國　立　交　通　大　學

電信工程學系碩士班

碩士論文

時空正交分頻多工調變之盲式通道估計

Blind Channel Estimation for Space-Time OFDM

研 究 生： 吳泰云

指導教授： 謝世福博士

中華民國九十三年七月

時空正交分頻多工調變之盲式通道估計

Blind Channel Estimation for Space-Time OFDM

研 究 生：吳泰云 Student：T. Y. Wu

指導教授：謝世福 博士 Advisor：Dr. S.F. Hsieh

國立交通大學

電信工程學系碩士班

碩士論文

A Thesis
Submitted to Department of Communication Engineering
College of Electrical Engineering and Computer Science
National Chiao Tung University
In Partial Fulfillment of the Requirements
For the Degree of
Master of Science
In
Electrical Engineering
July, 2004
Hsinchu, Taiwan, Republic of China

中華民國九十三年七月

時空正交分頻多工調變
之盲式通道估計

學生：吳泰云　　指導教授：謝世福

國立交通大學電信工程學系碩士班

摘要

時空正交分頻多工調變不僅可以達到高傳輸效率而且可以利用分集增益(diversity gain)，所以目前倍受矚目。在本篇論文中，我們採用了 Giannakis 所提出的時空正交分頻多工調變之盲式通道估計，並且為此估計子推導出均方誤差。不僅如此，我們更介紹了 decision direct (DD) 和 phase direct (PD) 兩個方法來使之前所提的通道估計更加趨於理想。 DD 是利用解得的決策信號來更新通道的估測，而 PD 是在我們得到通道的振幅響應之後，解得其相角響應，原來是只應用在一般的正交分頻多工上， 因為其通道的振幅響應很容易得到，但是在時空正交分頻多工調變上，因為接收到的信號是由兩個傳送天線傳來的，通道的振幅響應很難由其中解出。所以我們提出了一和差平方的演算法來的到通道的振幅響應進而就可以用 PD 來使通道的估計更趨於理想。並且在電腦模擬中驗證了我們的方法真的使通道的估計更完美。

Blind Channel Estimation for Space-Time OFDM

Student: T. Y. Wu Advisor: S. F. Hseieh

Department of Communication Engineering

National Chiao Tung University

Abstract

Space time (ST) orthogonal frequency division multiplexing (OFDM) has been well documented as an attractive means of achieving high data rate transmissions with diversity gains. In this thesis, we adopt a blind channel estimation algorithm proposed by Giannakis for ST OFDM, and derive the theoretical mean square error of the estimator. Moreover, we introduce phase direct (PD) and decision direct (DD) methods to further improve the performance of the estimator. DD and PD originally work on conventional OFDM, and PD is not suited for ST OFDM. Then we derive a new algorithm named sum-difference square method to make PD work on ST OFDM. DD is to update our estimated channel from the previous hard decision data, while PD is to solve the phase ambiguities after we've got the channel power response. Since the received data in ST OFDM is composed of two different transmitted data, the channel amplitude response is not easy to get. Hence, the aforementioned algorithm is about how to solve this problem. Furthermore, in computer simulations, we can see our algorithm really better the performance.

Acknowledgement

I would like to express my deepest gratitude to Dr. S. F. Hsieh for his enthusiastic guidance and great patient. I also wish to thank my friends for their concern and help. Finally, I would like to thank my parents, family about their consideration and encouragement.

Contents

List of Figures

Chapter 1

Introduction

In orthogonal frequency division multiplexing (OFDM) [1,2,3], the entire channel is divided into many narrow parallel subchannels, thereby increasing the symbol duration and reducing or eliminating the intersymbol interference (ISI) caused by the multipath environments. On the other hand, since the dispersive property of wireless channels causes frequency selective fading, there is higher error probability for those subchannels in deep fades. Hence, techniques such as error correction code and diversity have to be used to compensate for the frequency selectivity. In this thesis, we investigate transmitter diversity using space-time coding for OFDM systems.

Space-time codes (STC) [4,5,6,7] realize the diversity gains by introducing temporal and spatial correlation into the signals transmitted from different antennas without increasing the total transmitted power or transmission bandwidth. There is in fact a diversity gain that results from multiple paths between base station and user terminal, and a coding gain that results from how symbols are correlated across transmit antennas.

Transmitter diversity is an effective technique for combating fading in mobile wireless communications, especially when receiver diversity is expensive or impractical. Many researchers [8,9,10] have studied transmitter diversity for wireless systems. In this thesis, we focus on two transmit-antennas and one receive-antenna and use the well known Alamouti's block STC [11].

However, for most STC transceivers, *multichannel* estimation algorithms are needed. Training symbols are transmitted periodically in [12] for the receiver to acquire the multi-input multi-output (MIMO) frequency-flat channels (see also [13] for training-based estimation of frequency-selective channels in ST-OFDM). However, training sequences consume bandwidth and, thereby, incur spectral efficiency (and thus capacity) loss. For this reason, blind channel estimation methods receive growing attention, especially for estimating the MIMO channels corresponding to multiple transmit and receive antennas.

Only a few works, however, have been reported so far on blind MIMO and multi-input single-output (MISO) channel estimation that exploits the unique features of ST codes. Relying on nonredundant and nonconstant modulus precoding, blind channel estimation and equalization for OFDM-based multi-antenna systems has been proposed in [14] using cyclostationary statistics. Subspace-based blind method is proposed in [15] for estimating the channel responses of a multiuser and multiantenna OFDM uplink system. For ST-OFDM, a deterministic blind channel estimator was derived in [16] when the channel transfer functions are coprime (no common zeros) and the transmitted signals have constant-modulus (CM).

In this thesis, we deal with a linearly precoded ST-OFDM system with two transmit antennas and show a blind channel identification algorithm [17] for frequency-selective FIR channels through the subspace method. Moreover, as shown

in [17] with properly designed redundant precoders, the subspace-based method can estimate multiple channels simultaneously up to one scalar ambiguity.

Furthermore, based on the first-order perturbation theory [19], we also derive the theoretical mean square error of the estimator which shows the relationship with the simulation result.

To further improve the channel estimation, we can exploit the finite alphabet property to better the subspace-based channel estimates by applying the "Decision direct (DD)" and "Phase direct (PD)" methods. DD, as implied in the name, needs first to get the hard decision data and then use it to update our estimated channel, while PD is to solve the phase ambiguities after we've got the channel power response. DD originally works in conventional OFDM [21], which only requires simple scalar division. Based on the space time data matrix, we extend it to ST-OFDM, which corresponds to a matrix inverse and multiplication because the received data is composed of two different transmitted data.

The main idea of PD is to solve the phase ambiguities after we get the channel power response. For conventional OFDM system, it is very easy to get the channel power response. But in space-time OFDM, it is quite a different case, since the received data is composed of two different transmitted data; it's not easy to separate them. So, the main problem we face now is how to get the channel power response, which is hard to get in general. Hence, we only focus on BPSK system and exploit the transmitted data's time and temporal correlation to develop a new algorithm named sum-difference square method to solve this problem.

Moreover, in time varying channel, we also need to choose a best window size to get the channel power response and apply it to PD. As we all know, when the window is longer, we can suppress the noise, but then we can't follow the variation of the

channel. This is the trade off. However, the choice of the window size is dependent on how fast the channel changes. Precoder design is another issue behind the algorithm as will be discussed in chapter 4.

The rest of this paper is organized as follows. After presenting the system model in Chapter 2, we show our blind channel estimation algorithm in Chapter 3 and further improved methods in Chapter 4. Chapter 5 presents simulation results, and Chapter 6 gathers our conclusions.

Chapter 2

Space Time OFDM System Model

Fig. 2.1 depicts the space-time OFDM system considered in this thesis with two transmit antennas (there can be more antennas, but in this thesis we focus on two antennas) and one receive-antenna. Prior to transmission, the information bearing symbols $\underline{s}(n)$ are first grouped into super blocks of size $2K \times 1$, where we indicate the first K symbols as $\underline{s}^{(1)}(n)$ and last K symbols as $\underline{s}^{(2)}(n)$.

$$\underline{s}(n) = \begin{bmatrix} \underline{s}^{(1)}(n) \\ \underline{s}^{(2)}(n) \end{bmatrix} \tag{2.1}$$

Two different linear block precoders denoted by the tall $M \times K$ matrices $\mathbf{?}_1$ and $\mathbf{?}_2$ (one for $\underline{s}^{(1)}(n)$ and the other for $\underline{s}^{(2)}(n)$) are used to introduce redundancy ($M > K$). The corresponding $2M \times 1$ precoded block is

$$\begin{aligned} \underline{\tilde{s}}(n) &= \begin{bmatrix} \underline{\tilde{s}}^{(1)}(n) \\ \underline{\tilde{s}}^{(2)}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{?}_1 \underline{s}^{(1)}(n) \\ \mathbf{?}_2 \underline{s}^{(2)}(n) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{?}_2 \end{bmatrix} \begin{bmatrix} \underline{s}^{(1)}(n) \\ \underline{s}^{(2)}(n) \end{bmatrix} = \mathbf{T}\, \underline{s}(n) \end{aligned} \tag{2.2}$$

where

$$\mathbf{T} = \begin{bmatrix} \mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{?}_2 \end{bmatrix} \tag{2.3}$$

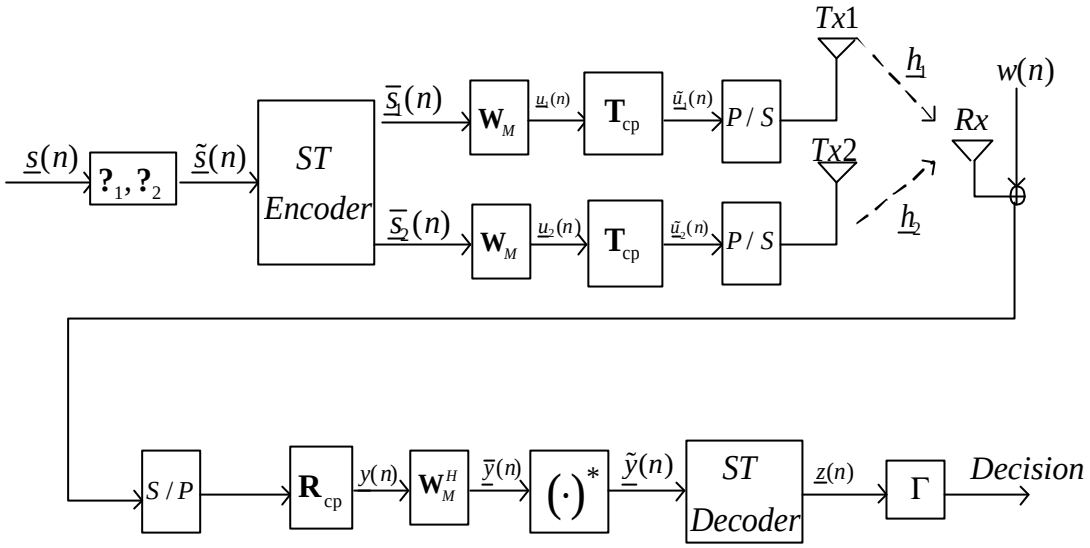


Fig. 2.1 Block precoded ST-OFDM transceiver model

$\underline{\tilde{s}}(n)$ is then fed to the space-time encoder. The encoder takes input two consecutive precoded blocks $\underline{\tilde{s}}^{(1)}(n)$ and $\underline{\tilde{s}}^{(2)}(n)$ to output the following $2M \times 2$ code matrix:

$$\left[\underline{\bar{s}}_1(n) \quad \underline{\bar{s}}_2(n)\right] = \begin{bmatrix} \underline{\bar{s}}_1^{(1)}(n) & \underline{\bar{s}}_2^{(1)}(n) \\ \underline{\bar{s}}_1^{(2)}(n) & \underline{\bar{s}}_2^{(2)}(n) \end{bmatrix} = \begin{bmatrix} \underline{\tilde{s}}^{(1)}(n) & \underline{\tilde{s}}^{(2)}(n) \\ -\underline{\tilde{s}}^{(2)^*}(n) & \underline{\tilde{s}}^{(1)^*}(n) \end{bmatrix} \tag{2.4}$$

where

$$\underline{\bar{s}}_1(n) = \begin{bmatrix} \underline{\bar{s}}_1^{(1)}(n) \\ \underline{\bar{s}}_1^{(2)}(n) \end{bmatrix} \quad \text{and} \quad \underline{\bar{s}}_2(n) = \begin{bmatrix} \underline{\bar{s}}_2^{(1)}(n) \\ \underline{\bar{s}}_2^{(2)}(n) \end{bmatrix} \tag{2.5}$$

Eq. (2.3) shows that the blocks $\underline{\tilde{s}}(n)$ is transmitted twice in two consecutive time intervals through two different channels.

Here we define a $M \times M$ IDFT matrix as

$$\mathbf{W}_M = \frac{1}{\sqrt{\mathrm{M}}}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & w_M^{-1} & \cdots & w_M^{-(M-1)} \\ 1 & w_M^{-2} & \cdots & w_M^{-2(M-1)} \\ \vdots & & & \vdots \\ 1 & w_M^{-(M-1)} & \cdots & w_M^{-(M-1)(M-1)} \end{bmatrix} \tag{2.6}$$

with $w_M = e^{j2p/M}$.

Then the OFDM symbol $\underline{s}_i^{(1)}(n)$ and $\underline{s}_i^{(2)}(n)$ of dimension M are then modulated by the IDFT matrix $\mathbf{W}_M$ to produce

$$\underline{u}_i(n) = \begin{bmatrix} \mathbf{W}_M \underline{s}_i^{(1)}(n) \\ \mathbf{W}_M \underline{s}_i^{(2)}(n) \end{bmatrix} \tag{2.7}$$

In order to eliminate inter-block-interference (IBI) caused by the channel, the "time domain" vector $\underline{u}_i(n)$ is then enlarged by a cyclic prefix (CP) of length L, resulting in a size $M+L$ vector, where CP insertion replicates the last L elements of the IDFT output vector in the front since we assume that the channel order is equal or less than L, as will be discussed latter. As shown in Fig. 2.1, the CP insertion can be described by $\mathbf{T}_{cp} = [\mathbf{I}_{cp}^T \ \mathbf{I}_M^T]^T$, where $\mathbf{I}_{cp}$ is formed by the last L rows of the $M \times M$ identity matrix $\mathbf{I}_{cp}$, and its output

$$\tilde{\underline{u}}_i(n) = \begin{bmatrix} \mathbf{T}_{cp}\mathbf{W}_M \underline{s}_i^{(1)}(n) \\ \mathbf{T}_{cp}\mathbf{W}_M \underline{s}_i^{(2)}(n) \end{bmatrix} \tag{2.8}$$

is finally sent sequentially through transmit antenna i.

We assume in what follows that the channels between the two transmit antennas and the receive antenna are frequency selective and that their baseband equivalent effect in discrete time is captured by an FIR linear time-invariant filter with impulse response vector

$$\underline{h}_i = [h_i(0), \ldots, h_i(L)] \quad i = 1,2 \tag{2.9}$$

where L is an upper bound for the channel orders of $\underline{h}_1$ and $\underline{h}_2$, i.e., $L \geq \max(L_1, L_2)$ if L_1 is the channel order for $\underline{h}_1$ and L_2 for $\underline{h}_2$.

Accordingly, the FIR channel can be described by the $(M+L)\times(M+L)$ Toeplitz matrix $\mathbf{H}_i$ with *(k,l)*th entry $h_i(k-l)$.

$$\mathbf{H}_i = \begin{bmatrix} h_i(0) & & & & & & \\ h_i(1) & h_i(0) & & & & & \\ h_i(2) & h_i(1) & h_i(0) & & & & \\ & & \ddots & & & & \\ & & & h_i(L) & \cdots & h_i(0) & \\ & & & & h_i(L) & \cdots & h_i(0) \end{bmatrix} \tag{2.10}$$

At the receiver end, the CP is simply removed. The operation of discarding the first L received symbols can be described by the matrix $\mathbf{R}_{cp} = [\mathbf{0}_{M\times L} \ \mathbf{I}_M]$. Let

$$\tilde{\mathbf{H}}_i = \mathbf{R}_{cp}\mathbf{H}_i\mathbf{T}_{cp} \tag{2.11}$$

denote the equivalent channel matrix after eliminating the IBI. The $2J\times 1$ IBI-free block $\underline{y}(n)$ in Fig. 2.1 is given by:

$$\begin{aligned} \underline{y}(n) &= \begin{bmatrix} \underline{y}^{(1)}(n) \\ \underline{y}^{(2)}(n) \end{bmatrix} = \mathbf{R}_{cp}\mathbf{H}_1\underline{\tilde{u}}_1(n) + \mathbf{R}_{cp}\mathbf{H}_2\underline{\tilde{u}}_2(n) + \mathbf{R}_{cp}\underline{w}(n) \\ &= \begin{bmatrix} \mathbf{R}_{cp}\mathbf{H}_1\mathbf{T}_{cp}\mathbf{W}_M\underline{S}_1^{(1)}(n) \\ \mathbf{R}_{cp}\mathbf{H}_1\mathbf{T}_{cp}\mathbf{W}_M\underline{S}_1^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{R}_{cp}\mathbf{H}_2\mathbf{T}_{cp}\mathbf{W}_M\underline{S}_2^{(1)}(n) \\ \mathbf{R}_{cp}\mathbf{H}_2\mathbf{T}_{cp}\mathbf{W}_M\underline{S}_2^{(2)}(n) \end{bmatrix} + \mathbf{R}_{cp}\underline{w}(n) \\ &= \begin{bmatrix} \tilde{\mathbf{H}}_1\mathbf{W}_M\underline{S}_1^{(1)}(n) \\ \tilde{\mathbf{H}}_1\mathbf{W}_M\underline{S}_1^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{H}}_2\mathbf{W}_M\underline{S}_2^{(1)}(n) \\ \tilde{\mathbf{H}}_2\mathbf{W}_M\underline{S}_2^{(2)}(n) \end{bmatrix} + \mathbf{R}_{cp}\underline{w}(n) \end{aligned} \tag{2.12}$$

where *w(n)* is the additive white Gaussian noise vector. Given $\underline{y}(n)$, the retrieval of the information blocks $\underline{s}(n)$ at the receiver proceeds in three steps as follows. First, the DFT matrix $\mathbf{W}_M^H$ is performed on $\underline{y}(n)$ to obtain $\underline{\bar{y}}(n)$,

$$\begin{aligned}\underline{\bar{y}}(n) &= \begin{bmatrix}\underline{\bar{y}}^{(1)}(n)\\ \underline{\bar{y}}^{(2)}(n)\end{bmatrix} = \begin{bmatrix}\mathbf{W}_M^H \underline{y}^{(1)}(n)\\ \mathbf{W}_M^H \underline{y}^{(2)}(n)\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{W}_M^H & 0\\ 0 & \mathbf{W}_M^H\end{bmatrix}\begin{bmatrix}\underline{y}^{(1)}(n)\\ \underline{y}^{(1)}(n)\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{W}_M^H\tilde{\mathbf{H}}_1\mathbf{W}_M\underline{\bar{s}}_1^{(1)}(n)\\ \mathbf{W}_M^H\tilde{\mathbf{H}}_1\mathbf{W}_M\underline{\bar{s}}_1^{(2)}(n)\end{bmatrix}+\begin{bmatrix}\mathbf{W}_M^H\tilde{\mathbf{H}}_2\mathbf{W}_M\underline{\bar{s}}_2^{(1)}(n)\\ \mathbf{W}_M^H\tilde{\mathbf{H}}_2\mathbf{W}_M\underline{\bar{s}}_2^{(2)}(n)\end{bmatrix}+\begin{bmatrix}\mathbf{W}_M^H & 0\\ 0 & \mathbf{W}_M^H\end{bmatrix}\mathbf{R}_{cp}\underline{w}(n)\end{aligned} \tag{2.13}$$

By defining $\underline{\tilde{y}}(n) = \begin{bmatrix}\underline{\bar{y}}^{(1)}(n)\\ \underline{\bar{y}}^{(2)^*}(n)\end{bmatrix}$, then $\underline{\tilde{y}}(n)$ is processed by the space-time decoder to produce the block $\underline{z}(n)$ with diversity gains. Finally, the equalizer Γ is employed to recover $\underline{s}(n)$. Then we can plot a frequency domain version of Fig. 2.1 in Fig.2.2.

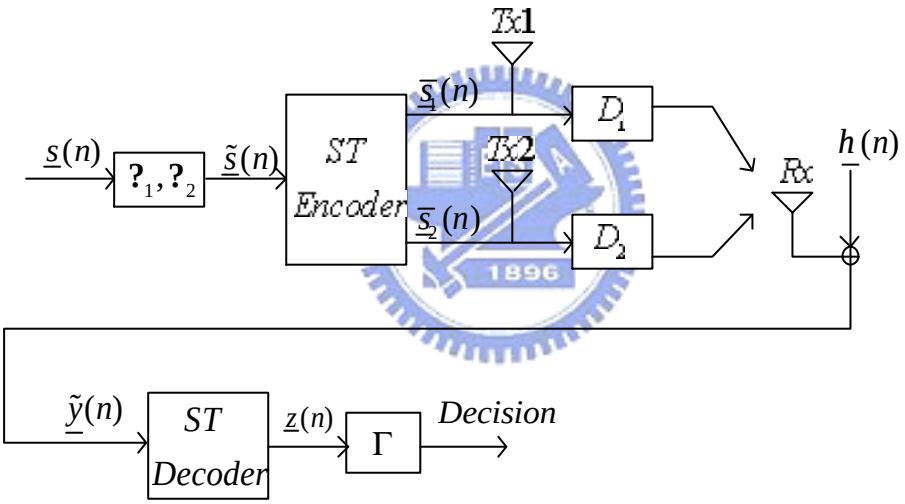


Fig. 2.2 Frequency domain version of ST-OFDM transceiver

Here, we want to start with the following fact [18]:

Fact 1: The matrix $\tilde{\mathbf{H}}_i$ can be diagonalized by pre- and post-multiplication with $\mathbf{W}_M^H$ and $\mathbf{W}_M$, i.e.,

$$\mathbf{W}_M^H\tilde{\mathbf{H}}_i\mathbf{W}_M = \mathbf{D}(\underline{H}_i), \tag{2.14}$$

where $\mathbf{D}(\underline{H}_i)$ stands for the diagnal matrix with the vector $\underline{H}_i$ on its diagonal, and $H_i(r)=\sum_{l=0}^{L} h_i(l)\, r^{-l}$ is the frequency response of channel *h(l)* at point r which is different from the Toeplitz matrix $\mathbf{H}_i$ in Eq.(2.10). Moreover, we can write

$$\underline{H}_i = \mathbf{V}\underline{h}_i = [H_i(e^{j0}),..., H_i(e^{j2p(M-1)/M})]^T \tag{2.15}$$

with $\mathbf{V}$ denoting submatrix consisting of the first *L+1* column of $\mathbf{W}_M^H$. ?

Exploiting *Facts 1* and substitute Eq.(2.4), the vector $\underline{\bar{y}}(n)$ in Eq.(2.13) can be written as:

$$\begin{aligned}
\underline{\bar{y}}(n) &= \begin{bmatrix} \mathbf{W}_M^H \tilde{\mathbf{H}}_1 \mathbf{W}_M \underline{\bar{s}}_1^{(1)}(n) \\ \mathbf{W}_M^H \tilde{\mathbf{H}}_1 \mathbf{W}_M \underline{\bar{s}}_1^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{W}_M^H \tilde{\mathbf{H}}_2 \mathbf{W}_M \underline{\bar{s}}_2^{(1)}(n) \\ \mathbf{W}_M^H \tilde{\mathbf{H}}_2 \mathbf{W}_M \underline{\bar{s}}_2^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{W}_M^H & 0 \\ 0 & \mathbf{W}_M^H \end{bmatrix} \mathbf{R}_{cp} \underline{w}(n) \\
&= \begin{bmatrix} \mathbf{D}(\underline{H}_1)\underline{\bar{s}}_1^{(1)}(n) \\ \mathbf{D}(\underline{H}_1)\underline{\bar{s}}_1^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{D}(\underline{H}_2)\underline{\bar{s}}_2^{(1)}(n) \\ \mathbf{D}(\underline{H}_2)\underline{\bar{s}}_2^{(2)}(n) \end{bmatrix} + \underline{v}(n) \\
&= \begin{bmatrix} \mathbf{D}(\underline{H}_1)\underline{\tilde{s}}^{(1)}(n) + \mathbf{D}(\underline{H}_2)\underline{\tilde{s}}^{(2)}(n) \\ -\mathbf{D}(\underline{H}_1)\underline{\tilde{s}}^{(2)^*}(n) + \mathbf{D}(\underline{H}_2)\underline{\tilde{s}}^{(1)^*}(n) \end{bmatrix} + \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)}(n) \end{bmatrix}
\end{aligned} \tag{2.16}$$

where

$$\underline{v}(n) = \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{W}_M^H & 0 \\ 0 & \mathbf{W}_M^H \end{bmatrix} \mathbf{R}_{cp} \underline{w}(n) \tag{2.17}$$

Then,

$$
\begin{aligned}
\underline{\tilde{y}}(n) &= \begin{bmatrix} \underline{\bar{y}}^{(1)}(n) \\ \underline{\bar{y}}^{(2)^*}(n) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{D}(\underline{H}_1)\underline{\tilde{s}}^{(1)}(n) + \mathbf{D}(\underline{H}_2)\underline{\tilde{s}}^{(2)}(n) \\ -\mathbf{D}(\underline{H}_1^*)\underline{\tilde{s}}^{(2)}(n) + \mathbf{D}(\underline{H}_2^*)\underline{\tilde{s}}^{(1)}(n) \end{bmatrix} + \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)^*}(n) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & -\mathbf{D}(\underline{H}_1^*) \end{bmatrix} \begin{bmatrix} \underline{\tilde{s}}^{(1)}(n) \\ \underline{\tilde{s}}^{(2)}(n) \end{bmatrix} + \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)^*}(n) \end{bmatrix} \\
&= \bar{\mathbf{D}}\Theta\underline{s}(n) + \underline{h}(n) \\
&= \mathbf{A}\underline{s}(n) + \underline{h}(n) \\
&= \underline{x}(n) + \underline{h}(n)
\end{aligned}
\tag{2.18}
$$

where $\mathbf{H}$, $\underline{x}(n)$ and $\underline{h}(n)$ are defined, respectively, as

$$
\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & \text{-}\mathbf{D}(\underline{H}_1^*) \end{bmatrix}, \quad \underline{h}(n) = \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)^*}(n) \end{bmatrix} \tag{2.19}
$$

$$
\mathbf{A} = \bar{\mathbf{D}}\Theta, \ \underline{x}(n) = \mathbf{A}\underline{s}(n), \tag{2.20}
$$

When the channel matrices $\underline{H}_1$ and $\underline{H}_2$ become available at the receiver, it is possible to demodulate $\underline{\tilde{y}}(n)$ with diversity gains by a simple matrix multiplication

$$
\begin{aligned}
\underline{z}(n) &= \bar{\mathbf{D}}^H \underline{\tilde{y}}(n) \\
&= \begin{bmatrix} \mathbf{D}(\underline{H}_1^*) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & \text{-}\mathbf{D}(\underline{H}_1) \end{bmatrix} \begin{bmatrix} \mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & \text{-}\mathbf{D}(\underline{H}_1^*) \end{bmatrix} \Theta\underline{s}(n) + \bar{\mathbf{D}}^H \underline{h}(n) \\
&= \begin{bmatrix} \mathbf{D}_{12}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12}\mathbf{?}_2 \end{bmatrix} \underline{s}(n) + \underline{x}(n)
\end{aligned}
\tag{2.21}
$$

where

$$
\mathbf{D}_{12} = \mathbf{D}(\underline{H}_1^*)\mathbf{D}(\underline{H}_1) + \mathbf{D}(\underline{H}_2^*)\mathbf{D}(\underline{H}_2) \tag{2.22}
$$

$$
\underline{x}(n) = \bar{\mathbf{D}}^H \underline{h}(n) \tag{2.23}
$$

We infer that multiantenna diversity of order two has been achieved because

$$
\mathbf{D}_{12} = diag\left(\sum_{i=1}^{2} |H_i(0)|^2, \ldots, \sum_{i=1}^{2} |H_i(M-1)|^2 \right) \tag{2.24}
$$

Eq. (2.19) implies that zeros-forcing recovery of $\underline{s}(n)$ from $\underline{z}(n)$ requires that the

inverse of $\begin{bmatrix} \mathbf{D}_{12}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12}\mathbf{?}_2 \end{bmatrix}$, which means the matrices $\mathbf{D}_{12}\mathbf{?}_i$, $i \in [1,2]$ need to be full column rank. We can adopt the following design condition on the block lengths and the linear precoders to make sure the full rank of $\mathbf{D}_{12}\mathbf{?}_i$, $i \in [1,2]$.

Condition 2.1) $M > K + L$.

Condition 2.2) $\mathbf{?}_i, i \in [1,2]$ is designed so that any K rows of $\mathbf{?}_i$ are linearly independent.

To select the appropriate precoders, we can construct them as Walsh code matrices. Then the soft decision data can be computed as:

$$\underline{s}(n) = \Gamma \cdot \underline{z}(n) = inv\left(\begin{bmatrix} \mathbf{D}_{12}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12}\mathbf{?}_2 \end{bmatrix}\right)\underline{z}(n) \tag{2.24}$$

where

$$\Gamma = inv\left(\begin{bmatrix} \mathbf{D}_{12}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12}\mathbf{?}_2 \end{bmatrix}\right) \tag{2.25}$$

Then we can project $\underline{s}(n)$ onto the finite alphabet to obtain the hard decision discrete values $\hat{\underline{s}}(n)$.

However, we can adopt the precoder like $[\mathbf{I}_{pre}^T \ \mathbf{I}_k^T]^T$ where $\mathbf{I}_{pre}$ is formed by any *M-K* rows of the $K \times K$ identity matrix $\mathbf{I}_k$, as will be used in chapter 4.

In Eq.(2.20), we assume the channel information is already known at the receiver end, next we want to show how the channel become available in chapter 3.

Chapter 3

Blind Channel Estimation

In this chapter, we rely on $M \times K$ redundant linear precoders Θ_1 and Θ_2 ($M > K$ to introduce redundancy) to show a blind channel estimation algorithm for Space-Time OFDM transmissions [17], and based on this channel estimation algorithm, we derive the theoretical mean square error of the estimator in section 3.2.

3.1 Subspace-based multichannel estimation

3.1.1 Noise free case

Before addressing the noisy case, we will start from the noiseless vectors $\underline{x}(n)$ in Eq.(2.18):

$$\underline{x}(n)=\mathbf{A}\underline{s}(n) \tag{3.1}$$

To estimate the channels, the receiver collects N blocks of $\underline{x}(n)$ and forms a

$2M \times N$ matrix $\mathbf{X}_N = [\underline{x}(0), \ldots, \underline{x}(N-1)]$, thus

$$\mathbf{X}_N = \mathbf{A}\mathbf{S}_N \tag{3.2}$$

with $\mathbf{S}_N = [\underline{s}(0), \ldots, \underline{s}(N-1)]$. At the receiver end, we also select the following [17]:

Condition 3.1) Note here the number of blocks N is large enough $(\geq 2K)$ so that $\mathbf{S}_N$ has full rank *2K*.

Condition 3.1) expresses the standard "persistence of excitation" assumption that is satisfied by all signal constellations for *N* sufficiently large.

Condition 2.1) and *Condition 2.2),* together with *Condition 3.1),* and that $\underline{s}(n)$ is a $2K \times 1$ independent vector, imply that rank($\mathbf{X}_N$)=2*K*, and the range space $Range(\mathbf{X}_N \mathbf{X}_N^H) = Range(\mathbf{A})$, and the nullity of $\mathbf{X}_N$ is $2M$ - $2K$. Further, the singular value decomposition (SVD)

$$\mathbf{X}_N = \mathbf{A}\mathbf{S}_N = [\mathbf{U}_x \ \mathbf{U}_n] \begin{bmatrix} \Sigma_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_x^H \\ \mathbf{V}_n^H \end{bmatrix} \tag{3.3}$$

with Σ_x =diag($s_1^2, \ldots, s_{2k}^2$) and $s_1^2 \geq \ldots \geq s_{2k}^2$, yielding the $2M \times (2M - 2K)$ matrix $\mathbf{U}_n$, whose columns span the null space $N(X_N)$, which is caused by the redundant precoders.

Because $N(X_N)$ is orthogonal to $Range(\mathbf{X}_N) = Range(\mathbf{A})$, it follows that

$$\underline{u}_k^H \mathbf{A} = 0_{2k \times 1}^T \quad \text{for } k \in [1, 2M - 2K] \tag{3.4}$$

where $\underline{u}_k$ is the null space vector standing for the *k*th column of $\mathbf{U}_n$.

Let us now split the null space vector $\underline{u}_k$ into its upper and lower parts as

$$\underline{u}_k = \begin{bmatrix} \hat{\underline{u}}_k \\ \breve{\underline{u}}_k \end{bmatrix}, \tag{3.5}$$

where $\hat{\underline{u}}_k$ and $\breve{\underline{u}}_k$ are $M \times 1$ vectors. Using Eq. (3.4) and Eq.(2.17), we can factor

$\underline{u}_k^H \mathbf{A}$ as

$$[\hat{\underline{u}}_k^H \ \breve{\underline{u}}_k^H]\begin{bmatrix}\mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2)\\ \mathbf{D}(\underline{H}_2^*) & -\mathbf{D}(\underline{H}_1^*)\end{bmatrix}\begin{bmatrix}\mathbf{?}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{?}_2\end{bmatrix}=0^T \tag{3.6}$$

Since for any $M\times 1$ vectors $\underline{a}$ and $\underline{b}$ it holds that

$$\underline{a}^H\mathbf{D}(\underline{b})=\underline{b}^T\mathbf{D}(\underline{a}^*), \tag{3.7}$$

Eq. (3.5) can be rewritten as

$$[\underline{H}_1^T \ \underline{H}_2^H]\underbrace{\begin{bmatrix}\mathbf{D}(\hat{\underline{u}}_k^*) & -\mathbf{D}(\breve{\underline{u}}_k)\\ \mathbf{D}(\breve{\underline{u}}_k^*) & \mathbf{D}(\hat{\underline{u}}_k)\end{bmatrix}}_{=\mathbf{G}(\underline{u}_k)}\underbrace{\begin{bmatrix}\mathbf{?}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{?}_2^*\end{bmatrix}}_{=\Psi}=0^T \tag{3.8}$$

where

$$\Psi=\begin{bmatrix}\mathbf{?}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{?}_2^*\end{bmatrix},\qquad \mathbf{G}(\underline{u}_k)=\begin{bmatrix}\mathbf{D}(\hat{\underline{u}}_k^*) & -\mathbf{D}(\breve{\underline{u}}_k)\\ \mathbf{D}(\breve{\underline{u}}_k^*) & \mathbf{D}(\hat{\underline{u}}_k)\end{bmatrix} \tag{3.9}$$

Plugging the Eq.(2.15) into Eq.(3.8), we have

$$[\underline{h}_1^T \ \underline{h}_2^H]\underbrace{\begin{bmatrix}\mathbf{V}^T & 0\\ 0 & \mathbf{V}^H\end{bmatrix}}_{=\mathbf{F}}\underbrace{\begin{bmatrix}\mathbf{D}(\hat{\underline{u}}_k^*) & -\mathbf{D}(\breve{\underline{u}}_k)\\ \mathbf{D}(\breve{\underline{u}}_k^*) & \mathbf{D}(\hat{\underline{u}}_k)\end{bmatrix}}_{=\mathbf{G}(\underline{u}_k)}\Psi=0^T \tag{3.10}$$

where

$$F=\begin{bmatrix}\mathbf{V}^T & 0\\ 0 & \mathbf{V}^H\end{bmatrix} \tag{3.11}$$

In the above, we use Eq.(3.11) in time domain instead of Eq.(3.8) in frequency domain since it can make $G(\underline{u}_k)\Psi$ a smaller matrix which can reduce computation, and in Eq.(3.8) $G(\underline{u}_k)\Psi$ cannot find the actual channel since it is not full rank.

Stacking Eq.(3.10) for each null space vector $\underline{u}_k$ with $k\in[1,\ldots,2M-2K]$, we obtain

$$\underbrace{[\underline{h}_1^T \ \underline{h}_2^H]}_{\tilde{\underline{h}}^H}\underbrace{F\left[\mathbf{G}(\underline{u}_1)\Psi,\ldots,\mathbf{G}(\underline{u}_{2J-2K})\Psi\right]}_{\mathbf{Q}}=0^T \tag{3.12}$$

where

$$\mathbf{Q} = F\left[\mathbf{G}(\underline{u}_1)\Psi, \ldots, \mathbf{G}(\underline{u}_{2J-2K})\Psi\right]. \tag{3.13}$$

$$\underline{\tilde{h}}^H = [\underline{h}_1^T \ \underline{h}_2^H] \tag{3.14}$$

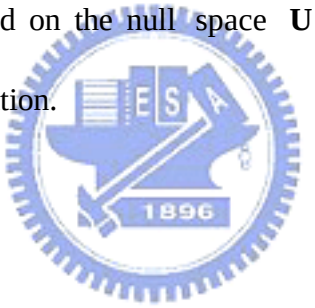
Hence,

$$\| \underline{\tilde{h}}^H \mathbf{Q} \|^2 = \underline{\tilde{h}}^H \mathbf{Q}\mathbf{Q}^H \underline{\tilde{h}} = 0 \tag{3.15}$$

Then, we can find the estimated channel $\underline{\hat{\tilde{h}}}$ as the eigenvector corresponding to the smallest eigenvalue of $\mathbf{Q}\mathbf{Q}^H$:

$$\begin{aligned} \underline{\hat{\tilde{h}}} &= \arg\min_{\|\underline{\tilde{h}}\|=1} \underline{\tilde{h}}^H \mathbf{Q}\mathbf{Q}^H \underline{\tilde{h}} \\ &= \text{eigenvector corresponding to the smallest eigenvalue of } \mathbf{Q}\mathbf{Q}^H \end{aligned} \tag{3.16}$$

The above algorithm is based on the null space $\mathbf{U}_n$, so we name this algorithm subspace-based channel estimation.



3.1.2 Noisy case

In the presence of white noise with variance s_w^2, we replace $\mathbf{X}_N$ in Eq.(3.3) by $\mathbf{Y}_N$, whose SVD has the following form:

$$\mathbf{Y}_N = [\tilde{\mathbf{U}}_x \ \tilde{\mathbf{U}}_n] \begin{bmatrix} \tilde{\Sigma}_x & \mathbf{0} \\ \mathbf{0} & \Sigma_n \end{bmatrix} \begin{bmatrix} \mathbf{V}_x^H \\ \mathbf{V}_n^H \end{bmatrix} \tag{3.17}$$

However, for large N, $\mathbf{Y}_N$ is replaced by the sample covariance matrix:

$$\mathbf{R}_{\tilde{y}} = \frac{1}{N}\sum_{n=1}^{N} \underline{\tilde{y}}(n)\underline{\tilde{y}}^H(n) = [\bar{\mathbf{U}}_x \ \bar{\mathbf{U}}_n] \begin{bmatrix} \tilde{\Sigma}_x^2 & \mathbf{0} \\ \mathbf{0} & \Sigma_n^2 \end{bmatrix} \begin{bmatrix} \bar{\mathbf{U}}_x^H \\ \bar{\mathbf{U}}_n^H \end{bmatrix} \tag{3.18}$$

to simplify the computation. Note here $N \geq 2K$ is a necessary condition.

Whether the channel identifiability can be guaranteed is summarized in the

following theorem [17]:

Theorem 1: Suppose *Condition 2.1), Condition 2.2),* and *Condition 3.1)* hold true; let $\bar{\mathbf{D}}$ denote any diagonal matrix with unit amplitude diagonal entries, and let $\bar{\Theta}_1, \bar{\Theta}_2$ be formed by any $J-L$ rows of Θ_1, Θ_2, respectively. If $\bar{\Theta}_1$ *and* $\bar{\Theta}_2$ satisfy $\bar{\mathbf{D}}\bar{\Theta}_1 \notin R(\bar{\Theta}_2)$, the solution of $[h_3^T, h_4^H]Q = 0^T$ is unique up to a constant, and thus, channel identifiability within one scalar is guaranteed:

$$\begin{bmatrix} \underline{h}_1 \\ \underline{h}_2 \end{bmatrix} = \begin{bmatrix} a\mathbf{I} & 0 \\ 0 & a^*\mathbf{I} \end{bmatrix} \begin{bmatrix} \underline{h}_3 \\ \underline{h}_4 \end{bmatrix} \tag{3.19}$$

Here we've assumed that $(\underline{h}_3, \underline{h}_4)$ is a pair of channel satisfying Eq.(3.16). ?

We summarize the proposed subspace-based channel estimation algorithm [17] in the following steps.

Step 1) Collect the received data blocks *y(n),* and compute $\mathbf{R}_{\tilde{y}}$ in Eq. (3.18).

Step 2) Determine the eigenvectors $\underline{u}_k, k = 1, \ldots 2M - 2K$ corresponding to the smallest eigenvalues of the matrix $\mathbf{R}_{\tilde{y}}$.

Step 3) Build $\mathbf{Q}$ in Eq.(3.13).

Step 4) Determine the eigenvector corresponding to the smallest eigenvalue of $\mathbf{QQ}^H$ as our estimate in Eq.(3.16).

3.2 Performance Analysis of Mean Square Error

A performance analysis is conducted on the proposed estimator to derive the theoretical mean square error (MSE). Here we start from *Theorem 2.*

Theorem 2: Assuming that both noise and the signals are zero mean i.i.d. random variables with variance s_s^2 and s_w^2, respectively, an approximation for the channel estimate' s MSE in Eq.(3.16) for high sample SNR and large sample size is

$$E(\| \hat{\underline{h}} - \underline{h} \|^2) \approx \frac{s_w^2 \| \mathbf{Q}^+ \|^2}{s_s^2 N} \tag{3.20}$$

In the following, we want to prove Eq.(3.20) based on the first-order perturbation theory given in [19] in high SNR condition (small perturbation).

Lemma 1[19] *:* Assuming a matrix $\mathbf{P}$ permits the SVD

$$\mathbf{P} = [\mathbf{U}_p \ \mathbf{U}_g] \begin{bmatrix} \Sigma_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_p^H \\ \mathbf{V}_g^H \end{bmatrix} \tag{3.21}$$

and the perturbed matrix $\tilde{\mathbf{P}}$ can be written as

$$\tilde{\mathbf{P}} = \mathbf{P} + \Delta\mathbf{P} = [\tilde{\mathbf{U}}_p \ \tilde{\mathbf{U}}_g] \begin{bmatrix} \tilde{\Sigma}_p & \mathbf{0} \\ \mathbf{0} & \Sigma_g \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{V}}_p^H \\ \tilde{\mathbf{V}}_g^H \end{bmatrix} \tag{3.22}$$

where

$$\tilde{\mathbf{U}}_n = \mathbf{U}_n + \Delta\mathbf{U}_n \tag{3.23}$$

g is a vector of $\mathbf{U}_g$, then

$$g^H \mathbf{P} = 0, \tag{3.24}$$

the first-order approximation of the perturbation to the vector g due to additive perturbation $\Delta\mathbf{P}$ to g is

$$\Delta g = -\mathbf{P}^+ \Delta\mathbf{P}^H g \tag{3.25}$$

where

$$\mathbf{P} = \mathbf{U}_p \Sigma_p^{-1} \mathbf{V}_p^H \tag{3.26}$$

Deducing from *lemma 1*, since $\underline{u}_k$ is the null space of $\mathbf{X}_N$, $\underline{u}_k^H \mathbf{X}_N = 0$, then

$$\triangle \underline{u}_k = \mathbf{X}_N^+ \Delta \mathbf{X}_N^H \underline{u}_k \tag{3.27}$$

where

$$\mathbf{X}_N^+ = \mathbf{U}_x \sum_x^{-1} \mathbf{V}_x^H \tag{3.28}$$

Moreover, from $\tilde{\underline{h}}^H \mathbf{Q} = 0$ in Eq.(3.12.), the perturbation of the channel estimate is

$$\triangle \underline{h} = -\mathbf{Q}^+ \Delta \mathbf{Q}^H \underline{h} \tag{3.29}$$

where the perturbation $\triangle \mathbf{Q}$ to $\mathbf{Q}$ is additive and can be formed

$$\Delta Q = F\left[\mathbf{D}(\triangle \underline{u}_1)\Psi, \ldots, \mathbf{D}(\triangle \underline{u}_{2J-2K})\Psi\right] \tag{3.30}$$

since F and Ψ are both deterministic.

In addition, the structure of $\triangle \mathbf{Q}$ gives

$$\begin{aligned}
\left(\triangle \mathbf{Q}^H \underline{h}\right)^H &= \underline{h}^H \triangle \mathbf{Q} \\
&= [\underline{h}_1^T \ \underline{h}_2^H] \underbrace{F\left[\mathbf{G}(\triangle \underline{u}_1), \ldots, \mathbf{G}(\triangle \underline{u}_{2M-2K})\right]}_{=\Delta Q} \Psi \\
&= [\underline{H}_1^T \ \underline{H}_2^H] \left[\mathbf{G}(\triangle \underline{u}_1), \ldots, \mathbf{G}(\triangle \underline{u}_{2M-2K})\right] \Psi \\
&= [\underline{H}_1^T \ \underline{H}_2^H] \left\{ \begin{bmatrix} \mathbf{D}(\hat{\underline{u}}_1^*) & -\mathbf{D}(\breve{\underline{u}}_1) \\ \mathbf{D}(\breve{\underline{u}}_1^*) & \mathbf{D}(\hat{\underline{u}}_1) \end{bmatrix}, \ldots, \begin{bmatrix} \mathbf{D}(\hat{\underline{u}}_{2M-2K}^*) & -\mathbf{D}(\breve{\underline{u}}_{2M-2K}) \\ \mathbf{D}(\breve{\underline{u}}_{2M-2K}^*) & \mathbf{D}(\hat{\underline{u}}_{2M-2K}) \end{bmatrix} \right\} \Psi \\
&= \left\{ \triangle \underline{u}_1^H \begin{bmatrix} \mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & -\mathbf{D}(\underline{H}_1^*) \end{bmatrix}, \ldots, \triangle \underline{u}_{2M-2K}^H \begin{bmatrix} \mathbf{D}(\underline{H}_1) & \mathbf{D}(\underline{H}_2) \\ \mathbf{D}(\underline{H}_2^*) & -\mathbf{D}(\underline{H}_1^*) \end{bmatrix} \right\} \Psi \\
&= [\triangle \tilde{\underline{u}}_1^H, \ldots, \triangle \tilde{\underline{u}}_{2M-2K}^H] \bar{\mathbf{D}} \Psi
\end{aligned} \tag{3.31}$$

The deduction of Eq.(3.31) is based on Eq.(3.6),Eq. (3.8), and Eq.(3.10), then

$$
\begin{aligned}
{}_{\triangle}\mathbf{Q}^H \underline{h} &= \begin{bmatrix} \Psi^H \bar{\mathbf{D}}^H {}_{\triangle}\underline{u}_1 \\ \\ \Psi^H \bar{\mathbf{D}}^H {}_{\triangle}\underline{u}_{2M-2K} \end{bmatrix} \\
&= \begin{bmatrix} -\Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ \Delta\mathbf{X}_N^H \underline{u}_1 \\ \\ -\Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ \Delta\mathbf{X}_N^H \underline{u}_{2M-2K} \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ & & \\ & \ddots & \\ & & \Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \Delta\mathbf{X}_N^H \underline{u}_1 \\ \vdots \\ \Delta\mathbf{X}_N^H \Delta\mathbf{X}_N^H \underline{u}_{2M-2K} \end{bmatrix}}_{\mathbf{d}}
\end{aligned}
\tag{3.32}
$$

where

$$
\begin{aligned}
\mathbf{B} &= \begin{bmatrix} \Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ & & \\ & \ddots & \\ & & \Psi^H \bar{\mathbf{D}}^H \mathbf{X}_N^+ \end{bmatrix} \\
\mathbf{d} &= \begin{bmatrix} \Delta\mathbf{X}_N^H \underline{u}_1 \\ \vdots \\ \Delta\mathbf{X}_N^H \Delta\mathbf{X}_N^H \underline{u}_{2M-2K} \end{bmatrix}
\end{aligned}
\tag{3.33}
$$

Substituting Eq.(3.32) into Eq.(3.29) leads to

$$
\begin{aligned}
{}_{\triangle}\underline{h} &= -\mathbf{Q}^+ {}_{\triangle}\mathbf{Q}^H \underline{h} \\
&= \mathbf{Q}^+ \mathbf{B}\mathbf{d}
\end{aligned}
\tag{3.34}
$$

Next, before computing the channel MSE, we prove the following lemma.

Lemma 2: Assume $\mathbf{N}$ is a matrix where each element is zero-mean i.i.d. random variable with variance s^2. Also assume $\mathbf{J}$ is an $m \times m$ deterministic matrix. Then $E(\mathbf{N}^H \mathbf{J}\mathbf{N}) = s^2 trace(\mathbf{J})\mathbf{I}_n$ where $\mathbf{I}_n$ is an $n \times n$ identity matrix and $trace(\cdot)$ gives the trace of the matrix. ?

Proof: Define $\mathbf{F} = \mathbf{N}^H \mathbf{J}\mathbf{N}$. Then

$$
\begin{aligned}
E(\mathbf{F}_{ij}) &= E\Big(\sum_{l=1}^{m}\Big[\Big(\sum_{k=1}^{m}\mathbf{N}_{ki}^{*}\cdot\mathbf{J}_{kl}\Big)\cdot\mathbf{N}_{lj}\Big]\Big) \\
&= \sum_{l=1}^{m}\Big[\sum_{k=1}^{m}\mathbf{J}_{kl}\cdot E(\mathbf{N}_{ki}^{*}\cdot\mathbf{N}_{lj})\Big] \\
&= \begin{cases} \sum_{l=1}^{m}\mathbf{J}_{ll}s^{2}, & i=j \\ 0, & \textit{otherwise.} \end{cases}
\end{aligned}
\tag{3.35}
$$

Therefore,

$$
E(\mathbf{N}^{H}\mathbf{J}\mathbf{N}) = s^{2}\sum_{l=1}^{m}\mathbf{J}_{ll}\mathbf{I}_{n} = s^{2}trace(\mathbf{J})\mathbf{I}_{n}. \tag{3.36}
$$

By using *Lemma 2*, it is easy to see that for the considered problem, we have

$$
\begin{aligned}
E(\Delta\mathbf{X}_{N}^{H}\underline{u}_{i}\underline{u}_{j}^{H}\Delta\mathbf{X}_{N}) &= s^{2}trace(\underline{u}_{i}\underline{u}_{j}^{H})\mathbf{I} \\
&= s_{n}^{2}d(i-j)\mathbf{I}
\end{aligned}
\tag{3.37}
$$

with $d(\cdot)$ being the Kronecker Delta function and, consequently

$$
E(\mathbf{d}\mathbf{d}^{H}) = s_{n}^{2}\mathbf{I} \tag{3.38}
$$

Then we can write estimated channel error covariance matrix

$$
\begin{aligned}
E[\triangle\underline{h}\cdot\triangle\underline{h}^{H}] &= E[\mathbf{Q}^{+}\mathbf{B}\mathbf{d}\mathbf{d}^{H}\mathbf{B}^{H}\mathbf{Q}^{+^{H}}] \\
&= \mathbf{Q}^{+}\mathbf{B}E[\mathbf{d}\mathbf{d}^{H}]\mathbf{B}^{H}\mathbf{Q}^{+^{H}} \\
&= s_{n}^{2}\mathbf{Q}^{+}\mathbf{B}\mathbf{B}^{H}\mathbf{Q}^{+^{H}}
\end{aligned}
\tag{3.39}
$$

where

$$
\mathbf{B}\mathbf{B}^{H} = \begin{bmatrix} \Psi^{H}\mathbf{D}^{H}\mathbf{X}_{N}^{+}\mathbf{X}_{N}^{+^{H}}\mathbf{D}\Psi & & \\ & \ddots & \\ & & \Psi^{H}\mathbf{D}^{H}\mathbf{X}_{N}^{+}\mathbf{X}_{N}^{+^{H}}\mathbf{D}\Psi \end{bmatrix} \tag{3.40}
$$

Here we assume that $\mathbf{d}$ is the only random variable.

In addition, to begin the derivation we recall that the unperturbed data $\mathbf{X}_{N}$ can be written as $\mathbf{X}_{N} = \mathbf{D}?\,\mathbf{S}_{N}$ in Eq.(3.2) by setting $\Psi = \Theta$,which means we need to select $?_{2} = ?_{2}^{*}$. We define the data covariance matrix $\mathbf{R}_{\mathbf{x}}$ as

$$\begin{aligned}\mathbf{R}_{\mathbf{x}} &= \mathbf{X}_N \mathbf{X}_N^H \\ &= \mathbf{D?}\ \mathbf{S}_N \mathbf{S}_N^H \mathbf{?}^{\,H} \mathbf{D}^H\end{aligned} \tag{3.41}$$

The generalized inverse of $\mathbf{R}_{\mathbf{x}}$ is

$$\begin{aligned}\mathbf{R}_{\mathbf{X}}^{+} &= \mathbf{X}_N^{+} \mathbf{X}_N^{+^H} \\ &= (\mathbf{?}^{\,H} \mathbf{D}^H)^{+} (\mathbf{S}_N \mathbf{S}_N^H)^{-1} (\mathbf{D?})^{+}\end{aligned} \tag{3.42}$$

Hence, we have

$$\begin{aligned}&\Psi^H \mathbf{D}^H \mathbf{X}_N^{+} \mathbf{X}_N^{+^H} \mathbf{D}\Psi \\ &= \Psi^H \mathbf{D}^H (\mathbf{?}^{\,H} \mathbf{D}^H)^{+} \mathbf{R}_s^{-1} (\mathbf{D?})^{+} \mathbf{D}\Psi \\ &= (\mathbf{S}_N \mathbf{S}_N^H)^{-1}\end{aligned} \tag{3.43}$$

Therefore, Eq.(3.37) becomes

$$\mathbf{B}\mathbf{B}^H = \begin{bmatrix} (\mathbf{S}_N \mathbf{S}_N^H)^{-1} & & \\ & \ddots & \\ & & (\mathbf{S}_N \mathbf{S}_N^H)^{-1} \end{bmatrix} \tag{3.44}$$

For large N (number of data blocks), approximation

$$\mathbf{S}_N \mathbf{S}_N^H = N s_s^2 \mathbf{I} \tag{3.45}$$

is reasonable, since we already assume that the signals are zero mean i.i.d. random variables. Thus

$$\mathbf{B}\mathbf{B}^H \approx \frac{1}{N s_s^2} \mathbf{I} \tag{3.46}$$

Then Eq.(3.36) becomes

$$E[\triangle \underline{h} \cdot \triangle \underline{h}^H] \approx \frac{s_n^2 \mathbf{Q}^{+} \mathbf{Q}^{+^H}}{s_s^2 \cdot N} \tag{3.47}$$

As a consequence, the MSE of the channel estimate is

$$
\begin{aligned}
E(\|\Delta \underline{h}\|^2) &= E[trace(\Delta \underline{h} \cdot \Delta \underline{h}^H)] \\
&\approx \frac{\sigma_n^2 \cdot trace(\mathbf{Q}^+ \mathbf{Q}^{+^H})}{\sigma_s^2 N} \\
&= \frac{\sigma_n^2 \cdot \|\mathbf{Q}^+\|^2}{\sigma_s^2 N}
\end{aligned}
\tag{3.48}
$$

The closed form MSE expression Eq.(3.48) is compact and enables us to study the estimator's performance dependence on the key system parameters—such as the input SNR, the number of data blocks. As expected, the MSE decreases with increasing input SNR and the number of received data blocks. This is later verified by computer simulation in Chapter 5.

CHAPTER 4

Improved Subspace Methods

To further improve the channel estimation, we can exploit the finite alphabet property to better the subspace-based channel estimates.

In this chapter, we discuss two different methods: decision direct (DD) [21] and phase direct (PD) [20]. DD, as implied in the name, needs first to get the hard decision data and then use it to update our estimated channel, while PD is to solve the phase ambiguities after we've got the channel power response. DD originally works in conventional OFDM, which only requires simple scalar division. Based on the space time data matrix, we extend it to ST-OFDM, which corresponds to a matrix inverse and multiplication because the received data is composed of two different transmitted data.

The main idea of PD is to solve the phase ambiguities after we get the channel power response. For conventional OFDM system, it is very easy to get the channel power response. But in space-time OFDM, it is quite a different case, since the received data is composed of two different transmitted data; it's not easy to separate

them. So, the main problem we face now is how to get the channel power response. However, in general case the channel power response is hard to obtain. Hence, we only focus on BPSK system and exploit the transmitted data's time and temporal correlation to develop a new algorithm named sum-difference square method to solve this problem.

Moreover, in time varying channel, we also need to choose a best window size to get the channel power response and apply it to PD. As we all know, when the window is longer, we can suppress the noise, but then we can't follow the variance of the channel. This is the trade off. However, the choice of the window size is dependent on how fast the channel changes. Precoder design is another issue behind the algorithm as will be discussed in section 4.2.4.

4.1 Decision Direct (DD)

4.1.1 DD in ST-OFDM

In [21], Giannakis has shown how DD works in conventional OFDM, and we want to extend it to space time OFDM by using the space time data matrix.

There are four steps shown as follows:

Step (1) Set j=0, find an initial time domain channel estimate $\hat{\underline{h}}_{1,(j)}$ and $\hat{\underline{h}}_{2,(j)}$ using Eq.(3.16), calculate the frequency response $\hat{\underline{H}}_{1,(j)}$ and $\hat{\underline{H}}_{2,(j)}$ by Eq.(2.15).

Step (2) Since now we have $\mathbf{D}_{1,(j)}$ and $\mathbf{D}_{2,(j)}$, it is possible to detect $\underline{s}(n)$ with diversity gains by a simple matrix multiplication as Eq.(2.21) by substituting $\mathbf{D}(\underline{H}_1)$ and $\mathbf{D}(\underline{H}_2)$.

$$\begin{aligned}\underline{z}(n) &= \bar{\mathbf{D}}^H \underline{\tilde{y}}(n) \\ &= \begin{bmatrix} \mathbf{D}(\underline{H}^*_{1,(j)}) & \mathbf{D}(\underline{H}_{2,(j)}) \\ \mathbf{D}(\underline{H}^*_{2,(j)}) & \text{-}\mathbf{D}(\underline{H}_{1,(j)}) \end{bmatrix} \underline{\tilde{y}}(n) + \bar{\mathbf{D}}^H \underline{h}(n) \\ &= \begin{bmatrix} \mathbf{D}_{12,(j)}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12,(j)}\mathbf{?}_2 \end{bmatrix} \underline{s}(n) + \underline{x}(n)\end{aligned} \tag{4.1}$$

Then we can get the soft decision $\underline{s}(n)$ as Eq.(2.24):

$$\underline{s}(n) = \Gamma \cdot \underline{z}(n) = inv\left(\begin{bmatrix} \mathbf{D}_{12,(j)}\mathbf{?}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{12,(j)}\mathbf{?}_2 \end{bmatrix} \right) \underline{z}(n) \tag{4.2}$$

then we can project onto the finite alphabet to obtain the hard decision discrete values $\underline{\hat{s}}_{(j)}(n)$.

Eq.(4.2) shows how we collect soft decision data which corresponds to

$$\hat{s}_{(j)}(n,m) = y(n,m) / \hat{H}_{(j)}(r_m) \tag{4.3}$$

in [20] for conventional OFDM which is only a simple one tap equalizer. However in ST OFDM, we need first to get $\underline{z}(n)$, the ST decoder output, and then feed it to equalizer Γ which compensates not only channel gain but also the precoder effect, hence it is more complicate.

Where $y(n,m)$ indicates the data for the mth subcarrier on the nth received data block, $\hat{s}(n,m)$ is the decision data corresponding to the mth subcarrier on the nth block, $\hat{H}(r_m)$ is the estimated frequency channel of the mth subcarrier.

Step (3) Here, we separate step 3 to two parts. Part a) is to update the estimated channel from the decision data, and part b) is to update the detected data from the updated channel.

(a) Here, we rewrite Eq. (2.16) again.

$$\underline{\bar{y}}(n)=\begin{bmatrix}\mathbf{D}_1\underline{\tilde{s}}^{(1)}(n)+\mathbf{D}_2\underline{\tilde{s}}^{(2)}(n)\\ -\mathbf{D}_1\underline{\tilde{s}}^{(2)^*}(n)+\mathbf{D}_2\underline{\tilde{s}}^{(1)^*}(n)\end{bmatrix}+\begin{bmatrix}\underline{v}^{(1)}(n)\\ \underline{v}^{(2)}(n)\end{bmatrix} \tag{4.4}$$

Let $\mathbf{D}(v)$ stand for the diagonal matrix with the vector v, and define

$$\begin{aligned}\tilde{\mathbf{S}}^{(1)}(n)&=\mathbf{D}(\underline{\tilde{s}}^{(1)}(n))\\ \tilde{\mathbf{S}}^{(2)}(n)&=\mathbf{D}(\underline{\tilde{s}}^{(2)}(n))\end{aligned} \tag{4.5}$$

Since for $J\times 1$ vectors $\underline{a}$ and $\underline{b}$ it holds that $\mathbf{D}(\underline{a})\underline{b}=\mathbf{D}(\underline{b})\underline{a}$, Eq.(4.3) can be rewritten as

$$\underline{\bar{y}}(n)=\begin{bmatrix}\tilde{\mathbf{S}}^{(1)}(n)\underline{H}_1+\tilde{\mathbf{S}}^{(2)}(n)\underline{H}_2\\ -\tilde{\mathbf{S}}^{(2)^*}(n)\underline{H}_1+\tilde{\mathbf{S}}^{(1)^*}(n)\underline{H}_2\end{bmatrix}+\begin{bmatrix}\underline{v}^{(1)}(n)\\ \underline{v}^{(2)}(n)\end{bmatrix} \tag{4.6}$$

Apply the hard decision data to Eq.(4.6):

$$\underline{\bar{y}}(n)=\begin{bmatrix}\hat{\mathbf{S}}^{(1)}_{(j)}(n) & \hat{\mathbf{S}}^{(2)}_{(j)}(n)\\ -\hat{\mathbf{S}}^{(2)^*}_{(j)}(n) & \hat{\mathbf{S}}^{(1)^*}_{(j)}(n)\end{bmatrix}\begin{bmatrix}\underline{H}_{1(j)}\\ \underline{H}_{2(j)}\end{bmatrix}+\begin{bmatrix}\underline{v}^{(1)}(n)\\ \underline{v}^{(2)}(n)\end{bmatrix} \tag{4.7}$$

where

$$\hat{\mathbf{S}}^{(1)}_{(j)}(n)=\mathbf{D}(?_1\underline{\hat{s}}_{1,(j)}(n)) \text{ and } \hat{\mathbf{S}}^{(2)}_{(j)}(n)=\mathbf{D}(?_2\underline{\hat{s}}_{2,(j)}(n)) \tag{4.8}$$

Then we can simply get $\begin{bmatrix}\underline{H}_{1(j)}\\ \underline{H}_{2(j)}\end{bmatrix}$ though Eq.(4.7)

$$\begin{bmatrix}\underline{H}_{1(j)}\\ \underline{H}_{2(j)}\end{bmatrix}=inv\left(\begin{bmatrix}\hat{\mathbf{S}}^{(1)}_{(j)}(n) & \hat{\mathbf{S}}^{(2)}_{(j)}(n)\\ -\hat{\mathbf{S}}^{(2)^*}_{(j)}(n) & \hat{\mathbf{S}}^{(1)^*}_{(j)}(n)\end{bmatrix}\right)\underline{\bar{y}}(n) \tag{4.9}$$

Update the time domain channel estimates

$$\underline{h}_{i,(j)}=\mathbf{V}^H\underline{H}_{i,(j)} \tag{4.10}$$

and their frequency response using

$$\underline{\hat{H}}_{i,(j)}=\mathbf{V}\underline{\hat{h}}_{i,(j)} \tag{4.11}$$

Note that here Eq. (4.10) means to perform an M-point inverse DFT on $\underline{H}_{i,(j)}$ and truncate the output by keeping only the first L+1 entries (since

we have assumed the channel delay spread is equal or less then L); Eq. (4.11) amounts to performing an M-point DFT on an vector formed after zero-padding, we call Eqs.(4.10) and (4.11) denoising.

Eq.(4.9) is to update the estimated channel from the decision data which corresponds to

$$\hat{H}(r_m) = y(n,m) / \hat{s}(n,m) \tag{4.12}$$

in [20] for conventional OFDM.

(b) Check if $\underline{\hat{h}}_{i,(j)} \cong \underline{\hat{h}}_{i,(j-1)}$, if yes, stop the iteration, if no , in each successive iteration, j is added by 1.

We have updated $\underline{\hat{H}}_{i,(j)}$ in (a), then We can form $\mathbf{D}_{1,(j)}$ and $\mathbf{D}_{2,(j)}$ from $\mathbf{D}_{1,(j)}=\mathbf{D}(\underline{\hat{H}}_{1,(j-1)})$ and $\mathbf{D}_{2,(j)}=\mathbf{D}(\underline{\hat{H}}_{1,(j-1)})$. Then, recompute the symbol estimates by Eqs.(4.1) and (4.2) to get the soft decision, and project again onto the finite alphabet to obtain hard decisions $\underline{\hat{s}}_{(j)}(n)$.

Step (4) Repeat Step (3).

See Fig.4.1 for the Signal-flow graph of DD in space-time OFDM system

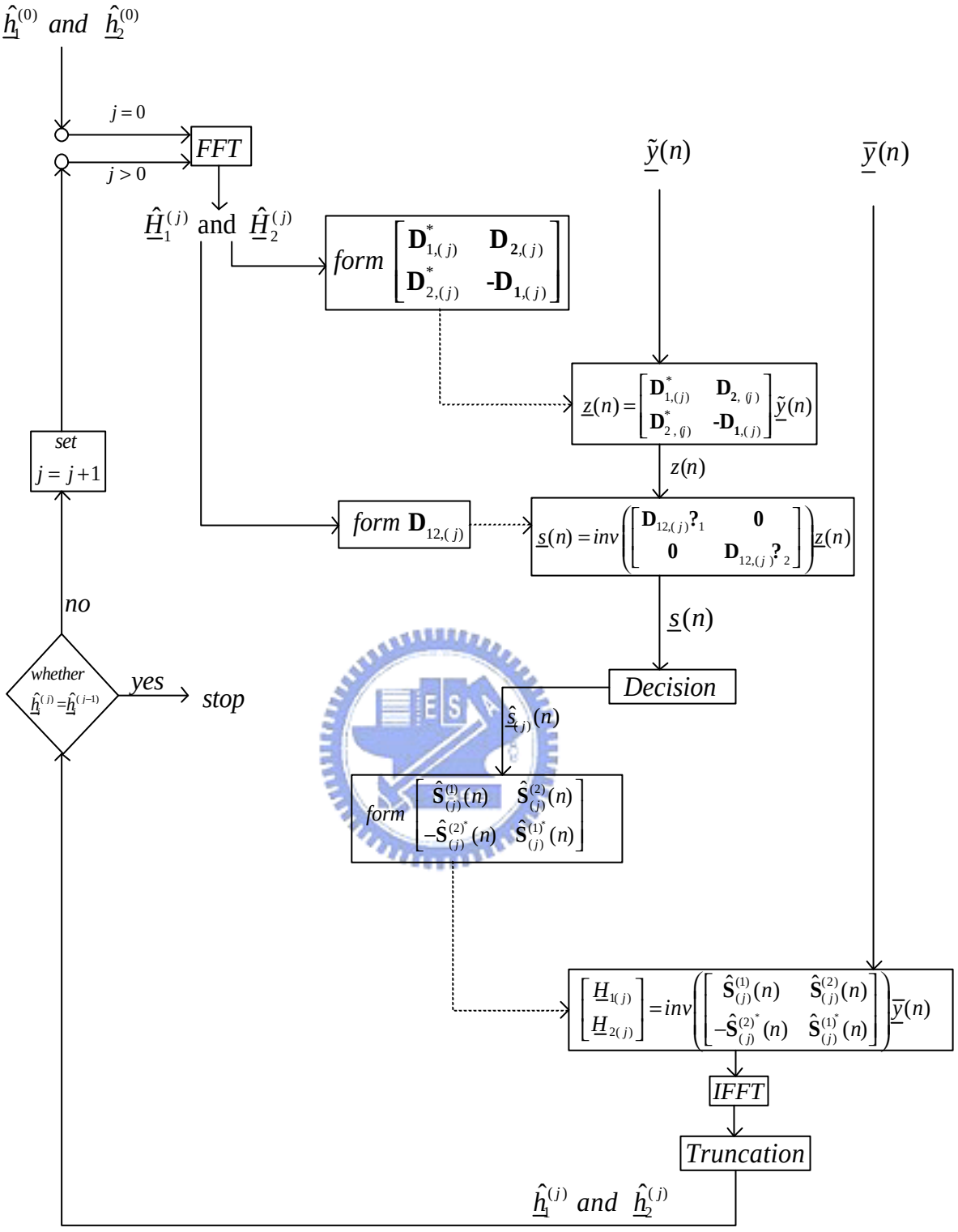


Fig. 4.1 Signal-flow graph of DD in space-time OFDM

4.1.2 Comparison with conventional OFDM

We have shown how DD works in space time OFDM system, and the main equations are Eq.(4.2) and Eq.(4.9), while the former is to get the soft decision data, and the latter is to update the channel response while we have the decision data, both of which are corresponding to matrix inverse and multiplication. However, for DD in conventional OFDM [21], these two equations just become a simple scalar division, which are

$$\begin{aligned} \hat{s}(n,m) &= y(n,m) / \hat{H}(r_m) \\ \hat{H}(r_m) &= y(n,m) / \hat{s}(n,m) \end{aligned} \tag{4.13}$$

where $y(n,m)$ indicates the data for mth subcarrier on the nth received data block, $\hat{s}(n,m)$ is the decision data corresponding to the mth subcarrier on the nth block, $\hat{H}(r_m)$ is the estimated frequency channel of the mth subcarrier. See Fig.4.2 for Signal-flow graph of DD in conventional OFDM.

Eq.(4.13) is based on the well-known formula [1],

$$y(n,m) = H(r_m)s(n,m) + n(r_m) \tag{4.14}$$

which leads to Eq.(4.13) a simple scalar division, but in space time OFDM, all the formulas are in matrix form which makes Eq.(4.2) and Eq.(4.9) both matrix inverse and multiplication. Note here that in Eq.(4.12) Γ corresponds to a psuedo inverse since its size is $2M \times 2K$, and in Eq.(4.9) $\begin{bmatrix} \hat{\mathbf{S}}_{(j)}^{(1)}(n) & \hat{\mathbf{S}}_{(j)}^{(2)}(n) \\ -\hat{\mathbf{S}}_{(j)}^{(2)^*}(n) & \hat{\mathbf{S}}_{(j)}^{(1)^*}(n) \end{bmatrix}$ is full rank of size $2M \times 2M$.

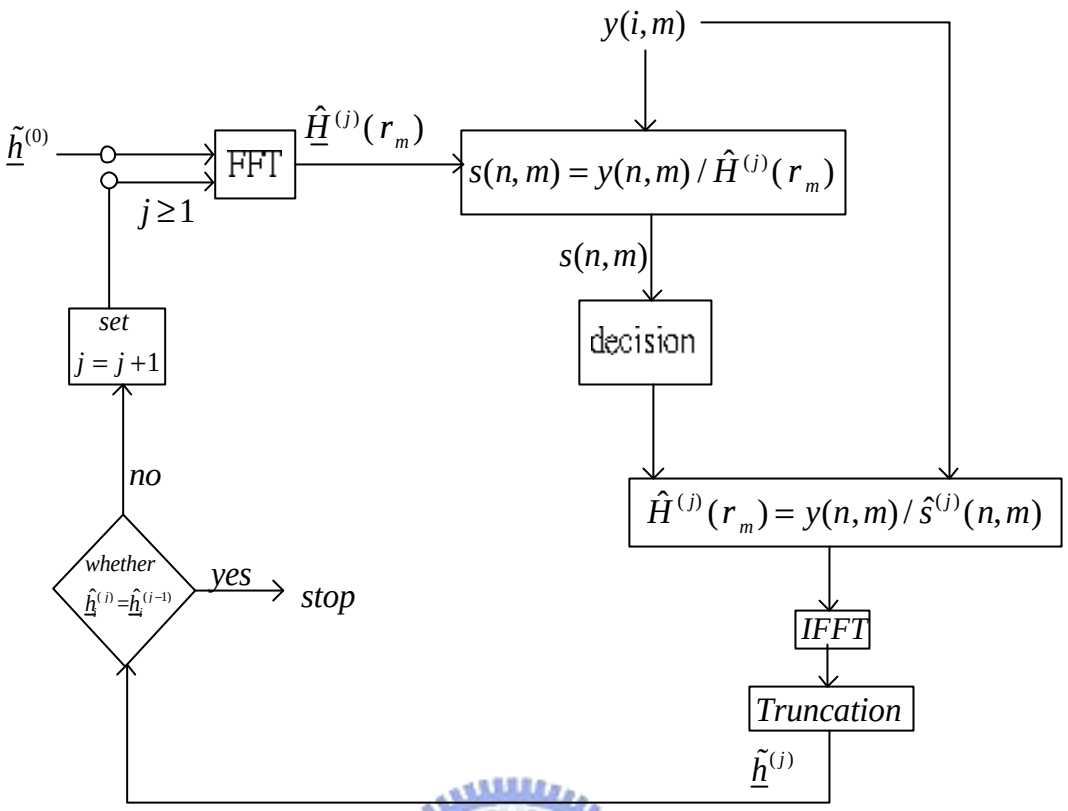


Fig. 4.2 Signal-flow graph of DD in conventional OFDM

4.2 Phase Direct (PD)

4.2.1 Introduction of PD

Before addressing PD [20] to space time OFDM, we first show how it woks in conventional OFDM. Now, we start this part from Eq.(4.14), and focus on PSK constellation of size P: $\left\{ z_P = e^{j2p\,p/P} \right\}_{p=1}^{P}$,

We take the power of P to Eq. (4.14), omit noise for simplicity, and get the expectation

$$E\{y^P(n,m)\} = E\{[H(r_m)s(n,m)]^P\} = H^P(r_m)E\{s^P(n,m)\} \tag{4.15}$$

In practice, $E\{y(i,m)^P\}$ is replaced by sample averages and thus is estimated as

$$H^P(r_m) = \frac{\frac{1}{N}\sum_{n=0}^{N-1} y^P(n,m)}{E\{s^P(n,m)\}} \tag{4.16}$$

where N is the received data block number.

Since we focus on PSK constellation of size P, we have $E\{s^O(n,m)\} = 1$. Hence, Eq.(4.16) becomes

$$H^P(r_m) = \frac{1}{N}\sum_{n=0}^{N-1} y^P(n,m)\,. \tag{4.17}$$

Hence, we've got the channel power response. Next, we only need to get the channel phase response, which means for each $m \in [0, M-1]$ (assuming total M subcarriers), we have

$$\hat{H}(r_m) = l_m[H^P(r_m)]^{1/P} \tag{4.18}$$

where $l_m \in \left\{ e^{j(2p/P)p},\quad p = 1,\ldots,P \right\}$ (4.19)

is the corresponding phase ambiguity in taking the Pth root.

For each $m \in [0, M-1]$, we can resolve the phase ambiguity by searching over candidate phase values

$$l_m = \arg\min_{l_m} \left\| H_{est}(r_m) - \hat{l}_m[H^P(r_m)]^{1/P} \right\|^2 \tag{4.20}$$

where $H_{est}(r_m)$ will be discussed in the following.

Therefore, we can improve channel estimation accuracy through what we term Phase Directed (PD) steps that we describe next:

Step (1) Set j=0, find an initial estimate $\underline{\hat{h}}_{(0)}$ using any estimated method , and calculate the frequency response $\underline{\hat{H}}_{(0)}$, then we can get $H_{est}(r_m) = \hat{H}_{(0)}(r_m)$ for *m=0,1,2,….,M-1.*

Step (2) In each successive iteration, j is added by 1, and

(a) Resolve phase ambiguities by replacing $H_{est}(r_m)$ with $H_{(j-1)}(r_m)$ in (4.20), and then form the vector

$$\underline{H}_{temp} = \left\{ l_0[H^P(r_0)]^{1/P}, \ldots\ldots, l_{M-1}[H^P(r_{M-1})]^{1/P} \right\} \tag{4.21}$$

(b) Update time domain channel estimates

$$\underline{\hat{h}}_{(j)} = \mathbf{V}^H \underline{H}_{temp} \tag{4.22}$$

and their frequency response using

$$\underline{\hat{H}}_{(j)} = \mathbf{V}\ \underline{\hat{h}}_{(j)} \tag{4.23}$$

Step 3) Repeat Step2 several times, or continue until $\underline{\hat{h}}_{(j)} \cong \underline{\hat{h}}_{(j-1)}$ within some tolerance.

See Fig.4.3 for the signal-flow graph of phase direct in conventional OFDM.

The main difference between PD and DD is that DD alternates between channel estimation and symbol detection while PD avoids symbol estimation by decoupling channel estimation from symbol recovery. Therefore, PD is immune to the well-known error propagation phenomenon that is present in DD iterations.

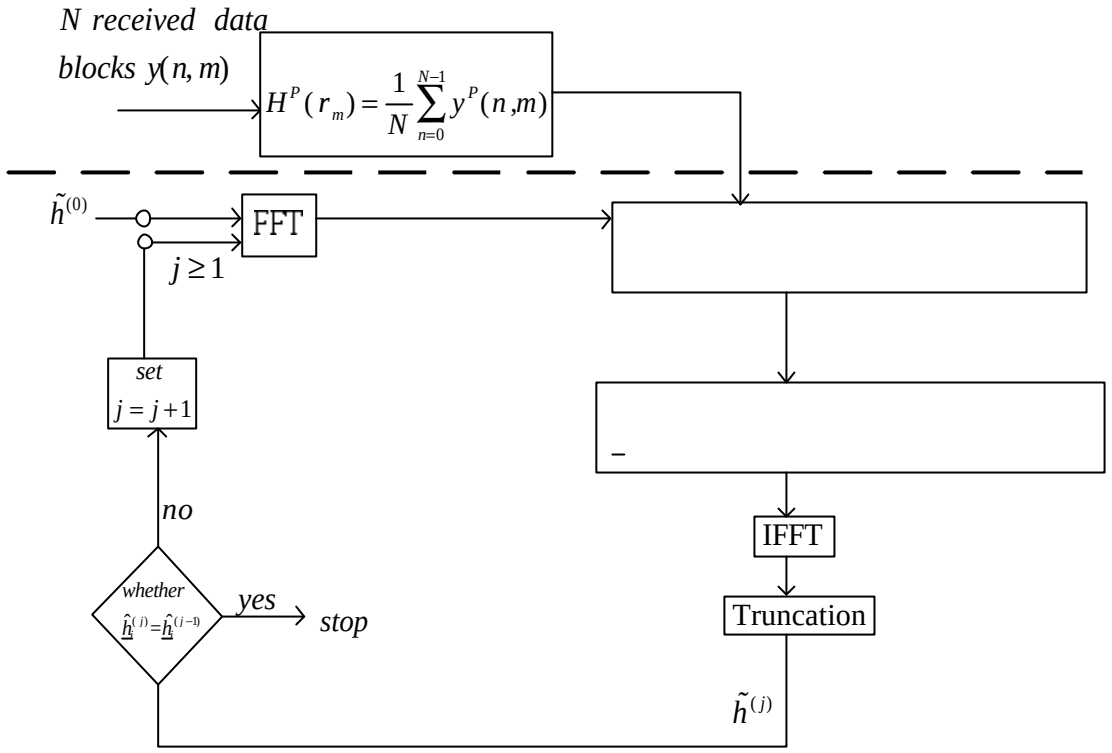


Fig. 4.3 Signal-flow graph of phase direct in conventional OFDM

4.2.2 PD in ST-OFDM based on sum-difference square algorithm

Here, we want to apply PD to the Space-Time OFDM system. As described earlier, the main idea of PD is to solve the phase ambiguities in Eq.(4.20) after we get the channel power response Eq.(4.17). For conventional OFDM system, it is very easy to get the channel power response from (4.14). But here it is quite a different case, since the received data is composed of two different transmitted data as in Eq.(2.13); it's not easy to separate them. So, here we derive an algorithm to get the channel power response.

Let's start from (2.14) and only focus on BPSK baseband system which means we

only use ± 1 data in baseband since for other system it's hard to solve the channel power response. Then we will get

$$\begin{bmatrix} \underline{\bar{y}}^{(1)}(n) \\ \underline{\bar{y}}^{(2)}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 \underline{\tilde{s}}^{(1)}(n) + \mathbf{D}_2 \underline{\tilde{s}}^{(2)}(n) \\ -\mathbf{D}_1 \underline{\tilde{s}}^{(2)}(n) + \mathbf{D}_2 \underline{\tilde{s}}^{(1)}(n) \end{bmatrix} + \begin{bmatrix} \underline{v}^{(1)}(n) \\ \underline{v}^{(2)}(n) \end{bmatrix} \tag{4.24}$$

For simplification, we only see the mth data part and omit noise.

$$\begin{aligned} &y_m^{(1)}(n) = m\text{th data of } \underline{\bar{y}}^{(1)}(n),\ y_m^{(2)}(n) = m\text{th data of } \underline{\bar{y}}^{(2)}(n) \\ &s_m^{(1)}(n) = m\text{th data of } \underline{\tilde{s}}^{(1)}(n),\ s_m^{(2)}(n) = m\text{th data of } \underline{\tilde{s}}^{(2)}(n) \\ &H_1(r_m) = \mathbf{D}_1(m,m),\ H_2(r_m) = \mathbf{D}_2(m,m) \end{aligned} \tag{4.25}$$

Then we get

$$\begin{aligned} y_m^{(1)}(n) &= H_1(r_m) s_m^{(1)}(n) + H_2(r_m) s_m^{(2)}(n) \\ y_m^{(2)}(n) &= -H_1(r_m) s_m^{(2)}(n) + H_2(r_m) s_m^{(1)}(n) \end{aligned} \tag{4.26}$$

In the following, our purpose is to get $H_1^2(r_m)$ and $H_2^2(r_m)$, the square of the channel frequency response. Square (4.26) and note that $\left(s_m^{(1)}(n)\right)^2 = \left(s_m^{(2)}(n)\right)^2 = 1$ for BPSK

$$\begin{aligned} \left(y_m^{(1)}(n)\right)^2 &= H_1^2(r_m) + H_2^2(r_m) + 2H_1(r_m)H_2(r_m) s_m^{(1)}(n) s_m^{(2)}(n) \\ \left(y_m^{(2)}(n)\right)^2 &= H_1^2(r_m) + H_2^2(r_m) - 2H_1(r_m)H_2(r_m) s_m^{(1)}(n) s_m^{(2)}(n) \end{aligned} \tag{4.27}$$

Here we adopt BPSK since we can have $\left(s_m^{(1)}(n)\right)^2 = \left(s_m^{(2)}(n)\right)^2 = 1$. However if we adopt any PSK constellation of size *P*, we have to take the power of *P* to Eq. (4.26) to make $\left(s_m^{(1)}(n)\right)^P = \left(s_m^{(2)}(n)\right)^P = 1$. Then it becomes very hard to solve, hence we only focus on BPSK.

By taking their sum and difference to Eq.(4.27), we have

$$\left(y_m^{(1)}(n)\right)^2 + \left(y_m^{(2)}(n)\right)^2 = 2\left(H_1^2(r_m) + H_2^2(r_m)\right) \tag{4.28a}$$

$$\begin{aligned} \left(y_m^{(1)}(n)\right)^2 - \left(y_m^{(2)}(n)\right)^2 &= 4H_1(r_m)H_2(r_m) s_m^{(1)}(n) s_m^{(2)}(n) \\ &= \pm 4H_1(r_m)H_2(r_m) \end{aligned} \tag{4.28b}$$

From Eq. (4.28b)

$$H_2(r_m) = \frac{\left(y_m^{(1)}(n)\right)^2 - \left(y_m^{(2)}(n)\right)^2}{\pm 4H_1(r_m)} \tag{4.29}$$

Squaring both sides of Eq.(4.29), we have

$$H_2^2(r_m) = \frac{\left[\left(y_m^{(1)}(n)\right)^2 - \left(y_m^{(2)}(n)\right)^2\right]^2}{16H_1^2(r_m)} \tag{4.30}$$

Substitute Eq. (4.30) to Eq. (4.28a), then

$$\frac{\left(y_m^{(1)}(n)\right)^2 + \left(y_m^{(2)}(n)\right)^2}{2} = H_1^2(r_m) + \frac{\left[\left(y_m^{(1)}(n)\right)^2 - \left(y_m^{(2)}(n)\right)^2\right]^2}{16H_1^2(r_m)} \tag{4.31}$$

Then, we have

$$16H_1^4(r_m) - 8\left[\left(y_m^{(1)}(n)\right)^2 + \left(y_m^{(2)}(n)\right)^2\right]H_1^2(r_m) + \left[\left(y_m^{(1)}(n)\right)^2 - \left(y_m^{(2)}(n)\right)^2\right]^2 = 0 \tag{4.32}$$

Finally,

$$H_1^2(r_m) = \frac{\left[y_m^{(1)}(n) - y_m^{(2)}(n)\right]^2}{4} \quad \text{or} \quad \frac{\left[y_m^{(1)}(n) + y_m^{(2)}(n)\right]^2}{4} \tag{4.33}$$

Then we can get $H_1^2(r_m)$ from the square of the sum or difference of the received data, so we name this algorithm sum-difference square method. By the same way, we can get $H_2^2(r_m)$ the same form as Eq. (4.33), and we assume that the two channels we use are different, so we can know that if $H_1^2(r_m)$ is one of Eq. (4.33) then $H_2^2(r_m)$ is another one.

The question here is how to know which one of (4.33) is $H_1^2(r_m)$. We only need to compare it with the estimated channel via subspace-based method to see which one has smaller Euclidean distance.

$$H_1^2(r_m) = \arg\min_{H_1^2(r_m)} \left\| H_{est}^2(r_m) - \frac{\left[y_m^{(1)}(n) \pm y_m^{(2)}(n) \right]^2}{4} \right\|^2 \tag{4.34}$$

where $H_{est}(r_m)$ is the estimated channel via subspace-based method. See Fig.4.4 for the Signal-flow graph of finding $H_1^2(r_m)$ and $H_2^2(r_m)$.

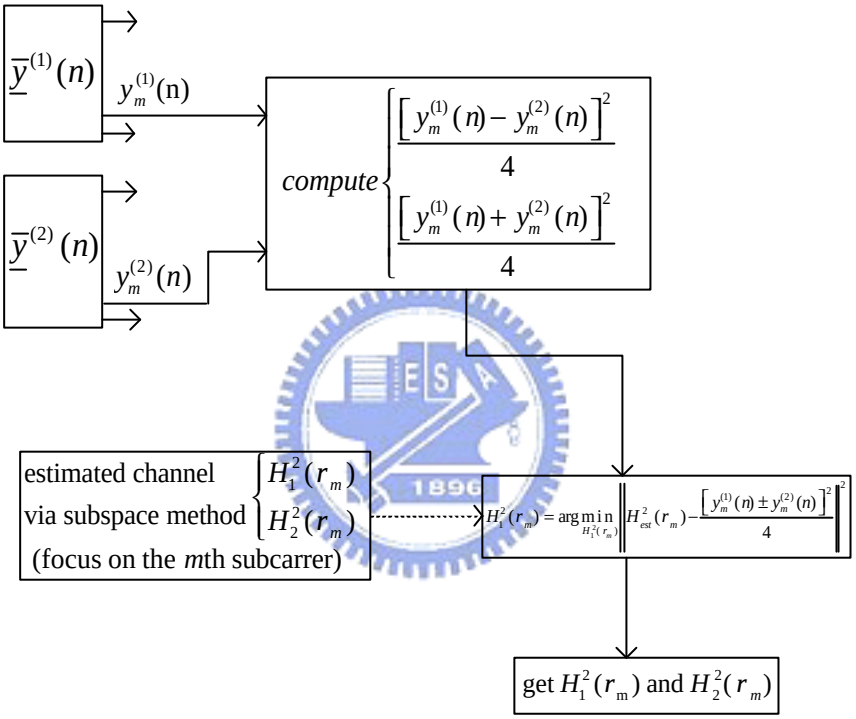


Fig. 4.4 Signal-flow graph of sum-difference square algorithm to find $H_1^2(r_m)$ and $H_2^2(r_m)$

4.2.3 Choice of the window size

We have shown above how to get $H_1^2(r_m)$ and $H_2^2(r_m)$, but they are only corresponding to the two channels' mth subcarrier, so we need to further get all other

subcarrirers' frequency channel power

$$\underline{H}_i^2 = [H_i^2(r_0), H_i^2(r_1), ..., H_i^2(r_{M-1})] \tag{4.35}$$

by using the same method as discussed above. Furthermore, in static channel if we have received more than one data block (assume *N*), for every block *n* we need to get the frequency channel power $\underline{H}_{i,(n)}^2$ and then average all of them to get $\underline{H}_i^2$.

$$\underline{H}_i^2 = \frac{1}{N}\sum_{n=0}^{N-1} \underline{H}_{i,(n)}^2 \tag{4.36}$$

After we've got $\underline{H}_i^2$, simply apply to the Eq. (4.15) by setting *Q=2* and follow the phase direct steps as described earlier to further improve the estimated channel. See Fig.4.4 for the Signal-flow graph of PD on space-time OFDM in static channel.

However, in time-varying channel we do not average all received block to get $\underline{H}_i^2$, instead we need to test how long the window size should be to get the better $\underline{H}_i^2$. As we all know, when the window is longer, we can suppress the noise, but then we can't follow the variance of the channel. This is the trade off. However, the choice of the window size is dependent on how fast the channel changes. See Fig.4.6, signal-flow graph of phase direct on space-time OFDM in time varying channel for the window size is $\frac{N}{4}$.

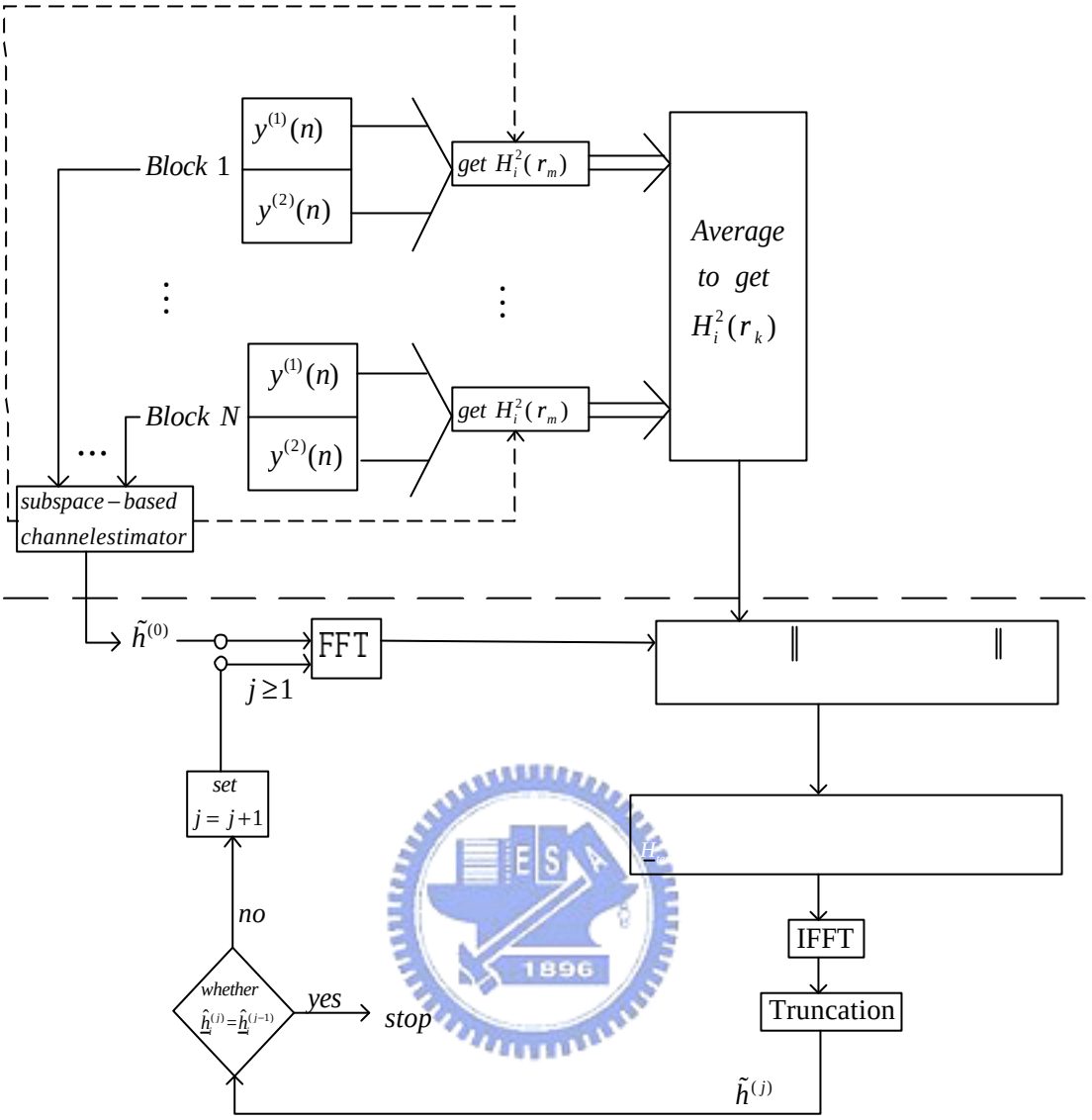


Fig.4.5 Signal-flow graph of phase direct on space-time OFDM in static channel

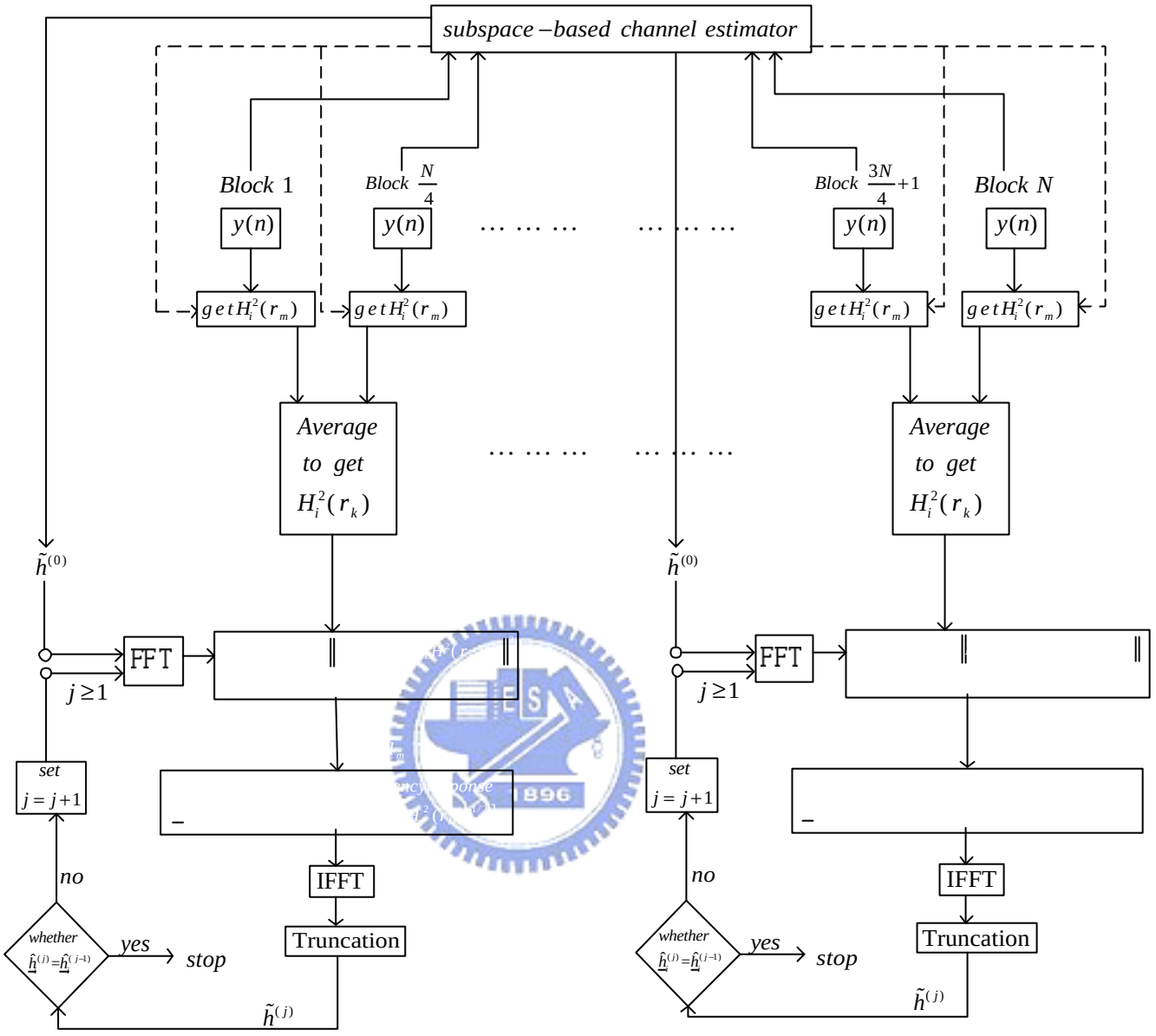


Fig.4.6 Signal-flow graph of phase direct on space-time OFDM in time varying channel

4.2.4 Precoder design

The above algorithm is based on $s_m^{(1)}(n), s_m^{(2)}(n) = \pm 1$, and from Fig.2.1 we know

that $\underline{\tilde{s}}^{(1)}(n), \underline{\tilde{s}}^{(2)}(n)$ are the precoder output, and if we apply the precoder at will, then $\underline{\tilde{s}}^{(1)}(n), \underline{\tilde{s}}^{(2)}(n)$ may not equal to ± 1. So there should be some constrains on the precoder, we simply use the precoder like $[\mathbf{I}_{pre}^T \ \mathbf{I}_k^T]^T$ where $\mathbf{I}_{pre}$ is formed by any *M-K* rows of the $K \times K$ identity matrix $\mathbf{I}_k$ just as the form of Fig 4.7.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \Big\} M-K \text{ rows} \qquad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Big\} M-K \text{ rows}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Big\} K \text{ rows} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \Big\} K \text{ rows}$$

Fig. 4.7 Forms of precoder

In this chapter, we have shown how PD and DD improve the subspace based channel estimator. In next chapter, we want to have some computer simulation to verify our algorithms.

Chapter 5

Computer simulations

In this chapter, we will use computer simulations to verify the algorithm discussed in chapter 3 and chapter 4. We first test the subspace-based estimator discussed in chapter 3 in section 5.1, then see how DD and PD perform in section 5.2. Next we want to see how these methods perform in time varying channel in section 5.3.

We illustrate the performance of our channel estimators through simulations. The figure of the performance for channel estimation is the normalized mean-squares channel error (NMSCE) defined in the frequency domain as:

$$\frac{\| \underline{h} - \hat{\underline{h}} \|^2}{\| \underline{h} \|^2} = \frac{\| \Delta \underline{h} \|^2}{\| \underline{h} \|^2} \tag{5.1}$$

5.1 Subspace-based method

We test our Subspace-based estimator in this section as a function of the input SNR in 5.1.1 and the number of received data blocks in 5.1.2.

Note that there is a complex scalar ambiguity inherent in the blind channel estimator. During the simulations, the power ambiguity is handled by assuming the true channel vector to unit norm and similarly normalizing the estimate. Without further processing, the phase ambiguity cannot be resolved. In our work, this phase ambiguity is determined from $h_i(0)/\hat{h}_i(0)$ and used to compensate the channel estimate prior to the NMSCE computations.

5.1.1 Estimator error V.S. SNR

In this section, we examine the estimator error described in chapter 3 as a function of the input SNR by using the following setup:

- BPSK system,
- K=24, J=32
- Rayleigh fading channel
- L=4 (five-ray channels),
- N=100

The simulation result of estimator error is shown below and is compared with the results from numerical analysis in Eq.(3.46). Solid line stands for the simulation result and dash line is the theoretical result. It shows good agreement of NMSCE (when SNR>20 dB) obtained from simulation and Eq.(3.46) since Eq.(3.46) is based on the assumption of high SNR condition(small perturbation) [19].

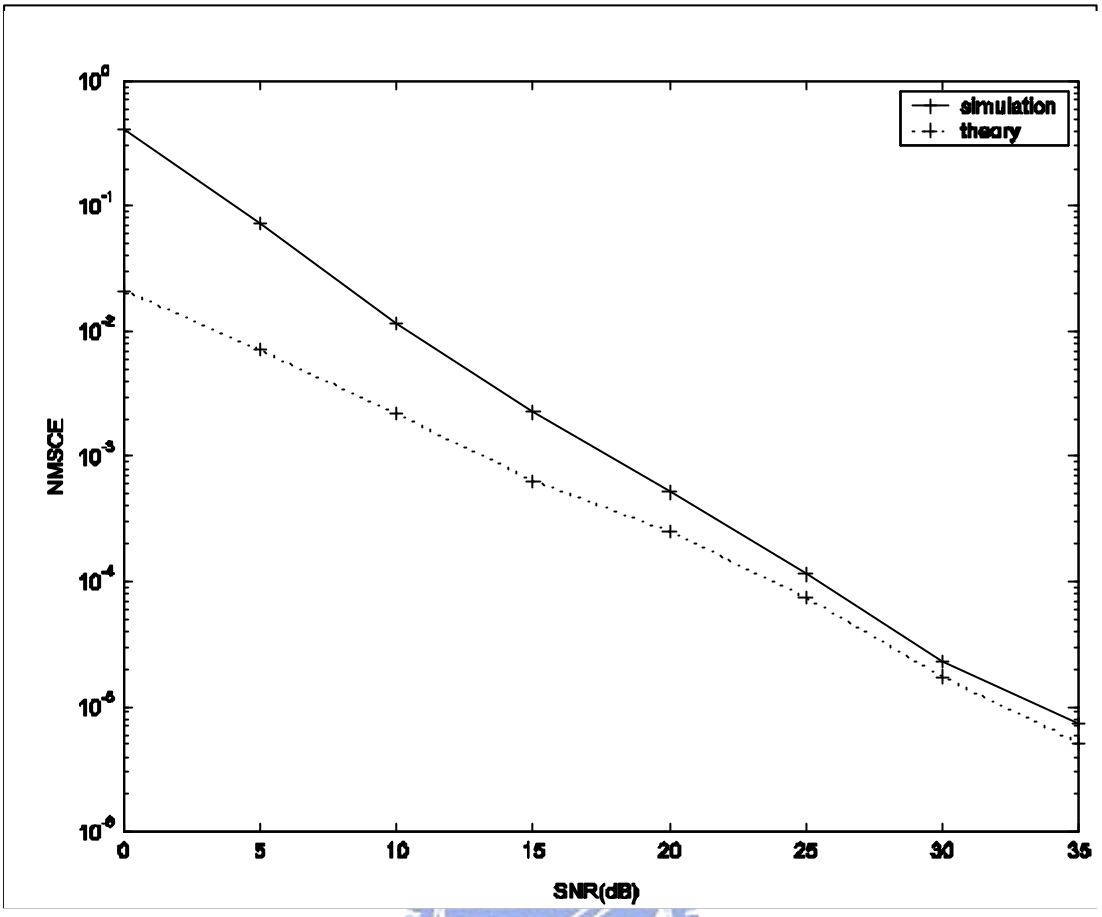


Fig.5.1. Channel error of simulation result and theory V.S *SNR*

5.1.2 Estimator error v.s. data block length

Following 5.1.1, here we illustrate the estimator error as a function of the number of received data blocks for *SNR*=15 dB by using the same setup as section 5.1.1.

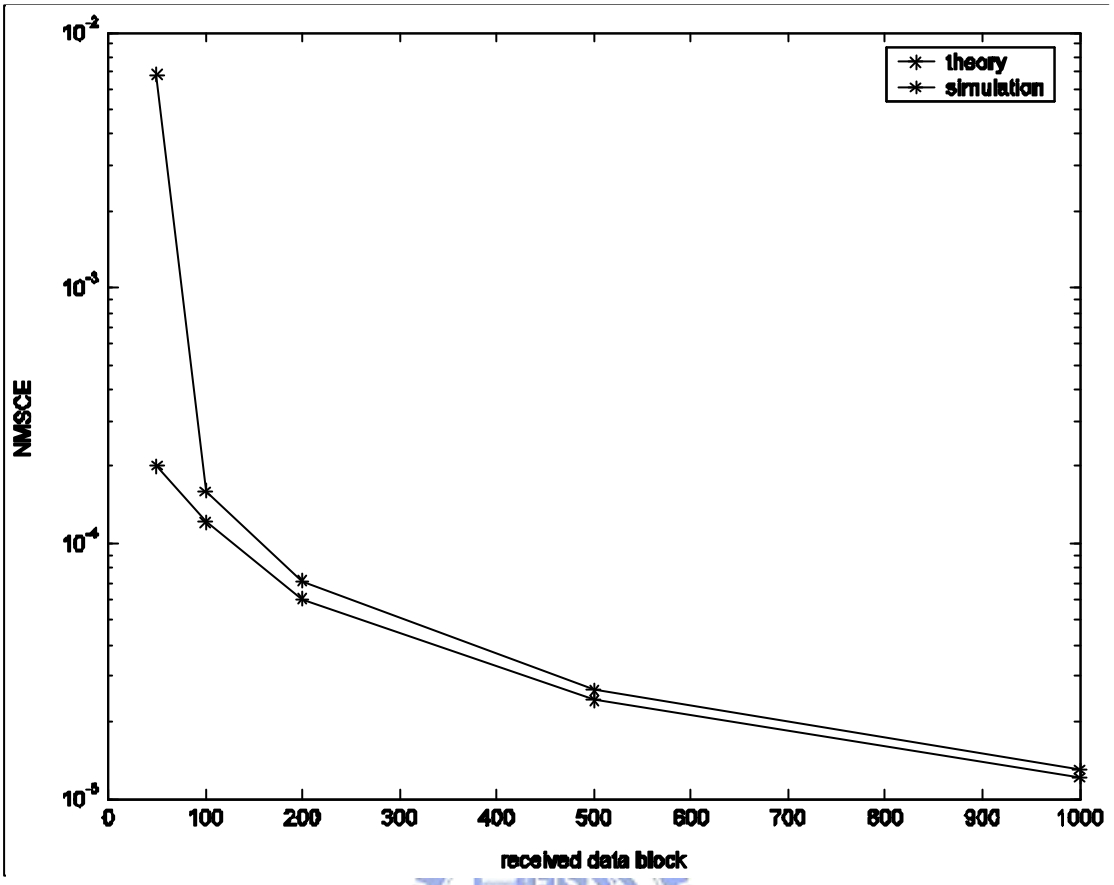


Fig.5.2. Channel error of simulation result and theory V.S. received data block

In Fig. 5.2, we can see that the estimator's performance is better as the received data block is more since the covariance matrix is much closer to the ideal one, and we can see that when the number of received data blocks is 50, the NMSCE diverges obviously since in this situation the covariance we've got is not precise enough.

We can make a conclusion from Fig.5.1 and Fig.5.2 that the estimator's performance is dependent on the key system parameters—such as the input SNR, the length of received data blocks.

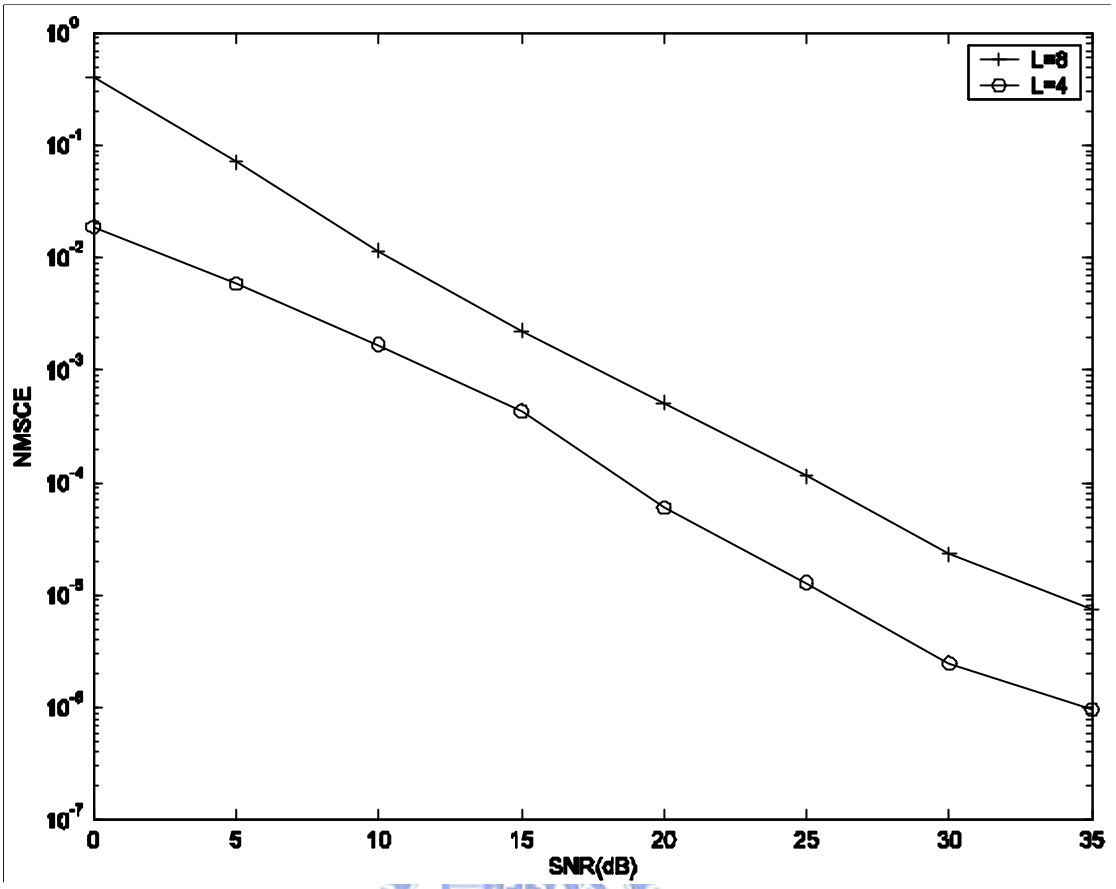


Fig.5.3. Channel error of different multipath length

In Fig.5.3, we test the subspace-based channel estimator for different multipath length. As expected, as the multipath length is larger, the performance is worse.

5.2 Performance of PD and DD

First we want to show how DD and PD work with and without denoising.

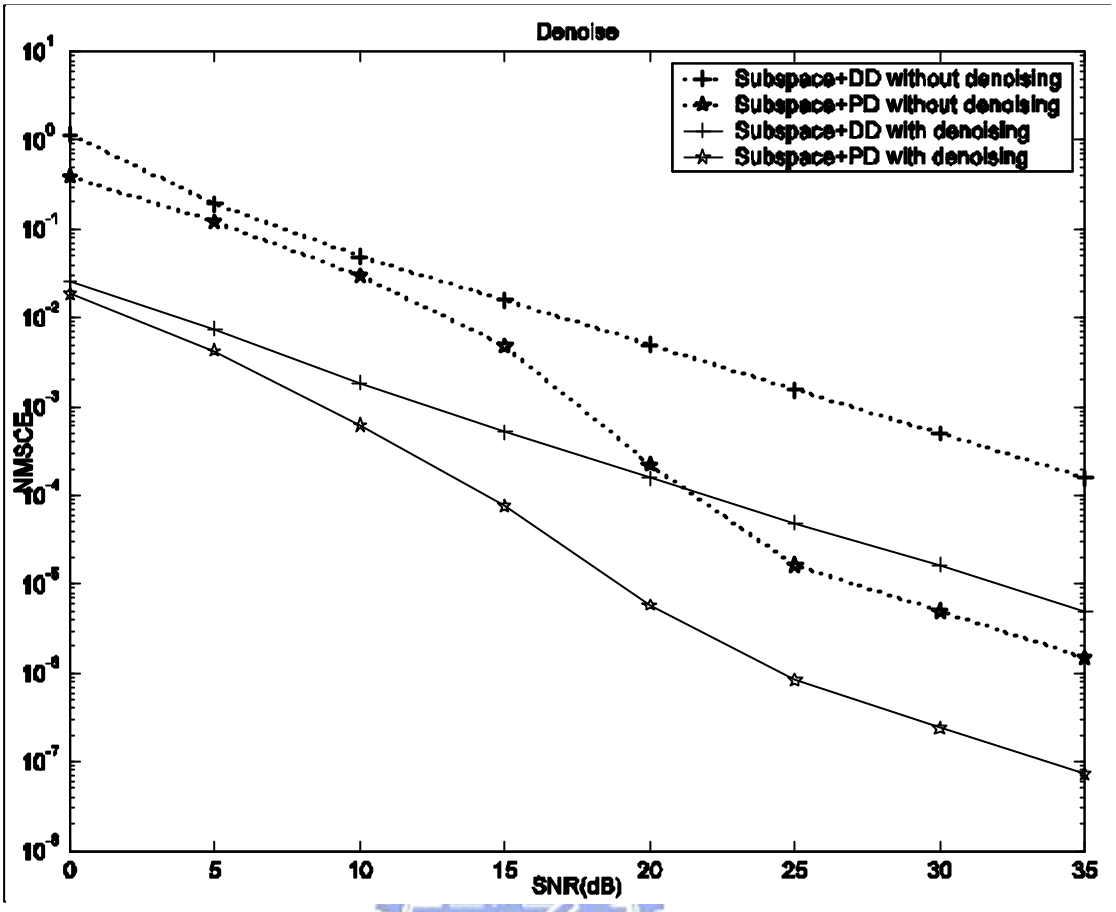


Fig.5.4. Channel error of PD and DD with denoising and without denoising

With denoising, the performance improves since we truncate the last M-L point.

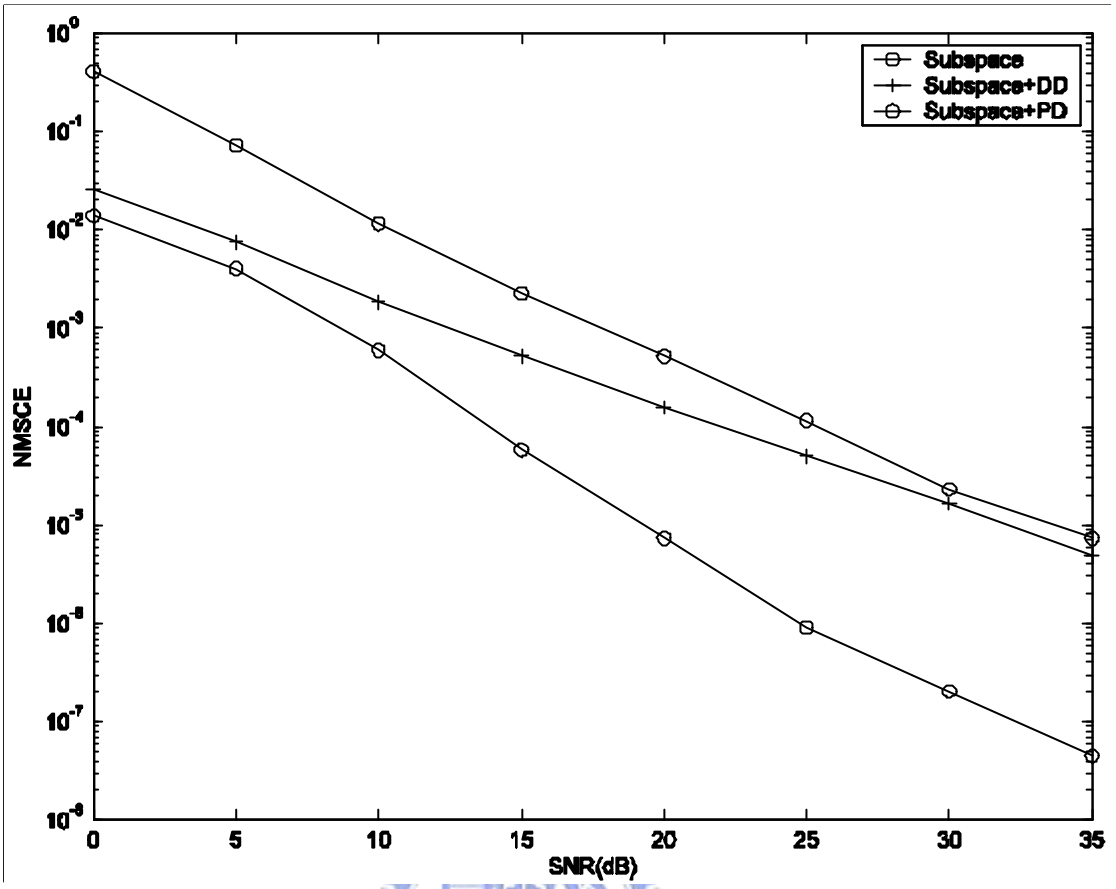


Fig.5.5 PD and DD v.s subspace-based estimator

From Fig.5.5, we can see that DD and PD with denoisng better the subspace-based estimator. Moreover, PD improves significantly.

Next we want to see DD and PD in different multipath length. In Fig.5.6, both DD and PD are immune to multipath length.

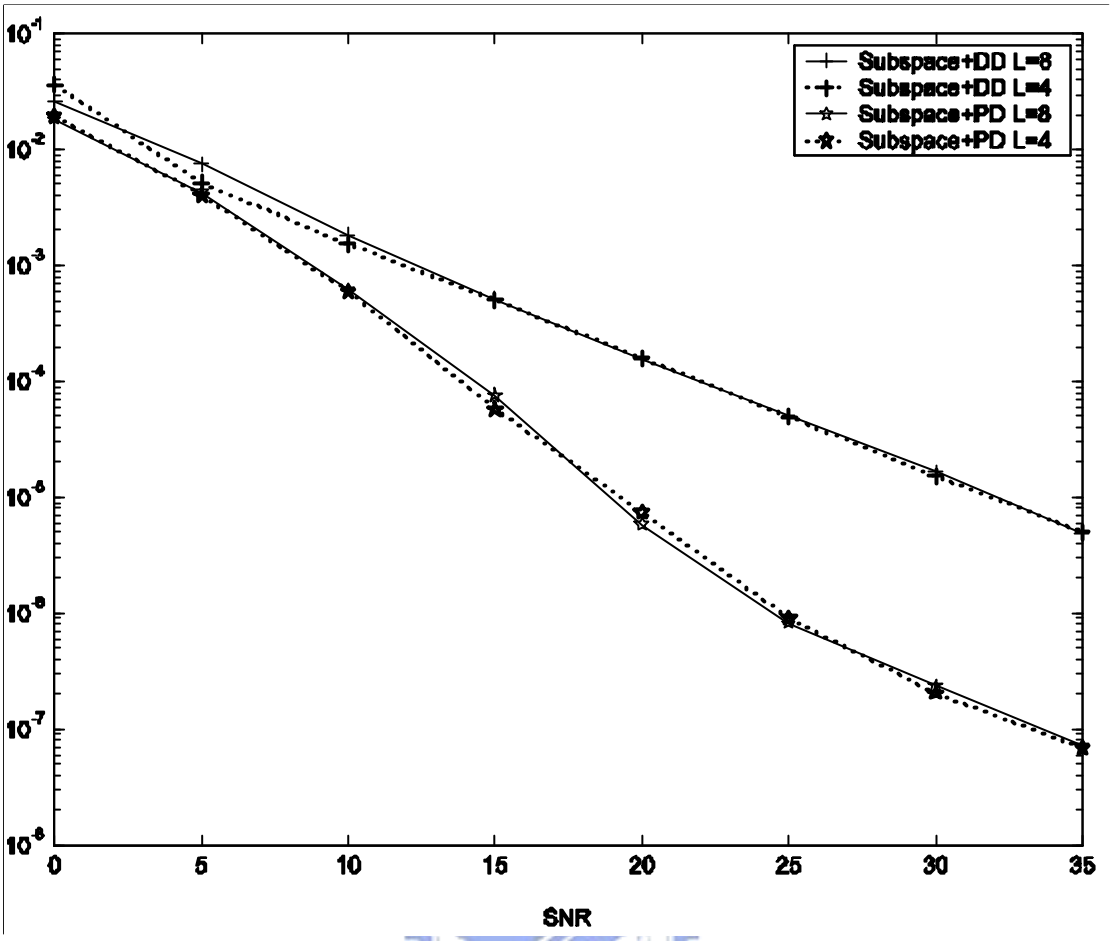


Fig.5.6. Testing DD and PD initialized with subspace method for different multipatj length

In the above, we use BPSK system to simulate both DD and PD. Here we want to test DD in different data constellation, such as QPSK, 16QAM, and 64QAM.

We can see from Fig.5.7 that both BPSK and QPSK improve the performance. However, 16QAM and 64QAM worsen the performance since it is easy to make wrong decision for them.

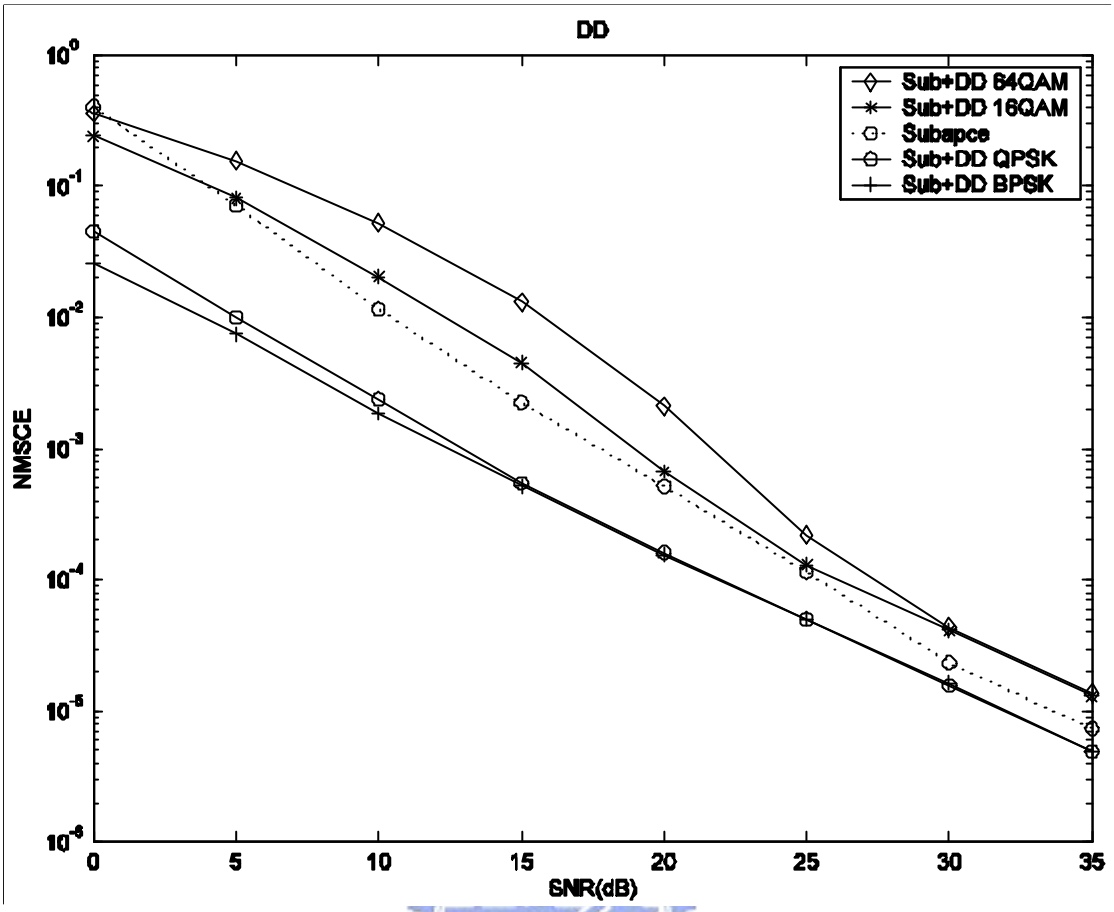


Fig.5.7. Testing DD in different data constellation

5.3 Time-varying channel estimation

5.3.1 Subspace-based method

Here, we test our proposed method in time variant channel to see how they behave. Each tap of the time-varying FIR channels varies according to Jakes' model, and the sample rate is 1 MHz.

First, the subspace-based channel estimator is tested and shown in Fig.5.8 in time varying channel for different maximum Doppler frequencies, which is equal to 10Hz, 60Hz, 100Hz, and 200Hz. We can see that as the maximum Doppler frequency is higher, the estimator is worse and all of them lead to error floor.

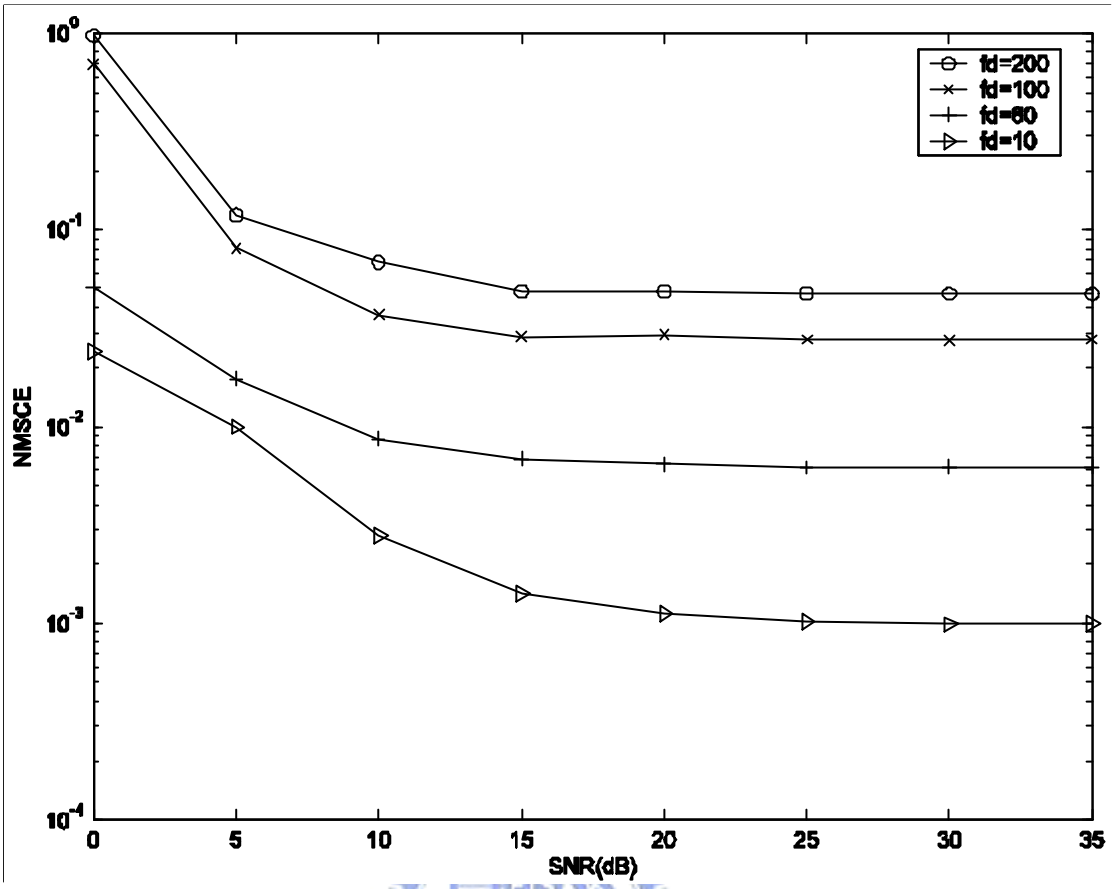


Fig.5.8. Testing of the subspace-based channel estimator in different Doppler frequencies for received blocks equaling to 100

In the above, we use received blocks equaling to 100, however in the 100 blocks the channel is already different, hence next we want to test the situation, which data block equal to 50.

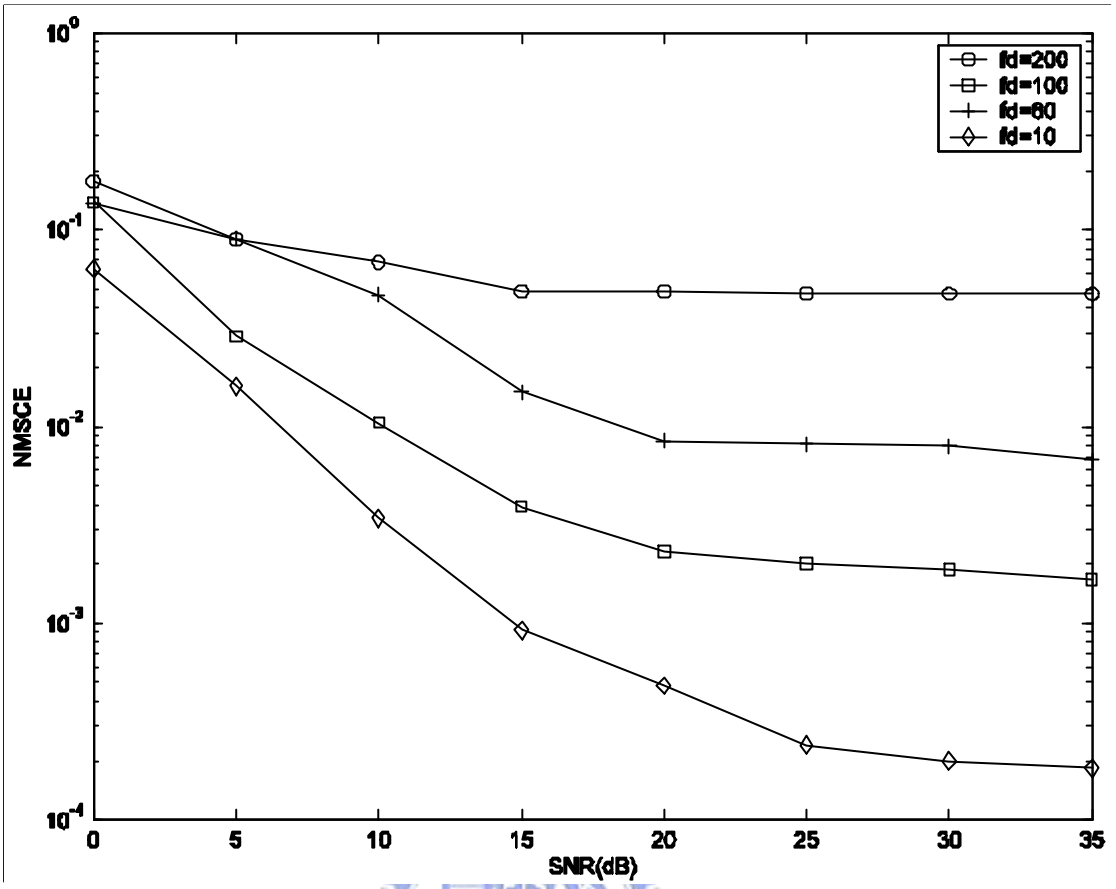


Fig.5.9. Testing of the subspace-based channel estimator in different Doppler frequencies for received blocks equaling to 50

Comparing Fig.5.8 and Fig.5.9, we can see that Fig.5.9 has got better performance since the received data blocks is less. As we all know, the estimator is to get a channel from the received data which have the information of channels, which best suits all the channels. However, the channels are all different. Hence, more blocks leads to poorer performance.

Fig.5.10 show the NMSCE for each block in the 50 received data blocks in fd=60Hz and SNR=15dB. As we can see, around the 25^{th} block, we can get best performance since the channel can be seen as changing as a straight line in the 50 block, the estimator is about to estimate one channel which has the smallest difference

to all the 50 blocks. Hence we' ve got this result.

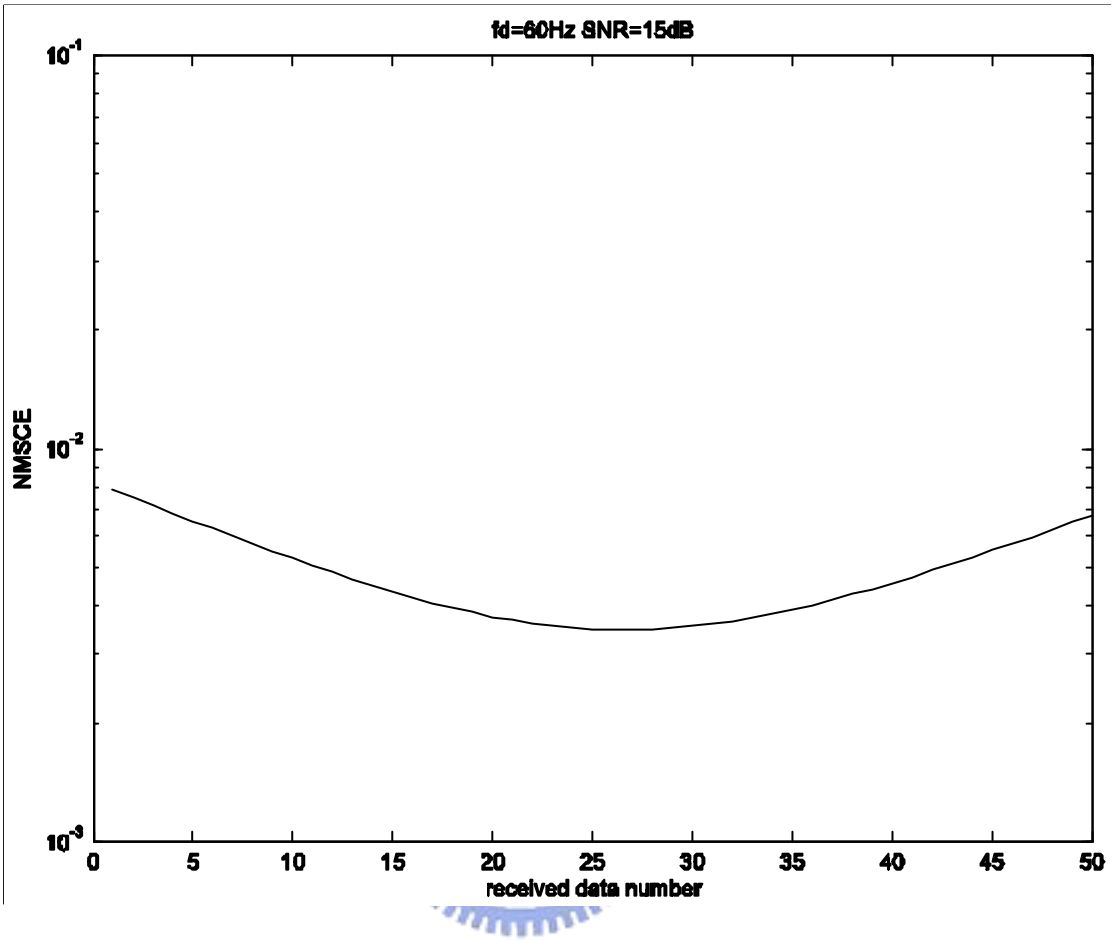


Fig.5.10 Channel error for each block

5.3.2 Performance of PD and DD

DD and PD are tested in the following. For PD we do not average all received block to estimate $H_i^2(r_k)$, instead we need to test how long the window size should be to get the better estimate, which is because as the window is longer, we can suppress more noise, but then we can't follow the variation of the channel.

We can see in Fig.5.11 that both PD and DD improve the performance moreover they resolve the error floor problem occurred in subspace-based estimator. Moreover,

in PD, we can see that “window size=1” outperforms, which is because in fd=50 Hz, the channel change fast, we need to trace the variation by using smaller window size.

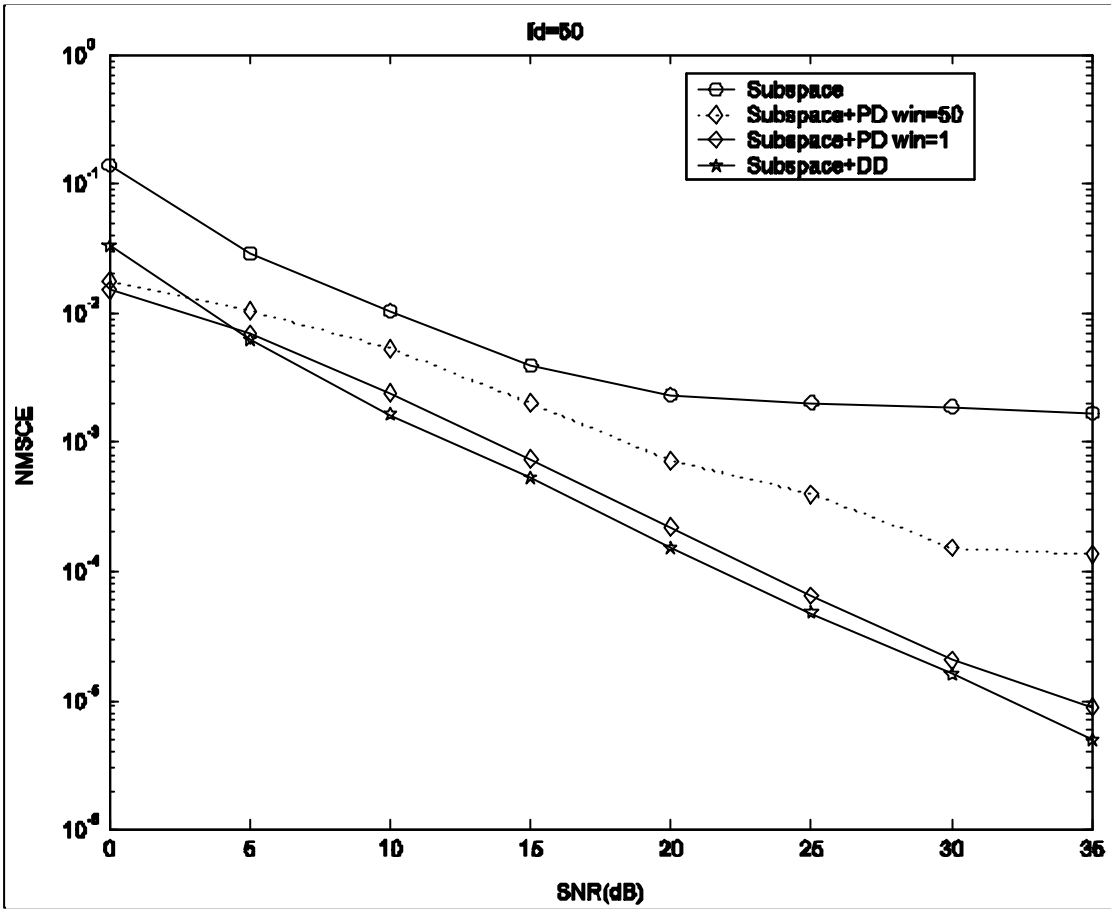


Fig.5.11 Test of DD and PD in fd=100Hz

Performance of DD in different data constellation is shown in Fig.5.12.We can see that BPSK, QPSK, and 16QAM improve the performance. However, 64QAM worsen the performance in low SNR since it is easy to make wrong decision.

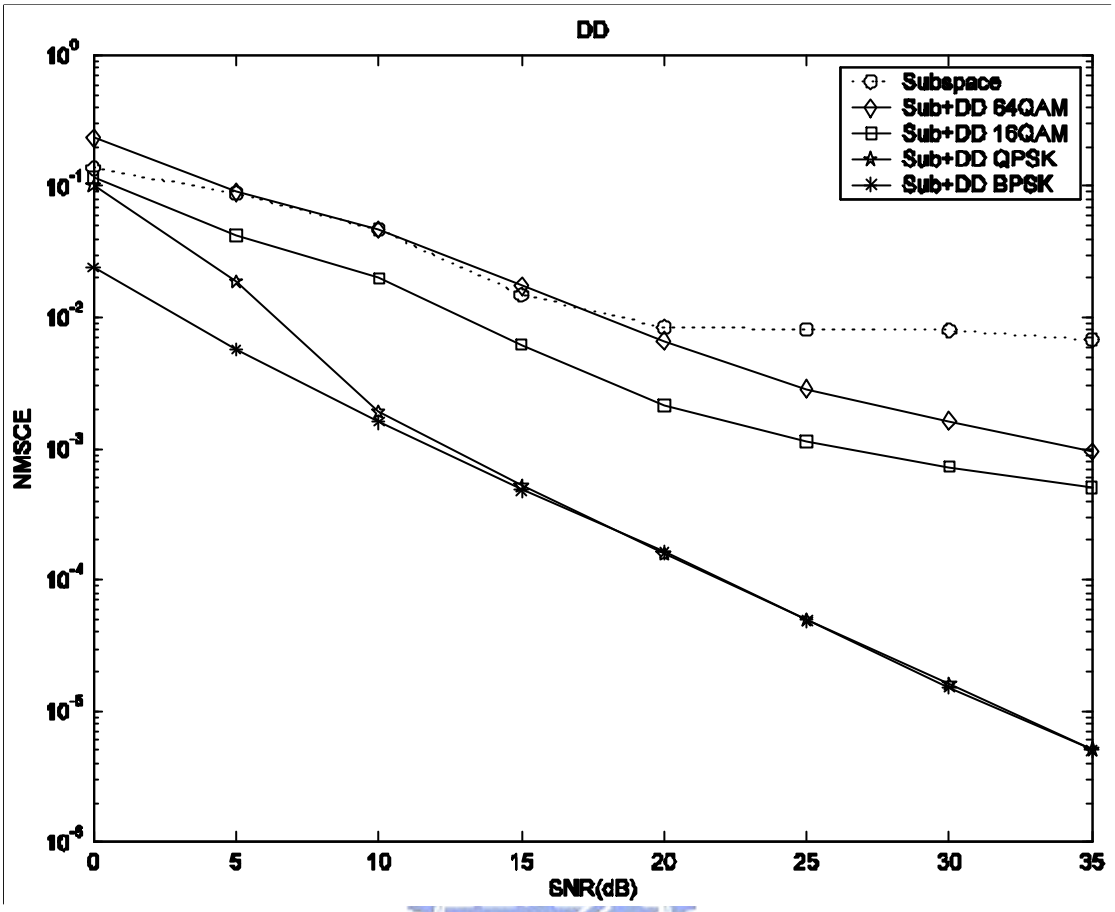
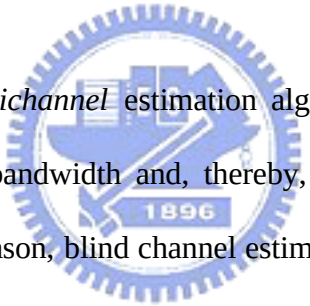


Fig.5.12 Test of DD for different data constellation in fd=50

Chapter 6

Conclusion



For STC transceivers, *multichannel* estimation algorithms are needed. However, training sequences consume bandwidth and, thereby, incur spectral efficiency (and thus capacity) loss. For this reason, blind channel estimation methods receive growing attention. We have shown a subspace-based blind channel estimation algorithm for ST OFDM transmissions and develop the theoretical mean square error of the estimator. To further improve the channel estimation, we can exploit the finite alphabet property to better the channel estimates. We discuss two different methods, DD and PD, and apply them to ST OFDM system. DD, as implied in the name, needs first to get the hard decision data and then use it to update our estimated channel, while PD is to solve the phase ambiguities after we've got the channel power response. However, in ST OFDM the channel power response is hard to get since the received data is composed of two different transmitted data. Hence, we only focus on BPSK system and exploit the transmitted data's time and temporal correlation to develop a sum-difference square algorithm to solve this problem. Moreover, in time-varying

channel we can choose a best window size to get the channel power. Computer simulations have shown that the PD and DD method really improve the NMSCE in static channel and time-varying channel and further resolve the error floor problem occurred in time-varying channel.

However, our PD method for Space-Time OFDM is only utilized in BPSK system, we shall try to apply it to other systems such as QPSK, QAM in the future.

Bibliography

[1] Marc Engels, ”Wireless OFDM systems ~How to make them work?” Kluwer academic publishers, 2002.

[2] John Terry and Juha Heuskala ”OFDM wireless LAN: a theoretical and practical guide ” 201 West 103rd St. Indianapolis, Indiana, USA, 2002.

[3] Richard van Nee, Ramjee Presad “OFDM wireless multimedia communications” Artech House Boston London, 2000

[4] Shengli Zhou, Georgios B. Giannakis “Space time coding with maximum diversity gains over frequency-selective fading channels” *IEEE Signal Processing Letters*, vol 8, pp.269-272, October 2001

[5] D. Mihai Ionescu, “On Space–Time Code Design” *IEEE Trans. on Wireless Communications*, vol. 2, pp. 20-28, January 2003

[6] Z. Liu and Georgios. B. Giannakis, “Space-time block coded multiple access through frequency-selective fading channels,” *IEEE Trans. Communications*, vol. 49, pp. 1033–1044, June 2001.

[7] Z. Liu, G. B. Giannakis, B. Muquet, and S. Zhou, “Space-time coding for broadband wireless communications,” *Wireless Communications and Mobile Computing*. *New York: Wiley*, vol. 1, pp.33–53, January–March 2001,

[8] Li, Justin C. Chuang, and N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high-rate data wireless networks " *IEEE J. Select. Areas Communications*, vol. 17, pp.1233-1243, July 1999

[9] Ye , Nambirajan Seshadri, and Sirikiat Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels" *IEEE J. Select. Areas Communications*, vol. 17, pp.461-471, March 1999

[10] Yongzhe Xie, and Costas N. Georghiades, " Two EM-Type channel estimation algorithms for OFDM with transmitter diversity" *IEEE Trans. Communications*, vol. 51, pp.106-115, January 2003

[11] S. M. Alamouti, "A simple transmit diversity technique for wireless communications ," *IEEE J. Select. Areas Communications*, vol. 16, pp.1451–1458, October 1998.

[12] B. Hassibi and B. Hochwald " How much training is needed in multiple-antenna wireless links?." *IEEE Trans. Inform. Theory*, vol. 49, pp.951-963, April 2003

[13] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Communications*, vol. 17, pp. 461–471, March 1999.

[14] H. Bölcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," *IEEE Trans. Signal*

Processing, vol. 50, pp. 96–109, January 2002.

[15] Yonghong Zeng, and Tung-Sang Ng, "Semi-Blind channel estimation method for multiuser multiantenna OFDM Systems" *IEEE Trans. Signal Processing*, vol 2, pp.1419-1429, May 2004

[16] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Transmit antenna space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE J. Select. Areas Communications*, vol. 19, pp.1352–1364, July 2001.

[17] Shengli Zhou, and Georgios B. Giannakis, "Subspace-Based (Semi-) Blind Channel Estimation for Block Precoded Space-Time OFDM ," *IEEE Trans. on signal processing*, vol. 50, pp.1215-1228, May 2002

[18] G. H. Golub and C.F. Van Loan, "*Matrix computations*" John Hopkins University Press, 3rd edition, 1996

[19] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: further unification, simplification, and observations,"

IEEE Trans. Aerosp., Electron. Syst., vol. 29, pp. 1170–1184, October 1993.

[20] Shengli Zhou, and Georgios B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. on Communications*, vol. 49, pp.1402-1415,August 2001

[21] Shengli Zhou, Georgios B. Giannakis and Anna Scaglione, “Long codes for generalized FH-OFDMA through unknown multipath channels” *IEEE Trans. on Communications,* vol. 49, pp 721-733, April 2001