CHANNEL ESTIMATION AND EQUALIZATION FOR WIRELESS MULTICARRIER TRANSMISSIONS

THESIS SUBMITTED IN ACCORDANCE WITH THE REQUIREMENTS
OF THE UNIVERSITY OF LIVERPOOL FOR THE DEGREE OF
DOCTOR IN PHILOSOPHY BY YI MA

THE UNIVERSITY OF LIVERPOOL
DEPARTMENT OF
ELECTRICAL ENGIEERING AND ELECTRONICS

The undersigned hereby certify that they have read and recommend to the Faculty of Engineering for acceptance a thesis entitled **"Channel Estimation and Equalization for Wireless Multicarrier Transmissions"** by **Yi Ma** in accordance with the requirements for the degree of **Doctor of Philosophy.**

Dated: May 2005

External Examiner: _____
Dr. H. Ghafouri-Shiraz

Research Supervisor: _____
Dr. Yi Huang

Examing Committee: _____
Dr. H. Ghafouri-Shiraz

_____
Prof. A. K. Nandi

BLANK PAGE IN ORIGINAL

To My Parents and Wife.

Table of Contents

Abstract

With the demand for higher date rates and quality of service for future wireless communications, the need to cope with the shared wireless medium increases along with the challenges. For example, frequency-selective multipath induced fading, time-selective mobility induced impairment, and the design of high performance transceivers. Orthogonal Frequency Division Multiplexing (OFDM) systems rely on Fourier transform based block transmission and redundant precoding to offer the promised high data rate transmission with high bandwidth efficiency over frequency selective channels. It has therefore been widely used for digital audio broadcasting (DAB), digital video broadcasting (DVB) and wireless local area networks (WLAN). More recently, OFDM has been considered to be the physical layer of the fourth generation mobile communication systems in Europe.

This thesis investigates OFDM systems in the presence of unknown frequency selective fading and even over rapidly time-varying multipath channels. A number of self symbol-recovery OFDM transceivers are designed to mitigate unknown multipath fading by exploiting subcarrier correlations. Relying on distinct block precodings, a novel multi-antenna based OFDM transmission scheme is introduced to enable a low complexity blind multichannel estimator. To further reduce the channel estimation complexity, a superimposed-training based OFDM transceiver is developed without loss of the bandwidth efficiency. A pilot-assisted OFDM transceiver is investigated for combating rapidly time-varying multipaths.

Our research on OFDM over time and frequency selective fading channels offers the potential benifit of enhanced performance, capacity, and high speed information

transmission to broadband wireless communication systems. The ultimate goal is to have advanced wireless services, such as mobile computing, high-speed internet access, and other personal communication services, made accessible to more people.

Acknowledgements

I would like to thank Dr. Yi Huang, my supervisor, for his suggestions, constant support and encouragement during this research. Without his help, this research would not be so smooth. It is really a happy time in my PhD study.

Dr. Waleed Al-Nuaimy showed his interest in my work and gave me useful advices in academic writing. Dr. Mansor Nakhkash gave me great help in my first research stage. He provided me with many useful references and friendly encouragement.

It is my pleasure of meeting Dr. Vicente Zarzoso. He is a wonderful guy, and his help improved my research ability significantly.

I would like to thank Mr. Zhengyan Guo, a math teacher in my high school. He encouraged me to attend math Olympic, and brought me into the realm of science.

I am pleased to thank the financial support from British council (ORS award) and the University of Liverpool (University Scholarship) for the financial support, which was awarded to me for the period 2001–2004, was crucial to the successful completion of my research.

Many thanks the following colleagues: Yiyuan Xiong, Jietao Zhang, H.M. Looe, Wenhu Tang, Fei Liu, and Martin Evans (from ARM Ltd.). Special thanks to Professor A. K. Nandi, Professor Michael Fang and Dr. K. G. Evans for the useful discussions and suggestions.

Finally and also the most important, I am grateful to my parents for their unconditional support in my life. They gave the meaning of my life. Also, I would like to thank my wife, Na Yi. She made my life so wonderful and comfortable.

Acronyms

ADSL	Asynchronous digital subscriber line
AWGN	Additive white Gaussian noise
BER	Bit error rate
BP	Block precoding
CAB	Channel average bias
CCE	Combined channel estimator
CDMA	Code division multiple access
CFO	Carrier frequency offset
CIR	Channel impulse response
CP	Cyclic prefix
CSI	Channel state information
DAB	Digital audio broadcasting
DD	Decision directed
DFT	Discrete Fourier transform
DRL	Data record length
DSP	Digital signal processor
DVB	Digital video broadcasting
ETSI	European telecommunications standard institute
FFT	Fast Fourier transform
FIR	Finite impulse response
ICI	Inter-carrier interference
IFFT	Inverse Fast Fourier transform
ISI	Inter-symbol interference
LE	Linear equation
LMMSE	Least minimum mean square error
LS	Least square

MAN	Metropolitan area networks
MD	Minimum distance
MIMO	Multiple inputs multiple outputs
MISO	Multiple inputs single output
ML	Maximum Likelihood
(M)MSE	(Minimum) mean square error
MUI	Multiple user interference
OFDM	Orthogonal frequency-division multiplexing
PD	Phase directed
PS	Parallel to serial
PSK	Phase shift key
RC	Repetition code
SD	Sign directed
SIR	Signal to interference ratio
STO	Symbol timing offset
SNR	Signal to noise ratio
STC	Space time coding
STBC	Space time block coding
TIR	Training to information ratio
WLAN	Wireless local area networks
ZF	Zero forcing
ZP	Zero padding

List of Figures

Notations

Throughout this thesis, lower case boldface symbols are used to denote column vectors (e.g. $\mathbf{s}$) while upper case boldface symbols are used to denote matrices (e.g. $\mathbf{F}$). Their sizes are emphasized using subscripts: for instance $\mathbf{s}_M$ indicates a size $M \times 1$ vector; $\mathbf{F}_M$ denotes a size $M \times M$, or $\mathbf{F}_{P\times M}$ indicates a size $P \times M$. Elements of vectors or matrices are expressed as $s(m)$ or $F(n, m)$.

$(\cdot)^*$, $(\cdot)^T$, $(\cdot)^{\mathcal{H}}$ and $(\cdot)^\dagger$ denotes complex conjugate, transpose, Hermitian transpose and pseudo inverse respectively. $\odot$, $\otimes$ and $\star$ indicate Hadamard product, Kronecker product and convolution respectively.

$\mathcal{E}\{\cdot\}$ stands for expectation, $\|\cdot\|$ for the norm of a vector (or a matrix), $\mathrm{Rank}\{\cdot\}$ for the rank of a matrix, and $\mathrm{Tr}\{\cdot\}$ for the trace of a matrix.

$\mathbf{I}_M$ denotes the identity matrix of size M, and $\mathbf{0}_{K\times P}$ denotes a zero matrix of size $K \times P$; $\mathbf{D}(\mathbf{s})$ denotes a diagonal matrix with $\mathbf{s}$ on its diagonal.

Chapter 1

Introduction to multicarrier transmission

This chapter provides an introduction to orthogonal frequency-division multiplexing (OFDM) transmission over frequency-selective fading channels. The presentation does not intend to be comprehensive, but the aim is to present all materials necessary to fully understand the following chapters. A more general and comprehensive coverage of multicarrier and OFDM systems can be found in [1].

Historically, OFDM was introduced in the 60's [2] as an inter-symbol interference (ISI) resilient scheme. However, it was hardly considered for practical applications due to its high computation complexity. With the rapid development of digital signal processing (DSP) technologies, OFDM, which requires an inverse fast Fourier transform (IFFT)/FFT component, can now be simply implemented in a high-speed embedded system. It receives a growing interest [3] and has now been adopted in many national and international standards. Specifically, it has been chosen as a solution for digital audio and video broadcasting in Europe (DAB [4] and DVB [5]). It has also been chosen for high-speed modem transmission over copper wires (ADSL and VDSL [6]). Recently, it has been adopted in the standards for WLAN both in Europe

(HIPERLAN/2 [7]-[8]) and in America (IEEE802.11a [9]). Europe and America are considering OFDM for wireless metropolitan area networks (Wireless MAN) [10].

The rest of this chapter is organized as follows: Section 1.1 overviews typical OFDM systems; Section 1.2 presents an OFDM system in the presence of multipath fading; Inter-carrier interference (ICI) in OFDM is introduced in Section 1.3; and, further understanding of OFDM is presented in Section 1.4. Finally, Section 1.5 presents the objective of our research.

1.1 Typical OFDM systems

This thesis only focuses on the discrete-time baseband aspects of OFDM systems. Typically, the OFDM systems are classified into two categories according to the type of time-domain guard: cyclic prefixed (CP) OFDM and zero padding (ZP) OFDM. Figures 1.1 and 1.2 depict the discrete-time equivalent models of CP-OFDM and ZP-OFDM systems respectively. Conventionally, OFDM modulators employ an IFFT component to map information-bearing symbols onto IFFT/FFT grids (orthogonal subcarriers) for the block transmission. The subcarrier spacing Δf is a reciprocal of the block duration T_b. Then, the parallel output of the IFFT is converted into the serial stream by employing a parallel-to-serial (P/S) component. Normally, a time-domain guard with the duration T_g is inserted between two consecutive OFDM blocks for mitigating the inter-block interference (IBI) caused by the time-dispersive channel and introducing the circulant property. Practically, most of OFDM-based communication systems use cyclic prefix as the time-domain guard (for instance [4]-[5], [7], [9]-[10]). At the receiving end, the CP-OFDM receiver first discards CPs, and then groups the received data into blocks with the block duration T_b. Those

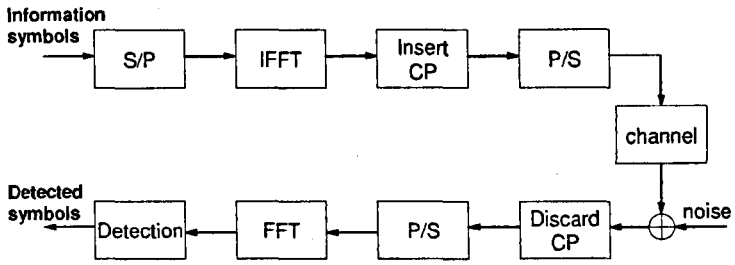

Figure 1.1: Discrete-time equivalent model of CP-OFDM transceiver.

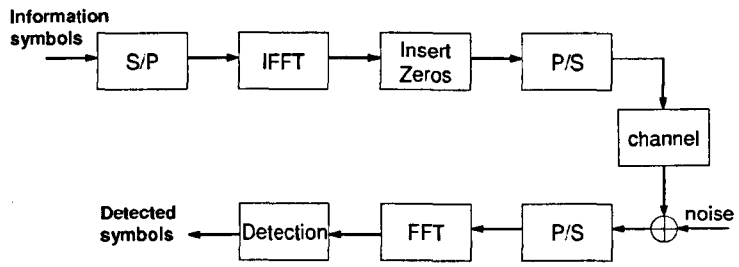

Figure 1.2: Discrete-time equivalent model of ZP-OFDM transceiver.

blocks are then fed into a FFT component to generate so called *frequency-domain signal.* Channel equalization and signal detection are then performed in the frequency domain.

Recently, it was proposed to replace CP by ZP (see [11]-[14]). The time-domain guard between two consecutive OFDM blocks is filled with zero symbols. Differed from the CP-OFDM receiver, as shown in Figure 1.2, the ZP-OFDM receiver does not remove the guard interval. It groups the received signal into blocks with the block duration $T_b + T_g$. Then, those blocks are fed to a FFT component to generate the frequency-domain signal. Due to the redundancy introduced by the guard time, [12] shows that the ZP-OFDM system guarantees the symbol recovery in the noiseless case regardless of channel zero locations. However, the ZP-OFDM receiver increases the equalization complexity.

The main advantage of OFDM is that the OFDM modem can convert a frequency-selective fading channel into a number of flat-fading subchannels, such that equalization of the frequency-selective channel becomes very simple. Besides this, OFDM

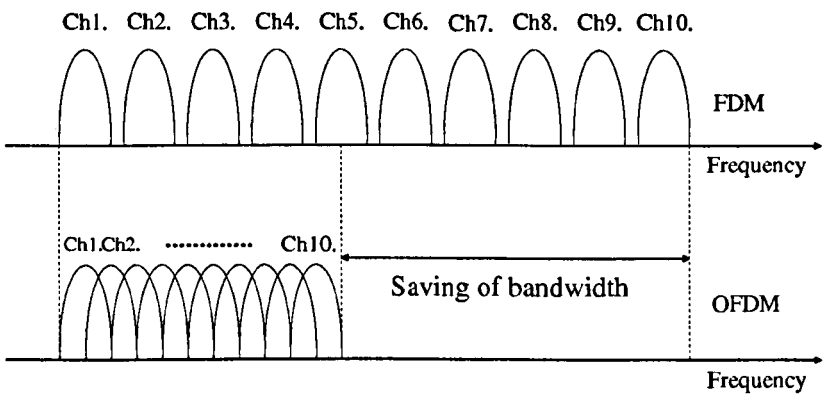

Figure 1.3: Comparison between conventional FDM and OFDM

is a spectral efficient transmission scheme. As an example, Figure 1.3 shows that OFDM can save almost fifty percent bandwidth compared with the conventional FDM scheme by overlapping orthogonal subcarriers. The working philosophy of OFDM in the frequency-selective channel is presented in Chapter 1.2.

1.2 OFDM in the multipath fading channel

The presence of reflecting objects and scatters in the channel creates a constantly changing environment that dissipates the signal energy in amplitude, phase, and time. These effects result in multiple versions of the transmitted signal that arrive at the receiving antenna. The random phase and amplitude of the different multipath components cause fluctuations in signal strength, thereby inducing small-scale fading and signal distortion. Due to the propagation delay, multipath propagation often lengthens the time required for the baseband signal to reach the receiver and causes inter-symbol interference (ISI) [15].

1.2.1 Propagation channel model

Figure 1.4 depicts the propagation channel model, which consists of three parts:

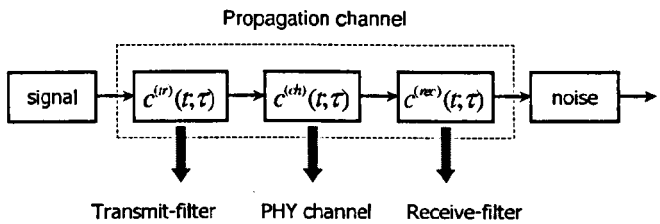

Figure 1.4: Continuous-time time-varying communication system.

transmit-filter, receive-filter and physical channel (also see [16]). The continuous channel impulse response $h_c(t,\tau)$ is a stochastic process in the time variable t. It is common practice to assume the channel to be wide sense stationary for a fixed lag τ and uncorrelated for different lags [17]. As shown in Figure 1.4, $h_c(t,\tau)$ is the convolution of the transmit-filter $c^{(tr)}(t,\tau)$, the physical channel $c^{(ch)}(t,\tau)$, and the receive-filter $c^{(rec)}(t,\tau)$. Defining $\psi(t,\tau) = c^{(tr)}(t,\tau) \star c^{(rec)}(t,\tau)$, the propagation channel model can be mathematically described as (see [18])

$$h_c(t,\tau) = \sum_{\nu} \psi(t, \tau - \tau_\nu(t)) \sum_{\mu} a_{\mu,\nu}(t) e^{j\phi_{\mu,\nu}(t)} e^{j2\pi f_{\mu,\nu}(t)t} \tag{1.2.1}$$

where subscript $_c$ denotes continues time, $a_{\mu,\nu}(t)$ and $\phi_{\mu,\nu}(t)$ denote the random amplitude and random phase for the μth ray of the νth lag, $f_{\mu,\nu}(t)$ denotes the Doppler frequency. If the channel is time-invariant, we can omit t in (1.2.1) and let $f_{\mu,\nu}(t) = 0$. Then, (1.2.1) becomes the frequency-selective channel model as

$$h_c(\tau) = \sum_{\nu} \psi(\tau - \tau_\nu) \bar{a}_v \tag{1.2.2}$$

where $\bar{a}_v = \sum_\mu a_{\mu,\nu} e^{j\phi_{\mu,\nu}}$. When the information-bearing symbols $s(\ell)$ goes through the channel, we can relate input and output as

$$z_c(t) = \sum_{\ell=-\infty}^{\infty} s(\ell) h_c(t - \ell T_s) \tag{1.2.3}$$

where T_s is the symbol duration. Once the output signal $z_c(t)$ is sampled at the symbol rate T_s, the discrete-time version of the output signal is a function of the sample index i

$$z(iT_s) = \sum_{\ell=-\infty}^{\infty} s(\ell) h(iT_s - \ell T_s) \tag{1.2.4}$$

We define $z(i) = z(iT_s)$ and are able to rewrite (1.2.4) into

$$z(i) = \sum_{\ell=-\infty}^{\infty} s(\ell) h(i - \ell) \tag{1.2.5}$$

Figure 1.5 demonstrates the relationship (1.2.5) using a flow graph. We can see that it is a moving average process [19], such that the frequency-selective channel can be modelled as a FIR filter with complex coefficients [21]. Since the upper bound of channel length L_u is experimentally measurable, the sampled channel impulse response (CIR), h, can be expressed by a $L_u \times 1$ vector as: $\mathbf{h} = [h(0),\ h(1),\ \cdots,\ h(L_u-1)]^T$. The channel frequency response is therefore given by

$$H(m) = \sum_{l=0}^{L_u-1} h(l) W_M^{ml} \tag{1.2.6}$$

where $W_M^{ml} = \exp\{-j2\pi ml/M\}$, and M denotes the number of DFT grids and also equals the number of subcarriers in the OFDM system. Later on, we would like to investigate the output $z(i)$ by employing the CP-OFDM and ZP-OFDM signals as the input. Also, we intend to exhibit the outstanding features of typical OFDM systems.

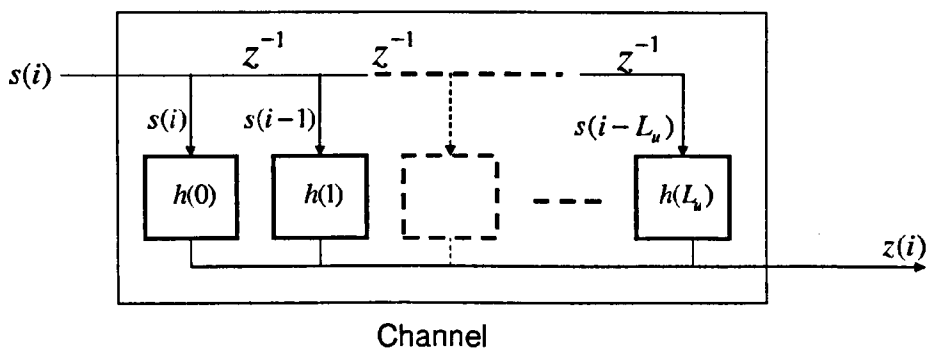

Figure 1.5: Input-output relationship of the discrete-time channel.

1.2.2 CP-OFDM system model

Prior to transmission, the information-bearing symbols are grouped into blocks $\mathbf{s}(n) = [s(n,0),\ s(n,1),\ \cdots,\ s(n,M-1)]^T$ with the size $M \times 1$. As shown in Figure 1.6, these blocks are fed into IFFT to generate OFDM blocks $\bar{\mathbf{s}}(n) = \mathbf{F}_M^{\mathcal{H}}\mathbf{s}(n)$, where $\mathbf{F}_M$ is the normalized DFT matrix with the (m,k)th entry $W_M^{mk}/\sqrt{M}$. Then, CPs are prepended to the OFDM blocks by duplicating the last L_{cp} samples of $\bar{\mathbf{s}}(n)$ to produce the transmitted blocks $\mathbf{x}(n)$ as

$$\mathbf{x}(n) = \Theta_{\text{cp}}\mathbf{s}(n) \tag{1.2.7}$$

where L_{cp} is the length of CP, Θ_{cp} is a $J \times M$ matrix formed by $\Theta_{\text{cp}} = [\mathbf{F}_{\text{cp}},\ \mathbf{F}_M]^{\mathcal{H}}$, and the $M \times L_{\text{cp}}$ matrix $\mathbf{F}_{\text{cp}}$ is formed by the last L_{cp} columns of $\mathbf{F}_M$.

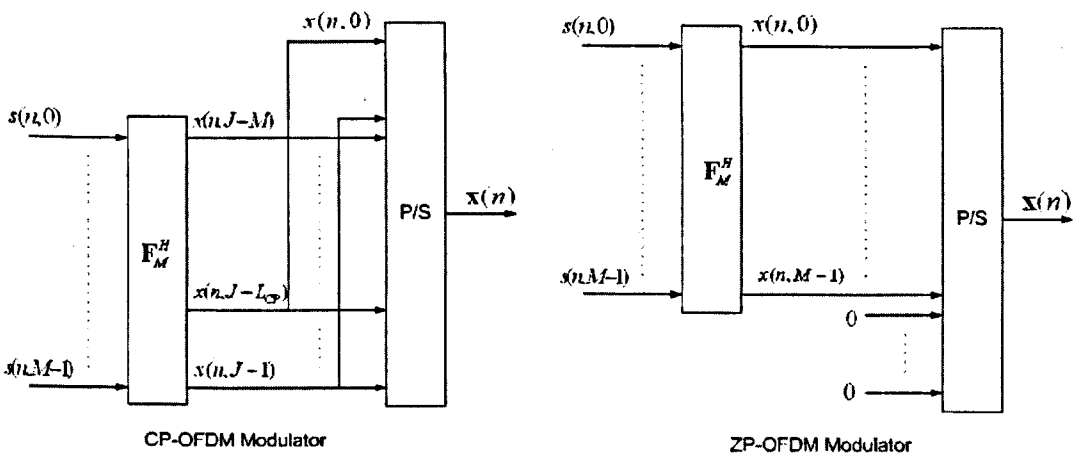

Figure 1.6: OFDM modulator with cyclic prefix and zero-padding

When the blocks $\mathbf{x}(n)$ go through the propagation channel, the input-output relationship is given by (see [12])

$$\mathbf{z}_J(n) = \mathbf{\Delta}_{\text{low}}\mathbf{\Theta}_{\text{cp}}\mathbf{s}(n) + \underbrace{\mathbf{\Delta}_{\text{up}}\mathbf{\Theta}_{\text{cp}}\mathbf{s}(n-1)}_{\textbf{IBI part}} \tag{1.2.8}$$

where $\mathbf{\Delta}_{\text{low}}$ is the $J \times J$ lower triangular Toeplitz filtering matrix with the first column $[h(0), \cdots, h(L_u - 1),\ 0,\ \cdots, 0]^T$; $\mathbf{\Delta}_{\text{up}}$ is the $J \times J$ upper triangular Toeplitz filtering matrix with the first row $[0,\ \cdots, 0,\ h(L_u - 1), \cdots, h(1)]$. Normally, the CP length, L_{cp}, is designed to be larger than the upper bound of channel order, therefore the IBI part can be easily eliminated by discarding the first L_{cp} elements of $\mathbf{z}_J(n)$. The remaining part of $\mathbf{z}_J(n)$ can be expressed as [12]

$$\mathbf{z}_{\text{cp}}(n) = \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) \tag{1.2.9}$$

where $\mathbf{\Gamma}_M$ is an $M \times M$ circulant matrix with the first row $[h(0),\ 0\ , \cdots, 0,\ h(L_u - 1),\ \cdots,\ h(1)]$. Since every circulant matrix can be diagonalized by using eigen-decomposition [12], [14], [20], equalization of CP-OFDM transmission can be performed in the frequency domain. Performing a FFT on $\mathbf{z}_{\text{cp}}(n)$, we can obtain the frequency-domain version of $\mathbf{z}_{\text{cp}}(n)$ as

$$\begin{aligned} \tilde{\mathbf{z}}_{\text{cp}}(n) &= \mathbf{F}_M \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) \\ &= \mathcal{D}_M \mathbf{s}(n) \end{aligned} \tag{1.2.10}$$

where $\mathcal{D}_M = \text{diag}\{H(0), H(1),\ \cdots,\ H(M-1)\}$. Once $\mathcal{D}_M$ is available at the receiver, equalization can be implemented using a simple matrix multiplication

$$\hat{\mathbf{s}}(n) = \mathcal{D}_M^{\dagger} \tilde{\mathbf{z}}_{\text{cp}}(n) \tag{1.2.11}$$

where $\hat{\mathbf{s}}(n)$ is the estimated version of the information symbol block. However, the symbol recovery is guaranteed only when the diagonal channel matrix $\mathcal{D}_M$ does not

comprise zeros on its diagonal. Practically, channel code is employed to introduce enough code redundancy [17]. From the above discussions, we can draw the following conclusions:

c1.1. By employing the cyclic prefix, (1.2.10) shows that the channel matrix can be diagonalized. Symbols in the block $\mathbf{s}(n)$ are not interfered each other. Therefore, CP-OFDM can mitigate the ISI.

c1.2. Since $\mathcal{D}_M$ may comprise at most $L_u - 1$ zeros on its the diagonal (see [22], [23]), channel code has to be used to guarantee the symbol recovery.

1.2.3 ZP-OFDM system model

The right-hand part of Figure 1.6 depicts the discrete-time equivalent model of the ZP-OFDM modulator. Differed from CP-OFDM, the relationship between $\mathbf{x}(n)$ and $\mathbf{s}(n)$ is given by

$$\mathbf{x}(n) = \mathbf{\Theta}_{\text{zp}}\mathbf{s}(n) \tag{1.2.12}$$

where $\mathbf{\Theta}_{\text{zp}} = [\mathbf{F}_M,\ \mathbf{0}_{M\times L_{\text{zp}}}]^{\mathcal{H}}$, $L_{\text{zp}} = L_{\text{cp}}$ is the number of zero symbols. Then, the output of the propagation channel is given by

$$\mathbf{z}_J(n) = \mathbf{\Delta}_{\text{low}}\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n) + \underbrace{\mathbf{\Delta}_{\text{up}}\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n-1)}_{\textbf{IBI part}} \tag{1.2.13}$$

Since $\mathbf{\Delta}_{\text{up}}\mathbf{\Theta}_{\text{zp}} = \mathbf{0}$ [14], the IBI part does not exist actually, and (1.2.13) becomes

$$\begin{aligned}\mathbf{z}_J(n) &= \mathbf{\Delta}_{\text{low}}\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n)\\ &= \mathbf{\Delta}_{\text{low}}\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n) + \mathbf{\Delta}_{\text{up}}\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n)\\ &= \mathbf{\Gamma}_J\mathbf{\Theta}_{\text{zp}}\mathbf{s}(n)\end{aligned} \tag{1.2.14}$$

where $\mathbf{\Gamma}_J = \mathbf{\Delta}_{\text{low}} + \mathbf{\Delta}_{\text{up}}$ is the $J \times J$ circulant channel matrix.

Similar to the CP-OFDM, equalization of ZP-OFDM transmission is also carried out in the frequency domain. Let $\mathbf{F}_J$ be a $J \times J$ normalized DFT matrix. The frequency-domain blocks can be expressed as

$$\begin{aligned} \tilde{\mathbf{z}}_J(n) &= \mathbf{F}_J \mathbf{z}_J(n) \\ &= \mathbf{F}_J \mathbf{\Gamma}_J \mathbf{F}_J^{\mathcal{H}} \mathbf{F}_J \mathbf{\Theta}_{\text{zp}} \mathbf{s}(n) \\ &= \mathcal{D}_J \underbrace{\mathbf{G}_{\text{zp}} \mathbf{s}(n)}_{\text{ISI part}} \end{aligned} \tag{1.2.15}$$

where $\mathcal{D}_J = \mathbf{F}_J \mathbf{\Gamma}_J \mathbf{F}_J^{\mathcal{H}} = \text{diag}\{H(0), H(1), \cdots, H(J-1)\}$, and $\mathbf{G}_{\text{zp}} = \mathbf{F}_J \mathbf{\Theta}_{\text{zp}}$ with the size $J \times M$. It can be observed that $J = L_{\text{zp}} + M$ and $\mathbf{G}_{\text{zp}}$ has full column rank. The symbol recovery is therefore guaranteed only when the channel frequency response contains at most L_{zp} zeros. Chapter 1.2.2 has mentioned that, $\mathcal{D}_J$ has at most $(L_{\text{u}} - 1) < L_{\text{zp}}$ zeros on its diagonal, therefore the following zero-forcing equalization method can guarantee the symbol recovery

$$\hat{\mathbf{s}}(n) = (\mathcal{D}_J \mathbf{G}_{\text{zp}})^{\dagger} \tilde{\mathbf{z}}_J(n) \tag{1.2.16}$$

From the above analysis, we can conclude that:

c1.3. The results (1.2.15) and (1.2.16) show that the tall matrix $\mathbf{G}_{\text{zp}}$ introduces enough redundancy to guarantee the symbol recovery regardless of channel zero locations. Hence, for the uncoded case, ZP-OFDM outperforms CP-OFDM in the frequency-selective channel.

c1.4 Due to the presence of $\mathbf{G}_{\text{zp}}$, (1.2.15) also shows that information symbols are interfered each other. The receiver complexity needs to be paid for removing the ISI.

1.3 Inter-carrier interference in OFDM

One of the main disadvantages of OFDM is its high sensitivity to carrier frequency offset (CFO), which is caused by oscillator mismatch between the transmitter and the receiver or the Doppler shift (see [26]-[27]). If CFO is an integer multiple of the subcarrier spacing Δf, then subcarriers are still mutually orthogonal, but the received symbols mapped to DFT grids are in the wrong position. If CFO is not an integer multiple of the subcarrier spacing, then energy is spilling over between subcarriers, which results in inter-carrier interference (ICI) [26].

Most of previous literatures consider the ICI effect in additive white Gaussian noise (AWGN) or flat-fading channels [27]-[29]. This thesis analyzes the ICI effect in a frequency-selective fading channel using the matrix formulation. To simplify the discussion, here, we just take CP-OFDM as an example.

As shown in [30], [31], the noise-free version of received signal with CFO is expressible as

$$z(i) = e^{j2\pi f_D T_s i} \sum_{\ell=-\infty}^{\infty} s(\ell)h(i-\ell) \tag{1.3.1}$$

where f_D denotes the CFO. This model considers the communication channel to be static (or slowly time-varying)[1], i.e. the CFO is mainly from the oscillator mismatch. Consider the transmitted signal to be CP-OFDM format, (1.3.1) can be expressed as the following matrix form

$$\mathbf{z}_J(n) = \mathbf{\Omega}_J(n)\mathbf{\Delta}_{\text{low}}\mathbf{\Theta}_{\text{cp}}\mathbf{s}(n) + \mathbf{\Omega}_J(n)\underbrace{\mathbf{\Delta}_{\text{up}}\mathbf{\Theta}_{\text{cp}}\mathbf{s}(n-1)}_{\textbf{IBI part}} \tag{1.3.2}$$

where $\mathbf{\Omega}_J(n)$ is a $J \times J$ diagonal matrix with the (m,m)th entry

$$\Omega_J(n;m,m) = \exp(j\frac{2\pi}{M}\varepsilon_d(nJ+m)) \tag{1.3.3}$$

[1]For the slowly time-varying channel, the doppler shift is very small and ignorable.

and ε_d denotes the normalized CFO with $\varepsilon_d = f_D M T_s$ [2]. When CP is removed, the remaining part of $\mathbf{z}_J(n)$ becomes

$$\mathbf{z}_{\text{cp}}(n) = \mathbf{\Omega}_M(n)\mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) \tag{1.3.4}$$

where $\mathbf{\Omega}_M(n)$ is a $M \times M$ diagonal matrix with the (m,m)th entry

$$\Omega_M(n;m,m) = \exp(j\frac{2\pi}{M}\varepsilon_d(nJ + m + L_{\text{cp}})) \tag{1.3.5}$$

The DFT of (1.3.4) is therefore given by

$$\begin{aligned} \tilde{\mathbf{z}}_{\text{cp}}(n) &= \mathbf{F}_M \mathbf{\Omega}_M(n)\mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) \\ &= \Omega_M(n;0,0)\mathcal{D}_M \mathbf{s}(n) + \mathbf{v}_{\text{ICI}}(n) \end{aligned} \tag{1.3.6}$$

where

$$\mathbf{v}_{\text{ICI}}(n) = \mathbf{F}_M[\mathbf{\Omega}_M(n) - \Omega_M(n;0,0)\mathbf{I}_M]\mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) \tag{1.3.7}$$

is the ICI part. (1.3.6) shows that CFO introduces not only the constant phase rotation $\Omega_M(n;0,0)$, but also the additive noise $\mathbf{v}_{\text{ICI}}(n)$. Figure 1.7 gives an example to illustrate the result of (1.3.6). In this example, all subcarriers for the CP-OFDM transmission are QPSK modulated. It shows that the ICI power increases with the increase of CFO. The signal to interference ratio (SIR) becomes very low when the normalized CFO $\varepsilon_d = 0.5$. To mitigate the ICI power, a number of ICI self-cancelling schemes [29] and CFO estimation algorithms [30], [32]-[36] have been proposed. Details of these schemes will be introduced in the following chapters.

[2] The block duration $T_b = MT_s$, and the subcarrier spacing $\Delta f = 1/T_b$, so this result is equivalent to $\varepsilon_d = f_D/\Delta f$.

1.4 Block precoded OFDM

Observing (1.2.9) and (1.2.14), the CP-OFDM and ZP-OFDM transmissions can be rewritten as the following matrix form,

$$\mathbf{z}_{\mathrm{cp}}(n) = \mathbf{\Gamma}_M \mathbf{\Theta}_{\mathrm{cp}} \mathbf{s}(n) \tag{1.4.1}$$

$$\mathbf{z}_{\mathrm{zp}}(n) = \mathbf{\Gamma}_J \mathbf{\Theta}_{\mathrm{zp}} \mathbf{s}(n) \tag{1.4.2}$$

where $\mathbf{\Theta}_{\mathrm{cp}} = \mathbf{F}_M^{\mathcal{H}}$ and $\mathbf{\Theta}_{\mathrm{zp}}$ can be regarded as linear block precoders. Then, CP-OFDM and ZP-OFDM have the following unified form

$$\mathbf{z}(n) = \mathbf{\Gamma}\mathbf{\Theta}\mathbf{s}(n) \tag{1.4.3}$$

where $\mathbf{\Gamma}$ is a circulant channel matrix, and $\mathbf{\Theta}$ is a linear block precoder. If we let $\mathbf{\Theta} = \mathbf{F}_M^{\mathcal{H}}\mathbf{W}$, where $\mathbf{W}$ is a linear block precoder with the full column rank, then (1.4.3) becomes the block precoded (BP) OFDM model addressed in [25]. It can be observed that typical OFDM is a special case of the BP-OFDM, when $\mathbf{W} = \mathbf{I}_M$. In the following contents, we intend to introduce other members of BP-OFDM family, for instance multicarrier code-division multiple access (MC-CDMA) and generalized MC-CDMA (GMC-CDMA).

1.4.1 Conventional MC-CDMA

Originally, MC-CDMA was presented as a combined framework with OFDM and CDMA [37]. It has been shown that MC-CDMA may save almost fifty percent bandwidth of a direct sequence (DS) CDMA [1], [37]. A typical MC-CDMA transmitter for single user is depicted in Figure 1.8. It shows that the MC-CDMA transmitter spreads the original signal using a given spreading sequence over the

frequency domain. Usually, the spreading sequence is selected from a $K_u \times K_u$ Walsh matrix $\mathbf{W}_{K_u}$ [38]. Meanwhile, K_u also denotes the total number of users. Let $\mathbf{s}_\mu(n) = [s_\mu(n,0), \cdots, s_\mu(n,P-1)]^T$ be the information bearing blocks of the μth user, where P denotes the number of transmitted symbols per block. Using the CP-OFDM transmitter, the received signal can be expressed as

$$\mathbf{z}_\mu(n) = \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} [\mathbf{w}_\mu \otimes \mathbf{s}_\mu(n)] \tag{1.4.4}$$

where $\mathbf{w}_\mu$ is the μth column of the Walsh matrix $\mathbf{W}_{K_u}$, and the number of subcarriers $M = P \times K_u$. If there has totally K_a active users, the received signal for the downlink is given by

$$\begin{aligned} \mathbf{z}_{\mathrm{d}}(n) &= \sum_{\mu=1}^{K_a} \mathbf{z}_\mu(n) \\ &= \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \sum_{\mu=1}^{K_a} [\mathbf{w}_\mu \otimes \mathbf{s}_\mu(n)] \end{aligned} \tag{1.4.5}$$

Note that (1.4.4) and (1.4.5) introduce the MC-CDMA models in the traditional way. Next, those models are represented in the view of block precoded systems.

First, let us form an $M \times P$ ($M = K_u \times P$) matrix $\boldsymbol{\mathcal{U}}_\mu$ as

$$\boldsymbol{\mathcal{U}}_\mu = \begin{pmatrix} \mathbf{w}_\mu & & \\ & \ddots & \\ & & \mathbf{w}_\mu \end{pmatrix} \tag{1.4.6}$$

and (1.4.4) can be rewritten into

$$\begin{aligned} \mathbf{z}_\mu(n) &= \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \boldsymbol{\mathcal{U}}_\mu \mathbf{s}_\mu(n) \\ &= \mathbf{\Gamma}_M \bar{\boldsymbol{\mathcal{U}}}_\mu \mathbf{s}_\mu(n) \end{aligned} \tag{1.4.7}$$

where $\bar{\boldsymbol{\mathcal{U}}}_\mu = \mathbf{F}_M^{\mathcal{H}} \boldsymbol{\mathcal{U}}_\mu$ can be regarded as a linear block precoder. For the single user case, the signal recovery is promised by the following theorem.

Theorem 1.4.1. *The information bearing block $\mathbf{s}_\mu(n)$ is uniquely identifiable in (1.4.7) regardless of channel zero locations only when the upper bound of channel length L_u, and the length of spreading sequence K_u satisfy:*

$$L_u \leqslant K_u \tag{1.4.8}$$

Proof. See Appendix 1.A.

In the presence of unknown frequency-selective fading, Theorem 1.4.1 suggests the design criterion of MC-CDMA for the single-user case. For the multiuser case in the downlink, the following contents will show the signal recovery condition in the view of a block precoded system.

Without loss of generality, it is assumed that the K_a active users use spreading sequences $\mathbf{w}_\mu$, for $\mu \in \{1, \cdots, K_a\}$. Then, (1.4.5) can be rewritten into

$$\mathbf{z}_\mathrm{d}(n) = \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \boldsymbol{\mathcal{U}}_{K_a} \mathbf{s}_{M_u}(n) \tag{1.4.9}$$

where $\mathbf{s}_{M_u}(n)$ is an $M_u \times 1$ vector ($M_u = K_a \times P$), which contains multiuser information

$$\mathbf{s}_{M_u}(n) = [s_1(n,0), \cdots, s_{K_a}(n,0), s_1(n,1), \cdots, s_{K_a}(n,P-1)]^T \tag{1.4.10}$$

and $\boldsymbol{\mathcal{U}}_{K_a}$ is given by

$$\boldsymbol{\mathcal{U}}_{K_a} = \begin{pmatrix} \mathbf{W}_{K_a} & & \\ & \ddots & \\ & & \mathbf{W}_{K_a} \end{pmatrix} \tag{1.4.11}$$

and $\mathbf{W}_{K_a} = [\mathbf{w}_1, \cdots, \mathbf{w}_{K_a}]$ with size $K_u \times K_a$. Let $\bar{\boldsymbol{\mathcal{U}}}_{K_a} = \mathbf{F}_M^{\mathcal{H}} \boldsymbol{\mathcal{U}}_{K_a}$, indeed, the linear model (1.4.9) is also a block precoded system.

Theorem 1.4.2. *The information-bearing block $\mathbf{s}_{M_u}(n)$ is uniquely identifiable in (1.4.9) regardless of channel zero locations only when the number of active users K_a satisfy*

$$K_u - K_a \geqslant L_u - 1 \tag{1.4.12}$$

Proof. See Appendix 1.B.

Theorem 1.4.2 shows that the traditional MC-CDMA is sensitive to the channel zero locations. The signal recovery is guaranteed regardless of channel zero locations only when the bandwidth efficiency η_b is

$$\eta_b = \frac{K_a}{K_u} \leqslant \frac{K_u - L_u + 1}{K_u} \tag{1.4.13}$$

1.4.2 Generalized MC-CDMA

To improve the MC-CDMA performance in the frequency-selective fading channel, recently, a generalized MC-CDMA transceiver has been reported, which is resilient to the multiuser interference (MUI) and ISI [24]-[25]. Similar to the traditional MC-CDMA, the GMC-CDMA is also a special block precoded OFDM scheme.

To clarify the difference between MC-CDMA and GMC-CDMA, the downlink (G)MC-CDMA models are represented here

$$\text{MC} - \text{CDMA}: \quad \mathbf{z}_{\text{d}}(n) = \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \boldsymbol{\mathcal{U}}_{K_a} \mathbf{s}_{M_u}(n) \tag{1.4.14}$$

$$\text{GMC} - \text{CDMA}: \quad \mathbf{z}_{\text{d}}(n) = \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{\Theta}_{K_a} \mathbf{s}_{M_u}(n) \tag{1.4.15}$$

It is easy to find that GMC-CDMA replaces the precoder $\boldsymbol{\mathcal{U}}_{K_a}$ with $\mathbf{\Theta}_{K_a}$. To guarantee the signal recovery in MUI regardless of channel zero locations, the precoder $\mathbf{\Theta}_{K_a}$ should meet the following two conditions [22], [24]:

c1.4.1 the block precoder $\mathbf{\Theta}_{K_a}$ should always have full column rank for a K_a, and $M \geqslant PK_u + L_u - 1$

c1.4.2 any PK_a rows of $\mathbf{\Theta}_{K_a}$ are linearly independent.

Obviously, (1.4.14) does not fulfil the above two conditions, and thereby decreases the overall system performance in the frequency-selective fading channel. On the other hand, GMC-CDMA promises the signal recovery, but it has to pay extra receiver complexity.

1.5 Objective of the work

OFDM-based block transmission schemes provide promising high data-rate services over frequency-selective fading channels. On the other hand, most of OFDM systems require the channel state information (CSI) to carry out the coherent signal detection. The channel state information may be obtained at receivers by employing channel estimators. A good channel estimator is very important for improving the overall system performance.

In the presence of slowly time-varying channels, OFDM transceivers equipped with blind channel estimators are capable of self symbol recovery. This thesis proposes a number of blind OFDM transceivers to cope with unknown frequency-selective fading channels. The contributions of this work contain:

1) A differential ICI self-cancelling OFDM transceiver is developed for combating time-varying frequency-selective fading channels. This transceiver is equipped with a blind channel estimator and a semi-coherent signal detector to improve the overall system performance.

2) A novel blind channel estimation method is developed for block precoded OFDM systems by exploiting subcarrier correlation. This blind estimator only uses second-order joint moments to offer a relatively low-complexity approach.

3) A superimposed training based OFDM transceiver is proposed both for cases with single transmit antenna and multiple transmit antennas. The first-order statistics based algorithm offers a low-complexity channel estimator. Since the superimposed training does not cost an extra bandwidth, the transceiver using superimposed training can be regarded as a kind of self symbol recovery equipment.

4) A novel MIMO-OFDM transceiver is proposed by employing distinct block precoders, i.e., ZP-OFDM and CP-OFDM signals are transmitted from two uncorrelated transmit antennas. This transceiver enables a second-order cyclostationary based blind channel estimator, which can cope with subchannels individually and improve the channel estimate performance.

When the channel state information varies very rapidly, pilot-assisted OFDM transceivers may be a good candidate rather than a blind transceiver [39]-[40]. This thesis also analyzes the performance of existing pilot-assisted channel estimators, and proposes two combined channel estimation approaches to improve the channel estimate performance.

Finally, the carrier frequency offset induced by time-varying channels is also investigated in the presence of unknown frequency selective-fading channels. A novel blind CFO estimation algorithm has been developed. It is established that the CFO parameter is identifiable regardless of unknown channel fading and unknown symbol timing mismatch.

Appendix 1.A: Proof of Theorem 1.4.1

For a FIR channel with length L_u, Chapter 1.2 has shown that $\mathbf{\Gamma}_M$ has at most $L_u - 1$ zero eigenvalues. It means that

$$\text{Rank}\{\mathbf{\Gamma}_M\} \geqslant M - L_u + 1 \tag{1.5.1}$$

It can also be observed that

$$\text{Rank}\{\bar{\mathcal{U}}_\mu\} = \text{Rank}\{\mathcal{U}_\mu\} = P \tag{1.5.2}$$

Hence, (1.4.7) shows that the necessary condition of the identifiability of $\mathbf{s}_\mu(n)$ is

$$\min\{\text{Rank}\{\mathbf{\Gamma}_M\}\} \geqslant \text{Rank}\{\bar{\mathcal{U}}_\mu\} \tag{1.5.3}$$

which leads to

$$L_u - 1 \leqslant M - P \tag{1.5.4}$$

However, this condition is not sufficient for the signal identifiability.

Consider the frequency-domain form of $\mathbf{z}_\mu(n)$, which is given by

$$\begin{aligned} \tilde{\mathbf{z}}_\mu(n) &= \mathbf{F}_M \mathbf{z}_\mu(n) \\ &= \mathcal{D}_M \mathcal{U}_\mu \mathbf{s}_\mu(n) \end{aligned} \tag{1.5.5}$$

The signal recovery is guaranteed when $\text{Rank}\{\mathcal{D}_M \mathcal{U}_\mu\} = P$. Now, split the matrix $\mathcal{D}_M$ into P sub-matrices $\bar{\mathcal{D}}_i$ with size $K_u \times K_u$, and

$$\mathcal{D}_M = \text{diag}\{\bar{\mathcal{D}}_0 \ \cdots \ \bar{\mathcal{D}}_{P-1}\} \tag{1.5.6}$$

The condition $\text{Rank}\{\mathcal{D}_M \mathcal{U}_\mu\} = P$ holds only when

$$\bar{\mathcal{D}}_i \mathbf{w}_\mu \neq \mathbf{0} \tag{1.5.7}$$

Hence, $\bar{\mathcal{D}}_i$ needs to contain at least one non-zero entry along the diagonal, i.e.

$$K_u \geqslant L_u \tag{1.5.8}$$

(1.5.4) and (1.5.8) show that $\mathbf{s}_\mu(n)$ is uniquely identifiable regardless of channel zero locations only when

$$L_u \leqslant \min(M - P + 1,\ K_u) \tag{1.5.9}$$

For $M = K_u \times P$ and $K_u \geqslant L_u$, it can be easily deduced that

$$M - P + 1 = (K_u - 1)P + 1 \geqslant K_u \tag{1.5.10}$$

This theorem is therefore proved.

Appendix 1.B: Proof of Theorem 1.4.2

DFT of (1.4.9) is given by

$$\bar{\mathbf{z}}_{\mathrm{d}}(n) = \mathcal{D}_M \mathcal{U}_{K_a} \mathbf{s}_{M_u}(n) \tag{1.5.11}$$

Similar to the proof of Theorem 1.4.1, (1.5.11) infers that the signal recovery is guaranteed only when

$$\mathrm{Rank}\{\bar{\mathcal{D}}_i \mathbf{W}_{K_a}\} \geqslant K_a \tag{1.5.12}$$

which leads to the result of (1.4.12).

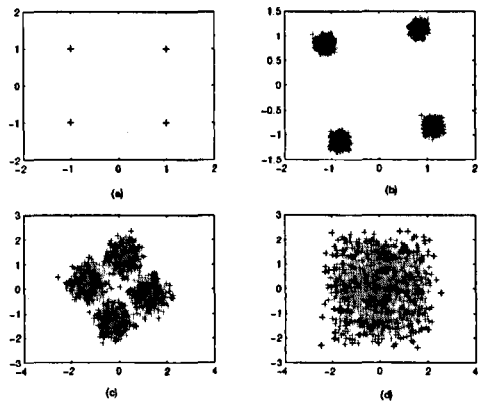

Figure 1.7: An example of ICI with the normalized Doppler shifts ε_d=: a) 0, b) 0.05, c) 0.2, d) 0.5. abscissa: real part of symbols, ordinate: imaginary part of symbols.

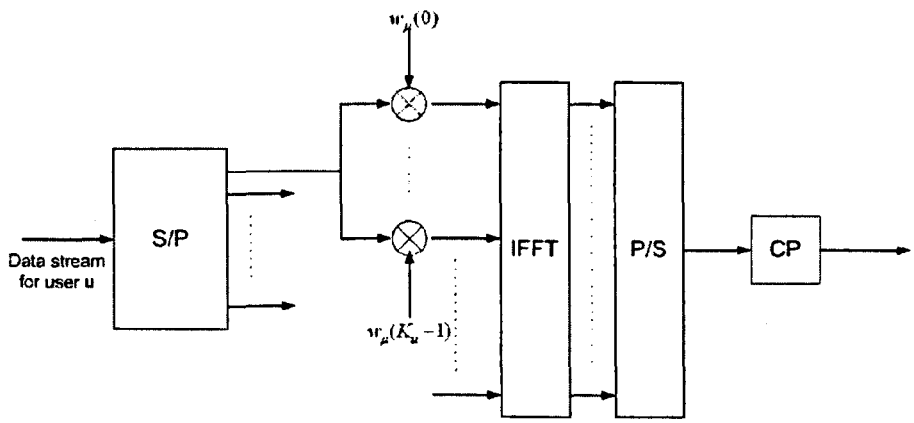

Figure 1.8: A MC-CDMA transmitter for single user.

Chapter 2

Channel estimation techniques for OFDM

Chapter 1 has shown that the channel estimation is very necessary for most of OFDM receivers in the presence of unknown multipath fading. Compared to noncoherent (DPSK [17]) and semi-coherent (differential space-time (ST) modulation [41]-[42]) modems, many papers have shown that, for the uncoded system, a receiver equipped with a channel estimator can improve the overall system performance significantly (e.g. [43]-[44]). This chapter is intended to briefly review channel estimation methods for OFDM systems.

2.1 Understanding of channel estimation

Figure 2.1 depicts the block diagram of channel estimation methods. Usually, channel estimation methods can be classified into two main categories: training based and blind approaches. The training based approach normally uses several training sequences ahead to track a time-varying channel [8]-[9], [49]. When the channel state information (CSI) varies rapidly, training sequences need to be frequently sent to maintain the overall system performance [45]. Alternatively, pilot assisted channel

estimation methods can be used to track the rapidly time-varying channel (see [1], [46]-[48]). These methods have already been considered in many standards, such as DVB [5] and wireless MAN [10]. Recently, blind techniques for identification of linear time-invariant systems have found widespread applications in time series modelling, econometrics, exploration seismology, and equalization of communication channels [16]. With no access to the input, as shown in Figure 2.1 (b), many blind methods have relied on stationary high-order statistics [41], [51]-[53], and cyclostationary or multivariate second-order statistics [21], [30], [54]-[57] of the output data. The blind schemes are important, for example, in digital broadcasting because transmission is not interrupted to train new users entering the cell. In wireless environments, the bandwidth is utilized efficiently when cold start-up is possible and in multipoint data networks throughput increases and management overhead drops when training is obviated [54].

2.2 Training based technique

Figure 2.2 illustrates the block diagram of the training based scheme for OFDM. To track the time-varying channel, training sequences need to be sent periodically. As shown in Figure 2.1 (a), when the input is known to the receiver, the CSI can be

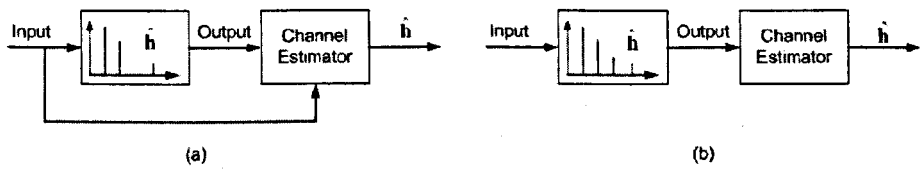

Figure 2.1: Block diagram of channel estimation. (a) training based estimation, (b) blind channel estimation

uniquely identified using well-known least square (LS) [58] and linear minimum mean square error (LMMSE) [49] algorithms. [49] has shown that LMMSE outperforms LS when the channel correlation matrix and noise power are known at the receiver. However, these conditions increase the receiver complexity and limit its application. Later on, we introduce the LS and LMMSE algorithm respectively, and also show their relationship.

2.2.1 LS algorithm and performance analysis

Let $\mathbf{t}(n_t) = [t(n_t, 0), \cdots, t(n_t, M-1)]^T$ be the training block on the n_tth block interval. For the CP-OFDM system, as shown in Section 1.2.2, the DFT-processed received signal (1.2.10) on the n_tth block interval can be expressed as

$$\tilde{\mathbf{z}}_{\text{cp}}(n_t) = \boldsymbol{\mathcal{D}}_M \mathbf{t}(n_t) \tag{2.2.1}$$

by replacing the symbol block $\mathbf{s}(n)$ with the training block $\mathbf{t}(n_t)$. Consider the presence of additive noise $\mathbf{v}(n_t)$, the received signal is given by

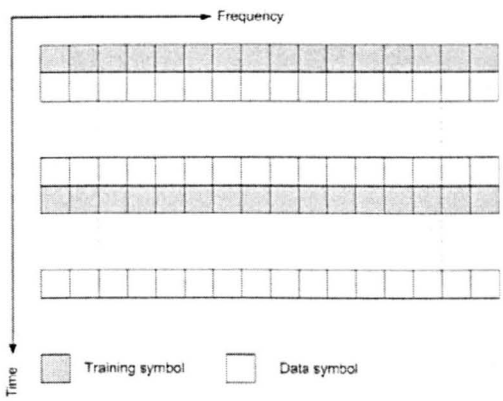

Figure 2.2: Structure of the training based channel estimation scheme.

$$\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) = \mathcal{D}_M \mathbf{t}(n_t) + \mathbf{v}(n_t) \tag{2.2.2}$$

Equivalently, (2.2.2) can be rewritten into

$$\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) = \mathbf{D}(\mathbf{t}(n_t))\tilde{\mathbf{h}} + \mathbf{v}(n_t) \tag{2.2.3}$$

where $\mathbf{D}(\mathbf{t}(n_t))$ stands for the diagonal matrix with the vector $\mathbf{t}(n_t)$ on its diagonal, and $\tilde{\mathbf{h}}$ is the frequency-response vector with size $M \times 1$. Collect the first L_u columns of the DFT matrix $\mathbf{F}_M$ to form a $M \times L_u$ matrix $\mathbf{F}_{\mathbf{h}}$. Eqn. (1.2.6) can be expressed as the following matrix form

$$\tilde{\mathbf{h}} = \sqrt{M}\mathbf{F}_{\mathbf{h}}\mathbf{h} \tag{2.2.4}$$

Substitutes this result into (2.2.3) to obtain that

$$\begin{aligned} \tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) &= \sqrt{M}\mathbf{D}(\mathbf{t}(n_t))\mathbf{F}_{\mathbf{h}}\mathbf{h} + \mathbf{v}(n_t) \\ &= \sqrt{M}\mathbf{Q}_{\mathbf{t}}\mathbf{h} + \mathbf{v}(n_t) \end{aligned} \tag{2.2.5}$$

where $\mathbf{Q}_{\mathbf{t}} = \mathbf{D}(\mathbf{t}(n_t))\mathbf{F}_{\mathbf{h}}$. It can be observed that the matrix $\mathbf{Q}_{\mathbf{t}}$ always has full column rank if the block $\mathbf{t}(n_t)$ contains at least L_u non-zero components. Hence, the CIR $\mathbf{h}$ can be uniquely identified using

$$\begin{aligned} \hat{\mathbf{h}} &= \frac{1}{\sqrt{M}}\mathbf{Q}_{\mathbf{t}}^{\dagger}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \\ &= \mathbf{h} + \frac{1}{\sqrt{M}}\mathbf{Q}_{\mathbf{t}}^{\dagger}\mathbf{v}(n_t) \end{aligned} \tag{2.2.6}$$

Usually, MSE is used to evaluate the channel estimate performance, which is given by

$$\begin{aligned} \epsilon &= \mathcal{E}\|\hat{\mathbf{h}} - \mathbf{h}\|^2 \\ &= \frac{1}{M}\mathrm{Tr}\{\mathbf{Q}_{\mathbf{t}}^{\dagger}\mathcal{E}\{\mathbf{v}(n_t)\mathbf{v}^{\mathcal{H}}(n_t)\}(\mathbf{Q}_{\mathbf{t}}^{\dagger})^{\mathcal{H}}\} \\ &= \frac{L_u}{M^2}\sum_{m=0}^{M-1}\frac{\sigma_{v,m}^2}{|t(n,m)|^2} \end{aligned} \tag{2.2.7}$$

where $\sigma_{v,m}^2$ denotes variance of the noise on the mth subcarrier. From the above analysis, we can find that:

r2.2.1 To guarantee the full column rank of $\mathbf{Q_t}$, the length of training sequence should be large enough ($M \geqslant L_u$).

r2.2.2 (2.2.7) shows that the channel estimation performance is related to the signal-to-noise ratio (SNR). So, in the low SNR (e.g. SNR= 0dB), the estimate performance is not as good as what we expect.

r2.2.3 The MSE, ϵ, is a function of $|t(n,m)|$. So, it is possible to minimize ϵ by designing the training sequences carefully. Previous work [40] has established the optimum training design for the case of white noise. Here, we study the optimum training design by considering the noise to be colour.

Theorem 2.2.1. *If the total energy of the training sequence is given, i.e.*

$$\sum_{m=0}^{M-1} |t(n,m)|^2 = \xi_t \tag{2.2.8}$$

then the MSE, ϵ, achieves its minimum

$$\epsilon_{\min} = \frac{L_u}{M\xi_t} \sum_{m=0}^{M-1} \sigma_{v,m}^2 \tag{2.2.9}$$

when the signal power is equally allocated along the subcarriers regardless of noise colour.

Proof. See Appendix 2.A.

When the noise is white, (2.5.4) shows that the power of the training sequence should be equally allocated for each subcarrier. The channel estimate error is therefore given by

$$\epsilon_{\min} = \frac{L_u \sigma_v^2}{\xi_t} \tag{2.2.10}$$

where $\sigma_v^2 = \sigma_{v,0}^2 = \cdots = \sigma_{v,M-1}^2$.

2.2.2 LS Vs LMMSE

Section 2.2.1 has presented the LS-based channel estimation algorithm by assuming that the upper bound of channel length, L_u, is known at the receiver. If we consider L_u to be the unknown parameter, then the LS approach can only estimate the channel frequency response $\tilde{\mathbf{h}}$ from (2.2.3) as

$$\hat{\tilde{\mathbf{h}}} = \mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \tag{2.2.11}$$

Then, the mean-square channel estimation error is given by

$$\begin{aligned} \mathcal{E}\|\hat{\mathbf{h}} - \mathbf{h}\|^2 &= \mathcal{E}\|\mathbf{F}_M^{-1}(\hat{\tilde{\mathbf{h}}} - \tilde{\mathbf{h}})\|^2 \\ &= \frac{1}{M}\sum_{m=0}^{M-1}\frac{\sigma_{v,m}^2}{|t(n,m)|^2} \end{aligned} \tag{2.2.12}$$

We can see that this MSE result equals to the result of (2.2.7) only when $L_u = M$. However, practical OFDM systems assure $L_u/M \approx 20\%$ (e.g. [4]). So, with the knowledge of L_u, the LS approach is able to improve the channel estimate performance.

The LMMSE algorithm for parameter estimation has been introduced in many textbooks (e.g. [59]). Using the LMMSE algorithm, [49] has proposed a channel estimation method for OFDM systems as

$$\hat{\tilde{\mathbf{h}}} = \mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}(\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} + \sigma_v^2(\mathbf{D}(\mathbf{t}(n_t))\mathbf{D}(\mathbf{t}(n_t))^{\mathcal{H}})^{-1})^{-1}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \tag{2.2.13}$$

where $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = \mathcal{E}\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathcal{H}}\}$ and σ_v^2 is the variance of white noise (i.e. $\sigma_{v,0}^2 = \sigma_{v,1}^2 = \cdots = \sigma_{v,M-1}^2 = \sigma_v^2$). Based on (2.2.13), we study the relationship between LMMSE and LS approaches.

`High SNR or noiseless case:` In this case, noise power is very low, such that

we can let $\sigma_v^2(\mathbf{D}(\mathbf{t}(n_t))\mathbf{D}(\mathbf{t}(n_t))^{\mathcal{H}})^{-1} = 0$. Then, (2.2.13) can be simplified as

$$\hat{\tilde{\mathbf{h}}} = \mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}(\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}})^{-1}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \tag{2.2.14}$$

Because the relationship between $\tilde{\mathbf{h}}$ and $\mathbf{h}$ is given by [60],

$$\tilde{\mathbf{h}} = \sqrt{M}\mathbf{F}_M[\mathbf{h}^T, \mathbf{0}_{1\times(M-L_u)}]^T \tag{2.2.15}$$

the correlation matrix $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ can be derived as

$$\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = M\mathbf{F}_M\mathrm{diag}\{\mathbf{R}_{\mathbf{hh}}, \mathbf{0}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}} \tag{2.2.16}$$

where $\mathbf{R}_{\mathbf{hh}} = \mathcal{E}\{\mathbf{h}\mathbf{h}^{\mathcal{H}}\}$. The inverse of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ is given by

$$\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}^{-1} = \frac{1}{M}\mathbf{F}_M\mathrm{diag}\{\mathbf{R}_{\mathbf{hh}}^{-1}, \mathbf{0}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}} \tag{2.2.17}$$

We plug (2.2.16) and (2.2.17) into (2.2.14) and obtain

$$\begin{aligned}\hat{\tilde{\mathbf{h}}} &= \mathbf{F}_M\mathrm{diag}\{\mathbf{I}_{L_u}, \mathbf{0}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \\ &= \mathbf{F}_M(\mathbf{D}(\mathbf{t}(n_t))\underbrace{\mathbf{F}_M\mathrm{diag}\{\mathbf{I}_{L_u}, \mathbf{0}_{M-L_u}\}}_{=\mathbf{F}_{\mathbf{h}}})^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t)\end{aligned} \tag{2.2.18}$$

Then, the channel impulse response $\mathbf{h}$ can be identified by

$$\begin{aligned}\hat{\mathbf{h}} &= \frac{1}{\sqrt{M}}\mathbf{F}_{\mathbf{h}}^{\dagger}\hat{\tilde{\mathbf{h}}} \\ &= \frac{1}{\sqrt{M}}(\mathbf{D}(\mathbf{t}(n_t))\mathbf{F}_{\mathbf{h}})^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t)\end{aligned} \tag{2.2.19}$$

This equation is the same as the LS-based approach (2.2.6). So, we can conclude that, in the high SNR (or noiseless) case, LMMSE and LS approaches lead to the same result if L_u is known at the receiver.

Using the optimum training sequence: The training sequence is designed according to Theorem 2.2.1, i.e. $\mathbf{D}(\mathbf{t}(n_t))\mathbf{D}(\mathbf{t}(n_t))^{\mathcal{H}} = \frac{\xi_t}{M}\mathbf{I}_M$. Let $\beta = \sigma_v^2 M/\xi_t$, (2.2.13) becomes

$$\hat{\tilde{\mathbf{h}}} = \mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}(\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} + \beta\mathbf{I}_M)^{-1}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \tag{2.2.20}$$

Because $\beta\mathbf{I}_M = M\mathbf{F}_M\frac{\beta}{M}\mathbf{I}_M\mathbf{F}_M^{\mathcal{H}}$, we can rewrite (2.2.20) into

$$\begin{aligned}\hat{\tilde{\mathbf{h}}} &= \frac{1}{M}\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}(\mathbf{F}_M\mathrm{diag}\{(\mathbf{R}_{\mathbf{hh}} + \beta/M)^{-1}, M/\beta\mathbf{I}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}})\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \\ &= \mathbf{F}_M\mathrm{diag}\{\mathbf{R}_{\mathbf{hh}}(\mathbf{R}_{\mathbf{hh}} + \beta/M)^{-1}, \mathbf{0}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \end{aligned} \tag{2.2.21}$$

Then, the channel impulse response $\hat{\mathbf{h}}$ can be obtained as

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{M}}\mathbf{F}_{\mathbf{h}}^{\dagger}\mathbf{F}_M\mathrm{diag}\{\mathbf{R}_{\mathbf{hh}}(\mathbf{R}_{\mathbf{hh}} + \beta/M)^{-1}, \mathbf{0}_{M-L_u}\}\mathbf{F}_M^{\mathcal{H}}\mathbf{D}(\mathbf{t}(n_t))^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t) \tag{2.2.22}$$

Due to $\mathbf{F}_{\mathbf{h}}^{\dagger}\mathbf{F}_M = \mathrm{diag}\{\mathbf{I}_{L_u}, \mathbf{0}_{M-L_u}\}$, we can finally obtain the following result

$$\hat{\mathbf{h}} = \mathbf{R}_{\mathbf{hh}}(\mathbf{R}_{\mathbf{hh}} + \beta/M)^{-1}\underbrace{\frac{1}{\sqrt{M}}(\mathbf{D}(\mathbf{t}(n_t))\mathbf{F}_{\mathbf{h}})^{-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n_t)}_{\text{LS approach}} \tag{2.2.23}$$

Since channel taps are random variables and uncorrelated each other, $\mathbf{R}_{\mathbf{hh}}$ is a diagonal matrix. Denoting $\sigma_{h,i}^2$ to be variance of the i tap and $\hat{\mathbf{h}}_{\mathrm{LS}}$ to be the estimated channel using the LS approach, (2.2.23) can be simplified as

$$\hat{\mathbf{h}} = \mathrm{diag}\{\frac{\sigma_{h,0}^2}{\sigma_{h,0}^2 + \beta/M}, \cdots, \frac{\sigma_{h,L_u-1}^2}{\sigma_{h,L_u-1}^2 + \beta/M}\}\hat{\mathbf{h}}_{\mathrm{LS}} \tag{2.2.24}$$

This result shows that $\hat{\mathbf{h}}$ and $\hat{\mathbf{h}}_{\mathrm{LS}}$ should have the same phase for each tap. Also, the LMMSE and LS approaches lead to the same result when the SNR is high (i.e. $\beta \approx 0$). Then, we examine the channel estimation error of the LMMSE approach.

Firstly, $\hat{\mathbf{h}}_{\mathrm{LS}}$ can be expressed as

$$\hat{\mathbf{h}}_{\mathrm{LS}} = \mathbf{h} + \mathbf{e}_{\mathrm{LS}} \tag{2.2.25}$$

where $\mathbf{e}_{\mathrm{LS}}$ is the channel estimation error for the LS approach. The variance of $\mathbf{e}_{\mathrm{LS}}$ has been given in (2.2.10). Let $\mathbf{\Lambda} = \mathrm{diag}\{\frac{\sigma^2_{h,0}}{\sigma^2_{h,0}+\beta/M}, \cdots, \frac{\sigma^2_{h,L_u-1}}{\sigma^2_{h,L_u-1}+\beta/M}\}$, the mean-square channel estimation error of the LMMSE approach is given by

$$\begin{aligned} \mathcal{E}\|\hat{\mathbf{h}} - \mathbf{h}\|^2 &= \mathcal{E}\|(\mathbf{\Lambda} - 1)\mathbf{h} + \mathbf{\Lambda}\mathbf{e}_{\mathrm{LS}}\|^2 \\ &= \mathrm{Tr}\{(1 - \mathbf{\Lambda})^2\mathbf{R}_{\mathbf{hh}}\} + \mathcal{E}\|\mathbf{\Lambda}\mathbf{e}_{\mathrm{LS}}\|^2 \end{aligned} \tag{2.2.26}$$

Differed from the LS approach, (2.2.26) shows that the channel estimation error for the LMMSE approach is related to the channel delay profile. This relationship becomes weak with the increase of SNR ($(1 - \mathbf{\Lambda}) \rightarrow 0$).

In order to describe performances for the LS and LMMSE approaches visually, we plot the theoretical MSE results in Figure 2.3 by employing two channel models. The first channel model is called *lab channel model*. Each channel tap is randomly generated with the variance of $1/L_u$. The second channel model is generated according

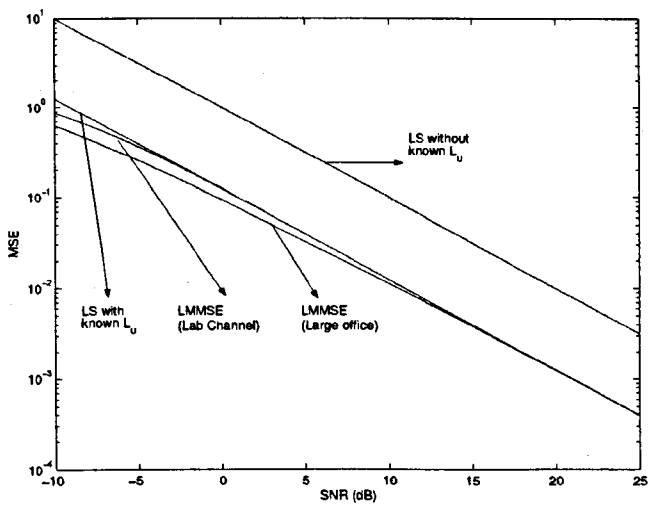

Figure 2.3: Channel estimation performance: LS Vs LMMSE.

to the measurement result of large office environment (approved by ETSI [67]). Figure 2.3 shows that the LS approach without knowledge of L_u is the worst among all approaches. Because the channel correlation matrix $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ contains the information of L_u, it is unfair to compare the LMMSE approach with the LS approach without the knowledge of L_u. For the fairness, we only compare the LMMSE approach with the LS approach with the knowledge of L_u. Figure 2.3 shows that the LMMSE approach is sensitive to the channel delay profile. Using the lab channel model, the LMMSE and LS approaches have almost the same MSE performance for SNR$\geqslant$ 0dB. Even though small difference exists for SNR$<$ 0dB, the MSE results are too large to be acceptable for practical applications. Using the large-office channel model, we can see the LMMSE approach outperforms the LS approach for SNR$<$ 11dB. The LMMSE approach pays more computation complexity, but does not improve the performance significantly. For the typical SNR (= 12dB), the performances for the LMMSE and LS approaches are almost the same.

2.3 Pilot assisted technique

When the channel varies rapidly, training sequences have to be sent frequently, because the channel estimate error increases considerably with increasing the training interval. In the broadcasting networks [5] and wireless MAN [10], a pilot assisted scheme is employed to track the rapidly time-varying channel.

Figure 2.4 illustrates an example of pilot assisted scheme for OFDM systems, where pilot symbols are uniformly arranged on DFT grids. [40] has shown that this kind of pilot arrangement is optimum in AWGN. This section is intended to briefly review the pilot assisted scheme.

Starting from (1.2.9), the received signal of the pilot assisted OFDM can be expressed as

$$\begin{aligned} \mathbf{z}_{\mathrm{cp}}(n) &= \mathbf{\Gamma}_M \mathbf{F}_M^{\mathcal{H}} \mathbf{s}(n) + \mathbf{v}(n) \\ &= \mathbf{\Gamma}_M \mathbf{F}_{\mathrm{d}}^{\mathcal{H}} \mathbf{s}_{\mathrm{d}}(n) + \mathbf{\Gamma}_M \mathbf{F}_{\mathrm{p}}^{\mathcal{H}} \mathbf{s}_{\mathrm{p}}(n) + \mathbf{v}(n) \end{aligned} \tag{2.3.1}$$

where $\mathbf{s}_{\mathrm{d}}(n)$ stands for a $M_d \times 1$ data block, $\mathbf{s}_{\mathrm{p}}(n)$ denotes a $M_p \times 1$ pilot block, $(M = M_d + M_p)$; The matrix $\mathbf{F}_{\mathrm{p}}$ is formed by uniformly collecting M_p rows from the DFT matrix $\mathbf{F}_M$, and the rest rows of $\mathbf{F}_M$ is used to construct $\mathbf{F}_{\mathrm{d}}$. Then, DFT of $\mathbf{z}_{\mathrm{cp}}(n)$ is given by

$$\tilde{\mathbf{z}}_{\mathrm{cp}}(n) = \underbrace{\mathcal{D}_M \mathbf{F}_{\mathrm{M}} \mathbf{F}_{\mathrm{d}}^{\mathcal{H}} \mathbf{s}_{\mathrm{d}}(n)}_{\mathrm{PartI}} + \underbrace{\mathcal{D}_M \mathbf{F}_{\mathrm{M}} \mathbf{F}_{\mathrm{p}}^{\mathcal{H}} \mathbf{s}_{\mathrm{p}}(n)}_{\mathrm{PartII}} + \tilde{\mathbf{v}}(n) \tag{2.3.2}$$

where $\tilde{\mathbf{v}}(n) = \mathbf{F}_M \mathbf{v}(n)$. Since Part I and Part II in (2.3.2) are orthogonal, the pilot information can be abstracted from the received signal as [1]

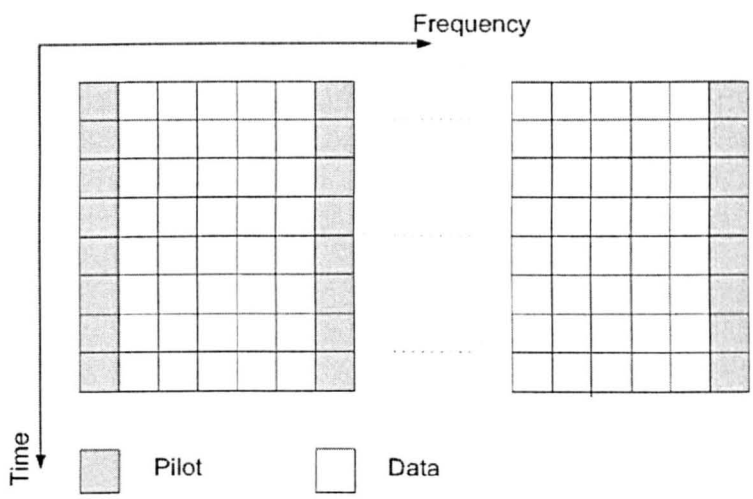

Figure 2.4: Structure of the pilot assisted channel estimation scheme.

$$\tilde{\mathbf{z}}_{\mathrm{p}}(n) = \mathcal{D}_{\mathrm{p}}\mathbf{s}_p(n) + \tilde{\mathbf{v}}_{\mathrm{p}}(n) \tag{2.3.3}$$

where $\tilde{\mathbf{z}}_{\mathrm{p}}(n)$, $\mathcal{D}_{\mathrm{p}}$ and $\mathbf{v}_{\mathrm{p}}(n)$ are formed by drawing the elements, corresponding to the pilot subcarriers, from $\tilde{\mathbf{z}}_{\mathrm{cp}}(n)$, $\mathcal{D}_M$ and $\tilde{\mathbf{v}}(n)$, respectively. Detailed contents can be found in [40]. Constructing $\bar{\mathbf{F}}_{\mathrm{p}}$ by collecting the first L_u columns of $\mathbf{F}_{\mathrm{p}}$, (2.3.3) can be represented as

$$\begin{aligned} \tilde{\mathbf{z}}_{\mathrm{p}}(n) &= \mathbf{D}(\mathbf{s}_p(n))\bar{\mathbf{F}}_{\mathrm{p}}\mathbf{h} + \tilde{\mathbf{v}}_{\mathrm{p}}(n) \\ &= \mathbf{Q}_{\mathrm{p}}\mathbf{h} + \tilde{\mathbf{v}}_{\mathrm{p}}(n) \end{aligned} \tag{2.3.4}$$

where $\mathbf{Q}_{\mathrm{p}} = \mathbf{D}(\mathbf{s}_p(n))\bar{\mathbf{F}}_{\mathrm{p}}$. Similar to the training based approach, the channel identifiability is guaranteed only when $\mathbf{Q}_{\mathrm{p}}$ has full column rank, and the CSI can be obtained as

$$\hat{\mathbf{h}} = \mathbf{Q}_{\mathrm{p}}^{\dagger}\tilde{\mathbf{z}}_{\mathrm{p}}(n) \tag{2.3.5}$$

The analysis of channel estimate performance may follow the way for the training based scheme, and the same conclusion can be drawn.

2.4 Blind channel estimation technique

Unlike the training based techniques, blind channel estimation methods cannot be unified to one solution. The blind estimator is designed only for a specified modulation scheme. As illustrated in Figure 2.1 (b), the estimator does not need the knowledge of the input data, but it needs to know the format of the input.

Originally, many blind channel estimators are based on the high-order statistics algorithms [50]-[52]. To offer a relatively fast convergence rate, second-order statistics based algorithms have been widely used to identify the FIR channel. Most of them

exploit the cyclostationarity of the output signal [13], [57] and [61]. The cyclostationarity can be introduced by many factors, for example, the pulse shaping filter [31], [62], [63], CPs [57], and the time domain precoding [30], [61]. The major advantages of the cyclostationary based approach are as follows: 1) the channel identifiability is guaranteed regardless of channel zero locations; 2) The present additive Gaussian noise is not necessarily white; 3) Some approaches using the pulse shaping filter or time-domain precoding does not rely on the coding redundancy. However, the existing approaches have their own drawbacks. The detail contents can be found in the above literature.

Another popularly used blind method is based on the subspace decomposition algorithm [21], [56], [64], [65]. The philosophy behind this method is that the signal subspace is orthogonal to the noise subspace [17], [19]. For a time-invariant or a slowly time-varying channel, subspace based approaches have demonstrated their excellent capability [64]. On the other hand, the subspace based approach has its drawbacks, such as: 1) this method needs singular value decomposition, which has relatively high computation complexity [65]; 2) the number of blocks collected for the channel estimation needs to be large enough to assure the full column rank of the autocorrelation matrix, otherwise, the channel identifiability cannot be guaranteed [64]. 3) the present Gaussian noise should be white.

Recently, a number of semi-blind algorithms have been proposed to improve the capability of the blind algorithms (e.g. [65]). This method is effective when the number of blocks available for the channel estimation are not large enough. To offer a relatively fast convergence rate, a training sequence is employed to obtain the initial channel estimate. The initial estimate is then updated in the blind estimation process.

Since the semi-blind approach can be regarded as an enhanced version of the blind methods, this thesis does not give a detailed presentation.

2.5 Conclusion

This chapter has presented various channel estimation methods for communication systems. It has also briefly reviewed existing channel estimators for OFDM systems and their applications. Some necessary performance analysis has been made for those estimators. From the next chapter, a number of novel channel estimation algorithms will be proposed for an OFDM receiver. Those estimators should improve the overall OFDM system performance in the presence of unknown multipath fading. Some new transmission techniques will also be developed to enable the proposed channel estimators.

Appendix 2.A: Proof of Theorem 2.2.1

Use Cauchy inequality, the following result can be deduced from (2.2.7),

$$\begin{aligned} \epsilon &= \frac{L_u}{M^2} \sum_{m=0}^{M-1} \frac{\sigma_{v,m}^2}{|t(n,m)|^2} \\ &\geqslant \frac{L_u}{M} \frac{\sqrt[M]{\prod_{m=0}^{M-1} \sigma_{v,m}^2}}{\sqrt[M]{\prod_{m=0}^{M-1} |t(n,m)|^2}} \end{aligned} \tag{2.5.1}$$

The equality reaches when

$$\frac{\sigma_{v,0}^2}{|t(n,0)|^2} = \cdots = \frac{\sigma_{v,M-1}^2}{|t(n,M-1)|^2} = \frac{\sum_{m=0}^{M-1} \sigma_{v,m}^2}{\sum_{m=0}^{M-1} |t(n,0)|^2} \tag{2.5.2}$$

Then, the minimum of ϵ can be derived from (2.5.1) and (2.5.2)

$$\epsilon_{\min} = \frac{L_u}{M\xi_t} \sum_{m=0}^{M-1} \sigma_{v,m}^2 \tag{2.5.3}$$

If the variance $\sigma_{v,m}^2$ is known, (2.5.2) also infers the power allocation strategy for the training sequence as

$$|t(n,m)|^2 = \xi_t \frac{\sigma_{v,m}^2}{\sum_{m=0}^{M-1} \sigma_{v,m}^2} \tag{2.5.4}$$

But, (2.5.2) is not the necessary condition to minimize MSE. If the power of training sequence is equally allocated along the subcarriers, i.e.,

$$|t(n,0)|^2 = \cdots = |t(n,M-1)|^2 = \xi_t/M \tag{2.5.5}$$

(2.5.3) can also be achieved by plugging (2.5.5) into (2.2.7). Theorem 2.2.1 is therefore proved.

Chapter 3

Exploiting subcarrier correlation for blind channel estimation (Part I): RC-OFDM

In the presence of unknown multipath fading, previous chapters have shown that most of OFDM transmission systems need a channel estimator to guarantee the signal recovery and improve the overall system performance. It has also been shown that a blind channel estimator could be very useful when the cold start-up occurs in wireless environments, or in where the training is obviated to improve the bandwidth efficiency.

In this chapter and Chapter 4, several blind channel estimation methods are proposed for various OFDM transmission schemes, e.g. repetition coded (RC) OFDM [29] and block precoded (BP) OFDM [22], [25]. These OFDM schemes have a common feature, i.e. part or all of subcarriers are correlated. Based on this feature, the CIR and the channel frequency response are blindly identifiable at the receiver.

The contents, *exploiting subcarrier correlation for blind channel estimation*, are divided into two parts. Part I presents blind channel estimation methods for RC-OFDM systems. A new RC-OFDM transmission scheme is also proposed to improve the overall system performance. Part II introduces a novel blind channel estimation

method for BP-OFDM systems.

3.1 Blind channel estimation for RC-OFDM

3.1.1 Introduction to RC-OFDM

The RC-OFDM transmission is originally proposed for an ICI self-cancellation scheme [29], where the output of the OFDM modulator can be expressed by

$$\mathbf{x}(n) = \mathbf{\Theta}_{\mathrm{cp}}[\mathbf{s}(n) \otimes \mathbf{w}_{2\times 1}] \tag{3.1.1}$$

where $\mathbf{w}_{2\times 1} = [1, \ -1]^T$ and $\otimes$ stands for the Kronecker product. At the presence of normalized frequency offset ε_d, the received signal on subcarrier m is given by [29]

$$\tilde{y}(n,m) = \sum_{m'=0}^{M/2-1} s(n,m')[\gamma(m'-m) - \gamma(m'+1-m)] + v(n,m) \tag{3.1.2}$$

and on subcarrier $m+1$ is

$$\tilde{y}(n,m+1) = \sum_{m'=0}^{M/2-1} s(n,m')[\gamma(m'-m-1) - \gamma(m'-m)] + v(n,m+1) \tag{3.1.3}$$

where $\gamma(m'-m)$ is defined as the ICI coefficient, which can be expressed as

$$\begin{aligned} \gamma(m'-m) &= \frac{\sin(\pi(m'+\varepsilon_d-m))}{M\sin(\frac{\pi}{M}(m'+\varepsilon_d-m))} \\ &\quad \cdot \exp(j\pi(1-\frac{1}{N})(m'-m+\varepsilon_d)) \end{aligned} \tag{3.1.4}$$

Since the difference between $\gamma(m'-m)$ and $\gamma(m'+1-m)$ is very small, [29] has shown that the coefficient $\gamma'(m'-m) = \gamma(m'-m) - \gamma(m'+1-m)$ is much smaller than the ICI coefficient $\gamma(m'-m)$, such that the SIR can be significantly improved (more than 10 dB improvement).

When the multipath fading is considered, the received signal on subcarrier m can be rewritten into

$$\begin{aligned}\tilde{y}(n,m) &= \sum_{m'=0}^{M/2-1} s(n,m')[H(m'-m)\gamma(m'-m) \\ &\quad -H(m'+1-m)\gamma(m'+1-m)] + v(n,m) \end{aligned} \tag{3.1.5}$$

Then, the coefficient $\gamma'(m'-m)$ is given by

$$\gamma'(m'-m) = H(m'-m)\gamma(m'-m) - H(m'+1-m)\gamma(m'+1-m) \tag{3.1.6}$$

It can be shown that SIR can be improved when the difference between two adjacent frequency responses are very small, i.e. $H(m) \approx H(m+1)$. It means that the ICI self-cancelling scheme works only when the subcarrier spacing Δf is small enough. In other words, this scheme is not applicable for broadband OFDM transmission, for instance WLAN.

3.1.2 Blind channel estimation

The RC-OFDM model (3.1.1) shows that each information symbol is loaded onto two subcarriers, such that the bandwidth efficiency is relatively low. Using the training-based channel estimation method, the bandwidth efficiency may be further reduced. So this section proposes a blind channel estimation method for the RC-OFDM transmission in the slowly time-varying channel. Part of this content has been presented in [83].

Start from the received signal model (3.1.5), which can be rewritten into

$$\begin{aligned} y(n,2m) &= s(n,m)(H(2m)\gamma(0) - H(2m+1)\gamma(1)) \\ &\quad +v_{\mathrm{ICI}}(n,2m) + v(n,2m) \end{aligned} \tag{3.1.7}$$

$$\begin{aligned} y(n,2m+1) &= s(n,m)(H(2m)\gamma(1) - H(2m+1)\gamma(0)) \\ &\quad + v_{\mathrm{ICI}}(n,2m+1) + v(n,2m+1) \end{aligned} \tag{3.1.8}$$

where v_{ICI} denotes the ICI part. As shown in Figure 1.7, v_{ICI} can be regarded as an additive noise with zero mean. For $H(2m) \approx H(2m+1)$, the signal models (3.1.7) and (3.1.8) can be further expressed as

$$y(n,2m) = s(n,m)H(2m)\gamma'(0) + v_{\mathrm{ICI}}(n,2m) + v(n,2m) \tag{3.1.9}$$

$$y(n,2m+1) = -s(n,m)H(2m+1)\gamma'(0) + v_{\mathrm{ICI}}(n,2m+1) + v(n,2m+1) \tag{3.1.10}$$

In a slowly time-varying channel, i.e. ε_d is small, it can be derived from (3.1.4) that the coefficient $\gamma'(0)$ is very close to 1. The correlation between subcarriers $2m$ and $2m+1$ is given by

$$\begin{aligned} C_y(2m) &= \mathcal{E}\{y(n,2m)y^*(n,2m+1)\} \\ &= -\sigma_s^2|\gamma'(0)|^2 H(2m)H^*(2m+1) + C_{\mathrm{ICI}}(2m) \end{aligned} \tag{3.1.11}$$

where

$$C_{\mathrm{ICI}}(2m) = \mathcal{E}\{v_{\mathrm{ICI}}(n,2m)v_{\mathrm{ICI}}^*(n,2m+1)\} \tag{3.1.12}$$

and σ_s^2 denotes the variance of information-bearing symbols. Previous subsection has shown that the ICI power can be significantly reduced by employing RC-OFDM, therefore the following result is true

$$\frac{C_{\mathrm{ICI}}(2m)}{\sigma_s^2} \approx 0 \tag{3.1.13}$$

If the information-bearing symbols are drawn from a finite alphabet, then the receiver has the knowledge of σ_s^2, and the following result can be derived

$$\begin{aligned} \tilde{\mathcal{R}}_y(2m) &= -\tilde{C}_y(m)/\sigma_s^2 \\ &= |\gamma'(0)|^2 H(2m)H^*(2m+1) \end{aligned} \tag{3.1.14}$$

Let $\bar{H}(m) = |\gamma'(0)|H(m)$, the coefficient $|\gamma'(0)|^2$ can be absorbed into the frequency response, such that,

$$\tilde{\mathcal{R}}_y(2m) = \bar{H}(2m)\bar{H}^*(2m+1) \tag{3.1.15}$$

If the number of subcarriers $M \geqslant 2L_u$, then an ideal interpolation can be used to estimate $\tilde{\mathcal{R}}_y(m)$ for $m \in [0,\ M-1]$ [40], [83]. The interpolation algorithm can be implemented using

$$\begin{aligned} \tilde{\mathcal{R}}_y(m) &= \frac{2}{M}\sum_{m}^{M-1}\sum_{m_1=0}^{M/2-1}\tilde{\mathcal{R}}_y(2m_1)W_M^{-2lm_1}W_M^{lm} \\ &= \bar{H}(m)\bar{H}^*(m+1) \end{aligned} \tag{3.1.16}$$

then, the autocorrelation for each subcarrier is given by

$$\begin{aligned} C'_y(m) &= \mathcal{E}\{y(m)y^*(m)\} \\ &= \sigma_s^2|\bar{H}(m)|^2 + \sigma_v^2 \end{aligned} \tag{3.1.17}$$

Assuming that the variance σ_v^2 is known, then we define

$$\tilde{\mathcal{R}}'_y(m) = (C'_y(m) - \sigma_v^2)/\sigma_s^2 \tag{3.1.18}$$

and the channel frequency-response can be obtained by

$$\hat{\bar{H}}(0) = \sqrt{\tilde{\mathcal{R}}'_y(0)} \cdot \exp(j\phi_0) \tag{3.1.19}$$

$$\hat{\bar{H}}(1) = \frac{\tilde{\mathcal{R}}_y(1)}{\hat{\bar{H}}(0)} \tag{3.1.20}$$

$$\hat{\bar{H}}(m) = \frac{\prod_{m'=0}^{m-1}\tilde{\mathcal{R}}_y(m')}{\hat{\bar{H}}(0)\prod_{m'=1}^{m-1}\tilde{\mathcal{R}}'_y(m')}, \text{ for } m \geqslant 2 \tag{3.1.21}$$

Remark 1: Like most of blind channel estimation algorithms, for instance [21], the proposed estimator can uniquely identify the CSI with a phase ambiguity ϕ_0. This ambiguity can be resolved using several pilot symbols.

Remark 2: This estimator requires the knowledge of σ_v^2, which can be easily obtained using the time-domain autocorrelation [57].

Remark 3: The proposed equations (3.1.19)-(3.1.21) are prone to error propagation. The estimation performance is the best when $|H(0)|$ becomes the largest among all subcarriers. To improve the performance, this estimator can be modified and summarized based on the following steps:

Step 1: Find the subcarrier $m_{\max}$ using

$$m_{\max} = \max(C'_y(m))^1, \text{ for } m \in [0, M-1] \tag{3.1.22}$$

Step 2: Estimate the frequency response $H(m_{\max})$ using

$$\hat{\tilde{H}}(m_{\max}) = \sqrt{\tilde{\mathcal{R}}'_y(m_{\max})} \cdot \exp(j\phi_0) \tag{3.1.23}$$

$$\hat{\tilde{H}}(m_{\max}+1) = \frac{\tilde{\mathcal{R}}_y(m_{\max}+1)}{\hat{\tilde{H}}(m_{\max})} \tag{3.1.24}$$

Step 3: Obtain the channel frequency response using

$$\hat{\tilde{H}}(m_{\max}+m) = \frac{\prod_{m'=0}^{m-1} \tilde{\mathcal{R}}_y(m_{\max}+m')}{\hat{\tilde{H}}(m_{\max}) \prod_{m'=1}^{m-1} \tilde{\mathcal{R}}'_y(m_{\max}+m')}, \text{for } m \geqslant 2 \tag{3.1.25}$$

[1] $\max(f(i))$ is to find the index $i_{\max}$ corresponding to the maximum of $f(i)$.

3.2 Differential RC-OFDM and semi-coherent detection

Using the ideal interpolation, (3.1.16) shows that the partial channel information can be blindly acquired at the receiver. If the information-bearing symbols can be recovered only using the partial channel information, then the channel estimation and equalization complexity can be reduced. This motivates us to propose the differential RC-OFDM modulation scheme and the semi-coherent detection technique.

3.2.1 Differential RC-OFDM modulation scheme

If each subcarrier is modulated using the phase shift key (PSK), then the temporal differential modulation scheme can be employed in OFDM systems [21], [68]. This section proposes a differential modulation scheme for RC-OFDM by exploiting the frequency diversity.

Recall the transmitted RC-OFDM signal model

$$\mathbf{x}(n) = \mathbf{\Theta}_{\mathrm{cp}}[\mathbf{s}(n) \otimes \mathbf{w}_{2\times 1}] \tag{3.2.1}$$

If the information-bearing symbols are PSK modulated [2], using the differential modulation scheme, the symbol blocks $\mathbf{s}(n)$ are updated by a new block $\mathbf{s}_d(n)$, whose mth element is given by

$$s_d(n,0) = s(n,0) \tag{3.2.2}$$

$$s_d(n,m) = s^*(n,m)s(n,m-1) \tag{3.2.3}$$

We can see that $\mathcal{E}\{s_d(n,m_1)s_d^*(n,m_2)\} = 0$. $\mathbf{s}_d(n)$ and $\mathbf{s}(n)$ have the same statistical property. Following the analysis in Chapter 3.1.1, it is easy to find that the differential RC-OFDM scheme is also an ICI self-cancelling scheme.

[2]The amplitude of PSK symbol is normalized.

Certainly, the differential scheme has its distinct feature, which enables the semi-coherent detection, i.e.

$$\begin{aligned} s_d(n,m)\cdot s_d^*(n,m+1) &= |s_d(n,m)|^2\cdot s(n,m+1) \\ &= s(n,m+1) \end{aligned} \tag{3.2.4}$$

3.2.2 Semi-coherent detection

The motivation of proposing the differential RC-OFDM scheme is to enable the semi-coherent detection technique by employing the partial channel knowledge obtained by (3.1.16). Similar to the RC-OFDM system, the received signal of the differential scheme can be expressed as

$$y(n,2m) = s_d(n,m)H(2m)\gamma'(0) + v_{\text{ICI}}(n,2m) + v(n,2m) \tag{3.2.5}$$

$$y(n,2m+1) = -s_d(n,m)H(2m+1)\gamma'(0) + v_{\text{ICI}}(n,2m+1) + v(n,2m+1) \tag{3.2.6}$$

Then, the product of $y(n,2m+1)$ and $y^*(n,2m+2)$ is given by

$$\begin{aligned} \mathcal{P}(n,m+1) &= y(n,2m+1)\cdot y^*(n,2m+2) \\ &= -s_d(n,m)s_d^*(n,m+1)\bar{H}(2m+1)\bar{H}(2m+2) \\ &\quad +v_{\mathcal{P}}(n,m+1) \\ &= -s(n,m+1)\underbrace{\bar{H}(2m+1)\bar{H}(2m+2)}_{\tilde{R}_y(m)} + v_{\mathcal{P}}(n,m+1) \end{aligned} \tag{3.2.7}$$

where $v_{\mathcal{P}}(n,m+1)$ is the interference term with zero mean. With the knowledge of $\tilde{R}_y(m)$, we can use the following equation to perform the partial channel equalization

$$\begin{aligned} \hat{s}(n,m+1) &= -\frac{\mathcal{P}(n,m+1)}{\tilde{R}_y(2m+1)} \\ &= s(n,m+1) - \underbrace{\frac{v_{\mathcal{P}}(n,m+1)}{\tilde{R}_y(2m+1)}}_{\text{Interference}} \end{aligned} \tag{3.2.8}$$

Remark 1: The differential RC-OFDM modulation scheme enables the semi-coherent detection method. It does not need full CSI and therefore simplifies the blind channel estimator.

Remark 2: For PSK symbols, it is easy to find that $|s(n,m)|^2 = |s_d(n,m)|^2 = \sigma_s^2$. Therefore, (3.1.16) and (3.1.21) show that the CSI and the partial CSI can be identified from one OFDM block.

Remark 3: (3.2.8) shows that $s(n,0)$ cannot be directly obtained using the semi-coherent detection. A simple way to solve this problem is to map pilot symbols on subcarrier 0.

Remark 4: Since the symbol detection is based on (3.2.8), the interference term $v_P(n,m)$ may affect the overall system performance considerably when the ICI and the additive noise are large. To improve the performance, the decision directed (DD) algorithm [42], [69] can be employed. The DD algorithm can be summarized as the following steps.

Step 1: Based on the initial estimate $\tilde{R}_y(m)$, the semi-coherent detection (3.2.8) can be carried out.

Step 2: For the ith iteration, full channel estimate $\hat{\mathbf{h}}^{(i)}$ is obtained by treating the estimated symbols as known symbols. Then, the channel frequency response $\hat{\tilde{H}}^{(i)}(m)$ can be calculated using DFT of $\hat{\mathbf{h}}^{(i)}$. $\hat{\tilde{H}}^{(i)}(m)$ is then used for equalization and constructing the estimate $\hat{\tilde{R}}_y^{(i)}(m) = \hat{\tilde{H}}^{(i)}(m)(\hat{\tilde{H}}^{(i)}(m+1))^*$.

Step 3: Repeat Step 2 until the Euclidean distance $|\tilde{R}_y(m) - \hat{\tilde{R}}_y^{(i)}(m)|$ is minimized.

Using the DD algorithm, $s(n,0)$ is not necessary known, and the full CSI can be obtained without any ambiguity. However, the DD algorithm diverges easily at low

SNR because symbol by symbol detection has poor performance. The smallest bit-error-rate (BER) will be that corresponding to known channels. In the high SNR case, the DD algorithm converges to the maximum likelihood solution of $|\tilde{R}_y(m) - \hat{\tilde{R}}_y^{(i)}(m)|$.

3.3 Simulations for RC-OFDM in the unknown multipath

In the simulations, the bit-error-rate (BER) is used to benchmark the overall system performance of RC-OFDM systems in the unknown multipath fading channel. In order to reflect the channel estimation performance, the RC-OFDM system does not employ the channel code. The system parameters are set as $L_u = 2$, $L_{\text{cp}} = 2$, $M = 32$, which is the typical setting of an OFDM-based cellular radio system. The information-bearing symbols are drawn from a QPSK constellation. The RC-OFDM scheme employs one pair of symbols for each block. The SNR is defined as the average received symbol energy to noise ratio σ_s^2/σ_v^2. The simulations test the RC-OFDM and differential RC-OFDM systems in random and time-varying channels.

`Test Case 1 (static random channels)`: Each tap of the FIR channel is randomly generated with variance $1/L_u$. The channels do not vary within one block duration, and therefore there is no ICI. The CIRs for different blocks are independent of each other. We perform the blind channel estimation based on each OFDM block. The simulation results are averaged over 600 channels. Figure 3.1 illustrates the BER performance of RC-OFDM and differential RC-OFDM systems for the SNR range from 0 dB to 25 dB. In order to compare with state-of-the-art method, we also plot the finite-alphabet (FA) based approach ([66]) in the figure. We can see that the BER performances between FA and differential RC-OFDM approaches are quite

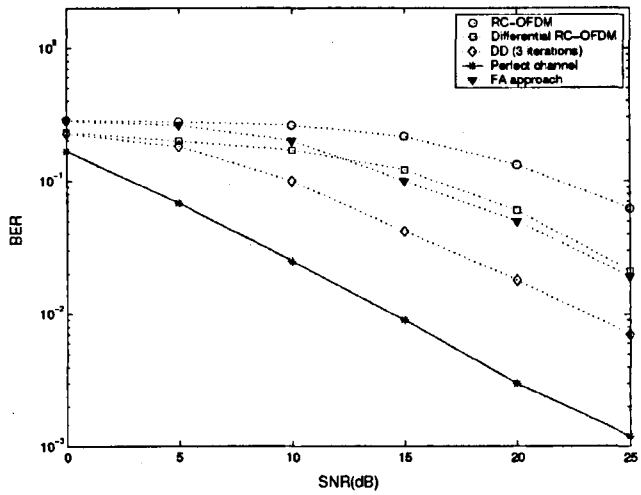

Figure 3.1: BER Vs SNR for RC-OFDM in random channels.

close. The differential RC-OFDM outperforms the RC-OFDM by using the semi-coherent detection technique. Its performance can be further improved by employing the decision-directed method. Since the channel estimate is blindly obtained only from one OFDM block, the result is not accurate enough, and thereby reduces the overall system performance.

`Test Case 2 (random channels with ICI)`: If the random channels vary within one block duration, the received signal block contains the ICI part. When the normalized Doppler shift $\varepsilon_d = 0.05$, Figure 3.2 depicts BER as a function of SNR for various RC-OFDM systems. Compared to Figure 3.1, it can be observed that the ICI reduces the BER performance particularly in the high SNR range ($20 - 25$ dB). Because RC-OFDM is capable of mitigating the ICI power, the loss of BER performance is not significant. Similar to the test case 1, the differential RC-OFDM still has better performance than that of RC-OFDM. Because we only select 6 out of 25 SNR points

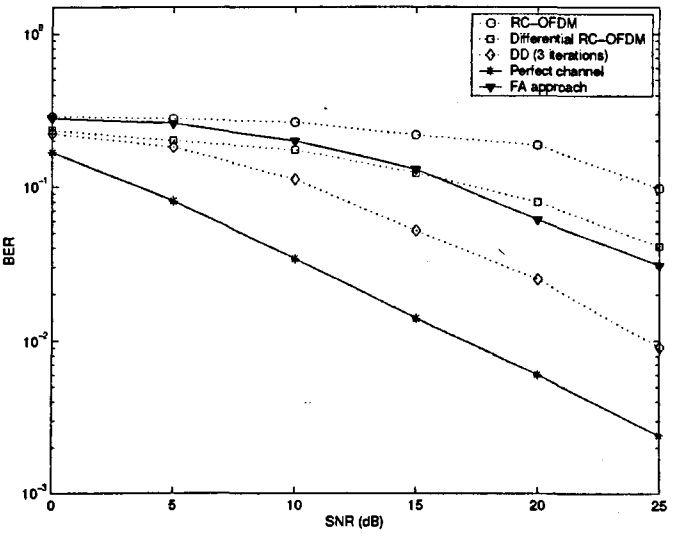

Figure 3.2: BER Vs SNR for RC-OFDM in random channels with ICI.

to plot the curve, a kink appears in the case with perfect channel knowledge at SNR = 5 dB.

`Test Case 3 (time varying channels):` The time-varying FIR channels are generated according to Jake's model [70] with a maximum normalized Doppler frequency $\varepsilon_d = 0.05$. It is assumed that each data burst (a frame) has $N = 64$ OFDM symbol blocks, which are all used for the blind channel estimation. The results are obtained by averaging over 500 Monte Carlo trials. Figure 3.3 shows that all approaches have the significantly improved BER performances compared with the test cases 1 and 2, because the blind channel estimator uses the second-order statistics algorithm to improve the channel estimate accuracy. It also shows that the DD algorithm (3 iterations) may make the overall system performance much close to the case with perfect CSI in the high SNR range (around 25 dB). Usually, blind approaches are used to track the static channel (or slowly time-varying channel). Therefore, in

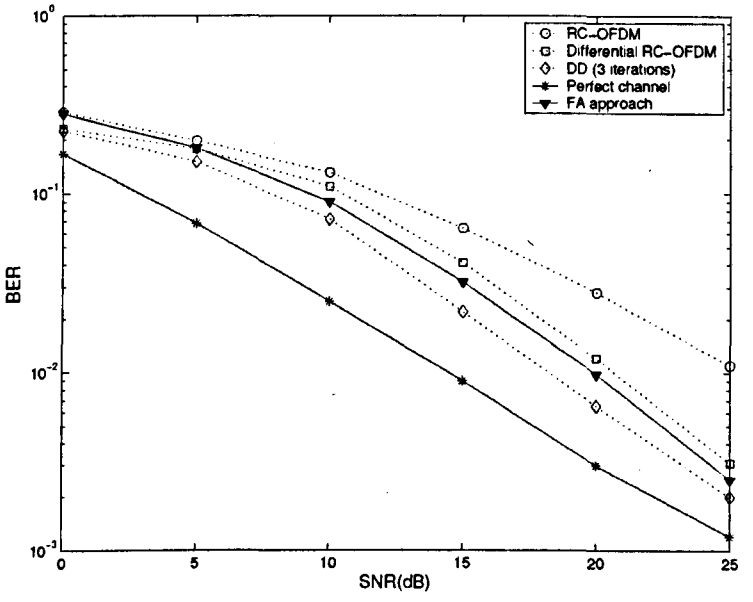

Figure 3.3: BER Vs SNR for RC-OFDM in time varying channels.

the moderately time-varying channel (e.g. $\varepsilon_d = 0.05$), their performance will not be very good.

3.4 Conclusion

This chapter has introduced several novel blind channel estimation and equalization methods for RC-OFDM system by exploiting the subcarrier correlation. A differential RC-OFDM modulation scheme has been developed for the PSK modulated OFDM systems. It has been shown that the differential scheme only needed the partial channel knowledge for the semi-coherent detection, and thereby reduced the channel estimation and equalization complexity. The overall system performance can be significantly improved by employing the decision directed algorithm. Simulation results were provided to confirm our theoretical analysis in static random and time

varying channels.

Chapter 4

Exploiting subcarrier correlation for blind channel estimation (Part II): BP-OFDM

4.1 Introduction

Relying on the precoding redundancy, symbol recovery for linear block precoded OFDM is independent with channel zero (deep fade) locations in frequency-selective fading channels. This chapter presents a second-order joint moment based blind channel estimation method for BP-OFDM systems by exploiting the subcarrier correlation. To enable the proposed method, a slight modification needs to be made at the transmitter (see Figure 4.1). Using the properly designed precoders, the channel identifiability is guaranteed regardless of channel zeros locations. The remaining sign ambiguity can be easily resolved using various blind and semi-blind algorithms. Relying on the exhaustive searching, the minimum distance (MD) algorithm always holds the best performance in both time invariant and slowly time varying channels. To reduce the complexity, a sign-directed (SD) algorithm can be used for the channel estimation. Its performance is subject to the accuracy of the initial channel estimate.

Unlike the subspace-based approach [65], the proposed blind channel estimator does not rely on the precoding redundancy. It works at the full rate transmission. Furthermore, the blind estimator adopts IFFT and sign directed algorithms, and thereby outperforms the eigendecomposition approach in complexity. The additive Gaussian noise considered here is not necessarily white. Part of this chapter has been presented in [84].

4.2 BP-OFDM Model, Modification, and Channel Identifiability

The BP-OFDM model is partly introduced in Chapter 1.4. The detailed introduction can be found in [14], [25]. Since the CP length is normally not smaller than the

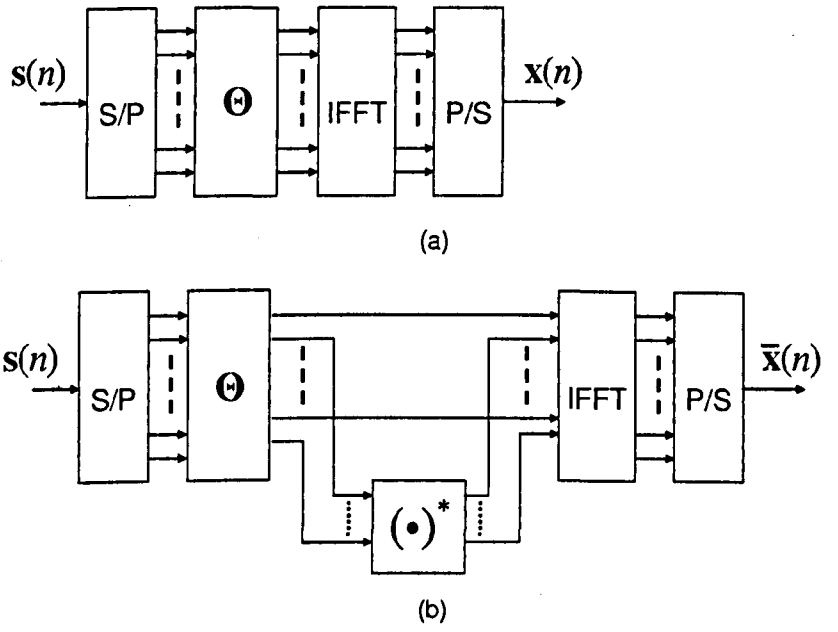

Figure 4.1: Discrete-time models of BP-OFDM transmitter: a) conventional BP-OFDM b) modified BP-OFDM

upper bound of the channel length, there is no IBI in the received block by removing CPs, and the channel equalization is normally implemented in the frequency domain. To simplify the discussion, this Chapter only considers the frequency domain signal model.

4.2.1 Briefly overview of the BP-OFDM model

Figure 4.1 depicts the discrete equivalent model of a BP-OFDM system [65]. Prior to transmission, the information-bearing symbols are first grouped into blocks $\mathbf{s}(n)$ of size $K \times 1$, where n denotes the block index. Then, these blocks are fed to a linear block precoder Θ of size $M \times K$ to introduce redundancy ($M > K$), and generate $\mathbf{x}(n) = \Theta\mathbf{s}(n)$. Let $\mathcal{D}_M$ be the $M \times M$ diagonal channel matrix, the entries along the diagonal are collected from the frequency response $\mathcal{D}_M = \text{diag}\{H(0) \cdots H(M-1)\}$. The received symbol block with the additive Gaussian noise $\mathbf{v}(n)$ of size $M \times 1$ can be expressed as

$$\begin{aligned} \mathbf{y}(n) &= \mathcal{D}_M\Theta\mathbf{s}(n) + \mathbf{v}(n) \\ &= \mathcal{D}_M\mathbf{x}(n) + \mathbf{v}(n) \\ &= \mathbf{z}(n) + \mathbf{v}(n) \end{aligned} \tag{4.2.1}$$

where $\mathbf{z}(n) = \mathcal{D}_M\mathbf{x}(n)$ denotes the noiseless version of the received blocks. When the channel matrix $\mathcal{D}_M$ is perfectly known, the channel equalization can be simply implemented using

$$\hat{\mathbf{s}}(n) = \mathcal{G}^\dagger\mathbf{z}(n) + \mathcal{G}^\dagger\mathbf{v}(n) \tag{4.2.2}$$

where $\mathcal{G} = \mathcal{D}_M\Theta$. In the noiseless case, the above zero-forcing method guarantees the signal recovery regardless of channel zero (deep fade) locations only when the

matrix $\mathcal{G}$ has full column rank. Otherwise, the channel coding technique is required to introduce the redundancy [17].

4.2.2 Subcarrier correlation

Unlike the uncoded OFDM system, subcarriers in the BP-OFDM system may be correlated with others. The autocorrelation of the transmitted blocks $\mathbf{x}(n)$ can be expressed as

$$\mathcal{E}\{\mathbf{x}(n)\mathbf{x}^{\mathcal{H}}(n)\} = \Theta\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\Theta^{\mathcal{H}} \tag{4.2.3}$$

Assuming that a1) the information-bearing symbols are drawn from a finite alphabet with equal probability, the autocorrelation of symbol blocks is given by

$$\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\} = \sigma_s^2\mathbf{I}_K, \tag{4.2.4}$$

where $\mathbf{I}_K$ denotes a $K \times K$ identity matrix. Hence, (4.2.3) can be represented as

$$\mathcal{E}\{\mathbf{x}(n)\mathbf{x}^{\mathcal{H}}(n)\} = \sigma_s^2\Theta\Theta^{\mathcal{H}} \tag{4.2.5}$$

If the matrix $\Theta\Theta^{\mathcal{H}}$ does not contain a zero entry, then any two subcarriers of BP-OFDM are correlated. To achieve this goal, the following assumption should be satisfied:

a2) any K rows of Θ are linearly independent, and for any $m_1 \neq m_2$, $\theta(m_1)\theta^{\mathcal{H}}(m_2) \neq 0$, where $\theta(m)$ denotes the mth row vector of the precoder Θ.

A simple example of the precoder satisfying a2) can be formed by an $M \times K$ Vandermonde matrix, whose (m,k)th entry is given by $\frac{1}{\sqrt{K}}W_M^{-mk}$. Using the subcarrier correlation, we next present a blind channel estimation method for BP-OFDM systems.

4.2.3 A modified BP-OFDM model

The modified BP-OFDM model is described in Figure 4.1(b). The modification is simply implemented by performing a conjugate operation on the $(2m+1)$th element of $\mathbf{x}(n)$ to generate the modified transmitted blocks $\bar{\mathbf{x}}(n)$, whose $2m$th and $(2m+1)$th elements are given by

$$\bar{x}(n, 2m) = \boldsymbol{\theta}(2m)\mathbf{s}(n) \tag{4.2.6}$$

$$\bar{x}(n, 2m+1) = \mathbf{s}^{\mathcal{H}}(n)\boldsymbol{\theta}^{\mathcal{H}}(2m+1) \tag{4.2.7}$$

and the $2m$th and $(2m+1)$th elements of the received block can be expressed as

$$\begin{aligned} y(n, 2m) &= H(2m)\boldsymbol{\theta}(2m)\mathbf{s}(n) + v(n, 2m) \\ &= z(n, 2m) + v(n, 2m) \end{aligned} \tag{4.2.8}$$

$$\begin{aligned} y^*(n, 2m+1) &= H^*(2m+1)\boldsymbol{\theta}(2m+1)\mathbf{s}(n) + v^*(n, 2m+1) \\ &= z^*(n, 2m+1) + v^*(n, 2m+1) \end{aligned} \tag{4.2.9}$$

where $z(n, m)$ denotes the mth element of the vector $\mathbf{z}(n)$, and $v(n, m)$ is the mth element of the noise vector $\mathbf{v}(n)$. Without loss of generality, let us assume M is even and construct an $M \times 1$ vector

$$\bar{\mathbf{y}}(n) = [y(n, 0),\ y^*(n, 1), \cdots, y(n, M-2), y^*(n, M-1)]^T \tag{4.2.10}$$

and a diagonal matrix

$$\bar{\boldsymbol{\mathcal{D}}}_M(n) = \mathrm{diag}\{H(0),\ H^*(1), \cdots, H(M-2), H^*(M-1)\}^T \tag{4.2.11}$$

The channel equalization can be implemented using the following matrix multiplication

$$\hat{\mathbf{s}}(n) = \bar{\mathcal{G}}^{\dagger}\bar{\mathbf{y}}(n) \tag{4.2.12}$$

where $\bar{\mathcal{G}} = \bar{\mathcal{D}_M}\Theta$. In the noiseless case, (4.2.12) guarantees the symbol recovery regardless of deep fade locations only when the matrix $\bar{\mathcal{G}}$ has full column rank. This is an equivalent condition as that addressed in the BP-OFDM model.

4.2.4 Channel identifiability

Besides the conditions a1) and a2), the channel identifiability is based on the modified BP-OFDM model (4.2.6)-(4.2.7) and the following assumptions:

a3) the Gaussian noise $v(n,m)$ is independent with the signal. It is uncorrelated with n's and m's and not necessary white.

a4) $M \geqslant 2L$.

Prior to addressing the channel identification issue, this section first investigates the second joint moment $\mathcal{M}_x(m_1,m_2) = \mathcal{E}\{x(n,m_1)x(n,m_2)\}$ of the transmitted blocks.

Result 4.2.1. *Define two sets* $\mathcal{A} = \{0,2,\cdots,M-2\}$ *and* $\mathcal{B} = \{1,3,\cdots,M-1\}$, *the second joint moment* $\mathcal{M}_x(m_1,m_2) \neq 0$ *holds only when the indices* m_1 *and* m_2 *are not in the same set.*

Proof. See Appendix 4.A.

Let us define an index $m_\tau \in \mathcal{B}$ and form $M \times 1$ blocks $\mathbf{z}_{m_\tau}(n)$ by circularly shifting the noiseless received blocks $\mathbf{z}(n)$ with the offset m_τ, such that

$$\mathbf{z}_{m_\tau}(n) = [z(n,m_\tau),\cdots,z(n,m_\tau-1)]^T \tag{4.2.13}$$

Using Result 4.2.1, the second joint moment $\mathcal{M}(\mathbf{z},\mathbf{z}_{m_\tau}) = \mathcal{E}\{\mathbf{z}(n)\odot\mathbf{z}_{m_\tau}(n)\}$ can be derived as

$$\begin{aligned}\mathcal{M}(\mathbf{z},\mathbf{z}_{m_\tau}) &= \mathcal{E}\{\mathbf{z}(n)\odot\mathbf{z}_{m_\tau}(n)\}\\ &= \sigma_s^2\Phi_{m_\tau}(\tilde{\mathbf{h}}\odot\tilde{\mathbf{h}}_{m_\tau})\end{aligned} \tag{4.2.14}$$

where $\odot$ stands for the Hadamard product; $\tilde{\mathbf{h}} = [H(0), \cdots, H(M-1)]^T$ and $\tilde{\mathbf{h}}_{m_\tau}$ is formed by circularly shifting $\tilde{\mathbf{h}}$ with the offset m_τ; $\mathbf{\Phi}_{m_\tau}$ is the diagonal matrix, which is given by

$$\mathbf{\Phi}_{m_\tau} = \text{diag}\{\boldsymbol{\theta}(0)\boldsymbol{\theta}^{\mathcal{H}}(m_\tau), \ \cdots, \boldsymbol{\theta}^*(0)\boldsymbol{\theta}^T(m_\tau - 1)\} \tag{4.2.15}$$

Define a $L_u \times 1$ vector $\mathbf{h}_{m_\tau} = [h(0), h(1)W_M^{m_\tau}, \cdots, h(L_u-1)W_M^{m_\tau(L_u-1)}]^T$ and construct a $(2L_u - 1) \times 1$ vector $\boldsymbol{\beta}_{m_\tau} = \mathbf{h} \star \mathbf{h}_{m_\tau}$, where $\star$ denotes the convolution operator. The circular convolution property in [60] shows that

$$\tilde{\mathbf{h}} \odot \tilde{\mathbf{h}}_{m_\tau} = \mathbf{F}\boldsymbol{\beta}_{m_\tau} \tag{4.2.16}$$

where $\mathbf{F}$ is formed by collecting the first $(2L_u - 1)$ columns of an $M \times M$ DFT matrix $\mathbf{F}_M$. Hence, (4.2.14) can be rewritten as

$$\mathcal{M}(\mathbf{z}, \mathbf{z}_{m_\tau}) = \sigma_s^2 \mathbf{\Phi}_{m_\tau} \mathbf{F} \boldsymbol{\beta}_{m_\tau} \tag{4.2.17}$$

As a2) and a4) confirm that $\mathbf{\Phi}_{m_\tau}$ has full rank and $M \geqslant 2L_u$, the result of $\mathbf{\Phi}_{m_\tau}\mathbf{F}$ has full column rank and the estimate of $\boldsymbol{\beta}_{m_\tau}$ can be expressed as

$$\hat{\boldsymbol{\beta}}_{m_\tau} = 1/\sigma_s^2 (\mathbf{\Phi}_{m_\tau}\mathbf{F})^\dagger \mathcal{M}(\mathbf{z}, \mathbf{z}_{m_\tau}) \tag{4.2.18}$$

This result shows that the vector $\boldsymbol{\beta}_{m_\tau}$ can be uniquely identified regardless of channel zeros locations, if the assumption a2) is satisfied. When $\boldsymbol{\beta}_{m_\tau}$ becomes available, the channel identifiability will be shown as follows.

Lemma 4.2.1. *Defining a polynomial*

$$f(x, m_\tau) = \sum_{l=0}^{2L_u-2} \beta_{m_\tau}(l) x^l \tag{4.2.19}$$

where $\beta_{m_\tau}(l)$ is the lth element of $\boldsymbol{\beta}_{m_\tau}$. Suppose x_1 to be one of roots of $f(x, m_\tau)$, then there must be another root x_2 such that $x_1 x_2^* = W_M^{\pm m_\tau} |x_1|^2$.

Proof. See Appendix 4.B.

Theorem 4.2.1. *If a2) and a4) holds. For $m_\tau \in \mathcal{B}$, and a5) $f(x, m_\tau)$ has $2L_u - 2$ distinct roots, the CIR* $\mathbf{h}$ *can be uniquely identified from $\boldsymbol{\beta}_{m_\tau}$ with a sign ambiguity.*

Proof. See Appendix 4.C.

The proof of Theorem 4.2.1 also suggests a root-based channel estimation algorithm, which is summarized as follows:

Step 1: Estimate the vector $\boldsymbol{\beta}_{m_\tau}$ using (4.2.18) and form the polynomial $f(x, m_\tau)$;

Step 2: Find the $2L_u - 2$ roots of $f(x, m_\tau)$, and determine the roots for $p(x)$, which is used to construct the matrix $\mathbf{V}$;

Step 3: Identify $h(0)$ using $\hat{h}(0) = \lambda\sqrt{\beta_{m_\tau}(0)}$, and determine $\boldsymbol{\alpha}$ using (4.5.9).

The remaining sign ambiguity λ can be easily resolved using several pilot symbols. However, the root based algorithm is restricted to the condition (a5). If (a5) does not hold, the following result will show the channel identifiability.

Lemma 4.2.2. *If $h(0), \cdots, h(L')$ are available, then the $(L'+1)$th tap $h(L'+1)$ can be uniquely identified from the $(L'+1)$th term of $f(x, m_\tau)$.*

Proof. See Appendix 4.D.

This result shows that $h(L'+1)$ is uniquely identifiable only when the term $1 + W_M^{L'+1} \neq 0$, which is guaranteed by the condition a4).

Theorem 4.2.2. *If $h(0)$ is available, the CIR* $\mathbf{h}$ *can be uniquely identified from $\boldsymbol{\beta}_{m_\tau}$.*

Proof. Lemma 4.2.2 has shown that the identifiability of $h(L'+1)$ depends on the knowledge of $h(0), \cdots, h(L')$. Using the linear equation (LE) algorithm addressed in [66], all channel taps are identifiable via (4.5.11) when $h(0)$ becomes available. However, the LE algorithm is sensitive to the noise and is prone to error propagation.

4.3 Blind channel estimation for BP-OFDM

Section 4.2 has established that the channel identifiability can be guaranteed with a sign ambiguity if the vectors $\boldsymbol{\beta}_{m_\tau}$ become available. As the proposed root-based and LE algorithms have their own drawbacks, this section introduces other blind algorithms to improve the channel estimate performance in the noisy case.

Theorem 4.3.1. *Define a polynomial $f_0(x) = p^2(x)$. If $m_\tau \in \mathcal{B}$ and a4) holds, $f_0(x)$ can be uniquely identified from $f(x, m_\tau)$.*

Proof. See Appendix 4.E.

Let $\boldsymbol{\beta}_0 = \mathbf{h} \star \mathbf{h}$, it can be observed that the lth element of $\boldsymbol{\beta}_0$ is actually the coefficient for the lth term of $f_0(x)$. It means that $\boldsymbol{\beta}_0$ can be uniquely identified when $\boldsymbol{\beta}_{m_\tau}$ is available. Hence, the CIR can be obtained by minimizing the Euclidean distance

$$\hat{\mathbf{h}} = \arg\min_{\hat{\mathbf{h}}} \|\hat{\boldsymbol{\beta}}_0 - \hat{\mathbf{h}} \star \hat{\mathbf{h}}\|^2 \tag{4.3.1}$$

In the presence of noise, the blind channel estimation method is presented as follows.

Step 1: Construct vectors $\mathbf{w}_{m_\tau}(n)$ and $\mathbf{y}_{m_\tau}(n)$ by circularly shifting $\mathbf{w}(n)$ and $\mathbf{y}(n)$ respectively with the offset m_τ. Then, the second joint moment $\mathcal{M}(\mathbf{y}, \mathbf{y}_{m_\tau}) = \mathcal{E}\{\mathbf{y}(n) \odot \mathbf{y}_{m_\tau}(n)\}$ is given by

$$\begin{aligned} \mathcal{M}(\mathbf{y}, \mathbf{y}_{m_\tau}) &= \mathcal{E}\{\mathbf{z}(n) \odot \mathbf{z}_{m_\tau}(n)\} + \mathcal{E}\{\mathbf{w}(n) \odot \mathbf{w}_{m_\tau}(n)\} \\ &= \mathcal{M}(\mathbf{z}, \mathbf{z}_{m_\tau}) \end{aligned} \tag{4.3.2}$$

where a2) makes $\mathcal{E}\{\mathbf{w}(n) \odot \mathbf{w}_{m_\tau}(n)\} = \mathbf{0}_{M\times 1}$;

Step 2: Find $\hat{\boldsymbol{\beta}}_{m_\tau}$ from (4.2.18), and obtain $\hat{\boldsymbol{\beta}}_0$ using (4.5.15) and (4.5.16);

Step 3: Estimate $\hat{\mathbf{h}}$ by solving (4.3.1).

The channel estimation issue by now becomes how to solve the equation (4.3.1). [66] has proposed a number of blind algorithms, including minimum distance (MD) and phase directed (PD) algorithms. This paper applies some of them to solve (4.3.1), and also proposes a new algorithm as the tradeoff between the estimate performance and the complexity.

4.3.1 Minimum distance algorithm

Let $\hat{\beta}_0(l)$ be the lth element of $\hat{\boldsymbol{\beta}}_0$, DFT of $\hat{\boldsymbol{\beta}}_0$ can be expressed as

$$X_0(m) = \sum_{l=0}^{2L-2} \hat{\beta}_0(l) W_M^{ml} = \hat{H}^2(m) \tag{4.3.3}$$

and the estimate $\hat{H}(m)$ can be obtained as the following with a sign ambiguity $\lambda \in \{1, -1\}$

$$\hat{H}(m) = \lambda \sqrt{X_0(m)} \tag{4.3.4}$$

So, the estimate $\hat{\tilde{\mathbf{h}}}$ has totally 2^M possible states. Let $\hat{\mathbf{h}} = \mathbf{F}_{\mathbf{h}}^{\dagger} \hat{\tilde{\mathbf{h}}}$, where $\mathbf{F}_{\mathbf{h}}$ collects the first L columns of the DFT matrix $\mathbf{F}_M$, (4.3.1) can be solved by exhaustively searching over all possible states of $\hat{\tilde{\mathbf{h}}}$. However, the complexity for the MD algorithm increases considerably with increasing M. In practical systems, e.g. WLAN, M is normally far larger than the channel length L_u. Selecting L_u out of M tones for the channel estimation, the computation complexity can be reduced to 2^{L_u}. This is called the modified MD (MMD) algorithm in [66].

4.3.2 Sign directed (SD) algorithm

To further reduce the complexity, a sign directed algorithm can be applied for avoiding the exhaustive searching. This is actually a modified version of the phase directed algorithm in [66]. Its realization is described as follows.

Step 1: Find the initial estimate $\hat{\mathbf{h}}_0$ and its frequency response $\hat{H}_0(m)$ using a low complexity algorithm, e.g. the LE algorithm;

Step 2: Resolve the sign ambiguity via

$$\hat{H}(m) = \arg\min_{\lambda} |\hat{H}_0(m) - \lambda\sqrt{X_0(m)}| \tag{4.3.5}$$

and identify $\hat{\mathbf{h}}$ using $\hat{\mathbf{h}} = \mathbf{F}_{\mathbf{h}}^{\dagger}\hat{\tilde{\mathbf{h}}}$;

Step 3: Let $\hat{\mathbf{h}}_0 = \hat{\mathbf{h}}$ and $\hat{\tilde{\mathbf{h}}}_0 = \mathbf{F}_{\mathbf{h}}\hat{\mathbf{h}}_0$, and repeat Step 2 several times. The number of iterations can be decided by users.

Using one iteration, the estimate performance for the LE-SD algorithm is relatively poor. [66] has demonstrated that the performance may be improved by conducting more than three iterations. Certainly, there must be a tradeoff between the estimate performance and the complexity. Alternatively, a training block can be employed to obtain the initial channel estimate $\hat{\mathbf{h}}_0$ accurately. However, in a time-varying channel, the training block has to be frequently used to keep a good estimation performance, which results in relatively low spectrum efficiency.

4.3.3 Modified SD algorithm

To achieve a good tradeoff between the complexity and the estimation performance, a modified SD (MSD) algorithm is proposed here.

Theorem 4.3.2. *If β_{m_τ}, for $m_\tau \in \mathcal{B}$, is available, then the CIR can be obtained using*

$$[\mathbf{h}^T, \mathbf{0}_{1\times(L_u-1)}]^T = \frac{2\lambda}{M\sqrt{\beta_0(0)}} \sum_{m_\tau\in\mathcal{B}} \beta_{m_\tau} \tag{4.3.6}$$

Proof. See Appendix 4.F.

Theorem 4.3.2 shows that the CIR can be uniquely identified from β_{m_τ}, for $m_\tau \in \mathcal{B}$ with a sign ambiguity λ. Unlike the LE algorithm, (4.3.6) is not prone to the error propagation, and thus improves the channel estimation performance. (4.3.6) also shows that the channel estimation performance is related to the energy of $h(0)$. It becomes the best only when the strongest path is well synchronized to the first path. Otherwise, the SD algorithm can further improve the performance when (4.3.6) is used to obtain the initial channel estimate.

4.4 Simulation results

In the simulation, the BER performance is used to benchmark the BP-OFDM system equipped with the proposed channel estimators in the unknown multipath fading channel. The system parameters are given by: $L_u = 4$, $K = 28$, $M = K + L_u = 32$, which are the same as HIPERLAN/2 [8]. Two pilot symbols are employed within each frame to resolve the sign ambiguity. The precoder Θ is formed by an $M \times K$ Vandermonde matrix, whose (m,k)th entry is given by $\frac{1}{\sqrt{K}} W_M^{-mk}$. The information-bearing symbols are drawn from a QPSK constellation with the equal probability. The proposed channel estimators are tested for static as well as slowly time-varying channels.

`Test Case 1 (static channels)` : The channel taps are randomly generated with equal variance $1/(L_u+1)$. The simulation results are averaged over 500 random channels. The proposed channel estimation algorithms, i.e. MD, MSD, LE-SD, Tr-SD and MMD-SD, are used to solve the equation (4.3.1). Those algorithms are evaluated for the data record length $N = 300$. Figure 4.2 shows that all the curves are very close for the SNR in the range of 0-25 dB but the LE-SD algorithm. Relying on the

exhaustive searching, the system equipped with the MD blind estimator demonstrates the best overall system performance. When the initial channel estimate is obtained using (4.3.6), it can be observed from Figure 4.2 that the system equipped with the MSD estimator only has 2 dB performance lost in the high SNR (> 25 dB). Next, we collect $N = 150$ OFDM blocks for the channel estimation. Figure 4.3 shows that the BER performance is worse than those illustrated in Figure 4.2, because the channel estimate accuracy is reduced with the decrease of the data record length. It also shows that all the curves are not close anymore. The MD estimator always offers the best overall system performance. As conclusions, we can further summarize that:

1. The proposed channel estimators are capable of identifying the time-invariant channel. Its channel estimate performance can be improved by increasing the data record length.

2. The MD and MMD-SD algorithms demonstrate their excellent performance, but the cost is the relatively high computation complexity. The training based SD and MSD algorithms may achieve a good tradeoff between complexity and the channel estimate performance.

3. The computation time for different curves are given by: MD (2.1 ms), MMD-SD (1.14ms), Tr-SD (0.7ms), LE-SD (0.611ms), MSD (0.39ms).

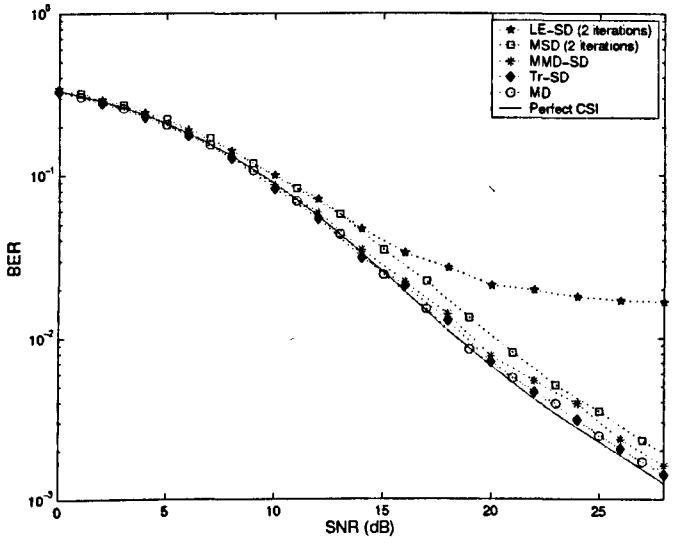

Figure 4.2: BER Vs SNR for BP-OFDM in static channels with N=300.

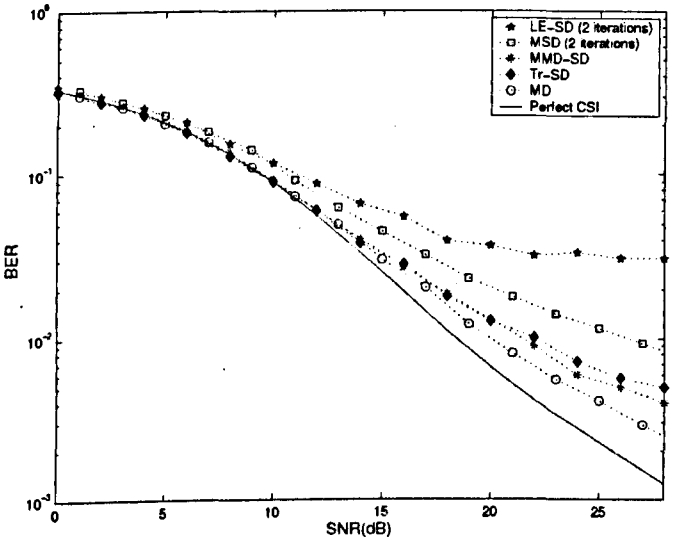

Figure 4.3: BER Vs SNR for BP-OFDM in static channels with N=150.

`Test Case 2 (slowly time-varying channels)`: The slowly time-varying FIR channels are generated according to the channel model A specified by ETSI for HIPERLAN/2 [67], where each tap varies according to Jakes' model with the maximum Doppler frequency of 52Hz corresponding to a typical terminal speed $v = 3$m/s and a carrier frequency of 5.2 GHz. The subcarrier spacing is set to 312.5KHz, which is the same as the specification of HIPERLAN/2 [8]. We assume that each data burst contains $N = 600$ OFDM blocks, where the first block is a training sequence. In the time-varying channel, the data-aided approach needs frequent training to track the CSI. Figure 4.4 shows that one training sequence is not enough for tracking the time-varying channel and results in relatively large error floor. This also affects the performance of the training based SD algorithm, and the MSD algorithm outperforms the training based SD algorithm in the time-varying channel. Similar to the simulations in static channels, the MD algorithm still has the best performance in all approaches. With decreasing the data record length to $N = 150$, Figure 4.5 shows that the overall system performance becomes worse. This simulation example shows that the proposed channel estimators can track the slowly time-varying channel if the data record length is large enough, e.g. $N = 300$. In order to compare with state-of-the-art approaches, I also plot BER performance by employing the subspace-based approach (see [64]) in Figures 4.4 and 4.5. We can see that the subspace-based approach does not perform well in the low SNR. This is because the subspace-based approach needs to resolve the residual scalar ambiguity. However, the scalar cannot be precisely estimated in the low SNR, particularly when the number of pilots is small. In the high SNR, the subspace-based approach has close BER performance to the MD approach.

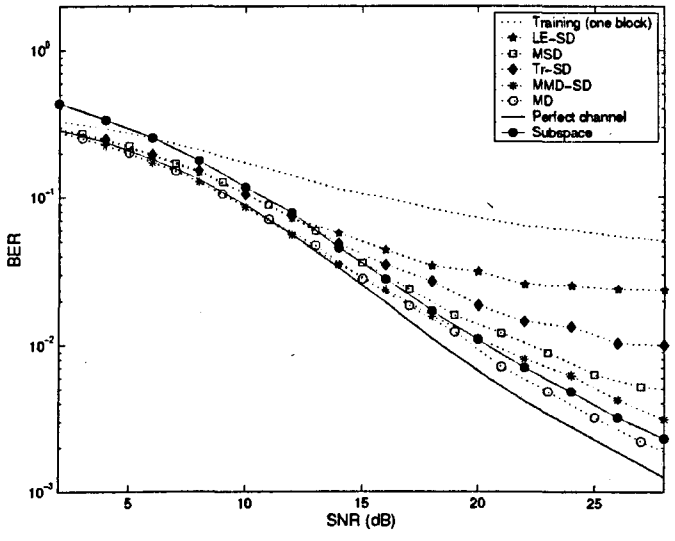

Figure 4.4: BER Vs SNR for BP-OFDM in the slowly time varying channel with N=300.

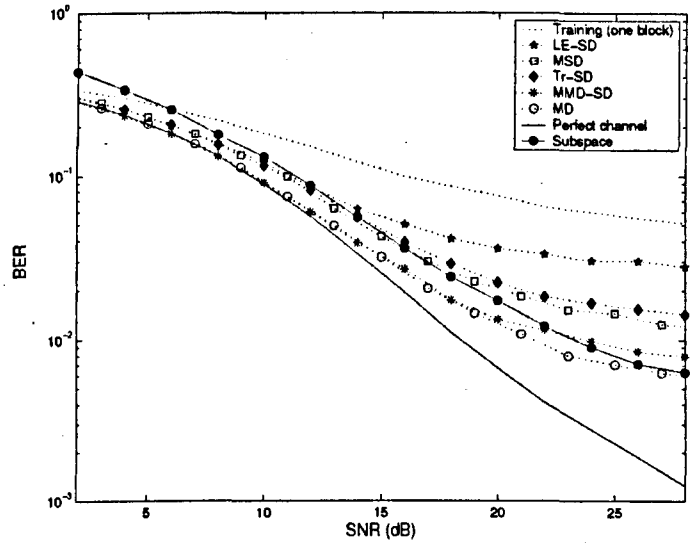

Figure 4.5: BER Vs SNR for BP-OFDM in the slowly time varying channel with N=150.

4.5 Conclusion

This chapter has introduced second-order joint moment based blind or semi-blind channel estimation algorithms for BP-OFDM systems. To enable the proposed algorithms, a modified BP-OFDM model has also been proposed. Exploiting the subcarrier correlation, it has been established that the channel identifiability was guaranteed regardless of channel zeros locations. Unlike the subspace approach, the proposed channel estimators did not rely on the precoding redundancy, and the present additive noise is not necessary white. Simulation results compared the overall system performance for all proposed channel estimators. It has been shown that the MD algorithm always had the best performance by paying the cost of high computation complexity. In the slowly time varying channel, the MSD algorithm may give the best tradeoff between the complexity and the channel estimate performance.

Appendix 4.A: Proof of Result 4.2.1

If m_1 and m_2 are in the same set, $\mathcal{M}_x(m_1, m_2)$ can be expressed as

$$\begin{aligned}\mathcal{M}_x(m_1, m_2) &= \mathcal{E}\{x(n, m_1)x(n, m_2)\} \\ &= \begin{cases} \boldsymbol{\theta}(m_1)\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^T(n)\}\boldsymbol{\theta}^T(m_2), & m_1, m_2 \in \mathcal{A} \\ \boldsymbol{\theta}^*(m_1)\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^T(n)\}^*\boldsymbol{\theta}^{\mathcal{H}}(m_2), & m_1, m_2 \in \mathcal{B} \end{cases}\end{aligned} \tag{4.5.1}$$

a1) indicates that $\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^T(n)\} = \mathbf{0}_K$, where $\mathbf{0}_K$ stands for a zero matrix of size $K \times K$, such that $\mathcal{M}_x(m_1, m_2) = 0$. When m_1 and m_2 are in different sets, without loss of generality, let $m_1 \in \mathcal{A}$ and $m_2 \in \mathcal{B}$, the function $\mathcal{M}_x(m_1, m_2)$ is given by

$$\begin{aligned}\mathcal{M}_x(m_1, m_2) &= \boldsymbol{\theta}(m_1)\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\boldsymbol{\theta}^{\mathcal{H}}(m_2) \\ &= \sigma_s^2\boldsymbol{\theta}(m_1)\boldsymbol{\theta}^{\mathcal{H}}(m_2)\end{aligned} \tag{4.5.2}$$

a2) assures $\theta(m_1)\theta^{\mathcal{H}}(m_2) \neq 0$, Result 4.2.1 is therefore proved.

Appendix 4.B: Proof of Lemma 4.2.1

For $\beta_{m_\tau} = \mathbf{h} \star \mathbf{h}_{m_\tau}$, the polynomial (4.2.19) can be represented as

$$f(x, m_\tau) = p(x)g(x, m_\tau) \tag{4.5.3}$$

where

$$p(x) = \sum_{l=0}^{L_u-1} h(l)x^l \tag{4.5.4}$$

$$g(x, m_\tau) = \sum_{l=0}^{L_u-1} h(l)(W_M^{m_\tau}x)^l \tag{4.5.5}$$

If x_1 is a root of $f(x, m_\tau)$, then it must be a root of either $p(x)$ or $g(x, m_\tau)$. Assume x_1 is a root of $p(x)$, then the following equation can be obtained

$$\begin{aligned} p(x_1) &= \sum_{l=0}^{L_u-1} h(l)x_1^l \\ &= \sum_{l=0}^{L_u-1} h(l)W_M^{-lm_\tau}(W_M^{m_\tau}x_1)^l \\ &= g(W_M^{-m_\tau}x_1, m_\tau) \\ &= 0 \end{aligned} \tag{4.5.6}$$

It means that $x_2 = W_M^{-m_\tau}x_1$ is also a root of $f(x, m_\tau)$, which results in $x_1 * x_2^* = W_M^{m_\tau}|x_1|^2$. Alternatively, if $g(x_1, m_\tau) = 0$, then it is easy to obtain $p(W_M^{m_\tau}x_1) = 0$. Let $x_2 = W_M^{m_\tau}x_1$, it can be found that $x_1 * x_2^* = W_M^{-m_\tau}|x_1|^2$.

Appendix 4.C: Proof of Theorem 4.2.1

Since a2) and a4) guarantee that the polynomial $f(x, m_\tau)$ is available, it is easy to find the $2L_u - 2$ distinct roots of $f(x, m_\tau)$. As shown in Lemma 4.2.1, $f(x, m_\tau)$ can

be factorized into two polynomials, $p(x)$ and $g(x, m_\tau)$. Let $x_{p,i}$ and $x_{g,i}$ be the ith roots of $p(x)$ and $g(x, m_\tau)$ respectively, Lemma 4.2.1 has shown that $x_{p,i} = W_M^{m_\tau} x_{g,i}$. Therefore, the roots for $p(x)$ can be determined and is used to form a $(L_u-1)\times(L_u-1)$ Vandermonde matrix as

$$\mathbf{V} = \begin{pmatrix} x_{p,1} & \cdots & x_{p,1}^{L-1} \\ \vdots & \ddots & \vdots \\ x_{p,L-1} & \cdots & x_{p,L-1}^{L-1} \end{pmatrix} \tag{4.5.7}$$

Let $\boldsymbol{\alpha} = [h(1), \cdots, h(L_u - 1)]^T$ and $\boldsymbol{\gamma} = [h(0), \cdots, h(0)]^T$, (4.5.4) and (4.5.5) show that

$$\boldsymbol{\gamma} = \mathbf{V}\boldsymbol{\alpha} \tag{4.5.8}$$

Since all roots are distinct, the Vandermonde matrix must be full rank. If $h(0)$ can be determined, then $\boldsymbol{\alpha}$ can be uniquely identified using

$$\boldsymbol{\alpha} = \mathbf{V}^\dagger \boldsymbol{\gamma} \tag{4.5.9}$$

Due to $\beta_{m_\tau}(0) = h^2(0)$, the first tap of the channel vector can be obtained by letting $\hat{h}(0) = \lambda\sqrt{\beta_{m_\tau}(0)}$, and $\lambda \in \{1, -1\}$. In this case, (4.5.9) shows that $\boldsymbol{\alpha}$ can be uniquely identified with a sign ambiguity.

Appendix 4.D: Proof of Lemma 4.2.2

(4.5.7)-(4.5.8) implies that the $(L'+1)$th term of $f(x, m_\tau)$ can be expressed as

$$\begin{aligned} \beta_{m_\tau}(L'+1)x^{L'+1} &= h(0)h(L'+1)(1 + W_M^{L'+1})x^{L'+1} \\ &+ f_{m_\tau}^{L'+1}(h(0), \cdots, h(L'), x) \end{aligned} \tag{4.5.10}$$

where $f_{m_\tau}^{L'+1}(h(0),\cdots,h(L'),x)$ is a function of the parameters $h(0),\cdots,h(L')$ and x. Let $x=1$, $h(L'+1)$ can be obtained using

$$h(L'+1)=\frac{\beta_{m_\tau}(L'+1)-f_{m_\tau}^{L'+1}(h(0),\cdots,h(L'),1)}{h(0)(1+W_M^{L'+1})} \tag{4.5.11}$$

Appendix 4.E: Proof of Theorem 4.3.1

Using (4.5.3) and (4.5.4), $f(x,m_\tau)$ can be rewritten into

$$f(x,m_\tau)=\sum_{l=0}^{L-1}p(x)h(l)W_M^{m_\tau l}x^l \tag{4.5.12}$$

For $m_\tau\in\mathcal{B}$, (4.5.12) can be expressed as the following matrix form

$$\mathbf{a}=\mathbf{V}_f\mathbf{b} \tag{4.5.13}$$

where $\mathbf{a}=[f(x,1),f(x,3),\cdots,f(x,M-1)]^T$, $\mathbf{b}=[p(x)h(0),\cdots,p(x)h(L-1)x^{L-1}]^T$, and $\mathbf{V}_f$ is a Vandermonde matrix of size $(M/2)\times L$

$$\mathbf{V}_f=\begin{pmatrix}1 & W_M^1 & \cdots & W_M^{L-1}\\ \vdots & \ddots & \ddots & \vdots\\ 1 & W_M^{M-1} & \cdots & W_M^{(M-1)(L-1)}\end{pmatrix} \tag{4.5.14}$$

Since a4) assures that $M/2\geqslant L$, $\mathbf{V}_f$ must have full column rank, and $\mathbf{b}$ can be found using

$$\mathbf{b}=\mathbf{V}_f^{\dagger}\mathbf{a} \tag{4.5.15}$$

The sum of all elements of $\mathbf{b}$ is given by

$$\sum_{l=0}^{L-1}p(x)h(l)x^l=p(x)\sum_{l=0}^{L-1}h(l)x^l=p^2(x) \tag{4.5.16}$$

Appendix 4.F: Proof of Theorem 4.3.2

Consider the right hand of (4.3.6)

$$\sum_{m_\tau \in \mathcal{B}} \beta_{m_\tau} = \mathbf{h} \star \sum_{m_\tau \in \mathcal{B}} \mathbf{h}_{m_\tau} \tag{4.5.17}$$

Chapter 4.2.4 has defined that

$$\begin{aligned} \mathbf{h}_{m_\tau} &= [h(0), h(1)W_M^{m_\tau}, \cdots, h(L_u - 1)W_M^{m_\tau(L_u-1)}]^T \\ &= \mathbf{D}(\mathbf{f}_\mathbf{h}(m_\tau))\mathbf{h} \end{aligned} \tag{4.5.18}$$

where $\mathbf{f}_\mathbf{h}(m_\tau)$ denotes the m_τth row vector of the matrix $\mathbf{F}_\mathbf{h}$. Substitute (4.5.18) into (4.5.17), we may obtain

$$\begin{aligned} \sum_{m_\tau \in \mathcal{B}} \beta_{m_\tau} &= \mathbf{h} \star \sum_{m_\tau \in \mathcal{B}} \mathbf{D}(\mathbf{f}_\mathbf{h}(m_\tau))\mathbf{h} \\ &= \mathbf{h} \star [\mathrm{diag}\{M/2,\ 0,\ \cdots, 0\}\mathbf{h}] \\ &= \frac{M}{2} h(0)[\mathbf{h}^T, \mathbf{0}_{1\times(L_u-1)}]^T \\ &= \frac{M}{2} \lambda \sqrt{\beta_0(0)}[\mathbf{h}^T, \mathbf{0}_{1\times(L_u-1)}]^T \end{aligned} \tag{4.5.19}$$

This is the equivalent result of (4.3.6).

Chapter 5

Superimposed training scheme for OFDM systems

Blind channel estimation algorithms usually need the second-order or high-order statistics, and thereby require a relatively large data record length. This chapter introduces a superimposed training scheme for the channel estimation in OFDM systems. This technique offers the first-order statistics based channel estimation method, and has the potential to shorten the required data record length and reduce the computation complexity. Part of this work is presented in [86].

5.1 Introduction

The superimposed training scheme was originally proposed to identify the frequency selective fading channel in single carrier systems [71]-[72]. The major advantages of this scheme are:

1. Like the blind channel estimation approaches, the superimposed training does not need extra bandwidth consumption, and is capable of tracking time-invariant or slowly time-varying channels.

2. The superimposed training scheme enables the first-order statistics based channel estimation, and thereby outperforms the second or high order statistics based channel estimators in computation complexity.

On the other side, this training scheme superimposes the periodic training sequences on the information-bearing symbols, and lower the effective signal to noise ratio. This drawback may eventually decrease the overall system performance.

In this chapter, we first apply this training scheme in multicarrier systems. A first-order statistics based channel estimation algorithm is developed for both CP-OFDM and ZP-OFDM. The mean square error (MSE) analysis is given to evaluate the channel estimation performance. Then, we investigate the superimposed training scheme in the space-time coding scenario. Simulation results are provided to confirm the theoretical analysis. Finally, future research on this work is also proposed.

5.2 Superimposed training models

Figure 5.1 illustrates the superimposed training scheme for OFDM systems. Differed from the conventional OFDM model, the data fed to the OFDM modulator is the sum of information-bearing symbols $\mathbf{s}(n)$ and the periodic training sequences $\mathbf{c}(n)$, i.e.

$$\bar{\mathbf{s}}(n) = \mathbf{s}(n) + \mathbf{c}(n) \tag{5.2.1}$$

When the energy of $\bar{\mathbf{s}}(n)$ is given, the energy allocation strategy can be reflected from the average signal to training ratio η, which is given by

$$\eta = \frac{\mathrm{Tr}\{\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\}}{\mathrm{Tr}\{\mathcal{E}\{\mathbf{c}(n)\mathbf{c}^{\mathcal{H}}(n)\}\}} \tag{5.2.2}$$

Since the training sequences $\mathbf{c}(n)$ are periodic signals, i.e. the blocks $\mathbf{c}(n)$ are constant for the block index n, then (5.2.2) becomes

$$\eta = \frac{\mathrm{Tr}\{\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\}}{\mathrm{Tr}\{\mathbf{c}(n)\mathbf{c}^{\mathcal{H}}(n)\}} \tag{5.2.3}$$

Next, it will be shown that η is a very important factor and affects the overall system performance. [86]

5.2.1 Superimposed training scheme for CP-OFDM

When the blocks $\bar{\mathbf{s}}(n)$ are fed to the CP-OFDM modulator, (1.2.7) infers that the transmitted blocks can be expressed as

$$\begin{aligned} \mathbf{x}(n) &= \Theta_{\mathrm{cp}}\bar{\mathbf{s}}(n) \\ &= \Theta_{\mathrm{cp}}\mathbf{s}(n) + \Theta_{\mathrm{cp}}\mathbf{c}(n) \end{aligned} \tag{5.2.4}$$

and the received post-FFT blocks are given by

$$\tilde{\mathbf{y}}_{\mathrm{cp}}(n) = \mathcal{D}_M\mathbf{s}(n) + \mathcal{D}_M\mathbf{c}(n) + \tilde{\mathbf{v}}(n) \tag{5.2.5}$$

If the channel matrix $\mathcal{D}_M$ is perfectly known, then the information symbols can be

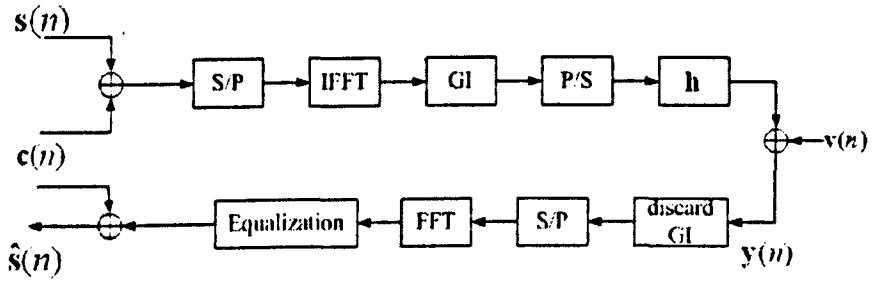

Figure 5.1: Block diagram of the superimposed training scheme for OFDM

recovered using

$$\hat{\mathbf{s}}(n) = \mathcal{D}_M^{\dagger}(\tilde{\mathbf{y}}_{\text{cp}}(n) - \mathcal{D}_M \mathbf{c}(n)) \tag{5.2.6}$$

It has been emphasized that the above ZF method is applicable only when $\mathcal{D}_M$ has full rank. Otherwise, technique using channel coding can be employed to guarantee the signal recovery.

To examine the equalization performance, (5.2.6) can be represented as

$$\hat{\mathbf{s}}(n) = \mathbf{s}(n) + \mathcal{D}_M^{\dagger}\tilde{\mathbf{v}}(n) \tag{5.2.7}$$

Hence, the effective signal-to-noise ratio (ESNR) is given by

$$\begin{aligned}
\text{ESNR}_{\text{cp}} &= \frac{\text{Tr}\{\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\}}{\text{Tr}\{\mathcal{D}_M^{\dagger}\mathcal{E}\{\mathbf{v}(n)\mathbf{v}^{\mathcal{H}}(n)\}(\mathcal{D}_M^{\dagger})^{\mathcal{H}}\}} \\
&= \frac{\sigma_s^2 \cdot M}{\sigma_v^2 \text{Tr}\{(\mathcal{D}_M \mathcal{D}_M^{\mathcal{H}})^{\dagger}\}} \\
&< \frac{\text{Tr}\{\mathcal{E}\{\bar{\mathbf{s}}(n)\bar{\mathbf{s}}^{\mathcal{H}}(n)\}\}}{\sigma_v^2 \text{Tr}\{(\mathcal{D}_M \mathcal{D}_M^{\mathcal{H}})^{\dagger}\}} \\
&= \frac{\sigma_s^2 \cdot M + \text{Tr}\{\mathbf{c}(n)\mathbf{c}^{\mathcal{H}}(n)\}}{\sigma_v^2 \text{Tr}\{(\mathcal{D}_M \mathcal{D}_M^{\mathcal{H}})^{\dagger}\}} \\
&= \frac{(1 + 1/\eta) M \cdot \sigma_s^2}{\sigma_v^2 \text{Tr}\{(\mathcal{D}_M \mathcal{D}_M^{\mathcal{H}})^{\dagger}\}} \\
&= \text{SNR}_{\text{cp}}
\end{aligned} \tag{5.2.8}$$

The difference between ESNR and SNR is given by

$$\begin{aligned}
\text{SNR}_d &= \text{SNR}_{\text{cp}} - \text{ESNR}_{\text{cp}} \\
&= \frac{M \cdot \sigma_s^2}{\eta \sigma_v^2 \text{Tr}\{(\mathcal{D}_M \mathcal{D}_M^{\mathcal{H}})^{\dagger}\}} \\
&= \frac{M \cdot \sigma_s^2}{\eta \sigma_v^2 \sum_{m=0}^{M-1} 1/|H(m)|^2}
\end{aligned} \tag{5.2.9}$$

which may lead to a considerable degradation of the overall system performance when the factor η is small.

5.2.2 Superimposed training scheme for ZP-OFDM

When the blocks $\mathbf{x}(n)$ are fed to the ZP-OFDM modulator, (1.2.15) infers that the received post-FFT blocks can be expressed as

$$\tilde{\mathbf{y}}_{\text{zp}}(n) = \boldsymbol{\mathcal{D}}_J \mathbf{G}_{\text{zp}} \mathbf{s}(n) + \boldsymbol{\mathcal{D}}_J \mathbf{G}_{\text{zp}} \mathbf{c}(n) + \tilde{\mathbf{v}}_J(n) \tag{5.2.10}$$

and the channel equalization can be simply implemented using

$$\begin{aligned} \hat{\mathbf{s}}(n) &= (\boldsymbol{\mathcal{D}}_J \mathbf{G}_{\text{zp}})^{\dagger} (\tilde{\mathbf{y}}_{\text{zp}}(n) - \boldsymbol{\mathcal{D}}_J \mathbf{G}_{\text{zp}} \mathbf{c}(n)) \\ &= \mathbf{s}(n) + (\boldsymbol{\mathcal{D}}_J \mathbf{G}_{\text{zp}})^{\dagger} \tilde{\mathbf{v}}_J(n) \end{aligned} \tag{5.2.11}$$

Therefore, ESNR for ZP-OFDM can be deduced by

$$\begin{aligned} \text{ESNR}_{\text{zp}} &= \frac{\text{Tr}\{\mathcal{E}\{\mathbf{s}(n)\mathbf{s}^{\mathcal{H}}(n)\}\}}{\text{Tr}\{\mathbf{G}_{\text{zp}}^{\dagger} \boldsymbol{\mathcal{D}}_J^{\dagger} \mathcal{E}\{\mathbf{v}_J(n)\mathbf{v}_J^{\mathcal{H}}(n)\} (\boldsymbol{\mathcal{D}}_J^{\dagger})^{\mathcal{H}} (\mathbf{G}_{\text{zp}}^{\dagger})^{\mathcal{H}}\}} \\ &= \frac{M \cdot \sigma_s^2}{\sigma_v^2 \text{Tr}\{\mathbf{G}_{\text{zp}}^{\dagger} \boldsymbol{\mathcal{D}}_J^{\dagger} (\boldsymbol{\mathcal{D}}_J^{\dagger})^{\mathcal{H}} (\mathbf{G}_{\text{zp}}^{\dagger})^{\mathcal{H}}\}} \end{aligned} \tag{5.2.12}$$

and SNR for ZP-OFDM can also be derived as

$$\begin{aligned} \text{SNR}_{\text{zp}} &= \frac{\text{Tr}\{\mathcal{E}\{\bar{\mathbf{s}}(n)\bar{\mathbf{s}}^{\mathcal{H}}(n)\}\}}{\sigma_v^2 \text{Tr}\{\mathbf{G}_{\text{zp}}^{\dagger} \boldsymbol{\mathcal{D}}_J^{\dagger} (\boldsymbol{\mathcal{D}}_J^{\dagger})^{\mathcal{H}} (\mathbf{G}_{\text{zp}}^{\dagger})^{\mathcal{H}}\}} \\ &= \frac{(1 + 1/\eta) M \cdot \sigma_s^2}{\sigma_v^2 \text{Tr}\{\mathbf{G}_{\text{zp}}^{\dagger} \boldsymbol{\mathcal{D}}_J^{\dagger} (\boldsymbol{\mathcal{D}}_J^{\dagger})^{\mathcal{H}} (\mathbf{G}_{\text{zp}}^{\dagger})^{\mathcal{H}}\}} \end{aligned} \tag{5.2.13}$$

So the difference between SNR and ESNR is given by

$$\begin{aligned} \text{SNR}_d &= \text{SNR}_{\text{zp}} - \text{ESNR}_{\text{zp}} \\ &= \frac{M \cdot \sigma_s^2}{\eta \sigma_v^2 \text{Tr}\{\mathbf{G}_{\text{zp}}^{\dagger} \boldsymbol{\mathcal{D}}_J^{\dagger} (\boldsymbol{\mathcal{D}}_J^{\dagger})^{\mathcal{H}} (\mathbf{G}_{\text{zp}}^{\dagger})^{\mathcal{H}}\}} \end{aligned} \tag{5.2.14}$$

Let $\bar{\mathbf{g}}(j)$ be column vectors of the matrix $\mathbf{G}_{\text{zp}}^{\dagger}$, (5.2.14) can be represented as

$$\text{SNR}_d = \frac{M \cdot \sigma_s^2}{\eta \sigma_v^2 \sum_{j=0}^{J-1} \|\bar{\mathbf{g}}(j)\|^2 / |H(j)|^2} \tag{5.2.15}$$

Next, we investigate the channel estimation algorithm for the superimposed training based OFDM systems.

5.3 First-order statistics based channel estimation

The superimposed training scheme enables the first-order statistics algorithm for the channel estimation. It is suitable for tracking time-invariant or slowly time-varying channels, and therefore is applicable in WLANs.

5.3.1 Channel estimation for CP-OFDM

The channel estimation is based on the expectation of the received post-FFT blocks (5.2.5), which is given by

$$\begin{aligned} \mathbf{m}_y &= \mathcal{E}\{\tilde{\mathbf{y}}_{\mathrm{cp}}(n)\} \\ &= \mathcal{D}_M \mathcal{E}\{\mathbf{s}(n)\} + \mathcal{D}_M \mathcal{E}\{\mathbf{c}(n)\} + \mathcal{E}\{\tilde{\mathbf{v}}(n)\} \\ &= \mathcal{D}_M \mathbf{c}(0) \end{aligned} \tag{5.3.1}$$

where $\mathbf{c}(0)$ denotes the 0th training block, and $\mathbf{c}(n) = \mathbf{c}(0)$. This result is quite like the training-based model (2.2.2) in Chapter 2. Hence, (5.3.1) can be rewritten into

$$\begin{aligned} \mathbf{m}_y &= \mathbf{D}(\mathbf{c}(0))\tilde{\mathbf{h}} \\ &= \sqrt{M}\mathbf{D}(\mathbf{c}(0))\mathbf{F}_{\mathbf{h}}\mathbf{h} \end{aligned} \tag{5.3.2}$$

Let $\mathbf{Q}_{\mathbf{c}} = \mathbf{D}(\mathbf{c}(0))\mathbf{F}_{\mathbf{h}}$, the channel estimate can be obtained using

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{M}}\mathbf{Q}_{\mathbf{c}}^{\dagger}\mathbf{m}_y \tag{5.3.3}$$

It shows that the successful channel estimation depends on that the matrix $\mathbf{Q}_{\mathbf{c}}$ has full column rank. This condition can be guaranteed by carefully designing the training blocks $\mathbf{c}(n)$, for instance, $\mathbf{c}(n)$ does not contain a zero element.

In practice, the expectation vector is replaced by the sample average based on finite (say, N) blocks

$$\begin{aligned} \mathbf{m}_y &= \frac{1}{N}\sum_{n=0}^{N-1}\tilde{\mathbf{y}}_{\mathrm{cp}}(n) \\ &= \sqrt{M}\mathbf{Q}_c\mathbf{h} + \mathbf{e}_s + \mathbf{e}_v \end{aligned} \tag{5.3.4}$$

where

$$\mathbf{e}_s = \frac{1}{N}\mathcal{D}_M\sum_{n=0}^{N-1}\mathbf{s}(n) \tag{5.3.5}$$

$$\mathbf{e}_v = \frac{1}{N}\sum_{n=0}^{N-1}\tilde{\mathbf{v}}(n) \tag{5.3.6}$$

Then, the channel estimate is given by

$$\begin{aligned} \hat{\mathbf{h}} &= \frac{1}{\sqrt{M}}\mathbf{Q}_c^{\dagger}\mathbf{m}_y \\ &= \mathbf{h} + \frac{1}{\sqrt{M}}\mathbf{Q}_c^{\dagger}\mathbf{e}_s + \frac{1}{\sqrt{M}}\mathbf{Q}_c^{\dagger}\mathbf{e}_v \end{aligned} \tag{5.3.7}$$

(5.3.7) reflects that the channel estimate error $\mathbf{e}_{\mathbf{h}} = \hat{\mathbf{h}} - \mathbf{h}$ is caused by the remaining interference $\mathbf{e}_s$ and $\mathbf{e}_v$.

Theorem 5.3.1. *Based on (5.3.7), the mean-square channel estimate error* $\mathcal{E}\|\mathbf{e}_{\mathbf{h}}\|^2$ *is given by*

$$\mathcal{E}\|\mathbf{e}_{\mathbf{h}}\|^2 = \frac{\sigma_s^2 L_u}{NM^2}\sum_{m=0}^{M-1}\frac{|H(m)|^2}{|c(m)|^2} + \frac{L_u}{NM^2}\sum_{m=0}^{M-1}\frac{\sigma_{v,m}^2}{|c(m)|^2} \tag{5.3.8}$$

Proof. See Appendix 5.A.

(5.6.4) also infers the following result

$$\lim_{\sigma_{v,m}^2\to 0}\mathcal{E}\|\mathbf{e}_{\mathbf{h}}\|^2 = \frac{\sigma_s^2 L_u}{NM^2}\sum_{m=0}^{M-1}\frac{|H(m)|^2}{|c(m)|^2} \tag{5.3.9}$$

It means that, if the data record length (N) is fixed, the mean-square channel estimate error is asymptotic to the bound (5.3.9) with increasing SNR. The estimate performance mainly depends on the data record length (N) rather than SNR.

Observing (5.6.4), we can find that the channel estimate error is also affected by another important factor $|c(m)|^2$. If the total energy of the training sequence is given, the optimum training sequence is designed as the following result.

Theorem 5.3.2. *If the total energy of the superimposed training sequence is*

$$\sum_{m=0}^{M-1} |c(m)|^2 = \xi_c \tag{5.3.10}$$

then the channel estimate error achieves its minimum

$$\epsilon_{\min} = \frac{L_u}{NM^2\xi_c}[\sigma_s^2 \sum_{m=0}^{M-1}(|H(m)|^2 + \sigma_{v,m}^2)] \tag{5.3.11}$$

when the signal energy is equally allocated along the subcarriers.

Proof. See Appendix 5.B.

This result implies that (5.6.7) is the optimum superimposed training sequence design regardless of the channel knowledge and noise color.

5.3.2 Channel estimation for ZP-OFDM

The channel estimation for ZP-OFDM is based on the expectation of (5.2.10)

$$\begin{aligned} \mathbf{m}_y &= \mathcal{E}\{\tilde{\mathbf{y}}_{\text{zp}}(n)\} \\ &= \mathcal{D}_J\mathbf{G}_{\text{zp}}\mathcal{E}\{\mathbf{s}(n)\} + \mathcal{D}_J\mathbf{G}_{\text{zp}}\mathcal{E}\{\mathbf{c}(n)\} + \mathcal{E}\{\tilde{\mathbf{v}}_J(n)\} \\ &= \mathcal{D}_J\mathbf{G}_{\text{zp}}\mathbf{c}(0) \\ &= \sqrt{J}\mathbf{D}(\mathbf{G}_{\text{zp}}\mathbf{c}(0))\bar{\mathbf{F}}_{\mathbf{h}}\mathbf{h} \end{aligned} \tag{5.3.12}$$

where $\bar{\mathbf{F}}_{\mathbf{h}}$ is formed by collecting the first L_u columns of $\mathbf{F}_J$. Let $\mathbf{Q}_z = \mathbf{D}(\mathbf{G}_{\text{zp}}\mathbf{c}(0))\bar{\mathbf{F}}_{\mathbf{h}}$, the CSI can be estimated using

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{J}}\mathbf{Q}_z^{\dagger}\mathbf{m}_y \tag{5.3.13}$$

The channel identifiability is guaranteed only when $\mathbf{Q}_z$ has full column rank.

Similar to the channel estimation for CP-OFDM, the expectation vector $\mathbf{m}_y$ is realized using the sample average based on N blocks

$$\begin{aligned}\mathbf{m}_y &= \frac{1}{N}\sum_{n=0}^{N-1}\tilde{\mathbf{y}}_{\rm cp}(n)\\ &= \sqrt{J}\mathbf{Q}_z\mathbf{h}+\bar{\mathbf{e}}_s+\bar{\mathbf{e}}_v\end{aligned} \tag{5.3.14}$$

where

$$\bar{\mathbf{e}}_s = \frac{1}{N}\mathcal{D}_J\mathbf{G}_{\rm zp}\sum_{n=0}^{N-1}\mathbf{s}(n) \tag{5.3.15}$$

$$\bar{\mathbf{e}}_v = \frac{1}{N}\sum_{n=0}^{N-1}\tilde{\mathbf{v}}_J(n) \tag{5.3.16}$$

The channel estimate is therefore given by

$$\begin{aligned}\hat{\mathbf{h}} &= \frac{1}{\sqrt{J}}\mathbf{Q}_z^\dagger\mathbf{m}_y\\ &= \mathbf{h}+\frac{1}{\sqrt{J}}\mathbf{Q}_z^\dagger\bar{\mathbf{e}}_s+\frac{1}{\sqrt{J}}\mathbf{Q}_z^\dagger\bar{\mathbf{e}}_v\end{aligned} \tag{5.3.17}$$

and the channel estimate error for ZP-OFDM is given in the following result.

Theorem 5.3.3. *Based on (5.3.17), the mean-square channel estimate error* $\mathcal{E}\|\mathbf{e}_{\rm h}\|^2$ *is given by*

$$\mathcal{E}\|\mathbf{e}_{\rm h}\|^2 = \frac{\sigma_s^2}{NJ}\sum_{j=0}^{J-1}\|\mathbf{u}(j)\|^2\lambda_e(j)+\frac{L_u}{NJ^2}\sum_{m=0}^{J-1}\frac{\sigma_{v,m}^2}{|\mathbf{g}_{\rm zp}(m)\mathbf{c}(0)|^2} \tag{5.3.18}$$

where $\mathbf{u}(j)$ *are column vectors, and* $\lambda_e(j)$ *are real, which are defined in the proof.*

Proof. See Appendix 5.C.

Similar to the CP-OFDM, (5.3.18) infers that

$$\lim_{\sigma_{v,m}^2\to 0}\mathcal{E}\|\mathbf{e}_{\rm h}\|^2 = \frac{\sigma_s^2}{NJ}\sum_{j=0}^{J-1}\|\mathbf{u}(j)\|^2\lambda_e(j) \tag{5.3.19}$$

So, the mean-square channel estimate error is asymptotic to the bound (5.3.19) with the increasing of SNR. The power allocation strategy for ZP-OFDM is actually the same as that for CP-OFDM. The proof will not be addressed here for its complexity.

5.4 Superimposed training for space-time block coded OFDM

Transmit diversity has been studied extensively for combating fading and improving capacity (see [73]-[74]). Specifically, transmit diversity schemes based on space-time (ST) coding exploit multiple transmissions to provide diversity gain. ST trellis codes were first proposed in [75] to achieve maximum diversity gain at the expense of increased complexity of optimal decoding at the receiver. ST block coding (STBC) emerged as an attractive alternative because it leads to ML decoding using only linear receiver processing [76], [78].

ST codes were originally designed for known slow flat fading channels [75], [76]. Unlike narrowband transmission, the flat-fading channel assumption is no longer justified in wideband communications, which motivated recent extension of ST coding to frequency selective channels for OFDM [22], [77]. With transmit diversity, ST coded OFDM can ameliorate fading effects caused by channel nulls. However, when transmit antennas are not well separated, the multiple channels become correlated, and symbol recovery is not guaranteed because nulls common to the multiple channels are possible to occur.

When the frequency-selective fading channels are not known to the transceiver, we are going to use the superimposed training scheme to conduct the channel estimation and signal recovery in STBC-OFDM.

5.4.1 STBC-OFDM model with superimposed training

The superimposed training scheme for STBC-OFDM is depicted in Figure 5.2, where the ST transceiver is equipped with two transmit antennas and one receive antenna as in [65]. Prior to transmission, the information bearing symbol blocks, $\mathbf{s}(n)$, are first fed to the ST encoder. The encoder takes two consecutive input blocks, $\mathbf{s}(2n)$ and $\mathbf{s}(2n+1)$, to output the following $2M \times 2$ code matrix

$$\begin{pmatrix} \bar{\mathbf{s}}_1(2n) & \bar{\mathbf{s}}_1(2n+1) \\ \bar{\mathbf{s}}_2(2n) & \bar{\mathbf{s}}_2(2n+1) \end{pmatrix} = \begin{pmatrix} \mathbf{s}(2n) & -\mathbf{s}^*(2n+1) \\ \mathbf{s}(2n+1) & \mathbf{s}^*(2n) \end{pmatrix} \tag{5.4.1}$$

Next, the periodic training blocks, $\mathbf{c}_1(n)$ and $\mathbf{c}_2(n)$, are superimposed on the ST coded blocks

$$\begin{pmatrix} \bar{\bar{\mathbf{s}}}_1(n) \\ \bar{\bar{\mathbf{s}}}_2(n) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{s}}_1(n) \\ \bar{\mathbf{s}}_2(n) \end{pmatrix} + \begin{pmatrix} \mathbf{c}_1(n) \\ \mathbf{c}_2(n) \end{pmatrix} \tag{5.4.2}$$

Each block is transmitted over successive time intervals with the blocks $\bar{\bar{\mathbf{s}}}_1(n)$ and $\bar{\bar{\mathbf{s}}}_2(n)$ sent from transmit antennas 1 and 2, respectively.

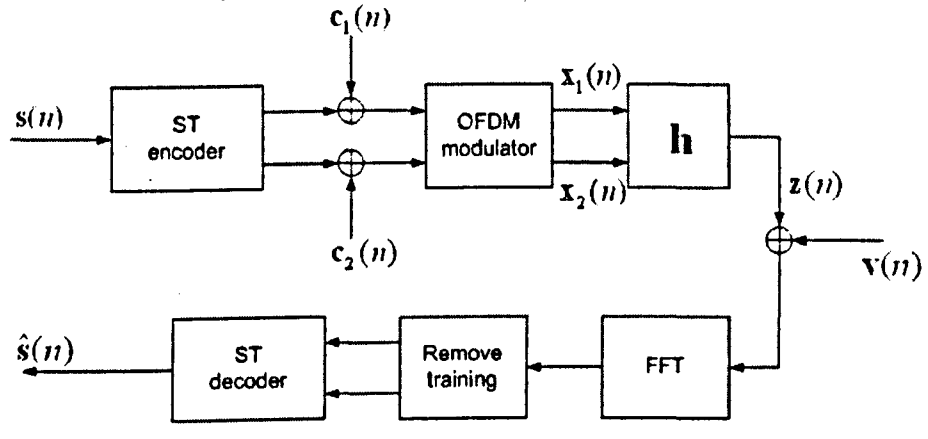

Figure 5.2: Block diagram of the superimposed training scheme for STBC-OFDM

To simplify our discussion, this chapter only considers the CP-OFDM transmission over the frequency-selective fading channels. Let $\mathcal{D}_i = \text{diag}\{H_i(0) \cdots H_i(M-1)\}$ be the channel matrices, the received post-FFT blocks $\mathbf{y}(n)$ can be expressed as

$$\mathbf{y}(n) = \sum_{i=1}^{2} \mathcal{D}_i \bar{\mathbf{s}}_i(n) + \mathbf{v}(n) \tag{5.4.3}$$

Using (5.4.1) and (5.4.2), the consecutive received blocks, $\mathbf{y}(2n)$ and $\mathbf{y}(2n+1)$, is given by

$$\mathbf{y}(2n) = \mathcal{D}_1 \mathbf{s}(2n) + \mathcal{D}_2 \mathbf{s}(2n+1) + \sum_{i=1}^{2} \mathcal{D}_i \mathbf{c}_i(2n) + \mathbf{v}(2n) \tag{5.4.4}$$

$$\mathbf{y}(2n+1) = -\mathcal{D}_1 \mathbf{s}^*(2n+1) + \mathcal{D}_2 \mathbf{s}^*(2n) + \sum_{i=1}^{2} \mathcal{D}_i \mathbf{c}_i(2n+1) + \mathbf{v}(2n+1) \tag{5.4.5}$$

Next, we define

$$\bar{\mathbf{y}}(n) = \mathbf{y}(n) - \sum_{i=1}^{2} \mathcal{D}_i \mathbf{c}_i(n) \tag{5.4.6}$$

$$\check{\mathbf{y}}(n) = [\bar{\mathbf{y}}^T(2n),\ \bar{\mathbf{y}}^{\mathcal{H}}(2n+1)]^T \tag{5.4.7}$$

$$\check{\mathbf{s}}(n) = [\mathbf{s}^T(2n),\ \mathbf{s}^T(2n+1)]^T \tag{5.4.8}$$

$$\mathcal{D} = \begin{pmatrix} \mathcal{D}_1 & \mathcal{D}_2 \\ \mathcal{D}_2^* & -\mathcal{D}_1^* \end{pmatrix} \tag{5.4.9}$$

Then, (5.4.4) and (5.4.5) can be represented as

$$\check{\mathbf{y}}(n) = \mathcal{D}\check{\mathbf{s}}(n) + \check{\mathbf{v}}(n) \tag{5.4.10}$$

where $\check{\mathbf{v}}(n) = [\mathbf{v}^T(2n),\ \mathbf{v}^T(2n+1)]^T$ is the corresponding noise vector. When the channel matrices $\mathcal{D}_1$ and $\mathcal{D}_2$ become available at the receiver, it is possible to demodulate $\check{\mathbf{y}}(n)$ by a simple matrix multiplication

$$\begin{aligned} \check{\mathbf{d}}(n) &= \mathcal{D}^{\mathcal{H}} \check{\mathbf{y}}(n) \\ &= \begin{pmatrix} \bar{\mathcal{D}}_{12} & 0 \\ 0 & \bar{\mathcal{D}}_{12} \end{pmatrix} \check{\mathbf{s}}(n) + \mathcal{D}^{\mathcal{H}} \check{\mathbf{v}}(n) \end{aligned} \tag{5.4.11}$$

where the diagonal matrix $\bar{\mathcal{D}}_{12} = \mathcal{D}_1^*\mathcal{D}_1 + \mathcal{D}_2^*\mathcal{D}_2$. Equation (5.4.11) reveals that ZF recovery of $\check{\mathbf{s}}(n)$ from $\check{\mathbf{d}}(n)$ requires the matrices $\bar{\mathcal{D}}_{12}$ to have full column rank. This condition is assured only when the matrices $\mathcal{D}_1$ and $\mathcal{D}_2$ do not share common zeros along the diagonal.

Based on (5.4.10), we next address the superimposed training based channel estimation algorithm and its performance analysis.

5.4.2 Channel estimation algorithm

The channel estimation algorithm is based on the first-order statistics of the signal model (5.4.10), which is given by

$$\mathcal{E}\{\check{\mathbf{y}}(n)\} = \mathcal{D}\mathcal{E}\{\check{\mathbf{s}}(n)\} + \mathcal{E}\{\check{\mathbf{v}}(n)\} = \mathbf{0} \tag{5.4.12}$$

From (5.4.6) and (5.4.7), it can be derived that

$$\begin{aligned} \mathcal{E}\{\bar{\mathbf{y}}(n)\} &= \mathcal{E}\{\mathbf{y}(n)\} - \sum_{i=1}^{2} \mathcal{D}_i \mathcal{E}\{\mathbf{c}_i(n)\} \\ &= \mathbf{m}_y - \sum_{i=1}^{2} \mathcal{D}_i \mathbf{c}_i(0) \\ &= \mathbf{0} \end{aligned} \tag{5.4.13}$$

which further leads to

$$\begin{aligned} \mathbf{m}_y &= \sum_{i=1}^{2} \mathcal{D}_i \mathbf{c}_i(0) \\ &= \sqrt{M} \sum_{i=1}^{2} \mathbf{D}(\mathbf{c}_i(0))\mathbf{F}_{\mathbf{h}}\mathbf{h}_i \end{aligned} \tag{5.4.14}$$

Let $\mathbf{Q}_s = [\mathbf{D}(\mathbf{c}_1(0))\mathbf{F}_{\mathbf{h}},\ \mathbf{D}(\mathbf{c}_2(0))\mathbf{F}_{\mathbf{h}}]$ and $\check{\mathbf{h}} = [\mathbf{h}_1^T,\ \mathbf{h}_2^T]^T$, (5.4.14) can be rewritten into the following matrix form

$$\mathbf{m}_y = \sqrt{M}\mathbf{Q}_s\check{\mathbf{h}} \tag{5.4.15}$$

It can be seen that the matrix $\mathbf{Q}_s$ has size $M \times 2L_u$. If the rank of $\mathbf{Q}_s$ is $2L_u$, then the channel vector $\breve{\mathbf{h}}$ can be uniquely identified using

$$\hat{\breve{\mathbf{h}}} = \frac{1}{\sqrt{M}}\mathbf{Q}_s^{\dagger}\mathbf{m}_y \tag{5.4.16}$$

To guarantee the full column rank of $\mathbf{Q}_s$, the training sequences $\mathbf{c}_1$ and $\mathbf{c}_2$ must be carefully designed.

5.4.3 Performance analysis

To analyze the channel estimate performance, we represent (5.4.13) as

$$\begin{aligned} \mathcal{E}\{\bar{\mathbf{y}}(n)\} &= \mathbf{m}_y - \sqrt{M}\mathbf{Q}_s\breve{\mathbf{h}} \\ &= \sum_{i=1}^{2} \mathcal{D}_i \mathcal{E}\{\bar{\mathbf{s}}_i(n)\} + \mathcal{E}\{\mathbf{v}(n)\} \end{aligned} \tag{5.4.17}$$

and the channel estimate error is given by

$$\begin{aligned} \frac{1}{\sqrt{M}}\mathbf{Q}_s^{\dagger}\mathcal{E}\{\bar{\mathbf{y}}(n)\} &= \hat{\breve{\mathbf{h}}} - \breve{\mathbf{h}} \\ &= \frac{1}{\sqrt{M}}\mathbf{Q}_s^{\dagger}\sum_{i=1}^{2} \mathcal{D}_i \mathcal{E}\{\bar{\mathbf{s}}_i(n)\} + \frac{1}{\sqrt{M}}\mathbf{Q}_s^{\dagger}\mathcal{E}\{\mathbf{v}(n)\} \end{aligned} \tag{5.4.18}$$

Similar to the SISO cases, $\mathcal{E}\{\bar{\mathbf{s}}_i(n)\}$ and $\mathcal{E}\{\mathbf{v}(n)\}$ are replaced by the assemble average

$$\mathcal{E}\{\bar{\mathbf{s}}_i(n)\} = \frac{1}{N}\sum_{n=0}^{N-1} \bar{\mathbf{s}}_i(n) \tag{5.4.19}$$

$$\mathcal{E}\{\mathbf{v}(n)\} = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{v}(n) \tag{5.4.20}$$

The mean-square channel estimate error is therefore derived as

$$\begin{aligned} \epsilon &= \mathcal{E}\|\hat{\breve{\mathbf{h}}} - \breve{\mathbf{h}}\|^2 \\ &= \epsilon_s + \epsilon_v \end{aligned} \tag{5.4.21}$$

where

$$\begin{aligned}\epsilon_s &= \frac{1}{M}\mathrm{Tr}\{\mathbf{Q}_s^\dagger \sum_{j=1}^{2}\sum_{i=1}^{2} \mathcal{D}_j \mathcal{E}\{\bar{\mathbf{s}}_j(n)\}\mathcal{E}\{\bar{\mathbf{s}}_i(n)\}^{\mathcal{H}}\mathcal{D}_i^{\mathcal{H}}(\mathbf{Q}_s^\dagger)^{\mathcal{H}}\} \\ &= \frac{\sigma_s^2}{NM}\mathrm{Tr}\{\mathbf{Q}_s^\dagger \sum_{i=1}^{2} \mathcal{D}_i\mathcal{D}_i^{\mathcal{H}}(\mathbf{Q}_s^\dagger)^{\mathcal{H}}\} \qquad (5.4.22)\end{aligned}$$

$$\begin{aligned}\epsilon_v &= \frac{1}{M}\mathrm{Tr}\{\mathbf{Q}_s^\dagger \mathcal{E}\{\mathbf{v}(n)\}\mathcal{E}\{\mathbf{v}(n)\}^{\mathcal{H}}(\mathbf{Q}_s^\dagger)^{\mathcal{H}}\} \\ &= \frac{1}{NM}\mathrm{Tr}\{\mathbf{Q}_s^\dagger \mathrm{diag}\{\sigma_{v,0}^2,\cdots,\sigma_{v,M-1}^2\}(\mathbf{Q}_s^\dagger)^{\mathcal{H}}\} \qquad (5.4.23)\end{aligned}$$

Define a diagonal matrix

$$\mathbf{\Lambda}_e = \sigma_s^2 \sum_{i=1}^{2} \mathcal{D}_i\mathcal{D}_i^{\mathcal{H}} + \mathrm{diag}\{\sigma_{v,0}^2,\cdots,\sigma_{v,M-1}^2\} \qquad (5.4.24)$$

The channel estimate error ϵ can be represented as

$$\epsilon = \frac{1}{NM}\mathrm{Tr}\{\mathbf{Q}_s^\dagger \mathbf{\Lambda}_e (\mathbf{Q}_s^\dagger)^{\mathcal{H}}\} \qquad (5.4.25)$$

5.4.4 Superimposed training sequences design

The objective of the superimposed training sequences design is to minimize the mean-square channel estimate error and reduce the computation complexity.

Theorem 5.4.1. *The channel estimate error (5.4.25) achieves its minimum*

$$\epsilon_{\min} = \frac{4L_u}{N\sum_{m=0}^{M-1}\frac{|c_1(m)|^2+|c_2(m)|^2}{\Lambda_e(m)}} \qquad (5.4.26)$$

only when the following condition is satisfied

$$\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^\dagger\mathbf{Q}_s = \frac{1}{2M}\sum_{m=0}^{M-1}\frac{|c_1(m)|^2+|c_2(m)|^2}{\Lambda_e(m)}\mathbf{I}_{2L_u}$$

where, $\Lambda_e(m) = \sigma_s^2(|H_1(m)|^2 + |H_2(m)|^2) + \sigma_{v,m}^2$ *denotes the* (m,m)*th entry of* $\mathbf{\Lambda}_e$.

Proof. See Appendix 5.D.

Theorem 5.4.2. *The condition (5.6.19) occurs when the superimposed training sequences* $\mathbf{c}_i(n)$ *satisfy*

$$\mathbf{D}(\mathbf{c}_1(n)) = \mathbf{D}_f^{\mathcal{H}}\mathbf{D}(\mathbf{c}_2(n)) \tag{5.4.27}$$

where $\mathbf{D}_f = \text{diag}\{1, W_M^{L_u}, \cdots, W_M^{L_u(M-1)}\}$.

Proof See Appendix 5.E.

Theorems 5.4.1 and 5.4.2 show that the mean-square channel estimate error can be minimized when the training sequences satisfy (5.4.27). Using this design, the matrix $\mathbf{Q}_s$ can be represented as

$$\begin{aligned} \mathbf{Q}_s &= [\mathbf{D}(\mathbf{c}_1(0))\mathbf{F}_\mathbf{h},\ \mathbf{D}_f\mathbf{D}(\mathbf{c}_1(0))\mathbf{F}_\mathbf{h}] \\ &= \mathbf{D}(\mathbf{c}_1(0))\bar{\mathbf{F}}_\mathbf{h} \end{aligned} \tag{5.4.28}$$

Substituting this result into (5.4.16), the channel estimation can be implemented as

$$\begin{aligned} \hat{\mathbf{h}} &= \frac{1}{\sqrt{M}}\bar{\mathbf{F}}_\mathbf{h}^\dagger\mathbf{D}^\dagger(\mathbf{c}_1(0))\mathbf{m}_y \\ &= \frac{1}{\sqrt{M}}\bar{\mathbf{F}}_\mathbf{h}^{\mathcal{H}}\mathbf{D}^\dagger(\mathbf{c}_1(0))\mathbf{m}_y \end{aligned} \tag{5.4.29}$$

From (5.4.16) and (5.4.29), we can see that the inverse of $\mathbf{Q}_s$ may introduce $\mathcal{O}(M^2 \log M)$ complexity into (5.4.16). By employing the carefully designed training sequences, (5.4.29) can be implemented using IFFT, and thus reduces the complexity to $\mathcal{O}(M \log M)$.

5.5 Simulations

The simulation results include two examples: 1) the mean-square channel estimate error and the overall system performance are examined in static channels, and 2) the overall system performance is also evaluated in slowly time varying channels. The

system parameters are set as $L_u = 4$, $M = 32$, which is the typical settings for OFDM-based WLAN (e.g. HIPERLAN). The information bearing symbols are drawn from the QPSK constellation with the equal probability. The frequency selective channels are modelled as a FIR filter. SNR for the superimposed training scheme is defined as the total transmitted energy from transmit antennas and noise ratio. Another important factor used in the simulations is the training to information ratio (TIR), which is defined by $\|\mathbf{c}(0)\|^2/M\sigma_s^2$.

`Test Case 1 (static channels):` The channel taps are randomly generated with equal variance $1/L_u$. The total variance for all taps of each channel is set to unity: $\mathcal{E}\{\|\mathbf{h}_i\|^2\} = 1$, $\forall i \in \{1,2\}$. The simulation results are averaged over 500 channels. We first investigate the channel estimate performance for the superimposed training scheme in the SISO scenario. The mean-square channel estimate error is defined as

$$\text{MSE} = \frac{1}{IL_u}\sum_{i=1}^{I}\|\hat{\mathbf{h}}_i - \mathbf{h}_i\|^2$$

where $I = 500$ denotes the number of random channels. Each channel estimation collects $N = 100$ symbol blocks for the first-order statistics. Using the optimum training sequences designed for minimizing the channel estimate error, Figure 5.3 illustrates the channel estimate performance for CP-OFDM systems. It shows that:

1) The simulation results fit very well with theoretical MSE bounds, particularly within the SNR range from $14 - 20$ dB.

2) When TIR is increased from 0.033 to 0.25, the channel estimate performance is significantly improved. This result confirms our theoretical MSE analysis. However, increasing of TIR will lower the effective SNR.

3) In static channels, the training-based scheme outperforms the superimposed

training scheme. However, the training approach needs extra bandwidth consumption. The performance of superimposed training can be further improved by collecting more symbol blocks for the estimation.

Using the estimated CSI, Figure 5.4 illustrates the equalization performance in various TIR cases. It shows that:

4) With the knowledge of the perfect CSI, the case with lower TIR (= 0.033) outperforms the case with higher TIR (= 0.25), because it has a higher effective SNR.

5) When TIR is low (e.g. 0.033), the BER performance is sensitive to the data record length, N. The performance can be significantly improved by increasing N. When TIR is high (e.g. 0.25), the channel estimation is very accurate even for $N = 100$, such that their BER performance is close to the case with the perfect CSI.

Next, we investigate the overall system performance for STBC-OFDM systems. The results reflected in Figure 5.5 are similar to 4) and 5) for uncoded CP-OFDM systems. Thanks to the STC technique, when the channel estimate is accurate (for TIR= 0.25), the BER performances are very close to the case with the perfect CSI. Since the training-based approach does not reduce the effective SNR, its BER performance is better than that of the superimposed training approach in static channels.

`Test Case 2 (slowly time-varying channels)`: The simulation environment used here is the same as that in Chapter 4.4. Since test case 1 has shown that uncoded OFDM and STBC-OFDM have many common features in the overall system performance, this test case only examines STBC-OFDM for concise presentation.

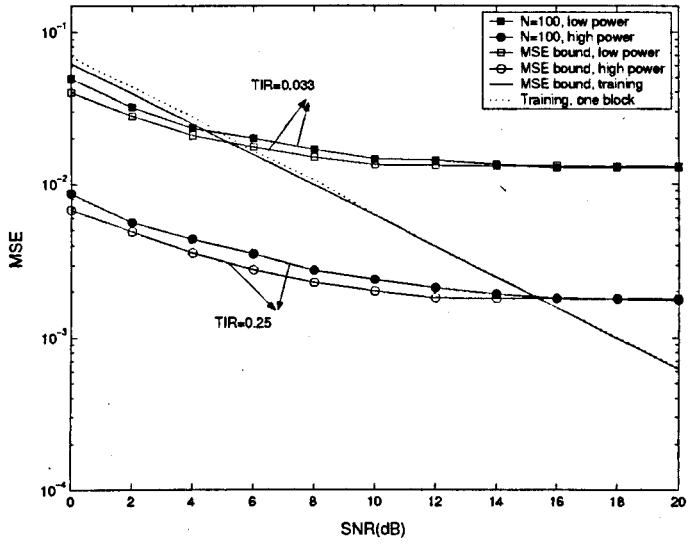

Figure 5.3: Mean-square channel estimate error in static channels

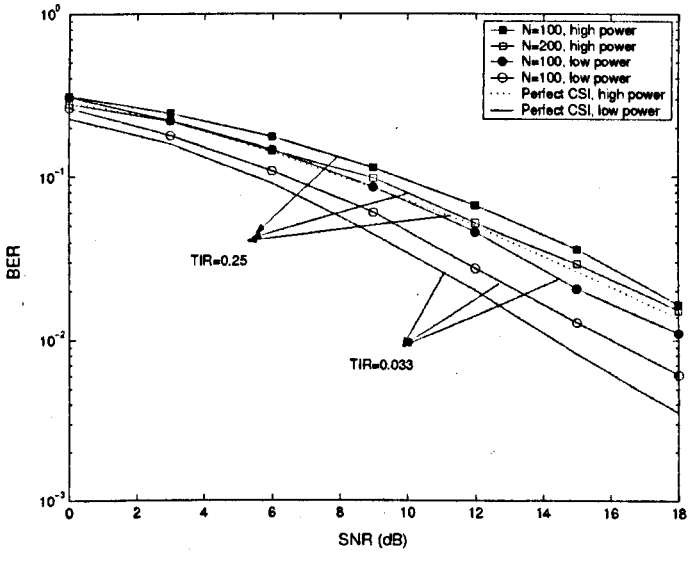

Figure 5.4: BER Vs SNR for CP-OFDM in static channels

Figure 5.6 shows that the training-based approach using one training block cannot track the time-varying channel and results in large error floor in relatively high SNR (e.g. > 12 dB). The superimposed training scheme demonstrates its good performance particularly when DRL is large, e.g $N = 200$. In fact, its performance can be further improved by increasing N. Similar to the case in static channels, TIR is also a very important factor to affect the channel estimate and overall system performance in time-varying channels.

One of major drawbacks of the superimposed-training based scheme is its relatively low effective SNR, which results in the loss of the overall system performance [71]-[72]. Hence, as shown in Figures 5.4 and 5.5, the BER performance is very sensitive to the TIR. Due to the low effective SNR, the BER is not good even for the case with the perfect channel knowledge. Although the superimposed-training scheme offers low-complexity channel estimator, it is very sensitive to channel variation. This is the reason why Figure 5.6 shows the poor performance. My future research will consider a superimposed-training scheme for BP-OFDM to improve the effective SNR. New channel estimation and equalization algorithm will also be considered for tracking the time-varying channel.

It should be noted here, the simulation results for ZP-OFDM are very similar to CP-OFDM. So, this document will not provide the simulations and analysis for ZP-OFDM.

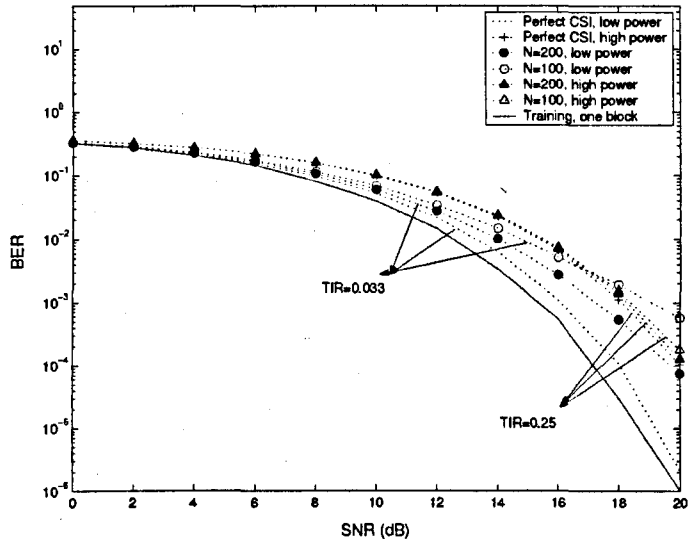

Figure 5.5: BER Vs SNR for STBC-OFDM in static channels

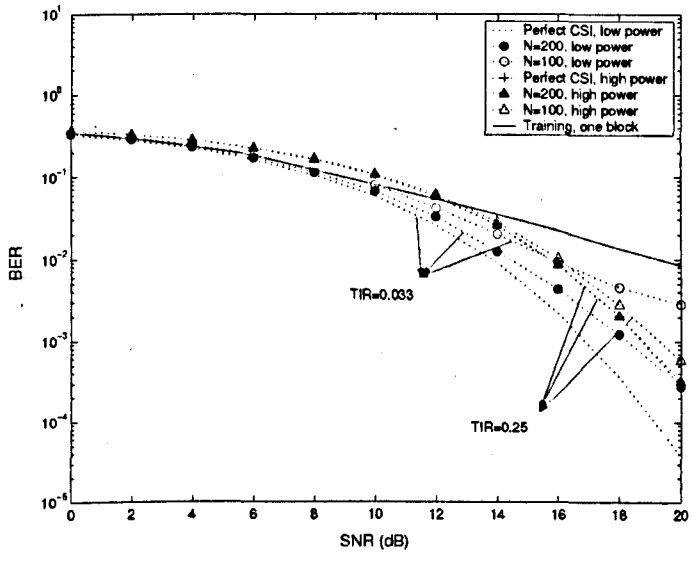

Figure 5.6: BER Vs SNR for STBC-OFDM in slowly time-varying channels

5.6 Conclusion

This chapter introduced a superimposed training scheme for various OFDM systems, including CP-OFDM, ZP-OFDM and STBC-OFDM, to track frequency selective fading channels. The channel estimator only used the first-order statistics and did not consume extra bandwidth. The mean-square channel estimate error was carefully analyzed, and MSE bounds for various OFDM systems were derived to examine the channel estimate performance. Optimum superimposed training sequences were generated to minimizing the MSE performances. Simulation results were provided to validate the theoretical analysis. It was shown that the major drawback for the superimposed training scheme is the loss of effective SNR. In my future research, this problem will be solved by employing linear block precoding in OFDM systems.

Appendix 5.A: Proof of Theorem 5.3.1

Based on (5.3.7), the MSE can be derived as

$$\begin{aligned}
\mathcal{E}\|\mathbf{e}_{\mathrm{h}}\|^2 &= \mathrm{Tr}\{\mathcal{E}\{\mathbf{e}_{\mathrm{h}}\mathbf{e}_{\mathrm{h}}^{\mathcal{H}}\}\} \\
&= \mathrm{Tr}\{\mathcal{E}\{(\mathbf{Q}_c^{\dagger}\mathbf{e}_s + \mathbf{Q}_c^{\dagger}\mathbf{e}_v)(\mathbf{Q}_c^{\dagger}\mathbf{e}_s + \mathbf{Q}_c^{\dagger}\mathbf{e}_v)^{\mathcal{H}}\}\} \\
&= \frac{1}{\sqrt{M}}\mathrm{Tr}\{\mathbf{Q}_c^{\dagger}\mathcal{E}\{\mathbf{e}_s\mathbf{e}_s^{\mathcal{H}}\}(\mathbf{Q}_c^{\dagger})^{\mathcal{H}}\} + \frac{1}{\sqrt{M}}\mathrm{Tr}\{\mathbf{Q}_c^{\dagger}\mathcal{E}\{\mathbf{e}_v\mathbf{e}_v^{\mathcal{H}}\}(\mathbf{Q}_c^{\dagger})^{\mathcal{H}}\} \quad (5.6.1)
\end{aligned}$$

For

$$\begin{aligned}
\mathcal{E}\{\mathbf{e}_s\mathbf{e}_s^{\mathcal{H}}\} &= \frac{1}{N^2}\mathcal{D}_M \sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} \mathcal{E}\{\mathbf{s}(n_1)\mathbf{s}^{\mathcal{H}}(n_2)\}\mathcal{D}_M^{\mathcal{H}} \\
&= \frac{\sigma_s^2}{N}\mathcal{D}_M\mathcal{D}_M^{\mathcal{H}} \quad (5.6.2)
\end{aligned}$$

$$\begin{aligned}\mathcal{E}\{\mathbf{e}_v\mathbf{e}_v^{\mathcal{H}}\} &= \frac{1}{N^2}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}\mathcal{E}\{\tilde{\mathbf{v}}(n_1)\tilde{\mathbf{v}}^{\mathcal{H}}(n_2)\} \\ &= \frac{1}{N}\mathrm{diag}\{\sigma_{v,0}^2, \cdots, \sigma_{v,M-1}^2\} \end{aligned} \tag{5.6.3}$$

(5.6.1) can be represented as

$$\begin{aligned}\mathcal{E}\|\mathbf{e}_\mathbf{h}\|^2 &= \frac{\sigma_s^2}{NM}\mathrm{Tr}\{\mathbf{Q}_c^\dagger\boldsymbol{\mathcal{D}}_M\boldsymbol{\mathcal{D}}_M^{\mathcal{H}}(\mathbf{Q}_c^\dagger)^{\mathcal{H}}\} + \frac{\sigma_v^2}{NM}\mathrm{Tr}\{(\mathbf{Q}_c^{\mathcal{H}}\mathbf{Q}_c)^\dagger\} \\ &= \frac{\sigma_s^2}{NM}\mathrm{Tr}\{\mathbf{F}_\mathbf{h}^\dagger\mathbf{D}^\dagger(\mathbf{c}(0))\boldsymbol{\mathcal{D}}_M\boldsymbol{\mathcal{D}}_M^{\mathcal{H}}(\mathbf{D}^\dagger(\mathbf{c}(0)))^{\mathcal{H}}(\mathbf{F}_\mathbf{h}^\dagger)^{\mathcal{H}}\} \\ &+ \frac{\sigma_v^2}{NM}\mathrm{Tr}\{\mathbf{F}_\mathbf{h}^\dagger\mathbf{D}^\dagger(\mathbf{c}(0))(\mathbf{D}^\dagger(\mathbf{c}(0)))^{\mathcal{H}}(\mathbf{F}_\mathbf{h}^\dagger)^{\mathcal{H}}\} \\ &= \frac{\sigma_s^2 L_u}{NM^2}\sum_{m=0}^{M-1}\frac{|H(m)|^2}{|c(m)|^2} + \frac{L_u}{NM^2}\sum_{m=0}^{M-1}\frac{\sigma_{v,m}^2}{|c(m)|^2} \end{aligned} \tag{5.6.4}$$

Appendix 5.B: Proof of Theorem 5.3.2

Rewrite the MSE result (5.3.8) into

$$\mathcal{E}\|\mathbf{e}_\mathbf{h}\|^2 = \frac{L_u}{NM^2}\sum_{m=0}^{M-1}\frac{\sigma_e^2(m)}{|c(m)|^2} \tag{5.6.5}$$

where

$$\sigma_e^2(m) = \sigma_s^2|H(m)|^2 + \sigma_{v,m}^2 \tag{5.6.6}$$

then, the MSE result for the superimposed training scheme has the similar form to Eqn. (2.2.7). Theorem 2.2.1 has shown that the optimum training sequence design for the model (5.6.5) is

$$|c(0)|^2 = \cdots = |c(M-1)|^2 = \xi_c/M \tag{5.6.7}$$

Plugging (5.6.7) into (5.6.4), the result (5.3.11) can be easily obtained, and Theorem 5.3.2 is therefore proved.

Appendix 5.C: Proof of Theorem 5.3.3

The mean-square channel estimate error can be derived from (5.3.17)

$$\begin{aligned}\mathcal{E}\|\mathbf{e_h}\|^2 &= \mathcal{E}\|\hat{\mathbf{h}}-\mathbf{h}\|^2 \\ &= \mathrm{Tr}\{\frac{1}{J}\mathbf{Q}_z^\dagger\mathcal{E}\{\bar{\mathbf{e}}_s\bar{\mathbf{e}}_s^{\mathcal{H}}\}(\mathbf{Q}_z^\dagger)^{\mathcal{H}}\}+\mathrm{Tr}\{\frac{1}{J}\mathbf{Q}_z^\dagger\mathcal{E}\{\bar{\mathbf{e}}_v\bar{\mathbf{e}}_v^{\mathcal{H}}\}(\mathbf{Q}_z^\dagger)^{\mathcal{H}}\} \qquad (5.6.8)\end{aligned}$$

Using (5.3.15) and (5.3.16), the following results can be obtained

$$\mathcal{E}\{\bar{\mathbf{e}}_s\bar{\mathbf{e}}_s^{\mathcal{H}}\}=\frac{\sigma_s^2}{N}\boldsymbol{\mathcal{D}}_J\mathbf{G}_{\mathrm{zp}}\mathbf{G}_{\mathrm{zp}}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_J^{\mathcal{H}} \qquad (5.6.9)$$

$$\mathcal{E}\{\bar{\mathbf{e}}_v\bar{\mathbf{e}}_v^{\mathcal{H}}\}=\frac{1}{N}\mathrm{diag}\{\sigma_{v,0}^2,\ \cdots,\sigma_{v,J-1}^2\} \qquad (5.6.10)$$

So, the error arising from the additive noise is given by

$$\begin{aligned}\epsilon_v &= \mathrm{Tr}\{\frac{1}{J}\mathbf{Q}_z^\dagger\mathcal{E}\{\bar{\mathbf{e}}_v\bar{\mathbf{e}}_v^{\mathcal{H}}\}(\mathbf{Q}_z^\dagger)^{\mathcal{H}}\} \\ &= \mathrm{Tr}\{\frac{1}{J}\bar{\mathbf{F}}_{\mathbf{h}}^\dagger\mathbf{D}^\dagger(\mathbf{G}_{\mathrm{zp}}\mathbf{c}(0))\mathcal{E}\{\bar{\mathbf{e}}_v\bar{\mathbf{e}}_v^{\mathcal{H}}\}(\mathbf{D}^\dagger(\mathbf{G}_{\mathrm{zp}}\mathbf{c}(0)))^{\mathcal{H}}(\bar{\mathbf{F}}_{\mathbf{h}}^\dagger)^{\mathcal{H}}\} \\ &= \frac{L_u}{NJ^2}\sum_{m=0}^{J-1}\frac{\sigma_{v,m}^2}{|\mathbf{g}_{\mathrm{zp}}(m)\mathbf{c}(0)|^2} \qquad (5.6.11)\end{aligned}$$

where, $\mathbf{g}_{\mathrm{zp}}(m)$ denotes the row vector of the matrix $\mathbf{G}_{\mathrm{zp}}$. Next, we calculate the error induced by the signal part

$$\begin{aligned}\epsilon_s &= \mathrm{Tr}\{\frac{1}{J}\mathbf{Q}_z^\dagger\mathcal{E}\{\bar{\mathbf{e}}_s\bar{\mathbf{e}}_s^{\mathcal{H}}\}(\mathbf{Q}_z^\dagger)^{\mathcal{H}}\} \\ &= \frac{\sigma_s^2}{NJ}\mathrm{Tr}\{\bar{\mathbf{F}}_{\mathbf{h}}^\dagger\underbrace{\mathbf{D}^\dagger(\mathbf{G}_{\mathrm{zp}}\mathbf{c}(0))\boldsymbol{\mathcal{D}}_J\mathbf{G}_{\mathrm{zp}}\mathbf{G}_{\mathrm{zp}}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_J^{\mathcal{H}}(\mathbf{D}^\dagger(\mathbf{G}_{\mathrm{zp}}\mathbf{c}(0)))^{\mathcal{H}}}_{\mathbf{H}_e}(\bar{\mathbf{F}}_{\mathbf{h}}^\dagger)^{\mathcal{H}}\} \qquad (5.6.12)\end{aligned}$$

It is easy to see that $\mathbf{H}_e$ is a Hermitian matrix. So, its eigendecomposition is given by

$$\mathbf{H}_e=\mathcal{U}^{\mathcal{H}}\Lambda_e\mathcal{U} \qquad (5.6.13)$$

where $\mathcal{U}$ is the unitary matrix, and $\mathbf{\Lambda}_e$ is the real diagonal matrix. Then, (5.6.12) can be rewritten into

$$\begin{aligned}\epsilon_s &= \frac{\sigma_s^2}{NJ}\mathrm{Tr}\{\bar{\mathbf{F}}_{\mathbf{h}}^{\mathcal{H}}\mathcal{U}^{\mathcal{H}}\mathbf{\Lambda}_e\mathcal{U}\bar{\mathbf{F}}_{\mathbf{h}}\} \\ &= \frac{\sigma_s^2}{NJ}\mathrm{Tr}\{\mathbf{U}\mathbf{\Lambda}_e\mathbf{U}^{\mathcal{H}}\}\end{aligned} \tag{5.6.14}$$

where $\mathbf{U} = \bar{\mathbf{F}}_{\mathbf{h}}^{\mathcal{H}}\mathcal{U}^{\mathcal{H}}$. Let $\mathbf{u}(j)$ be the jth column vector of $\mathbf{U}$, and $\lambda_e(j)$ be the (j,j)th entry of $\mathbf{\Lambda}_e$, then ϵ_s is given by

$$\epsilon_s = \frac{\sigma_s^2}{NJ}\sum_{j=0}^{J-1}\|\mathbf{u}(j)\|^2\lambda_e(j) \tag{5.6.15}$$

Plugging (5.6.11) and (5.6.15) into (5.6.8), the channel estimate error (5.3.18) is therefore obtained.

Appendix 5.D: Proof of Theorem 5.4.1

Consider the channel estimate error

$$\begin{aligned}\epsilon &= \frac{1}{NM}\mathrm{Tr}\{\mathbf{Q}_s^{\dagger}\mathbf{\Lambda}_e(\mathbf{Q}_s^{\dagger})^{\mathcal{H}}\} \\ &= \frac{1}{NM}\mathrm{Tr}\{(\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s)^{\dagger}\}\end{aligned} \tag{5.6.16}$$

(5.4.24) shows that $\mathbf{\Lambda}_e$ is a real diagonal matrix, so $\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s$ forms a $2L_u \times 2L_u$ Hermitian matrix. Let $\lambda_0, \cdots, \lambda_{2L_u-1}$ be eigenvalues of this Hermitian matrix, its trace is given by

$$\begin{aligned}\mathrm{Tr}\{\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s\} &= \sum_{l=0}^{2L_u-1}\lambda_l \\ &= \frac{L_u}{M}\sum_{m=0}^{M-1}\frac{|c_1(m)|^2 + |c_2(m)|^2}{\Lambda_e(m)}\end{aligned} \tag{5.6.17}$$

Next, the channel estimate error can be calculated as

$$\begin{aligned}\epsilon &= \frac{1}{NM}\mathrm{Tr}\{(\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s)^{\dagger}\} \\ &= \frac{1}{NM}\sum_{l=0}^{2L_u-1}\frac{1}{\lambda_l} \\ &\geqslant \frac{2L_u}{NM}\frac{1}{\sqrt[2L_u]{\lambda_0\cdots\lambda_{2L_u-1}}} \\ &\geqslant \frac{(2L_u)^2}{NM}\frac{1}{\sum_{l=0}^{2L_u-1}\lambda_l}\end{aligned} \tag{5.6.18}$$

The equality occurs only when all the eigenvalues $\{\lambda_l\}$ are identical and

$$\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s = \frac{1}{2M}\sum_{m=0}^{M-1}\frac{|c_1(m)|^2+|c_2(m)|^2}{\Lambda_e(m)}\mathbf{I}_{2L_u} \tag{5.6.19}$$

Thus, the error, ϵ, is possible to achieve its minimum

$$\epsilon_{\min} = \frac{4L_u}{N\sum_{m=0}^{M-1}\frac{|c_1(m)|^2+|c_2(m)|^2}{\Lambda_e(m)}} \tag{5.6.20}$$

Appendix 5.E: Proof of Theorem 5.4.2

Define an $M \times 2L_u$ matrix

$$\bar{\mathbf{Q}}_s = \sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{Q}_s \tag{5.6.21}$$

and

$$\lambda = \frac{1}{2M}\sum_{m=0}^{M-1}\frac{|c_1(m)|^2+|c_2(m)|^2}{\Lambda_e(m)} \tag{5.6.22}$$

(5.6.19) can be rewritten into

$$\begin{aligned}\mathbf{Q}_s^{\mathcal{H}}\mathbf{\Lambda}_e^{\dagger}\mathbf{Q}_s &= \lambda\mathbf{I}_{2L_u} \\ &= \bar{\mathbf{Q}}_s^{\mathcal{H}}\bar{\mathbf{Q}}_s \\ &= \lambda\bar{\mathbf{F}}_{\mathbf{h}}^{\mathcal{H}}\bar{\mathbf{F}}_{\mathbf{h}}\end{aligned} \tag{5.6.23}$$

where $\bar{\mathbf{F}}_{\mathbf{h}}$ is formed by collecting the first $2L_u$ columns of the DFT matrix $\mathbf{F}_M$. (5.6.23) reveals that, the condition (5.6.19) can be satisfied when

$$\begin{aligned}\sqrt{\lambda}\bar{\mathbf{F}}_{\mathbf{h}} &= \bar{\mathbf{Q}}_s \\ &= [\sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{D}(\mathbf{c}_1(0))\mathbf{F}_{\mathbf{h}},\ \sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{D}(\mathbf{c}_2(0))\mathbf{F}_{\mathbf{h}}] \\ &= [\sqrt{\lambda}\mathbf{F}_{\mathbf{h}}, \sqrt{\lambda}\mathbf{D}_f\mathbf{F}_{\mathbf{h}}] \end{aligned} \tag{5.6.24}$$

Thus, we have

$$\sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{D}(\mathbf{c}_1(0)) = \mathbf{D}_f^{\dagger}\sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{D}(\mathbf{c}_2(0)) \tag{5.6.25}$$

which results in

$$\begin{aligned}\mathbf{D}(\mathbf{c}_1(0)) &= \sqrt{\mathbf{\Lambda}_e}\mathbf{D}_f^{\dagger}\sqrt{\mathbf{\Lambda}_e^{\dagger}}\mathbf{D}(\mathbf{c}_2(0)) \\ &= \mathbf{D}_f^{\dagger}\mathbf{D}(\mathbf{c}_2(0)) \\ &= \mathbf{D}_f^{\mathcal{H}}\mathbf{D}(\mathbf{c}_2(0)) \end{aligned} \tag{5.6.26}$$

For $\mathbf{c}_i(n) = \mathbf{c}_i(0)$, Theorem 5.4.2 is therefore proved.

Chapter 6

Channel estimation for MIMO-OFDM using guard interval diversity

This chapter introduces a novel blind channel estimation method for MIMO-OFDM by exploiting guard interval diversity introduced by the transmitter. This blind channel estimator uses the second-order cyclostationary to yield unique channel estimates. To enable the proposed algorithm, the transmitter transmits uncorrelated CP-OFDM and ZP-OFDM signals from two different transmit antennas. Using this technique, the proposed blind channel estimator can deal with subchannels individually. Simulation results demonstrate the comparable channel estimate performance of this new method. This chapter is partly presented in [87].

6.1 Introduction

Deploying multiple antennas at both the transmitter and the receiver of OFDM systems has recently received considerable attentions to improve channel capacity (see [73]-[74]). Similar to other MIMO systems, signal recovery for MIMO-OFDM needs the knowledge of multichannel information. If signals from different antennas are

uncorrelated, [61] exploits non-constant antenna precoding (window) for the multichannel estimation. Thanks to the distinct precodings, the blind estimator can deal with subchannels individually. However, the channel estimation performance is not good enough. Recently, a subspace-based blind channel estimation algorithm has been proposed for block precoded ST-OFDM systems. It uses the precoding redundancy to yield a unique channel estimate. However, the subspace approach needs relatively large DRL and high computation complexity. It is only applicable for BP-OFDM systems. The objective of this chapter is to propose a new blind channel estimation algorithm for MIMO-OFDM systems, which has good channel estimate performance and fast convergence.

As introduced in Chapter 1, conventional OFDM has two transmission schemes, i.e. CP-OFDM and ZP-OFDM transmissions. The length of CP or ZP is normally longer than the upper bound of the channel order to avoid the ISI. Relying on the redundancy introduced by CPs or ZPs, a cyclostationary based blind channel estimation algorithm has been proposed for SISO systems (see [57]). The proposed estimator demonstrates its good estimate performance and fast convergence. We notice that the cyclostationarity for CP-OFDM and ZP-OFDM is different. If the transmitter equipped with two antennas can transmit CP-OFDM and ZP-OFDM signals simultaneously, it is possible to design a multichannel estimator, which may enable unique multichannel estimation. This motivates us to propose the following MIMO-OFDM transmission scheme with guard interval diversity.

6.2 MIMO-OFDM with guard interval diversity

Consider a wireless communication system equipped with two transmit antennas and two receive antennas. Figure 6.1 depicts the MIMO-OFDM transceiver with guard interval diversity. Prior to transmission, the information symbols are first grouped into blocks $\mathbf{s}_i(n) = [s_i(n,0), \cdots, s_i(n,M-1)]^T$, for $i = 1,2$, of size $M \times 1$. Then, symbol blocks $\mathbf{s}_1(n)$ are fed to CP-OFDM modulator, and $\mathbf{s}_2(n)$ are fed to ZP-OFDM modulator. So, the output of OFDM modulators are uncorrelated, and can be expressed as

$$\text{CP-OFDM}: \quad x_1(n,p) = \sum_{m=0}^{M-1} s_1(n,m) W_M^{-m(p-L_u)} \tag{6.2.1}$$

$$\text{ZP-OFDM}: \quad x_2(n,p) = (1 - \sum_{q=0}^{L_u-1} \delta(q-p)) \sum_{m=0}^{M-1} s_2(n,m) W_M^{-m(p-L_u)} \tag{6.2.2}$$

for $p = 0, \cdots, P-1$, and $P = L_u + M$, where $s_i(n,m)$ denotes the mth element of the nth block $\mathbf{s}_i(n)$.

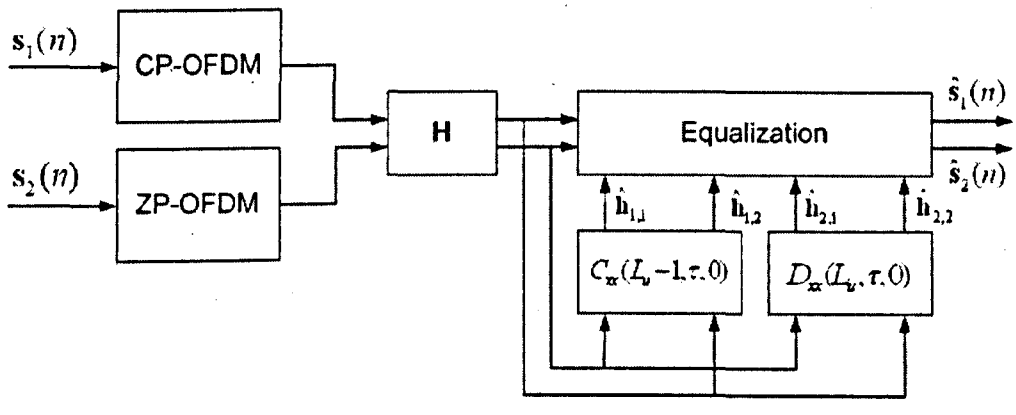

Figure 6.1: Transceiver for MIMO-OFDM with guard interval diversity

Due to the multipath propagation, the received signal at the jth receive antenna can be expressed as

$$y_j(n,p) = z_{1,j}(n,p) + z_{2,j}(n,p) + v_j(n,p) \tag{6.2.3}$$

where $v_j(n,p)$ is the additive white Gaussian noise with zero mean and variance σ_v^2, and uncorrelated for different n's and p's, and independent with the signal; $z_{1,j}(n,p)$ is the received CP-OFDM signal expressed by

$$\begin{aligned} z_{1,j}(n,p) &= \sum_{m=0}^{M-1} s_1(n,m) W_M^{-m(p-L_u)} \sum_{l=0}^{L_u-1} h_{1,j}(l) W_M^{ml} \\ &\quad \times \sum_{r=0}^{P-1} \delta(r-(p-l)) + \sum_{m=0}^{M-1} s_1(n-1,m) W_M^{-m(p-L_u)} \\ &\quad \times \sum_{l=0}^{L_u-1} h_{1,j}(l) W_M^{m(l-P)} \sum_{r=0}^{P-1} \delta(r-(p-l+P)) \end{aligned} \tag{6.2.4}$$

and $z_{2,j}(n,p)$ is the received ZP-OFDM signal given by

$$\begin{aligned} z_{2,j}(n,p) &= (1 - \sum_{q=0}^{L_u-1} \delta(q-p+l)) \sum_{m=0}^{M-1} s_2(n,m) W_M^{-m(p-L_u)} \\ &\quad \times \sum_{l=0}^{L_u-1} h_{2,j}(l) W_M^{ml} \sum_{r=0}^{P-1} \delta(r-(p-l)) \\ &\quad + (1 - \sum_{q=0}^{L_u-1} \delta(q-p+l-P)) \sum_{m=0}^{M-1} s_2(n-1,m) W_M^{m(p-L_u)} \\ &\quad \times \sum_{l=0}^{L_u-1} h_{2,j}(l) W_M^{m(l-P)} \sum_{r=0}^{P-1} \delta(r-(p-l+P)) \end{aligned} \tag{6.2.5}$$

The blind channel estimation algorithm is based on Eqn. (6.2.3). However, (6.2.3) brings difficulty for the description of channel equalization. To simplify the discussion on channel equalization, we next represent the received signal model in matrix form.

Prior to equalization, the guard intervals are first removed to eliminate ISI. The

remaining information for each block is given by

$$\bar{\mathbf{y}}_j(n) = [y_j(n, L_u), \ \cdots, y_j(n, P-1)]^T \tag{6.2.6}$$

According to the introduction in Chapter 1, $\bar{\mathbf{y}}_j(n)$ can be expressed as

$$\bar{\mathbf{y}}_j(n) = \sqrt{M}\boldsymbol{\Gamma}_{1,j}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}_1(n) + \sqrt{M}\boldsymbol{\Delta}_{2,j}\mathbf{F}_M^{\mathcal{H}}\mathbf{s}_2(n) + \bar{\mathbf{v}}_j(n) \tag{6.2.7}$$

where $\boldsymbol{\Gamma}_{1,j}$ denotes the circulant channel matrix of size $M \times M$; $\boldsymbol{\Delta}_{2,j}$ denotes the lower triangle channel matrix of size $M \times M$; $\bar{\mathbf{v}}_j(n)$ denotes the $M \times 1$ noise vector. With the knowledge of multichannel information, the original signal is possible to be recovered when at least two uncorrelated antennas are employed at the receiver. Let us consider the simplest case, i.e. $j = 1, 2$, and define a vector $\mathbf{y}(n) = [\bar{\mathbf{y}}_1^T(n), \ \bar{\mathbf{y}}_2^T(n)]^T$. The input signals to the demodulator can be expressed as

$$\mathbf{y}(n) = \sqrt{M}\begin{pmatrix} \mathbf{F}_M^{\mathcal{H}} & \\ & \mathbf{F}_M^{\mathcal{H}} \end{pmatrix}\begin{pmatrix} \boldsymbol{\Gamma}_{1,1} & \boldsymbol{\Delta}_{2,1} \\ \boldsymbol{\Gamma}_{1,2} & \boldsymbol{\Delta}_{2,2} \end{pmatrix}\mathbf{s}(n) + \mathbf{v}(n) \tag{6.2.8}$$

where, $\mathbf{s}(n) = [\mathbf{s}_1^T(n), \ \mathbf{s}_2^T(n)]^T$ and $\mathbf{v}(n) = [\bar{\mathbf{v}}_1^T(n), \ \bar{\mathbf{v}}_2^T(n)]^T$. If the multichannel matrix is invertible, then ZF method can be employed to recover the original signals

$$\hat{\mathbf{s}}(n) = \frac{1}{\sqrt{M}}\begin{pmatrix} \mathbf{F}_M & \\ & \mathbf{F}_M \end{pmatrix}\begin{pmatrix} \boldsymbol{\Gamma}_{1,1} & \boldsymbol{\Delta}_{2,1} \\ \boldsymbol{\Gamma}_{1,2} & \boldsymbol{\Delta}_{2,2} \end{pmatrix}^{\dagger}\mathbf{y}(n) \tag{6.2.9}$$

It can be observed that, the computation complexity for (6.2.9) is mainly arising from the matrix inverse, around $\mathcal{O}(4M^2 \log 2M)$. It means that MIMO-OFDM with guard interval diversity may increase the receiver complexity compared with conventional CP-OFDM and ZP-OFDM transmissions.

6.3 Blind channel estimation algorithm

The blind channel estimation algorithm uses cyclostationarity of the received signal (6.2.3). To clarify the analysis, we first investigate the autocorrelation for CP-OFDM

and ZP-OFDM.

According to the result in [57], the autocorrelation for transmitted CP-OFDM signals is given by

$$\begin{aligned} C_{x_1x_1}(nP+p\,;\tau) &= \mathcal{E}\{x_1(n,p)x_1^*(n,p+\tau)\} \\ &= \sigma_s^2 M\delta(\tau) + [\delta(\tau-M)\sum_{r=0}^{P-M-1}\delta(p-r) \\ &\quad +\delta(\tau+M)\sum_{r=M}^{P-1}\delta(p-r)] \end{aligned} \tag{6.3.1}$$

Using (6.2.2) and (6.3.1), the autocorrelation for transmitted ZP-OFDM signals can be derived as

$$C_{x_2x_2}(nP+p\,;\tau) = \sigma_s^2 M\delta(\tau)(1-\sum_{r=0}^{P-M-1}\delta(p-r)) \tag{6.3.2}$$

where σ_s^2 is the variance of information bearing symbols.

Based on (6.3.1) and (6.3.2), we next investigate the autocorrelation for the received signals. Due to $z_{1,j}(n,p)$, $z_{2,j}(n,p)$ and $v_j(n,p)$ are uncorrelated each other, the autocorrelation for (6.2.3) can be expressed as

$$\begin{aligned} C_{y_jy_j}(nP+p\,;\tau\,;q) &= C_{x_{1,j}x_{1,j}}(nP+p\,;\tau\,;q) + C_{x_{2,j}x_{2,j}}(nP+p\,;\tau\,;q) + \sigma_v^2\delta(\tau) \\ &= \sum_{l=0}^{L_u-1} h_{1,j}(l)h_{1,j}^*(l+\tau-q)C_{x_1x_1}(nP+p-l\,;q) \\ &\quad + \sum_{l=0}^{L_u-1} h_{2,j}(l)h_{2,j}^*(l+\tau-q)C_{x_2x_2}(nP+p-l\,;q) \\ &\quad +\sigma_v^2\delta(\tau) \end{aligned} \tag{6.3.3}$$

Let us define

$$A_i(l\,,\tau\,,q) = h_{i,j}(l)h_{i,j}^*(l+\tau-q) \tag{6.3.4}$$

and plug (6.3.1) and (6.3.2) into (6.3.3), we can obtain

$$\begin{aligned} C_{y_j y_j}(nP+p\,;\tau\,;q) &= M\sigma_s^2[\delta(q)\sum_{l=0}^{L_u-1}(A_1(l\,,\tau\,,q)+A_2(l\,,\tau\,,q)) \\ &\quad +\delta(q-M)\sum_{l=0}^{L_u-1}A_1(l\,,\tau\,,q)\sum_{r=0}^{P-M-1}\delta(p-r-l) \\ &\quad -\delta(q)\sum_{l=0}^{L_u-1}A_2(l\,,\tau\,,q)\sum_{r=0}^{P-M-1}\delta(p-r-l) \\ &\quad +\delta(q+M)\sum_{l=0}^{L_u-1}A_1(l\,,\tau\,,q)\sum_{r=M}^{P-1}\delta(p-r-l)] \\ &\quad +\sigma_v^2\delta(\tau) \end{aligned} \tag{6.3.5}$$

Now, we let $q=M$ and simplify (6.3.5) as

$$\begin{aligned} C_{y_j y_j}(nP+p\,;\tau\,;M) &= M\sigma_s^2\sum_{l=0}^{L_u-1}A_1(l\,,\tau\,,M) \\ &\quad \times\sum_{r=0}^{P-M-1}\delta(p-r-l)+\sigma_v^2\delta(\tau) \end{aligned} \tag{6.3.6}$$

If $q=0$, (6.3.5) can be rewritten into

$$\begin{aligned} C_{y_j y_j}(nP+p\,;\tau\,;0) &= M\sigma_s^2[\sum_{l=0}^{L_u-1}(A_1(l\,,\tau\,,0)+A_2(l\,,\tau\,,0)) \\ &\quad -\sum_{l=0}^{L_u-1}A_2(l\,,\tau\,,0)\sum_{r=0}^{P-M-1}\delta(p-r-l)] \\ &\quad +\sigma_v^2\delta(\tau) \end{aligned} \tag{6.3.7}$$

It can be observed that (6.3.6) only contains one subchannel information, which is possible to be blindly estimated. Assuming that the first subchannel information is obtained, it is possible to blindly identify another subchannel information from (6.3.7). Based on (6.3.6) and (6.3.7), the following results will show the multichannel identifiability.

Theorem 6.3.1. *(Sufficient condition) The channel impulse response* $\mathbf{h}_{1,j}$ *can be uniquely identified from (6.3.6) with a phase ambiguity when conditions 1)* $p = 0$, *and 2)* $M \leqslant \tau \leqslant M + L_u - 1$ *are satisfied.*

Proof. See Appendix 6.A.

Theorem 6.3.1 shows that the first subchannel can be uniquely identified with a constant phase ambiguity. This ambiguity can be resolved using several pilot symbols in the frequency domain. However, (6.5.9) implies that the channel estimate performance is related to the power of $h(0)$. The best performance will be achieved only when the strongest path is synchronized as the first path. Next, we introduce how to identify the other subchannel from (6.3.7).

Theorem 6.3.2. *(Sufficient condition) The channel impulse response* $\mathbf{h}_{2,j}$ *can be uniquely identified from (6.3.7) when the conditions 1)* $p = L_u$, *and 2)* $0 \leqslant \tau \leqslant L_u - 1$ *are satisfied.*

Proof. See Appendix 6.B.

By now, we have shown that the multichannel information is uniquely identifiable by employing guard interval diversity. Unlike the technique using nonconstant precoding, the proposed scheme here does not impose anything on OFDM transmissions, and therefore will not affect the overall system performance. Next, we use computer simulations to examine the channel estimate performance.

6.4 Simulation Results

To make a fair comparison with the nonconstant precoding approach, we set the system parameters as $M = 12$, $L_u = 4$, and $P = M + L_u = 16$, which are the same as the settings in [61]. Each subcarrier is QPSK modulated. The sampled channel impulse responses used for our tests are also the same as that used in [61].

The normalized MSE (NMSE) and the channel average bias (CAB) are both used to benchmark the channel estimation performance, which have been defined in [61] as the following

$$\text{NMSE} = \frac{1}{2L_u I}\sum_{i=1}^{I}\sum_{\mu=1}^{2}\frac{\|\hat{\mathbf{h}}_{\mu,i} - \mathbf{h}_\mu\|^2}{\|\mathbf{h}_\mu\|^2}$$

and the channel average bias (CAB) defined as

$$\text{CAB} = \frac{1}{2IL_u}\sum_{\mu=1}^{2}\sum_{l=0}^{L_u-1}\left|\sum_{i=1}^{I}(\hat{h}_{\mu,i}(l) - h_{\mu,i}(l))\right|$$

where I denotes the number of Monte Carlo trials.

`Test Case 1:` In the first test, we examine CAB and MSE of the proposed blind channel estimator as a function of SNR for the fixed DRLs ($N = 80,\ 500$). Figure 6.2 shows the following results:

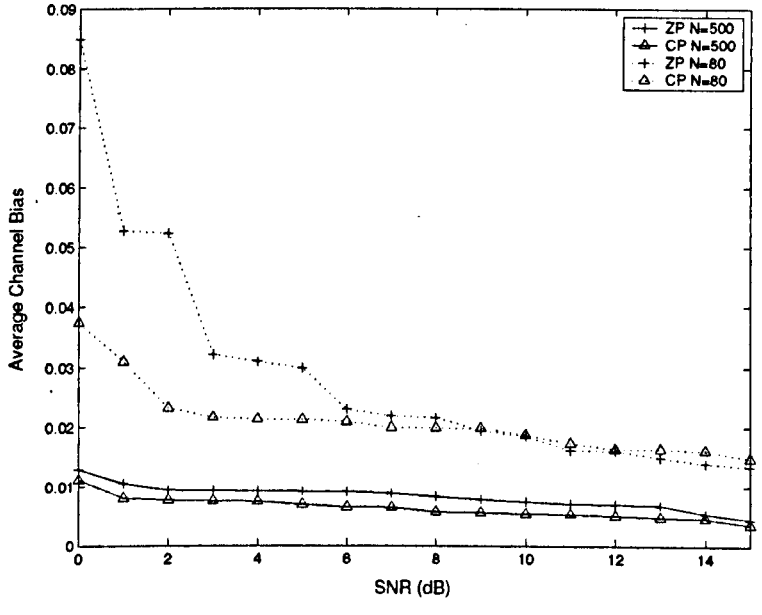

Figure 6.2: The average channel bias as a function of SNR

1) When the DRL is not long enough (e.g. $N = 80$), the noise effects cannot be mitigated significantly. So, the CAB performance is not good in low SNR (e.g. < 4 dB). With increasing SNR, the performance are generally improved.

2) When the DRL is much longer (e.g. $N = 500$), the CAB performance is significantly improved. Since the noise effects are well mitigated, the CAB is not very sensitive to SNR.

3) The channel estimate performance is also affected by transmission schemes. We can see that the channel estimate performance for the CP-OFDM transmission is better than that for the ZP-OFDM transmission in the proposed MIMO system, because zero sequences in the ZP-OFDM transmission brings no inter-antenna interference to the CP-OFDM transmission.

The MSE performance illustrated in Figure 6.3 also reflects the above features. In addition, we make a comparison between the proposed approach and the nonconstant precoding approach for $N = 500$. It is shown that the proposed approach outperforms the nonconstant precoding one within the SNR range $0 - 15$ dB. Since these two approaches are not sensitive to SNR, we do not plot more result for higher SNR.

`Test Case 2:` In this test, we benchmark the CAB and MSE as a function of DRL for the fixed SNR ($= 11,\ 15$ dB). Figure 6.4 shows that

1) The CAB performance generally improves with increasing the DRL. However, the improvement is not significantly when the DRL is longer than 200.

2) All curves are very close when the DRL is longer than 50.

The MSE performance illustrated in Figure 6.5 confirms the above results again and also shows that

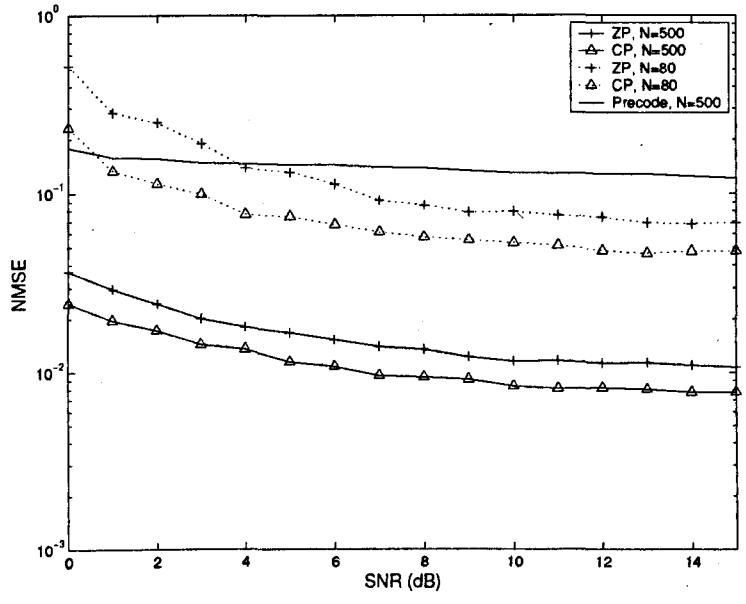

Figure 6.3: The mean square error as a function of SNR

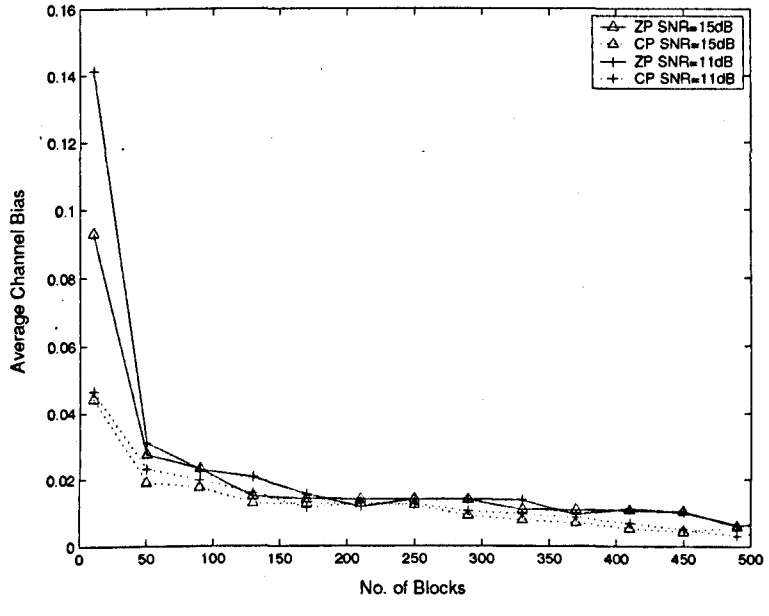

Figure 6.4: The average channel bias as a function of data record length

3) The MSE results degrade fast with increasing the DRL, for $N < 120$. After this, the MSE performances become very good and degrade slowly. It can be observed the proposed approach has much faster convergence than that of the nonconstant precoding approach.

4) For the DRL within the range $(50 - 500)$, the proposed approach demonstrates much better channel estimate performance than the nonconstant precoding approach when the SNR is 11 dB. Taking MSE= 0.2 as an example, the required DRL for the nonconstant precoding approach is around 400, but only 50 is needed for the proposed estimator.

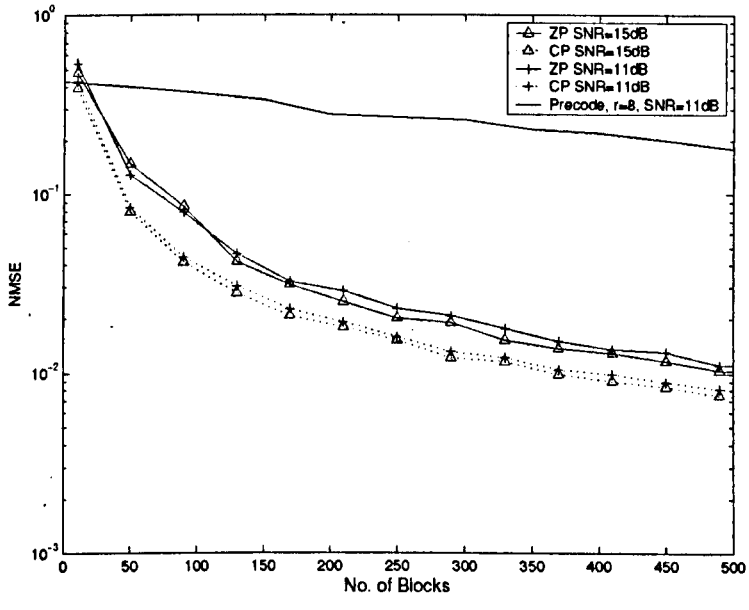

Figure 6.5: The mean square error as a function of data record length

6.5 Conclusion

This chapter has introduced a novel blind channel estimation method for MIMO-OFDM by exploiting guard interval diversity. This blind channel estimator exploited the second-order cyclostationary to yield unique channel estimates. Relying on the guard interval diversity, subchannels could be dealt with individually. The channel estimate performance has been examined for different SNR and DRL cases. Simulation results has shown that the proposed approach outperformed the nonconstant precoding approach in CAB and MSE performances. Based on the proposed MIMO-OFDM system, future research may focus on improving the blind channel estimator by exploiting Cholesky decomposition.

Appendix 6.A: Proof of Theorem 6.3.1

To assure that (6.3.6) contain enough information for channel identification, the following condition should be satisfied

$$\sum_{r=0}^{P-M-1} \delta(p-r-l) = 1 \tag{6.5.1}$$

Otherwise, (6.3.6) will only contains the noise part. The condition to fulfil (6.5.1) is given by

$$0 \leqslant l = p - r \leqslant L_u - 1 \tag{6.5.2}$$

For $r \in [0, L_u - 1]$, the range of p can be broadened as

$$0 \leqslant p \leqslant 4L_u - 2 \tag{6.5.3}$$

Based on this result, (6.3.6) can be rewritten into

$$C_{y_j y_j}(nP + p\,;\tau\,;M) = M\sigma_s^2 \sum_{l=0, l=p-r}^{L_u-1} h_{1,j}(l) h_{1,j}^*(l+\tau-M) + \sigma_v^2 \delta(\tau) \tag{6.5.4}$$

Now, we let $p = 0$. (6.5.1) shows that (6.5.4) will not be zero only when $l = 0$ and $r = 0$. Substituting these results into (6.5.4), we have

$$C_{y_j y_j}(nP\ ;\tau\ ;M) = M\sigma_s^2 h_{1,j}(0)h_{1,j}^*(\tau - M) + \sigma_v^2\delta(\tau) \tag{6.5.5}$$

In order to guarantee the channel identifiability, (6.5.1) shows that the following condition should be satisfied

$$0 \leqslant \tau - M \leqslant L_u - 1 \tag{6.5.6}$$

which results in

$$M \leqslant \tau \leqslant M + L_u - 1 \tag{6.5.7}$$

Based on (6.5.7), we can see that $\sigma_v^2\delta(\tau) = 0$, and the first channel tap $h(0)$ can be obtained by setting $\tau = M$,

$$\hat{h}_{1,j}(0) = \sqrt{\frac{C_{y_j y_j}(nP\ ;M\ ;M)}{M\sigma_s^2}}e^{j\phi} \tag{6.5.8}$$

where ϕ is the remained phase rotation. With the knowledge of $\hat{h}_{1,j}(0)$, the other taps can be uniquely identified by setting $M + 1 \leqslant \tau \leqslant M + L_u - 1$,

$$\hat{h}_{1,j}(\tau - M) = [\frac{C_{y_j y_j}(nP\ ;\tau\ ;M)}{M\sigma_s^2\hat{h}_{1,j}(0)}]^* \tag{6.5.9}$$

This theorem is therefore proved.

Appendix 6.B: Proof of Theorem 6.3.2

Eqn. (6.3.7) contains two items, $\sum_{l=0}^{L_u-1}(A_1(l\ ,\tau\ ,0) + A_2(l\ ,\tau\ ,0))$ and $\sigma_v^2\delta(\tau)$, which are independent of the index p. Hence, these two items can be removed from (6.3.7)

using the differential operation with respect to p, i.e.,

$$\begin{aligned}
\mathcal{D}_{y_jy_j}(nP+p\,;\tau\,;0) &= \mathcal{C}_{y_jy_j}(nP+p\,;\tau\,;0)-\mathcal{C}_{y_jy_j}(nP+p-1\,;\tau\,;0)\\
&= M\sigma_s^2[\sum_{l=0}^{L_u-1}A_2(l\,,\tau\,,0)\sum_{r=0}^{P-M-1}\delta(p-r-l-1)\\
&\quad -\sum_{l=0}^{L_u-1}A_2(l\,,\tau\,,0)\sum_{r=0}^{P-M-1}\delta(p-r-l)]\\
&= M\sigma_s^2\sum_{l=0}^{L_u-1}A_2(l\,,\tau\,,0)(\sum_{r=0}^{P-M-1}\delta(p-r-l-1)\\
&\quad -\sum_{r=0}^{P-M-1}\delta(p-r-l)) \qquad (6.5.10)
\end{aligned}$$

To assure $\mathcal{D}_{y_jy_j}(nP+p\,;\tau\,;0)\neq 0$, the following condition should be satisfied,

$$\sum_{r=0}^{P-M-1}\delta(p-r-l-1)-\sum_{r=0}^{P-M-1}\delta(p-r-l)\neq 0 \qquad (6.5.11)$$

which leads to two cases

$$\text{Case 1:}\begin{cases}\sum_{r=0}^{P-M-1}\delta(p-r-l-1)=0\\ \sum_{r=0}^{P-M-1}\delta(p-r-l)\neq 0\end{cases} \qquad (6.5.12)$$

$$\text{Case 2:}\begin{cases}\sum_{r=0}^{P-M-1}\delta(p-r-l-1)\neq 0\\ \sum_{r=0}^{P-M-1}\delta(p-r-l)=0\end{cases} \qquad (6.5.13)$$

Let $p=L_u=P-M$, Case 1 occurs only when

$$\begin{cases}0\leqslant r=L_u-l-1\leqslant L_u-1\\ L_u-l>L_u-1\,,\text{or } L_u-l<0\end{cases} \qquad (6.5.14)$$

The solution to (6.5.14) is $l=0$. Similarly, Case 2 occurs only when

$$\begin{cases}1\leqslant l\leqslant L_u-1\\ l<0\,,\text{or } l>L_u-1\end{cases} \qquad (6.5.15)$$

which results in $l = L_u$. For $l \in [0, L_u - 1]$, Case 2 will not occur. Next, we plug $p = L_u$ and $l = 0$ into (6.5.10) and obtain

$$\mathcal{D}_{y_j y_j}(nP + L_u \,;\tau \,;0) = M\sigma_s^2 h_{2,j}(0) h_{2,j}^*(\tau) \tag{6.5.16}$$

Hence, for $0 \leqslant \tau \leqslant L_u - 1$, the channel impulse response can be estimated using

$$h_{2,j}(0) = \sqrt{\frac{\mathcal{D}_{y_j y_j}(nP + L_u \,;0 \,;0)}{M\sigma_s^2}} e^{j\phi} \tag{6.5.17}$$

$$h_{2,j}(\tau) = [\frac{\mathcal{D}_{y_j y_j}(nP + L_u \,;\tau \,;0)}{M\sigma_s^2 h_{2,j}(0)}]^* \tag{6.5.18}$$

Theorem 2 is therefore proved.

Chapter 7

Performance on pilot-assisted channel estimation techniques

7.1 Introduction

Pilot-assisted channel estimation techniques are usually used in rapidly time-varying scenario, such as vehicle communications. Several OFDM-based standards have already applied the pilot-assisted technique to track rapidly time-varying channels (see [5] and [10]). By now, a number of pilot-assisted channel estimation algorithms have been reported, for instance LS-based ideal interpolation [40], transform domain interpolation [47] and model-based interpolation [46], [80]. Among all existing approaches, the LS-based ideal interpolation always holds the best channel estimate performance when the available pilot subcarriers are enough. Once the available pilot subcarriers are not enough, model-based approaches offer non-ideal interpolation algorithms. However, the interpolation error cannot be avoided and may decreases the overall system performance.

In this chapter, we intend to analyze performance of some classical pilot-assisted channel estimation approaches and their applications. To achieve the tradeoff between performance and bandwidth efficiency, a combined channel estimation method

is proposed by employing the decision-directed (DD) algorithm. A pilot allocation scheme is also addressed for improving the channel estimate performance. Part of this chapter has been presented in [85].

7.2 Pilot-assisted channel estimation method

7.2.1 Ideal interpolation: LS based method

The LS-based method has been partly introduced in Chapter 2.3. To clarify our analysis, we rewrite (2.3.4) here

$$\begin{aligned} \tilde{\mathbf{z}}_{\mathrm{p}}(n) &= \mathbf{D}(\mathbf{s}_p(n))\bar{\mathbf{F}}_{\mathrm{p}}\mathbf{h} + \tilde{\mathbf{v}}_{\mathrm{p}}(n) \\ &= \mathbf{Q}_{\mathrm{p}}\mathbf{h} + \tilde{\mathbf{v}}_{\mathrm{p}}(n) \end{aligned} \tag{7.2.1}$$

and it is possible to identify the CSI using

$$\hat{\mathbf{h}} = \mathbf{Q}_{\mathrm{p}}^{\dagger}\tilde{\mathbf{z}}_{\mathrm{p}}(n) \tag{7.2.2}$$

For the CIR, $\mathbf{h}$, with length L_u, the channel identifiability is guaranteed only when the matrix $\mathbf{Q}_{\mathrm{p}}$ has full column rank. It means that the size of $\mathbf{s}_p(n)$, which is denoted by M_p, should not be smaller than the channel length L_u. This is the major limitation for the LS-based algorithm. In this case, the maximum bandwidth efficiency for CP-OFDM is given by

$$\eta_{\max} = \frac{M - L_u}{M + L_u} \tag{7.2.3}$$

where, M denotes the total number of subcarriers. Usually, the ratio between channel length and OFDM symbol duration is around 0.2 (see [10]). Using this coefficient, the maximum bandwidth efficiency, $\eta_{\max}$, should be 0.67.

The channel estimate performance can be evaluated using the mean-square error, which can be derived from (7.2.2)

$$\begin{aligned} \epsilon &= \|\hat{\mathbf{h}} - \mathbf{h}\|^2 \\ &= \mathrm{Tr}\{\mathbf{Q}_\mathrm{p}^\dagger \mathrm{E}\{\tilde{\mathbf{v}}_\mathrm{p}(n)\tilde{\mathbf{v}}_\mathrm{p}(n)^\mathcal{H}\}(\mathbf{Q}_\mathrm{p}^\dagger)^\mathcal{H}\} \end{aligned} \tag{7.2.4}$$

(7.2.4) shows that the channel estimate error is caused only by the additive noise. If the additive noise is white, then (7.2.4) becomes

$$\epsilon = \sigma_v^2 \mathrm{Tr}\{(\mathbf{Q}_\mathrm{p}^\mathcal{H}\mathbf{Q}_\mathrm{p})^\dagger\} \tag{7.2.5}$$

[40] has shown that, if the pilots are uniformly distributed over the whole bandwidth, the MSE result can be minimized as

$$\epsilon_{\min} = \frac{L_u \sigma_v^2}{M_p \sigma_s^2} \tag{7.2.6}$$

We can see that the channel estimate performance is related to the number of pilots.

Since the transform-domain approach (see [47]) is actually a kind of LS algorithm, the above results are also suitable for it. Next, we introduce the model-based channel estimation method.

7.2.2 Polynomial interpolation: model-based channel estimation

Model-based channel estimation is actually a kind of curve fitting approach. The main idea of this approach is using polynomial interpolation to recover the channel transfer function. Mathematically, the time-frequency polynomial model is a 2-dimensional Taylor expansion of the channel transfer function, which can be expressed as

$$H_{m,n} = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} c_{m,n} \cdot n^i m^j + G_{I,J} \tag{7.2.7}$$

where $c_{m,n}$ is the model coefficient, I and J are the polynomial order for the frequency and time domain respectively, m and n are the subcarrier and block indices respectively, and $G_{I,J}$ is the modelling error. To simplify our discussion, we just consider 1-dimensional regression model in this document, i.e.

$$H_m = \sum_{i=0}^{I-1} c_m \cdot m^i + G_I \tag{7.2.8}$$

where $G_I = \frac{m^I}{I!}\frac{\partial^I H(f)}{\partial f^I}$. If the polynomial order, I, is determined, (7.2.8) can be presented as the following matrix form

$$\tilde{\mathbf{h}} = \mathbf{V}_M \mathbf{c} + \mathbf{g} \tag{7.2.9}$$

where $\tilde{\mathbf{h}}$ denotes the channel frequency-response vector of size $M \times 1$; the $(I+1) \times 1$ vector $\mathbf{c}$ contains the polynomial coefficients; $\mathbf{g}$ is the interpolation error vector of size $M \times 1$; $\mathbf{V}_M$ is the Vandermonde matrix given by

$$\mathbf{V}_M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{I-1} \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \cdots & \cdots & M^{I-1} \end{pmatrix} \tag{7.2.10}$$

Hence, if $\mathbf{c}$ is determined, then the channel frequency-response can be estimated using $\hat{\tilde{\mathbf{h}}} = \mathbf{V}_M \mathbf{c}$ with the estimate error $\mathbf{g}$.

Let p be the pilot spacing, the channel frequency-response on pilot subcarriers can be expressed as

$$\tilde{\mathbf{h}}_p = \mathbf{V}_{M_p} \mathbf{c} + \mathbf{g}_p \tag{7.2.11}$$

where

$$\tilde{\mathbf{h}}_p = [H(p_0), H(p_0 + p), \cdots, H(p_0 + (M_p - 1)p)]^T \tag{7.2.12}$$

$$\mathbf{V}_{M_p} = \begin{pmatrix} 1 & p_0 & \cdots & p_0^{I-1} \\ 1 & p_0+p & \cdots & (p_0+p)^{I-1} \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \cdots & \cdots & (p_0+(M_p-1)p)^{I-1} \end{pmatrix} \tag{7.2.13}$$

and $\mathbf{g}_p$ denotes error occurring on pilot subcarriers. To improve the channel estimate performance, $\mathbf{g}_p$ is expected to be zero. This case occurs only when

$$\tilde{\mathbf{h}}_p = \mathbf{V}_{M_p}\mathbf{c} \tag{7.2.14}$$

(7.2.14) leads to the following three results:

r1) If $M_p < I+1$, then the polynomial coefficients, $\mathbf{c}$, cannot be determined, because $\mathbf{V}_{M_p}$ losses the full column rank.

r2) If $M_p = I+1$, then $\mathbf{c}$ can be determined using

$$\mathbf{c} = \mathbf{V}_{M_p}^{\dagger}\tilde{\mathbf{h}}_p \tag{7.2.15}$$

and $\mathbf{g}_p = \mathbf{V}_{M_p}\mathbf{c} - \tilde{\mathbf{h}}_p = \mathbf{0}$.

r3) If $M_p > I+1$, then $\mathbf{c}$ can be determined using (7.2.15). Since $\mathbf{V}_{M_p}$ is no longer a square matrix, i.e. $\mathbf{V}_{M_p}\mathbf{V}_{M_p}^{\dagger} \neq \mathbf{I}_{M_p}$, errors on pilot subcarriers are given by

$$\begin{aligned} \mathbf{g}_p &= \mathbf{V}_{M_p}\mathbf{c} - \tilde{\mathbf{h}}_p \\ &= (\mathbf{V}_{M_p}\mathbf{V}_{M_p}^{\dagger} - \mathbf{I}_{M_p})\tilde{\mathbf{h}}_p \end{aligned} \tag{7.2.16}$$

So, to mitigate the modelling error, we can divide M_p pilots into several groups, and each group contains $I+1$ pilots. The vector $\mathbf{c}$ is then determined for different groups, and the channel frequency-response is identified group by group. The channel estimate error arising from additive noise can be found in [46].

7.3 Combined channel estimators

Chapter 7.2 has introduced that the LS-based channel estimator outperforms the model-based estimator when the number of pilots is not smaller than the channel length. The channel estimate performance can be increased by increasing the number of available pilots. However, increasing available pilots may reduce the bandwidth efficiency. If the available pilots are not enough ($M_p < L_u$), then the LS-based estimator does not work. To solve these problems, we propose combined channel estimators by employing the decision-directed (DD) algorithm.

7.3.1 Double-LS estimator

As illustrated in Figure 7.1 (a), the double-LS estimator employs two LS channel estimators to realize the DD algorithm with one iteration. It works only when the available pilots are large enough ($M_p \geqslant L_u$). The objective of this combined channel estimator (CCE) is to improve the channel estimate performance when M_p is specified.

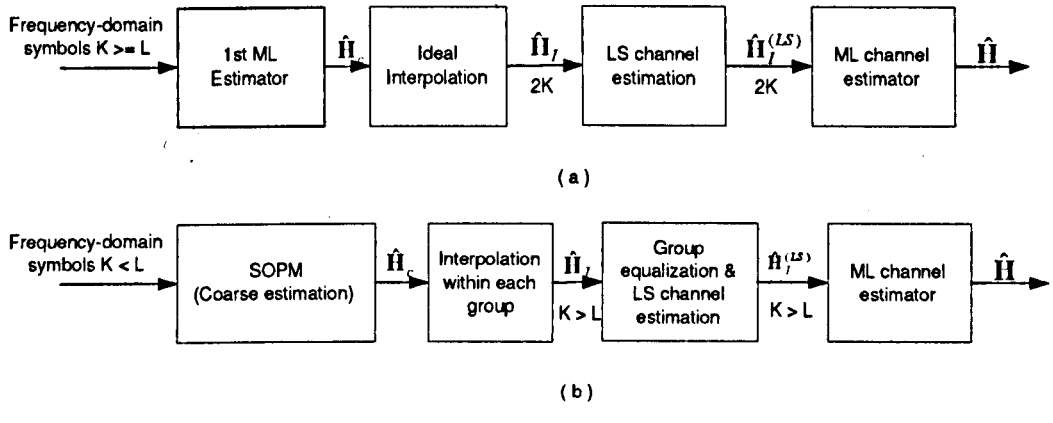

Figure 7.1: Block diagram of combined channel estimators (a) double LS estimator, (b) Polynomial+ML

For the double-LS estimator, the first LS estimator is used to find the initial channel estimate, which is used to recover the information-bearing symbols. Once the information-bearing symbols are determined, the second LS estimator updates the channel estimate by regarding the recovered symbols as pilot symbols. To improve the channel estimate performance, the second LS estimator does not use the subcarriers on deep fading, where the error probability is much higher than other subcarriers.

7.3.2 Polynomial+LS estimator

When the number of pilots $M_p < L_u$, most of pilot-assisted channel estimators do not work, for instance LS and transform domain interpolation approaches. Fortunately, the polynomial interpolation algorithm still works if the polynomial order $I < M_p$, even though its performance is not as good as that for ideal interpolation algorithm. Thus, for the case of $I < M_p < L_u$, we use polynomial+LS estimator to improve the channel estimate performance.

As illustrated in Figure 7.1 (b), the polynomial estimator is used to find the initial channel estimate, which is used to recover the symbols. Next, the LS estimator updates the channel estimate using the DD algorithm. [46] has shown that the channel estimate performance for the polynomial interpolation is affected by the pilot spacing. If the pilot spacing is too large, than the equalization performance may be considerably decreased. This may further affect the final channel estimate performance. Thus, we propose the following pilots arrangement for the CCE approach:

1. The pilots are divided into N_g groups, and each group contains N_p pilots with $N_p > I$.

2. The N_g groups are uniformly placed in the frequency band.

3. The N_p pilots within each group are uniformly allocated with the pilot-spacing p_g.

Based on this pilot arrangement scheme, the CCE estimator carries out the channel estimation as the following steps:

s1) The polynomial estimator first finds the channel frequency-response within each group. The channel information between two consecutive groups will not be obtained.

s2) Recover the symbols within each group using the partial channel estimate.

s3) The LS estimator utilizes the recovered symbols as the new pilots for the full channel estimation.

From the above discussion, we can see that the full channel information is identifiable only when the number of available pilots fed to the LS estimator is larger than the channel order. By employing the polynomial estimator, the total available pilots is increased from M_p to $M_p + p_g N_g$. Thus, if $M_p + p_g N_g \geqslant L_u$, the full channel information can be uniquely identified. The channel estimate performance is affected by the parameters, p_g, N_g and N_p.

7.4 Simulations

To investigate the pilot-assisted channel estimators, simulations examine the BER performance for OFDM systems with the following setup: $M = 512$ subcarriers, $L_u = 50$, information-bearing symbols are drawn from 16QAM constellations with equal probability. The multipath channel is modelled as a FIR filter. Each taps are generated according to the ETSI standard for digital cellular radios [79].

`Test Case 1:` In this experiment, we make the performance comparison between LS and model-based methods. The model-based estimator employs a second-order polynomial for the interpolation, i.e. $I = 2$. The normalized maximum Doppler shift, ε, is set to 4.16×10^{-3} and 10×10^{-3} respectively. The pilot spacing is set to $p = 5$. For a wide SNR range (0-50dB), Figure 7.2 shows the following results:

1) When the pilot spacing is fixed, the LS estimator outperforms the model-based estimator throughout the wide SNR range.

2) Within the low and middle SNR range (SNR< 40 dB), if the model-based estimator collects relatively more pilots, e.g. 9 pilots, for the channel estimation, then the overall system performance becomes much better. The reason here is that collecting more pilots can reduce the channel estimate error caused by additive Gaussian noise and ICI [46]. In the high SNR range (SNR> 40 dB)

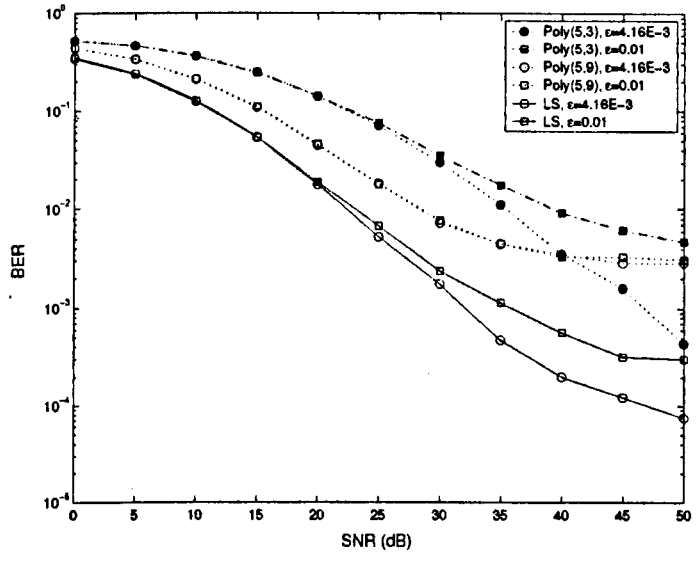

Figure 7.2: BER Vs SNR for the channel estimation in GSM channels (urban type).

with relatively small Doppler shift, the modelling error becomes the major interference, such that collecting less pilots, e.g. 3 pilots, may offer better BER performance.

3) The ICI power affects the overall system performance particularly in the high SNR range.

`Test Case 2:` This experiment evaluates the performance of CCE estimators when the number of pilots is lager than the channel length. In simulations, the number of pilots is set to 50, which is the same as the channel length. The normalized Doppler shift is set to 2.13×10^{-3}. Figure 7.3 shows that, the double LS approach demonstrates the best performance among all approaches. It improves around 2 dB performance in comparison with the LS estimator for the middle SNR range (= 10 $\sim$ 12 dB). The model-based approach has the worst performance because of its interpolation

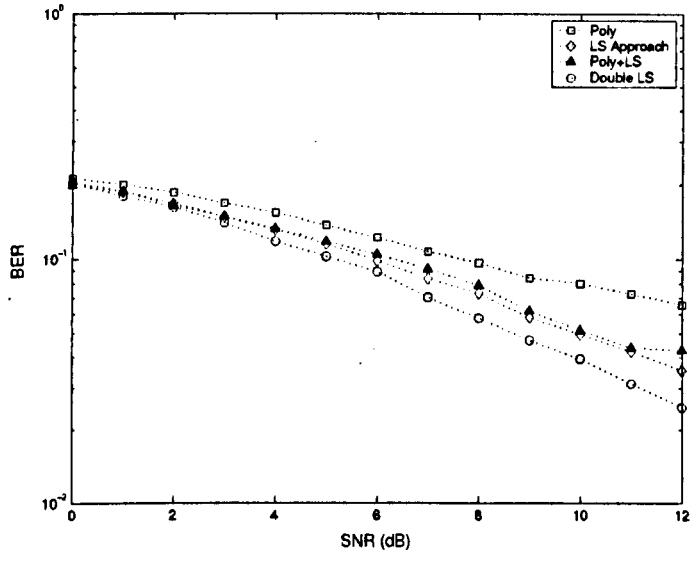

Figure 7.3: BER Vs SNR for CCE estimators in GSM channels (urban type), $\epsilon = 2.13 \times 10^{-3}$.

error. Its performance can be improved by employing a LS estimator as the second estimator in the CCE approach.

`Test Case 3:` When the number of pilots is smaller than the channel length, only model based and polynomial+LS estimators can be used for the channel estimation. To investigate the performance, the number of pilots is set to 48. For the model-based approach, the pilots are uniformly allocated. For the CCE approach, the pilots are divided into $N_g = 16$ groups, and the pilot spacing is given by $p_g = 5, 10, 15, 20, 25$ respectively. When the normalized Doppler shift is 4.16×10^{-3}, Figure 7.4 shows that the CCE approach may improve the BER performance for the SNR range (> 20 dB) when the pilot spacing is around 15. When the normalized Doppler shift is decreased to 2.13×10^{-3}, Figure 7.5 shows that the CCE approach may improve the BER performance when the pilot spacing $p \in [10, 20]$. It means that the CCE estimator is sensitive to ICI.

7.5 Conclusion

This chapter has investigated several pilot-assisted channel estimation approaches for OFDM systems in rapidly time-varying channels. It has been shown that the ideal interpolation algorithms were applicable only when the number of pilots was larger than the channel length. Otherwise, the polynomial interpolation approach could be employed for the channel estimation. To improve the channel estimate performance, two combine channel estimators have been proposed by employing the decision-directed algorithm. Simulation results have been provided to demonstrate the comparable performance of the CCE estimators.

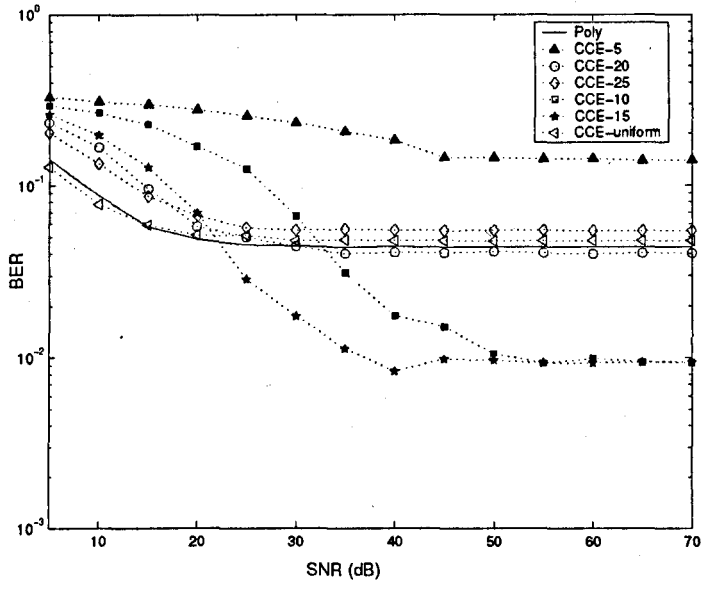

Figure 7.4: BER Vs SNR for CCE estimators in GSM channels (urban type), $\varepsilon = 4.16 \times 10^{-3}$.

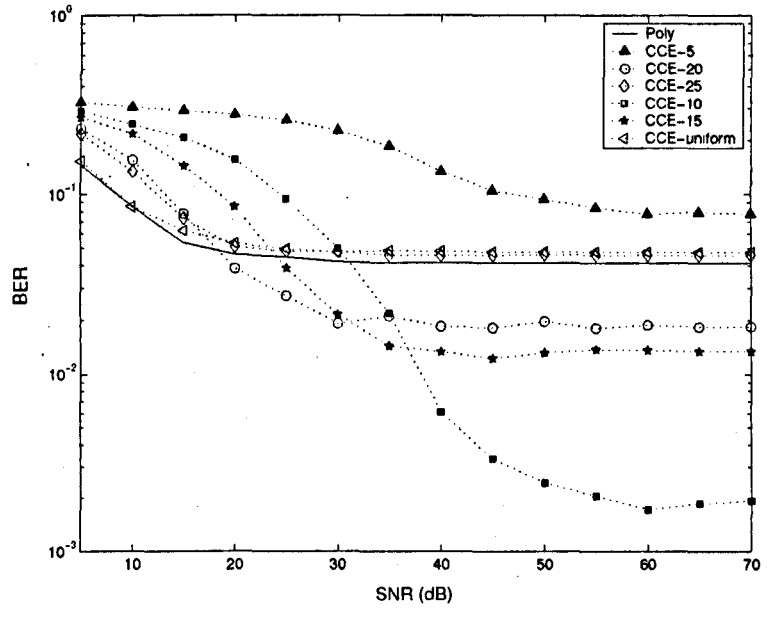

Figure 7.5: BER Vs SNR for CCE estimators in GSM channels (urban type), $\varepsilon = 2.13 \times 10^{-3}$.

Chapter 8

Blind estimation of frequency offset in unknown multipath fading channels

This chapter introduces a novel blind estimation method of carrier frequency offset (CFO) for OFDM in unknown multipath fading channels. This method is based on the second-order cyclostationary statistics induced by the pulse shaping filter and subcarrier weighting function. A special weighting function is employed to make the estimation of frequency offset efficient. This chapter is partly presented in our paper [82].

8.1 Signal Model

The baseband equivalent model of pulse-shaping OFDM signals can be expressed as (see [32]):

$$x(n) = \sum_{k=0}^{N-1} \sum_{i=-\infty}^{\infty} c_{k,l} w(k) g(n - iM) e^{j(2\pi/N)k(n-iM)} \tag{8.1.1}$$

where N is the number of subcarriers; M is the OFDM symbol length; $g(n)$ is the pulse-shaping filter; $w(k)$ is the subcarrier weighting function, and $c_{k,l}$ is the

information-bearing symbols. For $M > N$, the OFDM system is said to employ a time-frequency guard region. The guard region is usually introduced by cyclic prefix in the time domain and pilot subcarriers in the frequency domain.

The symbol timing offset (STO) is modelled as a time shift, and the unknown CFO is represented as a frequency shift. The received complex envelope of the OFDM signals with STO and CFO is given by

$$r(n) = e^{j(2\pi\theta_e n+\phi)} \int_0^1 X(f) e^{j2\pi v(n-n_e)} df + v(n) \tag{8.1.2}$$

where n_e denotes the STO; $\theta_e \in [-0.5,\ 0.5]$ denotes the normalized CFO; ϕ denotes the initial phase; $v(n)$ denotes the wide sense stationary noise; $X(f)$ denotes the frequency domain representation of the OFDM signals. In the presence of multipath fading, the received signal become much complicated

$$r(n) = e^{j(2\pi\theta_e n+\phi)} \sum_{l=0}^{L_h-1} \int_0^1 X(f) e^{j2\pi v(n-n_e-l)} df \cdot h(l) + v(n) \tag{8.1.3}$$

8.2 Blind estimation algorithm

To find the CFO, we consider the autocorrelation function of (8.1.3), which can be expressed as

$$\begin{aligned} R_r(n,\tau) &= \sigma_c^2 e^{j2\pi\theta_e\tau} W_N(\tau) \sum_{l=0}^{L_h-1} \sum_{m=0}^{L_h-1} h_l h_m^* \\ &\quad \times \sum_{i=-\infty}^{\infty} g(n-n_e-\tau_l-iM) \\ &\quad \times g^*(n-n_e-\tau_l-\tau-iM) + R_v(n,\tau) \end{aligned} \tag{8.2.1}$$

where σ_c^2 is the variance of symbols; τ is the autocorrelation delay; τ_l and τ_m are the multipath delay; $W_N(\tau) = \sum_{k=0}^{N-1} |w(k)|^2 e^{j2\pi/N k\tau}$. If τ is fixed, (8.2.1) shows

that the function $R_r(n,\tau)$ is periodic for the index n with the period M. The ℓ cyclic spectrum of $r(n,\tau)$ is defined as the kth slice of the two dimensional Fourier transform of $r(n,\tau)$, i.e.

$$\mathcal{Q}_r(\ell,\nu) = \sum_{n=0}^{M-1}\sum_{\tau} R_\tau(n,\tau)e^{-j(2\pi/M)\ell n_e}e^{-j\nu\tau} \tag{8.2.2}$$

It is tedious to prove but straightforward to show that

$$\begin{aligned}\mathcal{Q}_r(\ell,\nu) &= \frac{2\pi\sigma_c^2}{M}\delta(\nu-\theta_e)\star[e^{-j(2\pi/M)\ell n_e}\tilde{W}_N(\nu) \\ &\quad\times\sum_{l=0}^{L_h-1}\sum_{m=0}^{L_h-1}h_l h_m^* e^{j\nu(t_m-tau_l)}G(\nu+\ell\frac{2\pi}{M}) \\ &\quad\times G^*(\nu)] + R_v\delta(\ell)\end{aligned} \tag{8.2.3}$$

where $\tilde{W}_N(\nu) = \sum_\tau W_N(\tau)e^{-j\nu\tau}$. (8.2.3) shows that CFO only introduces a shift along the ν direction. If the weighting function is defined as

$$w(k) = \begin{cases}\lambda\ (\lambda >> 1), & \text{when } k=\xi \\ 1 \qquad\quad , & \text{otherwise}\end{cases} \tag{8.2.4}$$

where λ and ξ are constant, then we can obtain that

$$\tilde{W}_N(\nu) = \begin{cases}\lambda^2\ (\lambda >> 1), & \text{when } k=\xi \\ 1 \qquad\quad , & \text{otherwise}\end{cases} \tag{8.2.5}$$

Thus, $\tilde{W}_N(\nu)$ can be looked as an impulse function in the ν domain, and $\tilde{W}_N(\nu) = \lambda^2\delta(\nu-\xi)+1$. Let $\ell \neq 0$, (8.2.1) can be rewritten into

$$\begin{aligned}\mathcal{Q}_r(\ell,\nu) &= \frac{2\pi\sigma_c^2}{M}\delta(\nu-\theta_e)\star(A\delta(\nu-\xi)) \\ &= \frac{2\pi\sigma_c^2 A}{M}\delta(\nu-(\xi+\theta_e))\end{aligned} \tag{8.2.6}$$

where

$$A = e^{-j(2\pi/M)\ell n_e}\lambda^2\sum_{l=0}^{L_h-1}\sum_{m=0}^{L_h-1}h_l h_m^* e^{j\xi(\tau_m-\tau_l)}G(\xi+\ell(2\pi/M))G^*(\xi) \tag{8.2.7}$$

which is not related to the index ν. It can be observed from (8.2.7) that the maximum value of $\mathcal{Q}_r(\ell,\nu)$ with respect to ν is achieved when $\nu = \xi + \theta_e$. Let $F_{\max}(\mathcal{Q}_r(\ell,\nu))$ be this ν, the blind estimate of CFO is given by

$$\hat{\theta}_e = F_{\max}(\mathcal{Q}_r(\ell,\nu)) - \xi \tag{8.2.8}$$

Since the effects of the channel distortion and other interferences are absorbed in A, (8.2.7) and (8.2.8) show that this CFO estimator is not sensitive to the channel fading and other interferences. Thus, it can be employed for the blind CFO estimation in the presence of unknown multipath fading.

8.3 Simulations

The simulations examine the proposed blind CFO estimator in a pulse shaping 16-QAM OFDM system with 512 subcarriers and symbol duration of 562. The normalized CFO is set from 0.01 to 0.46. The multipath channel is modelled as a five-tap time-invariant FIR filter. The absolute error and MSE of the CFO estimate are illustrated in Figures 8.1 and 8.2, which demonstrate the good performance of the proposed estimator in the unknown multipath fading channel with SNR of 5 and 15 dB respectively. From those figures, we can conclude that

1) The absolute error and MSE are both very small even for small SNR case, such as SNR= 5 dB.

2) The absolute error is not very sensitive to the frequency offset. Its range is limited within $(0.001 \backsim 0.01)$.

3) The MSE performance is significantly improved with increasing CFO, particularly when the CFO is small.

4) For the small SNR (e.g. 5dB) and middle SNR (e.g. 10dB), the estimate performances have no significant difference. Thus, the proposed estimator is not very sensitive to the wide-sense stationary noise.

8.4 Conclusion

In this short chapter, we have introduced a novel blind CFO estimator for OFDM systems. Using the second-order cyclic statistics and a special subcarrier weighting function, the proposed estimator could identify the CFO parameter regardless of unknown multipath fading and unknown symbol timing mismatch. Simulation results have shown the excellent performance of this new CFO estimator.

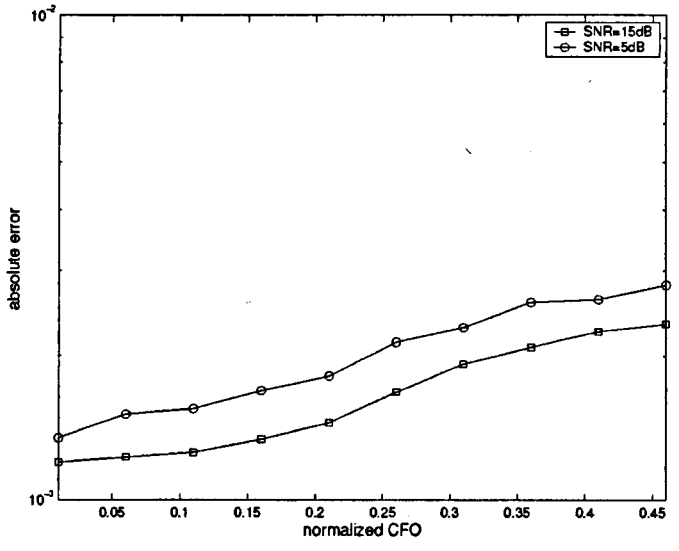

Figure 8.1: Absolute error for CFO estimate with SNR= 5 dB and 15 dB respectively.

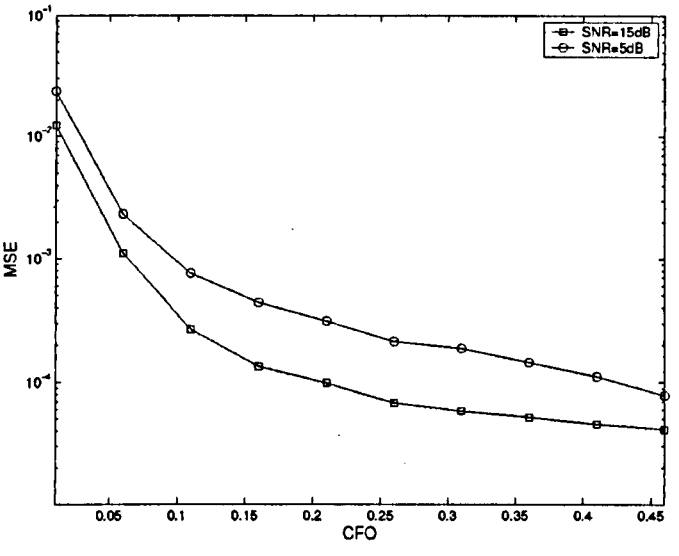

Figure 8.2: Mean square error for CFO estimate with SNR= 5 dB and 15 dB respectively.

Chapter 9

Conclusions and Future Work

9.1 Thesis Summary

In this thesis, we have investigated OFDM transmissions over unknown frequency-selective fading channels. It has been shown that, if the channel knowledge is unknown at receivers, most of OFDM systems need a channel estimator and a equalizer for the coherent or semi-coherent detection. When the channel is slowly time-varying, self-recovery equalization for OFDM systems is possible by employing blind channel estimators (addressed in Chapters 3, 4, and 6) and the superimposed training scheme (addressed in Chapter 5). When the channel is rapidly time-varying, pilot-assisted channel estimation method may be a good solution. Chapter 7 has investigated existing pilot-assisted channel estimator for OFDM and proposed combined channel estimators by utilizing the iterative algorithm. In time-varying channels, CFO induced by Doppler shift may reduce the channel estimation performance. Thus, in Chapter 8, we has developed a novel CFO estimator, which can identify CFO without need of the channel knowledge. Detail of my PhD work is summarized as the following:

Chapter 1 has briefly reviewed OFDM systems, including typical and BP-OFDM

transmissions. The symbol recovery conditions have been carefully investigated in frequency-selective fading channels. Several drawbacks of OFDM systems have also been addressed.

Chapter 2 has introduced existing channel estimation methods for OFDM systems, including training based, pilot-assisted and blind approaches. The channel estimation performance for training based and pilot assisted approaches have been theoretically evaluated using MSE analysis. The optimum training sequences have been designed to minimize MSE in color noise environment.

In Chapter 3, a novel self recovery scheme have been developed for the ICI self-cancelling OFDM system. To improve the overall system performance, we have proposed a differential ICI self-cancelling scheme equipped with a semi-coherent signal detector. This new scheme only needs partial channel information for the signal detection, and thus reduces complexity for blind channel estimation.

The blind channel estimation algorithms used for ICI self-cancelling scheme relies on the subcarrier correlation. In fact, the subcarrier correlation also exists in BP-OFDM systems. Based on this idea, Chapter 4 has proposed the modified BP-OFDM system to enable self recovery schemes. The modified BP-OFDM has demonstrated its very good performance in slowly time-varying channels.

In Chapter 5, we have investigated a superimposed training scheme for OFDM systems. A first-order statistics based channel estimation algorithm has been proposed for identifying time-invariant or slowly time-varying channels. The mean-square channel estimate error has been carefully analyzed for CP-OFDM, ZP-OFDM and space-time block coded OFDM systems. Optimum superimposed training sequences have been designed for OFDM with single transmit antenna and STBC-OFDM with two

transmit antennas. It has been established that the superimposed training scheme offered a low-complexity channel estimator, and thus reduced the complexity of the blind receiver.

Chapter 6 has developed a novel multi-antenna based OFDM transmission schemes. Two transmitted OFDM signals from two uncorrelated transmit antennas were generated by different OFDM modulators. In other words, one transmit antenna sent CP-OFDM signals, and the other sent ZP-OFDM. This kind of transmission enabled a second-order cyclostationary based blind channel estimator at the receiver. The multichannel information could be easily separated without imposing any additional information on the transmitted signals. Simulation results demonstrated the comparable performance of the proposed channel estimator.

Chapter 7 investigated performance of existing pilot-assisted channel estimation methods for OFDM systems, including ideal interpolation and polynomial interpolation approaches. Based on the performance analysis, two types of combined channel estimator have been developed using decision directed algorithms. It has been shown that the CCE approaches could improve the overall system performance in related to traditional approaches with increasing a little bit complexity.

In Chapter 8, a novel CFO estimation method has been developed for pulse-shaping OFDM systems. This method exploited the cyclic statistics induced by pulse-shaping filter and subcarrier weighting function. It has been shown that the CFO could be well identified without need of the channel knowledge and the symbol timing mismatch.

9.2 Future Plan

Recently, OFDM has been selected as a candidate for the fourth generation (4G) mobile communication systems. Due to the high mobility and large operation frequency (e.g. 5G-60GHz), propagation channels become very time-selective. The channel may varies even in one OFDM symbol. It brings big challenges in the channel estimation and equalization. This problem is possible to be solved by employing the pilot-assisted joint frequency offset and channel estimation method. However, very few results have been reported so far on this subject. This motivates me to start my future work on pilot-assisted joint frequency offset and channel estimation for the multiuser OFDM. Both downlink and uplink will be considered.

List of Publications

1) Y. Ma, and Y. Huang, "Blind estimation of carrier frequency offset for OFDM in unknown multipath channels," *IEE Electronics Letters*, vol. 39, pp. 128-130 Jan. 2003 .

2) Y. Ma, Y. Huang, F. Liu and N. Yi, "Blind channel estimation for repetition coded OFDM in block rayleigh fading," *Proceeding of IEEE Globecom 2003*, vol. 1, pp. 30-34, Dec. 2003.

3) Y. Ma, Y. Huang and N. Yi, "Exploiting Frequency Diversity for Blind Channel Estimation in Block Precoded OFDM Systems," *Proceeding of IEEE APCC 2004*, Beijing, China, Dec. 2004.

4) ——, "Combined channel estimator design for pilot-assisted OFDM systems," *Proceeding of IEEE VTC-Spring 2004*, Milan, Italy, May 2004.

5) Y. Ma, Y. Huang, X. Zhu and N. Yi, "Pilot embedded space-time block coded OFDM in unknown multipath fading channel," *Proceeding of IEEE VTC-Spring 2004*, Milan, Italy, May 2004.

6) ——, "Blind channel estimation for OFDM based multitransmitter systems using guard interval diversity," *Proceeding of IEEE VTC-Spring 2004*, Milan, Italy, May 2004.

7) Y. Ma, Y. Huang, W. Al-numaimy, and Y. Xiong, "Orthogonal multicarrier bandwidth modulaiton scheme for wireless communications," *Proceeding of IEEE PIMRC 2001*, vol. 5, pp. 2054-2058 , Sep. 2002.

Reference

[1] R. van Nee and R. Prasad, "OFDM for wireless multimedia communications," Artech House Publishers, Boston, USA, 2000.

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

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

[4] ETSI, "Radio broadcasting system: digital audio broadcasting (DAB) to mobile, portable and fixed receivers," *ETS 300 401*, Second Edition, France May 1997.

[5] ETSI, "Digital broadcasting systems for television, sound and data services; Framing structure, channel coding and modulation for digital terrestrial television," *ETS 300 744*, France 1996.

[6] ANSI, "The DWMT: a multicarrier transceiver for ADSL using M-band wavelets," Tech. Rep., ANSI Mar 1993.

[7] ETSI, "Broadband radio access networks (BRAN); high performance radio local area networks (HIPERLAN) Type 2: system overview," document ETR101683114, France Oct 1999.

[8] ETSI, "Broadband radio access networks (BRAN); high performance radio local area networks (HIPERLAN) Type 2: physical layer," document DTS0023003, France Dec 1999.

[9] IEEE, "Local and metropolitan area networks- Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications - high-speed physical layer in the 5GHz band," *IEEE std. 802.11a*, 1999.

[10] "Local and metropolitan area networks-Part 16: Standard air interface for fixed broadband wireless access systems - media access control modifications and additional physical layer for 2 - 11GHz," *IEEE Std 802.16 unapproved document, 2001.*

[11] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers-Part I: Unification and optimal designs and Part II: Blind channel estimation, synchronization and direct equalization," *IEEE Trans. on Signal Processing*, vol. 47, pp. 1988-2022, July 1999.

[12] B. Muquet, Z. Wang, G. B. Giannakis, M. Courville and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions," *IEEE Trans. on Communications*, vol. 50, pp. 2136-2148, Dec. 2002.

[13] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 184-187, June 1997.

[14] Z. Wang, and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Magzine*, vol. 17, pp. 29-48, May 2000.

[15] T. S. Rappaport "Wireless communications: Principles and Practice," Prentice Hall PTR, New Jersey, USA, 1996.

[16] G. B. Giannakis, and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, pp. 1969-1986, Oct. 1998.

[17] J. Proakis, *Digital Communications*, 4th ed., New York: McGraw-Hill, 2000.

[18] X. Ma, and G. B. Giannakis, "Maximum-diversity transmissions over doubly selective wireless channels," *IEEE Trans. on Information Theory*, vol. 49, pp. 1832-1840, July 2003.

[19] M. H. Hayes, "Statistical digital signal processing and modeling," John Wiley & Sons, inc, New York, USA, 1996.

[20] G. H. Golub, C. F. Van Loan "Matrix computations," The Johns Hopkins University Press, Baltimore and London, 3rd Edition, 1996.

[21] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. on Infomation Theory*, vol. 40, pp. 340-349, March 1994.

[22] Z. Liu, G. B. Giannakis, "Space-time block coded multiple access through frequency-selective fading channels," *IEEE Trans. on Communications*, vol. 49, pp. 1033-1044, June 2001.

[23] Z. Liu, G. B. Giannakis, S. Barbarossa and A. Scaglione, "Transmit-antennae space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE JSAC*, vol. 19, pp. 1352-1364, July 2001.

[24] G. B. Giannakis, Z. Wang, A. Scaglione, and S. Barbarossa, "AMOUR-Generalized multicarrier transceiver for blind CDMA regardless of multipath," *IEEE Trans. on Communications*, vol. 48, pp. 2064-2076, Dec. 2000.

[25] Z. Wang, and G. B. Giannakis, "Block precoding for MUI/ISI-resilient generalized multicarrier CDMA with multirate capabilities," *IEEE Trans. on Commun.*, vol. 49, pp. 2016-2027, Nov. 2001.

[26] T. Kelier and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *Proceedings of IEEE*, vol. 88, pp. 611-640, May 2000.

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

[28] Y. Li, and L. J. Cimini, Jr. "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. on Communications*, vol. 49, pp. 401-404, March 2001.

[29] Y. Zhao, and S. G. Haggman "Intercarrier interference self-cancellation scheme for OFDM mobile communication systems," *IEEE Trans. on Communications*, vol. 49, pp. 1185-1191, July 2001.

[30] E. Serpedin, A. Chevereuil, G. B. Giannakis, P. Loubaton, "Blind channel and carrier frequency offset estimation using periodic modulation precoders," *IEEE Trans. on Signal Processing*, vol. 48, pp.2389-2405, Aug. 2000.

[31] I. Kang, M. P. Fitz, and S. B. Gelfand, "Blind estimation of multipath channel parameters: A model analysis approach," *IEEE Trans. on Communications*, vol. 47, pp. 1140-1150, Aug. 1999.

[32] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Trans. on Communications*, vol. 49, pp. 988-998, June 2001.

[33] J. Li, G. Liu and G. B. Giannakis, "Carrier frequency offset estimation for OFDM-based WLAN," *IEEE Signal Processing Letters*, vol. 8, pp. 80-82, March 2001.

[34] J. v. de Beck, etc.al, "A time and frequency synchronization scheme for multiuser OFDM," *IEEE JSAC*, vol. 17, pp. 1900-1914, Nov. 1999.

[35] ——, "ML estimation of timing and frequency offset in OFDM systems," *IEEE Trans. on Signal Processing*, vol. 45, pp. 1800-1805, July 1997.

[36] T. M. Schmidl and D. C. Cox "Robust frequency and timing synchronization for OFDM," *IEEE Trans. on Communications*, vol. 45, pp. 1613-1621, Dec. 1997.

[37] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Communications Magzine*, vol. 35, pp. 126-133, Dec. 1997.

[38] I. Glover and P. M. Grant, "Digital Communications," Prentice Hall, Europe, 1998.

[39] L. Hanzo, M. Munster, B. J. Choi, and T. Keller, "OFDM and MC-CDMA for broadband multiuser communications, WLANs and broadcasting," IEEE Press, John Wiley & Sons, West Sussex, England, 2003.

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

[41] B. L. Hughes, "Differential space time modulation," *IEEE Trans. on Information Theory*, vol. 46, pp. 2567-2578, Nov. 2000.

[42] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE JSAC.*, vol. 18, pp. 1169-1174, July 2000.

[43] J. K. Cavers, "Pilot symbol assisted modulation and differential detection in fading and delay spread," *IEEE Trans. on Communications*, vol. 43, pp. 2206-2212, July 1995.

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

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

[46] M. Chang, and Y. Su, "Model-based channel estimation for OFDM signals in Rayleigh fading," *IEEE Trans. on Communications.*, vol. 50, pp. 540-544, April 2002.

[47] Y. Zhao and A. Huang, "A novel channel estimation method for OFDM mobile communication systems based on pilot signals and transformation-domain processing," *in Proc. IEEE 47th Vehicular Technology Conference*, Phoenix, AZ, pp. 2089-2093, May 1997.

[48] S. Coleri, M. Ergen A. Puri and A. Bahai, "Channel estimation techniques based on pilot arragement in OFDM systems" *IEEE Trans. on Broadcasting*, vol. 48, pp. 223-229, Sep. 2002.

[49] O. Edfors, M. Sandell, and J.-J. van de Beek, "OFDM channel estimation by sigular value decomposition," *IEEE Trans. on Communications*, vol. 46, pp. 931-939, July 1998.

[50] A. K. Nandi (Ed.), "Blind estimation using Higher-Order Statistics," Kluwer Academic Publishers, Amsterdam, Netherlands , 1999.

[51] J. K. Tugnait, "Identification of linear stochastic systems via second- and fourth-order cumulant matching," *IEEE Trans. on Infomation Theory*, vol. 33, pp. 393-407, May 1987.

[52] J. K. Tugnait, "Blind estimation and equalization of digital communication FIR channels using cumulant matching," *IEEE Trans. on Communications*, vol. 43, pp. 1240-1245, Feb/March/April 1995.

[53] B. Porate and B. Friedlander, "Blind equalization of digital communication channels using high-order statistics," *IEEE Trans. on Signal Processing*, vol. 39, pp. 522-526, 1991.

[54] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication systems," *IEEE Trans. on Communications*, vol. 28, pp. 1867-1875, 1980.

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

[56] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. on Signal Processing*, vol. 43, pp. 2982-2993, Dec. 1995.

[57] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationary for blind channel identification in OFDM systems," *IEEE Trans. on Signal Processing*, vol. 47, pp. 848-856, Mar. 1999.

[58] S. M. Kay, "Fundamentals of statistical signal processing: estimation theory," Prentice Hall, New Jersey, 1993.

[59] L. Ludeman, "Random processes filtering, estimation, and detection," John Wiley and Sons, New Jersey, 2003.

[60] A. V. Oppenheim and R. W. Schafer, "Discrete-time signal processing," Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1989.

[61] H. Bolcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," *IEEE Trans. on Signal Processing*, vol. 50, pp. 96-109, Jan. 2002.

[62] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels: A cyclostationary approach," *IEEE Trans. on Communications*, vol. 46, pp. 400-411, Mar. 1998.

[63] W. A. Gardner, "Introduction to random process: With applications to signals and systems" Macmillan Publishing Company, New York, 1986.

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

[65] S. Zhou, B. Muquet and G. B. Giannakis, "Subspace-based blind channel estimation for block precoded space-time OFDM," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1215-1227, May 2002.

[66] S. Zhou, G. B. Giannakis, and A. Scaglione, "Long codes for generalized FH-OFDMA through unknown multipath channels," *IEEE Trans. on Communications*, vol. 49, pp. 721-733, Apr. 2001.

[67] J. Medbo and P. Schramm, "Channel Models for Hiperlan/2 in Different Indoor Scenarios," *BRAN 3ERI085B, March 1998.*

[68] P. Dambacher, "Digital broadcasting," Institution of Electrical Engineers, London, 1996.

[69] P. Frenger and A. Svensson, "Decision-directed coherent detection in multicarrier systems on Rayleigh fading channels," *IEEE Trans. on Vehicel Technology* vol. 48, pp. 490-498, March 1999.

[70] W. C. Jakes, Jr, "Microwave mobile communications" John Wiley & Sons, New York, 1974.

[71] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Trans. on Broadcasting*, vol. 46, pp. 170-176, June 2000.

[72] J. K. Tugnait, and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Communication Letters*, vol. 7, pp. 413-415, Sep. 2003.

[73] A. F. Naguib, N. Seshadri, and R. Calderbank, "Space-time coding and signal processing for high data rate wireless communications," *IEEE Signal Processing Magzine*, vol. 17, pp. 76-92, May 2000.

[74] A. Paulraj, R. Nabar and D. Gore, "Introduction to space-time wireless communications," Cambridge university express, United Kingdom, 2003.

[75] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: perforamce criterion and code construction," *IEEE Transaction on Information Theory*, vol. 44, pp. 744-765, March 1998.

[76] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE JSAC*, vol. 16, pp. 1451-1458, Oct. 1998.

[77] Y. Li, J. C. Chung, and N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high rate data wireless networks," *IEEE JSAC*, vol. 17, pp. 1233-1243, July 1999.

[78] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transaction on Information Theory*, vol. 45, pp. 1456-1467, July 1999.

[79] ETSI, "European digital celluar telecommunication system (Phase 2): radio transmission and reception," document DTS0023003, France, July 2003.

[80] X. Wang and K. Liu, "Channel estimation for multicarrier modulation system using a time-frequency polynomial model," *IEEE Trans. on Comms.*, vol. 50, pp. 1045-1048, July 2002.

[81] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE JSAC*, vol. 17, pp. 461-471, March 1999.

[82] Y. Ma, and Y. Huang, "Blind estimation of carrier frequency offset for OFDM in unknown multipath channels," *IEE Electronics Letters*, vol. 39, pp. 128-130, Jan. 2003.

[83] Y. Ma, Y. Huang, F. Liu and N. Yi, "Blind channel estimation for repetition coded OFDM in block rayleigh fading," *Proceeding of IEEE Globecom 2003*, vol. 1, pp. 30-34, Dec. 2003.

[84] Y. Ma, Y. Huang and N. Yi, "Exploiting Frequency Diversity for Blind Channel Estimation in Block Precoded OFDM Systems," *Proceeding of IEEE APCC*, vol. 1, pp. 338-342, China, Dec. 2004.

[85] ——, "Combined channel estimator design for pilot-assisted OFDM systems," *Proceeding of IEEE VTC-Spring*, vol. 4, pp. 1888-1891, May 2004.

[86] Y. Ma, Y. Huang, X. Zhu and N. Yi, "Pilot embedded space-time block coded OFDM in unknown multipath fading channel," *Proceeding of IEEE VTC-Spring*, vol. 1, pp. 450-454, May 2004.

[87] ——, "Blind channel estimation for OFDM based multitransmitter systems using guard interval diversity," *Proceeding of IEEE VTC-Spring*, vol. 1, pp. 440-444, May 2004.

[88] Y. Ma, Y. Huang, W. Al-numaimy, and Y. Xiong, "Orthogonal multicarrier bandwidth modulaiton scheme for wireless communications," *Proceeding of IEEE PIMRC*, vol. 5, pp. 2054-2058, Sep. 2002.