ROBUST SYNCHRONIZATION AND CHANNEL ESTIMATION FOR MIMO-OFDM SYSTEMS

GAO FEIFEI

NATIONAL UNIVERSITY OF SINGAPORE

2007

ROBUST SYNCHRONIZATION AND CHANNEL ESTIMATION FOR MIMO-OFDM SYSTEMS

GAO FEIFEI

(*M.Eng., McMaster University*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2007

Dedications:

To my family

Acknowledgment

I would like to first thank Dr. Arumugam Nallanathan for his guidance and support throughout the past two and a half years and also thank for his kindly supervision and instruction on my work. His encouragement and patience were essential to the completion of this project.

I thank Dr. Yan Xin for being a great teacher and a friend. Dr. Xin's profound thinking, generosity and integrity will play an inspiring role in my future career. I thank Dr. Meixia Tao for many insightful discussions on the subject of space time coding and cooperative communications. I am deeply stimulated by her enthusiasm and integrity on research working.

I would like to thank Prof. Yide Wang in Ecole Polytechnique of University of Nantes, France, Dr. Yonghong Zeng in I^2R A-STAR, Singapore, Prof. Chintha Tellambura in University of Alberta, Canada, and Tao Cui in California Institute of Technology, USA, with whom I have had the good fortune to collaborate. Especial thank should be presented to Tao Cui from whom I have benefited a lot with hours of stimulating discussions and I also owe him a great deal for his friendship.

I am also fortunate to be in a research group whose members are always kind and have taught me many living tips in Singapore. The group members include Jinhua Jiang, Lan Zhang, Jianwen Zhang, Le Cao, Wei Cao, Yong Li, Yan Li, Yonglan Zhu, Qi Zhang, Jun He, Lokesh Bheema Thiagarajan, Hon Fah Chong, Anwar Halim and many others.

Last but not least, I would like to thank my parents for their love and support which played an instrumental role in the completion of this project.

Contents

Summary

The combination of multiple-input multiple-output (MIMO) transmission with orthogonal frequency division multiplexing (OFDM) technique is deemed as the candidate to the upcoming fourth generation (4G) wireless communication systems. This thesis addresses several initialization issues for MIMO OFDM systems. We answer the following questions: how to use a few pilot carriers to track the timing offset (TO) and the carrier frequency offset (CFO), how to apply the blind channel estimation when the number of the transmit antennas is greater than or equal to the number of receive antennas, how can we make the blind channel estimation more robust to parameter uncertainty. All these questions are interesting yet never answered or partly answered through the existing literatures.

Three main contributions are built from this thesis: First, a CFO tracking algorithm is developed by utilizing the scatter pilot tones (PT) and the virtual carriers (VC). The method not only shows the compatibility with most OFDM standards but also provides improved performance compared to the existing works. Furthermore, the algorithm is feasible for the synchronization initialization. Second, a robust re-modulation on MIMO OFDM is proposed such that the channel matrix possesses exciting properties. For example, the blind channel estimation after the system re-modulation is robust to the channel order over-estimation, and the channel estimation identifiability is guaranteed for random channel realization. Moreover, the method is applicable for MIMO OFDM systems with equal number of transceiver antennas, which is compatible to existing single-input single-output (SISO) OFDM standards and the upcoming 4G OFDM standards. Third, by

applying a non-redundant precoding, it is shown that the blind channel estimation is applicable even for the case where the number of the transmit antennas is greater than the number of receive antennas, e.g. multiple-input single-output (MISO) transmissions. This method exhibits great potential to be applied in the uplink cellular systems and the currently arising cooperative communications where there are, in general, multiple relays but one destination only.

List of Tables

List of Figures

List of Acronyms

OFDM	Orthogonal Frequency Division Multiplexing
4G	Fourth Generation
DAB	Digital Audio Broadcasting
DVB	Digital Video Broadcasting
HIPERLAN	High Performance Radio Local Area Network
SISO	Single-Input Single-Output
SIMO	Single-Input Multiple-Output
MISO	Multiple-Input Single-Output
MIMO	Multiple-Input Multiple-Output
TO	Timing Offset
CFO	Carrier Frequency Offset
TOFR	TO Failure Rate
CFOOP	CFO Outlier Probability
ML	Maximum Likelihood
AWGN	Additive White Gaussian Noise
DFT	Discrete Fourier Transform
IDFT	Inverse Discrete Fourier Transform
FFT	Fast Fourier Transform
SOS	Second Order Statistics
EVD	Eigenvalue Decomposition
SVD	Singular Value Decomposition
ISI	Inter-Symbol Interference

ICI	Inter-Carrier Interference
IBI	Inter-Block Interference
CSI	Channel State Information
CRB	Cramér-Rao Bound
ACRB	Approximated CRB
FIM	Fisher Information Matrix
MRC	Maximum Ratio Combining
SNR	Signal-to-Noise Ratio
BER	Bit Error Rate
CP	Cyclic Prefix
ZP	Zero Padding
MSE	Mean Square Error
MMSE	Minimum Mean Square Error
NMSE	Normalized Mean Square Error
RNMSE	Root Normalized Mean Square Error
ZF	Zero Forcing
BPSK	Binary Phase Shift Keying
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
PDF	Probability Density Function
STC	Space Time Coding
SD	Sphere Decoding
IEEE	Institute of Electrical and Electronics Engineers

List of Notations

a	lowercase letters are used to denote scalars
$\mathbf{a}$	boldface lowercase letters are used to denote column vectors
$\mathbf{A}$	boldface uppercase letters are used to denote matrices
$(\cdot)^T$	the transpose of a vector or a matrix
$(\cdot)^*$	the conjugate of a scalar or a vector or a matrix
$(\cdot)^H$	the Hermitian transpose of a vector or a matrix
$(\cdot)^{-1}$	the inversion of a matrix
$(\cdot)^\dagger$	the pseudo inverse of a matrix
$[\cdot]_{pq}$	the (p,q)th element of a matrix
$\lvert\cdot\rvert$	the absolute value of a scalar or the cardinality of a set
$\lVert\cdot\rVert$	the Euclidean norm of a vector
$\lVert\cdot\rVert_F$	the Frobenius norm of a matrix
$tr(\cdot)$	the trace of a matrix
$\mathrm{vec}(\cdot)$	the vectorization of a matrix
$\mathrm{diag}\{\mathbf{a}\}$	the diagonal matrix with the diagonal element built from $\mathbf{a}$
$\mathrm{E}\{\cdot\}$	the statistical expectation operator
$\angle(\cdot)$	the angle of a scalar
$\Re\{\}$	the real part of the argument
$\Im\{\}$	the imaginary part of the argument
$\otimes$	the Kronecker product
$\odot$	the Hadamard product

Chapter 1

Introduction

In this chapter, we provide overviews for OFDM systems, MIMO channels, as well as their integration—MIMO OFDM systems. We also briefly introduce initialization issues of the OFDM based transmission. In the end, we present our goals and list major contributions of this project.

1.1 Overview of OFDM

1.1.1 History of OFDM

The history of OFDM could be traced back to the mid 60's, when Chang presented his idea on the parallel transmissions of bandlimited signals over multi-channels [1]. He developed a principle for transmitting messages simultaneously through orthogonal channel that is free of both inter-channel interference (ICI) and inter-symbol interference (ISI).

Five years later, a breakthrough was made by Weinstein and Ebert who used the inverse discrete Fourier transform (IDFT) to perform base band modulation and used discrete Fourier transform (DFT) for the demodulation [2]. This model eliminates the need of subcarrier oscillator banks, and the symbols can be transmitted directly after the IDFT transform rather than being transmitted on different subcarriers. To this end, the physical meaning of OFDM, namely, signals

are transmitted through different frequency sub-bands, disappears. Nonetheless, the processing efficiency is greatly enhanced thanks to the development of fast Fourier Transform (FFT) algorithm. To combat ICI and ISI, Weinstein and Ebert used both guard space and raised cosine windowing in the time domain. Unfortunately, such an system could not obtain perfect orthogonality among subcarriers over a multi-path channel.

Another important contribution was made by Peled and Ruiz in 1980 [3], who suggested that a cyclic prefix (CP) that duplicated last portion of an OFDM block be inserted in the front the same OFDM block. This tricky way solves the orthogonality problem in the dispersive channel. In fact, as long as the cyclic extension is longer than the impulse response of the channel, the linear convolution between the channel and the data sequence becomes the cyclic convolution, which implies the perfect orthogonality among sub-channels. Although this CP introduces an energy loss proportional to the length of the CP, the orthogonality among sub-channels normally motivates this loss.

Currently, CP based OFDM is enjoying its success in many applications. It is used in European digital audio/video broadcasting (DAB, DVB) [4], [5], high performance local radio area network (HIPERLAN) [6], IEEE 802.11a wireless LAN standards [7], any may others. In fact, OFDM is also a fundamental technique that is adopted in the future fourth generation (4G) wireless communications [8], [9].

1.1.2 System Model of OFDM

The basic idea of OFDM is to divide the frequency band into several over-lapping yet orthogonal sub-bands such that symbols transmitted on each sub-band experiences only flat fading, which brings much lower computational complexity when performing the maximum likelihood (ML) data detection. A modern DFT based OFDM achieves orthogonality among sub-channels directly from the IDFT and the CP insertion. An example of such a block structure is shown in Fig. 1.1 [10]-[12]. Let K denote the number of the subcarriers in one OFDM block

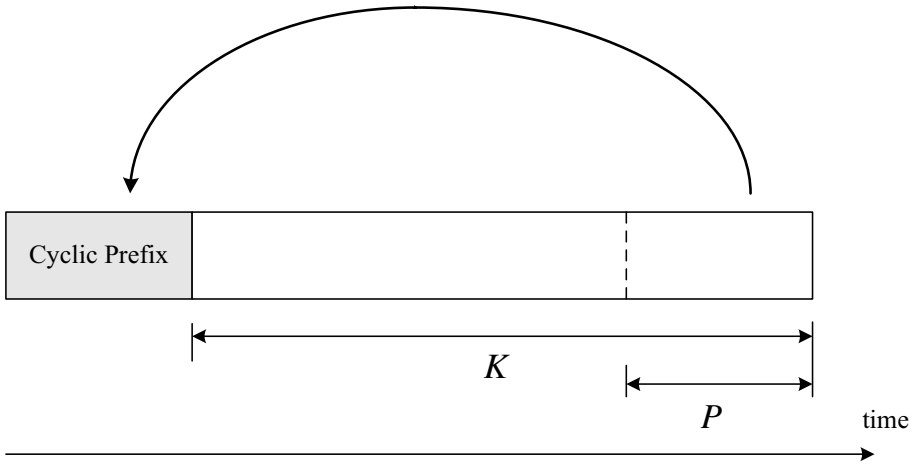

Figure 1.1: The OFDM block structure with cyclic prefix.

and $\mathbf{s}_i = [s_i(0), s_i(1), \ldots, s_i(K-1)]^T$ denote the signal block consisting of K symbols to be transmitted during the ith OFDM block. The time domain signal $\mathbf{x}_i = [x_i(0), x_i(1), \ldots, x_i(K-1)]^T$ is obtained from the IDFT of $\mathbf{s}_i$, which could be expressed as

$$\mathbf{x}_i = \mathbf{F}^H \mathbf{s}_i \tag{1.1}$$

where $\mathbf{F}$ is the normalized DFT matrix with the (a,b)th entry given by $\frac{1}{\sqrt{K}} e^{\frac{j2\pi(a-1)(b-1)}{K}}$. Assume the channel delay τ_h, after being normalized by the sampling interval T_s, is upper bounded by L. Throughout the whole thesis we only consider the constant channel during one frame transmission[1], so the equivalent discrete channel vector is written as $\mathbf{h} = [h(0), h(1), \ldots, h(L)]^T$. The length of CP, denoted by P, should be greater than or equal to L. After the CP insertion, the overall OFDM block of length $K_s = K + P$ is expressed as

$$\mathbf{u}_i = [x_i(K-P), \ldots, x_i(K-1), x_i(0), \ldots, x_i(K-1)]^T = \mathbf{T}_{\text{cp}} \mathbf{x}_i \tag{1.2}$$

where $\mathbf{T}_{\text{cp}}$ is the corresponding CP inserting matrix. The transmitted frame is composed of M consecutive OFDM blocks $\mathbf{u}_0, \ldots, \mathbf{u}_{M-1}$. A linear convolution between the frame and the channel is received at the destination, in the same time

[1]One frame consists of the heading and the information blocks.

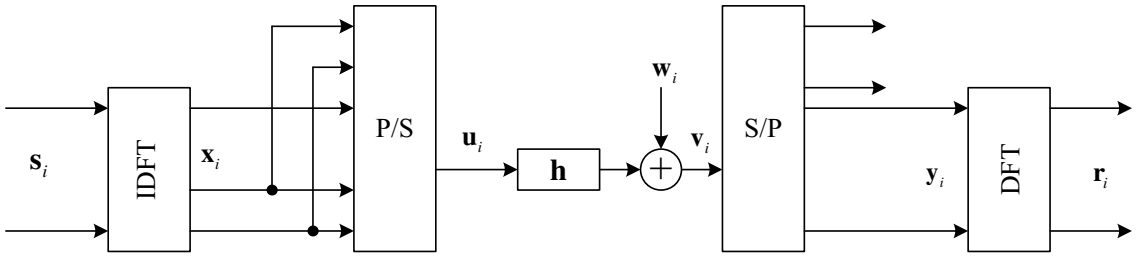

Figure 1.2: A based band OFDM system model.

with additive Gaussian white noise (AWGN) generated by thermal vibrations of atoms in antennas, shot noise, black body radiation from the earth or other warm objects.

A typical base band OFDM system diagram is shown in Fig. 1.2. Mathematically, the ith received block is given by

$$\mathbf{v}_i = \mathbf{H}_0\mathbf{u}_i + \mathbf{H}_1\mathbf{u}_{i-1} + \mathbf{w}_i \tag{1.3}$$

where $\mathbf{w}_i$ is an $K_s \times 1$ vector whose elements represent the AWGNs of variance σ_n^2 and

$$\mathbf{H}_0 = \begin{bmatrix} h(0) & 0 & 0 & \dots & 0 \\ \vdots & h(0) & 0 & \dots & 0 \\ h(L) & \dots & \ddots & \dots & \vdots \\ \vdots & \ddots & \dots & \ddots & 0 \\ 0 & \dots & h(L) & \dots & h(0) \end{bmatrix}, \quad \mathbf{H}_1 = \begin{bmatrix} 0 & \dots & h(L) & \dots & h(0) \\ \vdots & \ddots & 0 & \ddots & \vdots \\ 0 & \dots & \ddots & \dots & h(L) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}. \tag{1.4}$$

The second term $\mathbf{H}_1\mathbf{u}_{i-1}$ forms the so called inter-block interference (IBI). To remove the IBI, the first P elements in $\mathbf{v}_i$ is discarded and the remaining part is denoted by

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i + \mathbf{n}_i = \mathbf{H}\mathbf{F}^H\mathbf{s}_i + \mathbf{n}_i \tag{1.5}$$

where $\mathbf{n}_i$ is the last K elements in $\mathbf{w}_i$ and

$$\mathbf{H} = \begin{bmatrix} h(0) & \dots & 0 & h(L) & \dots & h(1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ h(L-1) & \ddots & h(0) & \dots & \dots & h(L) \\ h(L) & \ddots & \ddots & \ddots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & h(L) & h(L-1) & \dots & h(0) \end{bmatrix} \tag{1.6}$$

is the corresponding circulant channel matrix. The ML detection selects the optimal estimate $\hat{\mathbf{s}}_i$ to minimize the following objective function:

$$\hat{\mathbf{s}}_i = \arg\min_{\mathbf{b}} \|\mathbf{y}_i - \mathbf{H}\mathbf{F}^H\mathbf{b}\|^2 \tag{1.7}$$

where $\mathbf{b}$ is the trial variable whose elements are selected from the signal constellation. Generally, a computationally expensive K-dimensional search should be performed to arrive at the optimal detection.

It is known that any circulant matrix can be diagonalized by the normalized DFT matrix $\mathbf{F}$ [10]; namely, $\mathbf{H} = \mathbf{F}^H\mathbf{\Lambda}\mathbf{F}$ where $\mathbf{\Lambda}$ is a diagonal matrix with the kth diagonal element $\tilde{h}(k)$. Here, $\tilde{h}(k)$ is the kth element of $\tilde{\mathbf{h}}$, and $\tilde{\mathbf{h}}$ is the K-point DFT of $\mathbf{h}$. Applying the normalized DFT on $\mathbf{y}_i$ gives

$$\mathbf{r}_i = \mathbf{F}\mathbf{y}_i = \mathbf{\Lambda}\mathbf{s}_i + \underbrace{\tilde{\mathbf{n}}_i}_{\mathbf{F}\mathbf{n}_i}. \tag{1.8}$$

Note that $\tilde{\mathbf{n}}_i$ is a $K \times 1$ vector whose elements are also AWGNs of the variance σ_n^2. To see this, we can compute the covariance matrix of $\tilde{\mathbf{n}}_i$ as

$$\mathrm{E}\{\tilde{\mathbf{n}}_i\tilde{\mathbf{n}}_i^H\} = \mathrm{E}\{\mathbf{F}\mathbf{n}_i\mathbf{n}_i^H\mathbf{F}^H\} = \sigma_n^2\mathrm{E}\{\mathbf{F}\mathbf{F}^H\} = \sigma_n^2\mathbf{I}. \tag{1.9}$$

Since $\mathbf{\Lambda}$ is a diagonal matrix, the frequency selective channel is converted to K parallel flat fading subchannels for each element $s_i(k)$ with the equivalent channel coefficient $\tilde{h}(k)$. In this case, the ML detection of $s_i(k)$ could be separately obtained from

$$s_i(k) = \arg\min_{b} |r_i(k)/\tilde{h}(k) - b|^2. \tag{1.10}$$

This low- complexity one-step ML detection is a major advantage of using OFDM techniques.

1.2 Overview of MIMO System

Traditionally, multiple antennas are placed at one side of the wireless link to perform the interference cancelation through beamforming and to realize the diversity again or the array gain through different ways of combining. It is recently found that, adopting multiple antennas at both sides of the link offers additional benefits—spatial multiplexing gain, which is consistent with the direct goal in developing next-generation wireless communication systems, that is, to increase both the link throughput and the network capacity. Years early, it is normally considered that high data rate transmission can only be achieved by using more bandwidth. However, due to spectral limitations, it is often impractical or sometimes very expensive to increase the bandwidth. In this case, using multiple transmit and receive antennas for spectrally efficient transmission is an alternative but a very attractive solution. Meanwhile, MIMO technology can also enhance the link quality by introducing diversity scheme, e.g., space time coding (STC).

The MIMO channel has multiple links and operates on the same frequency band. One typical MIMO channel with N_t transmit antennas and N_r receive antennas is shown in Fig. 1.3. For ease of the illustration, we consider flat fading channel between different transceiver antennas and denote the corresponding channel coefficient as h_{pq} for $p = 1, \ldots, N_t$, $q = 1, \ldots, N_r$. The transmitted signal during the ith time slot is denoted by the $N_t \times 1$ vector $\mathbf{s}_i = [s_i(1), s_i(2), \ldots, s_i(N_t)]^T$ and the received signal is $\mathbf{r}_i = [r_i(1), r_i(2), \ldots, r_i(N_r)]^T$. Considering also the AWGN at the receiver, $\mathbf{r}_i$ could be represented as

$$\mathbf{r}_i = \mathbf{H}\mathbf{s}_i + \mathbf{n}_i \tag{1.11}$$

where $\mathbf{H}$ is the $N_r \times N_t$ channel matrix with the (q, p)th entry given by h_{pq} and $\mathbf{n}_i = [n_i(1), n_i(2), \ldots, n_i(N_r)]^T$ is the $N_r \times 1$ vector of noise whose elements have

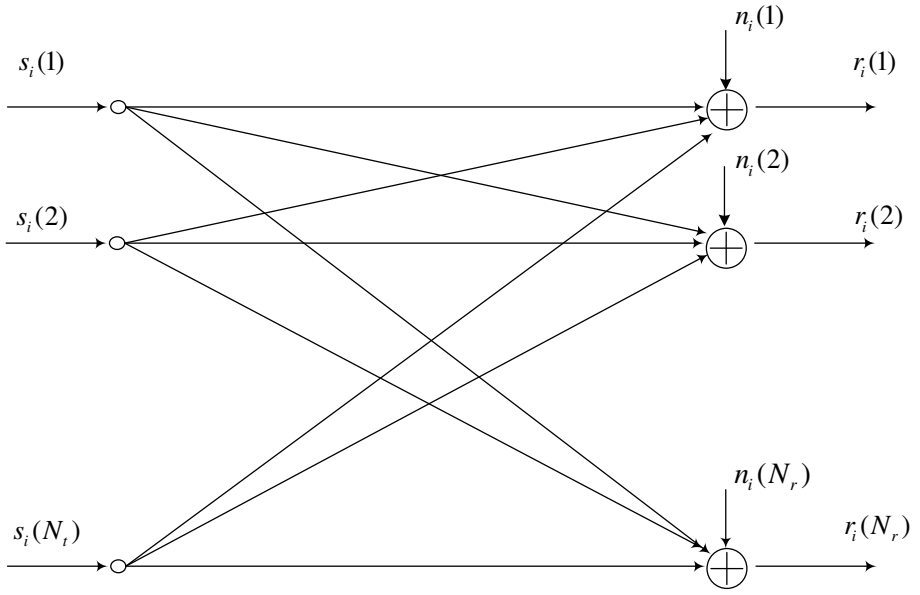

Figure 1.3: Block diagram of MIMO flat fading channels.

variance σ_n^2. Note that some notations are reused from the previous section due to the limited size of the alphabet. The capacity for such a MIMO channel has been derived by [13], [14]:

$$C = \max_{\mathbf{R}_s} \log \det(\mathbf{I} + \mathbf{H}\mathbf{R}_s\mathbf{H}^H/\sigma_n^2) \tag{1.12}$$

where $\mathbf{R}_s = \mathrm{E}\{\mathbf{s}_i\mathbf{s}_i^H\}$ is the covariance matrix of $\mathbf{s}_i$. The optimal $\mathbf{R}_s$ can be obtained from a water-filling procedure by considering the power constraint $tr(\mathbf{R}_s) \leq P_s$, where P_s is the maximum power consumed at the transmitter [13].

In fact, MIMO has gained its application in various standards. Table. 1.1 provides an overview of all current MIMO standards and their technologies.

1.3 MIMO-OFDM system

The signaling schemes in MIMO systems can be roughly grouped into two categories [15]: spatial multiplexing [16] which realizes the capacity gain, and STC [17] which improves the link reliability. Nonetheless, most MIMO systems possess both the spatial multiplexing and the diversity gain. A thorough study on the trade-off

Table 1.1: Current MIMO standards and the corresponding technologies.

Standard	Technology
WLAN 802.11n	OFDM
WiMAX 802.16-2004	OFDM/OFDMA
WiMAX 802.16e	OFDMA
3GPP Release 7	WCDMA
3GPP Release 8 (LTE)	OFDMA
802.20	OFDM
802.22	OFDM

between these two types of gains in flat fading MIMO channels is provided in [18].

It is noted that most performance studies, transmission schemes, and STC designs for MIMO are proposed under flat fading. However, practical wireless communications always contain multi-path fading, where the ISI degrades the system performance substantially and the ML detection can only be achieved with heavy computational burden. Due to the capability of the OFDM that could covert the time domain frequency selective channel to multiple flat fading subchannels, the combination of MIMO and OFDM becomes a natural solution to combat the multi-path fading and enhance the transmission throughput. Therefore, MIMO OFDM has attracted lots of attention and has been adopted in most current and future multi-antenna standards, as can be seen from Tab. 1.1.

Fig. 1.4 shows the MIMO OFDM system model that will be considered throughout the whole thesis. It is seen that MIMO OFDM is a straight combination of MIMO system and OFDM technique.

Assume that the equivalent discrete channel models for different links are expressed as $\mathbf{h}_{pq} = [h_{pq}(0), \ldots, h_{pq}(L_{pq})]^T$, where L_{pq} is the maximum channel delay between the pth transmit antenna and the qth receive antenna. Notations used here are basically the same as those used in subsection 1.1.2 but with the antenna index appearing on the superscript of different notations. For example, the time

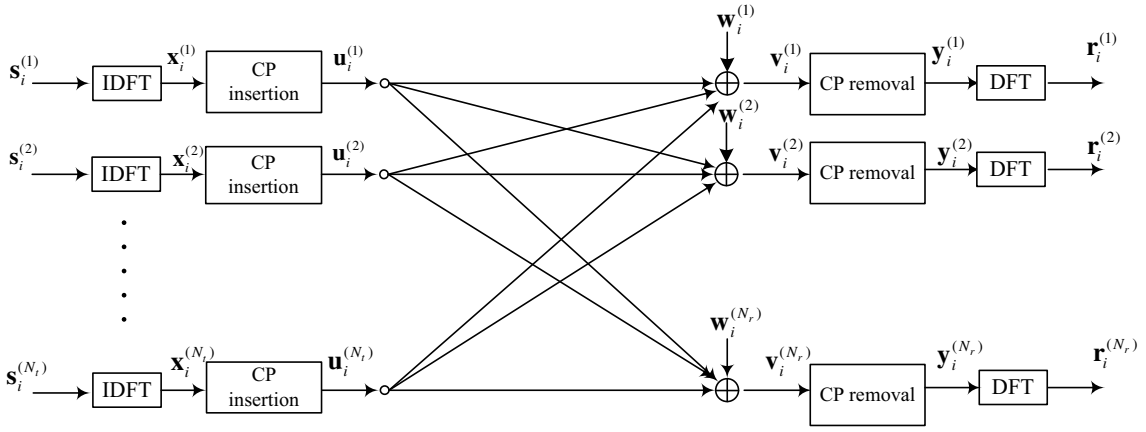

Figure 1.4: A base band MIMO-OFDM System.

domain signal block after IDFT from the pth antenna is $\mathbf{x}_i^{(p)}$ and the one after the CP insertion is $\mathbf{u}_i^{(p)}$. The received signal block on the qth receiver, after the removal of CP, is expressed as

$$\mathbf{y}_i^{(q)} = \sum_{p=1}^{N_t} \mathbf{H}_{pq}\mathbf{x}_i^{(p)} + \mathbf{n}_i^{(q)} \tag{1.13}$$

where $\mathbf{n}_i^{(q)}$ is the noise vector on the qth antenna during the ith signal block and $\mathbf{H}_{pq}$ is the circulant matrix built from $\mathbf{h}_{pq}$. The normalized DFT of $\mathbf{y}_i^{(q)}$ is

$$\mathbf{r}_i^{(q)} = \mathbf{F}\mathbf{y}_i^{(q)} = \sum_{p=1}^{N_t} \mathbf{\Lambda}_{pq}\mathbf{s}_i^{(p)} + \underbrace{\tilde{\mathbf{n}}_i^{(q)}}_{\mathbf{F}\mathbf{n}_i^{(q)}} \tag{1.14}$$

where $\tilde{\mathbf{n}}_{q,i}$ is the noise term after the normalized DFT and $\mathbf{\Lambda}_{pq}$ is the diagonal matrix whose diagonal elements are the K-point DFT of $\mathbf{h}_{pq}$, denoted as $\tilde{\mathbf{h}}_{pq}$. Due to the orthogonality among subcarriers, the signals on each subcarrier experience independent fading channel from each other. Therefore, we can build K different

signal vectors as

$$\begin{aligned}\boldsymbol{\xi}_k &= [r_i^{(1)}(k), r_i^{(2)}(k), \ldots, r_i^{(N_r)}(k)]^T \\ &= \underbrace{\begin{bmatrix} \tilde{h}_{11}(k) & \tilde{h}_{21}(k) & \ldots & \tilde{h}_{N_t 1}(k) \\ \tilde{h}_{12}(k) & \tilde{h}_{22}(k) & \ldots & \tilde{h}_{N_t 2}(k) \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{h}_{1N_r}(k) & \tilde{h}_{2N_r}(k) & \ldots & \tilde{h}_{N_t N_r}(k) \end{bmatrix}}_{\boldsymbol{\Xi}_k} \underbrace{\begin{bmatrix} s_i^{(1)}(k) \\ s_i^{(2)}(k) \\ \vdots \\ s_i^{(N_t)}(k) \end{bmatrix}}_{\boldsymbol{\zeta}_k} + \underbrace{\begin{bmatrix} \tilde{n}_i^{(1)}(k) \\ \tilde{n}_i^{(2)}(k) \\ \vdots \\ \tilde{n}_i^{(N_r)}(k) \end{bmatrix}}_{\boldsymbol{\eta}_k}. \end{aligned} \tag{1.15}$$

Therefore, the MIMO frequency selective fading channels are converted into K MIMO flat fading channels. Moreover each $\boldsymbol{\zeta}_k$ can be independently detected, since the new nose vectors $\boldsymbol{\eta}_k$ are independent across the index k. In this sense, the computational complexity is greatly reduced. The sphere decoding (SD) technique can be applied with the expected detection complexity $O(N_t^{e_c})$, where e_c is some constant related with the signal-to-noise ratio (SNR), the size of the lattice, and the number of the transceiver antennas. The SD algorithm has been intensively discussed in the literatures [19], [20] and the references therein.

1.4 System Initialization

1.4.1 Synchronization

Synchronization is the most important task for any digital communication system. To recognize it, consider a system with differential transmission and without any channel coding or source coding. Differential transmission eliminates the need of the channel estimation and introduces 3 dB loss in terms of SNR. Additionally, excluding coding introduces more SNR loss. Yet, the system could still work under high SNR if the perfect synchronization can be achieved. On the contrary, without the synchronization, the system fails even if the perfect channel knowledge is known and the most powerful coding is applied.

Synchronization normally includes time synchronization and frequency synchronization, of each there are both the requirements for initial estimation and

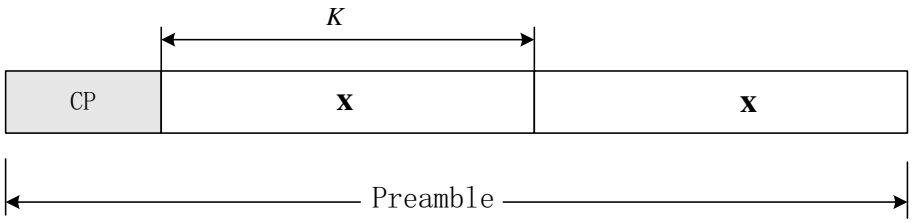

Figure 1.5: Preamble structure of most OFDM schemes.

tracking. Basically, initial estimation counts on the transmitter sending preamble to the receiver at the start of the transmission, whereas tracking requires sending several pilots during the data transmission.

In MIMO systems, antennas are close to each other on any side of the link and usually have a unique oscillator and sampling clock. As a result, the timing offsets (TO) and the carrier frequency offsets (CFO) between different transceiver pairs are normally the same [21]- [24]. In view of this, the synchronization for MIMO OFDM makes no difference from that for single-input single-output (SISO) OFDM.

Preamble

Most wireless communication systems are packet-switched systems with a random access protocol. This essentially indicates that a receiver has no *a priori* knowledge about the arrival time of any packet. The random nature of the arrival times and the high data rates require the synchronization to be completed shortly after the start of the reception of a packet. To facilitate "quick" synchronization, each data packet is equipped with a known sequence in the front, called the preamble. The preamble is designed to provide information for a good packet detection, synchronization, as well as channel estimation.

Channel estimation for the MIMO system normally requires orthogonal sequences for all transmit antennas to be included as parts of the preamble in order to achieve the optimal estimation [25]. To perform synchronization, a periodical structure in the preamble is preferred since the phase rotation between time-delayed versions of the same symbol is a measure for the CFO and this phase rotation does

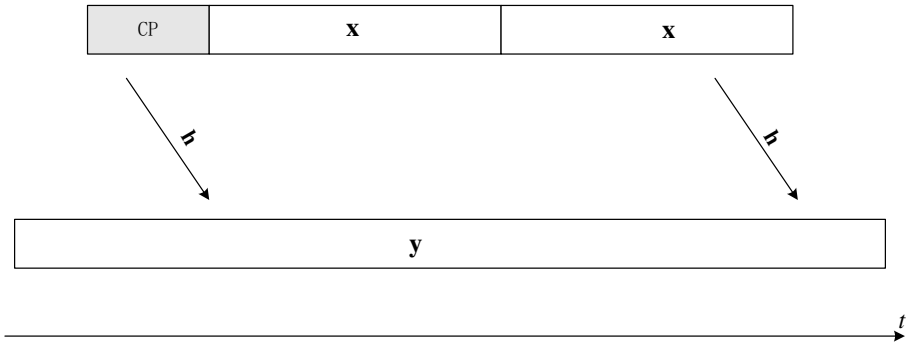

Figure 1.6: Receiving the preamble at the destination.

not affect the power of the received signals such that the frame detection and the TO estimation can be performed [26]. Therefore, many preambles consist of at least one concatenation of two identical training sequences per transmit antenna. Furthermore, to make the channel estimation less vulnerable to ISI, a CP with the length greater than the channel delay spread is added. A typical structure of the preamble is shown in Fig. 1.5. For different OFDM standards, there may exist minor alteration on the preamble structure. For example in IEEE 802. 11a, ten short identical training sequences are placed before two long identical training sequences.

Frame Detection and Time Synchronization

The task of the frame detection is to identify the preamble in order to detect the arrival of a packet. The frame detection algorithm can also be used as a time synchronization algorithm, since it inherently provides a rough estimate of the starting point of the packet.

Perhaps the most widely used algorithm is the one proposed by Schmidl and Cox in their early work [26]. The algorithm is based on the correlation between the two identical parts of the preamble. Define the received signal sequence as $\mathbf{y} = [y(0), y(1), \ldots, y(M)]^T$ as shown in Fig. 1.6. The timing metric could be written as

$$M(d) = \frac{|\sum_{k=0}^{K-1} y^*(d+k)y(d+k+K)|^2}{\sum_{k=0}^{K-1} |y(d+k+K)|^2}. \tag{1.16}$$

If $M(d)$ is greater than a pre-set threshold, the packet is supposed to be detected. Moreover, the value of d that maximizes the matrix $M(d)$ is deemed as the first symbol in the first training sequence.

Frequency Synchronization

The frequency synchronization mainly targets to correct the CFO, which is caused by the difference between oscillator center frequency at the transmitter and that at the receiver, or by the Doppler effect. The CFO can be estimated using the phase of the complex correlation between the two consecutive received training symbols [26]. A simple MIMO extension of [26] was proposed in [27], where it is assumed that there is a unique oscillator at either side of the MIMO system. This is a valid assumption if the antennas are co-located. The estimated CFO is then given by

$$\phi = \frac{1}{2\pi K} \angle \left(\sum_{k=0}^{K-1} y^*(d_{\text{op}} + k) y(d_{\text{op}} + k + K) \right) \tag{1.17}$$

where d_{op} is the optimal result from (1.16). It is also noted that the maximum estimation range of the CFO is limited to $(-0.5/K, 0.5/K]$ which equals one subcarrier spacing, because the angle that can be estimated without phase ambiguity is limited to $(-\pi, \pi]$. A larger range can be achieved by slightly changing the structure of the preamble, for example short periodical training are also adopted in IEEE 802.11a.

Frequency Offset Tracking

After the rough estimation by the preamble, the residue TO and the residue CFO usually lie in tolerable regions. Normally, the residue TO will not change from time to time if the sampling clock frequency is precise enough. However, the residue CFO may vary slowly due to the center frequency drifting of the oscillator that is caused by temperature changes, aging, and other effects [28]. Therefore, the residue CFO is not a static value but a rather random or time-varying process. Although this drift is slow relative to the symbol block period, it may hurt the performance in the

long term viewpoint. Therefore, the residue CFO must be tracked and compensated frequently during the data transmission. The training based methods that require sending continuous training symbols cannot deal with this issue well due to its bandwidth inefficiency. It is then better to design new approaches that can reliably track the frequency varying from either the pilot tone (PT) or the bind ways.

1.4.2 Channel Estimation

Channel estimation is one of the most important components for almost all the wireless communication systems. Knowing the channel state information (CSI) can not only facilitate the data detection but is also beneficial in power allocation and design of the capacity achieving schemes. Non-coherent detection, as an alternative, alleviates the requirement of channel estimation but suffers from 3 dB power loss compared to the coherent detection. In addition, not all transmission schemes have corresponding non-coherent detection techniques available. Consequently, channel estimation have been extensively studied over last two decades [29]-[49].

Training Based Channel Estimation

Training based channel estimation is adopted in almost all the current standards and applications, where either the preamble or the pilot are transmitted to help training the channels [29]-[34]. The advantages of the training based channel estimation is its capability to provide accurate estimation within a short period and require very low complexity. We give an example on training based channel estimation in SISO OFDM system. Suppose the training sequence is $\mathbf{s} = [s(0), s(1), \ldots, s(K-1)]^T$ and its normalized IDFT is $\mathbf{x}$. With perfect synchronization, the received signals $\mathbf{y}$ and its DFT $\mathbf{r} = [r(0), r(1), \ldots, r(K-1)]^T$ follow the similar expression in (1.5) and (1.8), respectively. Either the time domain channel vector $\mathbf{h}$ or the frequency domain vector $\tilde{\mathbf{h}}$ can be estimated. To achieve lower complexity, $\tilde{\mathbf{h}}$ is often chosen to be estimated and the ML channel estimation is given by

$$\tilde{h}(k) = r(k)/s(k), \quad k = 0, \ldots, K-1. \tag{1.18}$$

If the channel length or its upper bound is known as L, the denoising approach can be applied to increase the channel estimation accuracy [50]:

$$\tilde{\mathbf{h}} = \mathbf{F}\mathbf{F}^{\dagger}(:, 1:L+1)\tilde{\mathbf{h}}, \tag{1.19}$$

where $\mathbf{F}(:, 1:L+1)$ is the $K \times (L+1)$ matrix that contains the first $L+1$ columns of $\mathbf{F}$.

Blind Channel Estimation

Although training based channel estimation can provide reliable channel estimates, the spectrum efficiency is decreased since training should be transmitted from time to time or at least at every start of one packet. An alternative solution is the so called blind channel estimation which has received considerable attention during the past decade [35]-[49]. Blind channel estimation normally relies on the statistical information of transmitted signals, e.g., whiteness, circularity, etc. Although blind method has higher spectrum efficiency, it normally requires a longer observation of the received signals as well as a higher computational complexity. Therefore, blind method is not suitable for relatively fast fading channels. Nonetheless, for next generation wireless communications that aim at high data rate transmission, the channel could be reasonably considered constant during one packet transmission.

The first effort in blind channel estimation mainly focused on the higher-order statistics of the received symbols [35]-[38]. However, this procedure is computationally expensive and requires too long observation of data blocks. A major breakthrough was accomplished in [39] where a method allowing the blind identification of the channels using only second-order statistics (SOS) was proposed. Following this work, a promising family of blind channel estimation, so called subspace-based blind channel estimation algorithm (SS) was developed in [40]-[49] for either SISO, or single-input multi-output (SIMO) systems. In SS method, the observation space is separated into signal and noise subspace by applying eigen-value decomposition (EVD) on the covariance matrix of the received signals. By exploiting the inherent structure of the channel matrix, the channel vector can be estimated

from the noise subspace up to a complex scalar ambiguity. This ambiguity can be solved either by transmitting several training symbols [46], forming the so called semi-blind channel estimation, or by exploiting special symbol structure in the block transmissions [49].

1.5 Research Objectives and Main Contributions

In this thesis, we will develop robust CFO tracking algorithms as well as the blind channel estimation algorithms for MIMO-OFDM systems.

In terms of CFO tracking, we target at a new algorithm that could overcome the drawbacks of the existing methods, e.g., low accuracy, small estimation range, partial utilization of the existing resources, etc. We first develop a robust frequency tracking algorithm using PTs that are issued in almost all the standards and are embedded in each OFDM block. Identifiability of this pilot based algorithm is studied for the noise free case, and a constellation rotation strategy is proposed to eliminate the CFO ambiguity. To further improve the performance accuracy and enhance the algorithm robustness to the CFO ambiguity, we consider the combination from the virtual carriers (VC), that are also possessed in practical OFDM standards. For example, in IEEE 802.11a standard, the subcarriers with indices $\{0, 27, 28, \ldots, 36, 37\}$ are set as VCs, either to avoid the aliasing effect [51] or to be reserved for future use. The CFO estimation algorithm by exploiting VCs has been developed in [52]-[54]. Then, a weighted algorithm is proposed by exploiting both PTs and VCs. We show that in the weighted algorithm, the PT part increases the estimation accuracy, while the VC part reduces the outlier probability. Moreover, we derive the asymptotic mean square error (MSE) of our proposed algorithm, and the optimal weight is given in a closed-form. It turns out that, the proposed frequency tracking algorithm is also applicable to the synchronization initialization since the algorithm itself does not require the knowledge of the CSI and can provide the full range CFO estimation.

In terms of blind channel estimation, we develop a new SS algorithm that possesses the following advantages: robustness to channel order over-estimation,

guaranteeing the channel identifiability, applicability to the scenario where the number of the receive antennas is no more than the number of the transmit antennas ($N_r \leq N_t$), etc. Note that the last property is not possessed by the traditional SS algorithm. We first apply a re-modulation to the received signals such that the system model is converted to the one similar to zero-padding (ZP) based MIMO OFDM [55], which renders CP-OFDM all the advantages of ZP-OFDM. Besides, CP-OFDM is compatible to most existing OFDM standards or the further 4G MIMO-OFDM standards [8], [9]. We also provide thorough performance analysis for CP-OFDM and it is shown that the asymptotical channel estimation MSE agrees with the approximated asymptotical Cramér-Rao Bound (CRB). Since the re-modulation based SS algorithm is not applicable for the case with $N_r < N_t$, we further develop a non-redundant linear precoding based algorithm. The assumption that the symbols sent from different transmitters are independent and identically distributed (i.i.d.) allows this method to yield acceptable performance at low SNR region and to work even for the multiple-input single-output (MISO) transmission scenario. The method meets an error floor at high SNR which shows a reasonable trade-off as the method itself overcomes the very difficulty on the requirement of the number of the transceiver antennas. It is shown from the simulation that, acceptable performance can still be achieved with relatively short observation time. We also propose an approach to eliminate the multi-dimensional ambiguity that is known to exist for blind channel estimation under multi-transmitter scenarios.

1.6 Organization of the Thesis

The thesis is organized as follows. In Chapter 2, several existing CFO tracking algorithms for OFDM systems are introduced. The preliminary knowledge of SS method is also introduced in this chapter. In Chapter 3, the newly derived robust CFO tracking algorithm and its theoretical performance analysis are presented. In Chapter 4, we develop the system re-modulation to convert the CP based MIMO OFDM into a similar model of ZP based MIMO OFDM. Several analytical results

related to the channel estimation error are also derived. Chapter 5 provides the non-redundant precoding based channel estimation for MIMO OFDM systems, which is applicable for the case $N_r \leq N_t$. Finally, the concluding remarks are drawn in Chapter 6 and proofs of theorems are provided in Appendices.

Chapter 2

Review of Existing Techniques

In this chapter, we briefly introduce some current CFO tracking algorithms for OFDM systems. We point out that all the existing methods have their own drawbacks and may fail the CFO tracking under certain scenario. We then provide the preliminary knowledge of SS method and discuss the difficulties on extending the SS method to MIMO OFDM systems.

2.1 Convectional CFO Tracking Algorithms

The CFO tracking algorithms can be classified into three categories, i.e., PT-aided, CP-based, and VC-based schemes. PT-aided approach estimates CFO by periodically inserting pilots on particular subcarriers and correlating the received symbols with known pilots. CP-based method utilizes the periodicity created by the insertion of the CP. VC-based scheme, on the other side, makes use of the orthogonality between VCs and data modulated subcarriers. The principles of these three methods have been presented in [56]-[58], [52].

2.1.1 System Model

The notations from subsection 1.1.2 are adopted here. Since the CFO is in presence, the system model (1.5) should be modified accordingly. Denote the index sets

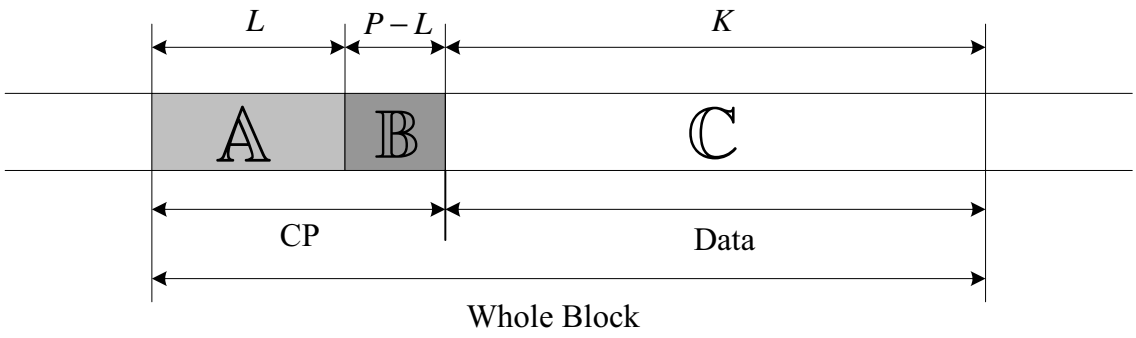

Figure 2.1: Structure of the received OFDM block.

for PTs and VCs as $\mathcal{P}$ and $\mathcal{V}$, respectively. The transmitted symbol on the kth subcarrier in the ith OFDM block is

$$s_i(k) = \begin{cases} p_i(k) \in \mathcal{C}_p & k \in \mathcal{P} \\ 0 & k \in \mathcal{V} \\ d_i(k) \in \mathcal{C}_d & \text{otherwise} \end{cases}, \quad k = 0, \ldots, K-1 \tag{2.1}$$

where $d_i(k)$ is the information symbol from the constellation $\mathcal{C}_d$, and $p_i(k)$ is the pilot symbol from the constellation $\mathcal{C}_p$.

At the receiver, there usually exist both CFO and TO, which must be estimated and compensated before the subsequent channel estimation and data detection. For simplicity we assume perfect time synchronization. Suppose the frequency offset is Δf and its normalization with respective to $1/T_s$ is $\phi = \Delta f T_s$. The received baseband signals before and after the CP removal are given by

$$\mathbf{v}_i = e^{j2\pi\phi((i-1)K_s)}\mathbf{\Psi}(\phi)(\mathbf{H}_0\mathbf{u}_i + \mathbf{H}_1\mathbf{u}_{i-1}) + \mathbf{w}_i \tag{2.2}$$

$$\mathbf{y}_i = e^{j2\pi\phi((i-1)K_s+P)}\mathbf{\Omega}(\phi)\mathbf{H}\mathbf{x}_i + \mathbf{n}_i \tag{2.3}$$

respectively, where $\mathbf{\Psi}(\phi)$ and $\mathbf{\Omega}(\phi)$ have the forms

$$\mathbf{\Psi}(\phi) = \text{diag}\left\{1, e^{j2\pi\phi}, \ldots, e^{j2\pi\phi(K_s-1)}\right\} \tag{2.4}$$

$$\mathbf{\Omega}(\phi) = \text{diag}\left\{1, e^{j2\pi\phi}, \ldots, e^{j2\pi\phi(K-1)}\right\}. \tag{2.5}$$

The structure of the received OFDM block is shown in Fig. 2.1, where it is divided into three regions: $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$. Symbols in region $\mathbb{A}$ are corrupted by IBI from

the previous block. Region $\mathbb{B}$ represents the part in CP that is IBI free. Region $\mathbb{C}$ denotes information symbols $\mathbf{y}_i$.

2.1.2 PT-Based Algorithm

Let $\mathbf{r}_i = [r_i(0), r_i(1), \ldots, r_i(K-1)]^T$ denote the K-point normalized DFT of $\mathbf{y}_i$. Then

$$\begin{aligned} r_i(k)e^{-j2\pi\phi((i-1)K_s+P)} &= \frac{1}{K}\sum_{v=0}^{K-1}\tilde{h}(v)s_i(v)e^{\frac{j(K-1)\pi(v+K\phi-k)}{K}}\frac{\sin(\pi(v+K\phi-k))}{\sin\left(\frac{\pi(v+K\phi-k)}{K}\right)}+\tilde{n}_i(k) \\ &= \frac{s_i(k)\tilde{h}(k)e^{j(K-1)\pi\phi}\sin(\pi K\phi)}{K\sin(\pi\phi)}+\mathrm{ICI}_i(k)+\tilde{n}_i(k) \end{aligned} \tag{2.6}$$

where

$$\mathrm{ICI}_i(k) = \sum_{v=0,\, v\neq k}^{K-1}\frac{\tilde{h}(v)s_i(v)e^{\frac{j(K-1)\pi(v+K\phi-k)}{K}}\sin\pi(v+K\phi-k))}{K\sin\left(\frac{\pi(v+K\phi-k)}{K}\right)} \tag{2.7}$$

is the inter-carrier interference (ICI). For noise free case and $\phi = 0$, $r_i(k) = \tilde{h}(k)s_i(k)$. A non-zero ϕ both introduces ICI and reduces the effective signal power by a factor of $\frac{e^{j(K-1)\pi\phi}\sin(\pi K\phi)}{K\sin(\pi K\phi)}$.

For a slow fading channel[1], the channel in $\nu = 2$ consecutive blocks can be assumed static. Based on this fact, the Classen&Meyr's method [56] is developed by using a few number of pilots. In fact, Classen&Meyr's method assumes a sufficiently small ϕ and a not high SNR, so that the ICI is much smaller than the noise and can thus be ignored. The CFO is then estimated from

$$\hat{\phi} = \frac{1}{2\pi K_s}\angle\left(\sum_{k\in\mathcal{P}} r_i^*(k)r_{i+1}(k)/\left(s_i^*(k)s_{i+1}(k)\right)\right). \tag{2.8}$$

Obviously, (2.8) is valid only when $\phi \ll \frac{1}{K_s}$. Therefore the coarse estimation during the CFO acquisition stage is crucial to the performance of (2.8). There also exist other problems: 1) the estimation accuracy of (2.8) is limited by ignoring the ICI; 2) At high SNR, since the ICI term is comparable to or even larger than the noise, the approximation in (2.8) is not valid any more.

[1]Most frame based transmissions assume a slow fading channel, e.g., IEEE 802.11a [7].

2.1.3 CP-Based Algorithm

Since CP is the copy of the last P data samples of the OFDM block, it is expected that $P-L$ symbols in the region $\mathbb{B}$ of Fig. 2.1 should be the same as the last $P-L$ symbols in the region $\mathbb{C}$ under the zero CFO case, namely,

$$v_i(L+t) = v_i(K+L+t) \tag{2.9}$$

for $t = 1, \ldots, P-L$. However, with the existence of CFO, the relationship changes to

$$v_i(L+t) = e^{-j2\pi K\phi} v_i(K+L+t). \tag{2.10}$$

Based on this observation, CP-based CFO tracking has been developed in [57], [58], which will also be named as Beek's method, as

$$\hat{\phi} = \frac{1}{2\pi K} \angle \left(\sum_{t=1}^{P-L} v_i^*(L+t) v_i(K+L+t) \right). \tag{2.11}$$

The advantages of the CP-based algorithm is its capability of tracking CFO within only one OFDM block whereas the PT-based algorithm requires two consecutive OFDM blocks. However, CP-based algorithm also exhibits many drawbacks: 1) The estimation range is still as small as one subcarrier spacing; 2) The knowledge of the channel length is crucial to the performance; say if $L = P$, then the CP-based algorithm cannot work at all.

2.1.4 VC-Based Algorithm

The algorithm purely relying on VCs has been studied in [52]-[54]. It is proven in [53] that, the ML CFO estimation based only on VCs is

$$\hat{\phi} = \min_{\varepsilon} \left\| \mathbf{F}_v^H \mathbf{\Omega}(-\varepsilon) \mathbf{y}_i \right\|^2 \tag{2.12}$$

where ε is the trial variable and $\mathbf{F}_v$ is the $K \times |\mathcal{V}|$ matrix whose columns are constructed from $\mathbf{f}_k$, $k \in \mathcal{V}$. The identifiability of (2.12) has been fully studied in [59].

VC-based algorithm exhibits many advantages: 1) The CFO tracking can be accomplished after receiving only one OFDM block; 2) The CFO estimation range reaches its maximum, i.e. $(-0.5, 0.5]$. Therefore, it can also be used for CFO acquisition at the start of the packet transmission; 3) The performance is not affected by the channel length. However, the bottle neck of this method is its high computational complexity since one dimensional searching of ε over the range $(-0.5, 0.5]$ is normally required. Although the complexity can be reduced by the polynomial rooting [54], it is still very high compared to the PT- or the CP-based method. Nonetheless, it is fine to use VC-based CFO tracking if the adaptive scheme is adopted, because after the CFO acquisition the residue CFO is not big and the local minimal converged from (2.12) is the true optimal with very high probability. Another drawback of VC-based algorithm is its lower accuracy, since this method is only a type of blind algorithm.

2.2 Conventional Subspace Based Channel Estimation Method

2.2.1 The Algorithm

In this subsection, we introduce the SS method for single transmit antenna systems, i.e. SISO, SIMO [40]. To provide a general discussion, we consider the pure mathematical approach and the system model is written as

$$\mathbf{r}_i = \mathbf{H}\mathbf{s}_i + \mathbf{n}_i \tag{2.13}$$

where $\mathbf{r}_i$ is the ith received signal block of dimension $N \times 1$; $\mathbf{s}_i$ is the ith transmitted signal block of dimension $M \times 1$; $\mathbf{n}_i$ is the $N \times 1$ noise vector whose elements are AWGNs with variance σ_n^2; $\mathbf{H}$ is the channel matrix whose elements are chosen from the multi-path channel vector $\mathbf{h} = [h(0), h(1), \ldots, h(L)]^T$ and should be constant over a certain period. As will be seen later, the structure of $\mathbf{H}$ is different for different systems. Nevertheless, the mth column of $\mathbf{H}$ could be represented as $\mathbf{C}_m\mathbf{h}$

where $\mathbf{C}_m$ is some appropriate matrix of dimension $N \times (L+1)$.

The covariance matrix of $\mathbf{r}_i$ is then calculated from

$$\mathbf{R}_r = \mathrm{E}\{\mathbf{r}_i \mathbf{r}_i^H\} = \mathbf{H}\mathbf{R}_s\mathbf{H}^H + \sigma_n^2\mathbf{I} \tag{2.14}$$

where $\mathbf{R}_s = \mathrm{E}\{\mathbf{s}_i \mathbf{s}_i^H\}$ is the covariance matrix of $\mathbf{s}_i$.

The subspace algorithm requires $N > M$ and $\mathbf{R}_s$ should also be a full rank matrix. The latter requirement is generally fulfilled since fully correlated symbols are seldom transmitted. Then, the term $\mathbf{H}\mathbf{R}_s\mathbf{H}^H$ in (2.14) can be eigen-decomposed as

$$\mathbf{H}\mathbf{R}_s\mathbf{H}^H = \mathbf{U}\boldsymbol{\Delta}\mathbf{U}^H = \begin{bmatrix}\mathbf{U}_s & \mathbf{U}_o\end{bmatrix} \begin{bmatrix}\boldsymbol{\Delta}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{U}_s^H \\ \mathbf{U}_o^H\end{bmatrix} \tag{2.15}$$

where the $M \times M$ diagonal matrix $\boldsymbol{\Delta}_s$ contains M non-zero eigen-values of $\mathbf{H}\mathbf{R}_s\mathbf{H}^H$ and the $N \times M$ matrix $\mathbf{U}_s$ spans the so called signal-subspace. In turn, the $N \times (N-M)$ matrix $\mathbf{U}_o$ spans the noise-subspace. It is not hard to know that $\mathbf{H}$ and $\mathbf{U}_s$ span the same subspace and the noise subspace is orthogonal to the signal subspace. Hence, the following equation holds:

$$\mathbf{U}_o^H\mathbf{H} = \mathbf{0}. \tag{2.16}$$

Since the noise is AWGN, the covariance matrix $\mathbf{R}_r$ is written as

$$\begin{aligned}\mathbf{R}_r &= \mathbf{H}\mathbf{R}_s\mathbf{H}^H + \sigma_n^2\mathbf{I} = \mathbf{U}\boldsymbol{\Delta}\mathbf{U}^H + \sigma_n^2\mathbf{U}\mathbf{U}^H \\ &= \mathbf{U}(\boldsymbol{\Delta} + \sigma_n^2\mathbf{I})\mathbf{U}^H = \begin{bmatrix}\mathbf{U}_s & \mathbf{U}_o\end{bmatrix} \begin{bmatrix}\boldsymbol{\Delta}_s + \sigma_n^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_n^2\mathbf{I}\end{bmatrix} \begin{bmatrix}\mathbf{U}_s^H \\ \mathbf{U}_o^H\end{bmatrix}.\end{aligned} \tag{2.17}$$

From the uniqueness of the EVD, $\mathbf{U}$ is also the eigen-matrix of $\mathbf{R}_r$. Therefore, even at the noisy case, $\mathbf{U}$ of $\mathbf{H}\mathbf{R}_s\mathbf{H}^H$ could still be obtained from the EVD of the signal covariance matrix $\mathbf{R}_r$. However, $\mathbf{U}_o$ should be obtained from the columns of $\mathbf{U}$ that corresponds to the eigen-values σ_n^2. Practically we can only construct the signal covariance matrix from the sample covariance matrix, namely:

$$\hat{\mathbf{R}} = \sum_i \mathbf{r}_i \mathbf{r}_i^H. \tag{2.18}$$

Then $\mathbf{U}_o$ should be obtained from the columns of $\mathbf{U}$ that corresponds to the $N-M$ smallest eigen-values.

From (2.16), there is

$$\mathbf{U}_o^H \mathbf{C}_m \mathbf{h} = \mathbf{0}, \quad m = 1, \ldots, M. \tag{2.19}$$

Define

$$\boldsymbol{\mathcal{K}} = [\mathbf{C}_1^H \mathbf{U}_o, \mathbf{C}_2^H \mathbf{U}_o, \ldots, \mathbf{C}_M^H \mathbf{U}_o]. \tag{2.20}$$

The channel vector can be obtained from

$$\boldsymbol{\mathcal{K}}^H \mathbf{h} = \mathbf{0}. \tag{2.21}$$

Therefore, the estimate of $\mathbf{h}$, denoted as $\hat{\mathbf{h}}$ is a basis vector of the null space of $\boldsymbol{\mathcal{K}}^H$.

It is important that the dimension of null space of $\boldsymbol{\mathcal{K}}^H$ is one such that $\hat{\mathbf{h}}$ is different from $\mathbf{h}$ by only an unknown complex scalar, namely $\hat{\mathbf{h}} = \alpha \mathbf{h}$. If this is satisfied, we call that the identifiability of the channel is guaranteed. The unknown scalar α exists in common to all the blind channel estimation method and could be removed by transmitting a few training symbols. The identifiability study is one of the most hardest problems for SOS based channel estimation. Different consideration should be assumed for different systems. If the identifiability is not guaranteed, the blind channel estimation would be meaningless.

2.2.2 Difficulties on Extending SS to MIMO OFDM

It is not such straightforward to extend SS method to MIMO OFDM systems. First of all, if the number of the receive antennas is smaller than or equal to the number of the transmit antennas, the channel matrix in MIMO OFDM is square or fat, and the dimension of the noise subspace is zero. In other words, it is impossible to obtain the channel vector from the SS method. Unfortunately, this undesirable scenario also includes the popular SISO OFDM, see equation (1.5). Different types of re-modulation on SISO OFDM have been proposed in [43], [45], [47], [48]. In [43], the existence of VCs are utilized to construct the noise subspace. However, the

method requires concatenating more than two OFDM blocks and yields a covariance matrix with very large dimension. Consequently, the computational complexity is increased. Besides, VCs are reserved for future use and purely counting on the existence of VCs may not be agreeable in future applications. In [45], an OFDM scheme without CP is proposed. The method not only destroys the orthogonality among subcarriers but also requires concatenation of several OFDM blocks. The concatenation of two OFDM blocks are also suggested in [47] where the CP between two OFDM blocks is used to construct the noise subspace. Besides increasing the computational complexity, this scheme is not well studied on the aspect of the identifiability issue. Moreover, how to extend this scheme to MIMO OFDM is not such straightforward. The redundant precoding based channel estimation combined with the space-time coding is proposed in [48]. Since the redundant precoding is applied, the transmission efficiency will be reduced.

A new OFDM scheme is proposed in [55] where, instead of using CP, consecutive zeros are padded at the end of each OFDM block. This scheme is called ZP-based OFDM as opposed to the traditional CP based OFDM. ZP OFDM exhibits many advantages. For example, it is applicable to MIMO system when $N_t \geq N_r$; it is robust to channel order over-estimation, its channel identifiability is guaranteed, etc. However, a major drawback that prevents ZP OFDM from wide application is its incompatibility to most existing OFDM standards [4]-[7] or the further 4G MIMO OFDM standards [8], [9].

2.3 Cramér-Rao Bounds

The Cramér-Rao bound (CRB) is a lower bound on the error variance of the best estimator for estimating a specific set of parameters with the given system. Denote the unknown parameters of a given system by a length-χ vector $\boldsymbol{\vartheta} = [\vartheta_1, \vartheta_2, \ldots, \vartheta_\chi]^T$ and the estimate of $\boldsymbol{\vartheta}$ is denoted by $\hat{\boldsymbol{\vartheta}}$. A lower bound on the accuracy of any

unbiased estimator is expressed as

$$\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) = \mathrm{E}\{(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta})(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta})^T\} \succcurlyeq \mathrm{CRB}(\boldsymbol{\vartheta}) \tag{2.22}$$

where "$\succcurlyeq$" is interpreted as that the matrix $\Big(\mathrm{Cov}(\hat{\boldsymbol{\vartheta}}) - \mathrm{CRB}(\boldsymbol{\vartheta})\Big)$ is positive semi-definite. The expression of CRB is given by the inverse of the Fisher Information Matrix (FIM) $\mathbf{J}(\boldsymbol{\vartheta})$, that is

$$\mathrm{CRB}(\boldsymbol{\vartheta}) = \mathbf{J}^{-1}(\boldsymbol{\vartheta}). \tag{2.23}$$

Let the noiseless measurement be denoted by $\bar{\mathbf{y}}(\boldsymbol{\vartheta})$ and the noisy measurement be denoted by $\bar{\mathbf{r}}$, respectively. The FIM is defined by

$$\mathbf{J}(\boldsymbol{\vartheta}) = \mathrm{E}\left\{\left(\frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right)\left(\frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right)^T\right\} \tag{2.24}$$

where $p(\bar{\mathbf{r}};\boldsymbol{\vartheta})$ is the probability density function (PDF) of $\bar{\mathbf{r}}$ and the derivative is evaluated at the true parameters. Assuming a zero mean white Gaussian noise with covariance matrix $\sigma_n^2\mathbf{I}$, the probability density function (PDF) for $\bar{\mathbf{r}}$ is given by

$$p(\bar{\mathbf{r}};\boldsymbol{\vartheta}) = \frac{1}{(2\pi\sigma_n^2)^{N/2}}\exp\left[-\frac{1}{2\sigma_n^2}(\bar{\mathbf{r}} - \bar{\mathbf{y}}(\boldsymbol{\vartheta}))^T(\bar{\mathbf{r}} - \bar{\mathbf{y}}(\boldsymbol{\vartheta}))\right] \tag{2.25}$$

where N is the length of the vector $\bar{\mathbf{r}}$. Here, the derivative of the scalar function $p(\bar{\mathbf{r}};\boldsymbol{\vartheta})$ to the vector $\boldsymbol{\vartheta}$ is defined as the vector

$$\frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} = \left[\frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \vartheta_1}, \frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \vartheta_2}, \ldots, \frac{\partial \ln p(\bar{\mathbf{r}};\boldsymbol{\vartheta})}{\partial \vartheta_\chi}\right]^T. \tag{2.26}$$

Chapter 3

Robust Synchronization for OFDM Systems

In this chapter, we propose a novel CFO tracking algorithm for OFDM systems by exploiting both PTs and VCs embedded in each OFDM block. The proposed tracking algorithm is able to achieve a full range CFO estimation, can be used before channel estimation, and could provide improved performance compared to existing algorithms. Moreover, reliable TO tracking can also be achieved at relatively high SNR. As other synchronization algorithms, the proposed one can be easily extended to MIMO OFDM systems.

3.1 Introduction

In OFDM systems, it is well known that a CFO, caused by oscillators' mismatch or Doppler effects, destroys the subcarrier orthogonality and results in a substantial bit error rate (BER) degradation [60]. Therefore, frequency synchronization should be performed before the channel estimation and the subsequent data detection.

Several CFO acquisition methods have been proposed in [26], [57], [61]-[67]. The Beek's method [57] exploits CP in front of each OFDM block for AWGN channels and is later extended to multi-path channels in [58]. Repeated training sequences

are used in [26], [61], [62], and several improvements are made in [63]-[66]. Optimal training sequence design for CFO estimation over frequency selective channel has been recently proposed in [67], where the periodic structures are, again, exploited. Hence, almost all the methods in [57]-[67] use the periodic nature of the time domain signal, either by utilizing CP or by designing training sequences with repeated parts. However, using the periodic nature greatly reduces the CFO estimation region. Furthermore, all these methods, except Beek's method, are only applicable in CFO acquisition stage because consecutive training blocks are required, which are only available when the transmission starts.

After the acquisition stage, there may exist a residue CFO, either because of the insufficient accuracy during the coarse estimation, or because of the time varying nature of the surrounding environment. The residue CFO, if not compensated, may still lead to performance degradation. Hence, many existing standards reserve a limited number of scattered pilot tones in each OFDM block to improve the system robustness. For example, in IEEE 802.11a WLAN standard [7], four pilots are placed at the subcarriers with indices $\{7, 21, 43, 57\}$ for the purpose of combating the residue CFO and the phase noise.

We have introduced in Chapter 2 several existing CFO tracking methods and listed their drawbacks. In this chapter, we develop a CFO tracking algorithm based on the limited number of PTs by assuming that the channel remains constant during two consecutive OFDM blocks. The method overcomes almost all the drawbacks of PT-based, CP-based and VC-based algorithms and could provide improved performances.

3.2 New CFO Tracking Algorithm

3.2.1 New Pilot-Based Tracking: p-Algorithm

We follow the system model introduced in subsection 2.1.1. Let ε be the trial variable for the unknown CFO. After compensating $\mathbf{y}(m)$ by a diagonal matrix $\mathbf{\Omega}(-\varepsilon)$, the

symbol on the kth subcarrier can be written as

$$r_i(k)e^{-j2\pi\phi((i-1)K_s+P)} = \frac{s_i(k)\tilde{h}(k)e^{j(K-1)\pi(\phi-\varepsilon)}\sin(\pi K(\phi-\varepsilon))}{K\sin\big(\pi(\phi-\varepsilon)\big)} + \mathrm{ICI}_i(k) + \tilde{n}_i(k). \tag{3.1}$$

In the absence of the noise, if $\varepsilon = \phi$, $r_i(k)e^{-j2\pi\varepsilon((i-1)K_s+P)}$ reduces to $s_i(k)\tilde{h}(k)$, and $r_{i+1}(k)e^{-j2\pi\varepsilon(iK_s+P)}$ reduces to $s_{i+1}(k)\tilde{h}(k)$. Assuming that the channel is constant over two consecutive OFDM blocks, the metric

$$\sum_{k\in\mathcal{P}} \left\| r_i(k)e^{-j2\pi\varepsilon((i-1)K_s+P)}/s_i(k) - r_{i+1}(k)e^{-j2\pi\varepsilon(iK_s+P)}/s_{i+1}(k) \right\|^2 \tag{3.2}$$

will be zero at $\varepsilon = \phi$. Note that, at least one $\tilde{h}(k)$, $k \in \mathcal{P}$ should be non-zero, which is also assumed by Classen&Meyr's method. Define

$$\begin{aligned} g_p(\varepsilon) =& \sum_{k\in\mathcal{P}} \left\| e^{-j2\pi\varepsilon((i-1)K_s+P)}\mathbf{f}_k^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_i/s_i(k) - e^{-j2\pi\varepsilon(iK_s+P)}\mathbf{f}_k^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+1}/s_{i+1}(k) \right\|^2 \\ =& \left\| \mathbf{S}_i\mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_i - \mathbf{S}_{i+1}\mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+1}e^{-j2\pi\varepsilon K_s} \right\|^2 \end{aligned} \tag{3.3}$$

where $\mathbf{f}_k$ is the kth column of $\mathbf{F}$; $\mathbf{F}_p$ is the $K\times|\mathcal{P}|$ matrix whose columns are obtained from $\mathbf{f}_k$, $k \in \mathcal{P}$; $\mathbf{S}_i$ and $\mathbf{S}_{i+1}$ are $|\mathcal{P}| \times |\mathcal{P}|$ diagonal matrices with the diagonal elements given by $1/s_i(k), 1/s_{i+1}(k), k \in \mathcal{P}$, respectively. A CFO estimator, called p-algorithm, at the noisy case is proposed as

$$\hat{\phi} = \min_{\varepsilon} g_p(\varepsilon). \tag{3.4}$$

If the channel is constant over $\nu > 2$ blocks, the estimator (3.4) can be modified to

$$\hat{\phi} = \min_{\varepsilon} \sum_{u=0}^{\nu-2} \left\| \mathbf{S}_{i+u}\mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+u} - \mathbf{S}_{i+u+1}\mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+u+1}e^{-j2\pi\varepsilon K_s} \right\|^2. \tag{3.5}$$

In the following discussion, we only focus on the case $\nu = 2$ since ν is preferred to be smaller in order to track a varying CFO. An example for the case $\nu > 2$ will be given in the simulation part.

In addition, it is interesting to find that Classen&Meyr's method (2.8) is equivalent to

$$\hat{\phi} = \min_{\varepsilon} \left\| \mathbf{S}_i\mathbf{F}_p^H\mathbf{y}_i - \mathbf{S}_{i+1}\mathbf{F}_p^H\mathbf{y}_{i+1}e^{-j2\pi\varepsilon K_s} \right\|^2. \tag{3.6}$$

Obviously, (3.4) possesses a much more reasonable structure than (3.6), since it firstly rotates $\mathbf{y}_i$, $\mathbf{y}_{i+1}$ by $\mathbf{\Omega}(-\varepsilon)$ and then estimates CFO by comparing the values on pilot carriers. On the contrary, the estimator (3.6) compares those values from the direct DFT of $\mathbf{y}_i$, $\mathbf{y}_{i+1}$, which may only give an acceptable approximation under certain conditions. A further benefit of (3.4) is that, it could possibly provide a full rage estimation of the CFO. A first look on this point is from the fact that (3.4) contains the term $e^{-j2\pi\varepsilon}$, then $\phi \in (-0.5, 0.5]$ is allowed. However, (3.6) only contains the term $e^{-j2\pi K_s\varepsilon}$ so the estimation range is limited to $(-0.5/K_s, 0.5/K_s]$.

3.2.2 Identifiability of p-Algorithm

Similar to [68], [69], we study the uniqueness of the estimator (3.4) under the noise free environment. The unknown CFO is assumed to be within the full region (-0.5, 0.5], and the trivial ambiguity $\hat{\phi} = \phi \pm b$, $b \in \mathcal{I}$ is excluded from the consideration. For the noise free case, (3.4) reduces to

$$\hat{\phi} = \{\varepsilon | g_p(\varepsilon) = 0\}. \tag{3.7}$$

Obviously, the true CFO ϕ is a solution to (3.7). The ambiguity appears if $\exists\ \bar{\phi} \neq \phi$ such that $g_p(\bar{\phi}) = 0$, which is equivalent to

$$\begin{aligned} 0 &= \sum_{v=0}^{K-1} \tilde{h}(v) \underbrace{\frac{e^{\frac{j(K-1)\pi(v+K\Delta\phi-k)}{K}} \sin(\pi(v+K\Delta\phi-k))}{K\sin\left(\frac{\pi(v+K\Delta\phi-k)}{K}\right)}}_{\alpha_{vk}} \\ &\quad \times \underbrace{\left(s_i(v)/s_i(k) - s_{i+1}(v)/s_{i+1}(k)e^{j2\pi\Delta\phi K_s}\right)}_{\beta_{vk}} \\ &= \sum_{v=0}^{K-1} \tilde{h}(v)\alpha_{vk}\beta_{vk}, \end{aligned} \tag{3.8}$$

for all $k \in \mathcal{P}$ and $\Delta\phi \neq 0$, where $\Delta\phi \triangleq \phi - \bar{\phi}$.

Case 1: $K\Delta\phi \notin \mathcal{I}$: In this case, $\alpha_{vk} \neq 0$ for all v. The discussion is further divided into two subcases.

1. Not all $\beta_{vk} = 0$: The ambiguity happens when $\sum_v \tilde{h}(v)\alpha_{vk}\beta_{vk} = 0$. This type of ambiguity will be referred as h-ambiguity. Since $\tilde{h}(v)$ is a linear combination

of continuous complex random variables h_l, the probability for h-ambiguity is zero. Therefore, h-ambiguity can be ignored.

2. All $\beta_{vk} = 0$: We call this kind of ambiguity as d-ambiguity. In order to avoid this type of ambiguity, the value on pilots can be properly designed such that β_{vk} is not zero for some $k \in \mathcal{P}$. For example, we can take $s_i(k_1) = 1$, $s_{i+1}(k_1) = 1$, while choose $s_i(k_2) = 1$, $s_{i+1}(k_2) = -1$. Then, $\beta_{k_1k_1}$ and $\beta_{k_1k_2}$ cannot be zero at the same time.

Case 2: $K\Delta\phi \in \mathcal{I}$ or more specifically, $K\Delta\phi \in \mathcal{I}_{K-1} \triangleq \{1, \ldots, K-1\}$[1]. Let $\tilde{v}_k = \big((k - K\Delta\phi) \mod K\big)$. Obviously, $\tilde{v}_k \neq k$ when $\Delta\phi \neq 0$. In this case, $\alpha_{\tilde{v}_k k} = 1$, and $\alpha_{vk} = 0$, $\forall v \neq \tilde{v}_k$. The ambiguity happens if $\beta_{\tilde{v}_k k} = 0$ or $\tilde{h}(\tilde{v}_k) = 0$. Since the latter can be equivalently considered as if the $\tilde{v}_k$th carrier is a virtual carrier, we incorporate the discussion on $\tilde{h}(\tilde{v}_k) = 0$ into the discussion on $\beta_{\tilde{v}_k k} = 0$, i.e.

$$s_i(\tilde{v}_k)/s_i(k) - s_{i+1}(\tilde{v}_k)/s_{i+1}(k)e^{j2\pi\Delta\phi K_s} = 0, \quad \text{for all } k \in \mathcal{P}. \tag{3.9}$$

Case 2 is divided into three subcases.

1. All $\tilde{v}_k$'s$\in \mathcal{P}$: The ambiguity under this subcase is called p-ambiguity. Two different methods can be used to avoid the p-ambiguity. One is to choose the pilot indices in a way that all pilot carriers are not equi-spaced. Then, the p-ambiguity cannot happen since any $K\Delta\phi \in \mathcal{I}_{K-1}$ could not make all $\tilde{v}_k$'s$\in \mathcal{P}$. However, since equi-spaced pilots can be used for optimal training in some OFDM systems [70] (if the number of the pilot carriers is greater than L), this approach to avoid p-ambiguity is not always recommended. The other way is to design the pilot symbols such that equation (3.9) does not hold for one or more $\tilde{v}_k \in \mathcal{P}$. If the pilot values are allowed to be arbitrary, then the design is quite trivial. However, pilot value should be restricted to the constellation $\mathcal{C}_p$. Here, we will put special attention on pilot values belonging

[1]We only need to consider this subset since $K\Delta\phi' = K\Delta\phi + bK$ only provides a trivial ambiguity in $\Delta\phi' = \Delta\phi + b$.

to $\{+1, -1\}$ as adopted in IEEE 802.11a standards [7]. Then, p-ambiguity can be avoided by properly consider three pilot carriers, say $\{k_1, k_2, k_3\}$. We can set $s_i(k_1) = s_{i+1}(k_1) = s_i(k_3) = s_{i+1}(k_3) = 1$, while taking $s_i(k_2) = 1$, $s_{i+1}(k_2) = -1$. Then, (3.9) does not hold simultaneously for k_1, k_2 and k_3.

2. All $\tilde{v}_k \notin \mathcal{P}$: This subcase can be further divided into two sub-subcases.

   (a) At least one $\tilde{v}_k, k \in \mathcal{P}$ does not belong to $\mathcal{V} \bigcup \mathcal{N}$, where $\mathcal{N}$ denotes the subcarrier index set for channel nulls. Without loss of generality, we denote this specific k and $\tilde{v}_k$ as k_1 and $\tilde{v}_{k_1}$, respectively. The ambiguity under this sub-subcase is called c-ambiguity. Since the values of $s_i(\tilde{v}_{k_1})$, $s_{i+1}(\tilde{v}_{k_1})$ are selected from a finite alphabet, all the possible values of $\Delta\phi$ in (3.9) should belong to the set

$$\Psi = \left\{ \frac{1}{2\pi K_s} \arg\left(\frac{s_1}{s_2}\right) + \frac{\omega}{K_s} \pm \frac{\iota}{K_s} \middle| \forall s_1, s_2 \in \mathcal{C}_d, \iota \in \mathcal{I} \right\}, \tag{3.10}$$

   where

$$\omega = \frac{1}{2\pi} \arg\left(\frac{s_{i+1}(k_1)}{s_i(k_1)}\right). \tag{3.11}$$

   Let

$$\mu \in \mathcal{A} \triangleq \left\{ \frac{1}{2\pi} \arg\left(\frac{s_1}{s_2}\right) \middle| \forall s_1, s_2 \in \mathcal{C}_d \right\} \tag{3.12}$$

   represent all the possible phase differences for a certain signal constellation $\mathcal{C}_d$. The c-ambiguity can be excluded if

$$\frac{(\mu + \omega + \iota)K}{K_s} \notin \mathcal{I}_{K-1}, \quad \forall \mu \in \mathcal{A}. \tag{3.13}$$

   The constellation $\mathcal{C}_p$, again, brings difficulties to fulfill (3.13). We still focus on the discussions when pilots can only be chosen from the set $\{+1, -1\}$ as adopted in IEEE 802.11a standard. Since the symmetric signal constellations are normally used, e.g., BPSK, QPSK, QAM, both $\frac{1}{2\pi} \arg\left(\frac{s_1}{s_2}\right)$ and $\frac{1}{2\pi}\left(\arg\left(\frac{s_1}{s_2}\right) + \pi\right)$ belongs to $\mathcal{A}$. Therefore, it is sufficient to consider only the case when $s_i(k_1) = s_{i+1}(k_1) = 1$, in which case ω is 0. Then, the c-ambiguity can be excluded if

$$\frac{(\mu + \iota)K}{K_s} \notin \mathcal{I}_{K-1}, \quad \forall \mu \in \mathcal{A}. \tag{3.14}$$

Instead of designing the pilots values, we need to properly choose the signal constellation $\mathcal{C}_d$ and the values of K, K_s. Note that, one value of μ must be zero, then $\frac{\iota K}{K_s} \notin \mathcal{I}_{K-1}$ is required. Therefore, K and K_s should be co-prime numbers.

We give an example to illustrate the relationship between c-ambiguity and system parameters. For simplicity, we consider $K = 16$, $P = 5$ and the QPSK constellation. Then, $\mathcal{A}$ is $\{0, 0.25, 0.5, 0.75\}$, and the only solutions to $K\Delta\phi \in \mathcal{I}_{K-1}$ are

$$\Delta\phi = \begin{cases} 0.25 & \mu = 0.25, \iota = 5 \\ 0.5 & \mu = 0.5, \iota = 10 \\ 0.75 & \mu = 0.75, \iota = 15 \end{cases} . \tag{3.15}$$

At the same time, if the symbol at the $\tilde{v}_k$th carrier satisfies $s_i(\tilde{v}_k) = s_{i+1}(\tilde{v}_k)e^{j2\pi\Delta\phi K_s}$ for any $\Delta\phi$ in (3.15), then $\beta_{\tilde{v}_k k} = 0$. In order to remove the ambiguity, we need to carefully reassign values of K, K_s such that (3.14) is satisfied. One thing to be mentioned is that, if the CFO can be restricted to the region (-0.125, 0.125], c-ambiguity is directly avoided, because the smallest $\Delta\phi$ in (3.15) is 0.25. Note that, this ambiguity free region is five times wider than that of Classen&Meyr's or Beek's method, which is only (-0.5/21,0.5/21] (one subcarrier spacing).

(b) All $\tilde{v}_k$'s, $k \in \mathcal{P}$ belong to $\mathcal{V} \bigcup \mathcal{N}$. We call this type of ambiguity as n-ambiguity. Since the index set for $\tilde{v}_k$ is actually $K\Delta\phi$ cyclic-shift from the set $\mathcal{P}$, we can properly design the number and the positions of pilot carriers such that at least one $\tilde{v}_k$ does not belong to $\mathcal{V} \bigcup \mathcal{N}$ for any cyclic integer shift $K\Delta\phi \in \mathcal{I}_k$. Due to the uncertainty of channel null set, an assuring way is to choose $|\mathcal{P}| > |\mathcal{V}| + |\mathcal{N}|$. However, one may find better choices if the CIR or at least the positions of channel nulls are known. It needs to be mentioned that, the OFDM structure provided in IEEE 802.11a is almost n-ambiguity free, since $\tilde{v}_k, k \in \mathcal{P}$ cannot simultaneously belong to $\mathcal{V}$, and $\mathcal{N}$ is an empty set with probability one.

3. Otherwise: The ambiguity for this subcase can be avoided by the methods in either of the previous two subcases.

Conclusion: Under the noise free condition, the CFO in the region (-0.5, 0.5] can be uniquely determined[2] from the estimator (3.4) by properly designing system parameters, i.e., $\mathcal{P}$, $p_i(k)$, K, K_s and $\mathcal{C}_d$. However, it is hard to arrive at an explicit design that could guarantee the ambiguity avoidance for an arbitrary system. Nevertheless, since we know the guidelines to deal with each type of ambiguity, we could easily derive an ambiguity free solution for any specific system.

Remark: The p-algorithm can be applied not only for CFO tracking but also for CFO acquisition because: 1) It could provide a full range estimation; 2) It can be applied before channel estimation.

3.2.3 Constellation Rotation: A Case Study for IEEE 802.11a WLAN

From the previous discussion, it is known that system parameters should be properly designed to eliminate the CFO estimation ambiguity. However, there exist several constraints that may bring inflexibility when designing some of these parameters. For example, K is generally taken as 2^p so that FFT can be implemented. Meanwhile, symmetric signal constellations are normally adopted, e.g., PSK, QAM.

An example here follows the IEEE 802.11a standard, where the parameters are chosen as $K = 64, P = 16, K_s = 80$, $\mathcal{V} = \{0, 27, \ldots, 37\}$, and $\mathcal{P} = \{7, 21, 43, 57\}$. Every pilot takes the value of $+1$ or -1. Obviously, the d-ambiguity can be removed by assigning ± 1 to different pilots. Meanwhile p-ambiguity and n-ambiguity do not exist due to the position of pilot carriers[3]. We only need to deal with the c-ambiguity that happens when

$$\frac{64}{80}(\mu + \iota) = \frac{4}{5}(\mu + \iota) \in \{1, \ldots, 63\}. \tag{3.16}$$

[2] h-ambiguity that happens with probability zero is ignored here.

[3] Channel nulls on subcarriers are not considered here.

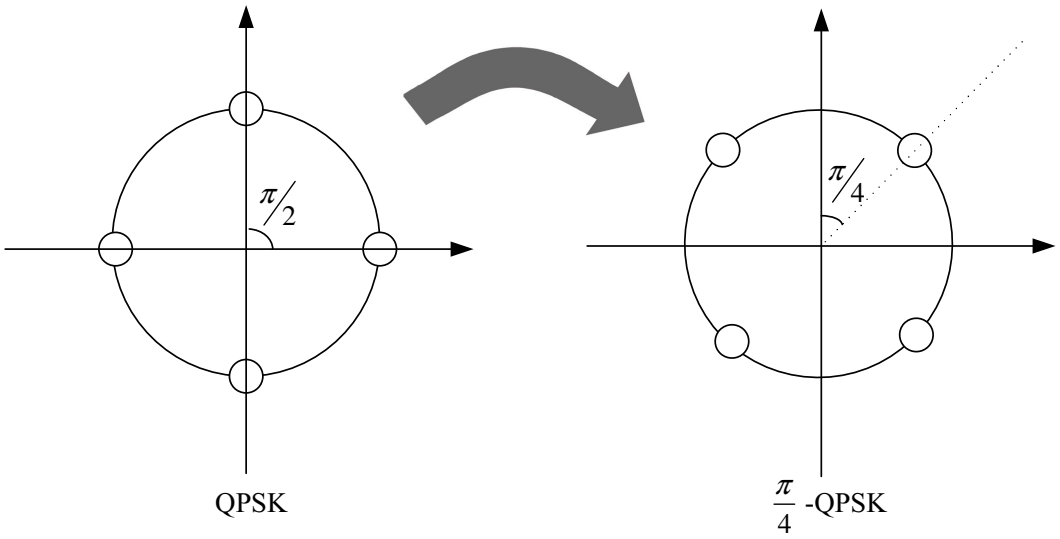

Figure 3.1: Constellation Rotation for QPSK.

It is seen that $\mu \in \{0, 0.25, 0.5, 0.75\}$ or equivalently $\arg\left(\frac{s_1}{s_2}\right) \in \{0, \pi/2, \pi, 3\pi/2\}$ are the only cases that may introduce c-ambiguity. Moreover, since K and K_s are not co-prime numbers, any signal constellation $\mathcal{C}_d$ may cause c-ambiguity. A way to resolve the c-ambiguity is to take double sets of modulations; namely, for OFDM blocks with odd indices, we use signal constellation $\mathcal{C}_{d1}$, whereas for OFDM blocks with even indices, we use signal constellation $\mathcal{C}_{d2}$. Let $\bar{s}_1$ be an arbitrary symbol in $\mathcal{C}_{d1}$ and $\bar{s}_2$ be an arbitrary symbol in $\mathcal{C}_{d2}$. Then, $\mathcal{C}_{d1}$ and $\mathcal{C}_{d2}$ should be designed such that $\arg\left(\frac{\bar{s}_1}{\bar{s}_2}\right) \notin \{0, \pi/2, \pi, 3\pi/2\}$.

To keep the system BER performance unaffected, we suggest a constellation rotation scheme, i.e., $\mathcal{C}_{d2}$ is a rotation from $\mathcal{C}_{d1}$ by a proper angle θ. For example, $\mathcal{C}_{d1}$ is taken as QPSK, while $\mathcal{C}_{d2}$ is taken as $\pi/4$-QPSK. As shown in Fig. 3.1, $\arg\left(\frac{\bar{s}_1}{\bar{s}_2}\right)$ only belongs to $\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$. Hence, the c-ambiguity can be totally removed for noise-free case. Actually, $\mathcal{C}_{d2}$ can be rotated from $\mathcal{C}_{d1}$ by an arbitrary $\theta \in (0, \pi/2)$ for a QPSK constellation under noise free conditions. However, an optimal selection of θ for noisy environment is not strictly derived yet. Due to the symmetry between the two constellations, a good choice may be $\theta = \pi/4$ because the minimum $\arg\left(\frac{\bar{s}_1}{\bar{s}_2}\right)$ is maximized, which may result in a larger distance between

$s_i(\tilde{v}_k)$ and $s_{i+1}(\tilde{v}_k)e^{j2\pi\Delta\phi K_s}$. We can also show that if $\theta = \pi/4$, c-ambiguity in noise free case is also removed for higher order constellations, e.g., 16-QAM or 64-QAM.

Remark: The above discussion is established in the noise free environment. The ambiguity, or more properly called the outlier, may happen in the presence of noise.

3.2.4 Virtual Carriers Based Tracking: v-Algorithm

The v-algorithm used in this chapter is written as

$$\hat{\phi} = \min_{\varepsilon} \left\|\mathbf{F}_v^H \mathbf{\Omega}(-\varepsilon)\mathbf{y}_i\right\|^2 + \left\|\mathbf{F}_v^H \mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+1}\right\|^2 = \min_{\varepsilon} g_v(\varepsilon), \tag{3.17}$$

where both the ith and the $(i+1)$th received blocks are considered for the consistence with p-algorithm.

3.2.5 Co-Consideration: pv-Algorithm

From previous subsections, we know that p-algorithm uses only pilot carriers, while v-algorithm uses only virtual carriers. Therefore the two algorithms work "independently" from each other. It is noted that, exploiting both pilot carriers and virtual carriers may offer several additional benefits: 1) Improve the estimation accuracy; 2) Reduce the probability of the ambiguity.

A reasonable combination of both p-algorithm and v-algorithm can be expressed as the weighted sum of the two corresponding cost functions. The combined cost function is given by

$$g_{pv}(\varepsilon) = g_p(\varepsilon) + \gamma g_v(\varepsilon) \tag{3.18}$$

where γ is the weight whose optimal value will be derived later.

The first benefit comes from the fact that $g_{pv}(\varepsilon)$ exploits more resources than either $g_p(\varepsilon)$ or $g_v(\varepsilon)$. The second benefit can be explained in the absence of the noise. Suppose $\exists\bar{\phi} \neq \phi$ such that $g_p(\bar{\phi}) = 0$. From intuition, since $g_p(\bar{\phi})$ and $g_v(\bar{\phi})$ are obtained through different approaches and possess different structures, the probability for $\bar{\phi}$ to also be the null point for $g_v(\varepsilon)$ is small. However, the true

CFO ϕ must be the null point for both $g_p(\varepsilon)$ and $g_v(\varepsilon)$. Therefore, after the addition, the false null in either estimator will be compensated by the other estimator, leaving only the true ϕ being the null of $g_{pv}(\varepsilon)$. Note that, the ambiguity is still possible to happen once ϕ is the common null of both $g_p(\varepsilon)$ and $g_v(\varepsilon)$. However, the probability may be greatly reduced compared with using either estimator. Furthermore, to eliminate the ambiguity for *pv*-algorithm, we only need to eliminate the ambiguity for *p*-algorithm (as in subsection 3.2.2), or to eliminate the ambiguity for *v*-algorithm [59].

An example regarding the second benefit is shown here. For simplicity, the parameters are taken as $L = 3$, $K = 16$, $P = 5$, $\phi = 0.25$, $\gamma = 1$, $\mathcal{P} = \{11\}$, $\mathcal{V} = \{13, 14, 15\}$. Symbols are generated from QPSK constellations and pilots are selected as $p_i(11) = p_{i+1}(11) = +1$. We also set $\tilde{h}(12) = 0$. From [59], the ambiguity for *v*-algorithm takes place at the point $\varepsilon = 0.1875$. For *p*-algorithm, the *c*-ambiguity may happen at the point $\varepsilon \in \{0, -0.25, 0.5\}$ as shown in (3.15). The *p*-ambiguity does not exist since there is only one pilot carrier. The *d*-ambiguity only happens when the two OFDM blocks are exactly the same and can be ignored in this example. Meanwhile, the *n*-ambiguity takes on place at $\varepsilon \in \{0.3750, 0.4375, 0.5\}$ because of the virtual carriers, and at $\varepsilon = 0.3125$ because of the channel null. It is seen that the *n*-ambiguity, caused by the channel nulls, in *p* algorithm shifts to the opposite direction as that in *v*-algorithm, i.e., $0.1875 = 0.25 - 1/16$ while $0.3125 = 0.25 + 1/16$. Hence, there is only one common null for both *p*-algorithm and *v*-algorithm, which corresponds to the true CFO value.

The cost function (3.4), (3.17), (3.18) of *p*-algorithm, *v*-algorithm, and *pv*-algorithm in the absence of the noise are shown in Fig. 3.2 and Fig. 3.3. From these two figures, although both *p*-algorithm and *v*-algorithm suffer from their respective ambiguity, the *pv*-algorithm has a unique null at the true CFO value, because the false null of either algorithm is compensated by the other algorithm.

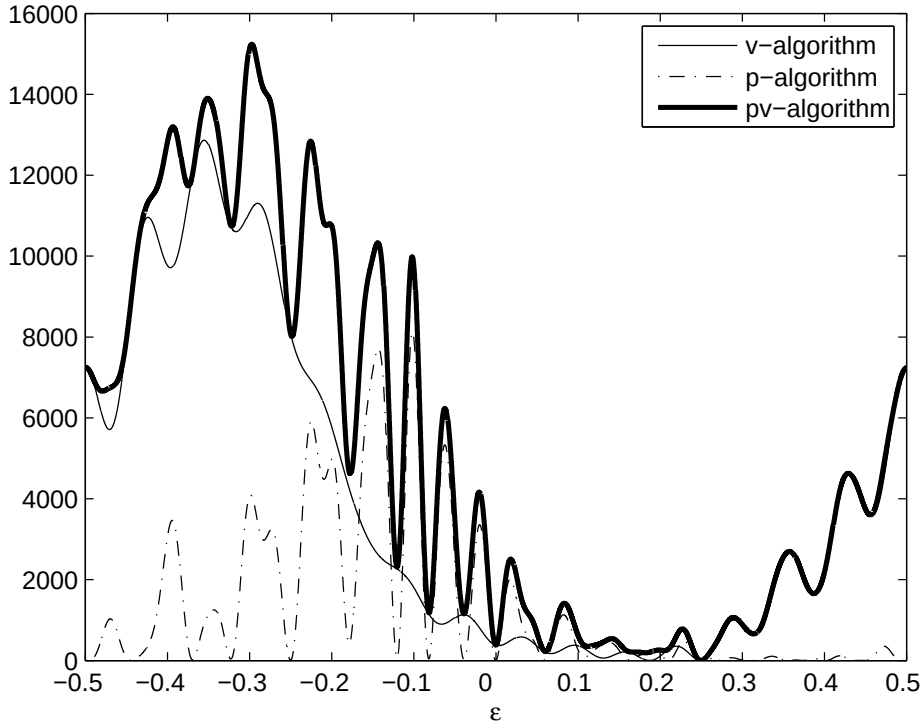

Figure 3.2: CFO pattern for p-algorithm, v-algorithm and pv-algorithm.

3.2.6 Ways to Obtain CFO from p- and pv-Algorithms

The direct way to estimate ϕ from either $g_p(\varepsilon)$ or $g_{pv}(\varepsilon)$ is the one dimensional searching. However, the complexity of such kind of searching, although acceptable in some applications, e.g., direction of arrival (DOA) estimation [71], is too high to be implemented in other real time applications. To avoid computationally expensive searching, several alternatives are considered. We will illustrate these methods for pv-algorithm whereas the extensions to p-algorithm would be an easier job.

FFT Based Method

The estimator (3.18) can be expanded as

$$g_{pv}(\varepsilon) = \rho_0 + 2\Re\left\{\sum_{i=1}^{K+K_s} \rho_i e^{-j2\pi i\varepsilon}\right\} \tag{3.19}$$

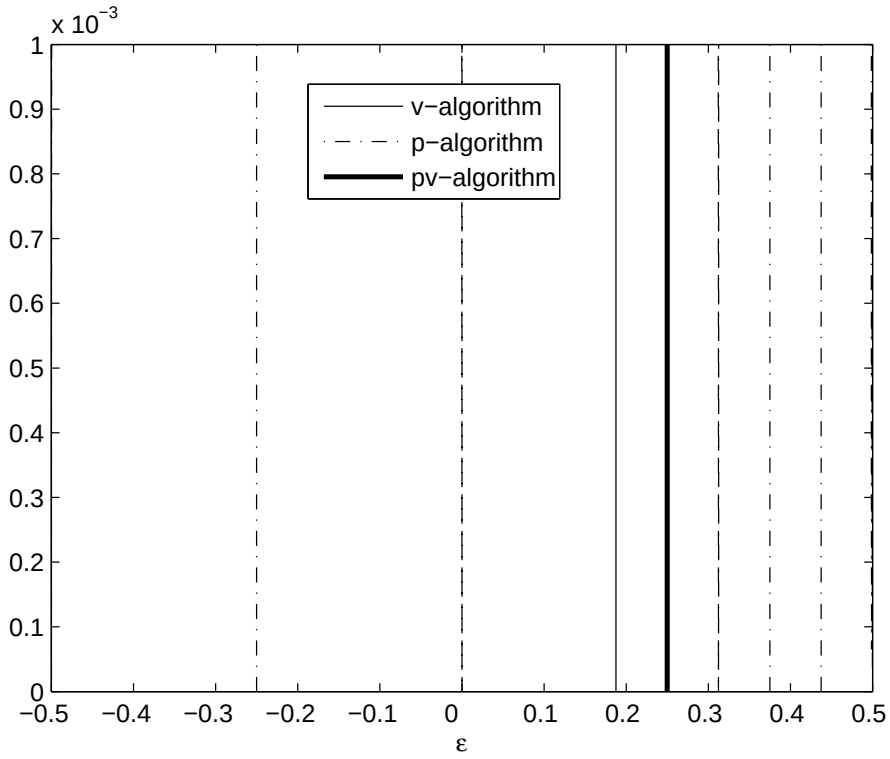

Figure 3.3: Scope-enlarged CFO pattern.

where ρ_i's are coefficients that can be obtained from (3.18) straightforwardly. From [62], we know that the minimization of the cost function can be achieved through FFT.

Polynomial Rooting

Polynomial rooting for a single ε has been proposed in [54]. This method is able to guarantee the achieving of the global minimum of $g_{pv}(\varepsilon)$, and its complexity can be approximated by $O((K + K_s)^3)$. Compared with the FFT based searching, the polynomial rooting is recommended for smaller K and K_s.

Adaptive Method

Adaptive algorithm can be utilized in certain cases in order to reduce the computational complexity. However, the major bottleneck of adaptive method is that, the initial point should be accurate enough to guarantee a global minimum. Nonetheless, the residue CFO after the coarse estimation is normally small and the tracking of the slow varying CFO could be reliably accomplished by the adaptive method.

3.3 Timing Offset Estimation

Although our major concern is the CFO tracking, we would like to also examine the capability of *pv*-algorithm on TO tracking. Let τ represent TO, the ith received OFDM block is then

$$\mathbf{y}_i = [v_i(P+1+\tau), \ldots, v_i(K_s-1), v_{i+1}(1), \ldots, v_{i+1}(\tau)]^T. \tag{3.20}$$

In this case, the joint TO and CFO estimation metric is written as

$$\{\hat{\phi}, \hat{\tau}\} = \min_{\varepsilon,\varrho \in \mathcal{I}} g_{pv}(\varepsilon, \varrho) \tag{3.21}$$

where $g_{pv}(\varepsilon, \varrho)$ is obtained in a similar way as (3.4) but using the new $\mathbf{y}_i$ instead.

Obviously, $g_{pv}(\varepsilon, \varrho) = 0$ at $\{\varepsilon, \varrho\} = \{\phi, \tau\}$ under the noise free case. Note that we need to perform a two dimensional search to solve (3.21), which usually has a high complexity. However τ is an integer and is limited to a short range after the acquisition stage. For each possible ϱ, we can solve (3.21) for ε. If the range of ϱ is short, the overall complexity can be kept reasonably low. Besides, the residue TO cannot varying from time to time, so the TO tracking does not need to be performed frequently. Similar to the TO estimation method in [26], there will be a plateau if the TO estimation error $\Delta\tau = \hat{\tau} - \tau$ stay in region $\mathbb{B}$ as shown in Fig. 2.1. This is because that the block starting in this region only causes a fixed phase rotation at each subcarriers. Then, the system performance will not be affected because the rotation is the same for the two consecutive OFDM blocks. Therefore, $\Delta\tau$ in

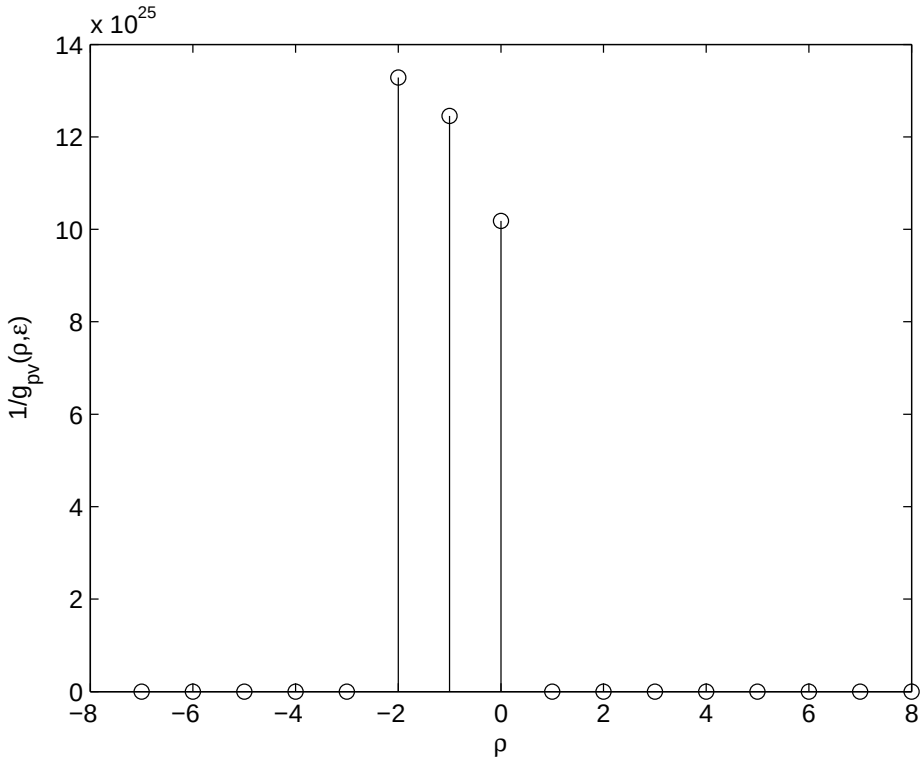

Figure 3.4: TO estimation metric versus the sample index, noiseless case.

$[L-P, 0]$ can be considered as a successful TO tracking. Further effort to obtain the perfect synchronization can be performed if the exact channel length is known [72].

An example of the proposed TO tracking is given here with parameters $K = 16$, $L = 3$, $P = 5$, $\phi = 0.25$. Signals are obtained from QPSK constellation, and noiseless environment is considered. We assume that during the tracking period, the TO is already restricted to the region $[-K/2, K/2]$ thanks to the TO acquisition at the very beginning of transmission. For simplicity, we only show the timing metric $1/\min_{\varepsilon} g_{pv}(\varepsilon, \varrho)$ versus sampling time index in Fig. 3.4. Clearly, TO is perfectly estimated within the region $[-2, 0]$.

For noisy environment, the performance of TO estimation can be evaluated by

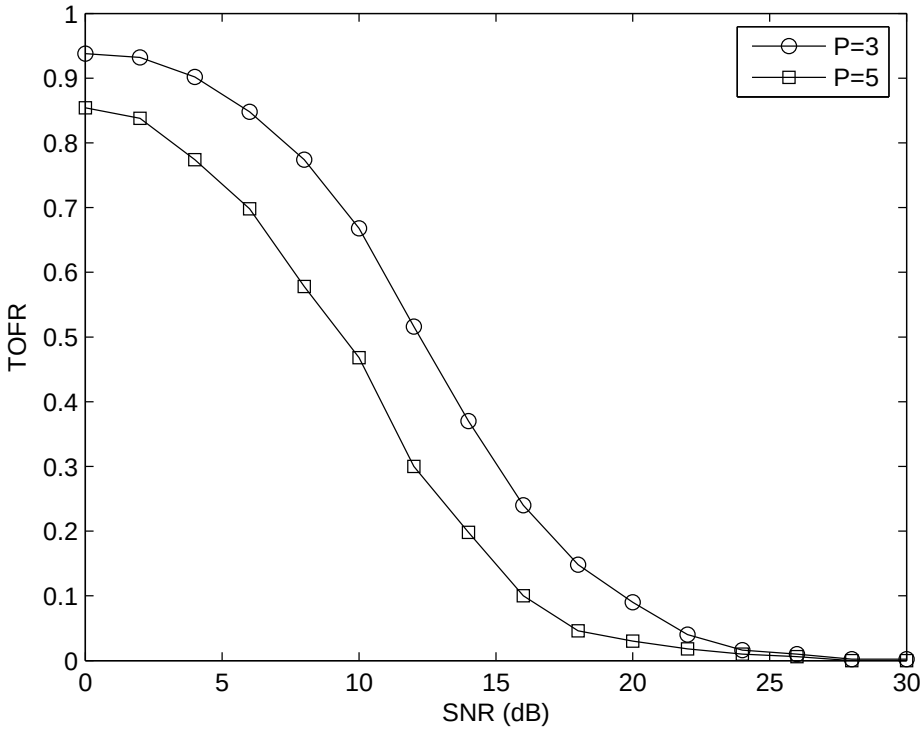

Figure 3.5: TOFRs versus the SNR in the presence of noise.

TO failure rate (TOFR) defined as

$$\text{TOFR} = \frac{\text{the number of runs that give } \Delta\tau \text{ out of the region } [-2, 0]}{\text{the total number of Monte Carlo runs}}. \tag{3.22}$$

We then show TOFR versus the SNR for different value of P in Fig. 3.5. Clearly, the proposed TO estimator (3.21) is effective at relatively high SNR.

3.4 Performance Analysis of CFO Tracking

For the ease of analysis, we assume that all pilots take values of $+1$. At high SNR, the expectation and the variance of the proposed estimator can be approximated by

[62]:

$$\mathrm{E}_{pv}\{\hat{\phi}\} \cong \phi - \frac{\mathrm{E}\{\dot{g}_{pv}(\phi)\}}{\mathrm{E}\{\ddot{g}_{pv}(\phi)\}} = \phi - \frac{\mathrm{E}\{\dot{g}_{p}(\phi) + \gamma\dot{g}_{v}(\phi)\}}{\mathrm{E}\{\ddot{g}_{p}(\phi) + \gamma\ddot{g}_{v}(\phi)\}} \tag{3.23}$$

$$\mathrm{var}_{pv}\{\hat{\phi}\} \cong \frac{\mathrm{E}\{[\dot{g}_{pv}(\phi)]^2\}}{[\mathrm{E}\{\ddot{g}_{pv}(\phi)\}]^2} = \frac{\mathrm{E}\{[\dot{g}_{p}(\phi) + \gamma\dot{g}_{v}(\phi)]^2\}}{[\mathrm{E}\{\ddot{g}_{p}(\phi) + \gamma\ddot{g}_{v}(\phi)\}]^2} \tag{3.24}$$

where $\dot{g}(\phi)$ and $\ddot{g}(\phi)$ represent the first and the second order derivatives of $g(\varepsilon)$ at $\varepsilon = \phi$. As derived in Appendix A, (3.23) and (3.24) can be simplified as

$$\mathrm{E}_{pv}\{\hat{\phi}\} \cong \phi \tag{3.25}$$

$$\mathrm{Var}_{pv}\{\hat{\phi}\} \cong \frac{\sigma^2}{8\pi^2}\frac{2Z_p + \gamma^2 Z_v}{(Z_p + \gamma Z_v)^2} \tag{3.26}$$

where

$$Z_p = \left\|\mathbf{P}_p(\mathbf{D}\Delta\boldsymbol{\eta} - K_s\boldsymbol{\eta}_{i+1}))\right\|^2 \tag{3.27}$$

$$Z_v = \sum_{m=i}^{i+1} \boldsymbol{\eta}_m^H \mathbf{D}\mathbf{P}_v\mathbf{D}\boldsymbol{\eta}_m \tag{3.28}$$

and all the variables are defined in Appendix A. Note that $\mathbf{P}_p\boldsymbol{\eta}_{i+1}$ is the term only related to the variable $\tilde{h}(k), k \in \mathcal{P}$, and $\mathrm{Var}_{pv}\{\hat{\phi}\}$ is not related to the specific CFO value. Meanwhile, the CFO estimation MSEs by using $g_p(\varepsilon)$ and $g_v(\varepsilon)$ are

$$\mathrm{Var}_{p}\{\hat{\phi}\} \cong \frac{\sigma^2}{4\pi^2 Z_p} \tag{3.29}$$

$$\mathrm{Var}_{v}\{\hat{\phi}\} \cong \frac{\sigma^2}{8\pi^2 Z_v} \tag{3.30}$$

respectively. It is interesting to find that

$$\mathrm{Var}_{pv}\{\phi\}|_{\gamma=0} = \mathrm{Var}_{p}\{\phi\} \tag{3.31}$$

$$\mathrm{Var}_{pv}\{\phi\}|_{\gamma\to\infty} = \mathrm{Var}_{v}\{\phi\}. \tag{3.32}$$

Therefore the value of γ controls the effective part of each single estimator.

A nice property of weighted sum of p-algorithm and v-algorithm is that, the closed form of the optimal weight γ can be obtained regardless of all other parameters. Taking the derivative of (3.26) with respect to γ, we arrive at that, the minimum value of $\mathrm{Var}_{pv}\{\hat{\phi}\}$ is always achieved at $\gamma = 2$.

3.5 Simulations

In this section, we examine the performance of the proposed estimators under various scenarios. All parameters are taken from IEEE 802.11a standard. The 4-ray channel model with an exponential power delay profile [73]

$$E\{|h_l|^2\} = \rho \exp(-l/10), \quad l = 0, \ldots, L \tag{3.33}$$

is used where ρ is the coefficient to normalize the overall channel gain. Each channel path is complex Gaussian. The normalized estimation mean square errors (NMSE) is defined as

$$\text{NMSE} = \frac{1}{M_q} \sum_{i=1}^{M_q} \frac{(\hat{\phi}_i - \phi)^2}{\phi^2} \tag{3.34}$$

where $M_q = 500$ Monte-Carlo runs are taken for average.

1) CFO Less than One Subcarrier Spacing.

In this example, different CFOs are taken from the region (-0.5/64,0.5/64]. The performance of p-algorithm, v-algorithm, pv-algorithm, Classen&Meyr's method and Beek's method are compared. We assume that the estimated channel length is $\hat{L} = 12$ in order to give a fair comparison between the Beek's method and our proposed algorithms[4]. Furthermore, QPSK constellation is used for all OFDM blocks. The NMSEs versus SNR for different algorithms are shown in Fig. 3.6, and the theoretical results for p-algorithm, v-algorithm, and pv-algorithms are given as well.

As is seen from the figure, Classen&Meyr's method can give a relatively satisfying performance at lower SNR with a normalized CFO 0.1 subcarrier spacing. However, at high SNR, Classen&Meyr's method has an error floor. Meanwhile, even when CFO is as small as 0.25 subcarrier spacing, the Classen&Meyr's method fails because the ICI term cannot be ignored any more. On the contrary, since our p-algorithm does not make any approximation, it does not have an error floor and is also valid for a large CFO value. Meanwhile, Beek's method, p-algorithm and pv-algorithm give comparable performance. However, our major concern is that the performance of Beek's method is greatly affected by the channel length, or the

[4]The length of the region $\mathbb{B}$ should be the same as $|\mathcal{P}|$ for fairness [58].

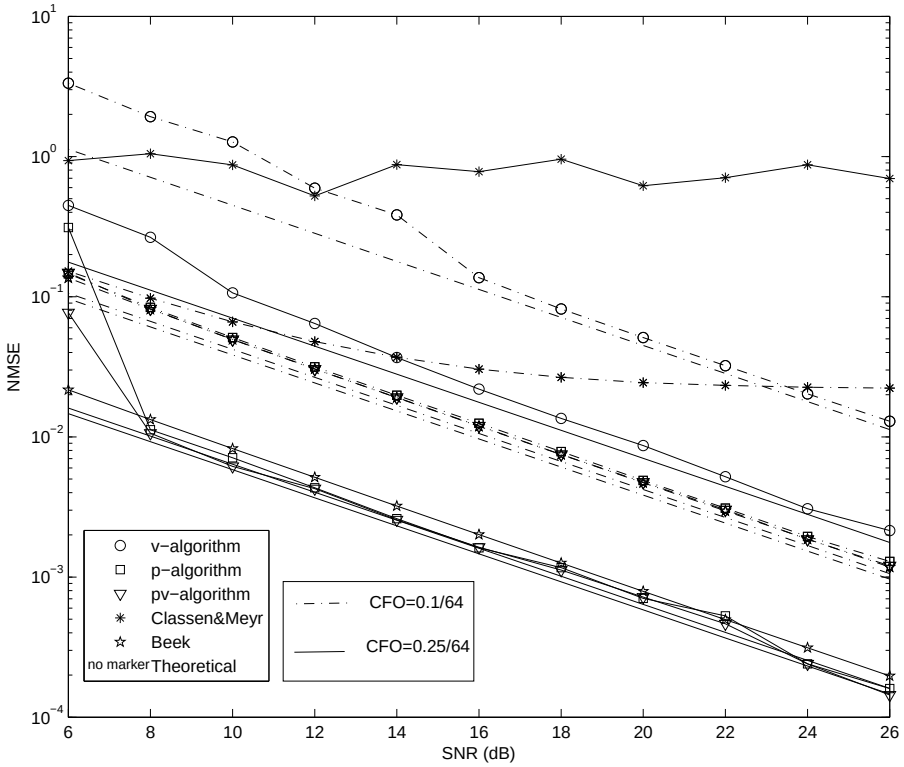

Figure 3.6: NMSEs versus SNR for different CFO estimation algorithm: CFO smaller than subcarrier spacing.

estimate of the channel length. For example, if $L = 16$, or if $L < 16$ but the estimated $\hat{L}$ is 16 due to the power leakage, then the Beek's method cannot even be applied. For v-algorithm, although no error floor is met, the performance is much worse than either p-algorithm or pv-algorithm. This is because that v-algorithm only consider the orthogonality between subcarriers and is actually a blind type CFO estimation method. From intuition, pilot aided algorithm outperforms blind algorithm. Note that there are three dashed lines without any marker. These lines, ordered from top down, represent the theoretical NMSE of v-algorithm, p algorithm, and pv-algorithm for CFO equaling to 0.1 subcarrier spacing, respectively. Similar discussions hold for three solid lines in the figure for CFO equaling 0.25 subcarrier spacing. We find that the numerical performance of p-algorithm and pv-algorithm

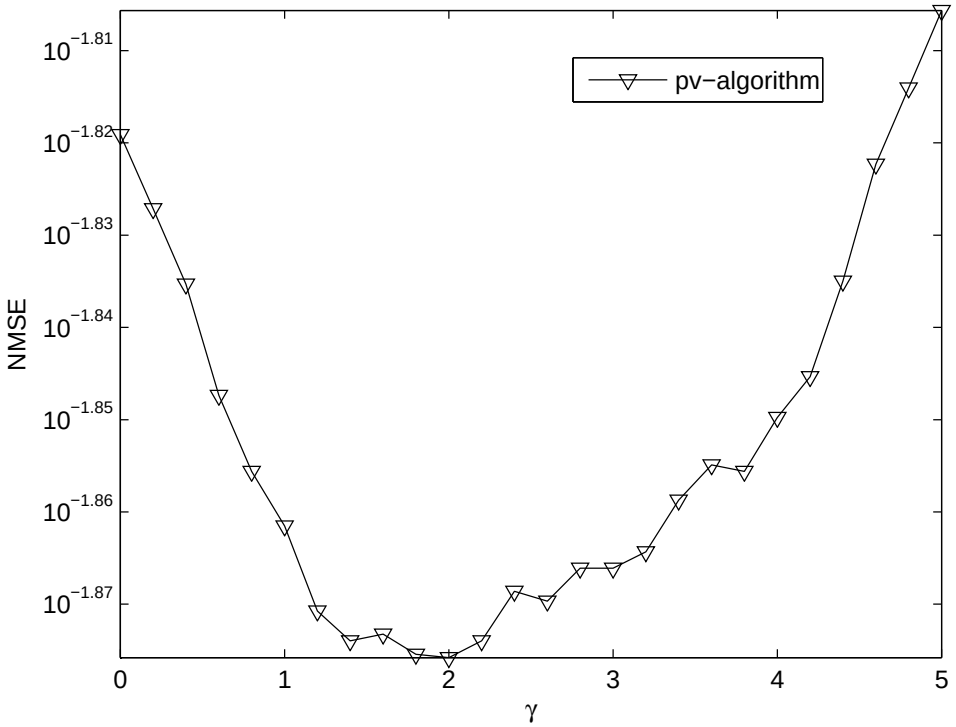

Figure 3.7: NMSEs for *pv*-algorithm under different weight γ.

agree with theoretical results very well, which corroborates our analytical studies.

We also verify the optimality of the choice for γ in Fig. 3.7, which shows the NMSE of *pv*-algorithm versus different value of γ at SNR=15 dB. Clearly, $\gamma = 2$ is the optimal weight, which agrees with the theoretical result. Since *p*-algorithm is the dominant component of *pv*-algorithm, the performance of *pv*-algorithm does not depend critically on the choice of γ. From Fig. 3.7, we find that the NMSE value does not change too much within the region $\gamma \in [0, 5]$.

2) CFO Larger than One Subcarrier Spacing

One important contribution of our proposed algorithm is its applicability for CFO greater than one subcarrier spacing. In this example, we consider the performance of *p*-algorithm, *v*-algorithm, *pv*-algorithm. Note that, Classen&Meyr's method and Beek's method are not included here because they are not applicable for this scenario. The constellation schemes with and without rotation are compared.

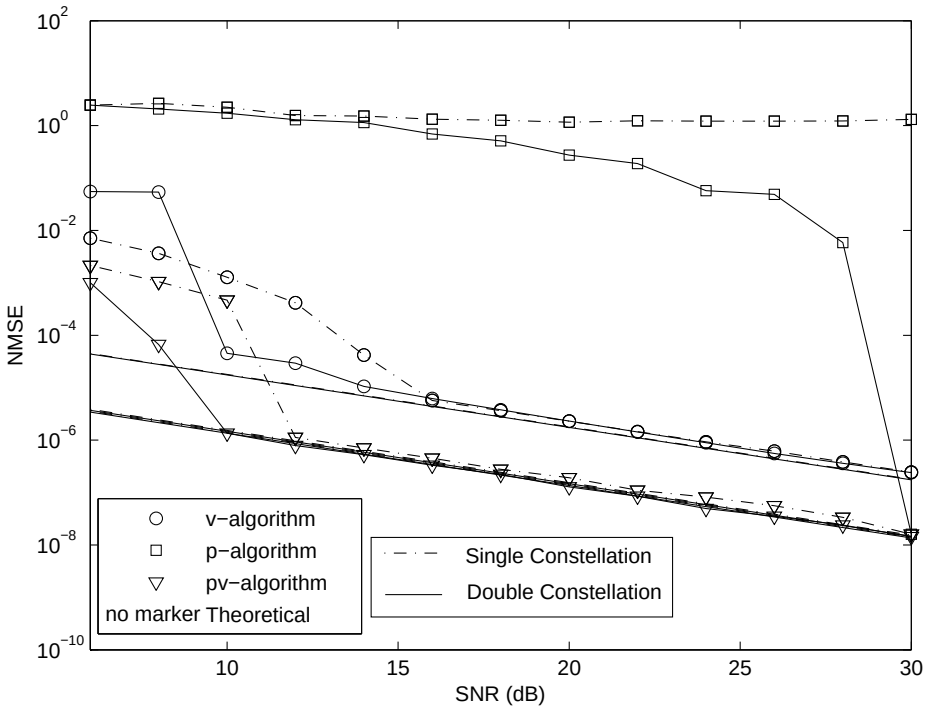

Figure 3.8: NMSEs versus SNR for different CFO estimation algorithm: CFO larger than subcarrier spacing.

For the former scheme, $\mathcal{C}_{d1}$ is QPSK and $\mathcal{C}_{d2}$ is $\pi/4$-QPSK, while for the latter scheme, QPSK constellation is used for all OFDM blocks. The CFO is taken as large as 0.25 of the total bandwidth, which is 16 subcarrier spacings. NMSEs versus SNR are shown in Fig. 3.8. It is seen from Fig. 3.8 that pv-algorithm is about 12 dB better than v-algorithm and gives accurate estimation over all SNRs. However, it is also noted that p-algorithm with constellation rotation cannot yield good performance for SNR$<$ 30 dB, and p-algorithm without constellation rotation fails at all SNR. In the simulations, we have observed that several Monte-Carlo runs give outlier on CFO estimation. From Fig. 3.2, we see that the noiseless cost function $g_p(\varepsilon)$ is close to zero at several locations. If the noise is present, those close-to-zero points may yield the minimum value in the cost function $g_p(\varepsilon)$, which causes CFO outlier. In section 3.2.2, we only provide the discussion on the ambiguity elimination for noise

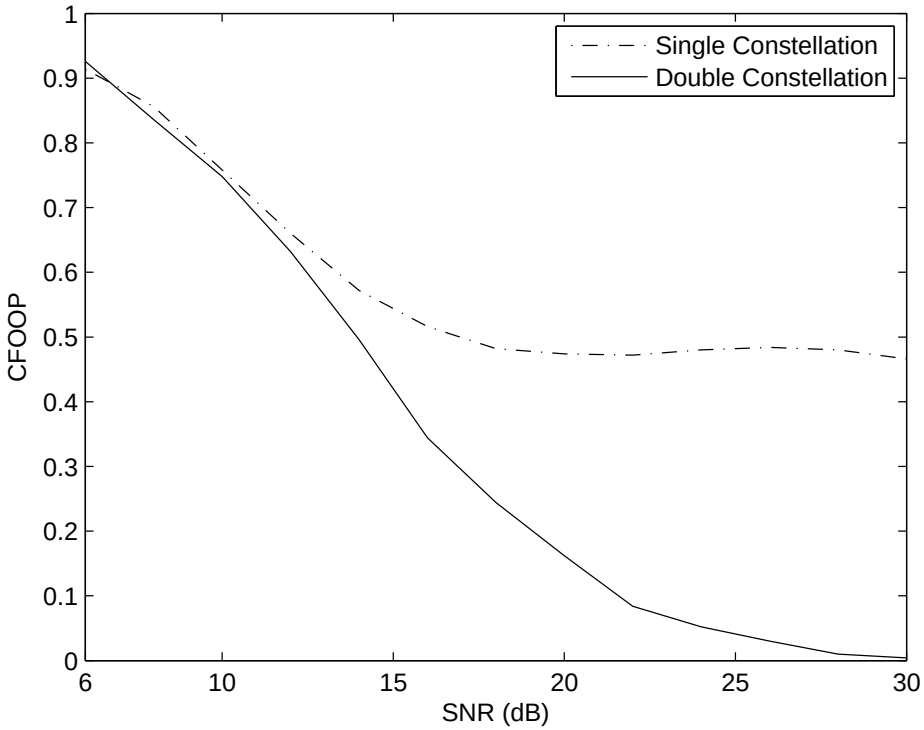

Figure 3.9: CFOOP versus SNR for p-algorithm: Comparison of two modulation schemes.

free environment. If the noise is present, the outlier may happen. Nevertheless, our pv-algorithm benefits from both algorithms. The p-algorithm part increases the estimation accuracy while the v-algorithm part reduces the outlier probability.

A reasonable way to evaluate the advantages of the constellation rotation scheme is to consider the CFO outlier probability (CFOOP), which is defined as

$$\text{CFOOP} = \frac{\text{the number of runs with outlier}}{\text{the total number of Monte Carlo runs}}. \tag{3.35}$$

where the outlier in the presence of the noise is defined as follows.

Definition: In noisy environment, the outlier occurs if the estimated $\hat{\phi}$ stays outside the region $[\phi - 0.5/K, \phi + 0.5/K]$.

The comparison of CFOOP for p-algorithm with different constellation schemes is shown in Fig. 3.9. Clearly, the CFOOP is reduced to zero at high SNR

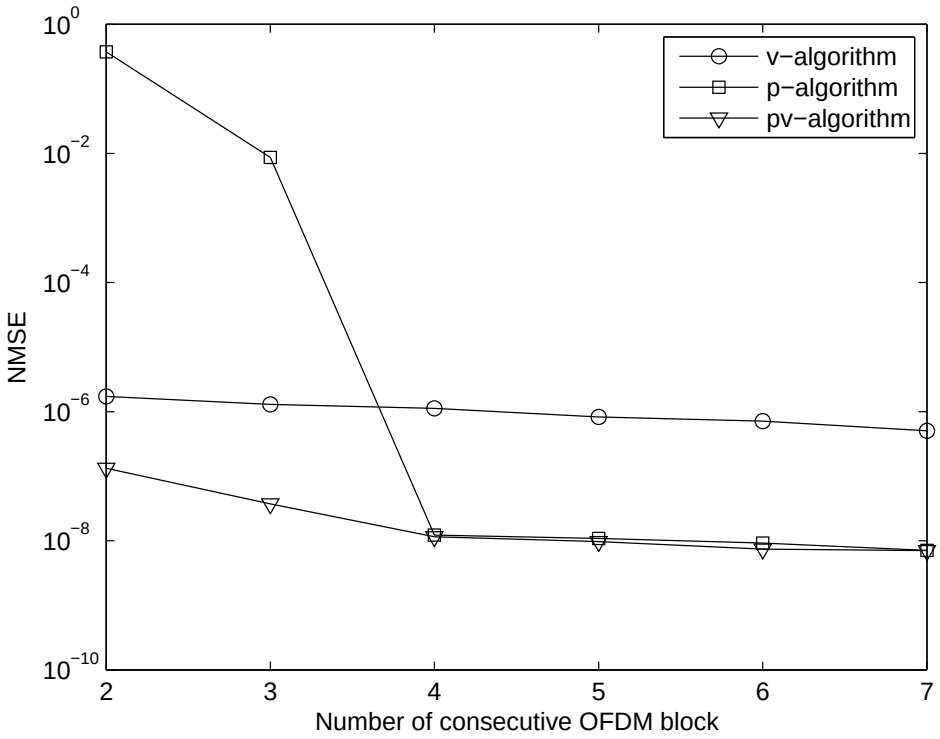

Figure 3.10: NMSEs versus number of the consecutive OFDM blocks: CFO lager than subcarrier spacing.

using the constellation rotation scheme whereas it meets an floor at high SNR without constellation rotation. As analyzed in subsection 3.2.2, the c-ambiguity for non-rotation scheme can never be removed, which consequently introduces non-zero CFOOP all the time.

It has been shown that the proposed p-algorithm, pv-algorithm can be readily modified when channel and CFO are constant over more than two consecutive OFDM blocks. Using the same parameters as in Fig. 3.8, we show the NMSEs of different algorithms versus the number of the consecutive OFDM blocks in Fig. 3.10. The SNR is fixed at 20 dB. From Fig. 3.10 we find that the outlier of p-algorithm in noisy environment can also be "removed" by increasing the number of blocks used. Actually, the probability of the occurrence of the outlier is reduced when more OFDM blocks are used.

3.6 Summery

In this chapter, a novel CFO tracking method was developed for practical OFDM systems. The proposed algorithm considers both the pilot carriers and the virtual carriers, hence is compatible to most practical standards. The ambiguity for pilot based algorithm was studied and several approaches to remove different types of the ambiguity were discussed. Since many parameters are already fixed in existing standards, a constellation rotation scheme was suggested to remove the c-ambiguity effect. Performance of the algorithms were analyzed, and numerous simulation results were presented to validate the theoretical results. From the simulation results, we found that the pilot based algorithm increases the estimation accuracy while the virtual carrier based algorithm reduces the outlier probability. Therefore, the proposed algorithm is not only an effective method for CFO tracking but is also applicable for CFO acquisition.

Chapter 4

Subspace Blind Channel Estimation for CP-Based MIMO OFDM Systems

A novel SS method for CP-based MIMO OFDM systems is proposed in this chapter with the aid of an appropriate re-modulation on the received signal blocks. The feature that the number of the receive antennas is allowed to be the same as the number of the transmit antennas shows great compatibility with the coming 4G wireless communication standards as well as most existing SISO OFDM standards.

4.1 Introduction

OFDM combined with multiple antennas at both the transmitter and the receiver sides has received considerable attention for its capability to combat the multipath fading and boost the system capacity [13], [14]. Consequently, MIMO OFDM appears as a promising candidate for the coming 4G wireless communications [74]. The channel estimation issues for the MIMO OFDM have been intensively studied in the recent literatures [75]-[80]. Several training based channel estimation methods were developed in [75]-[77]. However, the amount of the training increases

dramatically with the increment of the number of the transmit antennas [76], [77], which in turn, decreases the system bandwidth efficiency [81]. For packet based transmission, where the channel state information (CSI) is stable for certain number of blocks, the blind channel estimation could be applied to reduce the number of the training symbols while provides satisfactory performance. Therefore, blind methods have received considerable attention during the past decade.

A major blind approach is the direct inheritance from the so called SS algorithm which was originally developed in [40] for SIMO frequency selective channels. The SS method has simple structure and achieves good performance, but it meets several difficulties when being applied to MIMO OFDM systems [78]-[80]. First, it requires the number of receive antennas to be greater than the number of transmit antennas. Unfortunately, the symmetric links normally play a major role in the standard wireless transmission. For example the 2×2 MIMO is the mandatory operation for every IEEE 802.11n device [8], [9]. Besides, considering the equal number of the transceiver antennas is beneficial since it is also compatible with the current SISO OFDM transmission scheme, e.g., IEEE 802.11a standards [7]. Second, although the SS method can be directly applied when there are more receive antennas than transmit antennas, the precise knowledge of the channel order must be obtained, which is very difficult in practice. The order over-estimation may produce an ill-conditioned channel matrix which greatly deteriorates or sometimes fails the channel estimation. To solve this problem, a zero-padding (ZP) based MIMO OFDM has been suggested in [55]. Instead of using CP, consecutive zeros are padded at the end of each OFDM block. This method will be referred as ZPSOS throughout the thesis. Although ZPSOS exhibits many advantages, a major problem that prevents its application is the incompatibility to most existing OFDM standards or the further 4G MIMO-OFDM standards [9].

In this chapter, we develop a new SS method that is applicable for CP based MIMO OFDM systems through an appropriate re-modulation on the received signal blocks. The method will be, correspondingly, named as CPSOS. The proposed

method possesses all the advantages from ZPSOS but is also compatible with the current and the future OFDM standards. Consequently, CPSOS shows a great potential as a promising blind channel estimation candidate in MIMO OFDM systems. We also provide thorough performance analysis for CPSOS. It is shown that the asymptotical channel estimation MSE agrees with the asymptotically approximated Cramér-Rao bound (ACRB).

4.2 System Model of MIMO OFDM

We consider the $N_t \times N_r$ MIMO OFDM system shown in Fig. 1.4. Notations are the same as those in subsection 1.3. The information symbols are divided into N_t streams and each stream is transmitted from one transmitter. Specifically, each stream will be grouped into blocks of length K. Let

$$\mathbf{s}_i^{(p)} = [s_i^{(p)}(0), s_i^{(p)}(1), \ldots, s_i^{(p)}(K-1)]^T \quad p = 1, 2, \ldots, N_t, \\ i = 0, 1, \ldots, M-1$$

be the block symbol to be transmitted by the transmitter p during the ith OFDM block (before IDFT) and $\mathbf{x}_i^{(p)}$ be the normalized IDFT of $\mathbf{s}_i^{(p)}$.

Then, the overall ith transmitted OFDM block from user p is

$$\mathbf{u}_i^{(p)} = \begin{bmatrix} \mathbf{x}_{i,P}^{(p)} \\ \mathbf{x}_i^{(p)} \end{bmatrix} = \mathbf{T}_{\text{cp}} \mathbf{u}_i^{(p)} \tag{4.1}$$

where $\mathbf{x}_{i,P}^{(p)}$ is the CP that contains the last P entries of $\mathbf{x}_i^{(p)}$, and $\mathbf{T}_{\text{cp}}$ is the corresponding CP insertion matrix. Let

$$\mathbf{h}_{pq} = [h_{pq}(0), h_{pq}(1), \ldots, h_{pq}(L_{pq})]^T$$

be the channel response from the transmitter p to the receiver q, where L_{pq} is the corresponding channel order upper bounded by P. For convenience, we will pad $P - L_{pq}$ zeros at the end of $\mathbf{h}_{pq}$ such that they have a uniform length P. The

received ith block (of length $K_s = K + P$) on the qth receiver is then represented by

$$\mathbf{v}_i^{(q)} = \begin{bmatrix} \mathbf{y}_{i,P}^{(q)} \\ \mathbf{y}_i^{(q)} \end{bmatrix} = \sum_{p=1}^{N_t} \boldsymbol{\mathcal{H}}(\mathbf{h}_{pq}) \begin{bmatrix} \mathbf{x}_{i-1,L}^{(p)} \\ \mathbf{u}_i^{(p)} \end{bmatrix} + \mathbf{w}_i^{(q)} \tag{4.2}$$

where $\boldsymbol{\mathcal{H}}(\cdot)$ is the operation with respect to the argument inside the bracket:

$$\boldsymbol{\mathcal{H}}(\mathbf{h}_{pq}) = \underbrace{\left.\begin{bmatrix} h_{pq}(P) & \dots & h_{pq}(0) & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & h_{pq}(P) & \dots & h_{pq}(0) \end{bmatrix}\right\} K_s \text{ blocks}}_{(K_s+P) \text{ blocks}} \tag{4.3}$$

and $\mathbf{y}_{i,P}^{(q)}$, $\mathbf{y}_i^{(q)}$ have the structures

$$\mathbf{y}_{i,P}^{(q)} = [y_{i,P}^{(q)}(0), y_{i,P}^{(q)}(1), \dots, y_{i,P}^{(q)}(P-1)]^T \tag{4.4}$$

$$\mathbf{y}_i^{(q)} = [y_i^{(q)}(0), y_i^{(q)}(1), \dots, y_i^{(q)}(K-1)]^T. \tag{4.5}$$

The ith noise block on the qth receiver is denoted by

$$\mathbf{w}_i^{(q)} = [w_i^{(q)}(0), w_i^{(q)}(1), \dots, w_i^{(q)}(K_s-1)]^T \tag{4.6}$$

whose elements are zero mean complex AWGNs with the variance σ_n^2 and are both spatially and temporally independent from each other. We then group the transmitted and received signals into new variables by defining

$$\mathbf{x}_i(k) = [x_i^{(1)}(k), x_i^{(2)}(k), \dots, x_i^{(N_t)}(k)]^T, \quad k = 0, \dots, K-1 \tag{4.7a}$$

$$\mathbf{y}_{i,P}(l) = [y_{i,P}^{(1)}(l), y_{i,P}^{(2)}(l), \dots, y_{i,P}^{(N_r)}(l)]^T, \quad l = 0, \dots, P-1 \tag{4.7b}$$

$$\mathbf{y}_i(k) = [y_i^{(1)}(k), y_i^{(2)}(k), \dots, y_i^{(N_r)}(k)]^T \tag{4.7c}$$

$$\mathbf{h}^{(p)}(l) = [h_{1p}(l), h_{2p}(l), \dots, h_{N_r p}(l)]^T, \quad l = 0, \dots, P \tag{4.7d}$$

$$\mathbf{H}(l) = [\mathbf{h}^{(1)}(l), \mathbf{h}^{(2)}(l), \dots, \mathbf{h}^{(N_t)}(l)] \tag{4.7e}$$

$$\mathbf{w}_i(k) = [w_i^{(1)}(k), w_i^{(2)}(k), \dots, w_i^{(N_r)}(k)]^T, \quad k = 0, \dots, K_s-1 \tag{4.7f}$$

$$\mathbf{y}_i = [\mathbf{y}_i^T(0), \mathbf{y}_i^T(1), \dots, \mathbf{y}_i^T(K-1)]^T \tag{4.7g}$$

$$\mathbf{y}_{i,P} = [\mathbf{y}_{i,P}^T(0), \mathbf{y}_{i,P}^T(1), \dots, \mathbf{y}_{i,P}^T(P-1)]^T \tag{4.7h}$$

$$\mathbf{x}_i = [\mathbf{x}_i^T(0), \mathbf{x}_i^T(1), \dots, \mathbf{x}_i^T(K-1)]^T \tag{4.7i}$$

$$\mathbf{x}_{i,P} = [\mathbf{x}_i^T(K-P-1), \mathbf{x}_i^T(K-P), \ldots, \mathbf{x}_i^T(K-1)]^T \tag{4.7j}$$

$$\mathbf{u}_i = [\mathbf{x}_{i,P}^T, \mathbf{x}_i^T]^T \tag{4.7k}$$

$$\mathbf{v}_i = [\mathbf{y}_{i,P}^T, \mathbf{y}_i^T]^T \tag{4.7l}$$

$$\mathbf{w}_i = [\mathbf{w}_i^T(0), \mathbf{w}_i^T(1), \ldots, \mathbf{w}_i^T(K_s-1)] \tag{4.7m}$$

$$\mathbf{H} = [\mathbf{H}^T(0), \mathbf{H}^T(1), \ldots, \mathbf{H}^T(P)]^T. \tag{4.7n}$$

The signal blocks from all the N_r receivers, after proper entry permutation, can be re-expressed as

$$\mathbf{v}_i = \boldsymbol{\mathcal{H}}(\mathbf{H}) \begin{bmatrix} \mathbf{x}_{i-1,P} \\ \mathbf{u}_i \end{bmatrix} + \mathbf{w}_i = \boldsymbol{\mathcal{H}}(\mathbf{H})\boldsymbol{\mathcal{T}}_{\text{cp}} \begin{bmatrix} \mathbf{x}_{i-1,L} \\ \mathbf{x}_i \end{bmatrix} + \mathbf{w}_i \tag{4.8}$$

where $\boldsymbol{\mathcal{T}}_{\text{cp}}$ is the corresponding $N_t(K_s+P) \times N_tK_s$ matrix whose specific form is omitted for simplicity. Then, the SS method could be applied to (4.8) and the identifiability could be guaranteed if

1. The $N_rK_s \times N_tK_s$ matrix $\boldsymbol{\mathcal{H}}(\mathbf{H})\boldsymbol{\mathcal{T}}_{\text{cp}}$ is tall.
2. Matrix $\boldsymbol{\mathcal{H}}(\mathbf{H})\boldsymbol{\mathcal{T}}_{\text{cp}}$ is full rank.
3. Span of $\boldsymbol{\mathcal{H}}(\bar{\mathbf{H}})\boldsymbol{\mathcal{T}}_{\text{cp}}$ equals to span of $\boldsymbol{\mathcal{H}}(\mathbf{H})\boldsymbol{\mathcal{T}}_{\text{cp}}$ if and only if $\bar{\mathbf{H}} = \mathbf{HB}$, where $\mathbf{B}$ is an unknown constant matrix.

The first condition is satisfied only if $N_r > N_t$. Clearly, the direct modeling is not applicable to the scenarios with $N_r = N_t$, which includes both the popular SISO OFDM in IEEE 802.11a [7] and the 2×2 MIMO OFDM in IEEE 802.11n [9]. The second and the third conditions have not been studied yet in the existing literatures to the best of our knowledge. One obvious example that breaks condition 2, 3 is when $\mathbf{H}(P) = \ldots = \mathbf{H}(1) = \mathbf{0}$ but $\mathbf{H}(0)$ is full column rank. In this case, the matrix $\boldsymbol{\mathcal{H}}(\mathbf{H})\boldsymbol{\mathcal{T}}_{\text{cp}}$ becomes singular.

4.3 Proposed Algorithm and the Related Issues

4.3.1 System Re-Modulation

We find that, by properly remodulating the received signal block, the system model (4.8) could be converted to a similar model of ZPSOS proposed in [55]. Let us first divide the noise vector $\mathbf{w}_i$ into two components as $\mathbf{w}_{i1} = \mathbf{w}_i(1 : N_rP)$, and $\mathbf{w}_{i2} = \mathbf{w}_i(N_rP + 1 : N_rK_s)$. Construct a new vector $\acute{\mathbf{v}}_i = [\mathbf{y}_{i-1}^T, \mathbf{y}_{i,P}^T]^T$, which could be expressed as

$$\acute{\mathbf{v}}_i = \boldsymbol{\mathcal{H}}(\mathbf{H}) \begin{bmatrix} \mathbf{u}_{i-1} \\ \mathbf{x}_{i,P} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_{(i-1)2} \\ \mathbf{w}_{i1} \end{bmatrix}. \tag{4.9}$$

It can be verified that

$$\begin{aligned} \mathbf{z}_i &\triangleq \mathbf{v}_i - \acute{\mathbf{v}}_i \\ &= \boldsymbol{\mathcal{H}}(\mathbf{H}) \left(\begin{bmatrix} \mathbf{x}_{i-1,L} \\ \mathbf{u}_i \end{bmatrix} - \begin{bmatrix} \mathbf{u}_{i-1} \\ \mathbf{x}_{i,P} \end{bmatrix} \right) + \underbrace{\left(\mathbf{w}_i - \begin{bmatrix} \mathbf{w}_{(i-1)2} \\ \mathbf{w}_{i1} \end{bmatrix} \right)}_{\boldsymbol{\varsigma}_i} \\ &= \boldsymbol{\mathcal{H}}(\mathbf{H}) \begin{bmatrix} \mathbf{0}_{N_rP \times 1} \\ \mathbf{d}_i \\ \mathbf{0}_{N_rP \times 1} \end{bmatrix} + \boldsymbol{\varsigma}_i = \boldsymbol{\mathcal{G}} \mathbf{d}_i + \boldsymbol{\varsigma}_i \end{aligned} \tag{4.10}$$

where

$$\begin{aligned} \mathbf{d}_i &= \mathbf{u}_i(1 : N_tK) - \mathbf{x}_{i-1} \\ &= [\mathbf{x}_{i,P}^T, \mathbf{x}_i^T(0), \ldots, \mathbf{x}_i^T(K - P - 1)]^T - \mathbf{x}_{i-1} \end{aligned} \tag{4.11}$$

$$\boldsymbol{\mathcal{G}} = \begin{bmatrix} \mathbf{H}(0) & \ldots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{H}(P) & \ddots & \mathbf{H}(0) \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathbf{H}(P) \end{bmatrix}. \tag{4.12}$$

The newly defined noise vector $\boldsymbol{\varsigma}_i$ is colored and has the covariance matrix

$$\mathbf{R}_{\boldsymbol{\varsigma}} = \mathrm{E}\{\boldsymbol{\varsigma}_i \boldsymbol{\varsigma}_i^H\} = \sigma_n^2 \mathbf{R}_w, \tag{4.13}$$

with

$$\mathbf{R}_w = \begin{bmatrix} 2\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & -\mathbf{I}_{N_rP\times N_rP} \\ \mathbf{0} & 2\mathbf{I}_{N_r(K-P)\times N_r(K-P)} & \mathbf{0} \\ -\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & 2\mathbf{I}_{N_rP\times N_rP} \end{bmatrix}. \tag{4.14}$$

Although it appears that the noise power in $\boldsymbol{\varsigma}_i$ is increased by a factor of 2, the signal power in $\mathbf{d}_i$ is enlarged twice as well. Therefore, the effective SNR is not changed. Since $\boldsymbol{\mathcal{G}}$ is exactly the same as the channel matrix in ZPSOS [55], we get the following lemma:

Lemma 4.1 [55]: For $N_r \geq N_t$, if there exists an $l \in [0, L]$ such that $\mathbf{H}(l)$ is of full column rank, then $\boldsymbol{\mathcal{G}}$ is of full column rank.

The proof is obvious and is omitted for brevity. The full column rank property of $\mathbf{H}(l)$ is almost surely guaranteed because signal propagation from each of the K transmitters is most likely independent. In the following, we assume that this condition holds. Even if $\mathbf{H}(l)$ is not full column rank, it is still possible that $\boldsymbol{\mathcal{G}}$ is of full column rank, bearing in mind that Lemma 4.1 only provides a sufficient condition.

4.3.2 SS Algorithm

The standard SS method requires the covariance of the noise vector to be a scaled identity matrix. Therefore, we need to whiten the vector $\mathbf{z}_i$ by $\mathbf{R}_w^{-1/2}$ and obtain

$$\acute{\mathbf{z}}_i = \mathbf{R}_w^{-1/2}\mathbf{z}_i = \underbrace{\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{G}}}_{\boldsymbol{\mathcal{A}}}\mathbf{d}_i + \breve{\mathbf{w}}_i \tag{4.15}$$

where $\breve{\mathbf{w}}_i$ is the $N_rK_s \times 1$ white noise vector whose entries have variance σ_n^2. In addition, since $\mathbf{R}_w$ is a non-singular matrix, the new channel matrix $\boldsymbol{\mathcal{A}}$ is full column rank if $N_r \geq N_t$.

Due to the special structure of $\mathbf{R}_w$, $\mathbf{R}_w^{-1/2}$ can be calculated as

$$\mathbf{R}_w^{-1/2} = \begin{bmatrix} c_1\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & c_2\mathbf{I}_{N_rP\times N_rP} \\ \mathbf{0} & \frac{1}{\sqrt{2}}\mathbf{I}_{N_r(K-P)\times N_r(K-P)} & \mathbf{0} \\ c_2\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & c_1\mathbf{I}_{N_rP\times N_rP} \end{bmatrix} \tag{4.16}$$

where

$$c_1 = \sqrt{\frac{2/3 + \sqrt{1/3}}{2}}, \quad c_2 = \sqrt{\frac{2/3 - \sqrt{1/3}}{2}}$$

regardless of the N_t, K, P.

Proof: Noting that the square root of $\mathbf{R}_w$ is not unique, we may only focus on $\mathbf{R}_w^{1/2}$ that has the special structure as

$$\mathbf{R}_w^{1/2} = \begin{bmatrix} b_1\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & b_2\mathbf{I}_{N_rP\times N_rP} \\ \mathbf{0} & b_3\mathbf{I}_{N_r(K-P)\times N_r(K-P)} & \mathbf{0} \\ b_2\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & b_3\mathbf{I}_{N_rP\times N_rP} \end{bmatrix} \tag{4.17}$$

where b_i's are real scalars. From $\mathbf{R}_w = \mathbf{R}_w^{1/2}\mathbf{R}_w^{1/2}$, we obtained

$$b_1^2 + b_2^2 = 2, \quad 2b_1b_2 = -1, \quad b_3^2 = 2$$

or equivalently

$$b_1 = \frac{1+\sqrt{3}}{3}, \quad b_2 = \frac{1-\sqrt{3}}{3}, \quad b_3 = \sqrt{2}.$$

Similarly, by letting

$$\mathbf{R}_w^{-1/2} = \begin{bmatrix} c_1\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & c_2\mathbf{I}_{N_rP\times N_rP} \\ \mathbf{0} & c_3\mathbf{I}_{N_r(K-P)\times N_r(K-P)} & \mathbf{0} \\ c_2\mathbf{I}_{N_rP\times N_rP} & \mathbf{0} & c_1\mathbf{I}_{N_rP\times N_rP} \end{bmatrix} \tag{4.18}$$

and from $\mathbf{R}_w^{1/2}\mathbf{R}_w^{-1/2} = \mathbf{I}$, we obtain

$$c_1 = \sqrt{\frac{2/3 + \sqrt{1/3}}{2}}, \quad c_2 = \sqrt{\frac{2/3 - \sqrt{1/3}}{2}}, \quad c_3 = 1/\sqrt{2}.$$

Proof completed. ■

The covariance matrix of $\acute{\mathbf{z}}_i$ is derived from

$$\mathbf{R} = \mathrm{E}\{\acute{\mathbf{z}}_i\acute{\mathbf{z}}_i^H\} = \boldsymbol{\mathcal{A}}\mathbf{R}_d\boldsymbol{\mathcal{A}}^H + \sigma_n^2\mathbf{I}_{N_rK_s\times N_rK_s} \tag{4.19}$$

where $\mathbf{R}_d = \mathrm{E}\{\mathbf{d}_i\mathbf{d}_i^H\}$ is the source covariance matrix, which should be full rank if no two elements in $\mathbf{d}_i$ are fully correlated. The covariance matrix $\mathbf{R}$ can be eigen-decomposed as

$$\mathbf{R} = \mathbf{U}_s\boldsymbol{\Delta}_s\mathbf{U}_s^H + \sigma_n^2\mathbf{U}_o\mathbf{U}_o^H \tag{4.20}$$

where $\mathbf{\Delta}_s$ is the $N_tK \times N_tK$ diagonal matrix and the $N_rK_s \times N_tK$ matrix $\mathbf{U}_s$ spans the signal-subspace of $\mathbf{R}$. In turn, the $N_rK_s \times (N_rK_s - N_tK)$ matrix $\mathbf{U}_o$ spans the noise-subspace of $\mathbf{R}$. The standard SS method says that the matrix $\mathbf{U}_o$ is orthogonal to every column of $\boldsymbol{\mathcal{A}}$. This can be equivalently expressed as

$$\mathbf{U}_o^H \mathbf{R}_w^{-1/2} \mathbf{C}_k \mathbf{H} = \mathbf{0}, \quad k = 1, \ldots, K \tag{4.21}$$

where $\mathbf{C}_k$ is the $N_rK_s \times N_r(P+1)$ Toeplitz matrix with the first column $\mathbf{e}_{(n-1)N_r+1}$ and $\mathbf{e}_p$ is defined as the pth column of $\mathbf{I}_{N_rK_s}$. In addition, the first row of $\mathbf{C}_k$ is $[1, \mathbf{0}_{1\times(N_r(P+1)-1)}]$ for $k = 1$ and is $\mathbf{0}_{1\times N_r(P+1)}$ for $k \geq 2$.

Define

$$\boldsymbol{\mathcal{K}} = [\mathbf{C}_1^H \mathbf{R}_w^{-1/2} \mathbf{U}_o, \mathbf{C}_2^H \mathbf{R}_w^{-1/2} \mathbf{U}_o, \ldots, \mathbf{C}_K^H \mathbf{R}_w^{-1/2} \mathbf{U}_o]. \tag{4.22}$$

The channel matrix $\mathbf{H}$ could be estimated from

$$\boldsymbol{\mathcal{K}}^H \mathbf{H} = \mathbf{0}. \tag{4.23}$$

Therefore, the estimate of $\mathbf{H}$, denoted as $\hat{\mathbf{H}}$, is a basis matrix of the orthogonal complement space of $\boldsymbol{\mathcal{K}}$. We will show later that the dimension of the orthogonal complement space of $\boldsymbol{\mathcal{K}}$ is exactly N_t. Therefore, $\hat{\mathbf{H}}$ can be obtained from left singular vectors of $\boldsymbol{\mathcal{K}}$ and is away from the true $\mathbf{H}$ by an unknown matrix $\mathbf{B}$, namely

$$\hat{\mathbf{H}} = \mathbf{H}\mathbf{B}. \tag{4.24}$$

This matrix ambiguity could be easily resolved by transmitting training symbols as suggested in [55]. Note that the amount of the training needed to resolve the ambiguity is much smaller than that required by a direct training based channel estimation.

4.3.3 Channel Identifiability and Order Over-Estimation

Thanks to the proposed re-modulation, the channel matrix $\boldsymbol{\mathcal{A}}$ possesses the similar structure as that in [55], which greatly facilities the study of the identifiability issue.

Theorem 4.1: If $\mathbf{H}(0)$ is full column rank, then the matrix $\mathbf{H}$ is uniquely determined by span($\boldsymbol{\mathcal{A}}$) subject to a common $N_t \times N_t$ non-singular matrix ambiguity on each $\mathbf{H}(l)$.

Proof: Let $\bar{\boldsymbol{\mathcal{A}}} = \mathbf{R}_w^{-1/2}\bar{\boldsymbol{\mathcal{G}}}$, where $\boldsymbol{\mathcal{G}}$ is constructed from $\bar{\mathbf{H}}$ in a similar way as $\boldsymbol{\mathcal{G}}$ is constructed from $\mathbf{H}$. If span($\bar{\boldsymbol{\mathcal{A}}}$) =span($\boldsymbol{\mathcal{A}}$), then

$$\bar{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{P}} \tag{4.25}$$

where $\boldsymbol{\mathcal{P}}$ is a matrix with dimension $N_tK \times N_tK$. Since $\mathbf{R}_w^{-1/2}$ is a full rank matrix, it is not difficult to derive the following equality:

$$\bar{\boldsymbol{\mathcal{G}}} = \boldsymbol{\mathcal{G}}\boldsymbol{\mathcal{P}}. \tag{4.26}$$

Therefore, span($\bar{\boldsymbol{\mathcal{G}}}$) =span($\boldsymbol{\mathcal{G}}$). Following the same procedure in [55], we could obtain that

$$\bar{\mathbf{H}}(l) = \mathbf{H}(l)\mathbf{B}. \tag{4.27}$$

One good aspect of Theorem 4.1 is that $\mathbf{H}(0)$ is almost always guaranteed full rank as previously explained. From Theorem 4.1, we know that the dimension of the orthogonal complement space of $\boldsymbol{\mathcal{K}}$ must be N_t. It is also seen that an order over-estimation on each L_{pq} does not affect the channel identifiability of $\mathbf{H}$ because the estimate $\bar{\mathbf{H}}(l) = \mathbf{H}(l)\mathbf{B} = \mathbf{0}$ for $l = L_{j,k} + 1, \ldots, L$ is also correct. Therefore, the two restrictions on the SS method for general MIMO systems [80], that is, the requirement of exact channel order and the channel estimation identifiability, are simultaneously lifted in the re-modulated CP-based MIMO OFDM system. ■

4.3.4 Comparison with ZPSOS

Similarities

Similarities between these two methods mainly reside in the choice of system parameters and the model structures. For example, under the same transmission rate, namely, the same block length and the CP length, the channel matrix $\boldsymbol{\mathcal{G}}$ is exactly the same for both methods. The effective SNR, as discussed before,

is also the same. Similar channel estimation accuracy for both CPSOS and ZPSOS is observed in the later simulation. Moreover, problems like channel order over-estimation and the identifiability are lifted for both CPSOS and ZPSOS.

Differences

Despite many similarities, there do exist other differences that show the advantages of CPSOS over ZPSOS.

1. *Symbol Detection.* In ZP based OFDM, one needs to add the last P entries of $\mathbf{v}_i^{(p)}$ to its first P entries to remove the ICI effect. Then, similar relationship as in (1.15) could be derived for ZPSOS. Note that, $\boldsymbol{\eta}_i(k)$ in ZPSOS, is not independent for different k, although its entries $\eta_i^{(q)}(k)$ are independent with respect to q. Therefore, the ML detection requires the co-consideration of $\mathbf{r}_i(k)$ over all carriers. This results in an exponential increment in the detection complexity, which betrays the original purpose of adopting the MIMO OFDM systems. Although the low complexity Zero Forcing (ZF) detection is suggested in [55], it is well known that this linear detection suffers from considerable performance loss.

   We here suggest a suboptimal way that the detection still considers each subcarrier independently regardless of whether the noise is dependent across the carriers or not. It can be proved that the covariance matrix of $\boldsymbol{\eta}_i(k)$ in ZPSOS is

   $$\mathrm{E}\{\boldsymbol{\eta}_i(k)\boldsymbol{\eta}_i^H(k)\} = \left(1 + \frac{P}{K}\right)\sigma_n^2\mathbf{I}_{N_r\times N_r}. \tag{4.28}$$

   Therefore, the noise power, compared to CPSOS is increased by a factor of $(1+P/K)$, and the SNR loss is around $10\log(1+P/K)$ dB. In many standards, e.g., IEEE 802.11a, IEEE 802.11n, $K = 4P$ is adopted and the SNR loss is around 1 dB.

2. *Compatibility.* Obviously, CP-based OFDM has a much wider application than ZP-based OFDM. For example, CP-based OFDM has been well adopted into

European DAB, DVB, HIPERLAN, IEEE 802.11a WLAN standards and the coming IEEE 802.11n WLAN standards. However, to the best of the our knowledge, ZP based-OFDM has very limited applications.

4.4 Asymptotical Performance Analysis

4.4.1 Channel Estimation Mean Square Error

We provide a first-order performance analysis on the proposed estimator at high SNR similar to that in [44], [45].

Theorem 4.2: Assume that both noise and signals are zero-mean i.i.d. with variances σ_n^2 and σ_s^2, respectively, the mean and the covariance of the channel estimation error are approximated by

$$\mathrm{E}\{\Delta\mathbf{H}\} = \mathbf{0} \tag{4.29}$$

$$\mathrm{E}\{\mathrm{vec}(\Delta\mathbf{H})\mathrm{vec}^H(\Delta\mathbf{H})\} = \mathbf{I}_{N_t} \otimes \left(\frac{\sigma_n^2(\boldsymbol{\mathcal{K}}^H)^\dagger\boldsymbol{\mathcal{K}}^\dagger}{2M\sigma_s^2}\right). \tag{4.30}$$

Specifically, the error covariance matrix for pth transmit antennas is

$$\mathrm{E}\{\Delta\mathbf{H}(:,p)\Delta\mathbf{H}^H(:,p)\} = \frac{\sigma_n^2(\boldsymbol{\mathcal{K}}^H)^\dagger\boldsymbol{\mathcal{K}}^\dagger}{2M\sigma_s^2} \tag{4.31}$$

and the channel estimation MSE is

$$\mathrm{E}\{\|\Delta\mathbf{H}(:,p)\|^2\} = \frac{\sigma_n^2\|\boldsymbol{\mathcal{K}}^\dagger\|_F^2}{2M\sigma_s^2}. \tag{4.32}$$

See proof in Appendix B. Several insightful observations can be drawn from (4.31), for example, the MSE is proportional to the noise power but is inversely proportional to both the signal power and the number of the received signal block.

4.4.2 Deterministic Cramér-Rao-Bound

We consider the deterministic CRB [82] for CPSOS, where the observations are $\mathbf{z} = [\mathbf{z}_0^T, \ldots, \mathbf{z}_{M-1}^T]^T$, and the parameters to be estimated are $\boldsymbol{\vartheta} =$

$[\text{vec}(\mathbf{H}), \mathbf{d}_1, \ldots, \mathbf{d}_M, \sigma_n^2]$. In order to calculate CRB, we need the joint PDF of $\mathbf{z}$, denoted as $p(\mathbf{z}|\boldsymbol{\vartheta})$. Since $\boldsymbol{\varsigma}_i$ is correlated with both $\boldsymbol{\varsigma}_{i-1}$ and $\boldsymbol{\varsigma}_{i+1}$, the covariance matrix of $\mathbf{z}$, denoted as $\mathbf{R}_z$, is an $MN_rK_s \times MN_rK_s$ Toeplitz matrix with the main diagonal elements 2, the (N_rK+1)th, $-(N_rK+1)$th[1] diagonal elements -1, and all other elements 0. We note that the inverse of such a huge Toeplitz matrix is mathematically prohibitive.

To simplify the derivation and gain more insight into the proposed algorithm, we approximate $\mathbf{R}_z$ by

$$\mathbf{R}_z = \mathbf{I}_M \otimes \mathbf{R}_w \tag{4.33}$$

which equivalently says that we ignore the correlations among different $\boldsymbol{\varsigma}_i$'s. This approximation is also justified since the performance of SS method is only related to the auto-covariance of $\boldsymbol{\varsigma}_i$ and does not dependent on whether $\boldsymbol{\varsigma}_i$ are cross-correlated or not. The so derived CRB will be called as approximated CRB (ACRB). Since we relax the noise condition, ACRB should be greater than or equal to the CRB.

Define

$$\boldsymbol{\mathcal{D}}_i = [\mathbf{D}_i^{(1)}, \mathbf{D}_i^{(2)}, \ldots, \mathbf{D}_i^{(N_t)}] \tag{4.34}$$

where

$$\mathbf{D}_i^{(p)} = \sum_{k=1}^{K} \mathbf{C}_k d_i((k-1)N_t + p). \tag{4.35}$$

The ACRB is obtained as

$$\text{ACRB}_{\text{vec}(\mathbf{H})} = \sigma_n^2 \left(\sum_{i=0}^{M-1} \boldsymbol{\mathcal{D}}_i^H \mathbf{R}_w^{-1/2} \mathbf{P}_{\boldsymbol{\mathcal{A}}}^{\perp} \mathbf{R}_w^{-1/2} \boldsymbol{\mathcal{D}}_i \right)^{\dagger} \tag{4.36}$$

where $\mathbf{P}_{\boldsymbol{\mathcal{A}}}^{\perp}$ is the projection matrix onto the orthogonal complement space spanned by $\boldsymbol{\mathcal{A}}^{\perp}$. If signals are i.i.d. with variance σ_s^2, the asymptotical ACRB for large M is obtained as

$$\text{ACRB}_{\text{vec}(\mathbf{H})} = \mathbf{I}_{N_t} \otimes \left(\frac{\sigma_n^2 (\boldsymbol{\mathcal{K}}^H)^{\dagger} \boldsymbol{\mathcal{K}}^{\dagger}}{2M\sigma_s^2} \right) \tag{4.37}$$

and the asymptotical ACRB for each column of $\mathbf{H}$ can be, separately, obtained as

$$\text{ACRB}_{\text{vec}(\mathbf{H}(:,p))} = \frac{\sigma_n^2 (\boldsymbol{\mathcal{K}}^H)^{\dagger} \boldsymbol{\mathcal{K}}^{\dagger}}{2M\sigma_s^2}. \tag{4.38}$$

[1]Notations follow those in MATLAB.

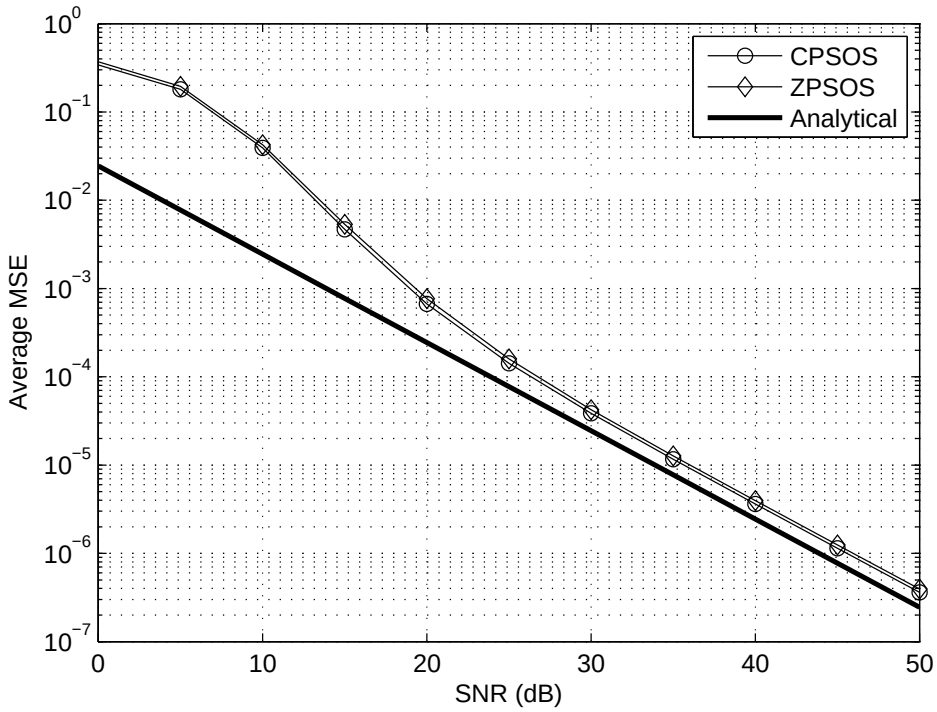

Figure 4.1: Channel estimation MSEs versus SNR with 200 received blocks.

See proof in Appendix C.

Interestingly, the asymptotical ACRB coincides with the asymptotical error covariance matrix (4.31). We may draw the conclusion that

1. The channel estimation MSE is greater than or equal to the CRB, which agrees with the intuition very well.

2. The SS method asymptotically achieves the optimal estimation (same performance as ML estimation), regardless of whether the noise is colored or white.

4.5 Simulations

In this section, we examine the performance of CPSOS for a 2×2 MIMO OFDM systems under various scenarios. The OFDM block length is taken as $K = 32$, and

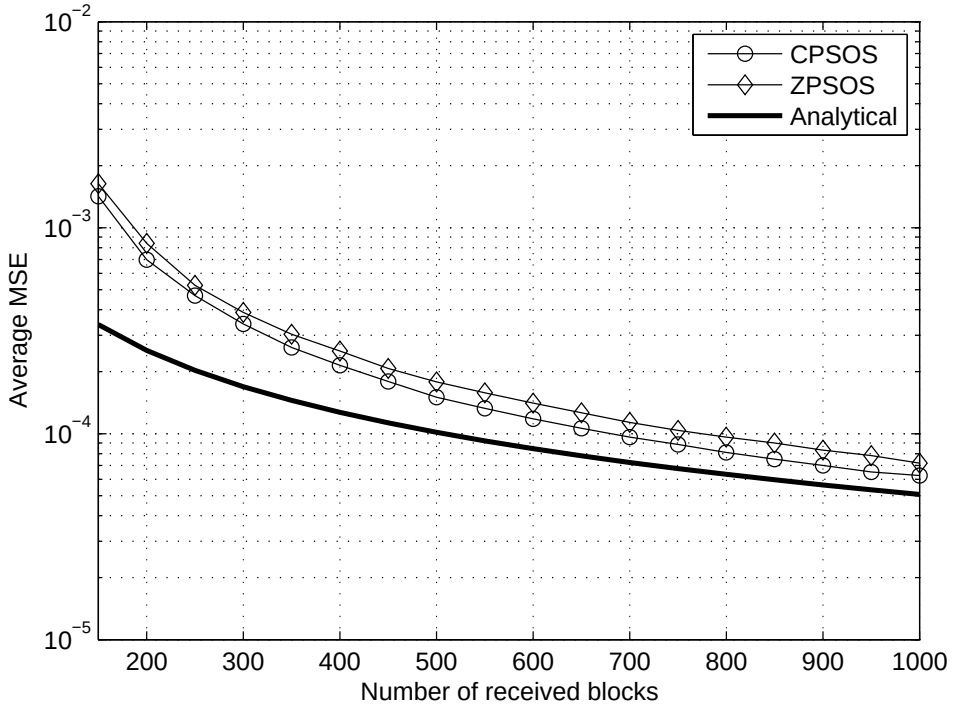

Figure 4.2: Channel estimation MSEs versus number of OFDM blocks for SNR= 20dB.

the CP length is taken as $P = 8$. The signal symbols are extracted from QPSK constellations. The 6-ray channel model with an exponential power delay profile

$$\mathrm{E}\{|h_{pq}(l)|^2\} = \rho \exp(-l/5), \quad l = 0, \ldots, 5 \tag{4.39}$$

is used where ρ is the coefficient to normalize the overall channel gain to $\|\mathbf{h}_{pq}\|^2 = 1$. The estimation MSE is defined as

$$\mathrm{MSE} = \frac{1}{N_t N_r} \|\hat{\mathbf{H}}\mathbf{B}^{-1} - \mathbf{H}\|_F^2 \tag{4.40}$$

where, for simulation purpose, the ambiguity matrix $\mathbf{B}$ is obtained according to [83]:

$$\mathbf{B} = \arg\min_{\mathbf{T}} \|\hat{\mathbf{H}}\mathbf{T}^{-1} - \mathbf{H}\|_F^2. \tag{4.41}$$

The number of the Monte-Carlo runs used for average is taken as 500.

We first fix the number of the OFDM blocks as 200 and compare two different blind channel estimators: CPSOS and ZPSOS. Note that 200 blocks is a common

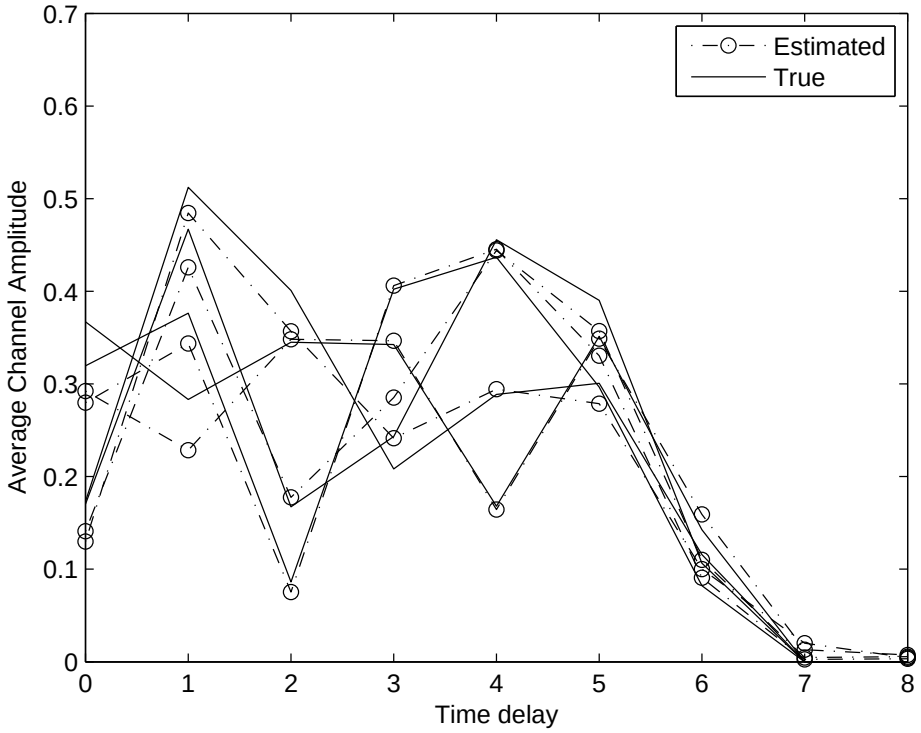

Figure 4.3: Amplitude estimation of channel taps at SNR= 12dB.

number for applying the SS algorithm. The channel estimation MSEs versus SNR for these two algorithms are shown in Fig. 4.1. The analytical performance derived from either the asymptotical MSE or ACRB is displayed as well. It is seen that ZPSOS and CPSOS give comparable performance over all SNR considered. Both methods are close to the analytical MSE after 20 dB. The gap between the analytical result and the simulation result is due to the usage of the asymptotical MSE here. One may expect a little bit better performance from CPSOS because the noise is colored with a known covariance matrix (implicating less independency).

Fig. 4.2 shows the performance of MSEs versus the number of the OFDM blocks for the two algorithms at SNR= 20 dB. As explained previously, the performance of CPSOS is a little bit better than ZPSOS and it may achieve the analytical asymptotical MSE when the number of the blocks becomes large.

Fig. 4.3 and Fig. 4.4 show the amplitude of the channel tap detection of four

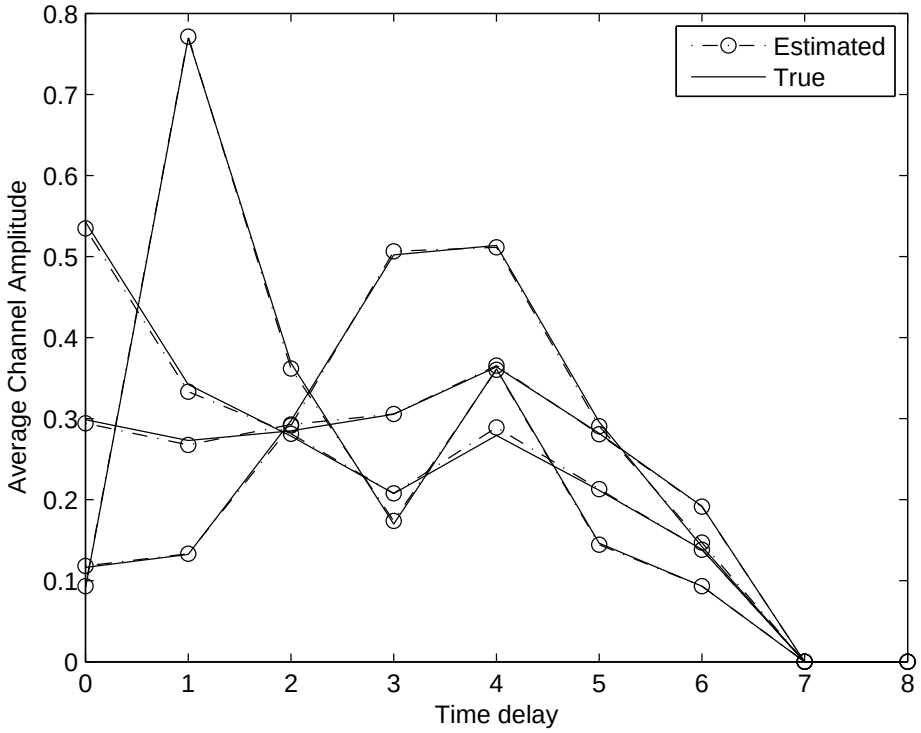

Figure 4.4: Amplitude estimation of channel taps at SNR= 20 dB.

random channel realizations at SNR= 12 dB and SNR= 20 dB, respectively. It is noted that SNR= 12 dB yields relatively good channel amplitude estimation, whereas SNR= 20 dB gives almost perfect estimation. Nevertheless, by noting that the channel amplitude at tap 7 and tap 8 are close to zero in both figures, we may also apply the channel order estimation method suggested in [55] under the proposed CPSOS.

To demonstrate the robustness of CPSOS to the channel order over-estimation, we assume the estimated channel order as $\hat{L} = 5, 6, 7$ respectively. The value $\hat{L} = 5$ corresponds to the correct channel order, and other values are those being over-estimated. The channel estimation MSEs versus SNR and block number for these three different orders are displayed in Fig. 4.5 and Fig. 4.6, respectively. As expected, order over-estimation only causes slight performance loss. This is reasonable that assuming more channel taps, even if those zero taps, may also

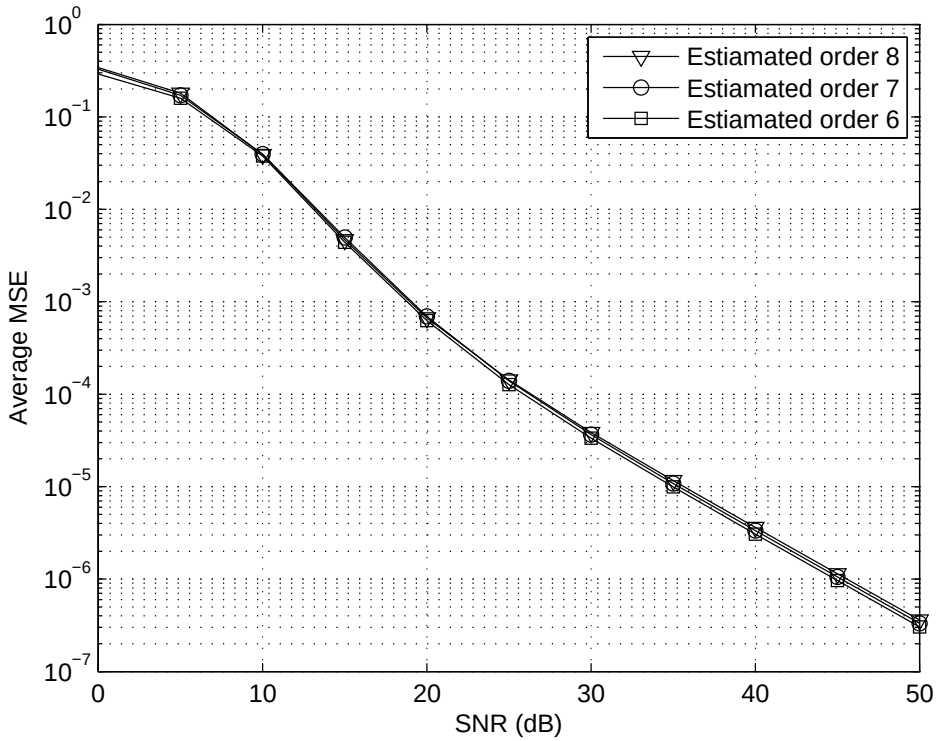

Figure 4.5: Channel estimation MSEs versus SNR for different estimated channel order.

contribute to the channel estimation error. Nevertheless, the largest loss, appearing when the order is taken the same as the CP length, is less than 1 dB.

Finally, we compare the bit error rate (BER) of the CPSOS and the ZPSOS when the number of OFDM blocks is taken as 200. The ML detection is adopted in CPSOS, while for ZPSOS, two different detections are used, e.g., sub-optimal detection and ZF detection [55]. Fig. 4.7 shows the BER performances versus SNR for the considered scenarios. Clearly, CPSOS outperforms ZPSOS within all the SNR region. The SNR gain (at high SNR) between CPSOS and ZPSOS with suboptimal detection meets our theoretical analysis, which is about 1 dB. Moreover, the gap between CPSOS and ZPSOS with ZF detection is considerably large.

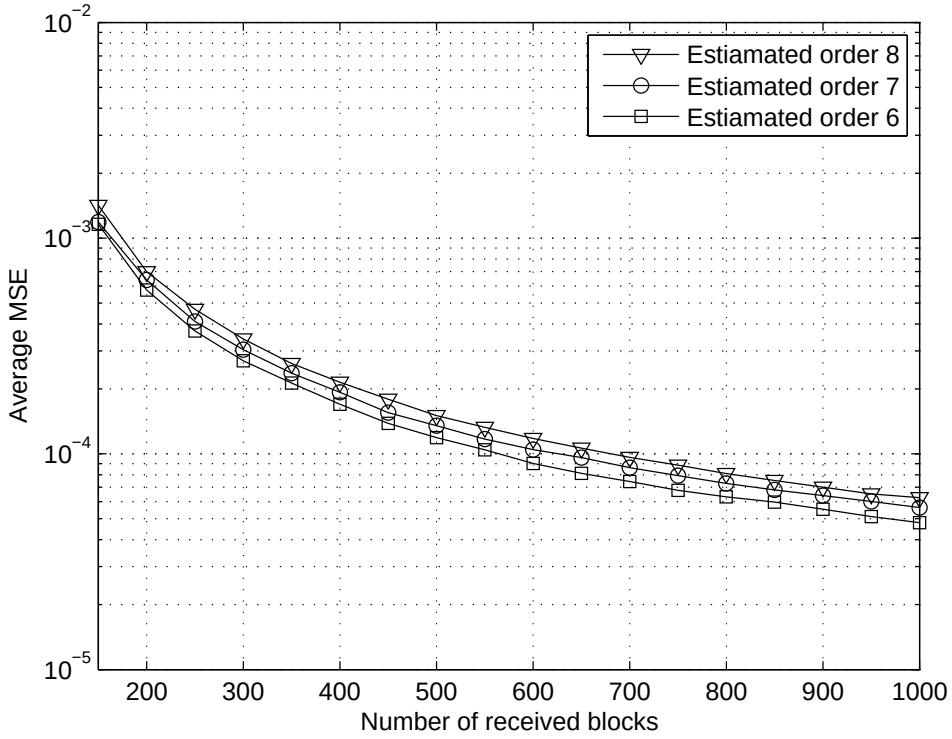

Figure 4.6: Channel estimation MSEs versus number of OFDM blocks for different estimated channel order.

4.6 Summery

In this chapter, we developed a new SS based blind channel estimation for MIMO OFDM systems. With an appropriate re-modulation on the received signals, an effective way has been found to apply the SS method for the CP based MIMO-OFDM system when the number of receive antennas is no less than the number of transmit antennas. Many issues related to the SS method have been studied for this newly proposed modulation, e.g., channel identifiability, order over-estimation, MSE of the channel estimation as well as the deterministic CRB on the channel estimation. Most importantly, since the proposed method allows blind channel estimation for the CP based MIMO OFDM, it is compatible with many existing standards and the coming 4G wireless communication standards.

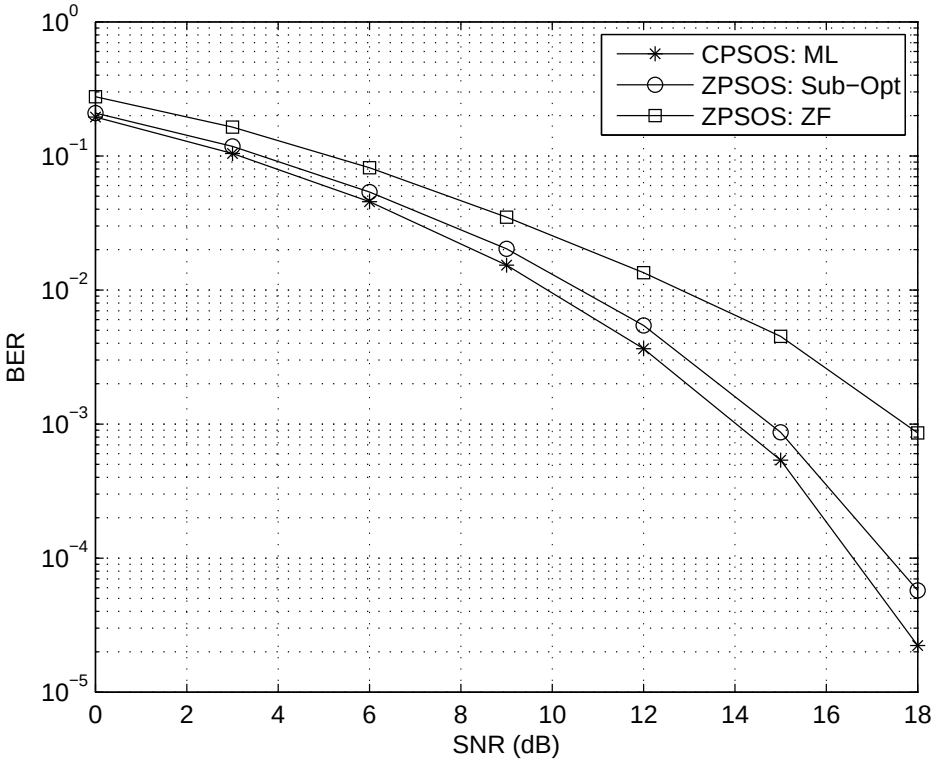

Figure 4.7: BERs versus SNR for CPSOS and ZPSOS.

Chapter 5

Non-Redundant Linear Precoding Based Blind Channel Estimation for MIMO OFDM Systems

In this chapter, we develop a simple blind channel estimation technique by utilizing non-redundant linear block precoding. The algorithm is executable as long as the number of the receive antennas is greater than a small portion of the number of the transmit antennas and it also shows other advantages over the traditional SS methods.

5.1 Introduction

From the previous chapters, we know that the traditional SS method cannot be directly applied if the number of antennas at the receiver is smaller than or equal to the number of antennas at the transmitter. Several solutions for SISO OFDM has been introduced in subsection 2.2.2 and the drawbacks of each of them were also listed. Although CPSOS developed in Chapter 4 is successfully applied to SISO OFDM and MIMO OFDM with $N_r \geq N_t$, it still remains unknown how to apply SOS blind channel estimation to MIMO OFDM with $N_r < N_t$. The problem is

tough and was not perfectly solved over the last two decades.

Precoded OFDM arose recently in [84], [85], where the information block is first precoded and is then sent to channels. The meaning of precoding is quite general, for example, if the precoding matrix is $\mathbf{T}_{\text{cp}}\mathbf{F}^H$, the resulting system is the so called CP-based OFDM. In addition, zero padding in [55] is also a kind of precoded OFDM. A redundant block precoding based channel estimation for SISO OFDM was proposed in [86] and a redundant precoding based space time coded OFDM was proposed in [48]. These techniques sacrifice at least one symbol in each block and are not spectrally efficient.

Based on the assumption that the transmitted symbols are independent and identically distributed to each other, a new type of blind channel estimation method for SISO-OFDM system has been reported in [87]-[91]. In [87] and [88], a non-redundant linear precoder is applied at the transmitter and the CSI is possessed in all entries of the signal covariance matrix. However, the authors purely extract the CSI from a predefined single column of the covariance matrix, which greatly limits the performance accuracy of the algorithm. In [89]-[91], cross correlation of two consecutive OFDM blocks are utilized. However, this approach requires twice the number of the received blocks to construct a similarly reliable correlation matrix as the one in [87], [88]. Since a major drawback of blind algorithms is the requirement for large number of the received signals, this cross correlation based method seems not as attractive as the one in [87], [88]. Most importantly, in all these works [87]-[91], how to apply the non-redundant linear precoding for the channel estimation in MIMO OFDM systems was not addressed. In fact, the generalization approach from SISO to MIMO is not a straightforward task.

In this chapter, we develop a new blind channel estimation method for both SISO- and MIMO-OFDM systems. A novel contribution of the proposed method lies in that it is even applicable for MISO systems. For SISO OFDM systems, the newly proposed method can improve the performance accuracy compared to the existing works. For MIMO-OFDM systems, we further propose an approach

to eliminate the multi-dimensional ambiguity that is known to exist for channel estimation under multi-transmitter scenarios. Finally, the stochastic CRB is derived and the simulation results are provided to show the effectiveness of our proposed algorithm.

5.2 System Model

We still consider the $N_t \times N_r$ MIMO OFDM system shown in Fig. 1.4 and the received signal model is written in (1.13). For convenience, we re-write it here as

$$\mathbf{y}_i^{(q)} = \sum_{p=1}^{N_t} \mathbf{H}_{pq}\mathbf{F}^H\mathbf{s}_i^{(p)} + \mathbf{n}_i^{(q)}, \quad i = 1, \ldots, M. \tag{5.1}$$

For most cases, the transmitted signals are independent both over the time and across the transmit antennas. That is, $\mathbf{s}_i^{(p_1)}$ is independent from $\mathbf{s}_i^{(p_2)}$ for any $p_1 \neq p_2$.

Define[1] new vectors

$$\mathbf{s}_i = \left[(\mathbf{s}_i^{(1)})^T, (\mathbf{s}_i^{(2)})^T, \ldots, (\mathbf{s}_i^{(N_t)})^T\right]^T \tag{5.2}$$

$$\mathbf{n}_i = \left[(\mathbf{n}_i^{(1)})^T, (\mathbf{n}_i^{(2)})^T, \ldots, (\mathbf{n}_i^{(N_r)})^T\right]^T \tag{5.3}$$

and let

$$\boldsymbol{\mathcal{H}}_{\mathbf{y}} = \begin{bmatrix} \mathbf{H}_{11} & \ldots & \mathbf{H}_{N_t 1} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1N_r} & \ldots & \mathbf{H}_{N_t N_r} \end{bmatrix} \tag{5.4}$$

denote the overall time domain channel matrix with its (q,p)th partitioned block given by the circulant matrix $\mathbf{H}_{pq}$. The overall received symbol vector $\mathbf{y}_i$ can be expressed as

$$\begin{aligned} \mathbf{y}_i &= \left[(\mathbf{y}_i^{(1)})^T, (\mathbf{y}_i^{(2)})^T, \ldots, (\mathbf{y}_i^{(N_r)})^T\right]^T \\ &= \boldsymbol{\mathcal{H}}_{\mathbf{y}}\boldsymbol{\mathcal{F}}^H\mathbf{s}_i + \mathbf{n}_i \end{aligned} \tag{5.5}$$

[1]The notations in this chapter will be the same as those defined in subsection 1.3 unless otherwise mentioned.

where $\boldsymbol{\mathcal{F}}$ is the block diagonal matrix with the form

$$\boldsymbol{\mathcal{F}} = \text{diag}\{\underbrace{\mathbf{F},\ \mathbf{F},\ldots,\ \mathbf{F}}_{N_t \text{ blocks}}\} = \mathbf{I}_{N_t} \otimes \mathbf{F}. \tag{5.6}$$

Let

$$\mathbf{h}_{pq} = [h_{pq}(0), h_{pq}(1), \ldots, h_{pq}(L)]^T$$

be the channel response from the transmitter p to the receiver q. The frequency domain response of $\mathbf{h}_{pq}$ is written as

$$\begin{aligned}\tilde{\mathbf{h}}_{pq} &= \text{DFT}(\mathbf{h}_{pq}) = \sqrt{K}\mathbf{F}(1:L+1)\mathbf{h}_{pq} \\ &= [\tilde{h}_{pq}(0), \tilde{h}_{pq}(1), \ldots, \tilde{h}_{pq}(K-1)]^T. \end{aligned} \tag{5.7}$$

The normalized DFT of $\mathbf{y}_i^{(q)}$ is written as

$$\begin{aligned}\mathbf{r}_i^{(q)} &= \mathbf{F}\mathbf{y}_i^{(q)} = \sum_{p=1}^{N_t} \mathbf{F}\mathbf{H}_{pq}\mathbf{F}^H\mathbf{s}_i^{(p)} + \mathbf{F}\mathbf{n}_i^{(q)} \\ &= \sum_{p=1}^{N_t} \boldsymbol{\Lambda}_{pq}\mathbf{s}_i^{(p)} + \tilde{\mathbf{n}}_i^{(q)} \end{aligned} \tag{5.8}$$

where $\boldsymbol{\Lambda}_{pq}$ is a diagonal matrix with the diagonal elements obtained from $\tilde{\mathbf{h}}_{pq}$. Since the noise is i.i.d. AWGN, the new noise vector $\tilde{\mathbf{n}}_i^{(q)}$ has the same statistical distribution as that of the vector $\mathbf{n}_i^{(q)}$.

Let

$$\boldsymbol{\mathcal{H}}_{\mathbf{r}} = \begin{bmatrix} \boldsymbol{\Lambda}_{11} & \ldots & \boldsymbol{\Lambda}_{N_t 1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Lambda}_{1N_r} & \ldots & \boldsymbol{\Lambda}_{N_t N_r} \end{bmatrix} \tag{5.9}$$

denote the overall frequency domain channel matrix. There is

$$\begin{aligned}\mathbf{r}_i &= \left[(\mathbf{r}_i^{(1)})^T, (\mathbf{r}_i^{(2)})^T, \ldots, (\mathbf{r}_i^{(N_r)})^T\right]^T \\ &= \boldsymbol{\mathcal{H}}_{\mathbf{r}}\mathbf{s}_i + \tilde{\mathbf{n}}_i \end{aligned} \tag{5.10}$$

where

$$\tilde{\mathbf{n}}_i = \boldsymbol{\mathcal{F}}\mathbf{n}_i \tag{5.11}$$

is the equivalent noise vector in the frequency domain. The source covariance matrix is, by assumption, expressed as

$$\mathbf{R_s} = \mathrm{E}\{\mathbf{s}_i \mathbf{s}_i^H\} = \sigma_s^2 \mathbf{I}_{KN_t} \tag{5.12}$$

where σ_s^2 is the transmitted signal power.

5.3 Blind Channel Estimation for SISO OFDM Systems

5.3.1 Generalized Precoding

For the SISO OFDM system, $\mathbf{s}_i$, $\boldsymbol{\mathcal{H}}_\mathbf{r}$ and $\tilde{\mathbf{n}}_i$ reduce to $\mathbf{s}_i^{(1)}$, $\boldsymbol{\Lambda}_{11}$ and $\tilde{\mathbf{n}}_i^{(1)}$, respectively. Suppose that the block symbol $\mathbf{s}_i^{(1)}$ is precoded by a predefined $K \times K$ matrix $\mathbf{W}_1$ before the IDFT operation. The received signal vector can be rewritten as

$$\mathbf{r}_i^{(1)} = \boldsymbol{\Lambda}_{11} \mathbf{W}_1 \mathbf{s}_i^{(1)} + \tilde{\mathbf{n}}_i^{(1)}. \tag{5.13}$$

As a result, the signal covariance matrix is

$$\begin{aligned}
\mathbf{R_r} &= \mathrm{E}\{\mathbf{r}_i^{(1)} (\mathbf{r}_i^{(1)})^H\} = \sigma_s^2 \boldsymbol{\Lambda}_{11} \mathbf{P}_1 \boldsymbol{\Lambda}_{11}^H + \sigma_n^2 \mathbf{I}_K \\
&= \begin{bmatrix} \sigma_s^2 [P_1]_{11} |\tilde{h}_{11}(0)|^2 + \sigma_n^2 & \dots & \sigma_s^2 [P_1]_{1K} \tilde{h}_{11}(0) \tilde{h}_{11}(K-1)^* \\ \vdots & \ddots & \vdots \\ \sigma_s^2 [P_1]_{K1} \tilde{h}_{11}(K-1) \tilde{h}_{11}(0)^* & \dots & \sigma_s^2 [P_1]_{KK} |\tilde{h}_{11}(K-1)|^2 + \sigma_n^2 \end{bmatrix} \\
&= \sigma_s^2 (\tilde{\mathbf{h}}_{11} \tilde{\mathbf{h}}_{11}^H) \odot \mathbf{P}_1 + \sigma_n^2 \mathbf{I}_{K \times K}
\end{aligned} \tag{5.14}$$

where

$$\mathbf{P}_1 = \mathbf{W}_1 \mathbf{W}_1^H \tag{5.15}$$

is the square of the precoding matrix.

Remark: As will be seen later, the precoder $\mathbf{W}_1$ in the proposed algorithm can be an arbitrary matrix. However, the precoding scheme in [87], [88] follows specific

formality. For example in [87], although the algorithm is not illustrated by using a precoding matrix, it in fact uses a precoding matrix with entries given by

$$[\mathbf{W}_1]_{ac} = \begin{cases} 1 & a = c \neq t \\ 2 & a = c = t \\ 1 & a \neq c,\ c = t \\ 0 & \text{otherwise} \end{cases} \tag{5.16}$$

where t is taken as $K/4$.

5.3.2 Blind Channel Estimation Algorithm

On the cth column of $\mathbf{R_r}$, except the diagonal element $[\mathbf{R_r}]_{cc}$, all the other $K-1$ elements are of the form

$$[\mathbf{R_r}]_{ac} = \sigma_s^2 [\mathbf{P}_1]_{ac} \tilde{h}_{11}(a-1)\tilde{h}_{11}(c-1)^*, \quad a \neq c. \tag{5.17}$$

Since $\mathbf{P}_1$ is known as *a prior*, $\sigma_s^2 \tilde{h}_{11}(a-1)\tilde{h}_{11}(c-1)^*$ can be obtained from

$$\sigma_s^2 \tilde{h}_{11}(a-1)\tilde{h}_{11}(c-1)^* = \frac{[\mathbf{R_r}]_{ac}}{[\mathbf{P}_1]_{ac}}, \quad a \neq c. \tag{5.18}$$

However, $\sigma_s^2 \tilde{h}_{11}(c-1)\tilde{h}_{11}(c-1)^*$ cannot be recovered since the (c,c)th entry of $\mathbf{R_r}$ is corrupted by the unknown noise variance.

As long as $K - 1 \geq L + 1$, as usually the case, both the $(L+1) \times 1$ time domain channel vector $\mathbf{h}_{11}$ and the $K \times 1$ frequency domain channel vector $\tilde{\mathbf{h}}_{11}$ can be obtained in the following way. Define

$$\begin{aligned} \breve{\mathbf{r}}_c &= \left[\frac{[\mathbf{R_r}]_{1c}}{[\mathbf{P}_1]_{1c}}, \ldots, \frac{[\mathbf{R_r}]_{(c-1)c}}{[\mathbf{P}_1]_{(c-1)c}}, \frac{[\mathbf{R_r}]_{(c+1)c}}{[\mathbf{P}_1]_{(c+1)c}}, \ldots, \frac{[\mathbf{R_r}]_{Kc}}{[\mathbf{P}_1]_{Kc}} \right]^T \\ &= [\sigma_s^2 \tilde{h}_{11}(0)\tilde{h}_{11}(c-1)^*, \ldots, \sigma_s^2 \tilde{h}_{11}(c-2)\tilde{h}_{11}(c-1)^*, \\ &\qquad \sigma_s^2 \tilde{h}_{11}(c)\tilde{h}_{11}(c-1)^*, \ldots, \sigma_s^2 \tilde{h}_{11}(K-1)\tilde{h}_{11}(c-1)^*]^T \end{aligned} \tag{5.19}$$

$$\mathbf{F}_c = [\mathbf{F}(1:c-1, 1:L+1)^T,\ \mathbf{F}(c+1:K, 1:L+1)^T]^T. \tag{5.20}$$

The estimate of $\mathbf{h}_{11}$ from $\breve{\mathbf{r}}_c$ can be derived as

$$\hat{\mathbf{h}}_{11}^c = \frac{1}{\sqrt{K}} \mathbf{F}_c^{\dagger} \breve{\mathbf{r}}_c = \sigma_s^2 \tilde{h}_{11}(c-1)^* \mathbf{h}_{11}, \quad c = 1, \ldots, K. \tag{5.21}$$

Consequently, the estimate of $\tilde{\mathbf{h}}_{11}$ is obtained as

$$\begin{aligned}\hat{\tilde{\mathbf{h}}}_{11}^{c} &= \mathrm{DFT}(\hat{\mathbf{h}}_{11}^{c}) = \sqrt{K}\mathbf{F}(1:L+1)\hat{\mathbf{h}}_{11}^{c} \\ &= \mathbf{F}(1:L+1)\mathbf{F}_{c}^{\dagger}\breve{\mathbf{r}}_{c} = \sigma_s^2\tilde{h}_{11}(c-1)^{*}\tilde{\mathbf{h}}_{11} \quad c=1,\ldots,K. \end{aligned} \tag{5.22}$$

In (5.21) and (5.22), the channel length is enforced in the time domain to obtain a truncated channel vector and then the K-point DFT is performed to get the channel frequency response. This process is known as denoising and could improve the estimation accuracy [50].

Remark: The estimated channel vectors from both (5.21) and (5.22) are different from the true one by a complex scalar, which is well known for almost all the blind channel estimation. Since the cth entry of $\hat{\tilde{\mathbf{h}}}_{11}^{c}$ is $\sigma_s^2|\tilde{h}_{11}(c-1)|^2$, the scaling ambiguity of the estimation can be eliminated if the transmitted signal power is known.

In [87], [88], the authors proposed a method that uses only one of $\hat{\tilde{\mathbf{h}}}_{11}^{c}$ as the final channel estimation, which greatly limits the performance of the algorithm. To obtain better performance, an averaging approach should be necessarily adopted. Since each $\hat{\tilde{\mathbf{h}}}_{11}^{c}$ is different from the true one by an unknown factor, it is not possible to apply the maximum ratio combining (MRC). Moreover the direct averaging may not be a good choice since some elements of the estimated channel vector may be canceled with each other, e.g., when $\tilde{h}_{11}(c_1-1)+\tilde{h}_{11}(c_2-1)=0$ for some $c_1, c_2 \in \{1,\ldots,K\}$. We now propose a new algorithm that jointly considers all entries of the matrix $\mathbf{R}_{\mathbf{r}}$.

The Algorithm:

- **Step 1**: From (5.22), obtain K estimates $\hat{\tilde{\mathbf{h}}}_{11}^{c}, c=1,\ldots,K$.
- **Step 2**: Form a new matrix

$$\mathbf{Q} = \left[\hat{\tilde{\mathbf{h}}}_{11}^{1}, \hat{\tilde{\mathbf{h}}}_{11}^{2}, \ldots, \hat{\tilde{\mathbf{h}}}_{11}^{K}\right]. \tag{5.23}$$

  It can be proved that

$$\mathbf{Q} = (\sigma_s\tilde{\mathbf{h}}_{11})(\sigma_s\tilde{\mathbf{h}}_{11})^{H}. \tag{5.24}$$

From the subspace based detection theory [40], the estimate of $\sigma_s\tilde{\mathbf{h}}_{11}$, denoted as $\hat{\tilde{\mathbf{h}}}_{11}$ can be obtained as the eigenvector of $\mathbf{Q}$ that corresponds to the largest eigenvalue.

- **Step 3**: Denoise the estimate vector $\hat{\tilde{\mathbf{h}}}_{11}$ and obtain the final estimation result.

Remark: Instead of using only one column of $\mathbf{R_r}$, the proposed method extracts the channel estimation from the whole covariance matrix. Hence the final estimate is more accurate than the one purely estimated from a single column [87], [88].

5.3.3 Criteria for the Design of Precoders

1. *Distortion Constraint:*

   In real applications, the number of the snapshots received within the channel coherent time is finite. Therefore, the received signal covariance matrix is replaced by the sample covariance matrix. As a result, each entry of $\mathbf{R_r}$ will contain the distortion due to both the noise and the lack of the number of snapshots. If one entry of $\mathbf{P}_1$ is much smaller than the other entries, the distortion in the corresponding entry of $\mathbf{R_r}$ will be greatly enlarged after the elimination of $\mathbf{P}_1$. Hence, the entry of $\mathbf{R_r}$ with possibly large distortion should be discarded or the corresponding entry of $\mathbf{P}_1$ should be assigned a relatively large value. However, no prior information of the distortion can be obtained due to all unknown factors. A reasonable way is to assign equal value to all the non-diagonal entries, at the same time as large as possible.

2. *Power Constraint:*

   To keep the transmission power unchanged, matrix $\mathbf{P}_1$ should be designed such that

$$\text{Transmitting Power} = \mathrm{E}\{(\mathbf{s}_i^{(1)})^H \mathbf{W}_1^H \mathbf{W}_1 (\mathbf{s}_i^{(1)})\} = \sigma_s^2 tr(\mathbf{P}_1) = K\sigma_s^2 \quad (5.25)$$

or

$$tr(\mathbf{P}_1) = K. \tag{5.26}$$

Since the proposed algorithm equally considers all entries of $\mathbf{R_y}$, there is no reason why the value of one diagonal entry should dominate another. Consequently, we consider each diagonal element to be 1. Combined with the distortion constraint, the (a, c)th entry of $\mathbf{P}_1$ takes the value

$$[\mathbf{P}_1]_{ac} = \begin{cases} 1 & a = c \\ \bar{p} & a \neq c \end{cases} \qquad a,\ c = 1, \ldots, K \tag{5.27}$$

where $\bar{p}$ is some non-zero value.

3. *Symbol Error Constraint:*

It is known that, $\mathbf{P}_1$ should be designed as a positive definite matrix such that the transmitted symbols can be restored after the channel is estimated. It can be proved that, one of the eigenvalues of $\mathbf{P}_1$ in (5.27) is $(K-1)\bar{p}+1$ and all the other eigenvalues are $1-\bar{p}$. Hence, the range of $\bar{p}$ is $-\frac{1}{M-1} < \bar{p} < 1$.

Suppose at the receiver, the minimum mean square error (MMSE) equalizer is applied for detecting the symbols [89]-[91]. That is

$$\hat{\mathbf{s}}_i^{(1)} = \mathbf{G}\mathbf{r}_i^{(1)} \tag{5.28}$$

where

$$\begin{aligned} \mathbf{G} &= \mathrm{E}\{\mathbf{s}_i^{(1)}(\mathbf{r}_i^{(1)})^H\}\mathrm{E}\{\mathbf{r}_i^{(1)}(\mathbf{r}_i^{(1)})^H\}^{-1} \\ &= \mathbf{W}_1^H\mathbf{\Lambda}_{11}^H(\mathbf{\Lambda}_{11}\mathbf{P}_1\mathbf{\Lambda}_{11}^H + \frac{\sigma_n^2}{\sigma_s^2}\mathbf{I})^{-1}. \end{aligned} \tag{5.29}$$

The MSE of the estimation can be obtained as

$$\begin{aligned} \mathrm{MSE} &= \mathrm{E}\{\|\mathbf{G}\mathbf{r}_i^{(1)} - \mathbf{s}_i^{(1)}\|^2\} \\ &= \sigma_s^2 tr(\mathbf{I} - (\mathbf{\Lambda}_{11}\mathbf{P}_1\tilde{\mathbf{H}}_{11}^H + \frac{\sigma_n^2}{\sigma_s^2}\mathbf{I})^{-1}\tilde{\mathbf{H}}_{11}\mathbf{P}_1\mathbf{\Lambda}_{11}^H) \\ &= \sigma_n^2 tr((\mathbf{\Lambda}_{11}\mathbf{P}_1\mathbf{\Lambda}_{11}^H + \frac{\sigma_n^2}{\sigma_s^2}\mathbf{I})^{-1}). \end{aligned} \tag{5.30}$$

Let us first provide the following lemma.

Lemma 5.1: For any positive definite $K \times K$ matrix $\mathbf{A}$, the following inequality holds [92]:

$$tr(\mathbf{A}^{-1}) \geq \sum_{a=1}^{K} ([\mathbf{A}]_{aa})^{-1} \tag{5.31}$$

and the equality holds if $\mathbf{A}$ is diagonal. Since $\mathbf{\Lambda}_{11}$ is a diagonal matrix, the minimum value of MSE can only be achieved when $\mathbf{P}_1$ is also a diagonal matrix. With the fact that $\mathbf{\Lambda}_{11}$ is unknown, a reasonable choice of $\mathbf{P}_1$ is to be an identity matrix. In this case, $\bar{p}$ is expected to be 0, which makes the channel estimation impossible. To reduce the symbol detection error, one may try to decrease the value of $|\bar{p}|$. However, this approach will enlarge the distortion effect as stated in the distortion constraint part.

From the above discussion, it is seen that there exists a compromise between channel estimation error and data detection error.

5.4 Blind Channel Estimation for MIMO Systems.

5.4.1 MIMO Channel Estimation with Ambiguity

For MISO and MIMO systems, each receiver receives a combination of the transmitted symbols from all the transmitters. Suppose the symbol from the pth transmitter is separately precoded by the matrix $\mathbf{W}_p$. The normalized DFT of the received signal vector from the qth receiver can be written from (5.13) as

$$\mathbf{y}_i^{(q)} = \sum_{p=1}^{N_t} \mathbf{\Lambda}_{pq} \mathbf{W}_p \mathbf{s}_i^{(p)} + \mathbf{n}_i^{(q)}. \tag{5.32}$$

Let $\boldsymbol{\mathcal{P}}$ be the block diagonal matrix with the form

$$\boldsymbol{\mathcal{P}} = \text{diag}\{\mathbf{P}_1,\ \mathbf{P}_2, \ldots,\ \mathbf{P}_{N_t}\} \tag{5.33}$$

where

$$\mathbf{P}_p = \mathbf{W}_p \mathbf{W}_p^H. \tag{5.34}$$

The signal covariance matrix $\mathbf{R_r}$ is then obtained as

$$\begin{aligned}\mathbf{R_r} &= \mathrm{E}\{\mathbf{r}_i \mathbf{r}_i^H\} = \sigma_s^2 \boldsymbol{\mathcal{H}}_\mathbf{r} \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{H}}_\mathbf{r}^H + \sigma_n^2 \mathbf{I}_{KN_r} \\ &= \begin{bmatrix} \mathbf{R}_{\mathbf{r},11} & \dots & \mathbf{R}_{\mathbf{y},1N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{\mathbf{r},N_r1} & \dots & \mathbf{R}_{\mathbf{r},N_rN_r} \end{bmatrix}\end{aligned} \tag{5.35}$$

with its (b,d)th sub-block given by

$$\begin{aligned}\mathbf{R}_{\mathbf{y},bd} &= \sigma_s^2 \sum_{p=1}^{N_t} \boldsymbol{\Lambda}_{pb} \mathbf{P}_p \boldsymbol{\Lambda}_{pd}^H + \delta(b-d)\sigma_n^2 \mathbf{I}_K \\ &= \sigma_s^2 \sum_{p=1}^{N_t} (\tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H) \odot \mathbf{P}_p + \delta(b-d)\sigma_n^2 \mathbf{I}_K, \quad b,\ d = 1, \dots, N_r\end{aligned} \tag{5.36}$$

where $\delta(\cdot)$ is the delta function that equals one at $b = d$ and zero otherwise.

From (5.35), it is hard to eliminate the effect of all $\mathbf{P}_p$'s for every entry of $\mathbf{R_r}$. A tricky way here is to take all the $\mathbf{P}_p$ same, namely:

$$\mathbf{P} = \mathbf{P}_1 = \dots = \mathbf{P}_{N_t} = \mathbf{W}\mathbf{W}^H. \tag{5.37}$$

With this effort, $\mathbf{R}_{\mathbf{r},bd}$ can be rewritten as

$$\mathbf{R}_{\mathbf{r},bd} = \sigma_s^2 \left(\sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H \right) \odot \mathbf{P} + \delta(b-d)\sigma_n^2 \mathbf{I}_K. \tag{5.38}$$

We will continue the discussion by considering the following two cases.

Case 1: $b \neq d$. In this case, $\mathbf{R}_{\mathbf{r},bd}$ simply has the form

$$\mathbf{R}_{\mathbf{r},bd} = \sigma_s^2 \left(\sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H \right) \odot \mathbf{P}. \tag{5.39}$$

Dividing the (a,c)th entry of $\mathbf{R}_{\mathbf{r},bd}$ by $[\mathbf{P}]_{ac}$, we obtain

$$\mathbf{Q}_{bd} = \sigma_s^2 \sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H, \quad b \neq d. \tag{5.40}$$

Case 2: $b = d$. For this case, $\mathbf{R}_{\mathbf{r},bd}$ is written as

$$\mathbf{R}_{\mathbf{r},bd} = \sigma_s^2 \left(\sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H \right) \odot \mathbf{P} + \sigma_n^2 \mathbf{I}_K. \tag{5.41}$$

Noting the similarity between (5.41) and (5.14), we can define

$$\breve{\mathbf{r}}_{bd,c} = \left[\frac{[\mathbf{R}_{\mathbf{r},bd}]_{1c}}{[\mathbf{P}]_{1c}}, \ldots, \frac{[\mathbf{R}_{\mathbf{r},bd}]_{(c-1)c}}{[\mathbf{P}]_{(c-1)c}}, \frac{[\mathbf{R}_{\mathbf{r},bd}]_{(c+1)c}}{[\mathbf{P}]_{(c+1)c}}, \ldots, \frac{[\mathbf{R}_{\mathbf{r},bd}]_{ac}}{[\mathbf{P}]_{ac}} \right]^T. \tag{5.42}$$

Similarly from (5.22), we can obtain

$$\sigma_s^2 \sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{h}_{pd}^*(c-1) = \mathbf{F}(:, 1:L+1) \mathbf{F}_c^\dagger \mathbf{r}_{bd,c}, \quad c = 1, \ldots, K \tag{5.43}$$

where $\mathbf{F}_c$ is defined in (5.20). Therefore

$$\mathbf{Q}_{bd} = \left[\sigma_s^2 \sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{h}_{pd}^*(0), \ldots, \sigma_s^2 \sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{h}_{pd}^*(K-1) \right] = \sigma_s^2 \sum_{p=1}^{N_t} \tilde{\mathbf{h}}_{pb} \tilde{\mathbf{h}}_{pd}^H \tag{5.44}$$

could also be constructed for $b = d$.

Define

$$\boldsymbol{\mathcal{Q}} = \begin{bmatrix} \mathbf{Q}_{11} & \ldots & \mathbf{Q}_{1N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{Q}_{N_r 1} & \ldots & \mathbf{Q}_{N_r N_r} \end{bmatrix} \tag{5.45}$$

$$\mathbf{U} = \begin{bmatrix} \tilde{\mathbf{h}}_{11} & \ldots & \tilde{\mathbf{h}}_{N_t 1} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{h}}_{1N_r} & \ldots & \tilde{\mathbf{h}}_{N_t N_r} \end{bmatrix}. \tag{5.46}$$

It can be verified that

$$\boldsymbol{\mathcal{Q}} = \sigma_s^2 \mathbf{U}\mathbf{U}^H = (\sigma_s \mathbf{U})(\sigma_s \mathbf{U})^H. \tag{5.47}$$

From the subspace detection theory, the matrix $\mathbf{U}$ can be estimated from the N_t eigenvectors of $\boldsymbol{\mathcal{Q}}$ which corresponds to the largest N_t eigenvalues. However, this estimate differs from the true value by an $N_t \times N_t$ matrix. Namely,

$$\hat{\mathbf{U}} = \mathbf{U}\mathbf{B} \tag{5.48}$$

where $\mathbf{B}$ is an $N_t \times N_t$ unknown matrix. This higher dimensional ambiguity commonly exists in the blind channel estimation when multiple transmit antennas are used [79]. As in Chapter 4, some training symbols should be transmitted or the finite alphabet property should be utilized to resolve this ambiguity. We will see in the next subsection that the higher dimensional ambiguity can also be eliminated by adopting a different scheme of precoding.

Remarks:

- It is seen that the proposed algorithm directly operates on the signal covariance matrix and the source covariance is assumed as $\mathbf{R}_s = \sigma_s^2\mathbf{I}$. Consequently, the channel ambiguity issue is automatically removed.

- Moreover, since the channel frequency response matrix $\mathbf{U}$ is estimated, the channel length constraint in the time domain is removed. That is, no matter what L is, the vector $\tilde{\mathbf{h}}_{pq}$ of length K is always estimated. So, the channel over-order estimation only degrades the performance of the channel estimation but will not fail the algorithm.

- Throughout the derivation, we do not put any restriction on the number of the transmit and the receive antennas. Therefore, the proposed algorithm is applicable for MIMO system when N_t is greater than or equal to N_r, in which case the traditional SS method [40]-[45] cannot be applied. However, the value of N_t should not exceed $N_r(L+1)-1$. The reason is illustrated as follows.

  Let

$$\mathbf{V} = \begin{bmatrix} \mathbf{h}_{11} & \dots & \mathbf{h}_{N_t 1} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{1N_r} & \dots & \mathbf{h}_{N_t N_r} \end{bmatrix} \tag{5.49}$$

  denote the $N_r(L+1) \times N_t$ matrix containing time domain channel response. The estimate of $\mathbf{V}$, denoted as $\hat{\mathbf{V}}$, can be obtained from $\hat{\mathbf{U}}$ via IDFT operation and is related with $\mathbf{V}$ by

$$\hat{\mathbf{V}} = \mathbf{V}\mathbf{B}. \tag{5.50}$$

If N_t is greater than or equal to $N_r(L+1)$, the estimation (5.50) would be meaningless. Because in this case, any $N_r(L+1) \times N_t$ matrix will be away from $\mathbf{V}$ by an $N_t \times N_t$ matrix $\mathbf{B}$. However, as can be seen in the next subsection, this upper bound does not exist if the higher dimensional ambiguity problem is resolved.

- By following the similar steps in section 5.3.3, the same design criterion for precoder $\mathbf{P}$ can be derived and $\mathbf{P}$ with the form (5.27) is also suggested as a good precoder for MIMO OFDM systems.

5.4.2 MIMO Channel Estimation with Scalar Ambiguity

We also propose the following scheme to eliminate the higher dimensional ambiguity directly through the channel estimation. Suppose at the $iN_t + \tau$th interval $\tau = 1, \ldots, N_t$, the symbol block from the pth transmitter is precoded by $\mathbf{W}_{p\tau}$ and the corresponding $\mathbf{P}_{p\tau}$ is defined as

$$\mathbf{P}_{p\tau} = \mathbf{W}_{p\tau}\mathbf{W}_{p\tau}^H, \quad p, \tau = 1, \ldots, N_t. \tag{5.51}$$

Define N_t covariance matrices as

$$\mathbf{R}_{\mathbf{r}\tau} = \mathrm{E}\{\mathbf{r}(iN_t + \tau)\mathbf{r}(iN_t + \tau)^H\}, \quad \tau = 1, \ldots, N_t. \tag{5.52}$$

We know from (5.35) and (5.36) that, the (b, d)th block of $\mathbf{R}_{\mathbf{r}\tau}$ has the form

$$\begin{aligned} \mathbf{R}_{\mathbf{r}\tau,bd} &= \sigma_s^2 \sum_{p=1}^{N_t} \mathbf{\Lambda}_{pb}\mathbf{P}_{p\tau}\mathbf{\Lambda}_{pd}^H + \delta(b-d)\sigma_n^2\mathbf{I}_K \\ &= \sigma_s^2 \sum_{p=1}^{N_t} (\tilde{\mathbf{h}}_{pb}\tilde{\mathbf{h}}_{pd}^H) \odot \mathbf{P}_{p\tau} + \delta(b-d)\sigma_n^2\mathbf{I}_K. \end{aligned} \tag{5.53}$$

Similar steps for estimating the channel response can be carried on as follows:
Case 1: $b \neq d$. For this case, the (a, c)th entry of $\mathbf{R}_{\mathbf{r}\tau,bd}$ can be expressed as

$$[\mathbf{R}_{\mathbf{r}\tau,bd}]_{ac} = \sigma_s^2 \sum_{p=1}^{N_t} [\mathbf{P}_{p\tau}]_{ac}\tilde{h}_{pb}(a-1)\tilde{h}_{pd}^*(c-1), \quad \tau = 1, \ldots, N_t. \tag{5.54}$$

Note that (5.54) contains N_t equations of N_t unknown parameters $\sigma_s^2\tilde{h}_{pb}(a-1)\tilde{h}_{pd}^*(c-1)$, $p=1,\ldots,N_t$. Therefore the unknown parameters can be obtained as

$$\begin{bmatrix} \sigma_s^2\tilde{h}_{1b}(a-1)\tilde{h}_{1d}^*(c-1) \\ \vdots \\ \sigma_s^2\tilde{h}_{N_tb}(a-1)\tilde{h}_{N_td}^*(c-1) \end{bmatrix} = \begin{bmatrix} [\mathbf{P}_{11}]_{ac} & \ldots & [\mathbf{P}_{N_t1}]_{ac} \\ \vdots & \ddots & \vdots \\ [\mathbf{P}_{1N_t}]_{ac} & \ldots & [\mathbf{P}_{N_tN_t}]_{ac} \end{bmatrix}^{-1} \begin{bmatrix} [\mathbf{R}_{\mathbf{y}1,bd}]_{ac} \\ \vdots \\ [\mathbf{R}_{\mathbf{y}N_t,bd}]_{ac} \end{bmatrix} \tag{5.55}$$

Since (5.55) holds if and only if the inverse item exists, we should no longer grant the same value to all $\mathbf{P}_{p\tau}$ as suggested in subsection 5.4.1. Instead, $\mathbf{P}_{p\tau}$ should be designed such that the square matrix in (5.55) is non-singular. By considering all pairs of (a,c) and properly reorganizing the results coming from (5.55), we can obtain N_t new matrices

$$\mathbf{Q}_{p,bd} = \sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{\mathbf{h}}_{pd}^H \quad \text{for} \quad p=1,\ldots,N_t. \tag{5.56}$$

Case 2: $b=d$. For this case, the diagonal entries of $\mathbf{R}_{\mathbf{r}\tau,bd}$ are corrupted by the unknown noise power. Therefore, we may consider only the entries with $a \neq c$. There is

$$[\mathbf{R}_{\mathbf{r}\tau,bd}]_{ac} = \sigma_s^2 \sum_{p=1}^{N_t} [\mathbf{P}_{p\tau}]_{ac}\tilde{h}_{pb}(a-1)\tilde{h}_{pd}^*(c-1) \tag{5.57}$$

for

$$\tau = 1,\ldots,N_t, \quad b=d, \quad a \neq c.$$

With a similar step in case 1, we can obtain the value of $\sigma_s^2\tilde{h}_{pb}(a-1)\tilde{h}_{pd}^*(c-1)$ from (5.55) for all $a \neq c$. Then, a new vector can be formed as follows

$$\begin{aligned} \breve{\mathbf{r}}_{bd,c} =& [\sigma_s^2\tilde{h}_{pb}(0)\tilde{h}_{pd}^*(c-1),\ldots,\sigma_s^2\tilde{h}_{pb}(c-2)\tilde{h}_{pd}^*(c-1), \\ & \sigma_s^2\tilde{h}_{pb}(c)\tilde{h}_{pd}^*(c-1),\ldots,\sigma_s^2\tilde{h}_{pb}(K-1)\tilde{h}_{pd}^*(c-1)]^T \end{aligned} \tag{5.58}$$

for each $c=1,\ldots,K$. Following similar steps through (5.21) to (5.22), the vector $\sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{h}_{pd}^*(c-1)$ can be obtained as

$$\sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{h}_{pd}^*(c-1) = \mathbf{F}(:,1:L+1)\mathbf{F}_c^{\dagger}\breve{\mathbf{r}}_{bd,c}. \tag{5.59}$$

Combining all the results of $\sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{h}_{pd}^*(c-1), c = 1, \ldots, K$, we can obtain the estimate

$$\mathbf{Q}_{p,bd} = [\sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{h}_{pd}^*(0), \ldots, \sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{h}_{pd}^*(K-1)] = \sigma_s^2\tilde{\mathbf{h}}_{pb}\tilde{\mathbf{h}}_{pd}^H \quad \text{for} \quad p = 1, \ldots, N_t. \tag{5.60}$$

Define a new matrix

$$\boldsymbol{\mathcal{Q}}_p = \begin{bmatrix} \mathbf{Q}_{p,11} & \cdots & \mathbf{Q}_{p,1N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{Q}_{p,N_r1} & \cdots & \mathbf{Q}_{p,N_rN_r} \end{bmatrix}. \tag{5.61}$$

It can be readily seen that,

$$\boldsymbol{\mathcal{Q}}_p = \sigma_s^2\mathbf{U}_p\mathbf{U}_p^H, \quad p = 1, \ldots, N_t \tag{5.62}$$

where

$$\mathbf{U}_p = [\tilde{\mathbf{h}}_{p1}^T, \tilde{\mathbf{h}}_{p2}^T, \ldots, \tilde{\mathbf{h}}_{pN_r}^T]^T \tag{5.63}$$

is the pth column of $\mathbf{U}$ and represents the $N_rK \times 1$ channel response vector purely from the pth transmitter to all the receivers. Again, $\mathbf{U}_p$ can be obtained from the eigenvector of $\boldsymbol{\mathcal{Q}}_p$ corresponding to the largest eigenvalue.

Therefore, by assigning different precoding matrix to different transmitter and taking N_t covariance matrices from different time slot, the higher dimensional ambiguity reduces to one scalar ambiguity for each $\mathbf{U}_p$, namely

$$\hat{\mathbf{U}}_p = \alpha_p\mathbf{U}_p \tag{5.64}$$

where α_p is an unknown complex scalar. In this case, there is no upper bound constraint on the number of the transmit antennas. Since N_t covariance matrices need to be constructed in real applications, the number of available snapshots seems to be critical to the performance of the algorithm. Roughly speaking, in order to build the comparably reliable covariance matrix as the one in the matrix ambiguity case, N_t times more snapshots need to be obtained.

Remark: A general guideline for designing $\mathbf{P}_{p\tau}$ is provided here. Similar to SISO case, the power constraint can be satisfied by choosing proper diagonal values for all $\mathbf{P}_{p\tau}$. However, special attention should be paid for the distortion constraint. One intuition is that $\mathbf{P}_{p\tau}$ should be designed such that each entry of the inverse

matrix in (5.55) takes on a value as small as possible for all pairs of (a, c) in order to reduce the distortion enhancement effect.

5.4.3 Symbol Detection

The proposed blind algorithm forms a novel contribution since it is a blind channel estimation method applicable for the MIMO OFDM with $N_t \geq N_r$. Common application may lie in downlink communications where there are multiple antennas at the base station, however, only one or two antennas are available at the mobile terminal. It is also possible to apply the proposed algorithm in the cooperative communications [93]-[97] where there are, in general, more relays but only one destination.

For the case $N_t > N_r$, the linear detection may not be reliable and the ML algorithm is usually carried out by an exhaustive search. Alternatively, specific consideration can be made in the structure of the transmitted symbols to guarantee a linear detection. The STC technique [17] will be a good choice.

For convenience, we only discuss the case with channel estimation ambiguity. We consider a 2×1 transmission system with the implementation of Alamouti codes [98] at the transmitter. Suppose in the $2i$th block interval, $\mathbf{s}_{2i}$ and $\mathbf{s}_{2i+1}$ are transmitted from the first and the second transmitters, respectively, and in the $(2i+1)$th block interval, $-\mathbf{s}_{2i+1}^*$ and $\mathbf{s}_{2i}^*$ are transmitted from the first and second transmitters, respectively. The received signals from the $(2i)$th block and the $(2i+1)$th block are given by

$$\mathbf{r}_{2i} = [\mathbf{\Lambda}_{11},\ \mathbf{\Lambda}_{21}] \begin{bmatrix} \mathbf{W}\mathbf{s}_{2i} \\ \mathbf{W}\mathbf{s}_{2i+1} \end{bmatrix} + \mathbf{n}_{2i} \tag{5.65}$$

$$\mathbf{r}_{2i+1} = [\mathbf{\Lambda}_{11}, \mathbf{\Lambda}_{21}] \begin{bmatrix} -\mathbf{W}\mathbf{s}_{2i+1}^* \\ \mathbf{W}\mathbf{s}_{2i}^* \end{bmatrix} + \mathbf{n}_{2i+1} \tag{5.66}$$

where $\mathbf{n}_{2i}$ and $\mathbf{n}_{2i+1}$ are the corresponding noise vectors. From our design of

precoder, $\mathbf{W} = \mathbf{P}^{1/2}$ is taken as a real matrix, therefore

$$\begin{bmatrix} \mathbf{r}_{2i} \\ \mathbf{r}^*_{2i+1} \end{bmatrix} = \begin{bmatrix} \mathbf{\Lambda}_{11} & \mathbf{\Lambda}_{21} \\ \mathbf{\Lambda}^*_{21} & -\mathbf{\Lambda}^*_{11} \end{bmatrix} \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{bmatrix} \begin{bmatrix} \mathbf{s}_{2i} \\ \mathbf{s}_{2i+1} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{2i} \\ \mathbf{n}^*_{2i+1} \end{bmatrix} \tag{5.67}$$

or equivalently

$$\bar{\mathbf{r}} = \bar{\mathbf{\Lambda}}\bar{\mathbf{W}}\bar{\mathbf{s}} + \bar{\mathbf{n}} \tag{5.68}$$

where $\bar{\mathbf{r}}$, $\bar{\mathbf{\Lambda}}$, $\bar{\mathbf{W}}$, $\bar{\mathbf{n}}$ representing the corresponding items in (5.67). It can be easily known that $\bar{\mathbf{\Lambda}}$ is a unitary matrix. Then, the transmitted symbols is simply detected from

$$\bar{\mathbf{s}} = \bar{\mathbf{W}}^{-1}\bar{\mathbf{\Lambda}}^H\bar{\mathbf{r}}. \tag{5.69}$$

Note that, $\bar{\mathbf{W}}$ is a pre-calculated term and can be stored in advance.

More general STC other than Alamouti code can also be utilized. In this case, $\mathbf{R_s}$ may not be a multiple of the identity matrix. However, as long as different sources send i.i.d. symbols, the value of $\mathbf{R_s}$ should be known in advance. Therefore, we need only to redefine the new $\mathbf{P}$ as

$$\mathbf{P} = \mathbf{W}\mathbf{R_s}\mathbf{W} \tag{5.70}$$

and the proposed algorithm still works.

5.5 Stochastic Cramér-Rao Bound

In this section we derive the CRB for the proposed channel estimation algorithm. For simplicity, we only focus on the discussion for the ambiguity case.

Since the channel matrix $\boldsymbol{\mathcal{H}}_\mathbf{r}$ or $\boldsymbol{\mathcal{H}}_\mathbf{y}$ in our model could be a square or a fat matrix, the deterministic CRB is no longer applicable. Instead, we consider the stochastic CRB [99], [100]. Here we do not use the covariance of $\mathbf{r}_i$, since $\tilde{\mathbf{h}}_{pq}$ has only $L+1$ degree of freedom. Namely, once $L+1$ elements of $\tilde{\mathbf{h}}_{pq}$ are fixed, then other elements of $\tilde{\mathbf{h}}_{pq}$ can be consequently determined. Therefore, the signal covariance

matrix of $\mathbf{y}_i$ is used and is parameterized by the $(N_t N_r(2L+2)+2)\times 1$ vector $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$ with

$$\boldsymbol{\theta}_1 = [\mathrm{Re}\{\mathbf{h}_{11}\}^T, \mathrm{Im}\{\mathbf{h}_{11}\}^T, \ldots, \mathrm{Re}\{\mathbf{h}_{pq}\}^T, \mathrm{Im}\{\mathbf{h}_{pq}\}^T, \ldots, \mathrm{Re}\{\mathbf{h}_{N_t N_r}\}^T, \mathrm{Im}\{\mathbf{h}_{N_t N_r}\}^T]^T \tag{5.71}$$

$$\boldsymbol{\theta}_2 = [\sigma_s^2, \sigma_n^2]^T. \tag{5.72}$$

The covariance matrix $\mathbf{R}_\mathbf{y}$ is then given by

$$\mathbf{R}_\mathbf{y} = \mathrm{E}\{\mathbf{y}_i \mathbf{y}_i^H\} = \sigma_s^2 \boldsymbol{\mathcal{H}}_\mathbf{y} \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{H}}_\mathbf{y}^H + \sigma_n^2 \mathbf{I}_{KN_r}. \tag{5.73}$$

Define $K \times K$ derivative matrices $\mathbf{T}_l$ with the (a,c)th entry

$$[\mathbf{T}_l]_{ac} = \begin{cases} 1 & \text{for } ((a-c) \bmod K) = l-1 \\ 0 & \text{otherwise} \end{cases}, \quad l = 1, \ldots, L+1 \tag{5.74}$$

and

$$\mathbf{D}_{pq,l} = \; q\text{th}\{ \underbrace{\begin{bmatrix} \mathbf{0} & \ldots & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathbf{T}_l & \ldots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \ldots & \mathbf{0} & \ldots & \mathbf{0} \end{bmatrix}}_{p\text{th.}} \quad \text{for} \quad \begin{matrix} p = 1, \ldots, N_t \\ q = 1, .., N_r \end{matrix} \tag{5.75}$$

Let

$$\mathbf{A}_{pq,l} = \sigma_s^2 \mathbf{R}_\mathbf{y}^{-1/2} \mathbf{D}_{pq,l} \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{H}}_\mathbf{y}^H \mathbf{R}_\mathbf{y}^{-1/2} \tag{5.76a}$$

$$\mathbf{G}_{pq}^R(:,l) = \mathrm{vec}(\mathbf{A}_{pq,l} + \mathbf{A}_{pq,l}^H) \tag{5.76b}$$

$$\mathbf{G}_{pq}^I(:,l) = \mathrm{vec}(j\mathbf{A}_{pq,l} - j\mathbf{A}_{pq,l}^H) \tag{5.76c}$$

$$\mathbf{v} = \mathrm{vec}(\mathbf{R}_\mathbf{y}^{-1/2} \boldsymbol{\mathcal{H}}_\mathbf{y} \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{P}} \boldsymbol{\mathcal{F}} \boldsymbol{\mathcal{H}}_\mathbf{y}^H \boldsymbol{\mathcal{R}}_\mathbf{y}^{-1/2}) \tag{5.76d}$$

$$\mathbf{u} = \mathrm{vec}(\mathbf{R}_\mathbf{y}^{-1}) \tag{5.76e}$$

$$\mathbf{G}_{pq} = [\mathbf{G}_{pq}^R \; \mathbf{G}_{pq}^I] \tag{5.76f}$$

$$\mathbf{G} = [\mathbf{G}_{11} \; , \ldots, \mathbf{G}_{pq} \; , \ldots, \mathbf{G}_{N_t N_r}] \tag{5.76g}$$

$$\boldsymbol{\Delta} = [\mathbf{v} \; \mathbf{u}]. \tag{5.76h}$$

Suppose M data blocks are received, the stochastic CRB of channel vector $\boldsymbol{\theta}_1$ is given by

$$\mathrm{CRB}_{\boldsymbol{\theta}_1} = \frac{1}{M}(\mathbf{G}^H \mathbf{P}_{\boldsymbol{\Delta}}^{\perp} \mathbf{G})^{\dagger} \tag{5.77}$$

where

$$\mathbf{P}_{\boldsymbol{\Delta}}^{\perp} = \mathbf{I} - \boldsymbol{\Delta}(\boldsymbol{\Delta}^H \boldsymbol{\Delta})^{-1} \boldsymbol{\Delta}^H \tag{5.78}$$

is the projection matrix onto the orthogonal complement space of $\boldsymbol{\Delta}$.

Proof: See Appendix D. ■

5.6 Simulations

In this section, we examine the performance of the proposed estimator under various scenarios. The 3-ray channel model with exponential power delay profile [73]

$$\mathrm{E}\{|h_{pq}(l)|^2\} = \exp(-l/10), \quad l = 0, \ldots, 2 \tag{5.79}$$

is used. The phase of each channel ray is uniformly distributed over $[0, 2\pi)$. For all numerical examples, the QPSK symbols are considered and the parameters are taken as $K = 32$, $L = 2$. In each simulation run, 150 blocks of signals are obtained and all the results are averaged over $N_w = 100$ Monte-Carlo runs.

The normalized estimation mean square errors (NMSE) is defined the same as in (4.40).

1) SISO OFDM Systems

We first illustrate the performance comparison of the proposed joint estimator with the single column estimator in [87]. For a relatively fair comparison, the value of $\bar{p}$ for the proposed precoder is obtained in a way that the MMSEs (5.30) for both the proposed precoder and the precoder in [87] are the same. In Fig. 5.1, the NMSEs of channel estimation versus SNR of these two different algorithms are displayed. Line 1 and line 2 show the comparison between the proposed estimator and single column estimator under the precoder suggested in [87]. We see that the joint estimator is better than the single column estimator since we considered full

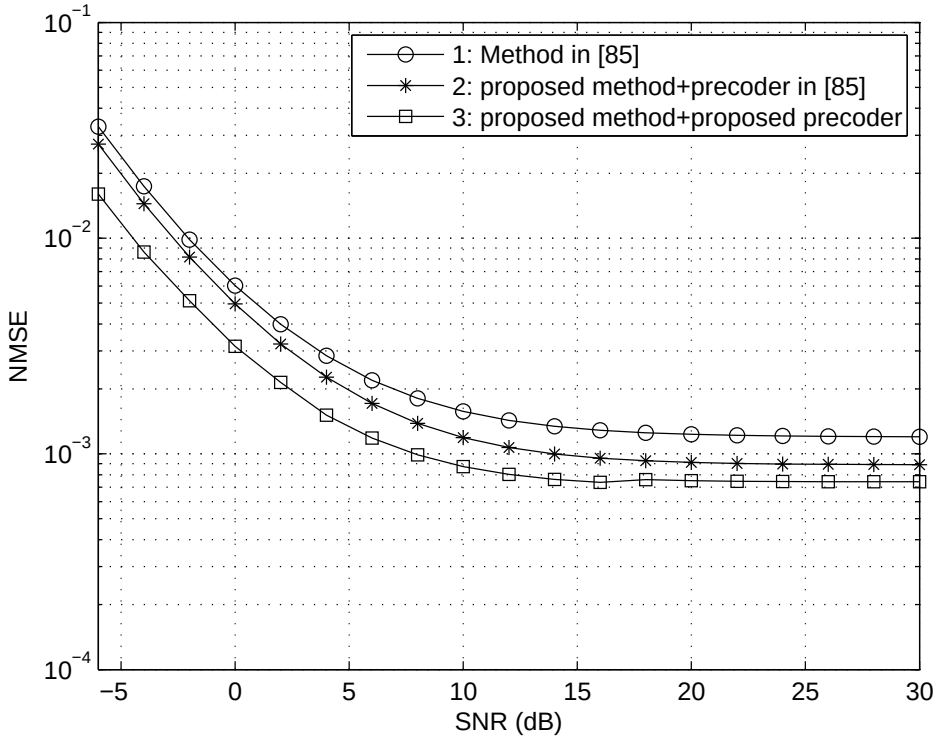

Figure 5.1: Comparison with the existing work in SISO OFDM.

exploitation of available information. At the same time, comparison between line 2 and line 3 clearly shows the difference of the proposed estimator under two different precoders, from which we see that our proposed precoder is much better than the precoder in [87].

Next we consider the performance variation for different values of $\bar{p}$. Three values of $\bar{p}$ as 0.25, 0.62, 1 are chosen and the corresponding NMSEs versus SNR are shown in Fig. 5.2. At the same time the three stochastic CRB curves are drawn from up to down representing the CRBs for 0.25, 0.62, and 1, respectively. We see that the value of $\bar{p}$ is critical to the performance of the proposed algorithm. As shown in section III, the increase in the value of $\bar{p}$ will cause the performance of the proposed algorithm to be better. At the same time, the gap between the proposed algorithm and their corresponding CRB becomes smaller. Moreover, we can also see that both the NMSE curves and the CRB curve for those with $\bar{p} \neq 1$ meet lower

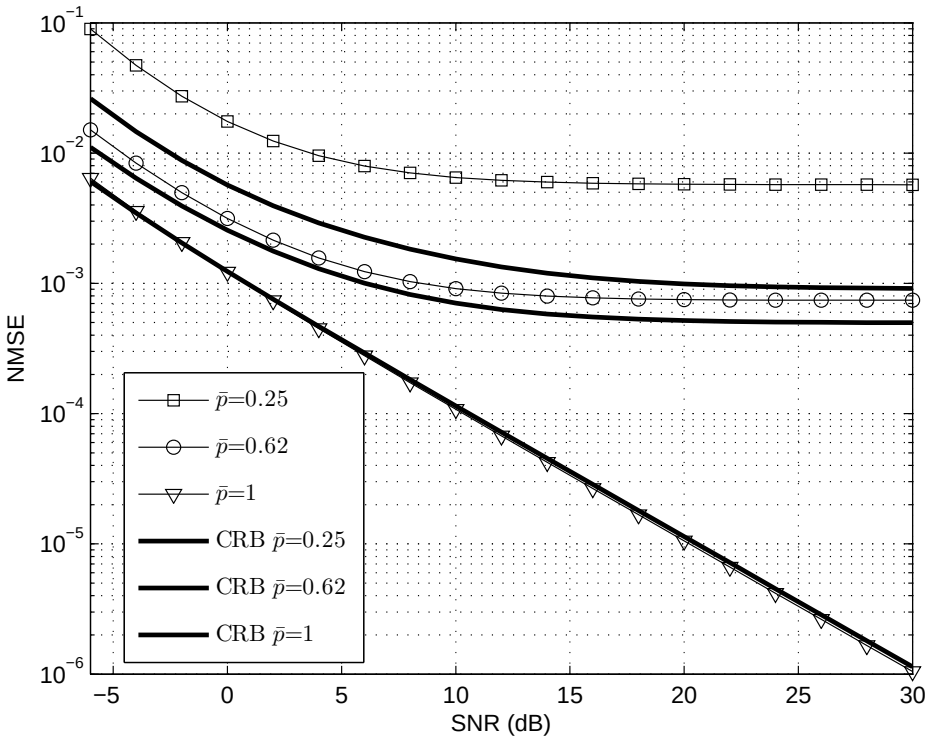

Figure 5.2: Performance of the proposed algorithm for SISO OFDM under different $\bar{p}$.

floors after the SNR reaches a certain level. Remember that the proposed algorithm is based on the assumption that the source covariance is a multiple of an identity matrix. Since only finite number of the snapshots is available in the simulation, an extra distortion (not related with SNR) will occur when we divide each entry of the covariance matrix by $\bar{p}$, causing the estimation floor irrelevant with SNR. However, for $\bar{p} = 1$, each entry of covariance matrix does not go through such a dividing and this extra distortion does not exist. Hence, no lower floor will be met for $\bar{p} = 1$.

Lastly, we examine BER of the MMSE detector with the channel response estimated from the proposed algorithm. In Fig. 5.3, the BERs versus SNR are plotted for $\bar{p}$ equals, 0.2, 0.4, 0.5, 0.6, 0.8, respectively. From Fig. 5.3, we numerically answer the question of how to make a good trade off when selecting the value of $\bar{p}$. Basically, $\bar{p}$ around 0.5 gives better performance and the BER reaches 10^{-5} at

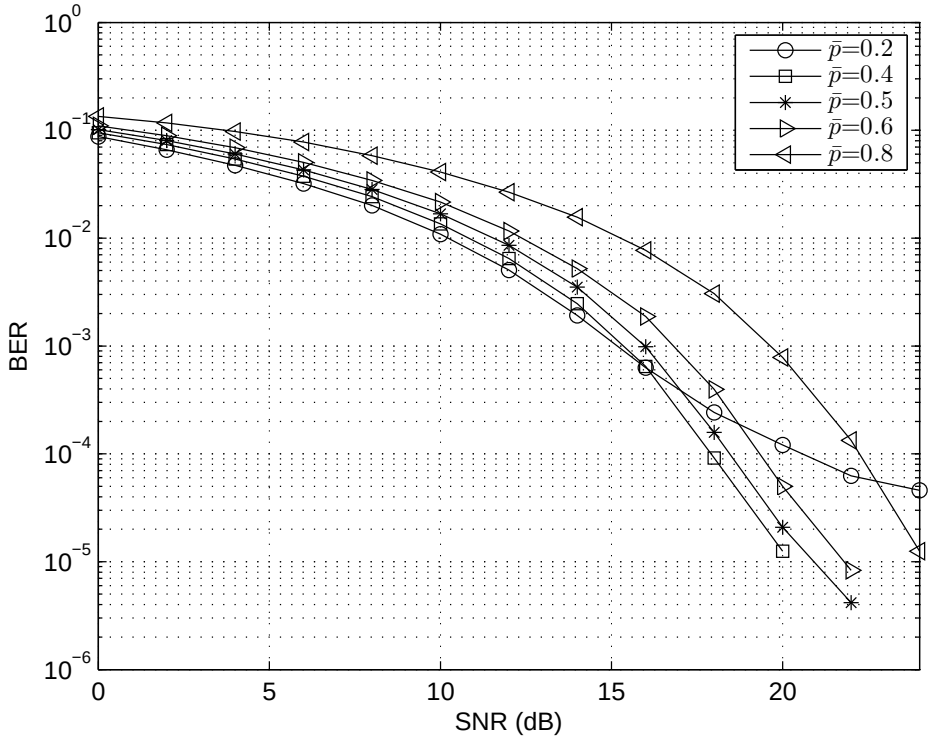

Figure 5.3: BERs for SISO OFDM under different $\bar{p}$.

20 dB.

2) MIMO OFDM Systems

In the second example, we show the performance results of the proposed algorithm for the MIMO OFDM system with two transmit antennas and two receive antennas.

Firstly, we consider the case where the channel estimation with matrix ambiguity. Three values of $\bar{p}$ as 0.25, 0.62, 1 are chosen and the corresponding NMSEs versus SNR are shown in Fig. 5.4. In the meantime, the three stochastic CRBs are also displayed from up to down representing the CRB for 0.25, 0.62, and 1, respectively. Similar to SISO scenario, larger $\bar{p}$ gives better performance.

Next, we show the performance NMSEs versus the number of the snapshots for the same MIMO system. It is expected that as the number of the signal block increases, the estimate of the covariance matrix $\mathbf{R_r}$ becomes accurate thus will result

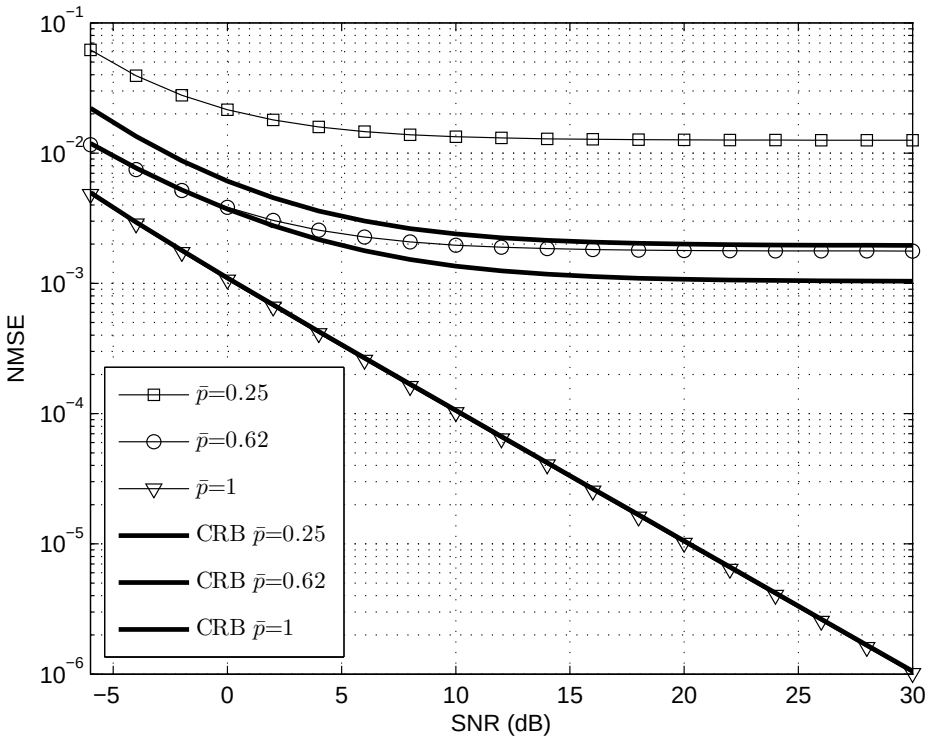

Figure 5.4: Performance NMSEs for MIMO OFDM versus SNR.

in a better estimation for $\tilde{\mathbf{h}}_{pq}$. Fig. 5.5 shows the NMSEs versus the number of blocks of these two estimators for the fixed SNR= 10 dB. The CRBs are displayed as well. It can be seen from this figure that estimation with block number greater than 150 can give relatively good performance.

The BER performance of the proposed method for MIMO systems is shown in Fig. 5.6. It is seen that small $\bar{p}$ performs similarly for low SNR region. Furthermore, small values of $\bar{p}$ perform better at low SNR whereas large values of $\bar{p}$ perform better at high SNR, say above 30 dB. We note that the BER for MIMO is a little bit worse than that for SISO case. Since the STC technique [17] is not applied and symbols are generated independently for each transmitter, the simulated MIMO scenario could only be considered as a spatial multiplexing system. Moreover, it is well known that linear receivers in spatial multiplexing systems suffer from the lack of diversity [101]. Actually, the diversity order for linear receivers can be approximated as $N_r - N_t + 1$,

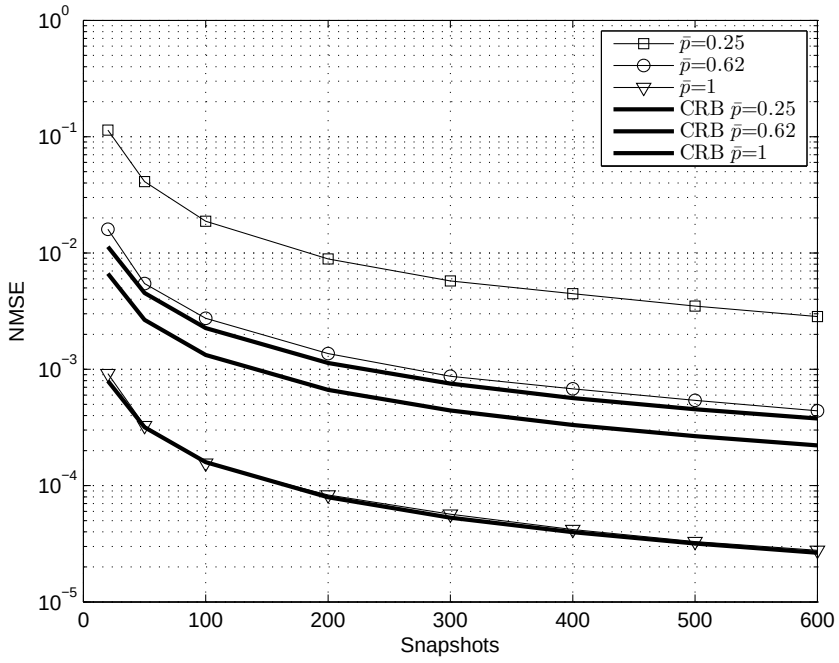

Figure 5.5: Performance NMSEs for MIMO OFDM versus number of snapshots.

which is only 1 for the simulated scenario. Therefore, no diversity is introduced here. Furthermore, we use the estimated channel for symbol detection. Note that, most diversity works are based on perfect channel knowledge. Therefore, the errors in channel estimation will directly affect the BER performance. One can see that the NMSEs of the channel estimation for MIMO is a little bit worse than that for SISO, which causes BER performance for MIMO to be a little bit worse than that for SISO under the same diversity order of 1. However, the BER performance can be improved if the STC or ML detection is applied.

Lastly, we illustrate the performance results of the proposed method with channel estimation scalar ambiguity. The total number of snapshots is taken as 300 such that each covariance matrix is still constructed with 150 samples. The

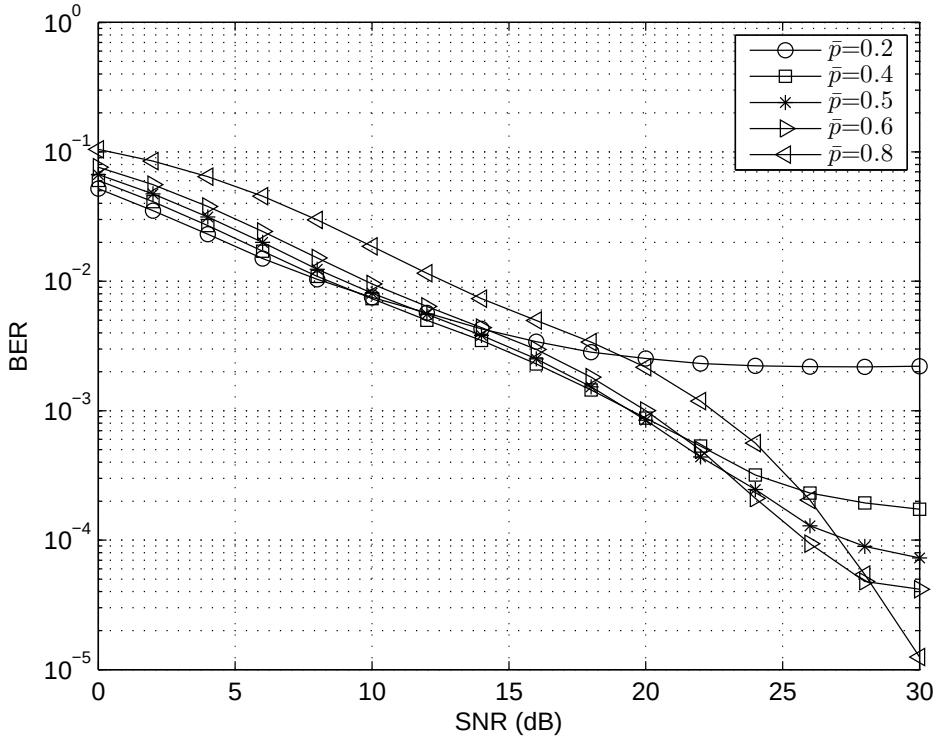

Figure 5.6: BERs for MIMO OFDM under different $\bar{p}$.

precoding matrix is taken as

$$[\mathbf{P}_{11}]_{ac} = [\mathbf{P}_{22}]_{ac} = \begin{cases} 1.2 & a = c = 1, \ldots, K/2 \\ 0.8 & a = c = K/2+1, \ldots, K \\ 2/3 & \text{otherwise,} \end{cases} \tag{5.80}$$

$$[\mathbf{P}_{12}]_{ac} = [\mathbf{P}_{21}]_{ac} = \begin{cases} 0.8 & a = c = 1, \ldots, K/2 \\ 1.2 & a = c = K/2+1, \ldots, K \\ 1/3 & \text{otherwise.} \end{cases} \tag{5.81}$$

The channel estimation NMSEs versus SNR for $\tilde{\mathbf{h}}_{11}$, $\tilde{\mathbf{h}}_{12}$, $\tilde{\mathbf{h}}_{21}$, $\tilde{\mathbf{h}}_{22}$ are shown in Fig. 5.7, separately. We see from the simulation results that the proposed method still works well for scalar ambiguity estimation.

3) MISO OFDM System

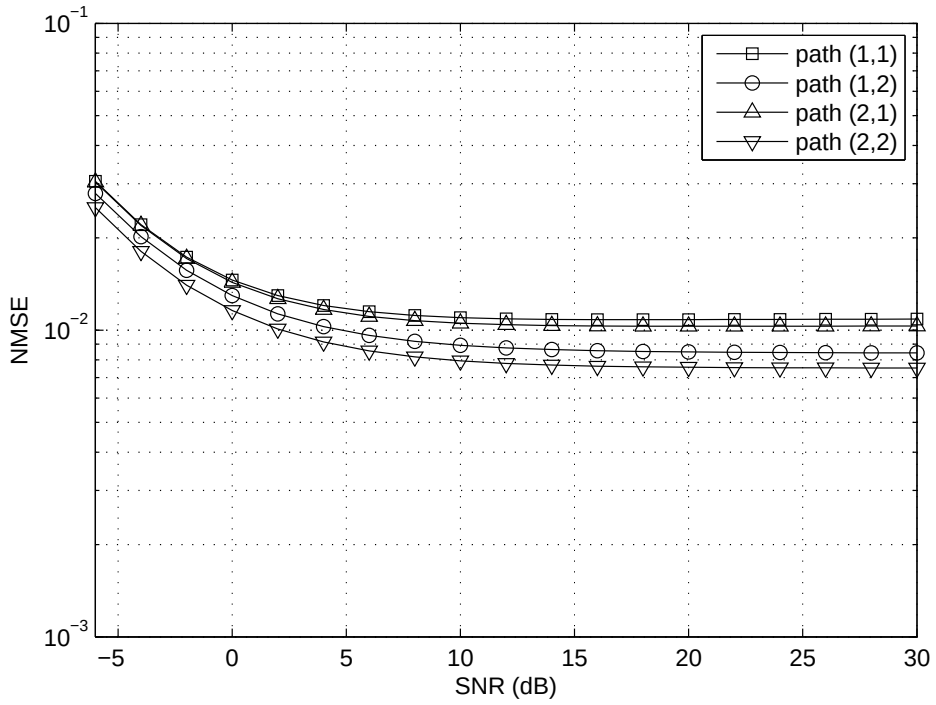

Figure 5.7: Performance NMSEs for MIMO OFDM versus SNR: with scalar ambiguity

In this example, we consider the performance results of the proposed algorithm for MISO OFDM system with two transmit and one receive antennas. Alamouti codes [98] are applied at the transmitter, and the channel estimation with matrix ambiguity is performed. We only illustrate the capability of the data detection of the proposed algorithm. The BER performance curves corresponding to different values of $\bar{p}$ are shown in Fig. 5.8. We see that the proposed method works well for the MISO case and the data detection is guaranteed by the application of Alamouti code at the transmitter. Comparing to Fig. 5.3 and Fig. 5.6, the diversity order is increased and BER curves drop in a much faster rate.

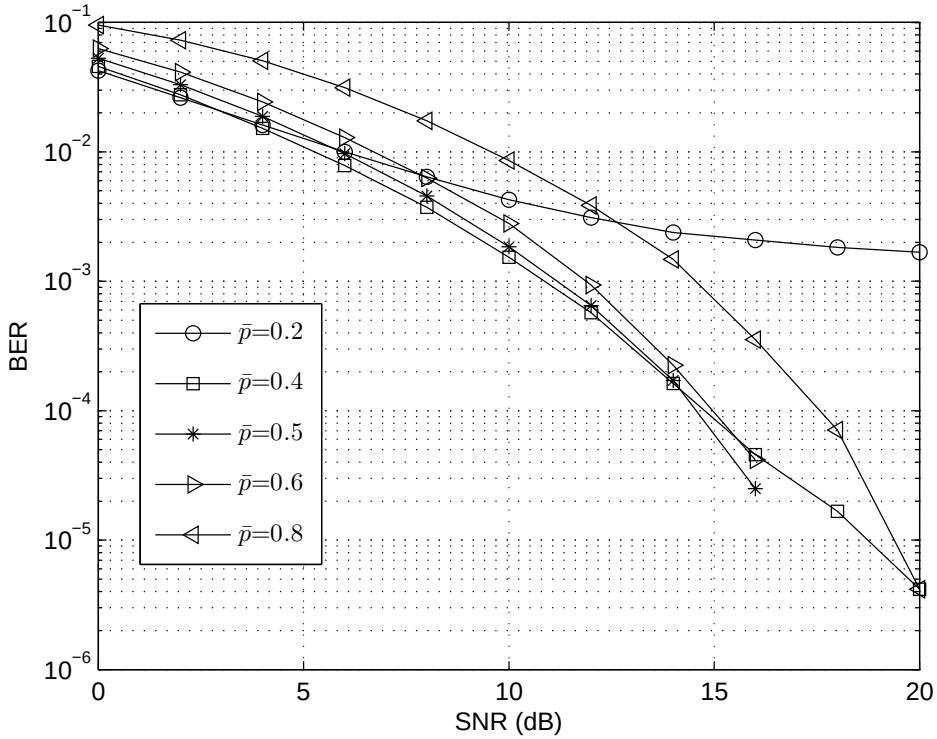

Figure 5.8: BERs for MISO OFDM with Alamouti code under different $\bar{p}$.

5.7 Summery

In this chapter, we developed a new blind channel estimation technique for SISO- and MIMO-OFDM systems based on the second order statistical analysis. One novel contribution of the newly proposed method is its capability under the scenario where the number of the transmit antennas is greater than or equal to the number of the receive antennas, in which case the traditional subspace based algorithm is not applicable. For SISO OFDM case, the proposed method is shown to perform better than the existing works. For MIMO OFDM case, the estimation methods with and without ambiguity are discussed. The trade off for eliminating the ambiguity is that more time slots during the channel coherent time are required to construct reliable covariance matrices. Finally, the stochastic CRB is derived and simulation results clearly show the effectiveness of the proposed method under various scenarios.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

The thesis provides solutions to several unsolved issues in synchronization and channel estimation for MIMO OFDM systems.

In terms of CFO tracking, we developed a new algorithm that could overcome the drawbacks of the existing methods and, meanwhile, exhibits many exciting properties, e.g., high estimation accuracy, full estimation range. The method is the first one that fully exploits all the resources in one OFDM blocks. We studied the identifiability of the tracking algorithm and showed that the ambiguity of CFO estimation could be removed by properly adjusting the system parameters. Nevertheless, with the exploitation on the virtual carriers, neither the ambiguity nor the outlier occurs frequently throughout the practical simulations. To be mentioned, the proposed frequency tracking algorithm is also applicable to TO tracking, although seldom needed during the data transmission, and to the synchronization initialization since the algorithm itself does not require the knowledge of the CSI and can provide the full range CFO estimation.

In terms of channel estimation, we proposed two new blind algorithms which fit the high speed transmission where the assumption that the channels are constant during each frame time becomes reasonably valid. We first developed a signal

re-modulated SS algorithm for CP-based MIMO OFDM with the following favorable properties: robustness to channel order over-estimation, guaranteeing the channel identifiability, and, most importantly, applicability to the scenario where the number of the receive antennas is equal to the number of the transmit antennas. The last property removes the most severe bottle neck on applying the traditional SS method to the MIMO OFDM system and is compatible to the coming 4G wireless communication systems. Afterwards, an even exciting blind channel estimation algorithm is developed that is applicable to the case where the number of the receive antennas is smaller than the number of the transmit antennas. A trade-off by doing this is the employment of a precoding matrix. This precoding matrix is not favorable when the ML estimation is required but is still feasible when the MMSE detection is adopted.

Various numerical results are provided to verify each of our studies.

6.2 Future Works

Recently, the relay based wireless transmission [93]-[97] attracted lots of attention due to its capability of enlarging the transmission range and enhancing the transmission diversity for size-limited terminals. In most relay networks, channel estimation is still an critical component for a reliable transmission. Training based channel estimation has been derived in [102], where it shows that the channel estimation in relay networks has fundamental differences from that in the traditional point-to-point communication systems. Therefore, it is of interest to rethink the blind channel estimation or the synchronization issues for the relay networks. All these works will be exploited in our future researches.

Bibliography

[1] R. W. Chang, "Synthesis of bandlimited orthogonal signals for multichannel data transmission," *Bell System Tech. J.*, vol. 45, pp. 1775-1796, Dec. 1966.

[2] S. B. Weinstein, and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. 19, pp. 628-634, Oct. 1971.

[3] A. Peled, and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," in *Proc. IEEE ICASSP'80*, Denver, CO, USA, vol. 5, pp. 964-967, Apr. 1980.

[4] "Radio broadcasing system, digital audio broadcasting (DAB) to mobile, portable, and fixed receiver," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETS 300 401, 1995-1997.

[5] "Digital broadcasting system television, sound, and data services; framing structure, channel coding, and modulation digital terrestrial television," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETS 300 744, 1996.

[6] "Broadband radio access networks (BRAN): high performance radio local area networks (HIPERLAN), type 2; Systems overview," Eur. Telecommun. Stand. Inst., Sophia-Antipolis, Valbonne, France, ETR 101 683 114, 1999.

[7] "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: high speed physical layer in the 5 GHZ band," IEEE802.11a, 1999.

[8] J. Lorincz and D. Begusic, "Physical layer analysis of emerging IEEE 802.11n WLAN standard," *Proc. ICACT'06*, Phoenix Park, Korea, vol. 1, pp. 189-194, Feb. 2006.

[9] TGn Sync Proposal Technical Specification IEEE 802.11-05/1095r2.

[10] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Mag.*, vol. 17, pp. 29-48, May 2000.

[11] J. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5-14, May 1990.

[12] W. Y. Zou and Y. Wu, "COFDM: an overview," *IEEE Trans. Broadcasting*, vol. 41, pp. 1-8, Mar. 1995.

[13] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Bell Labs Technical Memorandum*, 1995.

[14] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311-335, 1998.

[15] H. Bölcskei and E. Zurich, "MIMO-OFDM wireless systems: basics, perspectives, and challenges," *IEEE Wireless Commun.*, vol. 13, pp. 31-37, Aug. 2006.

[16] G. J. Foschini,"Layered space-time architecture for wireless communication in a fading environment when using multi-element antenna" *Bell Labs Tech. J.*, vol. 1, pp. 41-59, 1996.

[17] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 144-165, Mar. 1998.

[18] L. Zheng and D. N. C. Tse, "Diverstiy and multiplexing: a fundamental trade-off in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073-1096, May 2003.

[19] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm I: expected complexity," *IEEE Trans. Signal Processing*, vol. 53, pp. 2806-2818, Aug. 2005.

[20] H. Vikalo and B. Hassibi, "On the sphere-decoding algorithm II: generalizations, second-Order statistics, and applications to communications," *IEEE Trans. Signal Processing*, vol. 53, pp. 2819-2834, Aug. 2005.

[21] A. van Zelst and T. C. W. Schenk, "Implementation of a MIMO OFDM-Based Wireless LAN System," *IEEE Trans. Signal Processing*, vol. 52, pp. 483-494, Feb. 2004.

[22] Y. Yao and G. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, pp. 173-183, Jan. 2005.

[23] M. Minn, N. Al-Dhahir, and Y. Li, "Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise," *IEEE Trans. Commun.*, vol. 54, pp. 1754-1759, Oct. 2006.

[24] M. Sandell, D. McNamara, and S. Parker, "Analysis of frequency-offset tracking in MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 54, pp. 1481-1489, Aug. 2006.

[25] M. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1158-168, May 2006.

[26] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613-p1621, Dec. 1997.

[27] A. N. Mody and G. L. Stüber, "Synchronization for MIMO OFDM systems," in *Proc. GLOBECOM'01*, vol. 1, Nov. 2001, pp. 509-513.

[28] A. Hajimiri and T. H. Lee, *The Design of Low Noise Oscillators*, Norwell, MA: Kluwer, 1999.

[29] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, pp. 100-109, Feb. 1995.

[30] Y. Li, L. J. Cimini Jr, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902-915, July 1998.

[31] J. H. Manton, "Optimal training sequences and pilot tones for OFDM systems," *IEEE Commun. Lett.*, vol 5, pp. 151-153, Apr. 2001.

[32] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[33] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 67-75, Jan. 2002.

[34] H. Minn, D. I. Kim and V. K. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Trans. Commun.*, vol. 50, pp. 799-807, May 2002.

[35] D. Brillinger, "The identification of polynomial systems by means of higher-order spectra," *J. Sound Vib.*, vol. 20, pp. 301-313, 1970.

[36] G. Giannakis and J. Mendel, "Identification of nonminimum phase systems using higher order statistics," *IEEE Trans. Acout., Speech, Signal Processing*, vol. 37, pp. 360-377, Mar. 1989.

[37] D. Hatzinakos and C. Nikias, "Estimation of multipath channel response in frequency selective channels," *IEEE J. Select. Aeas Commun.*, vol. 7, pp. 12-19, Jan. 1989.

[38] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Trans. Signal Processing*, vol. 39, pp. 522-526, Feb. 1991.

[39] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf.*, Pacific Grove, USA, vol. 2, pp. 856-860, Nov. 1991.

[40] E. Moulines, P. Duhamel, J.F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516-525, Feb. 1995.

[41] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677-690, Mar. 1998.

[42] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cylic prefixing or zero padding for wireless multicarrier Transmissions," *IEEE Trans. Commun.*, vol. 50, pp. 2136-2148, Dec. 2002.

[43] C. Shin and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems using virtual carriers," in *Proc. GLOBECOM '04*, Dallas, USA, vol. 4, pp. 2465-2469, Nov. 2004.

[44] H. Liu and G. Xu, "A subspace method for signature waveform estimation in synchronous CDMA system," *IEEE Trans. Commun.*, vol. 44, pp. 1346-1354, Oct. 1996.

[45] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 572-579, Oct. 2002.

[46] V. Buchoux, O. Cappe, E. Moulines, and A. Gorokhov, "On the performance of semi-blind subspace-based channel estimation," *Trans. Signal Processing*, vol. 48, pp. 1750-1759, June 2000.

[47] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Processing*, vol. 50, pp. 1699-1712, July 2002.

[48] S. Zhou, B. Muquet, and G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for bloc kprecoded space-time OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 1215-1228, May 2002.

[49] F. Gao, A. Nallanathan, and C. Tellambura, "Blind channel estimation for cyclic prefixed single-carrier systems exploiting real symbol characteristics," *IEEE Trans. Veh. Technol.*, vol. 56, Sept. 2007.

[50] J.-J. van de Beek, O. Edfors, M. Sandell, S.K. Wilson, and P.O. Borjesson, "On channel estimation in OFDM system," in *Proc. VTC'95*, Chicago, USA, vol. 2, pp. 815-819, July 1995.

[51] G. Santella, "Frequency and symbol synchronization system of OFDM signals: architecture and simulation results," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 254-275, Jan. 2000.

[52] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE Commun. Lett.*, vol. 2, pp. 104-106, Apr. 1998.

[53] B. Chen, "Maximum likelihood estimation of OFDM carrier frequency offset," *IEEE Signal Processing Lett.*, vol. 9, pp. 123-126, Apr. 2002.

[54] F. Gao and A. Nallanathan, "Blind maximum likelihood CFO estimation for OFDM systems via polynomial rooting," *IEEE Signal Processing Lett.*, vol.13, pp. 73-76, Feb. 2006.

[55] Y. Zeng and T. S. Ng, "A semi-blind channel estimation method for multiuser multiantenna OFDM systems," *IEEE Trans. Signal Processing*, vol. 52, pp. 1419-1429, May, 2004.

[56] F. Classen and H. Meyr, "Frequency synchronization algorithms for OFDM systems suitable for communication over frequency selective fading channels," in *Proc. VTC'94*, Stockholm, Sweden, vol. 3, pp. 1655-1659, June 1994.

[57] J.-J van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Processing, vol. 45, pp. 1800-1805*, July 1997.

[58] H. Chen, and G. J. Pottie, "A comparison of frequency offset tracking algorithms for OFDM," in *Proc. GLOBECOM'03*, San Francisco, USA, pp. 1069-1073, Dec. 2003.

[59] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: identifiability, algorithms, and performance," *IEEE J. Sel. Areas Commun.*, vol. 19, pp. 2504-2515, Dec. 2001.

[60] T. Pollet, M. van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191-193, Feb. 1995.

[61] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908-2914, Oct. 1994.

[62] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, pp. 1580-1589, Sept. 2000.

[63] H. Minn, V. K. Bhargava, and K. B. Letaief, "A robut timing and frequency synchronization for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 822-839, July 2003.

[64] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.*, vol. 3, pp. 75-77, Mar. 1999.

[65] H. Minn and P. Tarasak, "Improved maximum likelihood frequency estimation based on likelihood metric design," in *Proc. ICC'05*, Seoul, Korea, vol. 4, pp. 2150-2156, May 2005.

[66] H. Minn, P. Tarasak, and V. K. Bhargava, "OFDM frequency offset estimation based on BLUE principle," in *Proc. VTC'02*, Vancouver, Canada, vol. 2, pp. 1230-1234, Sept. 2002.

[67] H. Minn, X. Fu, and V. K. Bhargava, "Optimal periodic training signal for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 54, pp. 1081-1096, June 2006.

[68] H. Minn, Y. Li, N. Al-Dhahir, and R. Calderbank, "Pilot designs for consistent frequency offset estimation in OFDM systems," in *Proc. ICC'06*, Istanbul, Turkey, vol. 10, pp. 4566-4571, June 2006.

[69] Y. Li, H. Minn, N. Al-Dhahir, and R. Calderbank, "Robust pilot design for consistent carrier frequency offset estimation," in *Proc. MILCOM'06*, Washington, DC, pp. 1-7, Oct. 2006.

[70] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, pp. 1122-1128, Aug. 1998.

[71] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: futher unification, simplification, and observations," *IEEE Trans. Aerosp., Electron. Syst.*, vol. 29, pp. 1170-1184, Oct. 1993.

[72] J. Li, G. Liu, and G. B. Giannakis, "Carrier frequency offset estimation for OFDM based WLANs," *IEEE Signal Processing Lett.*, vol. 8, pp. 80-82, Mar. 2001.

[73] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, pp. 3065-3073, Dec. 2001.

[74] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results," *IEEE Commun. Mag.*, pp. 143-149, Sept. 2002.

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

[76] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1615-1624, June, 2003.

[77] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1158-1168, May 2006.

[78] E. de Carvalho and D. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Trans. Signal Processing,* vol. 52, pp. 1053-1064, Apr. 2004.

[79] A. Medles and D. Slock, "Linear precoding for spatial multiplexing MIMO systems: blind channel estimation aspects," in *Proc. IEEE ICC'02*, New York, USA, vol. 1, pp. 401-405, 28 April-2 May 2002.

[80] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problem," *IEEE Trans. Inform. Theory*, vol. 32, pp. 499-511, Apr. 1997. 2000.

[81] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, pp. 1951-1968, Oct. 1998.

[82] E. de Carvalho, J. Cioffi, and D. Slock, "Cramér-rao bounds for blind multichannel estimation," in *Proc. GLOBECOM'00*, San Francisco, USA, vol. 2, pp. 1036-1040, Nov.

[83] H. Gazzah, P. A. Regalia, J. P. Delmas, and K. Abed-Meraim, "A blind multichannel identification algorithm robust to order overestimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1449-1458, June 2002.

[84] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless channel fades?" in *Proc. SPAWC'01*, pp. 267-270, Mar. 2001.

[85] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Trans. Commun.*, vol. 52, pp. 280-394, Mar. 2004.

[86] R. Zhang, "Blind OFDM channel estimation through linear precoding: a subspace approach," in *Proc. 36th Asilomar Conf.*, Pacific Grove, USA, vol. 1, pp. 631-633, Nov. 2002.

[87] A. P. Petropulu and R. Zhang, "Blind channel estimation for OFDM systems," in *Proc. DSP/SPE'02*, Georgia, USA, pp. 366-370, Oct. 2002.

[88] A. P. Petropulu, R. Zhang, R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. Wilress Commun.*, vol. 3, pp. 647-655. Mar. 2004.

[89] R. Lin and A. P. Petropulu, "Linear block precoding for blind channel estimation in OFDM systems," in *Proc. ISSPA'03*, Paris, France, vol. 2, pp. 395-398. July 2003.

[90] R. Lin and A. P. Petropulu, "Blind channel estimation for OFDM systems based on non-redundant linear precoding ," in *Proc. Statist. Processing*, St. Louis, USA, pp. 351-354, Oct. 2003.

[91] R. Lin and A. P. Petropulu, "Linear precoding assisted blind channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 983-995, May 2005.

[92] M. Biguesh and A. B. Gershman, "MIMO channel estimation: optimal training and tradeoffs between estimation techniques," in *Proc. ICC'04*, Paris, France, vol. 5, pp. 2658-2662, June 2004.

[93] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 572-584, Sept. 1979.

[94] J. N. Laneman and G. W. Wornell, "Distributed space time block coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415-2425, Oct. 2003.

[95] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062-3080, Dec. 2004.

[96] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: system description," *IEEE Trans. Commun.*, vol. 51, pp. 1927-1938, Nov. 2003.

[97] Y. Jing and B. Hassibi, "Distributed space time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 3524-3536, Dec. 2006.

[98] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[99] P. Stoica, E. G. Larsson, and A.B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Lett.*, vol. 8, pp. 148-150, May 2001.

[100] J-P. Delmas and H. Abeida, "Stochastic cramér-rao bound for noncircular signals with application to DOA estimation," *IEEE Trans. Signal Processing*, vol. 52, pp. 3192-3199, Nov. 2004.

[101] R. W. Heath and D.J. Love, "Multimode antenna selection for spatial multiplexing systems with linear receivers," *IEEE Trans. Signal Processing*, vol. 53, pp. 3042-3056, Aug. 2005.

[102] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay network", to appear *IEEE Trans. Wireless Commun.*

List of Publications

Journal Papers (published)

1. F. Gao, A. Nallanathan, and C. Tellambura "Blind Channel Estimation for Cyclic Prefixed Single-Carrier Systems Exploiting Real Symbol Characteristics", IEEE Transactions on Vehicular Technology, vol. 56, pp. 2487-2498, Sept. 2007.

2. F. Gao and A. Nallanathan, "Blind Channel Estimation for OFDM Systems via A Generalized Precoding", IEEE Transactions on Vehicular Technology, vol. 56, pp. 1155-1164, May 2007.

3. F. Gao and A. Nallanathan, "Reply to "A Comment on 'Blind Maximum Likelihood CFO Estimation for OFDM Systems via Polynomial Rooting'"", IEEE Signal Processing Letters, vol. 14, pp. 292-292, Apr. 2007.

4. F. Gao and A. Nallanathan, "Blind Channel Estimation for MIMO OFDM Systems via Non-Redundant Linear Precoding", IEEE Transactions on Signal Processing, vol. 55, pp. 784-789, Jan. 2007.

5. F. Gao and A. Nallanathan, "Identifiability of Data-Aided Carrier Frequency Offset Estimation over Frequency Selective Channels", IEEE Transactions on Signal Processing, vol. 54, pp. 3653-3657, Sept. 2006.

6. F. Gao and A. Nallanathan, "Blind Maximum Likelihood CFO Estimation for OFDM Systems via Polynomial Rooting", IEEE Signal Processing Letters, vol. 13, pp. 73-76, February 2006.

Journal Papers (accepted)

7. F. Gao, T. Cui, and A. Nallanathan, "Scattered Pilots and Virtual Carriers Based Frequency Offset Tracking for OFDM Systems: Algorithms, Identifiability, and Performance Analysis", IEEE Transactions on Communications.

8. F. Gao, Tao Cui, A. Nallanathan, and C. Tellambura, "ML based Estimation of Frequency and Phase Offsets in DCT OFDM Systems via Non-Circular Sources: Algorithms, Analysis and Comparisons", IEEE Transactions on Communications.

9. F. Gao and A. Nallanathan, "Resolving Multi-Dimensional Ambiguity in Blind Channel Estimation of MIMO-FIR Systems via Block Precoding", IEEE Transactions on Vehicular Technology.

10. F. Gao, T. Cui, and A. Nallanathan, "On Channel Estimation and Optimal Training Design for Amplify and Forward Relay Network", IEEE Transactions on Wireless Communications.

11. F. Gao, Y. Wang, and A. Nallanathan, "Improved MUSIC under The Coexistence of Both Non-Circular and Circular Sources", IEEE Transactions on Signal Processing.

12. F. Gao, T. Cui, A. Nallanathan, and C. Tellambura,, "Maximum Likelihood Detection for Differential Unitary Space-Time Modulation with Carrier Frequency Offset", IEEE Transactions on Communications .

13. F. Gao, Y. Zeng, A. Nallanathan, and T. S. Ng, "Robust Subspace Blind Channel Estimation for Cyclic Prefixed MIMO OFDM Systems: Algorithm, Identifiability and Performance Analysis", IEEE Journal On Selected Areas in Communications.

Conference Papers (published)

1. F. Gao and A. Nallanathan, "Blind Channel Estimation for OFDM Systems via a General Non-Redundant Precoding", in Proc. of IEEE ICC'06, Istanbul, Turkey, vol. 10, pp. 4612-4617, June 2006.

2. F. Gao and A. Nallanathan, "Subspace-Based Blind Channel Estimation for MISO and MIMO OFDM Systems", in Proc. of IEEE ICC'06, Istanbul, Turkey, vol. 7, pp. 3025-3030, June 2006.

3. F. Gao and A. Nallanathan, "Identifiability of Training Based CFO Estimation over Frequency Selective Channels", in Proc. of IEEE ICC'06, Istanbul, Turkey, vol.3, pp. 1421-1426, June 2006.

4. F. Gao and A. Nallanathan, "A Novel Subspace-Based Blind Channel Estimation for Cyclic Prefixed Single-Carrier Transmissions", in Proc. of IEEE WCNC'06, Las Vegas, U.S.A, vol. 3, pp. 1537-1542, vol. 3, pp. 1515-1518, Apr. 2006.

5. F. Gao and A. Nallanathan, "Polynomial Rooting Based Maximum Likelihood Carrier Frequency Offset Estimation for OFDM Systems", in Proc. of IEEE WCNC'06, Las Vegas, U.S.A, vol.2, pp. 1046-1049, Apr. 2006.

6. F. Gao and A. Nallanathan, "A Simple Subspace-Based Blind Channel Estimation for OFDM Systems", in Proc. of IEEE WCNC'06, Las Vegas, U.S.A, vol. 3, pp. 1515-1518, Apr. 2006.

7. F. Gao, and A. Nallanathan, "Improved MUSIC by Exploiting Both Real and Complex Sources", in Proc. of IEEE MILCOM'06, Washington DC, Oct. 2006.

Conference Papers (accepted)

8. T. Cui, F. Gao, and A. Nallanathan, "Optimal Training Design for Channel Estimation in Amplify and Forward Relay Networks", submitted to Proc. of IEEE GLOBECOM'07.

9. T. Cui, F. Gao, A. Nallanathan, and C. Tellambura, "Maximum Likelihood Detection and Optimal Code Design for Differential Unitary Space-Time Modulation with Carrier Frequency Offset", submitted to Proc. of IEEE GLOBECOM'07.

10. T. Cui, F. Gao, and A. Nallanathan, "Frequency Offset Tracking for OFDM Systems via Scattered Pilots and Virtual Carriers", in Proc. of IEEE ICC'07.

11. F. Gao, W. Wu, Y. Zeng and A. Nallanathan, "A Novel Blind Channel Estimation for CP-Based MIMO OFDM Systems", in Proc. of IEEE ICC'07.

12. T. Cui, F. Gao, A. Nallanathan, and C. Tellambura, "ML CFO and PO Estimation in DCT OFDM Systems under Non-Circular Transmissions", in Proc. of IEEE ICC'07.

13. F. Gao, and A. Nallanathan, "Higher-Dimensional Ambiguity Free Blind Channel Estimation for MIMO-FIR Systems via Linear Block Precoding", in Proc. of IEEE GLOBECOM'06.

Appendix A

Error Evaluation for CFO Tracking

Define new vectors

$$\breve{\mathbf{y}}_i = e^{-j2\pi\phi((i-1)K_s+P)}\mathbf{\Omega}(-\phi)\mathbf{y}_i = \underbrace{\mathbf{FHs}_i}_{\boldsymbol{\eta}_i} + \underbrace{e^{-j2\pi\phi(i-1)K_s}\mathbf{\Omega}(-\phi)\mathbf{n}_i}_{\breve{\mathbf{n}}_i} \tag{A.1}$$

$$\breve{\mathbf{y}}_{i+1} = e^{-j2\pi\phi(iK_s+P)}\mathbf{\Omega}(-\phi)\mathbf{y}_{i+1} = \underbrace{\mathbf{FHs}_{i+1}}_{\boldsymbol{\eta}_{i+1}} + \underbrace{e^{-j2\pi\phi iK_s}\mathbf{\Omega}(-\phi)\mathbf{n}_{i+1}}_{\breve{\mathbf{n}}_{i+1}}. \tag{A.2}$$

Obviously, $\breve{\mathbf{n}}_i$ has the same distribution as that of $\mathbf{n}_i$. Then $g_p(\varepsilon)$ is rewritten as

$$\begin{aligned} g_p(\varepsilon) &= \|\mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_i - \mathbf{F}_p^H\mathbf{\Omega}(-\varepsilon)\mathbf{y}_{i+1}e^{-j2\pi\varepsilon K_s}\|^2 \\ &= (\mathbf{y}_i^H - \mathbf{y}_{i+1}^H e^{j2\pi\varepsilon K_s})\mathbf{\Omega}(\varepsilon)\mathbf{P}_p\mathbf{\Omega}(-\varepsilon)\big(\mathbf{y}_i - \mathbf{y}_{i+1}e^{-j2\pi\varepsilon K_s}\big) \end{aligned} \tag{A.3}$$

where $\mathbf{P}_p = \mathbf{F}_p\mathbf{F}_p^H$ is the projection matrix onto the subspace spanned by $\mathbf{F}_p$. Bearing in mind that $\mathbf{F}_p^H\big(\boldsymbol{\eta}_i - \boldsymbol{\eta}_{i+1}\big) = \mathbf{0}$, $\dot{g}_p(\varepsilon)\,|_{\varepsilon=\phi}$ can be obtained as

$$\begin{aligned} \dot{g}_p(\phi) = {} & j2\pi\Delta\boldsymbol{\eta}\mathbf{D}\mathbf{P}_p\Delta\breve{\mathbf{n}} + j2\pi\Delta\breve{\mathbf{n}}^H\mathbf{D}\mathbf{P}_p\Delta\breve{\mathbf{n}} - j2\pi\Delta\breve{\mathbf{n}}^H\mathbf{P}_p\mathbf{D}\Delta\boldsymbol{\eta} - j2\pi\Delta\breve{\mathbf{n}}^H\mathbf{P}_p\mathbf{D}\Delta\breve{\mathbf{n}} \\ & - j2\pi K_s\boldsymbol{\eta}_{i+1}^H\mathbf{P}_p\Delta\breve{\mathbf{n}} - j2\pi K_s\breve{\mathbf{n}}_{i+1}^H\mathbf{P}_p\Delta\breve{\mathbf{n}} \\ & + j2\pi K_s\Delta\breve{\mathbf{n}}^H\mathbf{P}_p\boldsymbol{\eta}_{i+1} + j2\pi K_s\Delta\breve{\mathbf{n}}^H\mathbf{P}_p\breve{\mathbf{n}}_{i+1} \end{aligned} \tag{A.4}$$

where $\mathbf{D} \triangleq \text{diag}\{0, 1, ..., K-1\}$, $\Delta\boldsymbol{\eta} \triangleq \boldsymbol{\eta}_i - \boldsymbol{\eta}_{i+1}$ and $\Delta\breve{\mathbf{n}} \triangleq \breve{\mathbf{n}}_i - \breve{\mathbf{n}}_{i+1}$ are used for notation simplicity. The expectation of $\dot{g}_p(\phi)$ is

$$\begin{aligned}\mathrm{E}\{\dot{g}_p(\phi)\} &= (0) + j4\pi\sigma^2 tr(\mathbf{D}\mathbf{P}_p) - (0) - j4\pi\sigma^2 tr(\mathbf{P}_p\mathbf{D}) \\ &\quad + (0) + j2\pi K_s\sigma^2 tr(\mathbf{P}_p) + (0) - j2\pi K_s\sigma^2 tr(\mathbf{P}_p) = 0. \end{aligned} \tag{A.5}$$

After some manipulations, $\mathrm{E}\{(\dot{g}_p(\phi))^2\}$ and $\mathrm{E}\{\ddot{g}_p(\phi)\}$ can be obtained as

$$\mathrm{E}\{(\dot{g}_p(\phi))^2\} = 16\pi^2\sigma^2\|\mathbf{P}_p(\mathbf{D}\Delta\boldsymbol{\eta} - K_s\boldsymbol{\eta}_{i+1})\|^2 \tag{A.6}$$

$$\mathrm{E}\{\ddot{g}_p(\phi)\} = 8\pi^2\|\mathbf{P}_p(\mathbf{D}\Delta\boldsymbol{\eta} - K_s\boldsymbol{\eta}_{i+1})\|^2. \tag{A.7}$$

On the other hand, $g_v(\varepsilon)$ can be rewritten as

$$g_v(\varepsilon) = \sum_{m=i}^{i+1}\|\mathbf{F}_v^H\boldsymbol{\Omega}(-\varepsilon)\mathbf{y}_m\|^2 = \sum_{m=i}^{i+1}\mathbf{y}_m^H\boldsymbol{\Omega}(\varepsilon)\mathbf{P}_v\boldsymbol{\Omega}(-\varepsilon)\mathbf{y}_m \tag{A.8}$$

where $\mathbf{P}_v = \mathbf{F}_v\mathbf{F}_v^H$ is the projection matrix onto the subspace spanned by $\mathbf{F}_v$. Bearing in mind that $\mathbf{F}_v^H\boldsymbol{\eta}_i = \mathbf{0}$, $\dot{g}_v(\varepsilon)\,|_{\varepsilon=\phi}$ can be obtained as

$$\dot{g}_v(\phi) = j2\pi\sum_{m=i}^{i+1}\left(\boldsymbol{\eta}_m^H\mathbf{D}\mathbf{P}_v\breve{\mathbf{n}}_m + \breve{\mathbf{n}}_m^H\mathbf{D}\mathbf{P}_v\breve{\mathbf{n}}_m - \breve{\mathbf{n}}_q^H\mathbf{P}_v\mathbf{D}\boldsymbol{\eta}_m - \breve{\mathbf{n}}_q^H\mathbf{P}_v\mathbf{D}\breve{\mathbf{n}}_m\right). \tag{A.9}$$

It can be calculated that

$$\mathrm{E}\{\dot{g}_v(\phi)\} = j2\pi\sigma^2\sum_{m=i}^{i+1}\left(0 + tr(\mathbf{D}\mathbf{P}_v) + 0 - tr(\mathbf{P}_v\mathbf{D})\right) = 0. \tag{A.10}$$

Furthermore, $\mathrm{E}\{(\dot{g}_v(\phi))^2\}$ and $\mathrm{E}\{\ddot{g}_v(\phi)\}$ can be obtained as

$$\mathrm{E}\{(\dot{g}_v(\phi))^2\} = 8\pi^2\sigma^2\sum_{m=i}^{i+1}\boldsymbol{\eta}_m^H\mathbf{D}\mathbf{P}_v\mathbf{D}\boldsymbol{\eta}_m \tag{A.11}$$

$$\mathrm{E}\{\ddot{g}_v(\phi)\} = 8\pi^2\sum_{m=i}^{i+1}\boldsymbol{\eta}_m^H\mathbf{D}\mathbf{P}_v\mathbf{D}\boldsymbol{\eta}_m. \tag{A.12}$$

Lastly, we derive the expectation of $\dot{g}_p(\phi)\dot{g}_v(\phi)$ as

$$\mathrm{E}\{\dot{g}_p(\phi)\dot{g}_v(\phi)\} = o(\mathbf{n}_i^4) + o(\mathbf{n}_{i+1}^4) \tag{A.13}$$

where the property $\mathbf{P}_p^H\mathbf{P}_v = \mathbf{0}$ is used, and $o(\mathbf{n}_i^4)$ denotes the function at the order of $\mathbf{n}_i^4$. This term can be ignored at higher SNR compared to $\mathrm{E}\{\dot{g}_p(\phi)^2\}$ and $\mathrm{E}\{\dot{g}_v(\phi)^2\}$.

Therefore, the p-algorithm and v-algorithm can be considered as uncorrelated to each other.

Finally, substituting (A.5), (A.6), (A.7), (A.10), (A.11), (A.12) into (3.23), (3.24) yields (3.25), (3.26).

Appendix B

Channel MSE for Remodulated SS Algorithm

Firstly, we introduce the lemma provided in [71].

Lemma B.1 [71]: Denote the singular value decomposition (SVD) of

$$\acute{\mathbf{Z}} = \boldsymbol{\mathcal{A}}[\mathbf{d}_1, ..., \mathbf{d}_M] = \boldsymbol{\mathcal{A}}\mathbf{D} \tag{B.1}$$

as

$$\acute{\mathbf{Z}} = [\mathbf{U}_s \quad \mathbf{U}_o] \begin{bmatrix} \boldsymbol{\Delta}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^H \\ \mathbf{V}_o^H \end{bmatrix}. \tag{B.2}$$

The first order approximation of the perturbation to $\mathbf{U}_o$ due to the additive noise $\acute{\mathbf{W}} = [\breve{\mathbf{w}}_1, ..., \breve{\mathbf{w}}_M]$ is

$$\Delta\mathbf{U}_o = -\mathbf{U}_s\boldsymbol{\Delta}_s^{-1}\mathbf{V}_s^H\acute{\mathbf{W}}^H\mathbf{U}_o = -(\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{W}}^H\mathbf{U}_o. \tag{B.3}$$

Ideally, the channel matrix $\mathbf{H}$ is obtained from

$$\boldsymbol{\mathcal{K}}^H\mathbf{H} = \mathbf{0}. \tag{B.4}$$

However, we may only be able to obtain an orthonormal matrix $\hat{\mathbf{H}}$ from the left singular vectors of $\boldsymbol{\mathcal{K}}$. Therefore, $\mathbf{H}$ is expressed as $\mathbf{H} = \hat{\mathbf{H}}\mathbf{B}^{-1}$ for an unknown $\mathbf{B}$. By applying Lemma B.1 again, the perturbation of the channel estimate $\hat{\mathbf{H}}$ is

$$\Delta\hat{\mathbf{H}} = -(\boldsymbol{\mathcal{K}}^H)^\dagger\Delta\boldsymbol{\mathcal{K}}^H\hat{\mathbf{H}} \tag{B.5}$$

where

$$\begin{aligned}\Delta\boldsymbol{\mathcal{K}} &= [\mathbf{C}_1^H\mathbf{R}_w^{-1/2}\Delta\mathbf{U}_o,\ldots,\mathbf{C}_K^H\mathbf{R}_w^{-1/2}\Delta\mathbf{U}_o] \\ &= -[\mathbf{C}_1^H\mathbf{R}_w^{-1/2}(\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{W}}^H\mathbf{U}_o,\ldots,\mathbf{C}_K^H\mathbf{R}_w^{-1/2}(\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{W}}^H\mathbf{U}_o]. \end{aligned} \tag{B.6}$$

It then follows

$$\Delta\mathbf{H} = \Delta\hat{\mathbf{H}}\mathbf{B}^{-1} = -(\boldsymbol{\mathcal{K}}^H)^\dagger\Delta\boldsymbol{\mathcal{K}}^H\hat{\mathbf{H}}\mathbf{B}^{-1} = -(\boldsymbol{\mathcal{K}}^H)^\dagger\Delta\boldsymbol{\mathcal{K}}^H\mathbf{H}. \tag{B.7}$$

Note that (B.7) could not be directly derived from Lemma B.1 since Lemma B.1 is only applicable for perturbation in the eigen-space. Obviously,

$$\mathrm{E}\{\Delta\mathbf{H}\} = -(\boldsymbol{\mathcal{K}}^H)^\dagger\mathrm{E}\{\Delta\boldsymbol{\mathcal{K}}^H\}\mathbf{H} = \mathbf{0}. \tag{B.8}$$

From [45], we know that $\mathrm{E}\{\acute{\mathbf{W}}\mathbf{Q}\acute{\mathbf{W}}^H\} = \sigma_n^2 tr(\mathbf{Q})\mathbf{I}$. Therefore,

$$\begin{aligned}&\mathrm{E}\{\acute{\mathbf{W}}\acute{\mathbf{Z}}^\dagger\mathbf{R}_w^{-1/2}\mathbf{C}_m\mathbf{H}(:,p)\mathbf{H}^H(:,a)\mathbf{C}_b^H\mathbf{R}_w^{-1/2}(\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{W}}^H\} \\ =&\sigma_n^2 tr(\acute{\mathbf{Z}}^\dagger\mathbf{R}_w^{-1/2}\mathbf{C}_m\mathbf{H}(:,p)\mathbf{H}^H(:,a)\mathbf{C}_b^H\mathbf{R}_w^{-1/2}(\acute{\mathbf{Z}}^H)^\dagger)\mathbf{I} \\ =&\sigma_n^2 tr(\boldsymbol{\mathcal{A}}_{b,a}^H(\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{Z}}^\dagger\boldsymbol{\mathcal{A}}_{m,p})\mathbf{I} = \sigma_n^2 tr(\boldsymbol{\mathcal{A}}_{b,a}^H(\acute{\mathbf{Z}}\acute{\mathbf{Z}}^H)^\dagger\boldsymbol{\mathcal{A}}_{m,p})\mathbf{I} \\ =&\sigma_n^2 tr(\boldsymbol{\mathcal{A}}_{b,a}^H(\boldsymbol{\mathcal{A}}^H)^\dagger(\mathbf{D}\mathbf{D}^H)^{-1}\boldsymbol{\mathcal{A}}^\dagger\boldsymbol{\mathcal{A}}_{m,p})\mathbf{I} \\ =&\sigma_n^2 tr(\mathbf{e}_{(b-1)N_t+a}^H(\mathbf{D}\mathbf{D}^H)^{-1}\mathbf{e}_{(m-1)N_t+p})\mathbf{I} \end{aligned} \tag{B.9}$$

where $\boldsymbol{\mathcal{A}}_{m,p}$ is the $((m-1)N_t+p)$th column of $\boldsymbol{\mathcal{A}}$, and the property $(\acute{\mathbf{Z}}\acute{\mathbf{Z}}^H)^\dagger = (\acute{\mathbf{Z}}^H)^\dagger\acute{\mathbf{Z}}^\dagger$ is used. The term $\mathbf{D}\mathbf{D}^H$ is in fact the estimated signal covariance matrix $2\sigma_s^2 M\mathbf{I}$ for the asymptotically large M. Therefore, equation (B.9) can be well approximated by

$$\sigma_n^2 tr\left(\mathbf{e}_{(b-1)N_t+a}^H(\mathbf{D}\mathbf{D}^H)^{-1}\mathbf{e}_{(m-1)N_t+p}\right)\mathbf{I} = \frac{\sigma_n^2}{2M\sigma_s^2}\delta_{m-b}\delta_{a-p}\mathbf{I}. \tag{B.10}$$

Finally, the channel error covariance matrix can be obtained as

$$\begin{aligned}\mathrm{E}\{\mathrm{vec}(\Delta\mathbf{H})\mathrm{vec}^H(\Delta\mathbf{H})\} &= (\boldsymbol{\mathcal{K}}^H)^\dagger\mathrm{E}\{\Delta\boldsymbol{\mathcal{K}}^H\mathrm{vec}(\mathbf{H})\mathrm{vec}^H(\mathbf{H})\Delta\boldsymbol{\mathcal{K}}\}\boldsymbol{\mathcal{K}}^\dagger \\ &= \mathbf{I}_K\otimes\left(\frac{\sigma_n^2}{2M\sigma_s^2}(\boldsymbol{\mathcal{K}}^H)^\dagger\mathbf{U}_o^H\mathbf{U}_o\boldsymbol{\mathcal{K}}^\dagger\right) \\ &= \mathbf{I}_{N_t}\otimes\left(\frac{\sigma_n^2(\boldsymbol{\mathcal{K}}^H)^\dagger\boldsymbol{\mathcal{K}}^\dagger}{2M\sigma_s^2}\right). \end{aligned} \tag{B.11}$$

Appendix C

Deterministic CRB for Remodulated SS Algorithm

From approximation (4.33), it suffices to first consider $\mathbf{z}_i$, and the unknown parameters changes to $\boldsymbol{\theta} = [\text{vec}(\mathbf{H}), \mathbf{d}_i, \sigma_n^2]$. The exact FIM for $\boldsymbol{\vartheta} = [\text{vec}(\mathbf{H}), \mathbf{d}_i]$ can be expressed as [82]

$$\mathbf{J} = \frac{1}{\sigma_n^2}\boldsymbol{\Gamma}^H \mathbf{R}_w^{-1}\boldsymbol{\Gamma} \tag{C.1}$$

where

$$\boldsymbol{\Gamma} = \left[\frac{\partial(\boldsymbol{\mathcal{A}}\mathbf{d}_i)}{\partial \text{vec}(\mathbf{H})}, \frac{\partial(\boldsymbol{\mathcal{A}}\mathbf{d}_i)}{\partial \mathbf{d}_i}\right]. \tag{C.2}$$

It can be obtained straightforwardly that

$$\frac{\partial(\boldsymbol{\mathcal{A}}\mathbf{d}_i)}{\partial \text{vec}(\mathbf{H})} = \mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{D}}_i \tag{C.3}$$

$$\frac{\partial(\boldsymbol{\mathcal{A}}\mathbf{d}_i)}{\partial \mathbf{d}_i} = \boldsymbol{\mathcal{A}}. \tag{C.4}$$

From [82], we know that for blind channel estimation, the FIM is singular and its inverse does not exist. Then, some constraints should be utilized to make $\mathbf{J}$ a non-singular matrix. Instead of taking any specific constraint, we use the minimal constrained CRB defined as in [82].

Lemma C.1 [82]: Suppose the FIM for $\boldsymbol{\vartheta} = [\text{vec}(\mathbf{H}), \mathbf{d}_i]^T$ is

$$\mathbf{J} = \frac{1}{\sigma_n^2}\begin{bmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{bmatrix} \tag{C.5}$$

where $\mathbf{J}_{11}$ is of dimension $N_tN_r(P+1) \times N_tN_r(P+1)$ and assume $\mathbf{J}$ is singular but $\mathbf{J}_{22}$ is nonsingular. Then, the minimal constrained CRB for vec($\mathbf{H}$) is

$$\mathrm{CRB}_{\mathrm{vec}(\mathbf{H})} = \sigma_n^2[\mathbf{J}_{11} - \mathbf{J}_{12}\mathbf{J}_{22}^{-1}\mathbf{J}_{21}]^\dagger. \tag{C.6}$$

This is a particular constrained CRB that yields the lowest value for $tr\{\mathrm{CRB}\}$ among all lists of a minimal number of independent constraints.

Applying the above lemma, we obtain

$$\begin{aligned}
\mathrm{ACRB}_{\mathrm{vec}(\mathbf{H})} &= \sigma_n^2(\boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1}\boldsymbol{\mathcal{D}}_i - \boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}})^{-1}\boldsymbol{\mathcal{A}}^H\mathbf{R}_w^{-1/2}\mathbf{D}_i)^\dagger \\
&= \sigma_n^2(\boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1/2}(\mathbf{I} - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H)^{-1})\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{D}}_i)^\dagger \\
&= \sigma_n^2(\boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1/2}\mathbf{P}_{\boldsymbol{\mathcal{A}}}^{\perp}\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{D}}_i)^\dagger \\
&= \sigma_n^2(\boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1/2}\mathbf{U}_o\mathbf{U}_o^H\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{D}}_i)^\dagger.
\end{aligned} \tag{C.7}$$

From the approximation, the noise $\boldsymbol{\eta}_i$ can be considered independent for each i. Then, the ACRB, by observing $\mathbf{z}$, can be found directly from

$$\mathrm{ACRB}_{\mathrm{vec}(\mathbf{H})} = \sigma_n^2\left(\sum_{i=0}^{M-1}\boldsymbol{\mathcal{D}}_i^H\mathbf{R}_w^{-1/2}\mathbf{U}_o\mathbf{U}_o^H\mathbf{R}_w^{-1/2}\boldsymbol{\mathcal{D}}_i\right)^\dagger. \tag{C.8}$$

Asymptotically, we have

$$\begin{aligned}
&\sum_{i=0}^{M-1}(\mathbf{D}_i^{(k)})^H\mathbf{R}_w^{-1/2}\mathbf{U}_o\mathbf{U}_o^H\mathbf{R}_w^{-1/2}\mathbf{D}_i^{(p)} \\
=&\sum_{n=1}^{K}\sum_{m=1}^{K}\mathbf{C}_n^H\mathbf{R}_w^{-1/2}\mathbf{U}_o\mathbf{U}_o^H\mathbf{R}_w^{-1/2}\mathbf{C}_m\sum_{i=0}^{M-1}d_i^*((n-1)K+k)d_i((m-1)K+p) \\
\approx&2M\sigma_s^2\sum_{n=1}^{N}\mathbf{C}_n^H\mathbf{R}_w^{-1/2}\mathbf{U}_o\mathbf{U}_o^H\mathbf{R}_w^{-1/2}\mathbf{C}_n\delta_{k-p} \\
=&2M\sigma_s^2\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H\delta_{k-p}.
\end{aligned} \tag{C.9}$$

Equation (C.9) is obtained asymptotically for large M, bearing in mind that elements of $\mathbf{d}_i$ are i.i.d with variance $2\sigma_s^2$ if $s_i^{(p)}(n)$ are i.i.d with variance σ_s^2.

Therefore, the ACRB for vec($\mathbf{H}$) is

$$\begin{aligned}\mathrm{ACRB}_{\mathrm{vec}(\mathbf{H})} &= \sigma_n^2(\mathbf{I}_K \otimes (2M\sigma_s^2 \boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H))^\dagger \\ &= \frac{\sigma_n^2}{2M\sigma_s^2}\mathbf{I}_K \otimes (\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H)^\dagger \\ &= \mathbf{I}_K \otimes \left(\frac{\sigma_n^2(\boldsymbol{\mathcal{K}}^H)^\dagger \boldsymbol{\mathcal{K}}^\dagger}{2M\sigma_s^2}\right). \end{aligned} \tag{C.10}$$

Appendix D

Stochastic CRB for Precoded MIMO OFDM

For circular complex zero-mean Gaussian random variable $\mathbf{y}_i = \Re\{\mathbf{y}_i\} + j\Im\{\mathbf{y}_i\}$ with covariance matrices $\mathbf{R}_\mathbf{y} = \mathrm{E}\{\mathbf{y}_i\mathbf{y}_i^H\}$ parameterized by a real vector $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_\chi]$, the FIM of this vector $\boldsymbol{\theta}$ is given by [99],[100]

$$[\mathrm{FIM}]_{ab} = M\mathrm{tr}\left(\frac{d\mathbf{R}_\mathbf{y}}{d\theta_a}\mathbf{R}_\mathbf{y}^{-1}\frac{d\mathbf{R}_\mathbf{y}}{d\theta_b}\mathbf{R}_\mathbf{y}^{-1}\right), \quad \text{for } a,\ b = 1, ..., \chi \tag{D.1}$$

where M is the number of the available snapshots. For the proposed method, the signal covariance matrix is

$$\mathbf{R}_\mathbf{y} = \mathrm{E}\{\mathbf{y}_i\mathbf{y}_i^H\} = \sigma_s^2\boldsymbol{\mathcal{H}}_\mathbf{y}\boldsymbol{\mathcal{F}}^H\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{F}}\mathcal{H}_\mathbf{y}^H + \sigma_n^2\mathbf{I}_{KN_r\times MN_r} \tag{D.2}$$

and is parameterized by $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2]$. The following properties are very useful for the derivation of CRB

$$\mathrm{tr}(\mathbf{XY}) = \mathrm{vec}(\mathbf{X}^H)^H\mathrm{vec}(\mathbf{Y}) \tag{D.3}$$

$$\mathrm{vec}(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})\mathrm{vec}(\mathbf{Y}) \tag{D.4}$$

$$(\mathbf{X} \otimes \mathbf{Y})(\mathbf{Z} \otimes \mathbf{W}) = (\mathbf{XZ}) \otimes (\mathbf{YW}) \tag{D.5}$$

which hold for all matrices $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and $\mathbf{W}$. Using these properties, we can rewrite (D.1) as

$$\frac{1}{M}[\mathrm{FIM}]_{a,b} = \mathrm{vec}\left(\frac{d\mathbf{R}_\mathbf{y}}{d\theta_a}\right)^H(\mathbf{R}_\mathbf{y}^{-T} \otimes \mathbf{R}_\mathbf{y}^{-1})\mathrm{vec}\left(\frac{d\mathbf{R}_\mathbf{y}}{d\theta_b}\right) \tag{D.6}$$

or equivalently

$$\frac{1}{M}\text{FIM} = \left(\frac{d\mathbf{r}_\mathbf{y}}{d\boldsymbol{\theta}^T}\right)^H (\mathbf{R}_\mathbf{y}^{-T} \otimes \mathbf{R}_\mathbf{y}^{-1})\text{vec}\left(\frac{d\mathbf{r}_\mathbf{y}}{d\boldsymbol{\theta}^T}\right) \tag{D.7}$$

where

$$\mathbf{r}_\mathbf{y} = \text{vec}(\mathbf{R}_\mathbf{y}) = \sigma_s^2(\boldsymbol{\mathcal{H}}_\mathbf{y}^* \otimes \boldsymbol{\mathcal{H}}_\mathbf{y})\text{vec}(\boldsymbol{\mathcal{F}}^H\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{F}}) + \sigma_n^2\text{vec}(\mathbf{I}_{KN_r\times KN_r}). \tag{D.8}$$

Equation (D.7) can be partitioned as

$$\frac{1}{M}\text{FIM} = \begin{bmatrix} \mathbf{G}^H \\ \boldsymbol{\Delta}^H \end{bmatrix} [\mathbf{G}\ \boldsymbol{\Delta}] \tag{D.9}$$

where

$$[\mathbf{G} \quad | \quad \boldsymbol{\Delta}] = (\mathbf{R}_\mathbf{y}^{-T/2} \otimes \mathbf{R}_\mathbf{y}^{-1/2})\frac{d\mathbf{r}_\mathbf{y}}{d\boldsymbol{\theta}^T} = (\mathbf{R}_\mathbf{y}^{-T/2} \otimes \mathbf{R}_\mathbf{y}^{-1/2})\left[\frac{d\mathbf{r}_\mathbf{y}}{d\boldsymbol{\theta}_1^T} \quad | \quad \frac{d\mathbf{r}_\mathbf{y}}{d\boldsymbol{\theta}_2^T}\right]. \tag{D.10}$$

From Lemma 3, the minimal CRB for $\boldsymbol{\theta}_1$ can be written as

$$\text{CRB}_{\boldsymbol{\theta}_1} = \frac{1}{N}[\mathbf{G}^H\mathbf{G} - \mathbf{G}^H\boldsymbol{\Delta}(\boldsymbol{\Delta}^H\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}^H\mathbf{G}]^\dagger = \frac{1}{N}(\mathbf{G}^H\mathbf{P}_{\boldsymbol{\Delta}}^\perp\mathbf{G})^\dagger. \tag{D.11}$$

We need to evaluate the derivatives of $\mathbf{r}_\mathbf{y}$ with respect to $\boldsymbol{\theta}$. Let us firstly make the following partition

$$[\mathbf{G} \mid \boldsymbol{\Delta}] = [\mathbf{G}_{11}, ..., \mathbf{G}_{pq}, ..., \mathbf{G}_{N_tN_r} \mid \mathbf{v}\ \ \mathbf{u}] \tag{D.12}$$

where

$$\mathbf{G}_{pq} = [\mathbf{G}_{pq}^R \mid \mathbf{G}_{pq}^I] = (\mathbf{R}_\mathbf{y}^{-T/2} \otimes \mathbf{R}_\mathbf{y}^{-1/2})\left[\frac{d\mathbf{r}_\mathbf{y}}{d\Re\{\mathbf{h}_{pq}^T\}} \mid \frac{d\mathbf{r}_\mathbf{y}}{d\Im\{\mathbf{h}_{pq}^T\}}\right] \tag{D.13}$$

$$\mathbf{v} = (\mathbf{R}_\mathbf{y}^{-T/2} \otimes \mathbf{R}_\mathbf{y}^{-1/2})\frac{d\mathbf{r}_\mathbf{y}}{d\sigma_s^2} \tag{D.14}$$

$$\mathbf{u} = (\mathbf{R}_\mathbf{y}^{-T/2} \otimes \mathbf{R}_\mathbf{y}^{-1/2})\frac{d\mathbf{r}_\mathbf{y}}{d\sigma_n^2}. \tag{D.15}$$

After some algebraic manipulations, the following results can be obtained

$$\mathbf{G}_{pq}^R(:,l) = \text{vec}\left(\mathbf{R}_\mathbf{y}^{-1/2}\frac{d\mathbf{R}_\mathbf{y}}{d\Re\{h_{pq}(l-1)\}}\mathbf{R}_\mathbf{y}^{-1/2}\right) = \text{vec}(\mathbf{A}_{pq,l} + \mathbf{A}_{pq,l}^H) \tag{D.16}$$

$$\mathbf{G}_{pq}^I(:,l) = \text{vec}\left(\mathbf{R}_\mathbf{y}^{-1/2}\frac{d\mathbf{R}_\mathbf{y}}{d\Im\{h_{pq}(l-1)\}}\mathbf{R}_\mathbf{y}^{-1/2}\right) = \text{vec}(j\mathbf{A}_{pq,l} - j\mathbf{A}_{pq,l}^H) \tag{D.17}$$

$$\mathbf{v} = \text{vec}\left(\mathbf{R}_\mathbf{y}^{-1/2}\frac{d\mathbf{R}_\mathbf{y}}{d\sigma_s^2}\mathbf{R}_\mathbf{y}^{-1/2}\right) = \text{vec}(\mathbf{R}_\mathbf{y}^{-1/2}\boldsymbol{\mathcal{H}}_\mathbf{y}\boldsymbol{\mathcal{F}}^H\boldsymbol{\mathcal{P}}\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{H}}_\mathbf{y}^H\mathbf{R}_\mathbf{y}^{-1/2}) \tag{D.18}$$

$$\mathbf{u} = \text{vec}\left(\mathbf{R}_\mathbf{y}^{-1/2}\frac{d\mathbf{R}}{d\sigma_n^2}\mathbf{R}_\mathbf{y}^{-1/2}\right) = \text{vec}(\mathbf{R}_\mathbf{y}^{-1}). \tag{D.19}$$

Matrix $\mathbf{A}_{pq,l}$ is given in (5.76a). Proved.