transmitted symbol, drawn from a set of signal constellation points, at the i -th subcarrier of the k -th OFDM symbol. The modulated baseband signal for the k -th OFDM symbol can then be expressed as

$$x_k(t) = \sum_{i=0}^{N-1} \tilde{x}_k(i)\, \phi_i(t - kT_s), \quad kT_s \le t < (k+1)T_s \quad (2.\ 2)$$

where N is the number of subcarriers. The received signal $y(t)$ can be expressed as

$$y(t) = h(t,\tau) * x(t) + w(t) \quad (2.\ 3)$$

where $h(t,\tau)$ denotes the time-variant channel impulse response at time t, $x(t) = \sum_{k=-\infty}^{\infty} x^k(t)$ is the transmit signal, $w(t)$ is the additive white complex Gaussian noise, and $*$ denotes the operation of linear convolution.

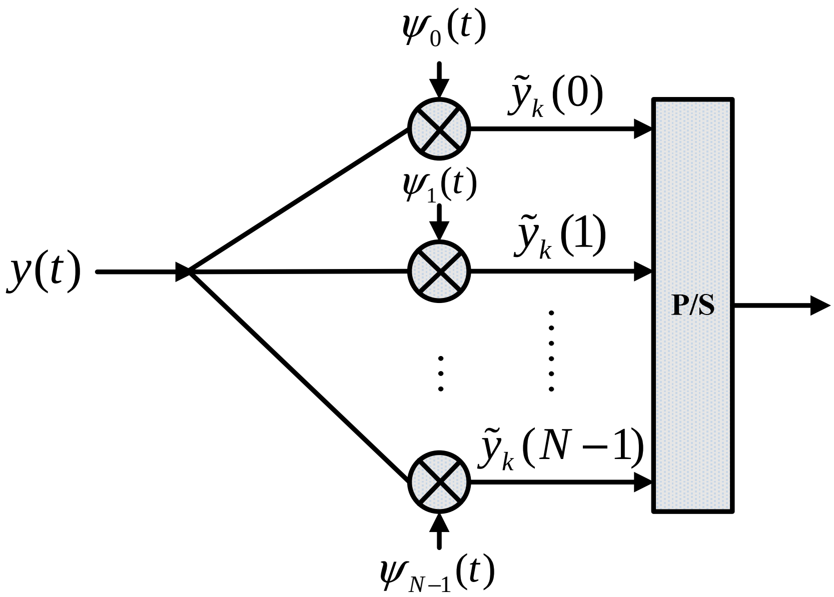

Figure 2- 4 Continuous-time OFDM baseband demodulator.

Figure 2- 4 shows a typical continuous-time OFDM baseband demodulator, in which $\psi_i(t)$ denotes the matched filter for the i-th subcarrier and $\tilde{y}_k(i)$ is the demodulated signal at i-th subcarrier for the k-th symbol. The matched filter is defined as:

$$\psi_i(t) = \begin{cases} \phi_i^*(T_s - t) & , t \in [0, T) \\ 0 & , otherwise \end{cases} \tag{2. 4}$$

2.1.2 Discrete-time OFDM signal model

Consider an OFDM symbol, the modulated baseband signal is given by

$$x(t) = \frac{1}{\sqrt{T}} \sum_{i=0}^{N-1} \tilde{x}_i e^{j\frac{2\pi it}{T}} \quad ,0 \le t \le T \tag{2. 5}$$

where $\tilde{x}_i$ is the transmitted data symbol. Now, sampling the signal $x(t)$ with the sampling period $T_d = \frac{T}{N}$, then (2. 5) can be rewritten as:

$$x[n] = x(t)\big|_{t=nT_d} = \frac{1}{\sqrt{NT_d}} \sum_{i=0}^{N-1} \tilde{x}_i e^{j2\pi i \frac{nT_d}{T}}$$
$$\propto \frac{1}{\sqrt{\mathrm{N}}} \sum_{i=0}^{N-1} \tilde{x}_i e^{\frac{j2\pi in}{N}} = IDFT_{\mathrm{N}}\{\tilde{x}_i\} \quad , 0 \le n \le N-1 \tag{2. 6}$$

For a noise-free system, the discrete demodulated signal $\tilde{y}[k]$ can be expressed as:

$$\tilde{y}[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y[n]\, e^{\frac{-j2\pi kn}{N}} = DFT_{\mathrm{N}}\{y[n]\} \quad , 0 \le k \le N-1 \tag{2. 7}$$

Equations (2. 6) and (2. 7) show that modulation and demodulation in OFDM systems can be conducted by inverse discrete Fourier transform (IDFT) and discrete Fourier transform (DFT), respectively. In practice, IDFT/DFT is implemented with inverse fast Fourier transform (IFFT)/fast Fourier Transform (FFT). Figure 2- 5 shows the discrete-time OFDM system model.

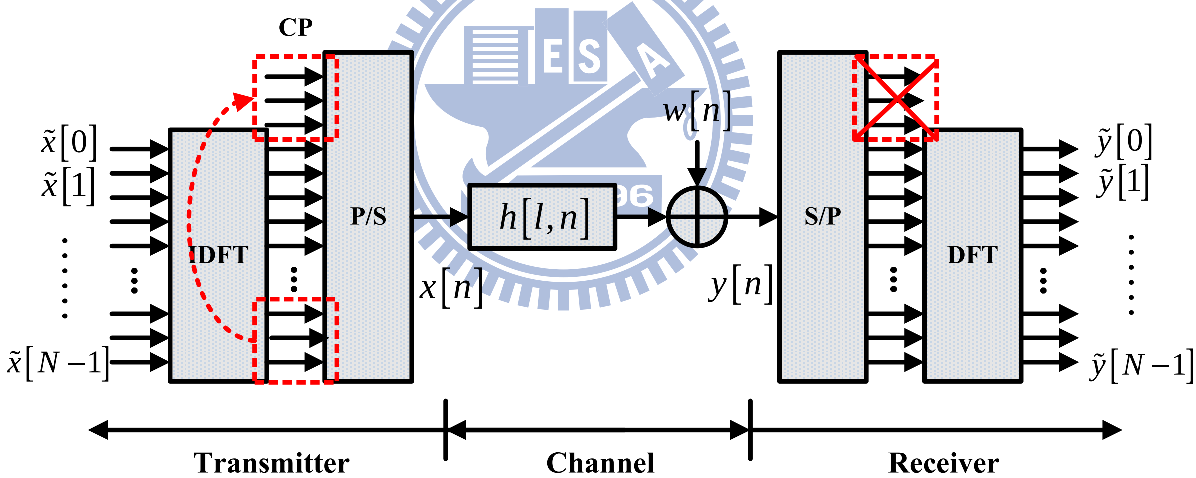

Figure 2- 5 Discrete-time OFDM system model.

Now, the modulation operation can then be summarized as follows. Data streams in the transmitter first modulate N subcarriers, which is performed by a N-point IDFT unit, and then a CP of length T_{CP} is added in the time-domain symbol. The resultant signal $x[n]$ is then passed through a time-variant multipath channel. Assuming that both timing and carrier frequency synchronization are perfect, we can express the received signal $y[n]$ at receiver as

$$y[n] = h[l,n] \otimes x[n] + w[n]$$
$$= \sum_{l=1}^{L} h[l,n]\, x\left[\left((n-l)\right)_N\right] + w[n] \tag{2.8}$$

where $h[l,n]$ is the channel impulse response of the l-th tap at n-th time instant, L is the number of channel taps, $(\ \cdot\ )_N$ represents a cyclic shift in the base of N, $\otimes$ is the circular convolution operator, and $w[n]$ is sampled additive white Gaussian noise (AWGN) with variance σ^2 .

Then a DFT is conducted for each symbol after CP removal, and the received signal in frequency-domain is given by

$$\tilde{y}[k] = \sum_{m=0}^{N-1} \tilde{h}[k,m]\tilde{x}[m] + \tilde{w}[k]$$
$$= \tilde{h}[k,k]\tilde{x}[k] + \tilde{w}[k] \qquad , 0 \le k \le N-1 \tag{2.9}$$

Equation (2. 9) can be expressed using a matrix equivalent model as

$$\begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_{N-1} \end{bmatrix} = \begin{bmatrix} \tilde{h}_{00} & 0 & 0 & \cdots & 0 \\ 0 & \tilde{h}_{11} & 0 & \cdots & 0 \\ 0 & 0 & \tilde{h}_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{h}_{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} \tilde{x}_0 \\ \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_{N-1} \end{bmatrix} + \begin{bmatrix} \tilde{w}_0 \\ \tilde{w}_1 \\ \tilde{w}_2 \\ \vdots \\ \tilde{w}_{N-1} \end{bmatrix} \tag{2.10}$$

where $[\tilde{x}_0, \tilde{x}_1, \cdots, \tilde{x}_{N-1}]^T$ is the frequency-domain transmitted data vector, $[\tilde{y}_0, \tilde{y}_1, \cdots, \tilde{y}_{N-1}]^T$ is the frequency-domain received data vector, $[\tilde{w}_0, \tilde{w}_1, \cdots, \tilde{w}_{N-1}]^T$ is the AWGN noise vector, and $diag\left\{\left[\tilde{h}_{00}, \tilde{h}_{11}, \cdots, \tilde{h}_{(N-1)(N-1)}\right]\right\}$ is an $N \times N$ matrix with $\left[\tilde{h}_{00}, \tilde{h}_{11}, \cdots, \tilde{h}_{(N-1)(N-1)}\right]$ as its diagonal elements denoting the channel frequency response.

2.1.3 Complete OFDM system

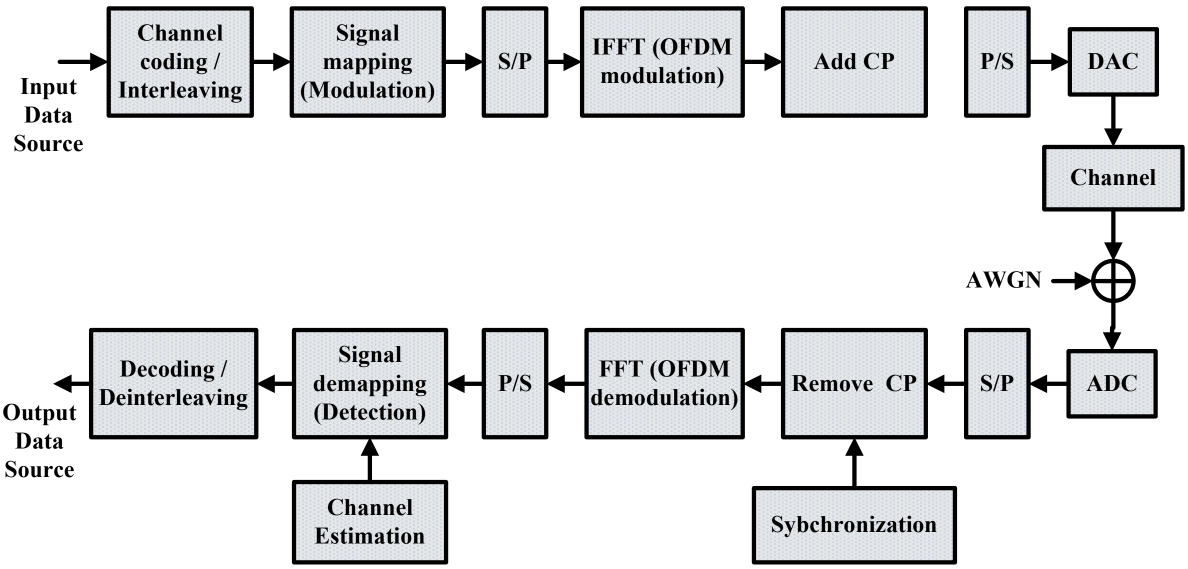

Figure 2- 6 Block diagram of complete OFDM system.

The block diagram of a complete OFDM system is shown in Figure 2- 6. The upper path denotes the transmitter chain, and the lower path is the receiver chain. At the transmitter, the data are first encoded by channel encoder, then interleaved and mapped onto QAM constellation. IFFT operation is then used as a modulator modulating each block of QAM symbols onto subcarriers. After that, a copy of the end portion of the symbol is added in front of each OFDM symbol as a CP. Finally, the baseband OFDM signal is passed to the digital-to-analog (D/A) converter, the RF circuit, and then transmitted. The receiver reverses the operations conducted at the transmitter. Note that synchronization and channel estimation have to be conducted firstly.

2.2 Compressive sampling

Conventionally, if we plan to reconstruct a sampled signal without any error, then a basic principle must be followed：The sampling rate must be at least twice the maximum frequency contained within the signal. The principle is known as Nyquist/Shannon sampling theory, which is one of the crucial theorems in signal processing. However, a novel sampling theory called compressive sampling (CS) goes against this common principle has recently emerged [3],[4]. It asserts that if we know the signal itself is sparse (the support of the coefficient sequence is in a small set) or compressible (the sequence is concentrated near a small set) by some known transformation, then it is possible to uniquely recover the signal from far fewer measurements with high probability. The idea is that for a *K*-sparse signal, which there are only *K* coefficients supported on the signal, the unknowns of the signal are actually the *K* non-zero positions and *K* values.

Now, consider a general problem of recovering a signal $\mathbf{x} \in \mathbb{R}^N$ from a noiseless measurement vector $\mathbf{y} = \left[\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k, \ldots, \mathbf{y}_m\right] \in \mathbb{R}^m$ where

$$\mathbf{y}_k = \langle \mathbf{x}, \varphi_k \rangle, \quad k = 1, \ldots, m. \tag{2. 11}$$

In other words, $\mathbf{x}$ is not directly observed. The measurements are obtained by correlating $\mathbf{x}$ with the waveforms $\varphi_k \in \mathbb{R}^N$. In general, the system is "underdetermined" ($m \ll N$) in the sense that the measurements are much less than the unknown signal values. Solving the ill-posed linear system of equations seems not possible. However, if signal $\mathbf{x}$ is sparse, which means the useful information content embedded in the signal is much smaller than its length/bandwidth, and the problem can be solved by the CS method, which exploits the sparsity and operates as we are directly capturing the information about the signal of importance.

2.2.1 CS recovery methods

Let $\mathbf{\Phi}$ be a matrix taking φ_k as its rows. The relation between the observation $\mathbf{y}$ and K-sparse signal vector $\mathbf{x}$ can be expressed as

$$\mathbf{y} = \mathbf{\Phi x}, \ \ \|\mathbf{x}\|_0 \leq K \tag{2. 12}$$

where $\mathbf{\Phi} \in \mathbb{R}^{m \times N}$ is referred to as the sensing matrix, and $\|\cdot\|_0$ denotes the ℓ_0-norm. It has been shown that one can recover the signal $\mathbf{x}$ by solving an ℓ_0-norm minimization problem:

$$\min \|\mathbf{x}\|_{\mathbf{0}} \ \ \textit{subject to} \ \ \mathbf{y} = \mathbf{\Phi x} \tag{2. 13}$$

However, this approach cannot be used in practical since it is NP-hard [5],[14], and the computational complexity will be very high.

A more computationally efficient strategy was then proposed. The signal recovering problem is now reformulated as a convex optimization problem:

$$\min \|\mathbf{x}\|_{\mathbf{1}} \ \ \textit{subject to} \ \ \mathbf{y} = \mathbf{\Phi x} \tag{2. 14}$$

where

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |\mathbf{x}_i| \tag{2. 15}$$

denotes the ℓ_1-norm of $\mathbf{x}$. This approach can be efficiently implemented by the standard process of linear programming (LP) [5],[15].

A Another way to estimate the sparse signals is the use of greedy algorithms such as Matching Pursuit (MP) [16],[17],[18], Orthogonal Matching Pursuit (OMP) [8],[9],[18], and Regularized Orthogonal Matching Pursuit (ROMP) [19], which iteratively decrease the approximation error by relaxing the sparsity constraint. These algorithms are operated as follows： Search for the supports of signal $\mathbf{x}$ by adding new candidates into the estimated support set and subtract their contribution from the measurement vector $\mathbf{y}$ successively. The objective is to minimize the residue vector $\mathbf{r}_j = \mathbf{y} - \mathbf{\Phi}_j \mathbf{x}_j$ at iteration j. The greedy algorithm provides an effective way to retrieve desired signals, referred as a small subset of vectors, from a large redundant set of vectors.

2.2.2 Robustness of CS theory

To study the reconstruction accuracy of CS, Restricted Isometry Principle (RIP) is introduced to describe the robustness of CS [5],[12],[20]. Let $\mathbf{\Phi}_T$ be a $m \times |T|$ matrix obtained by extracting the $|T|$ columns of $\mathbf{\Phi} \in \mathbb{R}^{m \times N}$ with $T \subset \{\mathbf{1}, \ldots, N\}$. Then matrix $\mathbf{\Phi}$ is said to satisfy the RIP if

$$(1-\delta_k)\|\mathbf{x}\|_2^2 \leq \|\mathbf{\Phi}_T \mathbf{x}\|_2^2 \leq (1+\delta_k)\|\mathbf{x}\|_2^2 \tag{2. 16}$$

for all coefficient sequence $\mathbf{x} \in \mathbb{R}^{|T|}$, $|T| \leq K$, where $K \leq m$ is the sparsity of signal $\mathbf{x}$, $\mathbf{0} \leq \delta_k \leq 1$ is the restricted isometry constant, and $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{N} |\mathbf{x}_i|^2\right)^{\frac{1}{2}}$ denotes the ℓ_2-norm. The principle coveys that when the RIP is held, the columns of sensing matrix $\mathbf{\Phi}$ are approximately orthogonal and the exact recovery achieves. A theorem has been proved in [5] that if signal $\mathbf{x}$ is K-sparse, and the restricted isometry constant satisfies $\delta_{2k} + \delta_{3k} < 1$, the solution of (2. 14) is exact. While the signal is just near sparse as a compressible signal, it has also been proved that the recovery error will be upper-bounded by

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq C \cdot \frac{\|\mathbf{x} - \mathbf{x}_K\|_1}{\sqrt{K}} \tag{2. 17}$$

for some positive constant C and restricted isometry constant $\delta_{3k} + \delta_{4k} < \mathbf{2}$, where $\hat{\mathbf{x}}$ is the solution of (2. 17), and $\mathbf{x}_K$ is the best K-sparse approximations obtained by keeping K

largest coefficients of $\mathbf{x}$ [13].

One may ask how to design a sensing matrix whose columns of size K are nearly orthogonal. For what value of K is this possible? We give some possible sensing matrices in the following：

1) Gaussian measurements: The entries of the sensing matrix is obtained by sampling independent and identically distributed (i.i.d) entries from the normal distribution with zero mean and variance $1/m$. In this case, if the sparsity K obeys, i.e.,

$$K \le C\frac{m}{\log(N/m)} \tag{2. 18}$$

where C is a constant related to the restricted isometry constant, then the probability of exact recovery can be expressed as $1-O(e^{-\gamma N})$ for some $\gamma > 0$ [20],[21].

2) Binary measurements: The entries of the $m\times N$ sensing matrix is obtained by sampling independently the symmetric Bernouli distribution $P(\mathbf{\Phi}_{ij} = \pm\frac{\mathbf{1}}{\sqrt{K}}) = \frac{\mathbf{1}}{\mathbf{2}}$. When (2. 18) is held, the probability of exact recovery is also proved to be $1-O(e^{-\gamma N})$ for some $\gamma > 0$ [20].

3) Fourier measurements: The $m\times N$ sensing matrix $\mathbf{\Phi}$ is obtained by selecting m rows from a $N\times N$ Fourier matrix randomly and the columns of $\mathbf{\Phi}$ are renormalized to have unit norms. Now, the constraint to the sparsity K is

$$K \le C\frac{m}{(\log N)^6} \tag{2. 19}$$

and is refined as

$$K \le C\frac{m}{(\log N)^4} \tag{2. 20}$$

to maintain an overwhelming probability of recovery [20],[22].

4) Incoherent measurements: The sensing matrix is obtained by selecting m rows from an $N\times N$ orthonormal matrix $\mathbf{U}$ randomly, and the columns are normalized to be unit-normed. The matrix $\mathbf{U} = \mathbf{\Phi\Psi}^*$ denotes a transform matrix that transforms the

signal from the $\mathbf{\Psi}$ domain to the $\mathbf{\Phi}$ domain. Then the exact recovery occurs if

$$K \le C \cdot \frac{1}{\mu^2} \cdot \frac{m}{(\log N)^4} \tag{2. 21}$$

where $\mu = \sqrt{N} \cdot \max_{i,j} \left| \left\langle \varphi_i , \psi_j \right\rangle \right|$ is the mutual coherence between $\mathbf{\Psi}$ and $\mathbf{\Phi}$.

In the real-world applications, noise is always present. As a result, (2. 12) becomes

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{z}, \quad \|\mathbf{x}\|_0 \le K \tag{2. 22}$$

where $\mathbf{z}$ is the noise vector with a bounded energy $\|\mathbf{z}\|_2^2 \le \sigma^2$. The problem we have now is to solve the equation shown below:

$$\min \ \|\mathbf{x}\|_1 \quad \textit{subject to} \quad \|\mathbf{y - \Phi x}\|_2 \le \sigma \tag{2. 23}$$

From the CS theory, it asserts that the solution of (2. 23), $\hat{\mathbf{x}}$, obeys

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 \le C \cdot K\sigma^2 \tag{2. 24}$$

for some constant C [13]. Therefore, the stability and robustness of CS are maintained.

The CS technique has been widely applied in many areas. For example, it is used in data compression, sensor networks, and error correcting codes. Recently, it has been applied in channel estimation, which is what we are concerned in this thesis.

Chapter 3 Channel Estimation

In OFDM systems, there are mainly two types of subcarriers allocated, which is shown in Figure 3- 1. The data subcarriers as what it named are used to transmit data symbols, and the pilot subcarriers are the subcarriers, spread uniformly in the frequency-domain, used to conduct channel estimation.

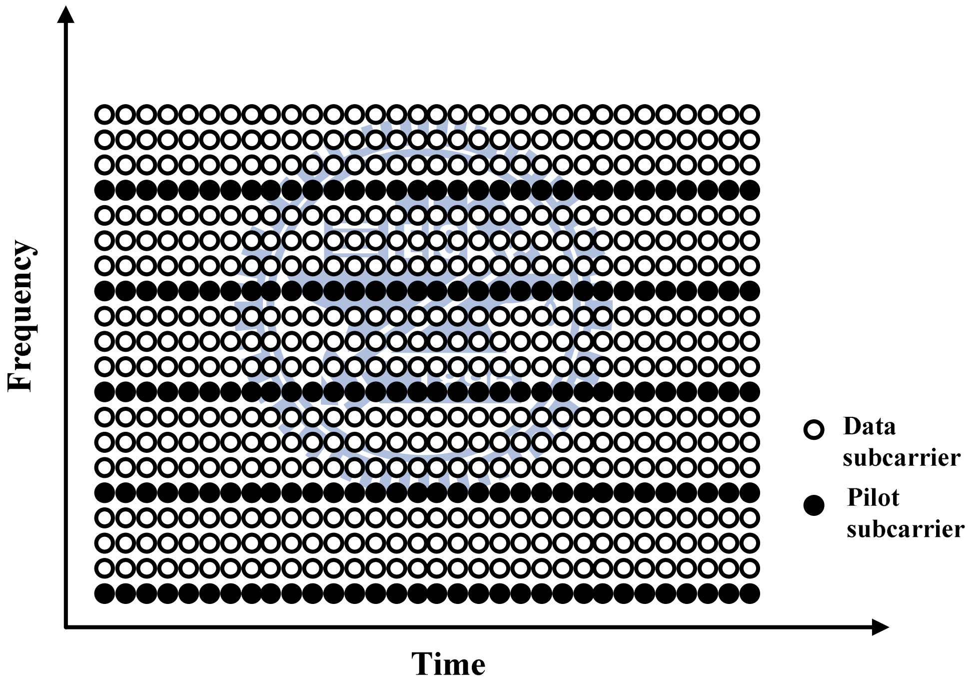

Figure 3- 1 Example of subcarrier allocation in OFDM systems.

3.1 Conventional least squares method

Typically, channel estimation can be performed either in the time-domain or the frequency-domain. The time-domain received signal (after CP removal) can be expressed as

$$y_n = h_n \otimes x_n + w_n \ , \ 0 \le n \le N-1 \tag{3. 1}$$

where N is the number of subcarriers, y_n is the time-domain received signal, x_n is the time-domain training symbols, h_n is the channel impulse response, and w_n is the AWGN noise. Reformulating (3. 1) in the matrix form, we can have

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} x_0 & x_{N-1} & \cdots & \cdots & x_{N-L+1} \\ x_1 & x_0 & \ddots & \cdots & x_{N-L+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{N-2} & x_{N-3} & \cdots & \cdots & x_{N-L-1} \\ x_{N-1} & x_{N-2} & \cdots & \cdots & x_{N-L} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{L-1} \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{N-1} \end{bmatrix} = \mathbf{Xh} + \mathbf{w} \tag{3. 2}$$

where L is the maximum channel delay. A conventional method for channel estimation is the least-squares (LS) method. The LS channel estimate minimizes the squared errors given by：

$$\left\| \mathbf{y} - \mathbf{X}\hat{\mathbf{h}}_{LS} \right\|^2 \tag{3. 3}$$

where $\mathbf{y} = [y_0, y_1, \ldots, y_{N-1}]^T$ is the time-domain received vector. The optimum estimate has been solved as

$$\hat{\mathbf{h}}_{LS} = \left(\mathbf{X}^H \mathbf{X}\right)^{-1} \mathbf{X}^H \mathbf{y} \tag{3. 4}$$

The time-domain LS method described above requires a time-domain training sequence. Another LS channel estimation method, described below, uses pilot subcarriers. Consider a received OFDM symbol in the i-th subcarrier:

$$\tilde{y}_i = \tilde{h}_i \tilde{x}_i + \tilde{w}_i \tag{3. 5}$$

where $\tilde{y}_i$ is the received signal in the frequency-domain, $\tilde{x}_i$ is the transmitted signal in the frequency-domain, $\tilde{h}_i$ is the channel frequency response, and $\tilde{w}_i$ is the corresponding AWGN noise. Note that $\tilde{h}_i$ can be expressed as $\tilde{h}_i = \mathbf{f} \cdot \mathbf{h}$, where is $\mathbf{h}$ the channel response in the time-domain expressed as a vector, and $\mathbf{f}$ is a row of the DFT matrix. Equation (3. 5) then becomes

$$\tilde{y}_i = \tilde{x}_i \mathbf{f}\mathbf{h} + \tilde{w}_i \tag{3. 6}$$

Considering the received signals in all pilot subcarriers, we can have following expression:

$$\begin{bmatrix} \tilde{y}_{p_0} \\ \tilde{y}_{p_1} \\ \tilde{y}_{p_2} \\ \vdots \\ \tilde{y}_{p_{M-1}} \end{bmatrix} = \tilde{\mathbf{X}}_p \begin{bmatrix} e^{-j2\pi\frac{p_0*0}{N}} & e^{-j2\pi\frac{p_0*1}{N}} & \cdots & e^{-j2\pi\frac{p_0*(L-1)}{N}} \\ e^{-j2\pi\frac{p_1*0}{N}} & e^{-j2\pi\frac{p_1*1}{N}} & \cdots & e^{-j2\pi\frac{p_1*(L-1)}{N}} \\ e^{-j2\pi\frac{p_2*0}{N}} & e^{-j2\pi\frac{p_2*1}{N}} & \cdots & e^{-j2\pi\frac{p_2*(L-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi\frac{p_{M-1}*0}{N}} & e^{-j2\pi\frac{p_{M-1}*1}{N}} & \cdots & e^{-j2\pi\frac{p_{M-1}*(L-1)}{N}} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{L-1} \end{bmatrix} + \begin{bmatrix} \tilde{w}_{p_0} \\ \tilde{w}_{p_1} \\ \tilde{w}_{p_2} \\ \vdots \\ \tilde{w}_{p_{M-1}} \end{bmatrix} \tag{3. 7}$$

$$= \tilde{\mathbf{X}}_p \mathbf{F}\mathbf{h} + \tilde{\mathbf{w}}_p$$

where

$$\tilde{\mathbf{X}}_p = \begin{bmatrix} \tilde{x}_{p_0} & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_{p_1} & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_{p_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_{p_{M-1}} \end{bmatrix} \tag{3. 8}$$

is a diagonal matrix with pilot signals as its diagonal elements, p_i , $0 \le i \le M-1$ is the index for pilot location and M is the total number of pilots, $\tilde{\mathbf{y}}_p = \left[\tilde{y}_{p_0}, \tilde{y}_{p_1}, \ldots, \tilde{y}_{p_{M-1}}\right]^T$ is the received frequency-domain vector on pilot locations, $\mathbf{F}$ is a partial DFT matrix obtained by selecting M rows from a N-point DFT matrix according to pilot positions and retaining the first L columns, $\mathbf{h} = \left[h_0, h_1, \ldots, h_{L-1}\right]^T$ is the time-domain channel impulse response, and $\tilde{\mathbf{w}}_p = \left[\tilde{w}_{p_0}, \tilde{w}_{p_1}, \ldots, \tilde{w}_{p_{M-1}}\right]^T$ is the noise vector. Using the LS channel estimator, the following squared error is minimized:

$$\left\| \tilde{\mathbf{y}}_p - \tilde{\mathbf{X}}_p \mathbf{F} \hat{\mathbf{h}}_{LS} \right\|^2 \tag{3. 9}$$

And the resultant time-domain LS channel estimation can then be derived as

$$\hat{\mathbf{h}}_{LS} = \left(\mathbf{F}^H \tilde{\mathbf{X}}_p{}^H \tilde{\mathbf{X}}_p \mathbf{F} \right)^{-1} \mathbf{F}^H \tilde{\mathbf{X}}_p{}^H \tilde{\mathbf{y}}_p \tag{3. 10}$$

The time-domain LS algorithm requires $O\left(L^3\right)$ arithmetic operations, so the computational complexity is high when L is large. But if only L' significant taps where L' is much less than L are taken into account, the computational complexity of the LS method can be reduced. For example, if only two taps in $\mathbf{h}$ are considered, say h_0 and h_{L-1}, then only the first and the last column of $\mathbf{F}$ are required to use. Then, we have

$$\begin{bmatrix} \tilde{y}_{p_0} \\ \tilde{y}_{p_1} \\ \tilde{y}_{p_2} \\ \vdots \\ \tilde{y}_{p_{M-1}} \end{bmatrix} = \tilde{\mathbf{X}}_p \begin{bmatrix} \boxed{\begin{matrix} e^{-j2\pi\frac{p_0*0}{N}} \\ e^{-j2\pi\frac{p_1*0}{N}} \\ e^{-j2\pi\frac{p_2*0}{N}} \\ \vdots \\ e^{-j2\pi\frac{p_{M-1}*0}{N}} \end{matrix}} & \begin{matrix} e^{-j2\pi\frac{p_0*1}{N}} \\ e^{-j2\pi\frac{p_1*1}{N}} \\ e^{-j2\pi\frac{p_2*1}{N}} \\ \vdots \\ e^{-j2\pi\frac{p_{M-1}*1}{N}} \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} e^{-j2\pi\frac{p_0*(L-1)}{N}} \\ e^{-j2\pi\frac{p_1*(L-1)}{N}} \\ e^{-j2\pi\frac{p_2*(L-1)}{N}} \\ \vdots \\ e^{-j2\pi\frac{p_{M-1}*(L-1)}{N}} \end{matrix}} \end{bmatrix} \begin{bmatrix} \boxed{h_0} \\ h_1 \\ h_2 \\ \vdots \\ \boxed{h_{L-1}} \end{bmatrix} + \begin{bmatrix} \tilde{w}_{p_0} \\ \tilde{w}_{p_1} \\ \tilde{w}_{p_2} \\ \vdots \\ \tilde{w}_{p_{M-1}} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{y}_{p_0} \\ \tilde{y}_{p_1} \\ \tilde{y}_{p_2} \\ \vdots \\ \tilde{y}_{p_{M-1}} \end{bmatrix} = \tilde{\mathbf{X}}_p \begin{bmatrix} e^{-j2\pi\frac{p_0*0}{N}} & e^{-j2\pi\frac{p_0*(L-1)}{N}} \\ e^{-j2\pi\frac{p_1*0}{N}} & e^{-j2\pi\frac{p_1*(L-1)}{N}} \\ e^{-j2\pi\frac{p_2*0}{N}} & e^{-j2\pi\frac{p_2*(L-1)}{N}} \\ \vdots & \vdots \\ e^{-j2\pi\frac{p_{M-1}*0}{N}} & e^{-j2\pi\frac{p_{M-1}*(L-1)}{N}} \end{bmatrix} \begin{bmatrix} h_0 \\ h_{L-1} \end{bmatrix} + \begin{bmatrix} \tilde{w}_{p_0} \\ \tilde{w}_{p_1} \\ \tilde{w}_{p_2} \\ \vdots \\ \tilde{w}_{p_{M-1}} \end{bmatrix} \tag{3. 11}$$

With the reduced DFT matrix in (3. 11), the matrix to be inversed in (3. 10) is reduced to a

2×2 matrix. As a result, the computational complexity is reduced.

3.2 Compressive sampling approach

For most wireless channels, the delay spread may be large, but the number of non-zero taps is generally small. That raises the subject in exploiting the sparsity of the time-domain channel response in the channel estimation problem. Since we are dealing with a sparse channel, the CS methods introduced in Chapter 2 are applicable. In the channel estimation of OFDM systems, the number of measurements indicates the number of inserted pilots. Now, we express the frequency-domain received signal vector on pilot positions as

$$\tilde{\mathbf{y}}_{D,p} = \mathbf{F}_{\Omega}\bar{\mathbf{h}} + \tilde{\mathbf{e}}_p \tag{3. 12}$$

where $\bar{\mathbf{h}} = \left[\mathbf{h}^T \quad 0_{1\times(N\text{-}L)}\right]^T$ with $0_{1\times(N\text{-}L)} \in \mathbb{R}^{1\times(N\text{-}L)}$ being a zero vector, $\mathbf{F}_{\Omega}$ is a matrix collecting the rows of a N-point DFT matrix according to pilot indices $\Omega = \{p_0, p_1, \ldots, p_{M-1}\}$, and $\tilde{\mathbf{y}}_{D,p}(i) = \tilde{y}_{p_i} / \tilde{x}_{pi}$, $\tilde{\mathbf{e}}_p(i) = \tilde{w}_{pi} / \tilde{x}_{pi}$, $0 \le i \le M-1$. Equation (3. 12) can be compared with (2. 22), indicating that $\mathbf{F}_{\Omega}$ is a sensing matrix. Thanks to the CS theory, we can estimate the time-domain channel impulse response with low pilot density.

3.2.1 Matching pursuit (MP) algorithm

MP algorithm was the first developed greedy algorithm applying in channel estimation, and it has been shown that MP is superior to the traditional LS method introduced in 3.1 either in the aspect of accuracy or the computational complexity [6],[7]. The block diagram of the MP algorithm is shown in Figure 3- 2.

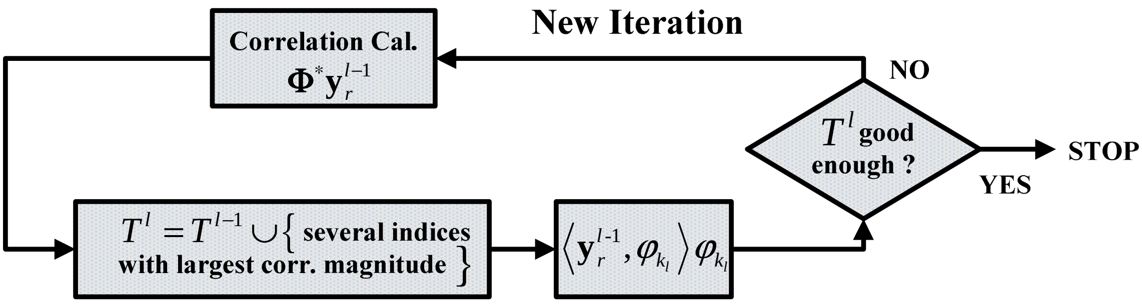

Figure 3- 2 Operations in MP algorithm.

We also summarize the operation of the MP algorithm in Table 3. 1.

Table 3. 1: Matching pursuit algorithm

Define $\mathbf{\Phi} = \left[\mathbf{F}_{\Omega}\left(:,(1:L)\right)\right]$ and $\mathrm{y} = \tilde{\mathbf{y}}_{D,p}$
Input: $\mathbf{\Phi}, \mathbf{y}$
Initialization:
1): $k_0 = \arg \max\limits_{n} \max\limits_{\varphi \in \mathbf{\Phi}} \left\| \langle \mathbf{y}, \varphi_n \rangle \right\|$, $n = 0,1,\ldots,N-1$, $\mathrm{T}^0 = \{k_0\}$.
2): $\mathbf{y}_r^0 = \mathbf{y} - \langle \mathbf{y}, \varphi_{k_0} \rangle \varphi_{k_0}$, $c_0 = \max \left\| \langle \mathbf{y}, \varphi_{k_0} \rangle \right\| \times sign\{\langle \mathbf{y}, \varphi_{k_0} \rangle\}$
Iteration: At the l-th iteration, go through the following steps.
1): $k_l = \arg \max\limits_{n} \max\limits_{\varphi \in \mathbf{\Phi}} \left\| \langle \mathbf{y}_r^{l-1}, \varphi_n \rangle \right\|$, $\mathrm{T}^l = \mathrm{T}^{l-1} \cup k_l$.
2): $\mathbf{y}_r^l = \mathbf{y}_r^{l-1} - \langle \mathbf{y}_r^{l-1}, \varphi_{k_l} \rangle \varphi_{k_l}$, $c_l = \max \left\| \langle \mathbf{y}_r^{l-1}, \varphi_{k_l} \rangle \right\| \times sign\{\langle \mathbf{y}_r^{l-1}, \varphi_{k_l} \rangle\}$
3): If the number of iterations $l = p$ or $\left\\| \mathbf{y}_r^l \right\\|_2 \le \varepsilon$, quit the iteration.
Output:
1): The estimated channel response, satisfying $\hat{\mathbf{h}}_{\{0,1,\ldots,N-1\}-T^l} = 0$ and the non-zero coefficients are stored in $\mathrm{C} = \{c_1, c_2, \ldots, c_l\}$.

The MP algorithm first finds the column vector of $\mathbf{\Phi}$ having the maximum correlation with the measurement vector $\mathbf{y}$. The column index is denoted by k_0 which is added into the index set T^0. Then the k_0-th column of $\mathbf{\Phi}$ is selected to compute the residue vector $\mathbf{y}_r^0$, and the correlation coefficient. The process is then repeated, and the index set during each iteration is updated as $\mathrm{T}^l = \mathrm{T}^{l-1} \cup k_l$ whenever $k_l \notin \mathrm{T}^{l-1}$, otherwise $\mathrm{T}^l = \mathrm{T}^{l-1}$. The algorithm is terminated when either the number of iteration exceeds a preset number, or the residue of the measurement vector is sufficiently small after l iterations.

3.2.2 Orthogonal Matching pursuit (OMP) algorithm

The MP algorithm searches all vectors in each iterative, leading to the re-selection problem. This problem makes the convergence of the algorithm slow. Moreover, MP only optimizes the coefficient of the last selected vector to minimize the error. To overcome these problems, the OMP method was introduced and it has been shown that OMP has the better performance [9]. Figure 3- 3 shows the block diagram of the OMP algorithm, and the corresponding operation is described in Table 3. 2.

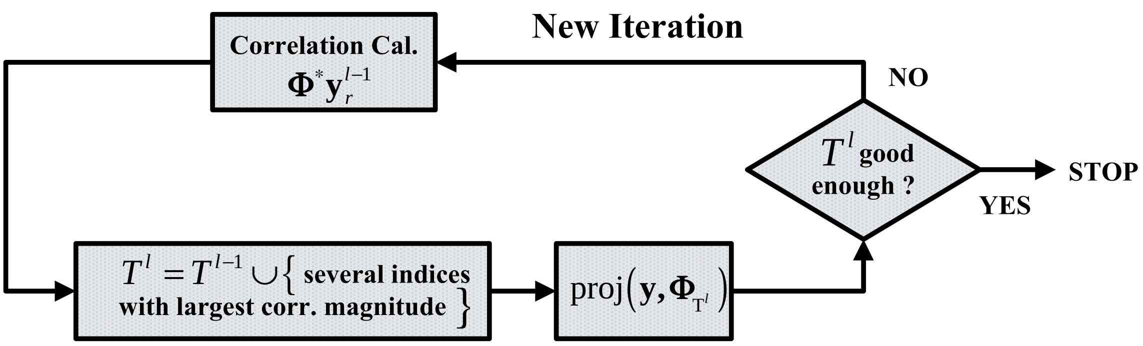

Figure 3- 3 Operations in OMP algorithm.

Table 3. 2: Orthogonal matching pursuit algorithm

Define $\mathbf{\Phi} = \left[\mathbf{F}_{\Omega}\left(:,(1:L)\right)\right]$ and $\mathrm{y} = \tilde{\mathbf{y}}_{D,p}$
Input: $\mathbf{\Phi}, \mathbf{y}$
Initialization:
1): $k_0 = \arg\ \max\left\|\mathbf{\Phi}^{*}\mathbf{y}\right\|$, $\mathrm{T}^{0} = \{k_0\}$.
2): $\mathbf{y}_r^0 = resid\left(\mathbf{y}, \mathbf{\Phi}_{\mathrm{T}^0}\right)$.
Iteration: At the l-th iteration, go through the following steps.
1): $k_l = \arg\ \max\left\|\mathbf{\Phi}^{*}\mathbf{y}_r^{l-1}\right\|$, $\mathrm{T}^{l} = \mathrm{T}^{l-1} \cup k_l$.
2): $\mathbf{y}_r^l = resid\left(\mathbf{y}, \mathbf{\Phi}_{\mathrm{T}^l}\right)$.
3): If the number of iterations $l = p$ or $\left\|\mathbf{y}_r^l\right\|_2 \le \varepsilon$, quit the iteration.
Output:
1): The estimated channel response, satisfying $\hat{\mathbf{h}}_{\{0,1,\ldots,N-1\}-T^l} = 0$ and the non-zero coefficients $\hat{\mathbf{h}}_{T^l} = \mathbf{\Phi}_{T^l}^{\dagger}\mathbf{y}$.
Note here $\mathbf{\Phi}^{\dagger} = \left(\mathbf{\Phi}^{H}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^{H}$, $resid\left(\mathbf{y}, \mathbf{\Phi}\right) = \mathbf{y} - \mathbf{\Phi}\left(\mathbf{\Phi}^{H}\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^{H}\mathbf{y}$ and $\mathbf{\Phi}_{T^l} = \mathbf{\Phi}\left(:, \{\mathrm{T}^l\}\right)$.

The main difference between the MP and OMP algorithm is that the OMP maintains a stored dictionary containing all the indices selected among the iterations. The corresponding vectors are used to compute the new residue vector, avoiding the re-selection problem occurred in MP. The coefficients estimated in OMP are optimized with a set of selected vectors, which results in a much smaller error while the computational complexity is higher.

Chapter 4 Proposed Subspace Pursuit Algorithm in Channel Estimation

4.1 Channel estimation in linear time-invariant (LTI) system

Recently, a new method, called subspace pursuit (SP), was developed for the CS signal reconstruction [10]. The computational complexity of the algorithm has been shown to be lower than that of OMP while the accuracy can approach that of LP. To the best of our knowledge, the SP method has not been used in the OFDM channel estimation problem. Here, we propose using the SP-based method in the channel estimation, and show that SP indeed outperforms the existing CS methods including LP, MP, and OMP by simulations. In this section, we only consider linear time-invariant (LTI) channels.

Recall that the received signal vector corresponding to the pilot sequence expressed as

$$\tilde{\mathbf{y}}_{D,p} = \mathbf{F}_{\Omega}\bar{\mathbf{h}} + \tilde{\mathbf{e}}_p \tag{4. 1}$$

with $\mathbf{F}_{\Omega}$ is the sensing matrix, $\bar{\mathbf{h}}$ the channel response, and $\tilde{\mathbf{e}}_p$ noise. Let $\mathrm{y} = \tilde{\mathbf{y}}_{D,p}$, a schematic diagram of the SP algorithm is shown in Figure 4- 1.

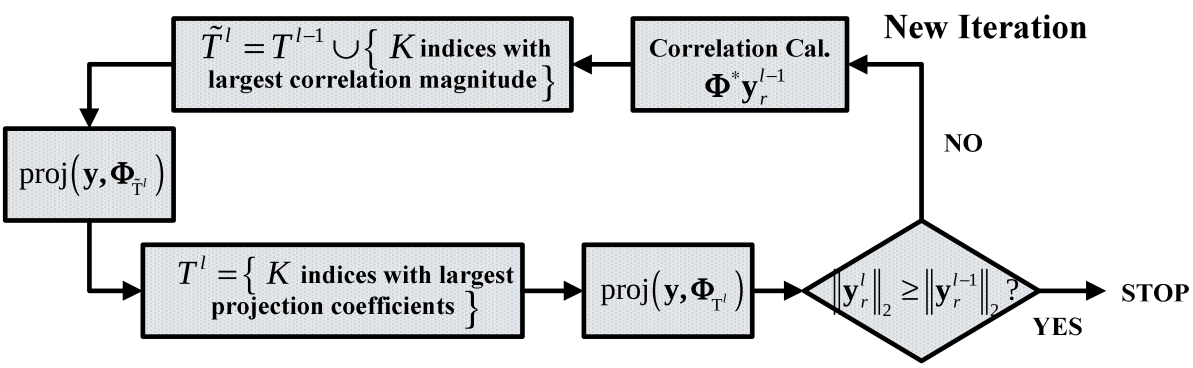

Figure 4- 1 Operations in SP algorithm.

The main operation steps of the SP algorithm are summarized in

Table 4. 1.

Table 4. 1 Subspace Pursuit Algorithm.

Define $\mathbf{\Phi} = \left[\mathbf{F}_{\Omega}\left(:,(1:L)\right)\right]$ and $\mathrm{y} = \tilde{\mathbf{y}}_{D,p}$
Input: $K, \mathbf{\Phi}, \mathbf{y}$
Initialization:
1): $T^0 = \{ K$ indices corresponding to the largest magnitude entries in the vector $\mathbf{\Phi}^* \mathbf{y} \}$.
2): $\mathbf{y}_r^0 = resid\left(\mathbf{y}, \mathbf{\Phi}_{\mathrm{T}^0}\right)$.
Iteration: At the l-th iteration, go through the following steps.
1): $\tilde{T}^l = T^{l-1} \cup \{ K$ indices corresponding to the largest magnitude entries in the vector $\mathbf{\Phi}^* \mathbf{y}_r^{l-1} \}$.
2): Set $\mathbf{h}_p = \mathbf{\Phi}_{\tilde{T}^l}^{\dagger} \mathbf{y}$.
3): $T^l = \{ K$ indices corresponding to the largest magnitude elements of $\mathbf{h}_p \}$.
4): $\mathbf{y}_r^l = resid\left(\mathbf{y}, \mathbf{\Phi}_{\mathrm{T}^l}\right)$.
5): If $\left\Vert \mathbf{y}_r^l \right\Vert_2 \geq \left\Vert \mathbf{y}_r^{l-1} \right\Vert_2$, let $T^l = T^{l-1}$ and quit the iteration.
Output:
1): The estimated channel response $\hat{\mathbf{h}}$, satisfying $\hat{\mathbf{h}}_{\{0,1,\ldots,N-1\}-T^l} = 0$ and $\hat{\mathbf{h}}_{T^l} = \mathbf{\Phi}_{T^l}^{\dagger} \mathbf{y}$.

In the SP algorithm, K indices corresponding to the largest magnitude entries in $\mathbf{\Phi}^* \mathbf{y}$

are first selected to form an index set T^0. Here, K denotes the channel sparsity, and the residue is computed with respect to $\mathbf{\Phi}_{T^0}$ obtained by collecting the K columns from $\mathbf{\Phi}$ according to the elements stored in T^0. During the iteration, an $2K$ index set is formed by K indices that maximizing the correlation between the columns of the sensing matrix and the residue vector united with the K indices found at the previous iteration. Then the size of this set is shrunk to K again by choosing K indices from the largest magnitude elements of $\mathbf{h}_p$, and this index set can be viewed as a refinement to the indices obtained previously. The SP algorithm is stopped when the residue vector derived is larger or equal to the preceding one. After convergence, the tap positions of the channel are then found with the index set, and the coefficient on the taps can be calculated by the LS method.

Due to the refinement, the indices that are mistakenly included in the index set can be removed in the following iterations, and so as the reliable candidates, which can be retrieved at any stage of the recovery process. This is different to the MP-based algorithms which generates the list of candidates sequentially without backtracking. So, more accurate indices of the sparse channel gains can be expected.

4.1.1 SP algorithm without information of tap numbers

The number of channel taps may be unknown to the receiver. Therefore, a revised version of the SP algorithm is needed. A simple idea is to conduct the SP method iteratively with an increasing tap number, and put a threshold on the estimated channel impulse responses. Since the gains of the insignificant taps are usually small, we expect that the number of estimated channel taps after thresholding will be the same whenever the tap number exceeds the channel sparsity K. This can then be used as the stop criterion to the SP algorithm. The flowchart of the proposed algorithm is depicted in Figure 4- 2.

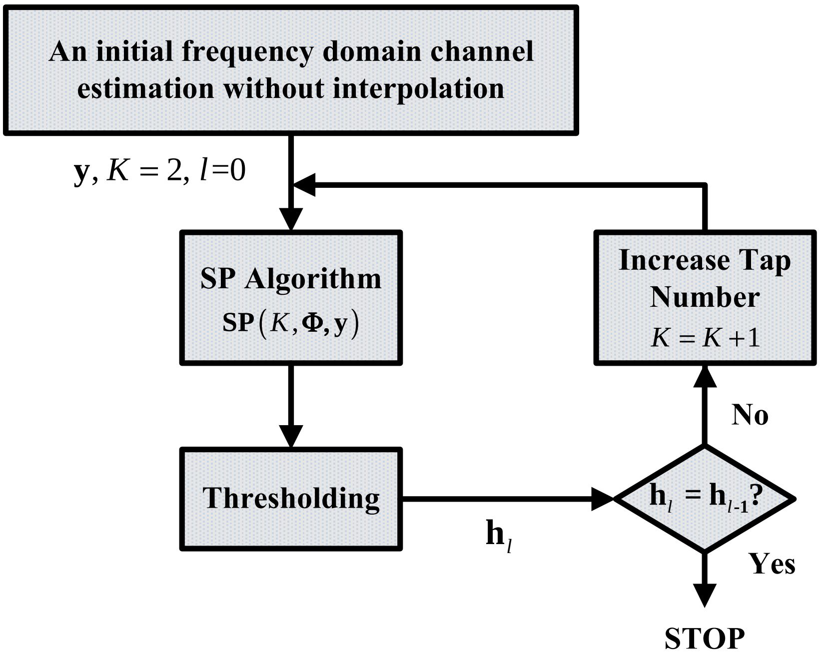

Figure 4- 2 Recursively conducted SP algorithm.

However, the number of the iteration of the algorithm proposed above can be large since the tap number of a channel can be large. If the statistics of the number of the channel taps were known a prior with the help of some statistical properties, we can then use the expected tap number as the first input instead of two used above. Next, we search forward and backward depending on whether the actual taps are more than or less than the expected tap number. In this way, the iteration needed for SP re-conduction will be dramatically reduced when tap numbers are large. This can be seen from the simulation result in chapter 5. Figure 4- 3 shows the flowchart of the modified recursively conducted algorithm.

The upper part of Figure 4- 3 is used to decide whether the true taps are more than or less than the expected ones. The idea is to run the SP algorithm twice. If the estimated channel response after thresholding process has the same tap number, then we know the tap number of the true channel is less than that of the assumed number. Then, go to the right section. Otherwise, go left.

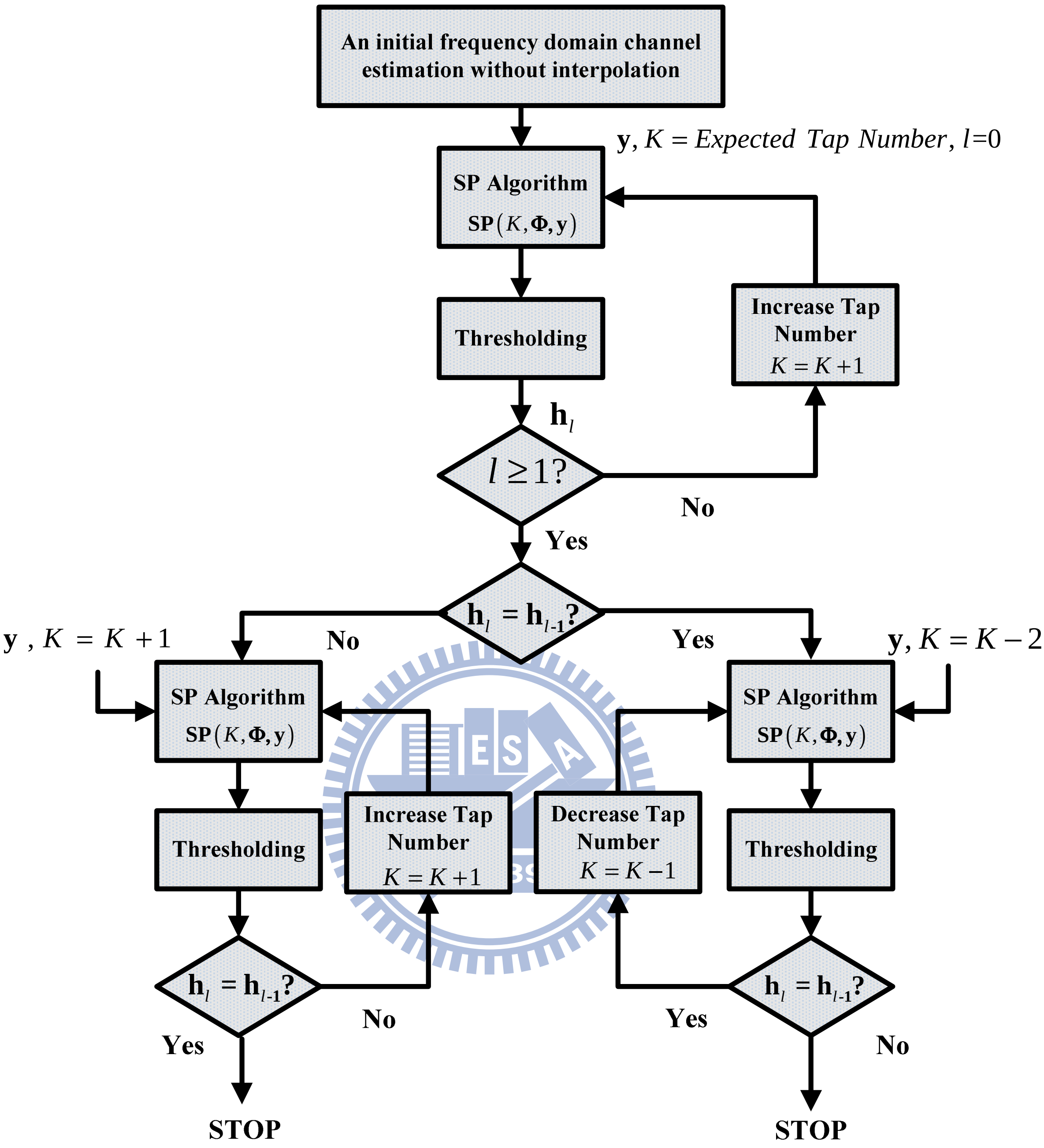

Figure 4- 3 Efficient recursively conducted SP algorithm.

4.1.2 Channel estimation with insufficient pilots

Now, if the pilot density is low, the channel estimate conducted by the SP method may be inaccurate. As mentioned in chapter 3, the pilot subcarriers are inserted uniformly in the spectrum. Since the channel is estimated by the pilot subcarriers, which is equivalent to conduct a sampling on the frequency-domain channel response. As a result, the channel response in time-domain can be seen as periodic. If the sampling rate is not high enough, an effect similar to aliasing will occur. The effect is depicted in Figure 4- 4.

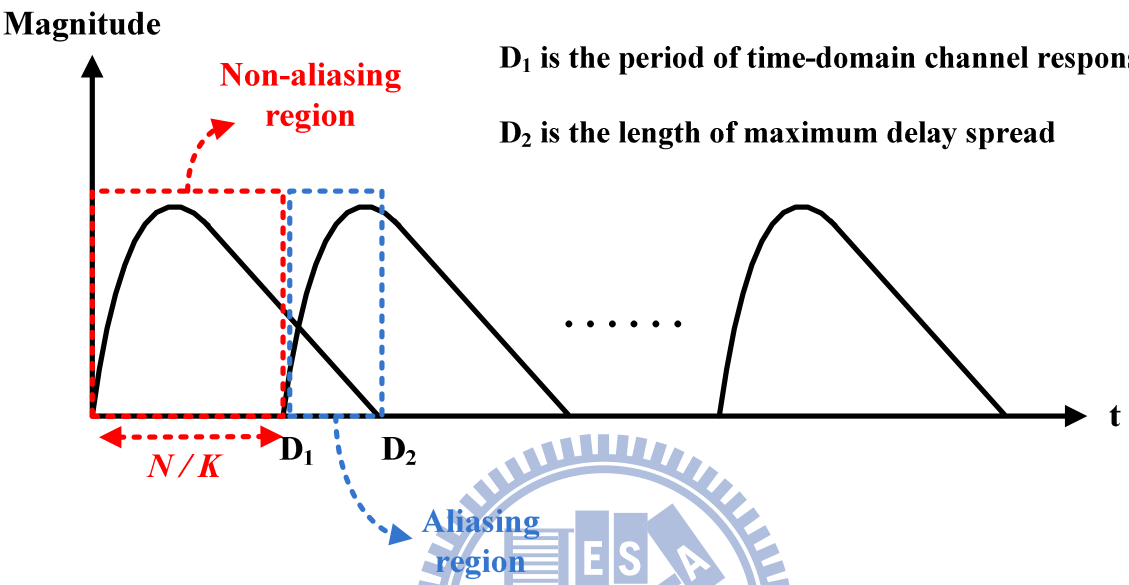

Figure 4- 4 Aliasing in initial channel estimation.

Suppose that the sampling period or we say the pilot interval is K, then the period of the time-domain channel estimate will be N/K, denoted as D_1, where N is the number of subcarriers. Let the maximum delay spread of the channel be D_2, it can be easily observed that the aliasing problem occurs when $D_2 \geq D_1$. And the response in the aliasing area cannot be recovered.

We propose a decision-feedback method to overcome this problem. In the first step, we conduct an initial time-domain channel estimate using the pilots and only take the response in the non-aliasing region. Using the estimated channel frequency response, we can use the SP algorithm to conduct a refined channel estimation and then symbol detection. Finally, using some detected symbols as additional pseudo pilots, we can conduct the SP method again to re-estimate the channel. We now describe the proposed method in detail. With the help of pilot subcarriers, we first obtain a frequency-domain channel response, and then transform the response to the time-domain. Selecting the response in the non-aliasing region and transform them back to the frequency domain, we then obtain a new frequency-domain channel estimate obtained. Figure 4- 5 shows the procedure. Since the aliasing area is not large, and the power of channel taps inside is usually small, the incomplete channel response can be used to recover data with an acceptable error probability.

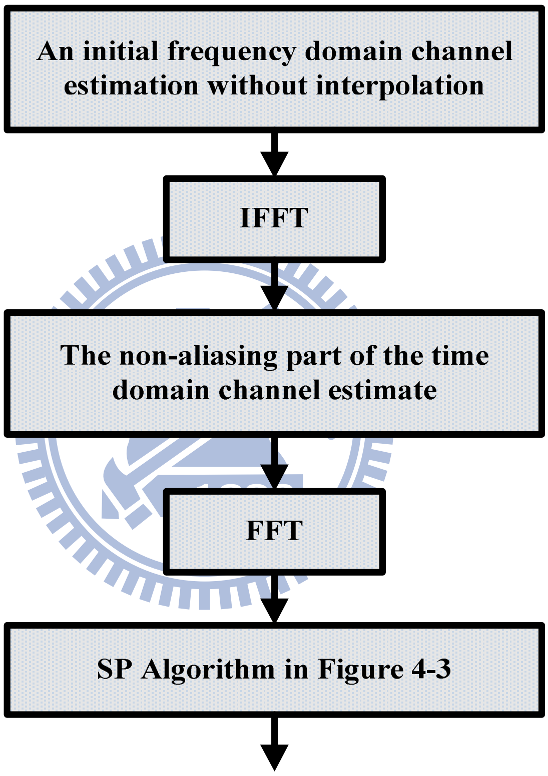

Figure 4- 5 Initial channel estimation method.

Let $\hat{\tilde{h}}_i$ denote the initial frequency-domain channel estimate. We then use the channel response to detect symbols at designated subcarriers. As known, the received signal in the frequency-domain at a subcarrier in subcarrier index i can be expressed as $\tilde{y}_i = \tilde{h}_i \tilde{x}_i + \tilde{w}_i$, where $\tilde{h}_i$ is the frequency-domain channel response, $\tilde{x}_i$ is the transmitted frequency domain signal, $\tilde{w}_i$ is AWGN noise. With the use of the zero-forcing (ZF) equalizer, the estimated symbol at the i-th subcarrier can be calculated by $\hat{\tilde{x}}_i = \frac{\tilde{y}_i}{\hat{\tilde{h}}_i}$. Then the data is sent to the decision device to recover the original transmitted symbol. Let the detected symbol be denoted as $\hat{\tilde{x}}_d$. The flowchart of this procedure is shown in Figure 4- 6.

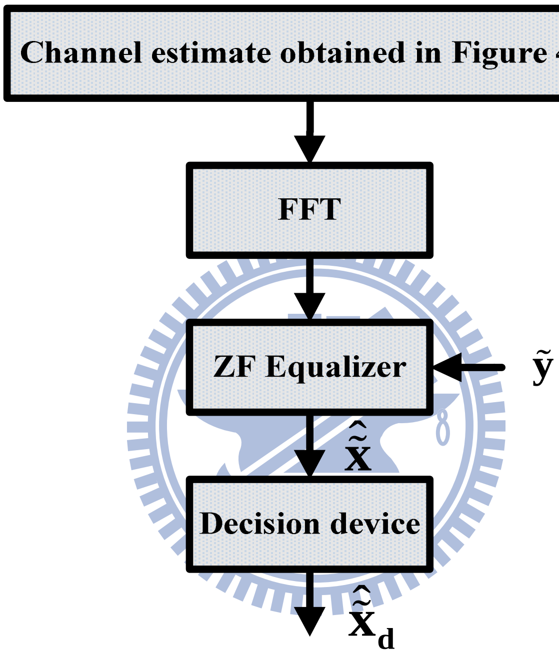

Figure 4- 6 Data detection procedure.

With detected data in hand, we can then choose some of them as additional pilots, increasing the pilot density. By choosing sufficient pseudo pilots, the period of the time-domain channel estimate D_1 will be larger than the maximum delay spread of the channel D_2. Therefore, aliasing will not occur. As long as aliasing does not occur, we can recover the whole channel response. The operation of the re-estimation can be performed as that of the channel estimation described previously, with the increased pilots that combining the original pilots and the pseudo pilots. To obtain better performance, the re-estimation can be conducted iteratively until a convergence is achieved. The complete flowchart of the proposed SP algorithm used with insufficient pilot measurements is shown in Figure 4- 7.

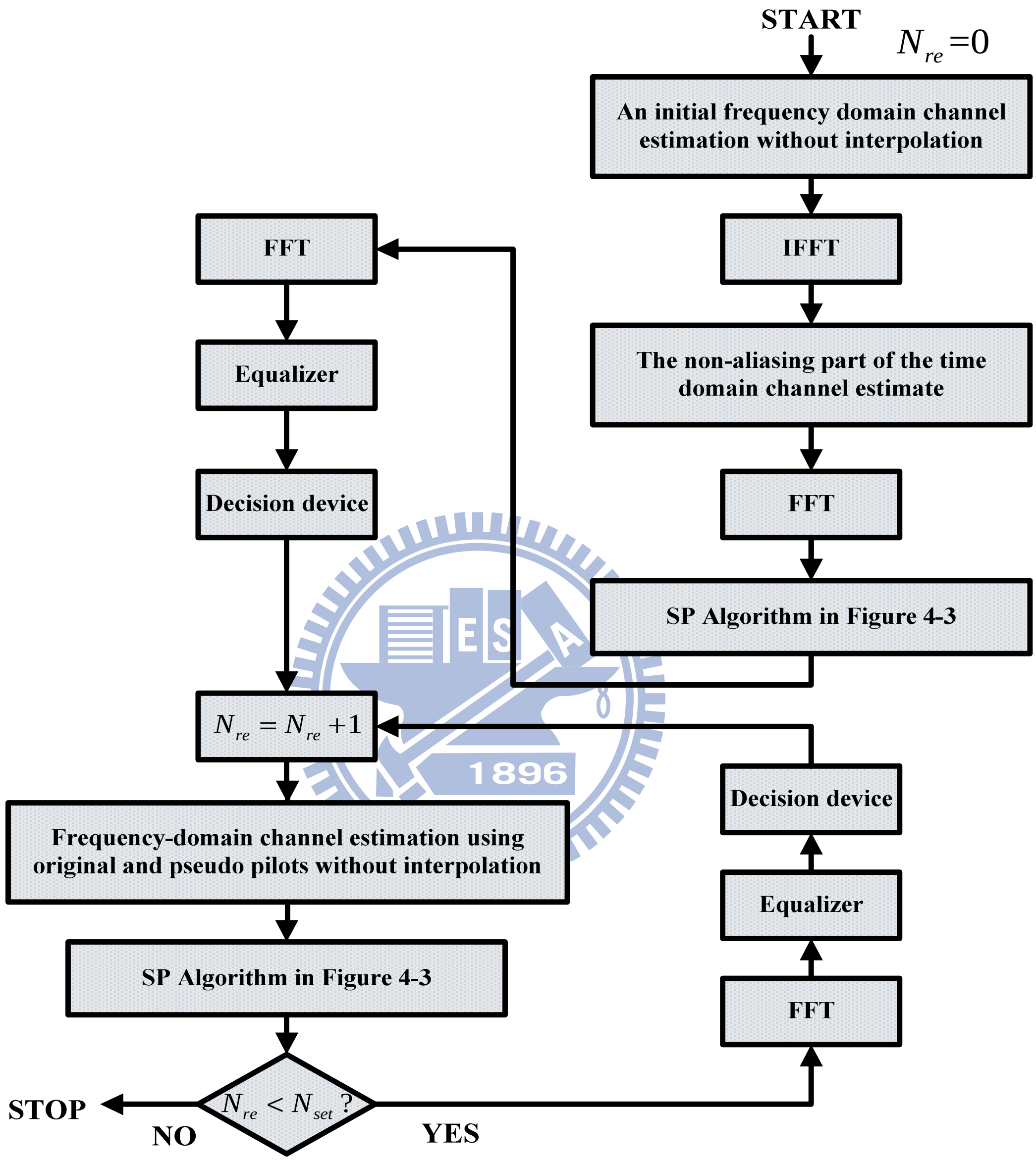

Figure 4- 7 The proposed SP algorithm with insufficient pilot measurements.

The operations conducted in Figure 4- 7 are summarized as follows：

1) Use the pilot subcarriers to obtain an initial frequency-domain channel estimate without interpolation.

2) Transform the frequency-domain estimate into the time-domain and obtain the

time-domain channel estimate.

3) Select the channel taps in the non-aliasing region, and then transform them back to the frequency-domain as a new frequency-domain channel estimate.

4) Conduct the SP algorithm in Figure 4- 3 to estimate the channel impulse response, and detect transmitted symbols with the corresponding estimated frequency-domain channel estimate. Then, make decisions at some designated subcarriers as pseudo pilots.

5) Using the original and pseudo pilots, perform the SP algorithm to re-estimate the whole channel response.

6) If the number of re-estimation N_{re} has not reached a preset value N_{set}, go to 3), where all the taps of channel are in the non-aliasing region when $N_{re}>0$.

4.2 Channel estimation in time-variant system

Nowadays, more and more applications are used in high-mobility wireless environments. In this case, the channel becomes time-variant within an OFDM symbol, violating the common assumption for conventional OFDM systems. This is shown in Figure 4- 8.

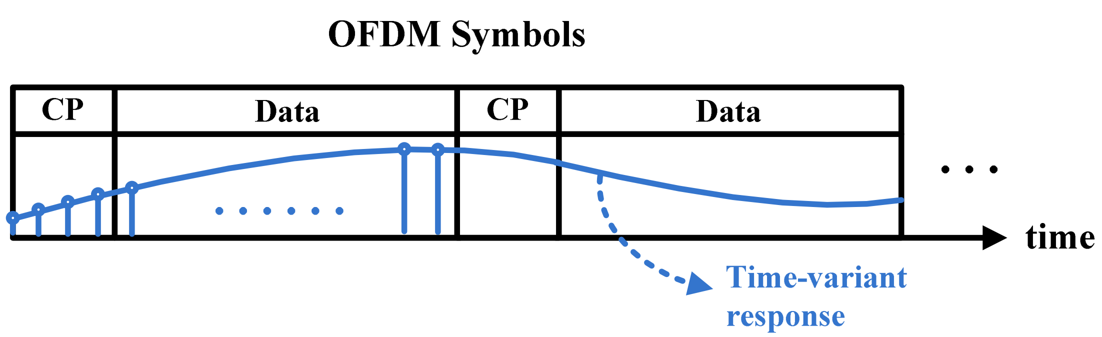

Figure 4- 8 One time-variant tap.

The channel variation rate depends on the Doppler frequency, which is proportional to the carrier frequency and the mobile speed. Denote the maximum Doppler frequency as

$$f_D = \frac{v \times f_C}{c} \tag{4. 2}$$

where v is the mobile speed, f_C is the carrier frequency, and c is the speed of light. The normalized Doppler frequency, defined as $f_D T_s$, is usually used as a parameter to indicate the variation rate. If the sampling period T_s and f_C stay constants, it is obviously that the faster the mobile speed, the larger the normalized Doppler frequency will be.

In time-variant channels, the performance of an OFDM system is degraded when the symbol duration is large. If the coherence time is small compared to the symbol duration, the channel response will change rapidly during one symbol. The orthogonality of the subcarriers in one OFDM symbol is no longer held and the ICI effect occurs. Thus, the frequency-domain channel response in (2. 9) will not be a diagonal matrix. Equation (2. 9) is now rewritten as

$$\begin{aligned}\tilde{y}[k] &= \sum_{m=0}^{N-1} \tilde{h}[k,m]\tilde{x}[m] + \tilde{w}[k] \\ &= \tilde{h}[k,k]\tilde{x}[k] + \sum_{m=0,m\neq k}^{N-1} \tilde{h}[k,m]\tilde{x}[m] + \tilde{w}[k] \quad , 0 \le k \le N-1\end{aligned} \tag{4. 3}$$

where the second term $\sum_{m=0,m\neq k}^{N-1} \tilde{h}[k,m]\tilde{x}[m]$ is known as the ICI term. In this section, we will extend the proposed SP channel estimation in the previous section to the time-variant environments.

4.2.1 Linear approximation method of time-variant channel

Since the variation of channel taps is usually small in typical channels, it is reasonable to model a variation of a channel tap in an OFDM symbol with a linear function. This is equivalent to use a straight line connecting the start and end points of a channel tap within one

OFDM symbol. In this way, only the start point of the channel tap and its variation slope are needed to be specified. The approximation has been shown to be good for the normalized Doppler up to 20% [23]. We show the approximation method in Figure 4- 9.

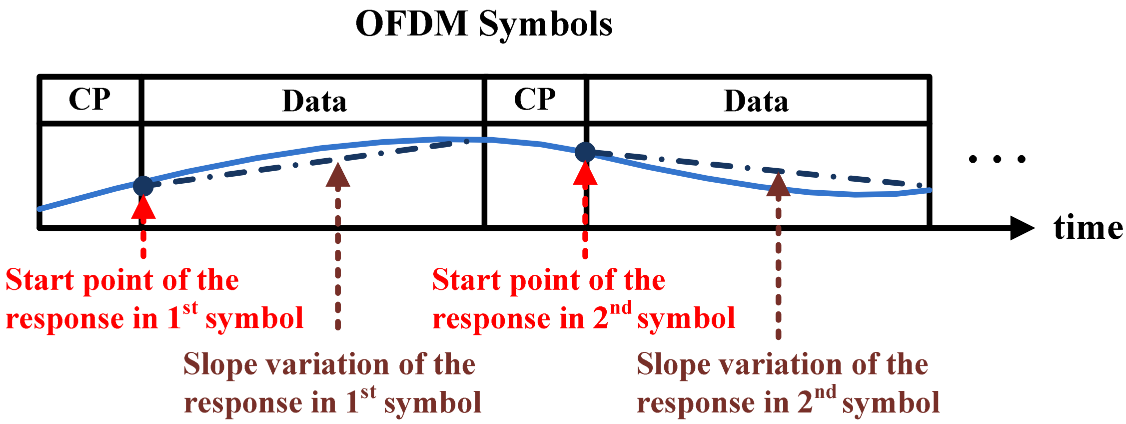

Figure 4- 9 Linear approximation of a time-variant channel tap.

By the linear model introduced above, the variation of a time-variant channel tap can then be expressed as

$$h_\ell[n] = h_\ell + n \times a_\ell \quad , \ 0 \le n \le N-1 \qquad (4.\ 4)$$

where $h_\ell[n]$ denotes the response of the ℓ-th channel tap at time instant n, h_ℓ is the initial value of the response which is set as the response of the start point, and a_ℓ is the variation slope of ℓ-th channel tap.

4.2.2 Time-domain LS estimator in time-variant channel

As known, the time-domain received OFDM symbol can be expressed as the transmitted symbol convolved with the channel impulse response as in (2. 8), and the operation can be formulated in the matrix form as

$$\mathbf{y}=\begin{bmatrix} h_0 & h_{N-1} & h_{N-2} & \cdots & h_1 \\ h_1+a_1 & h_0+a_0 & h_{N-1}+a_{N-1} & \cdots & h_2+a_2 \\ h_2+2a_2 & h_1+2a_1 & h_0+2a_0 & \cdots & h_3+2a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{N-1}+(N-1)a_{N-1} & h_{N-2}+(N-1)a_{N-2} & h_{N-3}+(N-1)a_{N-3} & \cdots & h_0+(N-1)a_0 \end{bmatrix}\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{bmatrix}+\mathbf{w} \quad (4.\ 5)$$

$$=\left(\begin{bmatrix} h_0 & h_{N-1} & \cdots & h_1 \\ h_1 & h_0 & \cdots & h_2 \\ h_2 & h_1 & \cdots & h_3 \\ \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & \cdots & h_0 \end{bmatrix}+\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & N-1 \end{bmatrix}\begin{bmatrix} a_0 & a_{N-1} & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_2 \\ a_2 & a_1 & \cdots & a_3 \\ \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & a_{N-2} & \cdots & a_0 \end{bmatrix}\right)\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{bmatrix}+\mathbf{w} \quad (4.\ 6)$$

For simplicity, we express equation (4. 6) as

$$\mathbf{y}=\left(\mathbf{H}+\mathbf{D}_{\mathbf{v}}\mathbf{A}\right)\mathbf{x}+\mathbf{w} \quad (4.\ 7)$$

where $\mathbf{y}$ is the time-domain received signal vector, $\mathbf{x}=\left[x_0,x_1,\ldots,x_{N-1}\right]^T$ is the time-domain transmitted signal vector, $\mathbf{w}$ is the AWGN noise, $\mathbf{H}$ is a circulant matrix with $\mathbf{h}=\left[h_0,h_1,\ldots,h_{N-1}\right]^T$ as its first column, $\mathbf{A}$ is a circulant matrix with $\mathbf{a}=\left[a_0,a_1,\ldots,a_{N-1}\right]^T$ as its first column, and $\mathbf{D}_{\mathbf{v}}$ is a diagonal matrix with $\mathbf{v}=\left[0,1,\ldots,N-1\right]^T$ as its diagonal elements.

We then transform the time-domain received signal vector to the frequency-domain with a $N\times N$ unitary DFT matrix $\mathbf{F}$. The result is given by

$$\tilde{\mathbf{y}}=\sqrt{N}\cdot\mathbf{F}\mathbf{y} \quad (4.\ 8)$$

where

$$\mathbf{F}=\frac{\mathbf{1}}{\sqrt{N}}\begin{bmatrix} e^{-j2\pi\frac{0*0}{N}} & e^{-j2\pi\frac{0*1}{N}} & \cdots & e^{-j2\pi\frac{0*(N-1)}{N}} \\ e^{-j2\pi\frac{1*0}{N}} & e^{-j2\pi\frac{1*1}{N}} & \cdots & e^{-j2\pi\frac{1*(N-1)}{N}} \\ e^{-j2\pi\frac{2*0}{N}} & e^{-j2\pi\frac{2*1}{N}} & \cdots & e^{-j2\pi\frac{2*(N-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi\frac{(N-1)*0}{N}} & e^{-j2\pi\frac{(N-1)*1}{N}} & \cdots & e^{-j2\pi\frac{(N-1)*(N-1)}{N}} \end{bmatrix} \quad (4.\ 9)$$

and $\mathbf{F}\mathbf{F}^{\mathbf{H}}=\mathbf{I}_N$. By (4. 7), the above equation becomes

$$\begin{aligned}\tilde{\mathbf{y}} &= \sqrt{N}\cdot\mathbf{F}\left[\left(\mathbf{H}+\mathbf{D}_{\mathbf{v}}\mathbf{A}\right)\mathbf{x}+\mathbf{w}\right]\\ &= \sqrt{N}\cdot\mathbf{F}\left(\mathbf{H}+\mathbf{D}_{\mathbf{v}}\mathbf{A}\right)\mathbf{F}^{\mathbf{H}}\mathbf{F}\mathbf{x}+\sqrt{N}\cdot\mathbf{F}\mathbf{w}\\ &= \left(\mathbf{F}\mathbf{H}\mathbf{F}^{\mathbf{H}}+\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{F}\mathbf{A}\mathbf{F}^{\mathbf{H}}\right)\tilde{\mathbf{x}}+\tilde{\mathbf{w}}\\ &= \left[\mathbf{D}_{\tilde{h}}+\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}\right]\tilde{\mathbf{x}}+\tilde{\mathbf{w}}\\ &= \tilde{\mathbf{M}}\tilde{\mathbf{x}}+\tilde{\mathbf{w}}\end{aligned} \tag{4. 10}$$

where $\tilde{\mathbf{y}}$ is the frequency-domain received signal vector, $\tilde{\mathbf{x}}$ is the frequency-domain transmitted signal vector, $\tilde{\mathbf{M}}$ is the frequency-domain ICI matrix, $\tilde{\mathbf{w}}$ is the corresponding frequency-domain noise, $\mathbf{D}_{\tilde{h}}=\mathbf{F}\mathbf{H}\mathbf{F}^{\mathbf{H}}$ and $\mathbf{D}_{\tilde{a}}=\mathbf{F}\mathbf{A}\mathbf{F}^{\mathbf{H}}$ are diagonal matrices with $\tilde{\mathbf{h}}=\sqrt{N}\cdot\mathbf{F}\mathbf{h}$ and $\tilde{\mathbf{a}}=\sqrt{N}\cdot\mathbf{F}\mathbf{a}$ as their diagonal elements respectively, and $\mathbf{D}_{\tilde{h}}$ can be viewed as the channel frequency response without ICI.

Next, we let the ICI matrix be $\tilde{\mathbf{M}}=\mathbf{D}_{\tilde{h}}+\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}$, and rearrange (4. 10) to a more compact form. Neglecting the noise term of (4. 10), we have

$$\tilde{\mathbf{y}}=\mathbf{D}_{\tilde{h}}\tilde{\mathbf{x}}+\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}\tilde{\mathbf{x}} \tag{4. 11}$$

The first term is rearrange as

$$\mathbf{D}_{\tilde{h}}\tilde{\mathbf{x}}=\sqrt{N}\begin{bmatrix}\mathbf{f}_1\mathbf{h} & 0 & 0 & \cdots & 0\\ 0 & \mathbf{f}_2\mathbf{h} & 0 & \cdots & 0\\ 0 & 0 & \mathbf{f}_3\mathbf{h} & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & \mathbf{f}_{\mathbf{N}}\mathbf{h}\end{bmatrix}\begin{bmatrix}\tilde{x}_0\\ \tilde{x}_1\\ \tilde{x}_2\\ \vdots\\ \tilde{x}_{N-1}\end{bmatrix}=\sqrt{N}\begin{bmatrix}\mathbf{f}_1\mathbf{h}\cdot\tilde{x}_0\\ \mathbf{f}_2\mathbf{h}\cdot\tilde{x}_1\\ \mathbf{f}_3\mathbf{h}\cdot\tilde{x}_2\\ \vdots\\ \mathbf{f}_{\mathbf{N}}\mathbf{h}\cdot\tilde{x}_{N-1}\end{bmatrix}$$

$$=\sqrt{N}\begin{bmatrix}\mathbf{f}_1\cdot\tilde{x}_0\\ \mathbf{f}_2\cdot\tilde{x}_1\\ \mathbf{f}_3\cdot\tilde{x}_2\\ \vdots\\ \mathbf{f}_{\mathbf{N}}\cdot\tilde{x}_{N-1}\end{bmatrix}\begin{bmatrix}h_0\\ h_1\\ h_2\\ \vdots\\ h_{N-1}\end{bmatrix} \tag{4. 12}$$

where $\mathbf{f}_i$ denotes the i-th row of the DFT matrix $\mathbf{F}$. The second term in (4. 11) is

$$\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}\tilde{\mathbf{x}} = \sqrt{N}\cdot\mathbf{F}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{N-1} \end{bmatrix}\mathbf{F}^{\mathbf{H}}\begin{bmatrix} \mathbf{f}_1\mathbf{a} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{f}_2\mathbf{a} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{f}_3\mathbf{a} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{f}_{\mathbf{N}}\mathbf{a} \end{bmatrix}\begin{bmatrix} \tilde{x}_0 \\ \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_{N-1} \end{bmatrix}$$

$$= \sqrt{N}\cdot\mathbf{F}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{N-1} \end{bmatrix}\mathbf{F}^{\mathbf{H}}\begin{bmatrix} \tilde{x}_0 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_1 & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{x}_{N-1} \end{bmatrix}\begin{bmatrix} \mathbf{f}_1\mathbf{a} \\ \mathbf{f}_2\mathbf{a} \\ \mathbf{f}_3\mathbf{a} \\ \vdots \\ \mathbf{f}_{\mathbf{N}}\mathbf{a} \end{bmatrix} \qquad (4.\ 13)$$

Let

$$\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}} = \mathbf{F}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{N-1} \end{bmatrix}\mathbf{F}^{\mathbf{H}}$$

$$= \begin{bmatrix} w_{00} & w_{01} & w_{02} & \cdots & w_{0(N-1)} \\ w_{10} & w_{11} & w_{12} & \cdots & w_{1(N-1)} \\ w_{20} & w_{21} & w_{22} & \cdots & w_{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{(N-1)0} & w_{(N-1)1} & w_{(N-1)2} & \cdots & w_{(N-1)(N-1)} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \vdots \\ \mathbf{w}_{\mathbf{N}} \end{bmatrix} \qquad (4.\ 14)$$

Then (4. 13) becomes

$$\mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}\tilde{\mathbf{x}} = \sqrt{N}\cdot\mathbf{F}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{N-1} \end{bmatrix}\mathbf{F}^{\mathbf{H}}\begin{bmatrix} \tilde{x}_0 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_1 & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{x}_{N-1} \end{bmatrix}\begin{bmatrix} \mathbf{f}_1\mathbf{a} \\ \mathbf{f}_2\mathbf{a} \\ \mathbf{f}_3\mathbf{a} \\ \vdots \\ \mathbf{f}_{\mathbf{N}}\mathbf{a} \end{bmatrix}$$

$$= \sqrt{N}\begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \vdots \\ \mathbf{w}_{\mathbf{N}} \end{bmatrix}\begin{bmatrix} \tilde{x}_0 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_1 & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{x}_{N-1} \end{bmatrix}\begin{bmatrix} \mathbf{f}_1\mathbf{a} \\ \mathbf{f}_2\mathbf{a} \\ \mathbf{f}_3\mathbf{a} \\ \vdots \\ \mathbf{f}_{\mathbf{N}}\mathbf{a} \end{bmatrix}$$

$$= \sqrt{N}\begin{bmatrix} \mathbf{w_1} \\ \mathbf{w_2} \\ \mathbf{w_3} \\ \vdots \\ \mathbf{w_N} \end{bmatrix}\begin{bmatrix} \tilde{x}_0 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_1 & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{x}_{N-1} \end{bmatrix}\mathbf{F}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix}$$

$$= \sqrt{N}\begin{bmatrix} \mathbf{v_0} \\ \mathbf{v_1} \\ \mathbf{v_2} \\ \vdots \\ \mathbf{v_{N-1}} \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix} \tag{4. 15}$$

with $\mathbf{v}_i$ as row vectors, and

$$\begin{bmatrix} \mathbf{v_0} \\ \mathbf{v_1} \\ \mathbf{v_2} \\ \vdots \\ \mathbf{v_{N-1}} \end{bmatrix} = \begin{bmatrix} v_{00} & v_{01} & v_{02} & \cdots & v_{0(N-1)} \\ v_{10} & v_{11} & v_{12} & \cdots & v_{1(N-1)} \\ v_{20} & v_{21} & v_{22} & \cdots & v_{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{(N-1)0} & v_{(N-1)1} & v_{(N-1)2} & \cdots & v_{(N-1)(N-1)} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{w_1} \\ \mathbf{w_2} \\ \mathbf{w_3} \\ \vdots \\ \mathbf{w_N} \end{bmatrix}\begin{bmatrix} \tilde{x}_0 & 0 & 0 & \cdots & 0 \\ 0 & \tilde{x}_1 & 0 & \cdots & 0 \\ 0 & 0 & \tilde{x}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{x}_{N-1} \end{bmatrix}\mathbf{F} \tag{4. 16}$$

For derivation simplicity, we ignore the noise term and rewrite (4. 11) as follows:

$$\tilde{\mathbf{y}} = \mathbf{D}_{\tilde{h}}\tilde{\mathbf{x}} + \mathbf{F}\mathbf{D}_{\mathbf{v}}\mathbf{F}^{\mathbf{H}}\mathbf{D}_{\tilde{a}}\tilde{\mathbf{x}} \Rightarrow$$

$$\begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_{N-1} \end{bmatrix} = \sqrt{N}\left(\begin{bmatrix} \mathbf{f_1}\cdot\tilde{x}_0 \\ \mathbf{f_2}\cdot\tilde{x}_1 \\ \mathbf{f_3}\cdot\tilde{x}_2 \\ \vdots \\ \mathbf{f_N}\cdot\tilde{x}_{N-1} \end{bmatrix}\begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{N-1} \end{bmatrix} + \begin{bmatrix} \mathbf{v_0} \\ \mathbf{v_1} \\ \mathbf{v_2} \\ \vdots \\ \mathbf{v_{N-1}} \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix}\right) \tag{4. 17}$$

Combining the two terms in the parenthesis of the above equation, we get

$$\begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_{N-1} \end{bmatrix} = \sqrt{N} \begin{bmatrix} \mathbf{f_1}\tilde{x}_0 & \mathbf{v_0} \\ \mathbf{f_2}\tilde{x}_1 & \mathbf{v_1} \\ \mathbf{f_3}\tilde{x}_2 & \mathbf{v_2} \\ \vdots & \vdots \\ \mathbf{f_N}\tilde{x}_{N-1} & \mathbf{v_{N-1}} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{N-1} \\ a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix} \tag{4. 18}$$

And then

$$\tilde{\mathbf{y}} = [\mathbf{U} \;\; \mathbf{V}] \times \mathbf{h}_a = \mathbf{Q} \times \mathbf{h}_a \tag{4. 19}$$

where $\tilde{\mathbf{y}} = [\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{N-1}]^T$ is the frequency-domain received signal vector, $\mathbf{Q} = [\mathbf{U} \;\; \mathbf{V}]$, $\mathbf{U} = \sqrt{N} \cdot [\mathbf{f_1}\tilde{x}_0, \mathbf{f_2}\tilde{x}_1, \ldots, \mathbf{f_N}\tilde{x}_{N-1}]^T$, $\mathbf{V} = \sqrt{N} \cdot [\mathbf{v_0}, \mathbf{v_1}, \ldots, \mathbf{v_{N-1}}]^T$, and $\mathbf{h}_a$ is an $2N \times 1$ vector containing the start values of the response and its variation slopes that we want to estimate. The LS algorithm can then be conducted by

$$\hat{\mathbf{h}}_a = \left(\mathbf{Q}^{\mathbf{H}}\mathbf{Q}\right)^{-1} \mathbf{Q}^{\mathbf{H}} \tilde{\mathbf{y}} \tag{4. 20}$$

While the data are available only on pilot subcarriers, so only the rows of $\mathbf{Q}$ that corresponding to the pilot positions are taken into account. And since the channel is sparse, many elements in $\mathbf{h}_a$ are actually zero, then only do the columns of $\mathbf{Q}$ corresponding to the non-zero tap positions are needed to perform the LS method. For example, if there are only two significant taps in $\mathbf{h}_a$, say h_0 and h_2, then only the first and third column of $\mathbf{U}$ and $\mathbf{V}$ need to be considered. We show this relationship in Figure 4- 10.

Let $\mathbf{Q} = [\mathbf{U}\ \ \mathbf{V}]$

$$= \begin{bmatrix} u_{00} & u_{01} & u_{02} & \cdots & u_{0(N-1)} & v_{00} & v_{01} & v_{02} & \cdots & v_{0(N-1)} \\ u_{10} & u_{11} & u_{12} & \cdots & u_{1(N-1)} & v_{10} & v_{11} & v_{12} & \cdots & v_{1(N-1)} \\ u_{20} & u_{21} & u_{22} & \cdots & u_{2(N-1)} & v_{20} & v_{21} & v_{22} & \cdots & v_{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{(N-1)0} & u_{(N-1)1} & u_{(N-1)2} & \cdots & u_{(N-1)(N-1)} & v_{(N-1)0} & v_{(N-1)1} & v_{(N-1)2} & \cdots & v_{(N-1)(N-1)} \end{bmatrix} \tag{4. 21}$$

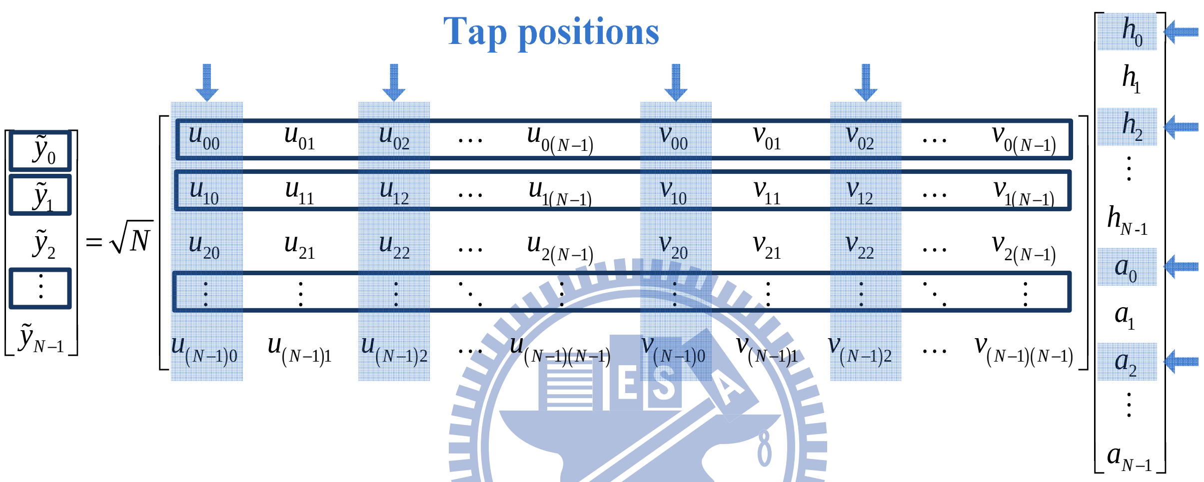

Figure 4- 10 The entries to be selected.

After removing the irrelevant elements in (4. 19), we then have a simplified form of (4. 19) as

$$\tilde{\mathbf{y}}_p = [\mathbf{U}_{pk}\ \ \mathbf{V}_{pk}] \times \mathbf{h}_{a_k} = \mathbf{Q}_{pk} \times \mathbf{h}_{a_k} \tag{4. 22}$$

where $\tilde{\mathbf{y}}_p = [\tilde{y}_{p_0}, \tilde{y}_{p_1}, \ldots, \tilde{y}_{p_{M-1}}]^T$ is the frequency-domain received signal vector on pilot subcarriers with p_i, $0 \le i \le M-1$ as pilot positions and M as the total number of pilots, $\mathbf{U}_{pk}$ and $\mathbf{V}_{pk}$ are sub-matrices of $\mathbf{U}$ and $\mathbf{V}$ respectively, whose rows are determined by pilot positions and the columns are selecting according to the non-zero tap positions, and $\mathbf{h}_{a_k} = \left[[h_{k_0}, h_{k_1}, \ldots, h_{k_{K-1}}], [a_{k_0}, a_{k_1}, \ldots, a_{k_{K-1}}] \right]^T$ is a parameter vector containing the K start values of the channel taps and the K variation slopes with k_j, $\mathbf{0} \le j \le K-1$ as its tap positions. From (4. 22), the LS solution for $\mathbf{h}_{a_k}$ is obtained as

$$\hat{\mathbf{h}}_{a_k} = \left(\mathbf{Q}_{pk}{}^{\mathbf{H}}\mathbf{Q}_{pk}\right)^{-1}\mathbf{Q}_{pk}{}^{\mathbf{H}}\tilde{\mathbf{y}}_p \tag{4. 23}$$

In the time-variant systems, the unknowns of the channel are the starting values of the channel taps and the corresponding variation slopes. Thus, the parameters we need to estimate are as twice as those in the time-invariant channels. As a result, the accuracy of estimation result can be affected. With the proposed algorithm, however, the problem can be alleviated since data decisions can be used as additional pseudo pilots as introduced in Section 4.1.2. In the re-estimation, the entries of the received signal vector and the rows of matrices $\mathbf{U}$ and $\mathbf{V}$ are then selected by the original and pseudo pilot positions.

4.2.3 Proposed method in time-variant channel estimation

We have discussed the proposed time-invariant channel estimation method in Section 4.1. The strategy for the time-variant channel estimation is the same：Use the SP algorithm to conduct tap searching. After the tap positions have been located, the values on the positions can then be computed with the LS estimator as described in previous section. The main idea for performance improvement is the application of decision feedback. The block diagram of the proposed scheme is shown in Figure 4- 11, and the detail operations are summarized in the procedure following the figure.

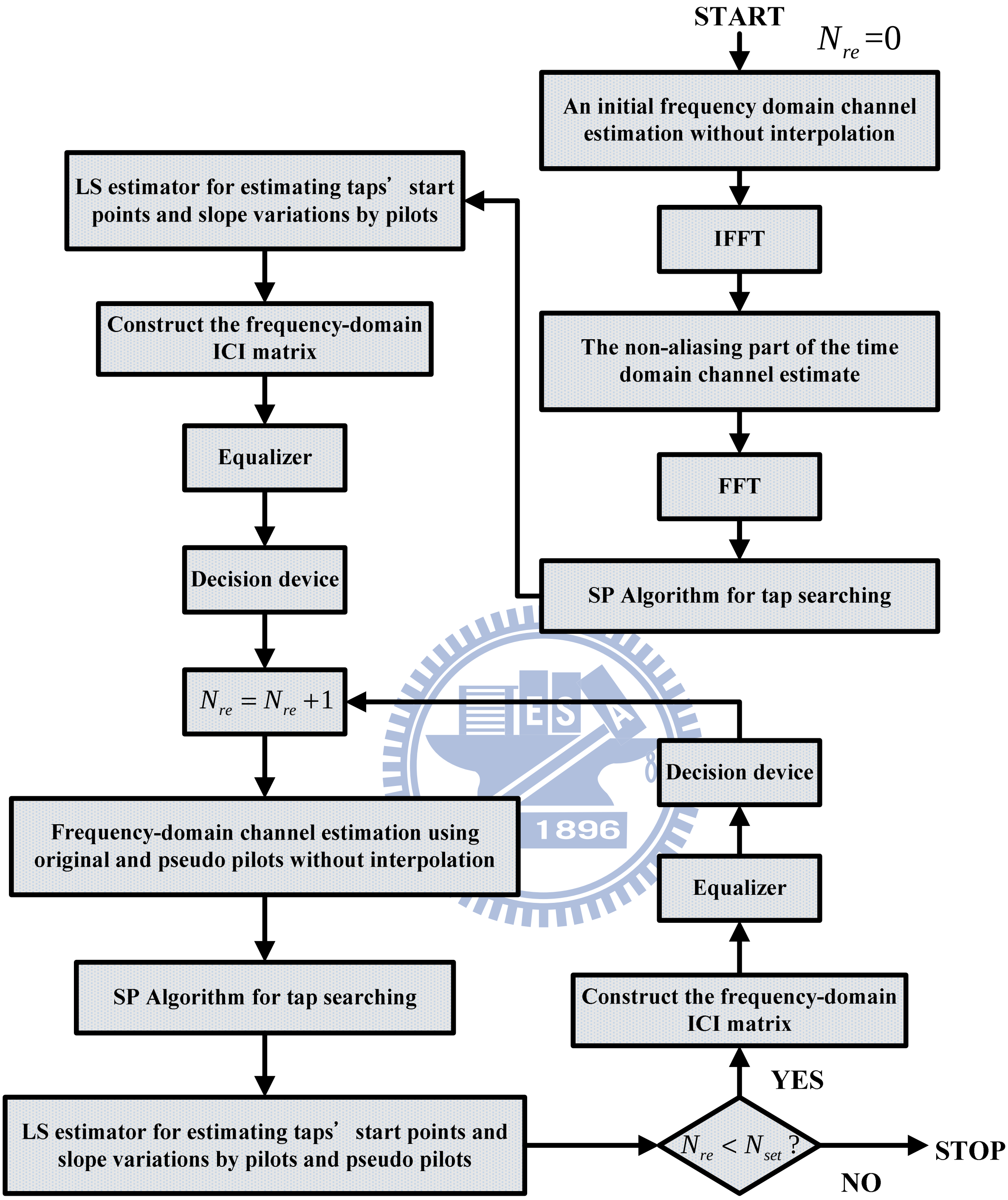

Figure 4- 11 Proposed method in time-variant channel.

1) Treat the channel as time-invariant and use the pilot subcarriers to obtain an initial frequency-domain channel estimate and transform it to the time-domain to obtain time-domain channel estimate.

2) Select the non-aliasing region of the estimated time-domain channel, and then transform it back to the frequency-domain to obtain a new frequency-domain channel estimate.
3) Conduct the SP algorithm for tap searching. Once the tap positions are located, the parameters including the start points and the variation slopes of the time-variant channel can then be obtained by the LS estimator described in Section 4.2.2.
4) Construct the ICI matrix $\tilde{\mathbf{M}}$ with the parameters estimated in 3), estimate the transmitted symbols by the zero-forcing equalizer $\left(\tilde{\mathbf{y}} = \tilde{\mathbf{M}}\tilde{\mathbf{x}} + \tilde{\mathbf{w}}\,,\ \hat{\tilde{\mathbf{x}}} = \tilde{\mathbf{M}}^{-1}\tilde{\mathbf{y}}\right)$, and make symbol decisions. Those at the designated subcarriers are then used as pseudo pilots.
5) Obtain the frequency-domain channel estimate again by using original and pseudo pilots. The whole channel response can be obtained by re-conducting the SP algorithm and the LS estimator.
6) If the number of re-estimation N_{re} is less than a preset value N_{set}, go to 2).

Chapter 5 Simulation Results

In this chapter, the simulation results are reported to demonstrate the performance of the proposed SP channel estimation algorithm. It is assumed that the synchronization of the receiver has been perfectly achieved. The OFDM system we used has 512 subcarriers and the CP length is 64, corresponding to $\frac{1}{8}$ of the symbol size. The pilot subcarriers are evenly allocated in the frequency-domain. For different simulation scenarios, the pilot density will be set to be different. Three modulation schemes including BPSK, QPSK, and 16-QAM are adopted for each subcarrier. Let the number of non-zero taps be 6, i.e., the sparsity of the channel is $K=6$. The tap values for the LTI channel are assumed to be independently identically-distributed (i.i.d.) with $CN(0,1)$, and the relative path power profiles are set as 0, -0.9, -4.9, -8, -7.8, and -23.9 (dB) as specified in ITU Ped-B channel [24]. And, the tap positions are uniformly distributed between 0 and $L-1$ where L is the maximum delay spread. An example of a 6-tap channel is shown in Figure 5- 1. In the time-varying system, the fading channel is generated by Jake's Model with various normalized Doppler frequency. The receive signal quality is indicated by the SNR defined as the received signal power divided by the noise power at the receiver. Also, the performance of the channel estimate is by the resultant bit-error-rate (BER).

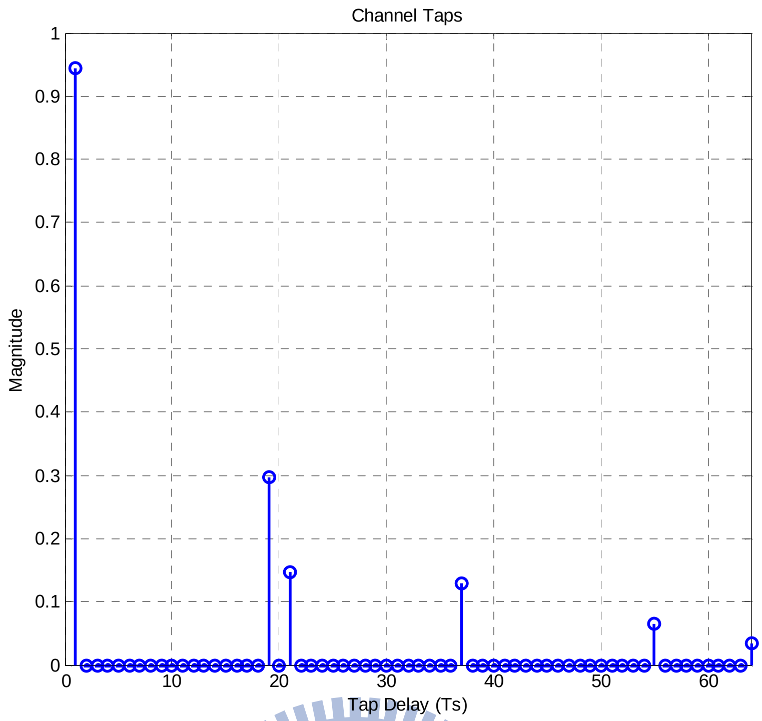

Figure 5- 1 An example of a 6-tap channel.

5.1 Results of LTI channel estimation

In this section, we report the performance of the proposed SP method in the LTI channel estimation problem. We first show that the proposed method can outperform other CS recovery methods. Then, we further explore the performance of the proposed channel estimator in different scenarios.

5.1.1 Performance of different CS methods in channel estimation

We compare the proposed method to the existing CS methods mentioned in Chapter 2, including LP, MP, and OMP algorithm. Figure 5- 2 shows the simulation results with the pilot density of 1/9 and the QPSK modulation scheme. As we can see, the performance of the SP algorithm is much better than MP and OMP, especially when the SNR is high. The performance of LP is very close to that of SP. However, the computational complexity of the LP algorithm is high.

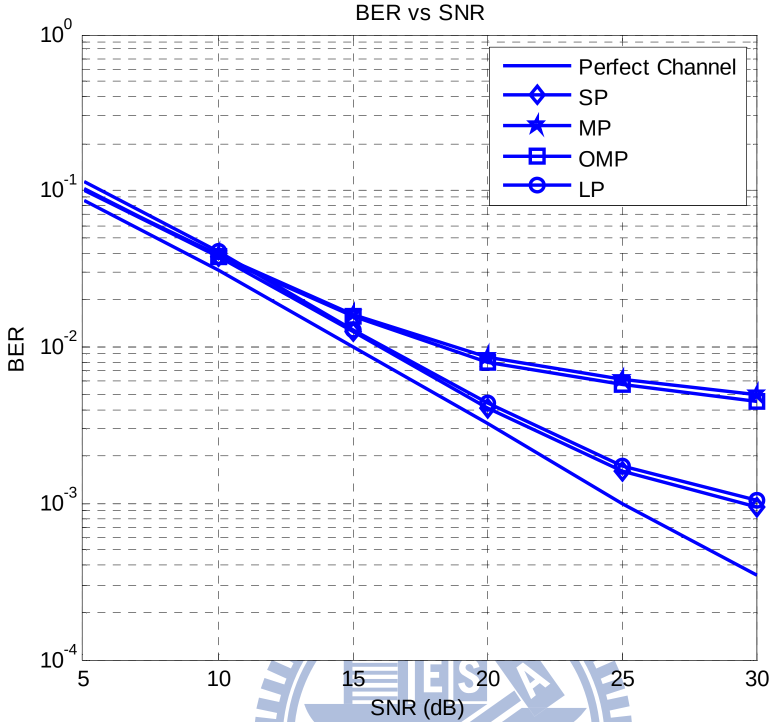

Figure 5- 2 Performance comparison of different CS recovery methods.

5.1.2 Simulation results of SP estimator with tap numbers are known

Figure 5- 3 and Figure 5- 4 show the estimation results for the pilot density of $\frac{1}{4}$ and $\frac{1}{8}$, respectively. Here, we assume that the number of taps is known as a prior. Comparing the results with perfect channel information denoted by BPSK perfect, QPSK perfect and 16-QAM perfect in Figure 5- 3, we see that the performance of the propose SP algorithm is good even for the 16-QAM scheme. When the pilot density is reduced, the performance, as shown in Figure 5- 4, is only slightly degraded. If this is not satisfactory, we can use the proposed decision-feedback algorithm to conduct the re-estimation and the performance can be further improved.

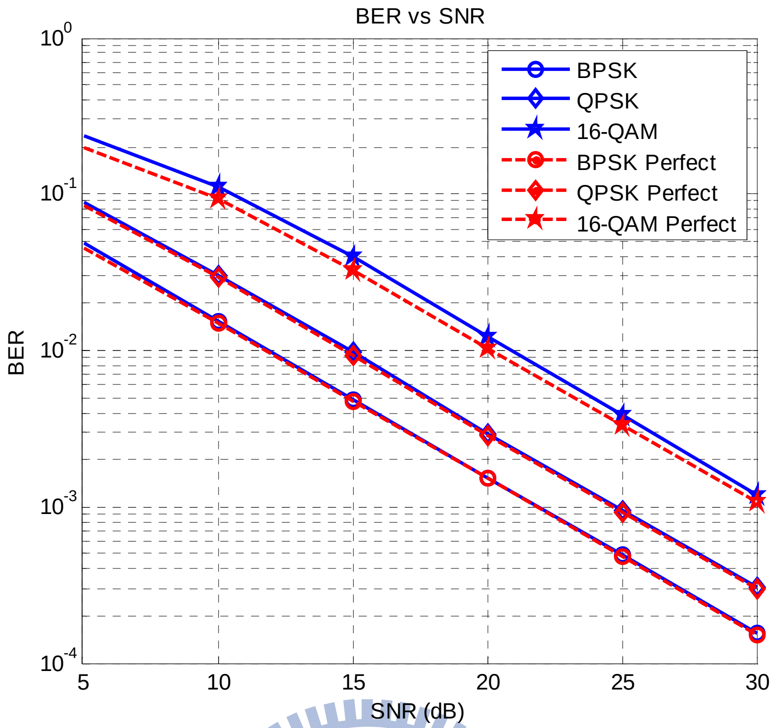

Figure 5- 3 BER performance of proposed channel estimator for BPSK, QPSK, and 16-QAM with pilot density of 1/4.

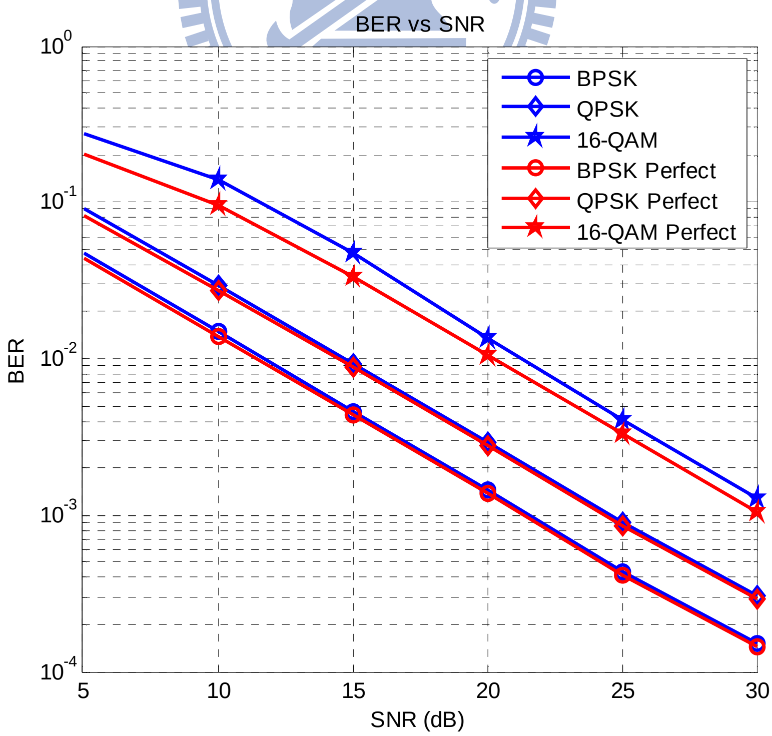

Figure 5- 4 BER performance of proposed channel estimator for BPSK, QPSK, and 16-QAM with pilot density of 1/8.

5.1.3 Simulation results of SP estimator with tap numbers are unknown

In this section, we will show the results of the proposed SP algorithm when the number of taps is unknown. As discussed in section 4.1.1, we propose two methods, specified in Figure 4- 2 and Figure 4- 3, to re-conduct SP algorithm when the tap number is unknown. Figure 5- 5 shows the iteration required for the SP algorithm in Figure 4- 2 and Figure 4- 3 vs. the expected channel tap. This figure clearly shows that the method depicted in Figure 4- 3 is much more efficient than the other one. Figure 5- 6 shows the BER performance with a pilot density of $\frac{1}{8}$. From Figure 5- 6, we see that the SP algorithm still works well even the information of the tap number is unknown.

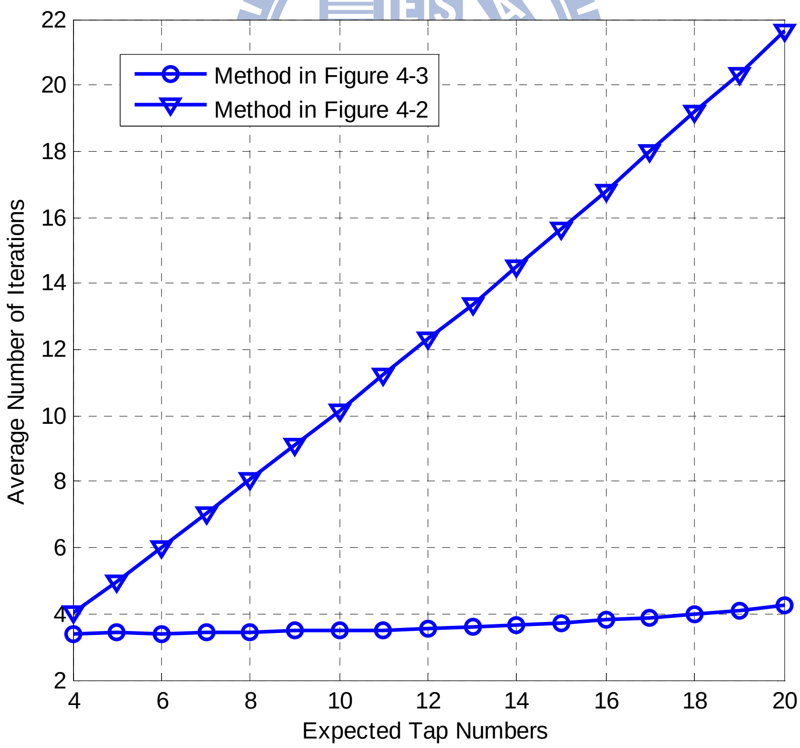

Figure 5- 5 Number of iterations required for SP re-conduction specified in Figure 4- 2 and Figure 4- 3.

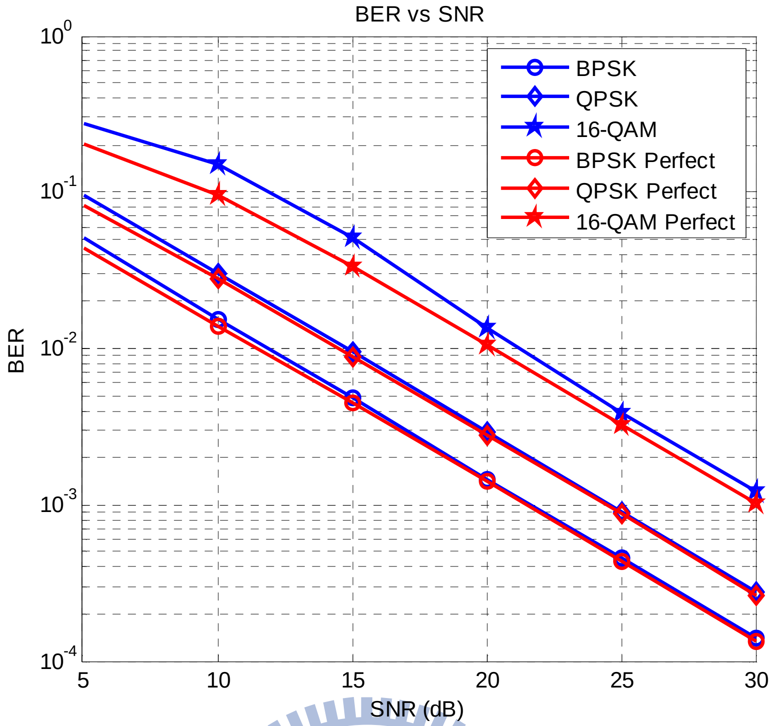

Figure 5- 6 Performance of proposed channel estimator with pilot density of 1/8 when tap numbers are unknown.

5.1.4 Simulation results of SP estimator with insufficient pilots

As discussed in Section 4.1.2, low pilot density may causes aliasing in the time-domain response. The proposed solution is first to deal with the response in the non-aliasing region and use it to conduct symbol detection and re-estimate the whole channel response. In this section, we use simulation results to show the number of the iterations required for the re-estimation in various modulation schemes and the resultant BER performance. Figure 5- 7 shows the performance vs. the number of iterations in the BPSK scheme with a pilot density of 1/12. From the figure, it is apparent that the performance for the second and third iteration is the same. Therefore, two iterations are sufficient for the proposed algorithm to obtain good performance. From Figure 5- 8 and Figure 5- 9, we can see that the numbers of iterations required for QPSK and 16-QAM are 2 and 4, respectively. Figure 5- 10 and Figure 5- 11 show the resultant BER performance for the cases with and without the known tap number, respectively. In Figure 5- 7 to Figure 5- 11, the pilot density (original pilots plus pseudo pilots) used for the re-estimation is set to 1/3.

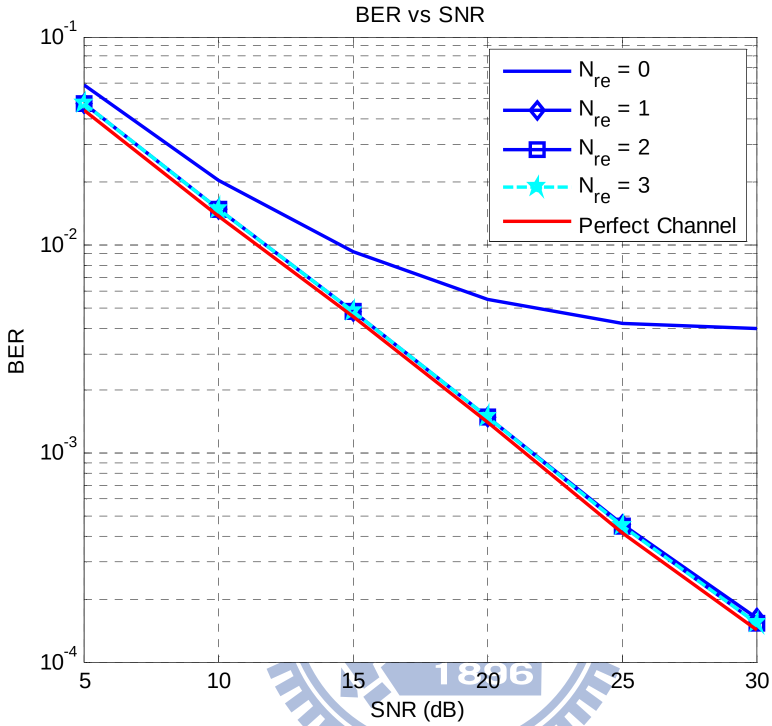

Figure 5- 7 Performance comparison of different number of iterations for channel re-estimation in BPSK.

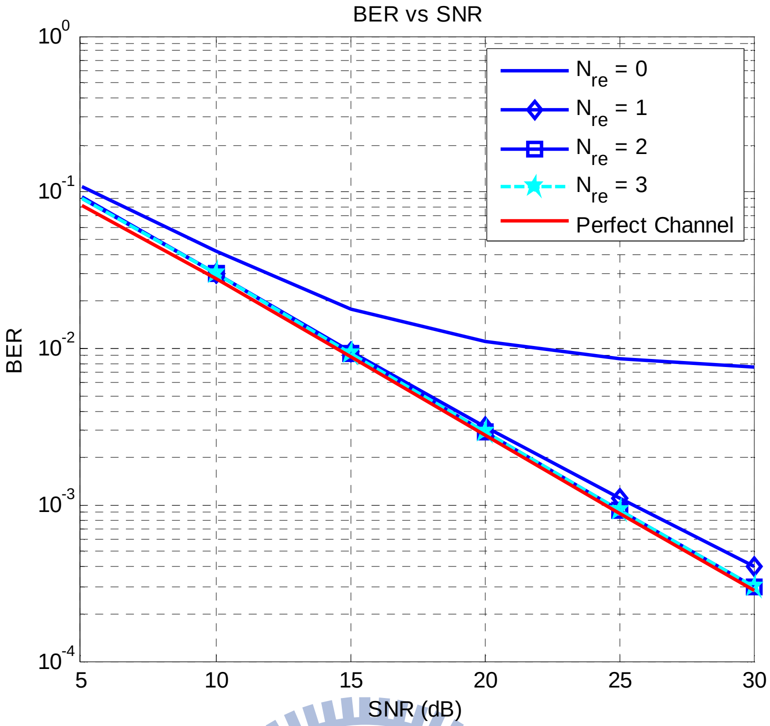

Figure 5- 8 Performance comparison of different number of iterations for channel re-estimation in QPSK.

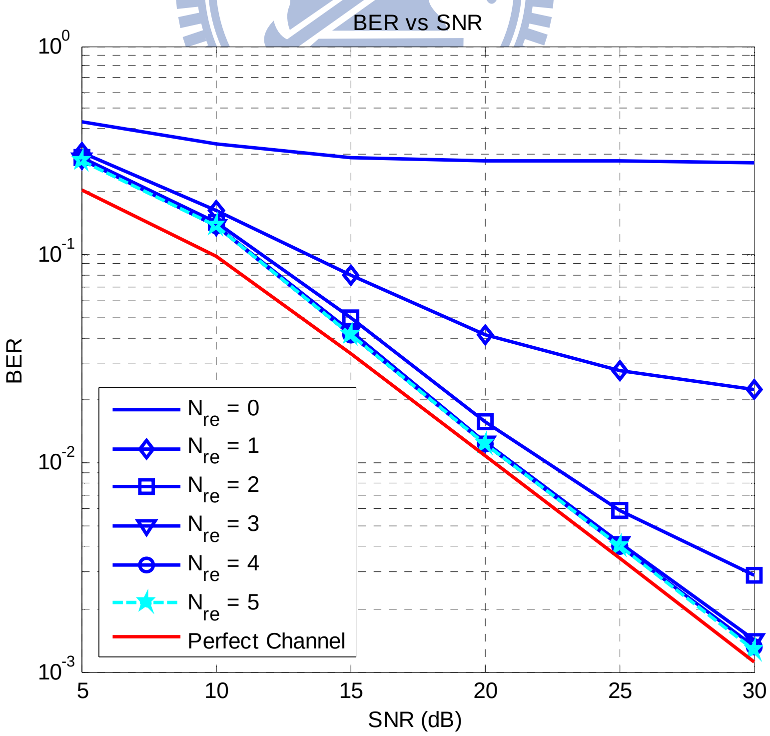

Figure 5- 9 Performance comparison of different number of iterations for channel re-estimation in 16-QAM.

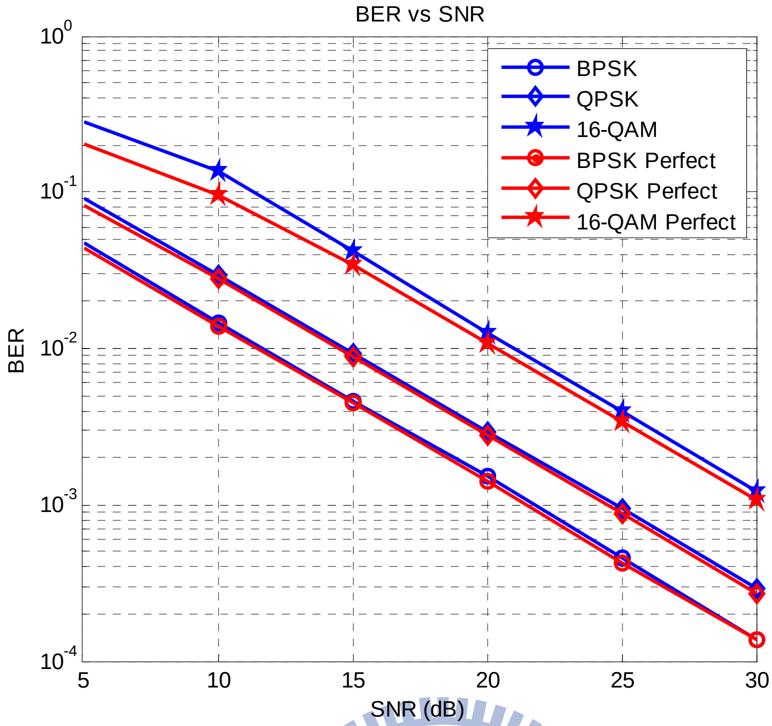

Figure 5- 10 BER performance of proposed channel estimator with pilot density of 1/12 when tap numbers are known.

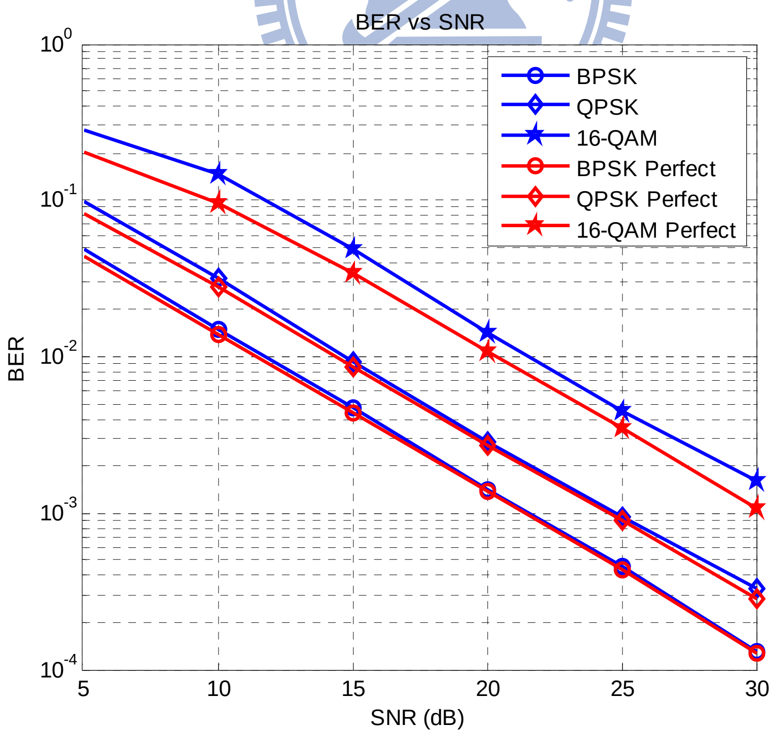

Figure 5- 11 BER performance of proposed channel estimator with pilot density of 1/12 when tap numbers are unknown.

5.2 Results of time-variant channel estimation

In the previous section, we have reported the performance of the proposed SP algorithm when the channel is assumed to be LTI. In this section, we consider the time-variant channel. Two scenarios will be discussed：The normalized Doppler frequency is 0.0244 and that is 0.1016 respectively.

5.2.1 Results of proposed time-variant channel estimator with normalized Doppler frequency of 0.0244

Figure 5- 12 shows the comparison of the BER performance of the proposed SP method with and without re-estimation procedure when the pilot density is $\frac{1}{8}$. From the figure, we see that the performance can indeed be improved with the re-estimation. Then, we show the simulation results for the cases when the tap number is known and unknown in Figure 5- 13 and Figure 5- 14, respectively. In the figures, the results for perfect channel estimation are also shown as the benchmarks. As we can see, the approximation errors of the proposed methods are small.

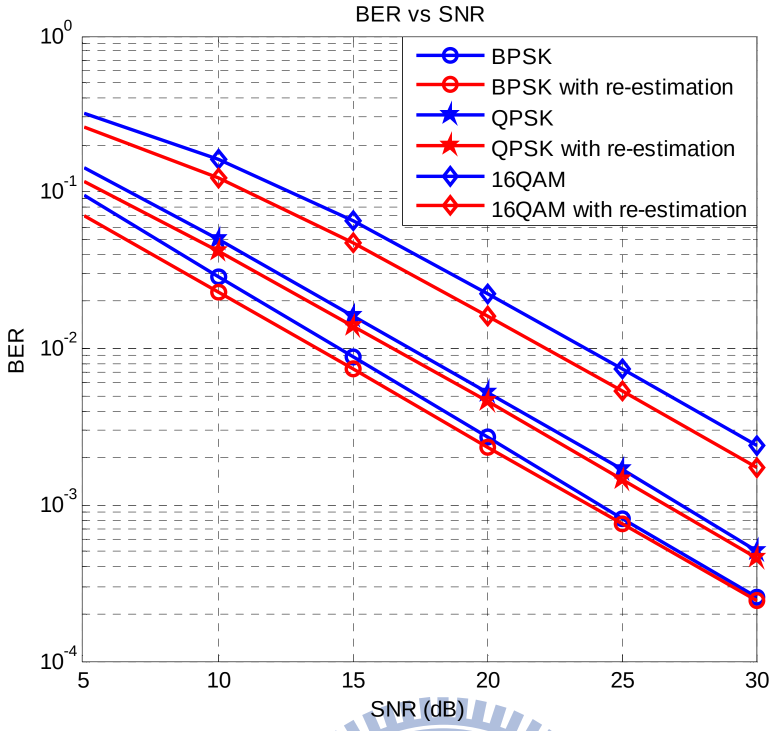

Figure 5- 12 Performance comparison of proposed channel estimator with and without re-estimation.

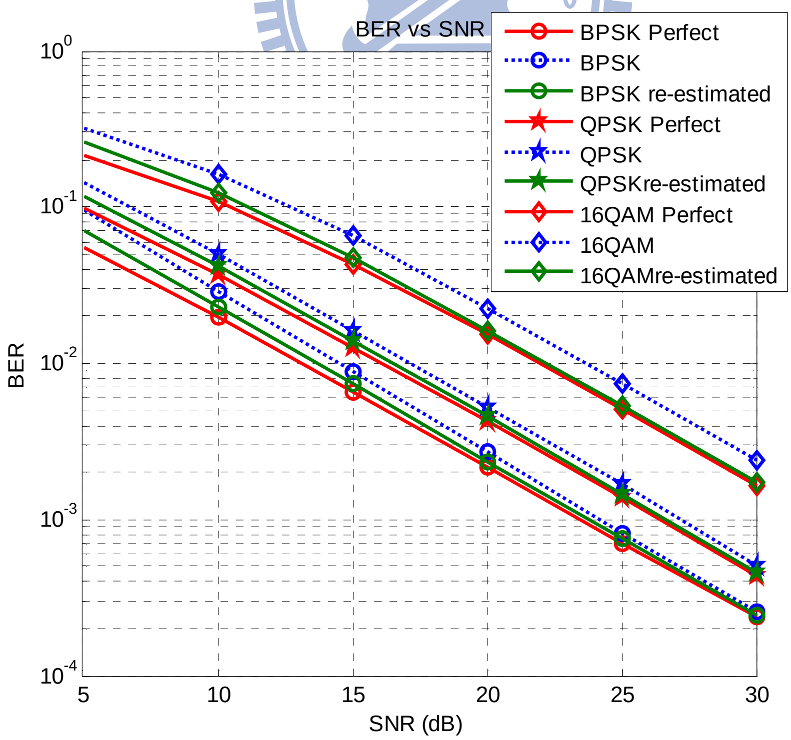

Figure 5- 13 Performance of proposed channel estimator with normalized Doppler frequency of 0.0244 for tap numbers are known.

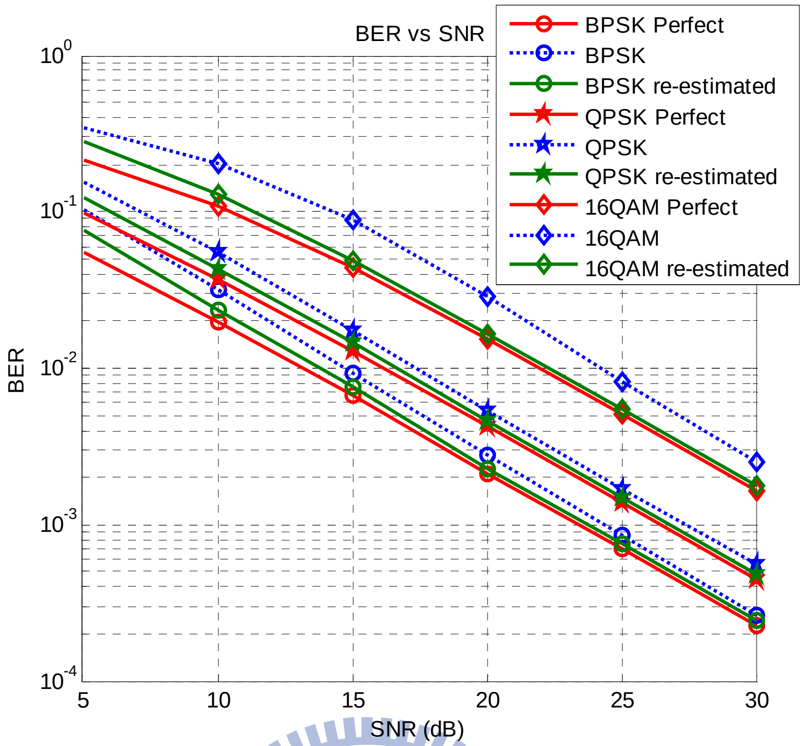

Figure 5- 14 Performance of proposed channel estimator with normalized Doppler frequency of 0.0244 for tap numbers are unknown.

5.2.2 Results of proposed time-variant channel estimator with normalized Doppler frequency of 0.1016

When the normalized Doppler frequency becomes 0.1016, the mobility speed becomes higher and the ICI effect is more severe. Figure 5- 15 and Figure 5- 16 show the simulation results. Notice that the error rate is raised because the ICI becomes larger. Nevertheless, by observing the two figures, we can see that the performance of the proposed SP algorithm still performs satisfactorily since the BER is close to that of the perfect channel.

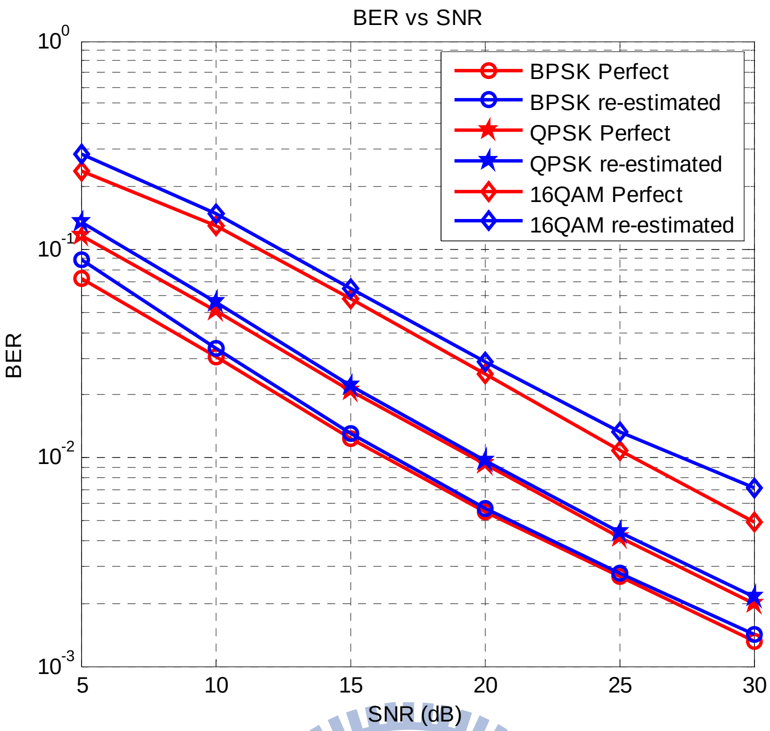

Figure 5- 15 Performance of proposed channel estimator with normalized Doppler frequency of 0.1016 for tap numbers are known.

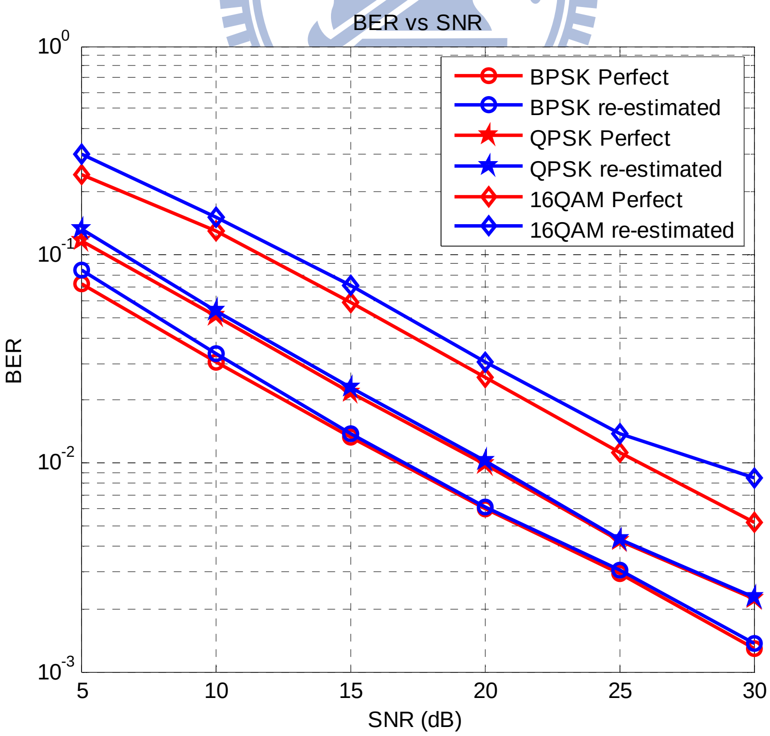

Figure 5- 16 Performance of proposed channel estimator with normalized Doppler frequency of 0.1016 for tap numbers are unknown.

Chapter 6 Conclusions and Future Works

In this thesis, we have applied the SP algorithm to the channel estimation problem in OFDM systems, where the channel impulse response is assumed to be sparse. Using simulation results, we first show that the SP algorithm is superior to the existing CS signal recovery methods. We then propose an iterative SP algorithm for the scenario that the number of channel taps is unknown. Simulation results show that the performance is close to the scenario that the number of channel taps is known. If the pilot density is low, aliasing will occur in the time-domain channel response and the performance of the SP method will be affected. We then further proposed a decision-feedback SP method in which some decisions are used as additional pseudo pilots to overcome the problem. Simulations show that the proposed decision-feedback SP algorithm still performs well even when the pilot density is very low. Finally, we apply the proposed SP channel estimators to time-variant channels. In the environments, ICI is introduced and the parameters to be estimated are doubled. And we show that the proposed method can yield good performance even when the mobile speed is high.

Through the entire thesis, we only focus on the single-input-single-output (SISO) systems. However, multiple-input-multiple-output (MIMO) systems are developed rapidly in recent years since they can provide higher data throughput, better coverage, and higher reliability. Thus, we may apply the proposed methods to the channel estimation problem in MIMO-OFDM systems. In equalization, we use a ZF equalizer to obtain the data decisions; however, there are many equalization methods which may provide better performance. This can also serve as a topic for further research.

References

[1] G. Auer and E. Karipidis, “Pilot aided channel estimation for OFDM: a separated approach for smoothing and interpolation” in *Proc. IEEE ICC* 2005.

[2] P. Y. Tsai and T. D. Chiueh, “Frequency-domain interpolation-based channel estimation in pilot-aided OFDM systems,” in *Proc. IEEE VTC 2004*.

[3] D. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp.1298-1306, Apr. 2006.

[4] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, March 2008.

[5] E. Candès and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.

[6] S.F. Cotter and B.D. Rao, “Matching pursuit based decision-feedback equalizers,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2000, vol. 5, pp. 2713–2716.

[7] Yang Liu and D.K. Borah, “Estimation of fading channels with large possible delay spreads,” *Electronic Letters*, vol. 39, no. 1, pp. 130–131, Jan. 2003.

[8] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Asilomar Conference on Signals, Systems and Computers*, 1993, vol. 1, pp. 40–44.

[9] G. Z. Karabulut, A. Yongacoglu, “Sparse cahnnel estimation using orthogonal matching pursuit algorithm” in *Proc. IEEE VTC-Fall 2004*.

[10] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230-2249, May 2009.

[11] M. Sharp and A. Scaglione, “Application of sparse signal recovery to pilot-assisted channel estimation,” in *Proc. IEEE ICASSP 2008*.

[12] E. J. Candès, “The restricted isometry property and its implications for compressive sensing,” *C. R. l’Academie des Sciences*, ser. I, no. 346, pp. 589-592, 2008.

[13] E. Candès, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug.2006.

[14] E. Candès, J. Romberg, and T. Tao, Robust Signal Recovery from Incomplete Observations," *Image Processing, 2006 IEEE International Conference on* ,pp.1281-1284, 8-11 Oct. 2006

[15] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “A metnod for large scale l_1- regulaized least sqaures,” *IEEE Trans. Selected Topics Signal Processing*, vol. 1, no. 4, pp. 606-617, Dec. 2007.

[16] S. Mallat and Z. Zhang, “Matching pursuit with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, Dec 1993.

[17] B. D. Rao, “Signal processing with the sparseness constraint,” in *Proc. IEEE ICASSP* 1998.

[18] A. Tropp, “Greed is good: algorithm results for spaese approximation,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp.2231-2242, Oct. 2004.

[19] D. Needell and R. Vershynin, “Signal Recovery From Incomplete and Inaccurate Measurements Via Regularized Orthogonal Matching Pursuit”, *Selected Topics in Signal Processing, IEEE Journal of* , vol.4, no.2, pp.310-316, April 2010

[20] E. Candès and T. Tao, “Near optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[21] M. Rudelson and R. Vershynin, “On sparse reconstruction from Fourier and Gaussian measurements,” submitted for publication.

[22] Rudelson, M., Vershynin, R., “Sparse reconstruction by convex relaxation: Fourier and Gaussian measurements. ” in *Proc. 2006 40th. Annu. Conf. Information Sciences and Systems*, , pp. 207–212, Mar. 2006.

[23] Y. Mostofi and D. C. Cox, “ICI Mitigation for Pilot-Aided OFDM Mobile Systems,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 765-774, Mar. 2005.

[24] T. B. Sorensen, P. E. Mogensen, and F. Frederiksen, “Extension of the ITU channel models for wideband (OFDM) systems,” in Proc. 2005 IEEE 62nd Vehicular Technology Conference, 2005.