BLIND CHANNEL ESTIMATION FOR MIMO OFDM COMMUNICATION SYSTEMS

CHEN XI

NATIONAL UNIVERSITY OF SINGAPORE

2009

BLIND CHANNEL ESTIMATION FOR MIMO OFDM COMMUNICATION SYSTEMS

CHEN XI

(*M.Sc., National University of Singapore, Singapore*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

Acknowledgments

I would like to express grateful appreciation and gratitude to my supervisor, Dr. A. Rahim Leyman, and co-supervisor, Dr. Liang Ying Chang, for their supports and guidance during the course of my studies. Their advice and guidance on ways of performing research are most invaluable and on many occasions, have served as a driving force for me to keep on going. I have learned enormously from them not only how to do research, but also how to communicate effectively. Especially, I am indebted to Dr. Leyman for his great concern in matters outside of academics during these years, and his readiness to assist me in my future road-map.

I would also like to thank all friends at Institute for Infocomm Research (I2R) and National University of Singapore (NUS) who have supported me and given me much joy during these years.

In addition, I am grateful to I2R for providing me with conductive environment and facilities needed to complete my course of studies.

The completion of this thesis would not be possible without the love and support of my mother. She has been there for me whenever I needed a helping hand. I also thank my father who passed away in August 1998. He is always my moral support and is always together with me in my heart.

Contents

Summary

The main contribution of this thesis is the development of three blind channel estimation and one blind source separation algorithms for MIMO-OFDM systems. The first proposed channel estimation algorithm is a subspace based method. We study the inherent structure of autocorrelation matrices of the system output and construct a new criterion function, minimizing which leads to a close form solution of the channel matrices. The second algorithms is based on the assistance of a non-redundant linear precoder, which brings in cross-correlations between the signals transmitted on different subcarriers. For the third one, we exploit the spectra correlation of the system output. It is shown that when the source signals have distinct spectra correlation, then the channel matrix can be estimated up to a complex scalar and column permutation. Therefore, the problem of the ambiguity matrix in many of the existing blind channel estimation algorithm can be avoided. The blind source separation algorithm proposed in this thesis is a geometric based non-iterative algorithm based on the assumption that the source signal has finite alphabet. The proposed algorithm compares favorably with the existing hyperplane-based and kurtosis-based algorithms.

List of Figures

List of Symbols

Chapter 1

Introduction

1.1 Towards Fourth Generation Mobile Systems

Comparing with the traditional wired communication technologies, wireless communication is an emerging field, which has seen enormous growth in the last several years. Market demands for higher cellular density in urban areas, broadband internet wireless, and better data security, while using a minimum amount of frequency spectrum is driving wireless developments forward at an amazing speed. Ubiquitous connectivity (i.e., connectivity anytime and anywhere) to the internet, to company's intranets, or to other data services is creating room for applications that might not even be thought of today.

The mobile communication systems are often categorized as different generations depending on the services offered. Figure 1.1 shows the evolution routine of the mobile communication systems. The first generation (1G) comprises the analog frequency-division multiplexing access (FDMA) systems such as the NMT (Nordic Mobile Telephone) [1] and AMPS (Advanced Mobile Phone Services) [2]. The second

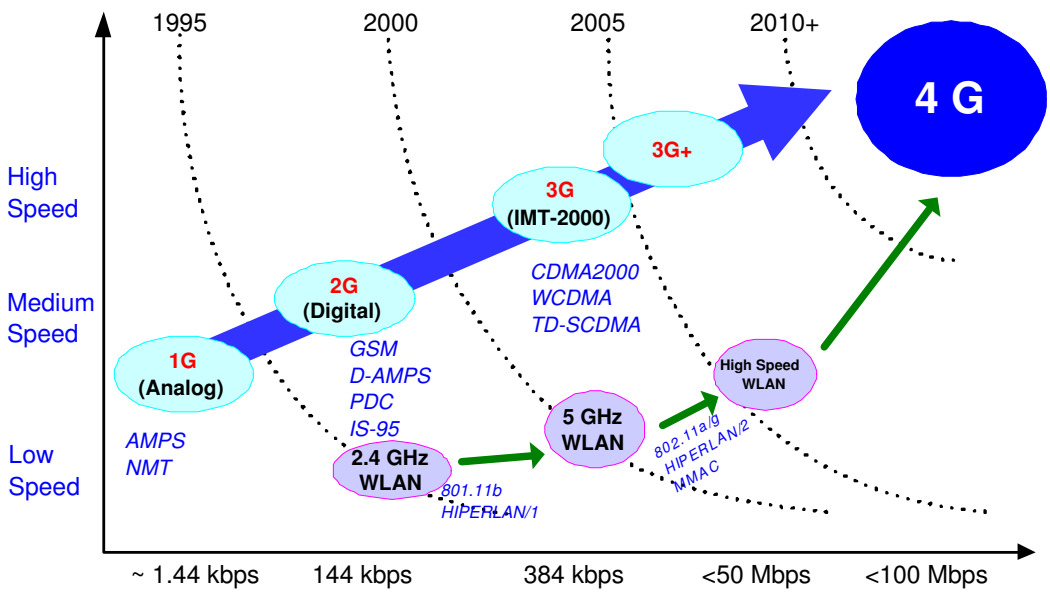

Figure 1.1: Current and Future Wireless Communication Systems

generation (2G) consists of the first mobile digital communication systems such as the time-division multiple access (TDMA) based GSM (Global System for Mobile Communication) [8], D-AMPS (Digital AMPS) [1], PDC (Pacific Digital Cellular) [2] and the code-division multiple access (CDMA) based system IS-95 [9]. In 1999, the International Telecommunication Union (ITU) approved an industry standard for third generation (3G) of mobile communication systems. This standard, called International Mobile Telecommunications-2000 (IMT-2000) [2], strives to provide higher data rates than current second-generation (2G) systems. 2G systems are mainly targeted at providing voice services, while 3G systems will be able to support a wide range of applications including internet access, voice communications and mobile videophones. In addition to this, a large number of new applications will emerge to utilize the permanent network connectivity, such as wireless appliances, notebooks with built in mobile phones, remote logging, wireless web cameras, car

navigation systems, and so forth [10]- [13].

In Europe auctions of 3G licenses of the radio spectrum began in 1999. In the United Kingdom, 90 MHz of bandwidth [12] was auctioned off for £22.5 billion [13]. In Germany the result was similar, with 100 MHz of bandwidth raising $46 billion (US) [11]. This represents a value of around $450 Million (US) per MHz. The length of these license agreements is 20 years [12] and so to obtain a reasonable rate of return of 8% on investment, $105 Million (US) per MHz must be raised per year. It is therefore vitally important that the spectral efficiency of the communication system is maximised, as this is one of the main limitations to providing a low cost high data rate service.

In parallel to the development of the 3G systems, there has been an increasingly interesting in high-speed wireless local area networks (WLANs). The WLAN systems do not offer the same wide area coverage as the 3G mobile systems do, but within their limited coverage area they provide much higher data rates.

Since the beginning of the 1990's, WLANs for the 900 MHz, 2.4 GHz and 5 GHz license-free ISM (Industrial, Scientific and Medical) bands have been available, based on a range of proprietary techniques [6]. In June 1997 the Institute of Electrical and Electronics Engineers (IEEE) defined an international interoperability standard, called IEEE 802.11 [34]. This standard specifies a number of Medium Access Control (MAC) protocols and three different Physical Layers (PHYs) which support data rates of 1 Mbps and optionally 2 Mbps. In July 1998 IEEE extended the IEEE 802.11 standard to IEEE 802.11b which describes a PHY providing a basic rate of 11 Mbps and a fall-back rate of 5.5 Mbps. Meanwhile, the European

Telecommunication Standards Institute (ETSI) specified the European WLAN standard, called HIPERLAN/1 [35], which defines data rates ranging from 1 Mbps to 20 Mbps. However, in contrast to the IEEE 802.11b, no commercial products have been developed that support the HIPERLAN/1 standard.

Motivated by the demand for even higher data rates, a new standard called IEEE 802.11a was ratified in 2000, which is based on the OFDM as the transmission technique for the newly available spectrum in the 5 GHz band. It defines data rates between 6 and 54 Mbps [59]. To make sure that these data rates are also available in the 2.4 GHz band, mid 2003 IEEE standardization group issued a similar standard for this band named IEEE 802.11g [34]. At the same time, the ETSI working group named Broadband Radio access Networks (BRAN) in Europe and Multimedia Mobile Access Communication (MMAC) group in Japan published their new generation of WLAN standards, called HIPERLAN/2 [5] and the MMAC [6] respectively. Following the selection of OFDM by the IEEE 802.11a standardization group, both the ETSI BRAN and MMAC working group adopted OFDM for their PHY.

While the roll-out of 3G systems is under progress, research activities on the fourth generation (4G) have already started [14]- [17]. According to the increasing demand of wireless data traffic, it is obvious that the main goal in developing next generations of wireless communication systems are increasing the link throughput (i.e., bit rate) and the network capacity. Few of the aims of 4G networks have yet been published, however it is likely that they will be to extend the capabilities of 3G networks, allowing a greater range of applications, and improved universal access.

Ultimately 4G networks should encompass broadband wireless services, such as High Definition Television (HDTV) (4-20 Mbps) and computer network applications (1-100 Mbps). This will allow 4G networks to replace many of the functions of WLAN systems. In fact, a popular vision suggests to combine WLAN systems for high peak data rates with cellular systems for wide area-coverage, and to allow inter-system handovers [18]. On the other hand, cost of service must be reduced significantly from 3G networks. The spectral efficiency of 3G networks is too low to support high data rate services at low cost. As a consequence one of the main focuses of 4G systems will be to significantly improve the spectral efficiency [17].

In addition to high data rates, future systems must support a higher Quality Of Service (QoS) than current cellular systems, which are designed to achieve 90-95% coverage [19], i.e. network connection can be obtained over 90-95% of the area of the cell. This will become inadequate as more systems become dependent on wireless networking. As a result 4G systems are likely to require a QoS closer to 98-99.5%. In order to achieve this level of QoS it will require the communication system to be more flexible and adaptive. In many applications it is more important to maintain network connectivity than the actual data rate achieved. If the transmission path is very poor, e.g. in a building basement, then the data rate has to drop to maintain the link. Thus the data rate might vary from as low as 1 kbps in extreme conditions, to as high as 20 Mbps for a good transmission path. Alternatively, for applications requiring a fixed data rate, the QoS can be improved by allocating additional resources to users with a poor transmission path.

1.2 OFDM

1.2.1 Wideband Air-interface Design Using OFDM

Multipath propagation is the primary issue in the air-interface design for wideband (high data-rate) communication systems. Multiple replicas of the transmitted signal arrive at the receiver with various propagation delays, due to reflections on all kinds of objects and obstacles in the environment. Therefore, if a high-rate data stream is transmitted on such a channel, multiple data symbols interfere with each other, making the data recovery difficult. This phenomenon is called "inter-symbol-interference" (ISI). The standard solution to the ISI problem is to design a linear filter at the receiver side that employs a means for compensating or reducing the ISI in the received signal. This compensation method is called equalization. The main challenge is to adapt the filter coefficients to the time-variant channel conditions. The adaptation could be computationally extremely demanding, particularly if long filters are required as in the case where the channel impulse response spans many data symbols.

Fortunately, Orthogonal Frequency Division Multiplexing (OFDM) can drastically simplify the equalization problem [6]. In OFDM, the high-rate serial data stream is split up into a number (several dozens up to a few thousand) of parallel data streams at a much lower (common) symbol rate, which are modulated on a set of subcarriers (frequency division multiplexing). High spectral efficiency is achieved by selecting a specific (orthogonal) set of subcarrier frequencies. Inter-carrier-interference is avoided due to the orthogonality, although the spectra of the subcarriers actually overlap (see Figure 1.2) [6]. The idea is to make the symbol

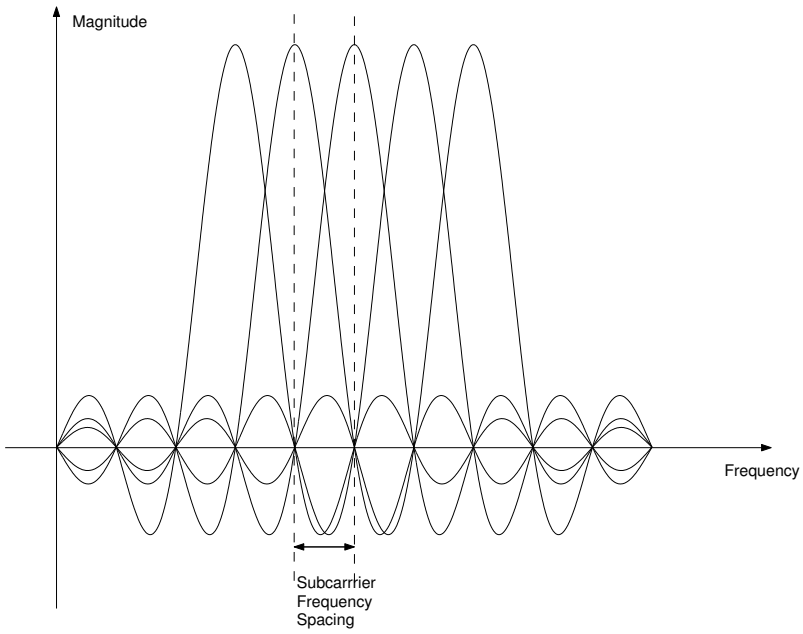

Figure 1.2: Spectrum overlap in OFDM

period long with respect to the channel impulse response in order to reduce ISI. This implies that the bandwidth of the subcarriers gets small (with respect to the channel's coherence bandwidth [25]), thus the impact of the channel is reduced to an attenuation and phase distortion of the subcarrier symbols ("flat fading"), which can be compensated by efficient one-tap equalization.

Thus, it is quite attractive in the robustness against frequency selective fading, especially for high-speed data transmission [26]. In practice, OFDM has already been used in European digital audio broadcasting (DAB), digital video broadcasting (DVB) systems and high performance radio local area network (HIPERLAN) [23]-[24], [27]. Furthermore, combined with Multiple-Input Multiple-Output (MIMO) wireless technology, OFDM has been recognized as one of the most promising techniques for the future 4G systems [10].

The first study of OFDM was published by Chang in 1966 [24]. He presents

a principle for transmitting messages simultaneously through a linear bandlimited channel without interchannel (ICI) and intersymbol interference (ISI). In 1971, a major contribution to OFDM was presented by Weinstain and Ebert [25], who used the discrete Fourier transform (DFT) to perform baseband modulation and demodulation. This technique involved assembling the input information into blocks of N_c complex symbols, one for each subchannel. An inverse FFT is performed on each block, and the resultant transmitted serially. At the receiver, the information is recovered by performing an FFT on the received block of signal samples. This form of OFDM is often referred to as Discrete Multi-Tone (DMT). The most significant advantage of this DMT approach is the the efficiency of the FFT algorithm. An N_c-point FFT requires only on the order of $N_c \log N_c$ multiplications, rather than N_c^2 as in a straightforward computation.

Another important contribution was due to Peled and Ruiz in 1980 [26], who introduced the cyclic prefix (CP) or cyclic extension, solving the orthogonality problem. Instead of using an empty guard space, they filled the guard space with a cyclic extension of the OFDM symbol. This effectively simulates a channel performing cyclic convolution, which implies orthogonality over dispersive channels when the CP is longer than the impulse response of the channel [24], [26]. This introduces an energy loss proportional to the length of CP, but the zero ISI generally motivates the loss.

1.2.2 Main Advantages and Disadvantages of OFDM

The advantages of OFDM, especially in the multipath propagation, interference and fading environment, make the technology a promising alternative in digital communications including mobile multimedia. The advantages of OFDM are:

- Efficient use of the available bandwidth since the subchannels are overlapping.
- Spreading out the frequency fading over many symbols. This effectively randomizes the burst errors caused by the Rayleigh fading, so that instead of several adjacent symbols (in time on a single-carrier) being completely destroyed, (many) symbols in parallel are only slightly distorted.
- The symbol period is increased and thus the sensitivity of the system to delay spread is reduced.

On the other hand, there are also problems associated with OFDM system design:

- OFDM signal is contaminated by non-linear distortion of transmitter power amplifier, because it is a combined amplitude-frequency modulation (it is necessary to maintain linearity).
- OFDM is very sensitive to carrier frequency offset caused by the jitter of carrier wave and Doppler effect caused by moving of the mobile terminal.

1.2.3 MIMO-OFDM

Research in the information theory, performed in the early 90's, has revealed that important improvement in spectral efficiency can be achieved when multiple antennae are applied at both the transmitter and receiver side, especially in rich-scattering

environments. This has been shown for wireless communication links in both narrowband channels [28] as well as wideband channels [29], and it initiated a lot of research activity to practical communication schemes that exploit this spectral-efficiency enhancement. The resulting multiple-transmit multiple-receive antenna, i.e., Multiple-Input Multiple-Output (MIMO), techniques can basically be split into two groups: Space-Time Coding (STC) [30]- [32] and Space Division Multiplexing (SDM) [28], [29], [33]. STC increases the robustness / performance of the wireless communication system by transmitting different representations of the same data stream (by means of coding) on the different transmitter branches, while SDM achieves a higher throughput by transmitting independent data streams on the different transmitter branches simultaneously and at the same carrier frequency.

The highest spectral efficiency gains are achieved when the individual channels from every transmit antenna to every receive antenna can be regarded to be independent. In practice this is the case in rich-scattering environments with no Line of Sight (LOS) path present between transmitter and receiver. So, especially for enhancement of the throughput of wireless applications in rich-scattering environment, MIMO techniques are appealing. In general, MIMO can be considered as an extension to any Single-Input Single-Output (SISO), Single-Input Multiple-Output (SIMO), i.e., receiver diversity, or Multiple-Input Multiple-Output (MISO), i.e., transmit diversity, system operating in these environments.

The WLAN standards IEEE 802.11b, IEEE 802.11a/g indicate that they are usually deployed in an indoor environment, while the probability of having no direct communication path between transmitter and receiver is high [34]. So, we can

conclude that the deployment conditions of WLAN systems are most favorable for applying MIMO. In fact, these standards are the WLAN standards that currently gain the most momentum. They are both based on OFDM. Thus the robustness of OFDM against frequency-selective fading and the favorable properties of indoor radio channels for MIMO techniques lead to the very promising combination of MIMO-OFDM as potential solution to satisfy the main goals in developing next generations of wireless communication systems. As such, MIMO-OFDM techniques are attractive candidates for high data rate extensions of the IEEE 802.11a and 802.11g standards. As an example the IEEE 802.11 Task Group 'n' (TGn) can be mentioned which is planning to define high-data rate WLAN extensions up to 250 Mbps [34].

1.3 Blind Channel Estimation

As mentioned in the previous section, multipath propagation is the primary issue in the wideband wireless communication systems. In order to recover the transmitted signal at the receiver, it is essential to know some information about the channel. The cancellation of channel effects is referred to as equalization. It is possible to construct the equalizer directly without explicitly estimating the channel, or indirectly, by first estimating the channel. In either case, the transmitter should send a signal known a priori by the receiver which is called training. However, most wireless devices will be battery powered. Hence the transmission of training signals will seriously affect the longevity of such devices. Moreover, training increases the overhead of the transmitted signal, thus reducing the net data transmission rate.

Thus, it is reasonable to use blind channel estimation methods to possibly reduce the amount of training required significantly. Typically, some special property of the transmitted signal is exploited for blind channel estimation.

Blind equalization methods provide attractive solutions since they do not require any known transmitted data for channel estimation and equalization purposes [4], [39]- [42]. Instead, they use the statistical and structural properties of the communication signals (Finite alphabet, constant modulus, sub-spaces orthogonality). Channel identification or equalization requires that information about both the channel amplitude and phase responses can be acquired from the received signal. A symbol rate sampled communications signal is typically wide sense stationary (WSS). Second order statistics from a WSS process contain no phase information. Hence one can not distinguish between minimum phase and non-minimum phase channels. Therefore, other statistical properties of the signal have to be used to extract the phase information.

The communication signals are typically non-Gaussian. Hence, the Higher Order Statistic (HOS) of the signals are non-zero and may also be exploited in equalization. HOS retain the phase information as well [36]. Early blind algorithms were either implicitly or explicitly based on HOS. In time domain, HOS are represented by higher than second order cumulants and moments. However, Higher order statistics and spectra may not provide a feasible approach for constructing practical equalizers. They have a large variance and consequently large sample sets are needed in order to obtain reliable channel estimates. This is a severe drawback, in particular in applications where the channel is time varying, data rates are high or low

computational complexity is needed [37].

In case a multiple-output model resulting from oversampling or employing multiple receivers is used, the received signal typically possesses the cyclostationarity property, i.e. signal statistics such as the autocorrelation function are periodic. Gardner discovered in [38] the fact that non-minimum phase channel equalization/identification may be obtained from the Second Order Statistics of the received signal because the cyclic autocorrelation function preserves the phase information. Hence, smaller sample sizes than for HOS are required for the convergence of the estimated statistics. The main drawback is that some channel types may not be identified [39]. In particular, the channel cannot be identified if the subchannels resulting from oversampling share common zeroes. If the mutiple-output model is obtained by using an antenna array at the receiver with antenna elements well separated this limitation is less severe. This is because the resulted sub-channels are uncorrelated.

The channel impulse responses can be blindly identified and equalized under certain conditions, usually up to a complex scalar ambiguity. The identification conditions, the inherent ambiguities as well as the slow convergence, may limit the applicability of blind equalization methods in practical communication systems. Because of their high potential in providing higher effective data rates it is of great interest to study their feasibility in particular communication systems. Using a limited number of training symbols may solve their problems. Limited training data in conjunction with blind algorithms leads to semi-blind methods. Semi-blind methods are a more feasible solution for practical communication systems since

they combine the benefits of both training based and blind methods. They usually achieve better performance than the traditional training based algorithms whilst using a smaller amount of training data [40]. They have a larger sample support, since both known symbols and statistical information are used. Consequently they exhibit a lower variance of the estimates.

1.4 Blind Source Separation

Blind source separation (BSS) refers to the problem of recovering signals from several observed linear mixtures. In a large number of cases, statistically independent sources are mixed through an unknown channel where only the channel outputs (observed signals) are measurable. The objective is: based on the information contained in observed signals design a separation network to extract the original sources. In this thesis, some channel estimation algorithms are introduced, which enables blind channel estimation of MIMO-OFDM systems up to a unitary ambiguity matrix, which need to be further removed by using source separation processes. Hence, BSS may be regarded as the extended work of blind channel estimation.

There appears to be something magical about blind source separation: we are estimating the original source signals without knowing the parameters of mixing and/or filtering processes. It is difficult to imagine that one can estimate this at all. In fact, without some a priori knowledge, it is not possible to uniquely estimate the original source signals. However, one can usually estimate them up to certain indeterminacies. In mathematical terms these indeterminacies and ambiguities can be expressed as arbitrary scaling, permutation and delay of estimated source signals

[55]. These indeterminacies preserve, however, the waveforms of original sources. Although these indeterminacies seem to be rather severe limitations, in a great number of applications these limitations are not essential, since the most relevant information about the source signals is contained in the temporal waveforms or time-frequency patterns of the source signals and usually not in their amplitudes or order in which they are arranged in the output of the system.

Although many different source separation algorithms are available, their principles can be summarized by the following four fundamental approaches:

- The most popular approach exploits as the cost function some measure of signals statistical independence, non-Gaussianity or sparseness. When original sources are assumed to be statistically independent without a temporal structure, the higher-order statistics (HOS) are essential (implicitly or explicitly) to solve the BSS problem. In such a case, the method does not allow more than one Gaussian source [43]- [45].

- If sources have temporal structures, then each source has non-vanishing temporal correlation, and less restrictive conditions than statistical independence can be used, namely, second-order statistics (SOS) are often sufficient to estimate the mixing matrix and sources. Along this line, several methods have been developed [46]- [50]. Note that these SOS methods do not allow the separation of sources with identical power spectra shapes or i.i.d. (independent and identically distributed) sources.

- The third approach exploits non-stationarity (NS) properties and second order statistics (SOS). Mainly, we are interested in the second-order non-stationarity

in the sense that source variances vary in time. The non-stationarity was first taken into account by [51] and it was shown that a simple decorrelation technique is able for wide class of source signals to perform the BSS task. In contrast to other approaches, the non-stationarity information based methods allow the separation of colored Gaussian sources with identical power spectra shapes. However, they do not allow the separation of sources with identical non-stationarity properties. There are some recent works on non-stationary source separation [52], [53]. Methods that exploit either the temporal structure of sources (mainly second-order correlations) and/or the non-stationarity of sources, lead in the simplest scenario to the second-order statistics BSS methods. In contrast to BSS methods based on HOS, all the second-order statistics based methods do not have to infer the probability distributions of sources or nonlinear activation (score) functions (see next sections).

- The fourth approach exploits the various diversities of signals, typically, time, frequency, (spectral or "time coherence") and/or time-frequency diversities, or more generally, joint space-time-frequency (STF) diversity. Such approach leads to concept of Time-Frequency Component Analyzer (TFCA) [20]. TFCA decomposes the signal into specific components in the time-frequency domain and computes the time-frequency representations (TFRs) of the individual components. Usually components are interpreted here as localized, sparse and structured signals in the time-frequency plain (spectrogram). In other words, in TFCA components are estimated by analyzing the time-frequency distribution of the observed signals. TFCA provides an elegant and promising

solution to suppression of some artifacts and interference via masking and/or filtering of undesired - components.

1.5 Outline

The structure of the thesis is organized as follows.

- In Chapter 2, first the principle of OFDM is explained. Second, the combination of MIMO and OFDM is described. The core idea is that the wideband frequency-selective MIMO channel by means of the MIMO-OFDM processing is transferred to a number of parallel flat-fading MIMO channels [6].

- In Chapter 3, we present a novel subspace based blind channel estimation algorithm for MIMO-OFDM systems driven by either white or colored source. This algorithm is efficient and works in ill conditioned environments.

- In Chapter 4, we design a nonredundant linear precoder for MIMO-OFDM which enables blind channel estimation. The identifiability of the proposed algorithm is guaranteed even when the channel matrices share common zeros at subcarrier frequencies.

- In Chapter 5, we propose a geometric based blind source separation method to resolve the ambiguity matrix which is yet to be removed by using the blind channel estimation methods proposed in the previous chapters. Moreover, this proposed separation method is also a general blind separation method for all flat fading channels.

- In Chapter 6, we exploit the second-order spectra correlations of the system output to blindly estimate the FIR channel matrix of the MIMO-OFDM systems, which are driven by stationary or cyclostationary and nonwhite inputs with distinct but known correlations. By using this method, the ambiguity matrix problem which indeed exists in many existing methods can be avoided.

- In the last chapter, we draw the conclusions of this thesis, and present some prospective work in order to address future key problems in MIMO-OFDM communication systems.

Chapter 2

MIMO-OFDM System Model

2.1 Introduction

In general, a MIMO system takes advantage of the spatial diversity obtained by spatially separated antennae in a dense multipath scattering environment. It may be implemented in a number of different ways to obtain either a diversity gain to combat signal fading or to obtain a capacity gain. One of the potential application areas is that of Wireless Local Area Networks (WLANs).

The current WLAN standards IEEE 802.11a and IEEE 802.11g [34] are based on Orthogonal Frequency Division Multiplexing (OFDM) [6], [56]. A high-data-rate extension of these standards could be based on Space Division Multiplex (SDM) [28]. That is, the OFDM-based transmission system can be extended to a MIMO architecture, which leads to the promising combination of the data rate enhancement of SDM with the relatively high spectral efficiency and the robustness against frequency-selective fading and narrowband interference of OFDM. An advantage of wireless LAN systems is that they are mainly deployed in indoor environments. These en-

vironments are typically characterized by richly scattered multipath. As explained in [57], this is good condition for having a high MIMO capacity.

In this chapter, we review the basic principles of the OFDM systems. The mathematical system model of the OFDM systems using the IFFT technique is derived. Then we extend it to the general MIMO-OFDM case. Although the blind channel estimation algorithms proposed in this thesis are mainly designed for the CP-OFDM systems, the ZP-OFDM system model is also studied in this chapter as a reference.

The rest of this chapter is organized as follows. First the multipath fading channel in a typical wireless communication system is discussed in Section 2.2, second the brief introduction to OFDM is given in Section 2.3. We review the block diagram of a "classic" OFDM system, which employs a guard interval to mitigate the impairments of the multipath channel. Then the combination of MIMO and OFDM is described in Section 2.4, where the relation between the transmitted and received MIMO-OFDM symbols are captured in matrix form.

2.2 The Multipath Fading Channel

Due to the presence of reflecting, scattering, relative motion between transmitters and receivers, etc., two or more versions of the signal waveforms transmitter by the receiver arrive at the receiver at slightly different times. This is known as multipath fading. Figure 2.1 shows the diagram of multipath fading in wireless communication systems. Consider the channel with a total of P paths. Each signal path has its own individual path length and complex valued gain. Since the resultant signal at

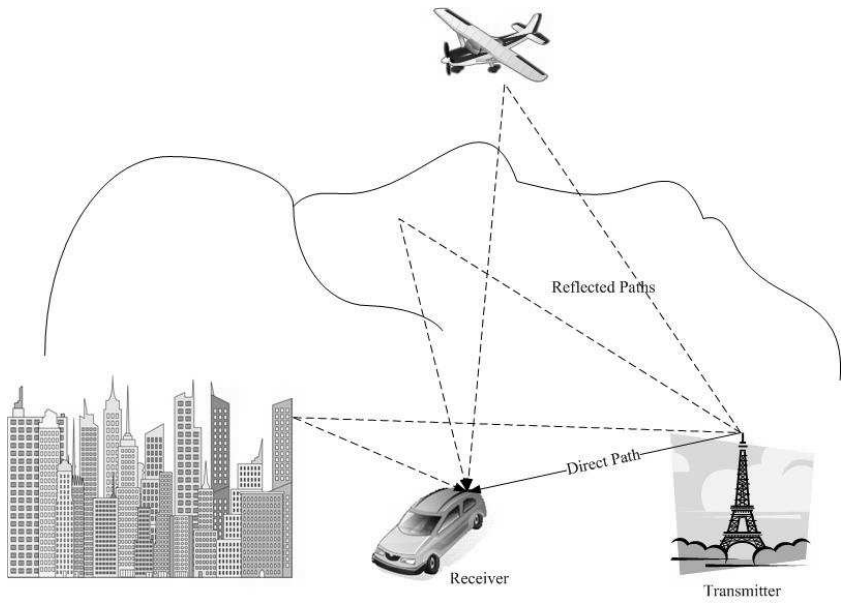

Figure 2.1: Diagram of Multipath Fading

the receiver is a superposition of the signals from all P paths, we may write the baseband impulse response of a multipath channel as [2]

$$h(t,\tau) = \sum_{p=1}^{P} a_p \exp\left\{j(2\pi f_c \tau_p(t) + \phi_p)\right\} \delta(\tau - \tau_p(t)) \tag{2.1}$$

where a_p, $\tau_p(t)$ and ϕ_p are gain coefficient, delay, and phase of the pth path respectively, while f_c is the carrier frequency. It is clearly seen that the baseband channel impulse response is not only affected by the properties of the delayed path but also the carrier frequency. Hence, under the worst case where one or more of the delays τ_p, $p = 1, 2, \cdots, P$ exceed the two sided bandwidth of the transmitted signal $2B$, and there is relative motion between the transmitter and receiver, then the baseband received signal can be written as

$$r(t) = h(t,\tau) * s(t) \tag{2.2}$$

where $s(t)$ and $r(t)$ are the baseband transmitted and received signal respectively, and $*$ denotes the convolution. In such case, the channel is commonly referred to as time- and frequency-selective, frequency-selective fading or dispersive fading channel.

In digital communication systems, it is convenient to use a discretized version of the channel model.The most common mathematical model used for such channel is the tapped delay line (TDL) model [2]. When the Nyquist sampling criterion [3] is satisfied, the kth sample of the received signal is expressed by

$$r[k] \triangleq r(kT_s) = h(kT_s, \tau) * s(kT_s) \tag{2.3}$$

where T_s is the sample duration. Since the continues time band limited source signal can be interpolated by the sinc function, i.e.

$$s(t) = \sum_{l=-\infty}^{\infty} s[l]\text{sinc}\left(\frac{t - lT_s}{T_s}\right) \tag{2.4}$$

where $s[l] \triangleq s(lT_s)$. Substitute Eqn.(2.4) to Eqn.(2.3), we get

$$r[k] = \sum_{l=-\infty}^{\infty} s[k-l]\left(h(kT_s, \tau) * \text{sinc}\left(\frac{t - lT_s}{T_s}\right)\right) \tag{2.5}$$

Hence we define

$$h[k, l] \triangleq h(kT_s, \tau) * \text{sinc}\left(\frac{t - lT_s}{T_s}\right) \tag{2.6}$$

and we may then write the received samples in terms of discrete transmitted samples and channel samples as

$$r[k] = \sum_{-\infty}^{\infty} h[k, l]s[k-l] \tag{2.7}$$

When channel is time invariant, the above Eqn.(2.7) can also be written as

$$r[l] = \sum_{-\infty}^{\infty} h[l]s[k-l] = \sum_{l=0}^{L} h[l]s[k-l] \tag{2.8}$$

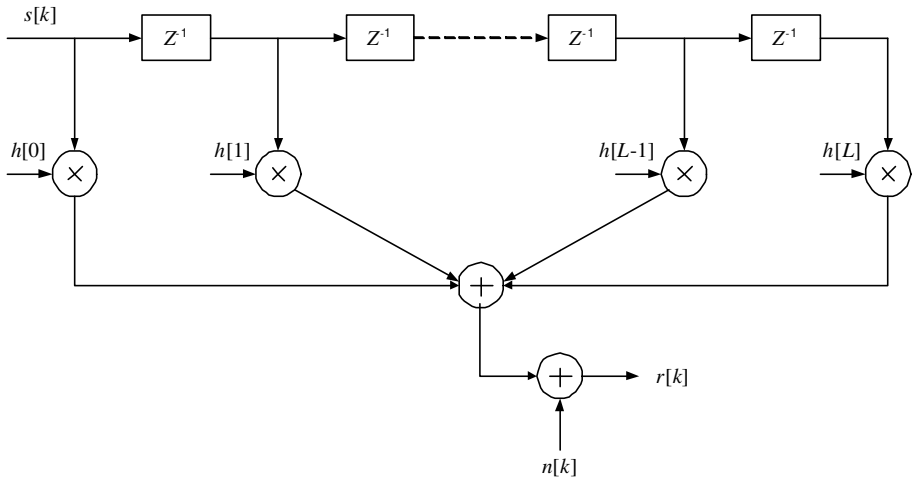

Figure 2.2: Discrete Time TDL Channel Model

where L is the channel order, which is defined as the maximum sample duration of the the channel delay.

2.3 Orthogonal Frequency Division Multiplexing

2.3.1 Background

In classical data systems in which more data rate was sought by exploiting the frequency domain, parallel transmissions were achieved by dividing the total signal frequency band into N_c non-overlapping frequency subchannels. This technique is referred to as Frequency Division Multiplexing (FDM). In this technique, each subchannel or subcarrier is modulated with a separate symbol and then the N_c subchannels are frequency multiplexed. Spectral overlap is avoided by putting enough guard space between adjacent subchannels. In this way Inter Carrier Interference (ICI) is eliminated. This method, however, leads to a very inefficient use of the available spectrum. A more efficient use of bandwidth can by obtained with parallel

transmissions if the spectra of the individual subchannels are permitted to partly overlap. This requires that specific orthogonality constraints are imposed to facilitate separation of the subchannels at the receiver.

Orthogonal Frequency Division Multiplexing (OFDM) is an example of a multicarrier technique that operates with specific orthogonality constraints between the subcarriers. With OFDM transmission, a high-rate serial data stream is split up into a set of low-rate sub-streams, each of which is modulated by a separate subcarrier. These multiple subcarriers overlap in the frequency domain, but do not cause ICI due to the orthogonal nature to the modulation. Hence, the orthogonal nature of the OFDM makes it very attractive by reducing the guard band required by normal FDM transmissions, greatly improving the spectral efficiency (see Figure 1.2 in Chapter 1). As a result, more and more systems that operate in the Gigahertz bands are based on OFDM, such as wireless LANs [58], [59], Digital Video Broadcasting (DVB) [60], and Digital Audio Broadcasting (DAB) [61].

2.3.2 Principles of OFDM

In an OFDM system, a block of N_c serial data symbols, each of duration T_s, is converted into a block of N_c parallel data symbols, each of duration $T = N_c T_s$. These N_c parallel data symbols modulate the N_c orthogonal subcarriers respectively. Consider a set of subcarrier frequencies $\{f_n\}$, where $0 \leq f_n \leq \frac{N_c-1}{T}$. Let one of the subcarrier signal be $\phi_{n_1}(t) = \exp\{j(2\pi f_{n_1} t) + \theta_{n_1}\}$ with the subcarrier frequency f_{n_1} and a random phase θ_{n_1}. Let another subcarrier signal be $\phi_{n_2}(t) = \exp\{j(2\pi f_{n_2} t + \theta_{n_2})\}$ with the subcarrier frequency f_{n_2} and a random phase θ_{n_2}. Then orthogonality

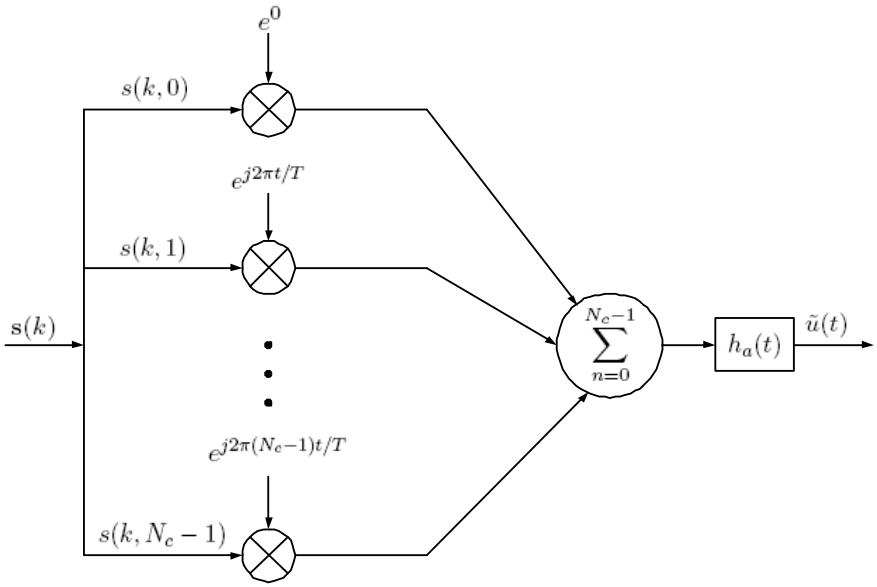

Figure 2.3: OFDM Modulation

in $[0,T]$ is achieved if

$$\begin{aligned} & \int_0^T \phi_{n_1}(t)\phi_{n_2}^*(t)dt = 0 \\ \Leftrightarrow \quad & \int_0^T e^{j(2\pi f_{n_1}t+\theta_{n_1})}e^{-j(2\pi f_{n_2}t+\theta_{n_2})}dt = 0 \\ \Leftrightarrow \quad & \frac{e^{j\left(2\pi(f_{n_1}-f_{n_2})T+(\theta_{n_1}-\theta_{n_2})\right)} - e^{j(\theta_{n_1}-\theta_{n_2})}}{j2\pi(f_{n_1}-f_{n_2})} = 0 \end{aligned} \tag{2.9}$$

When $2\pi(f_{n_1}-f_{n_2})T$ is a multiple of 2π, then Eqn.(2.9) will be true for any value of $\theta_{n_1}-\theta_{n_2}$. Thus, we choose the subcarrier frequencies separated by $1/T$ to guarantee the orthogonality with the presence of random phase offsets.

Hence the complex envelope of an OFDM signal is given by

$$\tilde{u}(t) = \sum_k h_a(t-kT_{total}) \sum_{n=0}^{N_c-1} s(k,n)e^{j2\pi(k-\frac{N_c-1}{2})(t-kT_{total})/T} \tag{2.10}$$

where $s(k,n)$ is element of the kth block of complex source symbols modulating the nth subcarrier, which are often chosen form a constellation such as QAM, PSK etc. [3], $e^{-j\pi(N_c-1)(t-kT_{total})/T}$ is the fixed frequency offset to make sure that the band

pass signal is centered about the center frequency, $h_a(t)$ is the pulse shaping function, and T_{total} is the symbol duration including the effective part of the OFDM symbol T, the guard interval (GI) T_g, and the window interval T_w for pulse shaping.

The guard interval, a cyclic prefix (CP), is a copy of the last part of the OFDM symbol, which is transmitted before the so-called "effective" part of the symbol. Its duration T_g is selected larger than the maximum excess delay of the radio channel. Therefore, the effective part of the received signal can be seen as the cyclic convolution of the transmitted OFDM symbol by the channel impulse response. This is attractive because the frequency selective channel is thus transferred to a set of parallel flat fading channels. However, the transmitted energy increases with the length of CP. The Signal-to-Noise Ratio (SNR) loss due to the insertion of CP is given by

$$\mathrm{SNR}_{loss} = -10\log_{10}\left(1 - \frac{T_g}{T + T_g}\right) \tag{2.11}$$

Also, the bandwidth efficiency is decreased to $\frac{T}{T+T_g}$ of that without CP. Hence, the CP should not be made longer than strictly necessary. Fortunately, when making T_g equal to the length of the channel impulse response, the relative length $\frac{T_g}{T+T_g}$ is typically small, so that the ISI free transmission motivates the small SNR loss.

To avoid out of band radiation, the pulse shaping (or equivalently called windowing) technique is deployed to fasten the roll off of side lobes. Raised cosine window is a wildly chosen option [34]. Figure 2.4 depicts schematically the implementation of the pulse shaping in an OFDM symbol [6]. The power spectrum of an OFDM signals with 48 subcarrier and different windowing length is simulated in Figure 2.5. The effective OFDM symbol length is $T = 4.8$ seconds, the GI length is $T_g = 1.6$

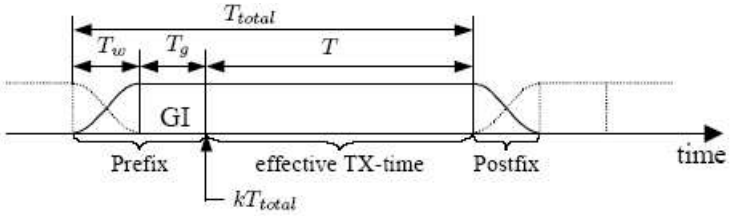

Figure 2.4: Cyclic extension and pulse shaping of the OFDM symbol

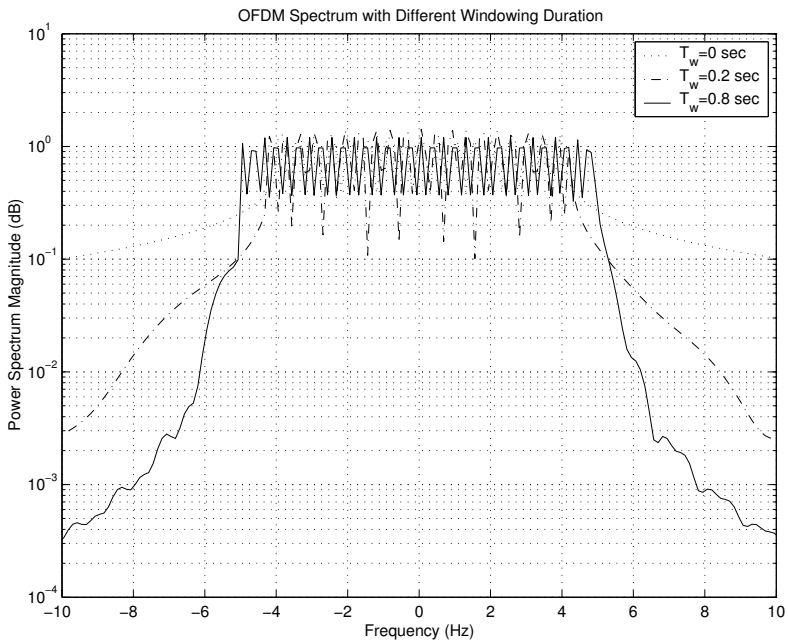

Figure 2.5: OFDM Power Spectrum with Different Window Length

seconds, and the windowing lengthes are set $T_w = 0$, 0.2, and 0.8 second respectively to compare the effecting of the windowing on OFDM power spectrums.

2.3.3 FFT Based OFDM and System Model

Although the continuous time system model in Eqn.(2.10) is conceptually simple and straightforward, they cannot be realized in its entirety in a digital system, especially a real time system, due to the computational cost. This bottleneck problem is

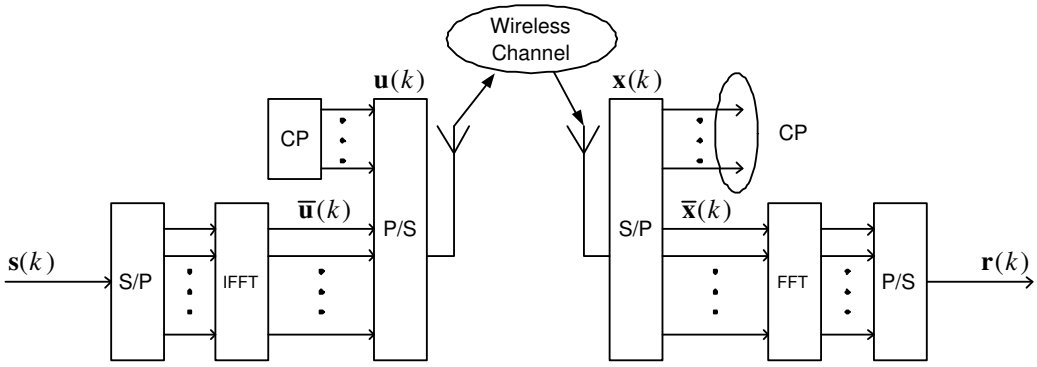

Figure 2.6: OFDM System Block Diagram

settled in 1971 by Weinstain and Ebert [25], who used the discrete Fourier transform (DFT) to perform baseband modulation and demodulation. A block diagram of the equivalent complex-valued baseband core of an OFDM system is depicted in Figure 2.6. The two main principles incorporated are:

- The inverse fast Fourier transform (IFFT) and the fast Fourier transform (FFT) are used, respectively, modulating and demodulating the data constellations on the orthogonal subcarriers [56]. The input of the IFFT, N_c data constellation points $\{s_i(k)\}$ are present, where N_c is the number of IFFT points, i is an index on the subcarrier, and k is an index on the OFDM symbol. These constellations can be taken according to any PSK or QAM signaling set.

  Usually, M is taken as an integer power of two, enabling the application of the highly efficient IFFT and FFT algorithm for modulation and demodulation.

- The introduction of a cyclic prefix as a guard interval (GI), whose length should exceed the maximum excess delay of the multipath propagation channel [6]. Due to the cyclic prefix, the transmitted signal becomes "periodic", and the effect of the time-dispersive multipath channel becomes equivalent to a

cyclic convolution, discarding the guard interval at the receiver. Due to the properties of the cyclic convolution, the effect of the multipath channel is limited to a point-wise multiplication of the transmitted data constellations by the channel transfer function, the Fourier transform of the channel impulse response, i.e., the subcarriers remain orthogonal [25], [26].

The equalization required for detecting the data constellation is an element-wise multiplication of the FFT output by the inverse of the estimated channel transfer function. For phase modulation schemes, multiplication by the complex conjugate of the channel estimate can do the equalization.

In the previous sections, we use the square brackets "$[\cdot]$" and the parentheses "$(\cdot)$" to represent the time index for the digital and analog systems respectively to distinguish these two systems. However, we will consider the digital system only in the rest of this thesis. Therefore, we replace the square brackets "$[\cdot]$" by the parentheses "$(\cdot)$" to make the notations in this thesis to be consistent with the popular literatures, e.g., Eqn.(2.8) will be rewritten as

$$r(l) = \sum_{-\infty}^{\infty} h(l)s(k-l) = \sum_{l=0}^{L} h(l)s(k-l) \tag{2.12}$$

Consider the source data sequence $s(0), s(1), s(2), \cdots$, which are often chosen form a constellation such as QAM, PSK etc. [3]. In the OFDM system where N_c subcarriers are used, every N_c data symbols are collected as a group for further processing. Define the kth block of source data symbols as

$$\mathbf{s}(k) \triangleq \begin{bmatrix} s(k,0) & \cdots & s(k,N_c-1) \end{bmatrix}^T \tag{2.13}$$

where

$$s(k,n) \triangleq s(kN_c + n), \quad n = 0, 1, \cdots, N_c - 1 \tag{2.14}$$

which is the element of the kth block of complex source symbols modulating the nth subcarrier. Since the elements of $\mathbf{s}(k)$ are amplitudes modulating subcarriers, they are often interpreted as frequency domain values. $\mathbf{s}(k)$ is therefore called frequency domain OFDM symbol or simply OFDM symbol [25]. As shown in Figure 2.6, the source data is serial-to-parallel (S/P) converted, and fed to an N_c-point FFT unit. The corresponding k^{th} time-domain vector is given by

$$\bar{\mathbf{u}}(k) = \frac{1}{\sqrt{N_c}} \text{IFFT}\left\{\mathbf{s}(k)\right\} = \mathbf{F}_{N_c}^H \mathbf{s}(k) \tag{2.15}$$

where $\mathbf{F}_{N_c}$ is the $N_c \times N_c$ FFT matrix which can be written as

$$\mathbf{F}_{N_c} = \frac{1}{\sqrt{N_c}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j2\pi\frac{1}{N_c}} & \cdots & e^{-j2\pi\frac{N_c-1}{N_c}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j2\pi\frac{N_c-1}{N_c}} & \cdots & e^{-j2\pi\frac{(N_c-1)^2}{N_c}} \end{bmatrix} \tag{2.16}$$

The resulting vector $\bar{\mathbf{u}}(k)$ is then appended with a Cyclic Prefix (CP) of length N_g, which is the copy of the last N_g elements of $\bar{\mathbf{u}}(k)$, resulting in a size $N = N_c + N_g$ signal vector

$$\mathbf{u}(k) \triangleq \mathbf{T}_{cp}\bar{\mathbf{u}}(k) = \mathbf{F}_{cp}\mathbf{s}(k) = \begin{bmatrix} u(k,0) & \cdots & u(k,N-1) \end{bmatrix}^T \tag{2.17}$$

where $\mathbf{T}_{cp}$ is the corresponding CP insertion matrix, which is defined as follows

$$\mathbf{T}_{cp} \triangleq \begin{bmatrix} \mathbf{0}_{N_g \times (N_c - N_g)} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_c} & \end{bmatrix} \tag{2.18}$$

where $\mathbf{I}_{N_c}$ and $\mathbf{I}_{N_g}$ are the $N_c \times N_c$ and $N_g \times N_g$ identity matrices respectively, and $\mathbf{F}_{cp}$ is constructed by appending the last N_g rows of $\mathbf{F}_{N_c}^H$ to its beginning, i.e.,

$$\mathbf{F}_{cp} \triangleq \mathbf{T}_{cp}\mathbf{F}_{N_c}^H \tag{2.19}$$

Then, the resulted signal is parallel-to-serial (P/S) converted, resulting in a serial data data sequence $u(0), u(1), u(2), \cdots$, where

$$u(kN+n) \triangleq u(k,n), \text{ for } n = 0, \cdots, N-1 \tag{2.20}$$

This serial data sequence is called the "OFDM chip sequence", and is pulse shaped to the corresponding continuous time signal, i.e.,

$$u_c(t) = \sum_{n=-\infty}^{\infty} u(n)\varphi_c(t-nT) \tag{2.21}$$

where T is the chip period and $\varphi_c(t)$ is the chip pulse. The transmitted waveform $u_c(t)$ propagates through a dispersive channel $h_c(t)$ and is filtered by the receive filter $\bar{\varphi}_c(t)$. Define the overall impulse response of the transmitter, channel, and receiver as

$$h_g(t) = \varphi_c(t) * h_c(t) * \bar{\varphi}_c(t) \tag{2.22}$$

Then, the received signal $x_c(t)$ sampled at the chip rate can be written as

$$\begin{aligned} x(n) &\triangleq x_c(t)|_{t=nT} \\ &= \int_{-\infty}^{\infty} h(\tau)u_c(nT-\tau)d\tau + v(nT) \\ &= \sum_{l=0}^{L} h(l)u(n-l) + v(n) \end{aligned} \tag{2.23}$$

where

$$h(l) \triangleq \int_{-\infty}^{\infty} h_g(\tau)\text{sinc}(\frac{lT-\tau}{T})d\tau \tag{2.24}$$

which is the equivalent discrete time channel impulse response, and $v(n) = v(nT)$ is the filtered additive white Gaussian noise (AWGN) with zero-mean and variance σ_v^2.

To recover the signal by FFT demodulation, every N received symbols are collected to form a group. Define the kth block of received signal as

$$\mathbf{x}(k) \triangleq \begin{bmatrix} x(k,0) & \cdots & x(k,N-1) \end{bmatrix}^T \tag{2.25}$$

where

$$x(k,n) \triangleq x(kN+n), \quad n = 0, \cdots N-1 \tag{2.26}$$

Based on the assumption that the length of CP is greater than or equal to the channel order L, and define the channel vector as

$$\mathbf{h} \triangleq [h(0), h(1), \cdots, h(L)]^T \tag{2.27}$$

Then the relations of the kth received and transmitted OFDM symbol can be derived as follows

$$\begin{aligned} \mathbf{x}(k) &= \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k-1) + \mathbf{v}(k) \\ &= \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{cp}\mathbf{s}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{cp}\mathbf{s}(k-1) + \mathbf{v}(k) \end{aligned} \tag{2.28}$$

where $\mathbf{v}(k)$ is the kth block of discrete AWGN noise. $\dot{\mathcal{T}}_N(\mathbf{h})$ and $\ddot{\mathcal{T}}_N(\mathbf{h})$ are the $N \times N$ lower and upper triangular Toeplitz matrices constructed by $\mathbf{h}$ respectively, i.e.

$$\dot{\mathcal{T}}_N(\mathbf{h}) \triangleq \begin{bmatrix} h(0) & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ h(L) & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h(L) & \cdots & h(0) \end{bmatrix} \tag{2.29}$$

$$\ddot{\mathcal{T}}_N(\mathbf{h}) \triangleq \begin{bmatrix} 0 & \cdots & 0 & h(L) & \cdots & h(1) \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & h(L) \\ \vdots & & & \ddots & & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix} \tag{2.30}$$

As shown in Figure 2.6, the received serial sequence is then S/P converted, and the CP (i.e. the first N_g elements of $\mathbf{x}(k)$) is removed. The remaining samples are collected into the N_c-dimensional vector $\bar{\mathbf{x}}(k)$:

$$\begin{aligned} \bar{\mathbf{x}}(k) &= \mathbf{R}_{cp}\mathbf{x}(k) \\ &= \mathbf{R}_{cp}\dot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{cp}\mathbf{s}(k) + \mathbf{R}_{cp}\mathbf{v}(k) \\ &= \mathcal{C}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H\mathbf{s}(k) + \bar{\mathbf{v}}(k) \end{aligned} \tag{2.31}$$

where $\mathbf{R}_{cp} = [\mathbf{0}_{N_c \times N_g}, \mathbf{I}_{N_c}]$, $\bar{\mathbf{v}}(k)$ is the $N_c \times 1$ truncated noise vector, and $\mathcal{C}_{N_c}(\mathbf{h})$ is the $N_c \times N_c$ circulant channel matrix constructed by $\mathbf{h}$, i.e.

$$\mathcal{C}_{N_c}(\mathbf{h}) \triangleq \begin{bmatrix} h(0) & 0 & \cdots & 0 & h(L) & \cdots & h(1) \\ h(1) & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & h(L) \\ h(L) & & \ddots & \ddots & \ddots & & 0 \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h(L) & \cdots & h(1) & h(0) \end{bmatrix} \tag{2.32}$$

According to the properties of the circulant matrix, $\mathbf{F}_{N_c}\mathcal{C}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H$ is a $N_c \times N_c$ diagonal matrix [62], i.e.,

$$\mathbf{F}_{N_c}\mathcal{C}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H = \mathcal{D}(\mathbf{H}) \tag{2.33}$$

where $\mathcal{D}(\mathbf{H})$ denotes the diagonal matrix with the elements of vector $\mathbf{H}$ along its diagonal, and the $N_c \times 1$ vector $\mathbf{H}$ is N_c-point FFT output of the channel vector $\mathbf{h}$, i.e.

$$\mathbf{H} \triangleq \begin{bmatrix} H(0) & H(1) & \cdots & H(N_c - 1) \end{bmatrix}^T = \mathrm{FFT}\{\mathbf{h}\} \tag{2.34}$$

Hence, the FFT output of the received symbols can be expressed as

$$\begin{aligned} \mathbf{y}(k) &= \mathbf{F}_{N_c}\bar{\mathbf{x}}(k) \\ &= \mathbf{F}_{N_c}\mathcal{C}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H\mathbf{s}(k) + \mathbf{F}_{N_c}\bar{\mathbf{v}}(k) \\ &= \mathcal{D}(\mathbf{H})\mathbf{s}(k) + \mathbf{n}(k) \end{aligned} \tag{2.35}$$

where $\mathbf{n}(k) = \mathbf{F}_{N_c}\bar{\mathbf{v}}(k)$.

Finally, the received OFDM symbols are P/S converted to the serial vector $y(n)$, where

$$y(kN_c + n) \triangleq y(k, n) \quad \text{for } n = 0, \cdots, N_c - 1 \tag{2.36}$$

and $y(k, n)$ is the nth element of the kth demodulated OFDM symbol $\mathbf{y}(k)$. Thus, by using the CP, the FIR convolutive channel is converted to parallel flat-fading sub-channels independent with each other.

2.3.4 Zero Padded-OFDM

Padding zero samples to the end of each modulated OFDM symbol is an alternative way to suppress the ISI caused by the FIR channel. This is referred to as the Zero Padded OFDM (ZP- OFDM). In this subsection, we briefly introduce the system model of ZP-OFDM as a comparison to the CP-OFDM. Note that through out this thesis, without any further annunciation, the term *OFDM* means CP-OFDM, and the zero padded-OFDM is denoted as ZP-OFDM.

In ZP-OFDM, the mathematical model for the source signal $\mathbf{s}(k)$ and the IFFT modulated OFDM symbol $\bar{\mathbf{u}}(k)$ are the same as in CP-OFDM system, which are described in Eqn.(2.13) and Eqn.(2.15) respectively. However, the modulated OFDM symbol $\bar{\mathbf{u}}(k)$ is padded with zero sequence of length $N_g \geq L$, resulting in a length $N = N_c + N_g$ vector as

$$\mathbf{u}(k) = \mathbf{F}_{zp}\mathbf{s}(k) \tag{2.37}$$

where $\mathbf{F}_{zp} \triangleq [\mathbf{F}_{N_c}, \mathbf{0}_{N_c \times N_g}]^H$, $\mathbf{F}_{N_c}$ is the FFT matrix defined by Eqn.(2.16), and $\mathbf{0}_{N_c \times N_g}$ is the $N_c \times N_g$ zero matrix.

At the receiver side, due to the effect of FIR channel and the AWGN noise, the kth block of received signal can be written as follows

$$\begin{aligned}
\mathbf{x}(k) &= \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k-1) + \mathbf{v}(k) \\
&= \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{zp}\mathbf{s}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{zp}\mathbf{s}(k-1) + \mathbf{v}(k) \\
&= \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{zp}\mathbf{s}(k) + \mathbf{v}(k) \\
&= \mathcal{T}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H\mathbf{s}(k) + \mathbf{v}(k) \qquad (2.38)
\end{aligned}$$

where $\dot{\mathcal{T}}_N(\mathbf{h})$ and $\ddot{\mathcal{T}}_N(\mathbf{h})$ are lower and upper triangular Toeplitz matrices defined by Eqn.(2.29) and Eqn.(2.30) respectively. $\mathcal{T}_{N_c}(\mathbf{h})$ is the Toeplitz matrix defined as

$$\mathcal{T}_{N_c}(\mathbf{h}) \triangleq \begin{bmatrix}
h(0) & & & & \\
\vdots & h(0) & & \mathbf{0} & \\
h(L) & \vdots & \ddots & & \\
 & h(L) & & \ddots & \\
 & & \ddots & & h(0) \\
 & \mathbf{0} & & & h(L) \\
0 & \cdots & & & 0 \\
\vdots & & & & \vdots \\
0 & \cdots & & & 0
\end{bmatrix} \qquad (2.39)$$

Note that the elements of the last $N_g - L$ row of $\mathcal{T}_{N_c}(\mathbf{h})$ are all zeros.

To construct the circulant channel matrix as in Eqn.(2.32), we add the last N_g elements of $\mathbf{x}(k)$ to its first N_g elements, through the matrix $\mathbf{R}_{zp} \triangleq [\mathbf{I}_{N_c}\ \mathbf{I}_{zp}]$, where $\mathbf{I}_{zp}$ denotes the matrix containing the first N_g columns of the $N_c \times N_c$ identity matrix

$\mathbf{I}_{N_c}$. Then we obtain:

$$\begin{aligned}\bar{\mathbf{x}}(k) &= \mathbf{R}_{zp}\mathbf{x}(k) = \mathbf{R}_{cp}\mathcal{T}(\mathbf{h})\mathbf{F}_{N_c}^H\mathbf{s}(k) + \mathbf{R}_{zp}\mathbf{v}(k) \\ &= \mathcal{C}_{N_c}(\mathbf{h})\mathbf{F}_{N_c}^H\mathbf{s}(k) + \bar{\mathbf{v}}(k) \end{aligned} \tag{2.40}$$

The notations in Eqn.(2.40) is the same as in Eqn.(2.31), except that $\bar{\mathbf{v}}(k)$ is slightly different, obtained by adding the last N_g elements of $\mathbf{v}(k)$ to its first N_g elements. This leads to the result that the noise in ZP-OFDM is cyclostationary, and noise power level is slightly higher than the CP-OFDM in a same transmission environment.

Once $\bar{\mathbf{x}}(k)$ is obtained through Eqn.(2.40), then the consequential steps are the same as CP-OFDM.

2.4 MIMO-OFDM System Model

2.4.1 Basic Concept

In recent years, key techniques related to the combination of MIMO with OFDM have been studied for larger capacity, higher data rate and better performance in wireless communications [63]- [68].

Figure 2.7 shows a simplified schematic representation of a MIMO-OFDM transmitter. The source bitstream is encoded by a forward error correction (FEC) encoder. After that, the coded bitstream is mapped to a constellation by the digital modulator, and encoded by a MIMO encoder. Then each of the parallel output symbol streams corresponding to a certain transmit antenna follows the same transmission process. Then the symbol sequence in the frequency domain is modulated

by IFFT to an OFDM symbol sequence. A cycle prefix (CP) is attached to every OFDM symbol to mitigate the effect of channel delay spread, and a preamble is inserted in every slot for timing. Finally, the constructed data frame is transferred to RF components for transmission.

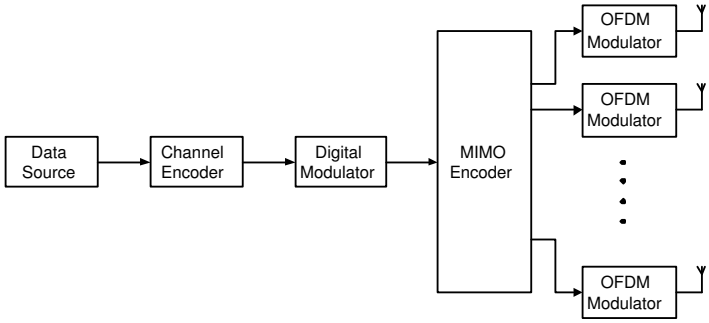

Figure 2.7: A Simplified Schematic Representation of a MIMO-OFDM Transmitter

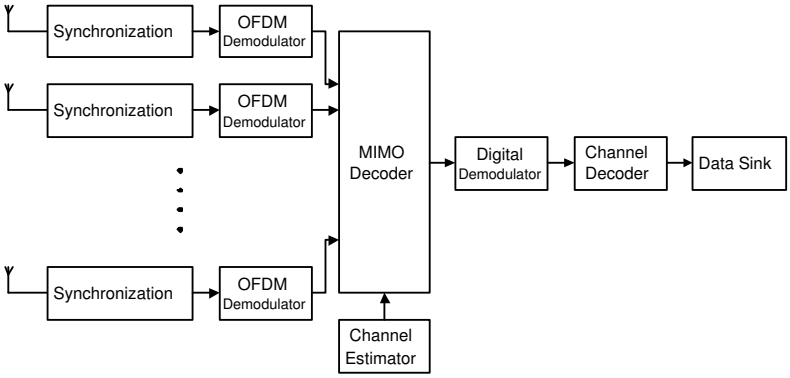

Figure 2.8: A Simplified Schematic Representation of a MIMO-OFDM Receiver

Figure 2.8 shows a simplified schematic representation of a MIMO-OFDM receiver. The received symbol stream from RF components over the receive antennae are first synchronized, including coarse frequency synchronization and timing aided by the preamble. After that, the preambles and CP are extracted from the received symbol stream, and the remaining OFDM symbol is demodulated by FFT. Then

frequency synchronization and channel estimation are carried out for the following processing. The estimated channel matrix aids the MIMO decoder in decoding the refined OFDM symbols. The estimated transmit symbols are then demodulated and decoded. Finally, the decoded source bitstreams are transmitted to the sink.

Finally, note that OFDM has as advantage that it introduces a certain amount of parallelism by means of its N_c subcarriers. This fact can be exploited by MIMO-OFDM. Namely, if MIMO detection is performed per subcarrier, then a given detector is allowed to work N_c times slower than the MIMO detector of an equivalent single carrier system with comparable data rate. Although in the case of MIMO-OFDM N_c of such detectors a required, they can work in parallel, which might ease the implementation.

2.4.2 MIMO-OFDM System Model

In this subsection, a signal model is introduced for a MIMO-OFDM system in which the relation between the transmitted and received MIMO-OFDM symbols is captured in matrix form. With this concise matrix notation we mathematically show that the signal model per subcarrier equals the narrowband signal model.

Consider a MIMO-OFDM system with M_t transmit (TX) and M_r receive (RX) antennae respectively, as shown in Figure 2.9. Similar to the SISO-OFDM system model described in Section 2.3, the OFDM system utilize a maximum of N_c subcarriers per antenna to deal with the frequency selectivity of the channel. To combat the ISI, a guard interval of N_g samples is added per OFDM symbol. Thus an OFDM including the guard interval consists of $N = N_c + N_g$ complex symbols. Based on

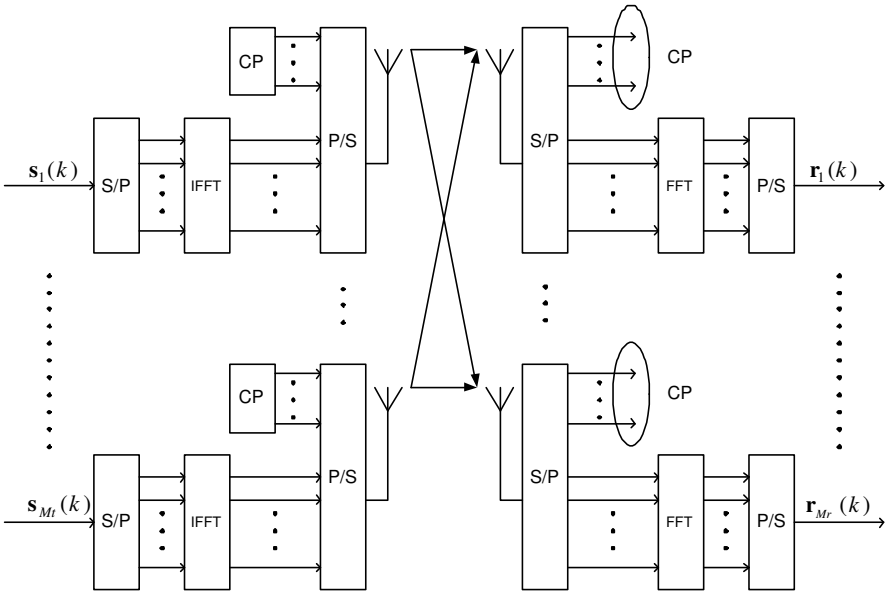

Figure 2.9: Block Diagram of a MIMO-OFDM

this, the collection of complex symbols to be sent on the kth MIMO-OFDM symbol can be denoted by the $N_c M_t \times 1$ vector $\mathbf{s}(k)$ as follows

$$\mathbf{s}(k) = \begin{bmatrix} \mathbf{s}(k,0) \\ \vdots \\ \mathbf{s}(k, N_c - 1) \end{bmatrix} \tag{2.41}$$

where the $M_t \times 1$ vector

$$\mathbf{s}(k,n) \triangleq \begin{bmatrix} s_1(k,n), & s_2(k,n), & \cdots, & s_{M_t}(k,n) \end{bmatrix}^T \tag{2.42}$$

represents the chips of the kth MIMO-OFDM symbol to be transmitted through the nth subchannel, and

$$s_i(k,n) \triangleq s_i(kN_c + n), \quad (n = 0, \cdots, N_c, \ i = 1, \cdots, M_t) \tag{2.43}$$

denotes the element of the kth block of the source signal carried by the nth subcarrier and transmitted by the ith transmit antenna.

Firstly, the IFFT is applied at each transmit antenna, which transforms the frequency domain vector $\mathbf{s}(k)$ into the time domain. Hence, the output of the IFFT modulation is denoted by

$$\bar{\mathbf{u}}(k) = (\mathbf{F}_{N_c}^H \otimes \mathbf{I}_{M_t})\mathbf{s}(k) \tag{2.44}$$

where $\otimes$ represents the Kronecker product, $\mathbf{F}_{N_c}$ is the FFT matrix defined by Eqn.(2.16), and $\mathbf{I}_{M_t}$ denotes the $M_t \times M_t$ identity matrix.

Secondly, the CP is added and P/S converted. This is done by taking the last $M_t N_c$ elements of $\bar{\mathbf{u}}(k)$ and stacking them on top of $\bar{\mathbf{u}}(k)$ to produce the vector $\mathbf{u}(k)$. In matrix notation this can be written as

$$\mathbf{u}(k) = (\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t})\, \bar{\mathbf{u}}(k) = (\mathbf{F}_{cp} \otimes \mathbf{I}_{M_t})\, \mathbf{s}(k) \tag{2.45}$$

where the CP insertion matrix $\mathbf{T}_{cp}$ and the OFDM modulation matrix $\mathbf{F}_{cp}$ are respectively defined by Eqn.(2.18) and Eqn.(2.19) in the previous section. Meanwhile, we define the modulated OFDM symbols as

$$\mathbf{u}(k) \triangleq \begin{bmatrix} \mathbf{u}^T(k,0) & \cdots & \mathbf{u}^T(k,N-1) \end{bmatrix}^T \tag{2.46}$$

where

$$\mathbf{u}(k,n) \triangleq \begin{bmatrix} u_1(k,n) & \cdots & u_{M_t}(k,n) \end{bmatrix}^T \tag{2.47}$$

These parallel data blocks are then P/S converted to M_t serial data sequences $u_i(n)$, $(i = 1, \cdots M_t,\ n = 0, 1, 2, \cdots)$ which are to be transmitted by the M_t transmit antennae respectively, where

$$u_i(kN+n) \triangleq u_i(k,n), \quad n = 0, \cdots, N-1, \quad i = 1, \cdots, M_t \tag{2.48}$$

Finally, the signal is transmitted over the dispersive channel. Define the time domain MIMO channel impulse response matrix as

$$\mathbf{h} \triangleq \begin{bmatrix} \mathbf{h}^T(0) & \mathbf{h}^T(1) & \cdots & \mathbf{h}^T(L) \end{bmatrix}^T \tag{2.49}$$

where

$$\mathbf{h}(l) = \begin{bmatrix} h_{1,1}(l) & \cdots & h_{1,M_t}(l) \\ \vdots & \ddots & \vdots \\ h_{M_r,1}(l) & \cdots & h_{M_r,M_t}(l) \end{bmatrix} \text{for} \quad l = 0, 1, \cdots L \tag{2.50}$$

while $h_{j,i}(l)$ represents the lth tap of tap coefficient of the FIR channel between the ith transmit antenna and the jth receive antenna.

Similar to the case of SISO-OFDM, at the MIMO-OFDM receiver side, every N data symbols received by each antenna are collected to form a group. The kth group of received data group, or namely k^{th} received OFDM symbol, can expressed as follows

$$\mathbf{x}(k) = \dot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\mathbf{u}(k-1) + \mathbf{v}(k) \tag{2.51}$$

where

$$\mathbf{x}(k) \triangleq \begin{bmatrix} \mathbf{x}^T(k,0) & \cdots & \mathbf{x}^T(k,N-1) \end{bmatrix}^T \tag{2.52}$$

$$\mathbf{x}(k,n) = \begin{bmatrix} x_1(k,n) & \cdots & x_{M_r}(k,n) \end{bmatrix}^T \tag{2.53}$$

$$x_j(k,n) = x_j(kN+n), \quad j = 1, \cdots, M_r \tag{2.54}$$

while $\dot{\mathcal{T}}_N(\mathbf{h})$ and $\ddot{\mathcal{T}}_N(\mathbf{h})$ are $NM_r \times NM_t$ block lower and upper Toeplitz triangular

matrices constructed by the channel matrix $\mathbf{h}$ respectively, i.e.

$$\dot{\mathcal{T}}_N(\mathbf{h}) \triangleq \begin{bmatrix} \mathbf{h}(0) & \mathbf{0}_{M_r\times M_t} & \cdots & \cdots & \cdots & \mathbf{0}_{M_r\times M_t} \\ \vdots & \ddots & \ddots & & & \vdots \\ \mathbf{h}(L) & \ddots & \ddots & \ddots & & \vdots \\ \mathbf{0}_{M_r\times M_t} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0}_{M_r\times M_t} \\ \mathbf{0}_{M_r\times M_t} & \cdots & \mathbf{0}_{M_r\times M_t} & \mathbf{h}(L) & \cdots & \mathbf{h}(0) \end{bmatrix} \tag{2.55}$$

$$\ddot{\mathcal{T}}_N(\mathbf{h}) \triangleq \begin{bmatrix} \mathbf{0}_{M_r\times M_t} & \cdots & \mathbf{0}_{M_r\times M_t} & \mathbf{h}(L) & \cdots & \mathbf{h}(1) \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & \ddots & \mathbf{h}(L) \\ \vdots & & & \ddots & & \mathbf{0}_{M_r\times M_t} \\ \vdots & & & & \ddots & \vdots \\ \mathbf{0}_{M_r\times M_t} & \cdots & \cdots & \cdots & \cdots & \mathbf{0}_{M_r\times M_t} \end{bmatrix} \tag{2.56}$$

where $\mathbf{0}_{M_r\times M_t}$ denotes the $M_r \times M_t$ zero matrix, and $\mathbf{v}(k)$ is the AWGN noise vector defined by

$$\mathbf{v}(k) \triangleq [\mathbf{v}^T(k,0)\cdots\mathbf{v}^T(k,N_c-1)]^T \tag{2.57}$$

where

$$\mathbf{v}(k,n) \triangleq [v_1(k,n)\cdots v_{M_r}(k,n)]^T \tag{2.58}$$

where $v_j(k,n) \triangleq v_j(kN+n)$; $(n = 0,\cdots,N-1)$ is the nth element of the kth block of AGWN noise sequence according to the jth receive antenna.

At each receive antenna, the received signal sequence is S/P converted, the CP is removed by discarding the first N_gM_r samples of $\mathbf{x}(k)$, the FFT demodulation is

performed to the resulted vector, and the resulted signal is P/S converted in the end. Together, this results in

$$\begin{aligned}\mathbf{y}(k) &= (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\mathbf{R}_{cp}\mathbf{x}(k) \\ &= (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r}) \left(\begin{bmatrix} \mathbf{0}_{N_c \times N_g} & \mathbf{I}_{N_c} \end{bmatrix} \otimes \mathbf{I}_{M_r} \right) \mathbf{x}(k) \end{aligned} \tag{2.59}$$

where $\mathbf{R}_{cp}$ and $\mathbf{y}(k)$ are respectively defined as

$$\mathbf{R}_{cp} \triangleq \begin{bmatrix} \mathbf{0}_{N_c \times N_g} & \mathbf{I}_{N_c} \end{bmatrix} \tag{2.60}$$

$$\mathbf{y}(k) \triangleq \begin{bmatrix} \mathbf{y}(k,0) \\ \vdots \\ \mathbf{y}(k, N_c - 1) \end{bmatrix} \tag{2.61}$$

where

$$\mathbf{y}(k,n) \triangleq [y_1(k,n) \cdots y_{M_r}(k,n)]^T \tag{2.62}$$

and

$$y_j(k,n) \triangleq y_j(kN_c + n), \quad n = 0, \cdots, N_c - 1 \tag{2.63}$$

which is the nth element of the kth block of demodulated signal associated with the jth receive antenna.

Combining all above steps and assuming that no ISI occurs on a MIMO-OFDM symbol basis (i.e., $L \leqslant N_g$), this leads to the following relation between $\mathbf{s}(k)$ and $\mathbf{y}(k)$:

$$\begin{aligned}\mathbf{y}(k) &= (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\mathbf{R}_{cp}\dot{\mathcal{T}}_N(\mathbf{h})\mathbf{F}_{cp}\mathbf{s}(k) + (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\mathbf{R}_{cp}\mathbf{v}(k) \\ &= (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\mathcal{C}_{N_c}(\mathbf{h})(\mathbf{F}_{N_c}^H \otimes \mathbf{I}_{M_t})\mathbf{s}(k) + \mathbf{n}(k) \end{aligned} \tag{2.64}$$

where $\mathcal{C}_{N_c}(\mathbf{h})$ is an $N_cM_r \times N_cM_t$ block circulant matrix constructed by the time domain channel matrix $\mathbf{h}$, i.e.

$$\mathcal{C}_{N_c}(\mathbf{h}) \triangleq \begin{bmatrix} \mathbf{h}(0) & \mathbf{0}_{M_r\times M_t} & \cdots & \mathbf{0}_{M_r\times M_t} & \mathbf{h}(L) & \cdots & \mathbf{h}(1) \\ \mathbf{h}(1) & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \mathbf{h}(L) \\ \mathbf{h}(L) & & \ddots & \ddots & \ddots & & \mathbf{0}_{M_r\times M_t} \\ \mathbf{0}_{M_r\times M_t} & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \mathbf{0}_{M_r\times M_t} \\ \mathbf{0}_{M_r\times M_t} & \cdots & \mathbf{0}_{M_r\times M_t} & \mathbf{h}(L) & \cdots & \mathbf{h}(1) & \mathbf{h}(0) \end{bmatrix} \tag{2.65}$$

and $\mathbf{n}(k)$ represents the frequency domain noise, which is given by

$$\mathbf{n}(k) = (\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\,\mathbf{R}_{cp}\mathbf{v}(k) \tag{2.66}$$

According to the properties of the block circulant matrix, $\mathcal{C}_{N_c}(\mathbf{h})$ in Eqn.(2.64) can be block-diagonalized, i.e.

$$(\mathbf{F}_{N_c} \otimes \mathbf{I}_{M_r})\mathcal{C}_{N_c}(\mathbf{h})(\mathbf{F}_{N_c}^H \otimes \mathbf{I}_{M_t}) = \mathcal{D}(\mathbf{H}) \tag{2.67}$$

where $\mathcal{D}(\mathbf{H})$ is the block diagonal matrix constructed by $\mathbf{H}$, and $\mathbf{H}$ is the frequency domain channel matrix defined as follows

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}^T(0) & \mathbf{H}^T(1) & \cdots & \mathbf{H}^T(N_c-1) \end{bmatrix}^T \tag{2.68}$$

where

$$\mathbf{H}(n) = \sum_{l=0}^{L} \mathbf{h}(l)e^{-j2\pi\frac{nl}{N_c}} \quad \text{for} \quad n = 0, 1, \cdots, N_c - 1 \tag{2.69}$$

Hence, (2.64) can be rewritten as

$$\mathbf{y}(k) \;=\; \mathcal{D}(\mathbf{H})\mathbf{s}(k) + \mathbf{n}(k) \tag{2.70}$$

and for the nth subcarrier we may write

$$\mathbf{y}(k,n) = \mathbf{H}(n)\mathbf{s}(k,n) + \mathbf{n}(k,n) \tag{2.71}$$

2.4.3 Zero Padded MIMO-OFDM

In this subsection, we briefly introduce the system model of zero padded-MIMO-OFDM (ZP-MIMO-OFDM) as the comparisons. Similar to the SISO case, the modulated MIMO-OFDM symbol with ZP can be written as follows

$$\mathbf{u}(k) = (\mathbf{F}_{zp} \otimes \mathbf{I}_{M_t})\mathbf{s}(k) \tag{2.72}$$

where $\mathbf{F}_{zp}$ and $\mathbf{s}(k)$ are defined by Eqn.(2.37) and Eqn.(2.41) respectively. Under the assumption that the length of ZP is larger than the maximum channel order, it can be verified that the received ZP-MIMO-OFDM symbol can be written as

$$\mathbf{x}(k) = \mathcal{T}_{N_c}(\mathbf{h})(\mathbf{F}_{N_c}^H \otimes \mathbf{I}_{M_t})\mathbf{s}(k) + \mathbf{v}(k) \tag{2.73}$$

where

$$\mathcal{T}_{N_c}(\mathbf{h}) \triangleq \begin{bmatrix} \mathbf{h}(0) & & & & \\ \vdots & \mathbf{h}(0) & & \mathbf{0} & \\ \mathbf{h}(L) & \vdots & \ddots & & \\ & \mathbf{h}(L) & & \ddots & \\ & & \ddots & & \mathbf{h}(0) \\ & \mathbf{0} & & & \mathbf{h}(L) \\ 0 & \cdots & & & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & & & 0 \end{bmatrix} \tag{2.74}$$

By adding the last $N_g M_r$ elements of $\mathbf{x}(k)$ to its first $N_g M_r$ elements, and perform the FFT demodulation to the received symbols according to the M_r receive antennae respectively, then we get $\mathbf{r}(k)$ which is defined the same as Eqn.(2.64), with only the slight difference that the noise before FFT is circular stationary.

2.5 Summary

In this chapter, we derived the basic theory behind OFDM and its applications. The mathematical system models for both SISO and MIMO cases are introduced. The derivation of the OFDM system model has confirmed that data symbols can be transmitted independently over multipath fading radio channels. It has to be assumed, however, that the channels maximum excess delay is shorter than the guard interval. On the other hand, it is essential that the channel is known at the receiver in order to perfectly recover the transmitted signal. In the following chapters, we present some novel blind channel estimation methods for MIMO-OFDM systems.

Chapter 3

Subspace-Based Blind Channel Estimation for MIMO-OFDM Systems

3.1 Introduction

In the previous chapter, we introduced the mathematical model of the MIMO-OFDM systems. In this and the following chapters, we present some blind channel estimation algorithms for MIMO-OFDM systems. Although by properly designing the cyclic prefix (CP), successive OFDM symbols will not interfere and can be reliably recovered at the receiver end, the channel impulse responses between transmit and receive antennae are still required for coherent signal detection. The channels can be estimated by sending training sequences. However, it is shown that estimating MIMO channels can require a significant amount of training sequences [68].

Furthermore, training is infeasible for some certain communication systems [84, 85]. Thus, blind channel estimation for MIMO-OFDM systems has been studied extensively in recent years.

Among various known blind channel estimation algorithms, the so called subspace (SS) based algorithm is most attractive due to its special properties [4]. This algorithm was originally developed in [69] for single-input multi-output (SIMO) systems. The SS method has simple structure and achieves good performance, but it also has some requirements, which make it difficult to be applied in general MIMO-OFDM systems. Firstly, the receive antennae must be strictly more than the transmit antennae [71, 78], which may not be satisfied by many existing standards, e.g. the IEEE 802.11n [76] standard defines the 2×2 transmit and receive antennae pairs. Besides, in the case of SISO systems, the equal number of transmit and receive antennae is obviously used, which does not satisfy the requirement. Secondly, the precise knowledge of the channel order must be obtained, which is very difficult in practice. The channel order over-estimation may cause significant performance degradation.

Recently, some SS methods have been proposed for OFDM systems. A subspace based method for SISO-OFDM systems was developed by utilizing the redundancy introduced by the CP insertion [72]. In [73], a SS based method was proposed which can be applied for OFDM systems without CP. Both of these two methods are designed for SISO-OFDM systems. In [74], a zero padding (ZP) OFDM system was suggested. Instead of using the CP, consecutive zeros are padded at the end of each OFDM symbol, and the general SS method is applied to blindly estimate the

channel. However, this method is not suitable for most of the existing MIMO-OFDM system which use CP as the guard interval. Another way to apply the SS method in MIMO-OFDM systems is to precode the source signal before transmission. In [81], a SS method is proposed for MIMO-OFDM systems with the assistance of the properly designed redundant linear precoder. In some practical cases, the OFDM systems are not fully loaded, i.e., some of the subcarriers are set to zero without any information [6]. These subcarriers are referred to as virtual subcarriers (VCs). In fact, this can also be viewed as a special case of redundant precoding and can be used for blind channel estimation [83]. In [70], the authors unified and generalized the SISO-OFDM SS methods to the case of MIMO-OFDM systems with any number of transmit and receive antennae, by exploiting the redundancy induced by the CP and/or VCs. However, this method still suffers the problem caused by the channel order over-estimation, and the computational inefficiency caused by the singular-value decomposition (SVD) of the large-sized correlation matrix of the channel outputs.

In this chapter, we present a novel SS method for blind channel estimation of MIMO-OFDM systems. We study the inherent structure of the correlation matrices of the channel output and develop a new criterion function, minimizing which leads to the estimate of the time domain channel matrix. The proposed estimation method is capable for MIMO-OFDM systems driven by either white or colored sources. It does not requires the number of the receive antennae to be strictly larger than the number of the transmit antennae either. Furthermore, it even doesn't require the length of the CP to be greater than the maximum channel order. This property makes the proposed channel estimator to be attractive in the ill conditioned envi-

ronments. Unlike the SS method in [70], where multiple consecutive OFDM symbols are collected and observed, the proposed algorithm is more efficient because the size of observation window is fixed to be exactly one single OFDM symbol.

The remaining of this chapter is organized as follows. First, the statistical system model described in Chapter 2 is recalled in Section 2.2. Then a subspace-based blind channel estimation algorithm is proposed in Section 3.3. The identifiability, comparison with the existing algorithm and the asymptotic performance of the proposed algorithm are discussed in Section 3.4. The numerical results by computer experiments are presented in Section 3.5. Lastly, the summary of this chapter is given in Section 3.6.

3.2 System Model and Basic Assumptions

Without unnecessarily repeating the system modeling, we directly recall the baseband representation of the MIMO-OFDM systems described by Eqn.(2.51) as follows

$$\mathbf{x}(k) = \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k-1) + \mathbf{v}(k) \tag{3.1}$$

where $\mathbf{s}(k)$ is the kth block of the MIMO source signal before OFDM modulation, $\mathbf{x}(k)$ is the kth block of the received signal before removing the CP, $\mathbf{v}(k)$ is the AWGN noise, $\mathbf{h}$ is the time domain channel matrix, while $\dot{\mathcal{T}}_N(\mathbf{h})$ and $\ddot{\mathcal{T}}_N(\mathbf{h})$ are respectively the lower and upper block triangular Toeplitz matrices constructed from $\mathbf{h}$, and

$$\tilde{\mathbf{F}}_{cp} \triangleq \mathbf{F}_{cp} \otimes \mathbf{I}_{M_t} \tag{3.2}$$

where $\mathbf{F}_{cp}$ is the IFFT and CP adding matrix defined by Eqn.(2.19), and $\otimes$ denotes the Kronecker product.

In the remaining of this chapter, we adopt the following assumptions:

A1) All the sources are wide sense stationary, and spatially uncorrelated.

A2) Additive noise are spatially and temporally white noise, and they are statistically independent of the sources.

A3) $\dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}$ and $\ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}$ are tall matrices, i.e., $M_r N > M_t N_c$.

A4) There exists an $l \in [0, L]$ such that $\mathbf{h}(l)$ is of full column rank.

3.3 Subspace-Based Blind Channel Estimator

3.3.1 Second Order Statistics of the MIMO-OFDM Symbols

Consider the autocorrelation matrix of the source signal with block lag κ, which is defined as follows

$$\mathbf{R}_{\mathbf{s}}(\kappa) \triangleq E\{\mathbf{s}(k)\mathbf{s}^H(k+\kappa)\} \tag{3.3}$$

Consequently, the autocorrelation matrix of the received signal before removing the CP can be expressed as

$$\begin{aligned}
\mathbf{R}_{\mathbf{x}}(\kappa) &\triangleq E\{\mathbf{x}(k)\mathbf{x}^H(k+\kappa)\} \\
&= \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) \\
&\quad +\dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa+1)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) \\
&\quad +\delta(\kappa)\sigma_v^2\mathbf{I}_{NM_r}
\end{aligned} \tag{3.4}$$

where σ_v^2 is the noise power.

Next, we exploit the structures of the autocorrelation matrices of the received signals to estimate the MIMO channel. Construct the following two matrices from the autocorrelation matrices of the transmitted and received symbols respectively

$$\mathbb{R}_{\mathbf{x}}(\kappa) \triangleq \sum_{j=-\kappa}^{\kappa} \mathbf{R}_{\mathbf{x}}(j) \tag{3.5a}$$

$$\mathbb{R}_{\mathbf{s}}(\kappa) \triangleq \sum_{j=-\kappa}^{\kappa} \mathbf{R}_{\mathbf{s}}(j) \tag{3.5b}$$

and substitute them to Eqn.(3.4), we have

$$\begin{aligned}\mathbb{R}_{\mathbf{x}}(\kappa) &= \sum_{j=-\kappa}^{\kappa} \Big(\dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(j)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(j)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) \\ &\quad +\dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(j-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(j+1)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h})\Big) + \sigma_v^2\mathbf{I}_{NM_r} \\ &= \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbb{R}_{\mathbf{s}}(\kappa-1)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) + \Phi_\kappa + \sigma_v^2\mathbf{I}_{NM_r} \end{aligned} \tag{3.6}$$

where

$$\begin{aligned}\Phi_\kappa &= \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(-\kappa)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) \\ &\quad +\ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(\kappa+1)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) + \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{R}_{\mathbf{s}}(-\kappa-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) \\ &\quad +\ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\left[\mathbf{R}_{\mathbf{s}}(\kappa) + \mathbf{R}_{\mathbf{s}}(-\kappa)\right]\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) \end{aligned} \tag{3.7}$$

and

$$\mathcal{C}_N(\mathbf{h}) \triangleq \dot{\mathcal{T}}_N(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h}) \tag{3.8}$$

It should be noted that $\mathcal{C}_N(\mathbf{h})$ is an $NM_r \times NM_t$ block circulant matrix constructed from the time domain channel matrix $\mathbf{h}$ by the similar way as $\mathcal{C}_{N_c}(\mathbf{h})$, which is defined by Eqn.(2.65).

Without loss of generality, we can assume that the source OFDM symbols are correlated only when the lag $\kappa \leqslant \kappa_{\max}$ (where $\kappa_{\max} \geqslant 0$), or in other words, $\mathbf{R}_{\mathbf{s}}(k) =$

$\mathbf{R_s}(-k) = \mathbf{0}$ for $\kappa > \kappa_{\max}$. Under this assumption, we have $\Phi_{\kappa_{\max}+1} = \mathbf{0}$. Thus, by substituting $\kappa = \kappa_{\max} + 1$ into Eqn.(3.6), we have

$$\mathbb{R}_{\mathbf{x}} = \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbb{R}_{\mathbf{s}}\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) + \sigma_v^2\mathbf{I}_{NM_r} \tag{3.9}$$

where we let $\mathbb{R}_{\mathbf{x}} \triangleq \mathbb{R}_{\mathbf{x}}(\kappa_{\max} + 1)$, and $\mathbb{R}_{\mathbf{s}} \triangleq \mathbb{R}_{\mathbf{s}}(\kappa_{\max})$.

Lemma 3.3.1 For $M_r \geq M_t$, if there exists an $l \in [0, L]$ such that $\mathbf{h}(l)$ is of full column rank, then $\mathcal{C}_N(\mathbf{h})$ is of full column rank.

The proof of this lemma is obvious and is omitted for brevity. It should be noted that Lemma 3.3.1 only provides a sufficient condition. In most cases, signal propagation from each of the transmitters is most likely independent, and hence the full column rank is almost surely guaranteed. Under assumption A4), we presume that the full column rank condition for $\mathcal{C}_N(\mathbf{h})$ holds in the rest sections of this chapter.

3.3.2 Proposed Channel Estimation Algorithm

Under the assumptions A1) $\sim$ A4), $\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}$ is a tall matrix with full column rank, and $\mathbb{R}_{\mathbf{s}}$ is of full rank. Thus, we can apply the MUSIC algorithm which exploits the special structure of $\mathbb{R}_{\mathbf{x}}$ to estimate the channel. Specifically, express the SVD of $\mathbb{R}_{\mathbf{x}}$ as follows

$$\mathbb{R}_{\mathbf{x}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{3.10}$$

where

$$\mathbf{\Lambda}_s = diag[\lambda_0 \cdots \lambda_{N_cM_t-1}] \tag{3.11a}$$

$$\mathbf{\Lambda}_n = diag[\lambda_{N_cM_t} \cdots \lambda_{NM_r-1}] \tag{3.11b}$$

and

$$\lambda_0 \geq \cdots \geq \lambda_{N_c M_t - 1} \geq \lambda_{N_c M_t} = \cdots = \lambda_{N M_r - 1} = \sigma_v^2 \tag{3.12}$$

The $N_c M_t \times N M_r$ matrix $\mathbf{U}_s$ spans the signal subspace of $\mathbb{R}_{\mathbf{x}}$, while the $(N M_r - N_c M_t) \times N M_r$ matrix $\mathbf{U}_n$ spans the noise subspace of $\mathbb{R}_{\mathbf{x}}$. Given $\mathbb{R}_{\mathbf{s}}$ is of full rank, then according to the standard subspace (SS) method, the matrix $\mathbf{U}_n$ is orthogonal to every column of $\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}$. This can be equivalently expressed as

$$\mathbf{U}_n^H \mathcal{C}_N(\mathbf{h})(\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t})(\mathbf{F}_N^H \otimes \mathbf{I}_{M_t}) = \mathbf{0} \tag{3.13}$$

Note that $(\mathbf{F}_N^H \otimes \mathbf{I}_{M_t})$ is obviously of full rank. Thus the above equation can be further derived as

$$\mathbf{U}_n^H \mathcal{C}_N(\mathbf{h})(\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t}) = \mathbf{0} \tag{3.14}$$

Meanwhile, the block circulant matrix $\mathcal{C}_N(\mathbf{h})$ can be written as

$$\mathcal{C}_N(\mathbf{h}) = [\ \bar{\mathbf{C}}_0\mathbf{h} \quad \bar{\mathbf{C}}_1\mathbf{h} \quad \cdots \quad \bar{\mathbf{C}}_{N-1}\mathbf{h}\ ] \tag{3.15}$$

where $\bar{\mathbf{C}}_n$, $n = 0, \cdots, N-1$, is defined as the matrix containing the first $(L+1)M_r$ columns of $\mathbf{C}_n$, where $\mathbf{C}_n$ is the $N M_r \times N M_r$ circulant matrix with the first column being $\mathbf{e}_{nM_r}$, and $\mathbf{e}_{nM_r}$ is defined as the $(nM_r)^{th}$ column of the identity matrix $\mathbf{I}_{NM_r}$. On the other hand, due to the structure of $\mathbf{T}_{cp}$, it can be verified that

$$\mathcal{C}_N(\mathbf{h})(\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t}) = [\ \mathbb{C}_0\mathbf{h} \quad \mathbb{C}_1\mathbf{h} \quad \cdots \quad \mathbb{C}_{N_c-1}\mathbf{h}\ ] \tag{3.16}$$

where

$$\mathbb{C}_n = \begin{cases} \bar{\mathbf{C}}_{n+N_g} & \text{for} \quad n = 0, \cdots, N_c - N_g - 1 \\ \bar{\mathbf{C}}_{n+N_g} + \bar{\mathbf{C}}_{n+N_g-N_c} & \text{for} \quad n = N_c - N_g, \cdots, N_c - 1 \end{cases} \tag{3.17}$$

Define

$$\boldsymbol{\mathcal{K}} \triangleq [\ \mathbb{C}_0^H \mathbf{U}_n \quad \mathbb{C}_1^H \mathbf{U}_n \quad \cdots \quad \mathbb{C}_{N_c-1}^H \mathbf{U}_n\ ] \tag{3.18}$$

and substitute to Eqn.(3.14), thus the channel matrix $\mathbf{h}$ can be estimated from

$$\boldsymbol{\mathcal{K}}^H \mathbf{h} = \mathbf{0} \tag{3.19}$$

Therefore, the channel matrix $\mathbf{h}$ can be estimated as the left singular vectors of $\boldsymbol{\mathcal{K}}$. In practice, the autocorrelation matrices of the received signal $\mathbf{R}_{\mathbf{x}}(\kappa)$ is unknown and must be estimated from the observed data via time averaging, i.e.

$$\mathbf{R}_{\mathbf{x}}(\kappa) \approx \hat{\mathbf{R}}_{\mathbf{x}}(\kappa) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}(k)\mathbf{x}^H(k+\kappa) \tag{3.20}$$

where K is the total number of samples collected to estimate $\mathbf{R}_{\mathbf{x}}(\kappa)$. We also note that $\mathbf{R}_{\mathbf{x}}(\kappa) = \mathbf{R}_{\mathbf{x}}^H(-\kappa)$. Hence the estimation of the matrix $\mathbb{R}_{\mathbf{x}}(\kappa)$, say $\hat{\mathbb{R}}_{\mathbf{x}}(\kappa)$, is calculated by

$$\hat{\mathbb{R}}_{\mathbf{x}}(\kappa) = \hat{\mathbf{R}}_{\mathbf{x}}(0) + \sum_{\kappa=1}^{\kappa_{\max}+1} \left[\hat{\mathbf{R}}_{\mathbf{x}}(\kappa) + \hat{\mathbf{R}}_{\mathbf{x}}^H(\kappa)\right] \tag{3.21}$$

Therefore, the SVD is applied to $\hat{\mathbb{R}}_{\mathbf{x}}$, resulting in

$$\hat{\mathbb{R}}_{\mathbf{x}} = [\ \hat{\mathbf{U}}_s \quad \hat{\mathbf{U}}_n\ ] \begin{bmatrix} \hat{\boldsymbol{\Lambda}}_s & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\Lambda}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix} \tag{3.22}$$

where $\hat{\mathbf{U}}_s$, $\hat{\mathbf{U}}_n$, $\hat{\boldsymbol{\Lambda}}_s$, and $\hat{\boldsymbol{\Lambda}}_n$ are the noisy estimates of $\mathbf{U}_s$, $\mathbf{U}_n$, $\boldsymbol{\Lambda}_s$, and $\boldsymbol{\Lambda}_n$, respectively. Consequently, the noisy estimates of $\boldsymbol{\mathcal{K}}$ is estimated accordingly, i.e.,

$$\hat{\boldsymbol{\mathcal{K}}} \triangleq [\ \mathbb{C}_0^H \hat{\mathbf{U}}_n \quad \mathbb{C}_1^H \hat{\mathbf{U}}_n \quad \cdots \quad \mathbb{C}_{N_c-1}^H \hat{\mathbf{U}}_n\ ] \tag{3.23}$$

Ideally, if $\hat{\boldsymbol{\mathcal{K}}} = \boldsymbol{\mathcal{K}}$, the eigenvalues of $\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H$ are positive except the smallest one, which is equal to 0. In such case, the channel matrix can be estimated exactly by

solving Eqn.(3.19). However, in practice, the estimation error in $\hat{\boldsymbol{\mathcal{K}}}$ may result in a positive perturbation in the smallest eigenvalue of $\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H$, and Eqn.(3.19) does not have a nontrivial solution. Hence, we consider the following optimization criterion function instead:

$$\hat{\mathbf{h}} = \arg\min_{\|\mathbf{h}\|=1} \|\mathbf{h}^H \boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H \mathbf{h}\|^2 \tag{3.24}$$

which is equivalent to obtain $\hat{\mathbf{h}}$ as the M_t eigen vectors according to the smallest eigenvalue of $\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H$.

Thus, we summarize our estimation algorithm as follows

1. Estimate the autocorrelation matrices of the received signal, $\mathbf{R}_{\mathbf{x}}(\kappa)$ ($\kappa = 0, \cdots, \kappa_{\max} + 1$), by Eqn.(3.20).

2. Calculate the correlation related matrix $\mathbb{R}_{\mathbf{x}}$ by Eqn.(3.21).

3. Apply the SVD to the calculated $\mathbb{R}_{\mathbf{x}}$, and construct $\hat{\boldsymbol{\mathcal{K}}}$ by Eqn.(3.23).

4. Apply SVD to $\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{K}}^H$, and take the estimated channel matrix as the M_t eigenvectors according to the M_t smallest eigenvalues.

3.4 Discussion

3.4.1 Identifiability

In the previous section, the subspace-based blind channel estimation method is proposed, and the criterion function has a close-form solution which can be obtained using the SVD method. In this subsection, we discuss the uniqueness of the proposed estimation algorithm. I.e., we consider the problem whether or not the solution to

to the criteria function Eqn.(3.24) is unique (up to an unitary ambiguity matrix) and under what condition the solution will be unique.

It is not difficult to verify that if there exists a matrix $\hat{\mathbf{h}}$ which satisfies Eqn.(3.24), then

$$\text{span}(\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}) = \text{span}(\mathcal{C}_N(\hat{\mathbf{h}})\tilde{\mathbf{F}}_{cp}) \tag{3.25}$$

which directly follows that

$$\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp} = \mathcal{C}_N(\hat{\mathbf{h}})\tilde{\mathbf{F}}_{cp}\boldsymbol{\mathcal{Q}} \tag{3.26}$$

where $\boldsymbol{\mathcal{Q}}$ is an unknown $NM_t \times NM_r$ ambiguity matrix.

Define the cyclic prefix adding matrix as

$$\mathbf{T}_{cp} \triangleq \begin{bmatrix} \mathbf{I}_{N_c}(N_c - N_g : N_c - 1, :) \\ \\ \mathbf{I}_{N_c} \end{bmatrix} \tag{3.27}$$

where $\mathbf{I}_{N_c}(N_c - N_g : N_c - 1, :)$ denotes the submatrix of $\mathbf{I}_{N_c}$ which contains its last N_g rows. Then, Eqn.(3.26) can be rewritten as

$$\mathcal{C}_N(\mathbf{h})\left[\left(\mathbf{T}_{cp}\mathbf{F}_{N_c}^H\right) \otimes \mathbf{I}_{M_t}\right] = \mathcal{C}_N(\hat{\mathbf{h}})\left[\left(\mathbf{T}_{cp}\mathbf{F}_{N_c}^H\right) \otimes \mathbf{I}_{M_t}\right]\boldsymbol{\mathcal{Q}} \tag{3.28}$$

Theorem 3.4.1 Let $\mathbf{h}$ be a $(L+1)M_r \times M_t$ "tall" matrix with full column rank, and $\mathcal{C}_N(\mathbf{h})$ be the $NM_r \times NM_t$ block circulant matrix constructed from $\mathbf{h}$. If $N \neq kN_g$ (k is an arbitrary positive integer greater than 1), and Eqn.(3.26) is satisfied, then

$$\boldsymbol{\mathcal{Q}} = \mathbf{I}_{N_c} \otimes \mathbf{Q} \tag{3.29}$$

where $\mathbf{Q}$ is an invertible ambiguity matrix.

Proof: According to the matrix multiplication property of the Kronecker product, Eqn.(3.28) can be rewritten as

$$\mathcal{C}_N^\dagger(\hat{\mathbf{h}})\mathcal{C}_N(\mathbf{h})\left(\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t}\right) = \left(\mathbf{T}_{cp} \otimes \mathbf{I}_{M_t}\right)\tilde{\mathbf{F}}_{N_c}^H\boldsymbol{\mathcal{Q}}\tilde{\mathbf{F}}_{N_c} \tag{3.30}$$

where $\tilde{\mathbf{F}}_{N_c} \triangleq \mathbf{F}_{N_c} \otimes \mathbf{I}_{M_t}$. Define

$$\mathcal{C} \triangleq \mathcal{C}_N^{\dagger}(\hat{\mathbf{h}})\mathcal{C}_N(\mathbf{h}) \tag{3.31}$$

We note that $\mathcal{C}$ is a block circulant matrix. Without loss of generality, we can assume $\mathcal{C}$ is constructed from the $NM_r \times M_t$ matrix $\boldsymbol{\vartheta}$, which can be divided into blocks with size $M_t \times M_t$, and we denote the nth $(n = 0, 1, \cdots, N-1)$ $M_t \times M_t$ block of $\boldsymbol{\vartheta}$ as $\boldsymbol{\vartheta}_n$. Eqn.(3.30) also indicates that the first $M_t N_g$ rows of of $\tilde{\mathcal{C}} \triangleq \mathcal{C}\mathbf{T}_{cp}$ are identical to last $M_t N_g$ rows of itself respectively. Thus, we have the following equations

$$\begin{cases} \boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_{1+N_g} = \boldsymbol{\vartheta}_{1+2N_g} = \cdots \\ \boldsymbol{\vartheta}_2 = \boldsymbol{\vartheta}_{2+N_g} = \boldsymbol{\vartheta}_{2+2N_g} = \cdots \\ \qquad\qquad \vdots \\ \boldsymbol{\vartheta}_{N_g} = \boldsymbol{\vartheta}_{2N_g} = \boldsymbol{\vartheta}_{3N_g} = \cdots \end{cases} \tag{3.32}$$

and

$$\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_{N-2N_g+1}; \quad \boldsymbol{\vartheta}_2 = \boldsymbol{\vartheta}_{N-2N_g+2}; \quad \cdots; \quad \boldsymbol{\vartheta}_{2N_g-1} = \boldsymbol{\vartheta}_{N-1} \tag{3.33}$$

Thus, it can be verified that if $N \neq kN_g$ (k is an arbitrary positive integer greater than 1), then $\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_2 = \cdots = \boldsymbol{\vartheta}_{N-1}$.

On the other hand, recall the condition that $N \geq L+1$ and $M_r \geq M_t$, which means that $\mathbf{h}_{N-1} = \hat{\mathbf{h}}_{N-1} = \mathbf{0}$. Hence, we have

$$\begin{aligned} &\mathbf{0}\boldsymbol{\vartheta}_0 + \mathbf{h}_{N-2}\boldsymbol{\vartheta}_1 + \cdots + \mathbf{h}_0\boldsymbol{\vartheta}_{N-1} = (\mathbf{h}_0 + \cdots + \mathbf{h}_{N_2})\boldsymbol{\vartheta}_1 = \mathbf{0} \\ \Rightarrow \quad &\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_2 = \cdots = \boldsymbol{\vartheta}_{N-1} = \mathbf{0} \end{aligned} \tag{3.34}$$

Substitute to Eqn.(3.31), we have

$$\mathcal{C}_N(\mathbf{h}) = \mathcal{C}_N(\hat{\mathbf{h}})(\mathbf{I}_N \otimes \boldsymbol{\vartheta}_1) = \mathcal{C}_N(\hat{\mathbf{h}})(\mathbf{I}_N \otimes \mathbf{Q}) = \mathcal{C}_N(\hat{\mathbf{h}})\boldsymbol{\mathcal{Q}} \tag{3.35}$$

where $\boldsymbol{\mathcal{Q}} = \mathbf{I}_N \otimes \boldsymbol{\vartheta}_1$, and $\mathbf{Q} = \boldsymbol{\vartheta}_1$. ■

According to Theorem 3.4.1, if $\hat{\mathbf{h}}$ the estimated channel matrix by the proposed algorithm, then it must be satisfied that

$$\mathcal{C}_N(\mathbf{h}) = \mathcal{C}_N(\hat{\mathbf{h}})(\mathbf{I}_N \otimes \mathbf{Q}) \tag{3.36}$$

which is equivalent to

$$\mathbf{h} = \hat{\mathbf{h}}\mathbf{Q} \tag{3.37}$$

Thus, we can conclude that the proposed estimation algorithm can uniquely estimate the channel matrix up to an unitary ambiguity matrix $\mathbf{Q}$. The ambiguity matrix indeed exists in all kinds of blind channel estimators and can be remedied by introducing extra constraints, e.g., using the blind source separation [86, 87]. Thus, we assume that the unitary ambiguity matrix is known exactly in the rest of this chapter.

3.4.2 Comparison with the Existing Algorithm

In this subsection, we compare the proposed algorithm with the existing subspace-based algorithm proposed in [70]. In common, both algorithms make use of the redundancy induced by the CP. Based on the fact that the channel matrix is orthogonal to the noise subspace of the correlation matrix of the channel output, the time domain channel matrix can be blindly estimated up to an ambiguity matrix.

However, despite the similarities mentioned above, the proposed algorithm differs from the previously existing algorithm in the following aspects:

Firstly, the way to construct the criterion functions are different. The existing algorithm collects multiple consecutive OFDM symbols, and calculates the correlation

matrix of the resulted "long" symbols. On the other hand, the proposed algorithm constructs the target matrix $\mathbb{R}_{\mathbf{x}}$ by taking the summation of the autocorrelation matrices of the received OFDM symbols with different delay lags (See Eqn.(3.9)).

Secondly, the computational costs are different. The existing algorithms require the correlation matrix of the collection of multiple consecutive OFDM symbols. Therefore, the size of the correlation matrix is $(KNM_r - LM_r) \times (KNM_r - LM_r)$, where $K \geq 2$ is the number of consecutive OFDM symbols. Since the length of the OFDM symbols N is usually quite large, it suffers the inefficiency caused by the SVD of the large-sized correlation matrix. However, the proposed algorithm fixes the size of the target matrix $\mathbb{R}_{\mathbf{x}}$ to be $NMr \times NMr$. Thus the computational costs of the existing algorithm for calculating the correlation matrix and the SVD decomposition are both higher than those of the proposed algorithm. Furthermore, according to Eqn.(3.9), for the proposed algorithm, we only need to calculate the autocorrelation matrices $\mathbf{R}_{\mathbf{x}}(\kappa)$ for $\kappa \geq 0$, since those for $\kappa < 0$ can be easily obtained by taking $\mathbf{R}_{\mathbf{x}}(-\kappa) = \mathbf{R}_{\mathbf{x}}^H(\kappa)$. Therefore, the computational cost can be further reduced.

Lastly, the conditions for estimation are different. To benefit from the CP redundancy, $KNM_r - LM_r > KN_cMt$ is required for the existing algorithm. Thus, the upper bound of the channel order must be known, and K must be selected such that

$$K > \frac{LM_r}{NM_r - N_cM_t} \tag{3.38}$$

If there exists channel order over estimation, or the transmission environment changes, then the existing algorithm, may not work properly. On the contrary, for the proposed algorithm, besides the other common conditions, we only need $N_g > 0$. Obvi-

ously, satisfying this condition is easier than satisfying Eqn.(3.38). Even when the channel order changes, it does not influence the proposed estimator.

3.4.3 Asymptotic Performance Analysis

In this subsection, we investigate the asymptotic performance of the proposed blind channel estimation algorithm due to the finite received data samples and the AWGN noise. Recall Eqn.(3.22), and rewrite the noisy estimates of $\mathbf{U}_s$ and $\mathbf{U}_n$ as

$$\hat{\mathbf{U}}_s \triangleq \mathbf{U}_s + \Delta\mathbf{U}_s \tag{3.39a}$$

$$\hat{\mathbf{U}}_n \triangleq \mathbf{U}_n + \Delta\mathbf{U}_n \tag{3.39b}$$

where $\Delta\mathbf{U}_s$ and $\Delta\mathbf{U}_n$ are the perturbation in the estimated signal and noise subspaces. From Eqn.(3.9) we have the following

$$\begin{aligned} \mathbf{\Xi} &= \mathbb{R}_{\mathbf{x}} - \sigma_v^2 \mathbf{I}_{NM_r} \\ &= \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \end{aligned} \tag{3.40}$$

where $\mathbf{\Xi} \triangleq \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbb{R}_{\mathbf{s}}\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h})$. By substituting the above estimated subspaces components into Eqn.(3.40), we obtain

$$\begin{aligned} \mathbf{\Xi} + \Delta\mathbf{\Xi} &= \hat{\mathbb{R}}_{\mathbf{x}} - \sigma_v^2 \mathbf{I} \\ &= \begin{bmatrix} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Lambda}}_s & \mathbf{0} \\ \mathbf{0} & \Delta\mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix} \end{aligned} \tag{3.41}$$

where $\Delta\mathbf{\Xi}$ denotes the perturbation of $\mathbf{\Xi}$ and σ_v^2 is the noise power. In a noise free case, we have $\Delta\mathbf{\Lambda}_n = \mathbf{0}$. Next, we consider the first order perturbation of the noise subspace of $\mathbb{R}_{\mathbf{x}}$. We first introduce the following lemma in [88].

Lemma 3.4.2 The perturbed noise subspace $\Delta\mathbf{U}_n$ is spanned by $\mathbf{U}_s\mathbf{W}$ and the perturbed signal subspace $\Delta\mathbf{U}_n$ is spanned by $\mathbf{U}_n\mathbf{T}$, where $\mathbf{W}$ and $\mathbf{T}$ are the matrices whose norms are of the order of $\|\Delta\Xi\|$. The matrix norm can be any submultiplicative norm such as the Euclidean 2-norm.

Proof: See [88]. ■

At high SNR, the first order perturbation expansion of $\Delta\mathbf{U}_s$ and $\Delta\mathbf{U}_n$ can be expressed as a linear approximation form. First, the perturbed signal subspace $\hat{\mathbf{U}}_s$ and the perturbed noise subspace $\hat{\mathbf{U}}_n$ can be expressed as follows respectively

$$\hat{\mathbf{U}}_s \triangleq \mathbf{U}_s + \mathbf{U}_n\mathbf{T} \tag{3.42a}$$

$$\hat{\mathbf{U}}_n \triangleq \mathbf{U}_n + \mathbf{U}_s\mathbf{W} \tag{3.42b}$$

Left multiplying both side of Eqn.(3.41) by $\hat{\mathbf{U}}_n^H$, we have

$$\begin{aligned}\hat{\mathbf{U}}_n^H(\Xi + \Delta\Xi) &= \hat{\mathbf{U}}_n^H(\hat{\mathbf{U}}_s\hat{\mathbf{\Lambda}}_s\hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n\Delta\mathbf{\Lambda}_n\hat{\mathbf{U}}_n^H) \\ &= \Delta\mathbf{\Lambda}_n\hat{\mathbf{U}}_n^H \end{aligned}\tag{3.43}$$

The second equality follows from the fact the $\hat{\mathbf{U}}_s \perp \hat{\mathbf{U}}_n$ and $\hat{\mathbf{U}}_n^H\hat{\mathbf{U}}_n = \mathbf{I}$. By substituting Eqn.(3.42a) and Eqn.(3.42b) into Eqn.(3.43), we get

$$(\mathbf{U}_n + \mathbf{U}_s\mathbf{W})^H(\Xi + \Delta\Xi) = \Delta\mathbf{\Lambda}_n(\mathbf{U}_n + \mathbf{U}_s\mathbf{W})^H \tag{3.44}$$

Neglect the second-order terms and use the fact that $\mathbf{U}_n^H\Xi = \mathbf{0}$, then we get

$$\mathbf{W} \doteq -\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H(\Delta\Xi)^H\mathbf{U}_n \tag{3.45}$$

and consequently,

$$\begin{aligned}\Delta\mathbf{U}_n &= -\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H\Delta\Xi^H\mathbf{U}_n \\ &= -\left(\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbb{R}_{\mathbf{s}}\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h})\right)^{\dagger}\left[\Delta\left(\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbb{R}_{\mathbf{s}}\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h})\right)\right]^H\mathbf{U}_n \end{aligned}\tag{3.46}$$

According to the proposed estimation algorithm, the channel matrix is taken as the M_t left singular vectors according to the smallest singular values. Thus, apply Lemma 3.4.2 again, we get

$$\Delta\mathbf{h} \triangleq \hat{\mathbf{h}} - \mathbf{h} = -(\boldsymbol{\mathcal{K}}^H)^\dagger \Delta\boldsymbol{\mathcal{K}}^H \mathbf{h} \tag{3.47}$$

where

$$\begin{aligned} \Delta\boldsymbol{\mathcal{K}} &= [\ \mathbb{C}_0^H \Delta\mathbf{U}_n \quad \mathbb{C}_1^H \Delta\mathbf{U}_n \quad \cdots \quad \mathbb{C}_{N_c-1}^H \Delta\mathbf{U}_n\ ] \\ &= -[\ \mathbb{C}_0^H \boldsymbol{\Xi}^\dagger \Delta\boldsymbol{\Xi}\mathbf{U}_n \quad \mathbb{C}_1^H \boldsymbol{\Xi}^\dagger \Delta\boldsymbol{\Xi}\mathbf{U}_n \quad \cdots \quad \mathbb{C}_{N_c-1}^H \boldsymbol{\Xi}^\dagger \Delta\boldsymbol{\Xi}\mathbf{U}_n\ ] \end{aligned} \tag{3.48}$$

Finally, the covariance matrix of the channel estimation error can be obtained as

$$\begin{aligned} & E\left\{\Delta\vec{\mathbf{h}}\Delta\vec{\mathbf{h}}^H\right\} \\ = & \left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] E\left\{\left(\mathbf{I}_{M_t} \otimes \Delta\boldsymbol{\mathcal{K}}^H\right) \vec{\mathbf{h}}\vec{\mathbf{h}}^H \left(\mathbf{I}_{M_t} \otimes \Delta\boldsymbol{\mathcal{K}}\right)\right\} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right] \end{aligned} \tag{3.49}$$

where $\vec{\mathbf{h}}$ denotes the column vector constructed by concatenating the columns of the matrix $\mathbf{h}$.

The covariance matrix of the channel estimation error expressed in Eqn.(3.49) is not easy to be resolved further. However, it can be assumed that it can be assumed that $\kappa_{\max} = 0$, which greatly simplifies the derivation. This assumption is reasonable because in many practical cases, the source signals are independent and identically distributed (i.i.d.), or colored signals which are block precoded from the i.i.d. source. Thus, we can model the source signal $\mathbf{s}(k)$ as

$$\mathbf{s}(k) = \mathbf{P}\mathbf{d}(k) \tag{3.50}$$

where $\mathbf{P}$ is the $M_tN_c \times M_tN_c$ precoding matrix, and $\mathbf{d}(k)$ is the i.i.d. source with unit power such that

$$\mathbf{R}_{\mathbf{d}}(\kappa) \triangleq E\left\{\mathbf{d}(k)\mathbf{d}^H(k+\kappa)\right\} = \mathbf{I}_{M_tN_c}\delta(\kappa) \tag{3.51}$$

The assumption that the source signal $\mathbf{d}(k)$ has unit power is reasonable, because the transmission power can be absorbed in the precoding matrix $\mathbf{P}$. Especially, when the transmitted source signal $\mathbf{s}(k)$ is i.i.d., the precoding matrix must satisfy that $\mathbf{P}\mathbf{P}^H = \sigma_s^2 \mathbf{I}_{M_t N_c}$, where σ_s^2 is the transmission power. Under this assumption, we derive the covariance matrix of the channel estimation error.

Theorem 3.4.3 Assume that the noise is zero-mean i.i.d with covariance σ_v^2, and the transmitted signal is modeled as $\mathbf{s}(k) = \mathbf{P}\mathbf{d}(k)$, where $\mathbf{d}(k)$ is the i.i.d. source signal with unit power, and $\mathbf{P}$ is the non-redundant precoding matrix. Then the covariance matrix of the channel estimation error is approximated by

$$E\left\{\Delta\vec{\mathbf{h}}\Delta\vec{\mathbf{h}}^H\right\} \approx \left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] \mathcal{E} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right] \tag{3.52}$$

and the channel estimation MSE is

$$E\left\{\|\Delta\vec{\mathbf{h}}\|^2\right\} = \mathrm{tr}\{\left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] \mathcal{E} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right]\} \tag{3.53}$$

where the corresponding notations are defined in Appendix A.

Proof: See Appendix A. ■

The channel estimation error is mainly caused by two factors: the AWGN noise and the estimation error of the correlation matrix. On one hand, The estimation error caused by the AWGN noise is common in the existing subspace-based channel estimation algorithms. On the other hand, the proposed estimation algorithm prospects $\mathbf{\Phi}_{\kappa_{\max}+1}$ in Eqn.(3.7) to be a zero matrix. However, its estimation by time averaging is perturbated by $\Delta\mathbf{\Phi}_{\kappa_{\max}+1}$, which trends to be zero when the number of the observation samples K is large.

3.5 Simulation Results

In this section, we consider three simulation examples to illustrate the performance of our proposed algorithm.

Example 1

In this example, we compare the NRMSE and BER performances of the proposed estimation algorithm with the existing subspace-based algorithm in [79] and the training based LS algorithm in [68]. The theoretical asymptotic NRMSE performance of the proposed algorithm discussed in Section 3.4.3 is also demonstrated. The system is driven by the white source signals which are extracted form the QPSK constellations. The number of subcarriers for each OFDM symbol is $N_c = 64$, and the length of the CP is $N_g = 4$ for the existing algorithm and the training based algorithm, which are labeled as "MUSIC" and "Training Based" respectively. For the proposed algorithm, we set the length of CP to be $N_g = 4$ and $N_g = 1$, to show the performances in the normal and ill conditioned environments respectively.

The simulated OFDM system is equipped with 2 transmit and 2 receive antennae respectively, and the channel is modeled as a 4-tap FIR filter with tap coefficients independently chosen from a white Gaussian process. We simulate 30 independent channels and each for 100 Monte Carlo runs. To evaluate the channel estimation error, we employ the normalized-root-mean-square-error (NRMSE), which is defined as

$$\text{NRMSE} = \sqrt{\frac{1}{N_1 N_2} \sum_{t_1=1}^{N_1} \sum_{t_2=1}^{N_2} \frac{\|\hat{\mathbf{h}}^{(t_1,t_2)} - \mathbf{h}^{(t_1)}\|_F^2}{\|\mathbf{h}^{(t_1)}\|_F^2}} \tag{3.54}$$

where $\|.\|_F$ denotes the Frobenius norm, N_1 and N_2 are the number random channels

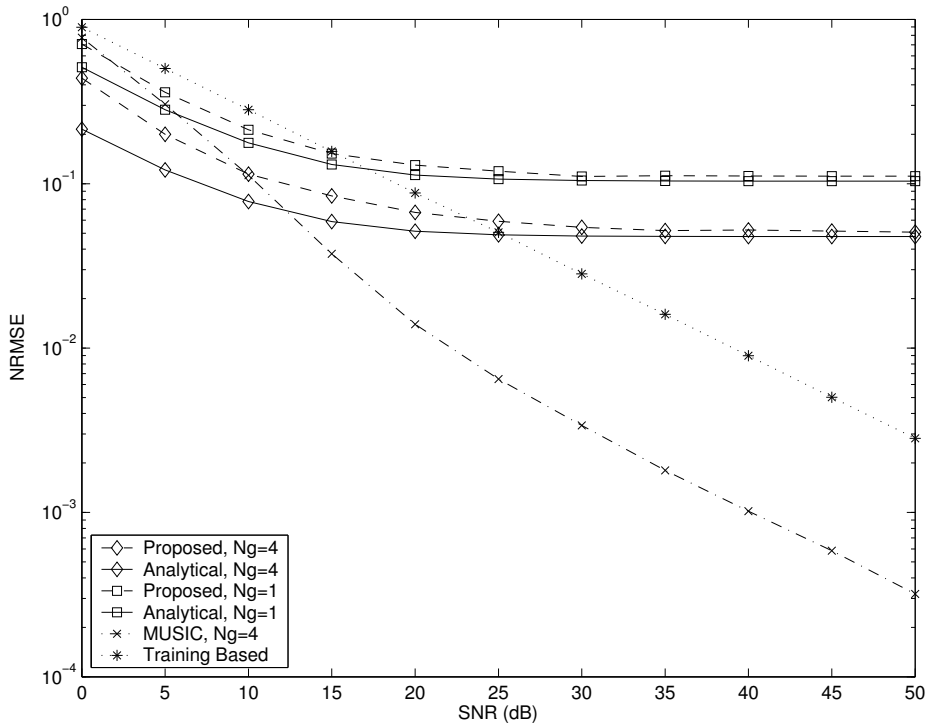

Figure 3.1: NRMSE performance as a function of SNR

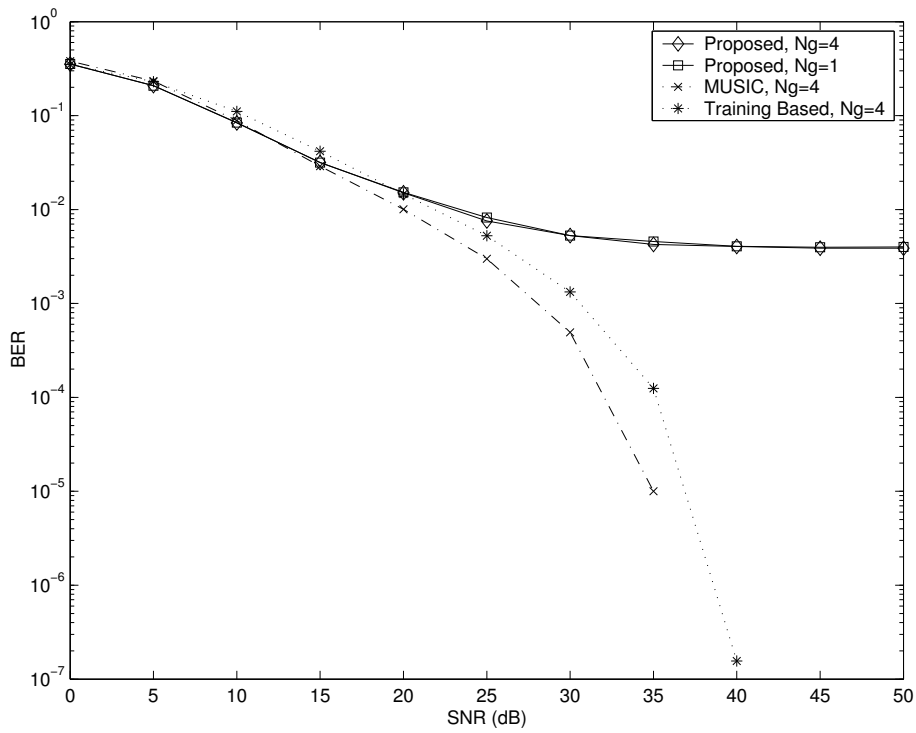

Figure 3.2: BER performance as a function of SNR

and the number of Monte Carlo runs for each simulated channel respectively. $\mathbf{h}^{(t_1)}$ is the true channel matrix, and $\hat{\mathbf{h}}^{(t_1,t_2)}$ is the estimation of $\mathbf{h}^{(t_1)}$ for the t_2^{th} Monte Carlo run.

Figure 3.1 illustrates the NRMSE as a function of the signal-to-noise ratio (SNR). For the subspace based algorithms, both existing and proposed, the number of the observed OFDM symbols is set to be 1000, while 2 optimized training OFDM symbols are used for the training based LS algorithm. The number of the consecutive OFDM symbols for the existing algorithm is $K = 2$. From the figure, we see that performances of these algorithms are very close to each other. Although the performance of the proposed algorithm has an error floor when SNR is high, it does not impact the motivation of the proposed algorithm which relies on the computational efficiency and the identifiably. In this example, K is set to be 2 for the existing algorithm. Therefore, when $N_g = 1$, Eqn.(3.38) is not satisfied, and the estimation fails. However, from the figure, we see that the proposed algorithm works properly regardless the length of the CP. The similar conclusion can be drawn from Figure 3.2, which illustrates the BER as a function of the SNR. The detected symbols are the outputs of the Zero Forcing equalizers using the estimated channel coefficients.

Example 2

In this example, we evaluate the NRMSE and BER performances of the proposed algorithm subject to different SNR and number of observed symbols (NOS), as shown in Figure 3.3 and Figure 3.4 respectively. The parameter setting is the same as Example 1, except that the length of CP, $N_g = 4$ only. Similar to the previous example, the theoretical asymptotic NRMSE performance is as shown. We illustrate

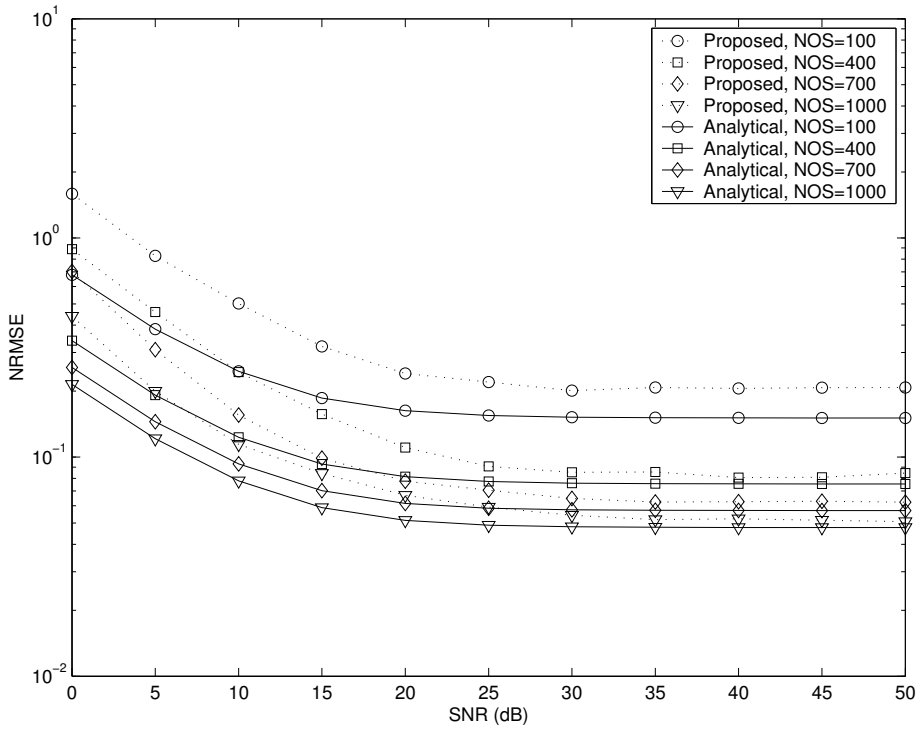

Figure 3.3: NRMSE performance as a function of SNR

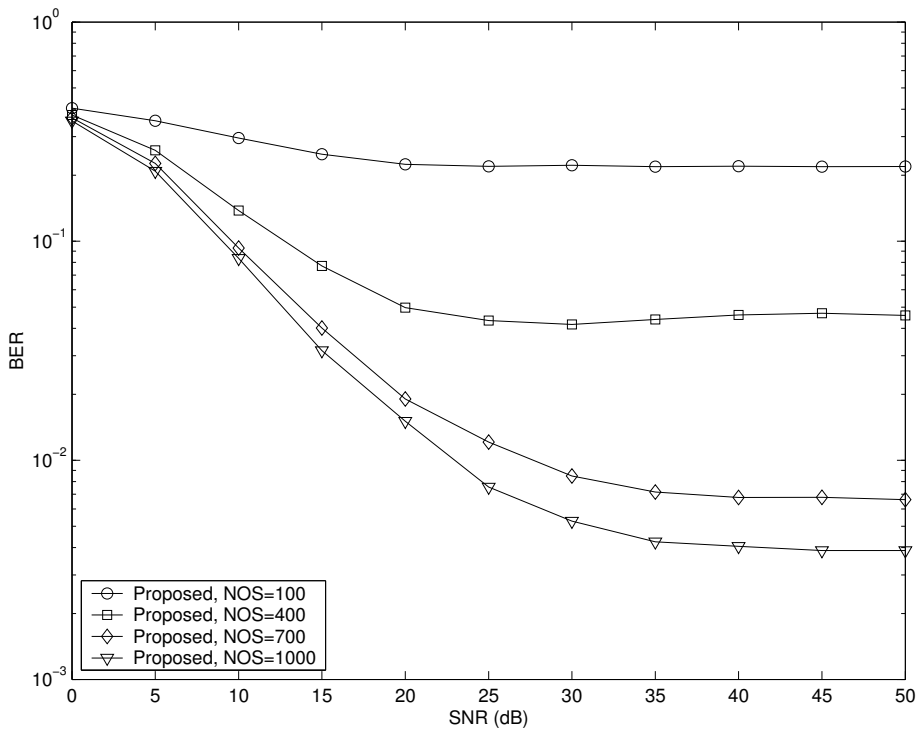

Figure 3.4: BER performance as a function of SNR

the performances where the number of the observed OFDM symbols being 100, 400, 700 and 1000, and the SNR from 0 to 50 dB. From these two figures, it is obvious that the proposed estimator achieves better performances when more number of received OFDM symbols are observed, i.e., NOS is higher. The reason is simple. When more OFDM symbols are observed, then the estimate of the target matrix $\mathbb{R}_{\mathbf{x}}$ is more accurate, and consequently we get better NRMSE and BER performances.

Example 3

In this example, the simulated system is driven by the colored source. The colored source symbols $s(n)$ are drawn from a 4-QAM constellation according to the following rule. Let b_n be the input stream of independent and identically distributed bits, i.e., $b_n \in \{0, 1\}$. Then

$$s(n) = \begin{cases} -\frac{\sqrt{(2)}}{2} + \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (0 \; 0) \\ +\frac{\sqrt{(2)}}{2} + \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (0 \; 1) \\ -\frac{\sqrt{(2)}}{2} - \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (1 \; 0) \\ +\frac{\sqrt{(2)}}{2} - \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (1 \; 1) \end{cases} \tag{3.55}$$

This generates a colored symbol sequence with autocorrelation

$$E\{s(n)s^*(n+\tau)\} = \begin{cases} 1 & \text{if} \quad \tau = 0 \\ \mp\frac{1}{2}j & \text{if} \quad \tau = \pm 1 \\ 0 & \text{else} \end{cases} \tag{3.56}$$

Therefore, in this example, $\kappa_{\max} = 1$, and hence

$$\hat{\mathbb{R}}_{\mathbf{x}} \triangleq \hat{\mathbb{R}}_{\mathbf{x}}(\kappa_{\max} + 1) = \sum_{\kappa=-2}^{2} \hat{\mathbf{R}}_{\mathbf{x}}(\kappa) \tag{3.57}$$

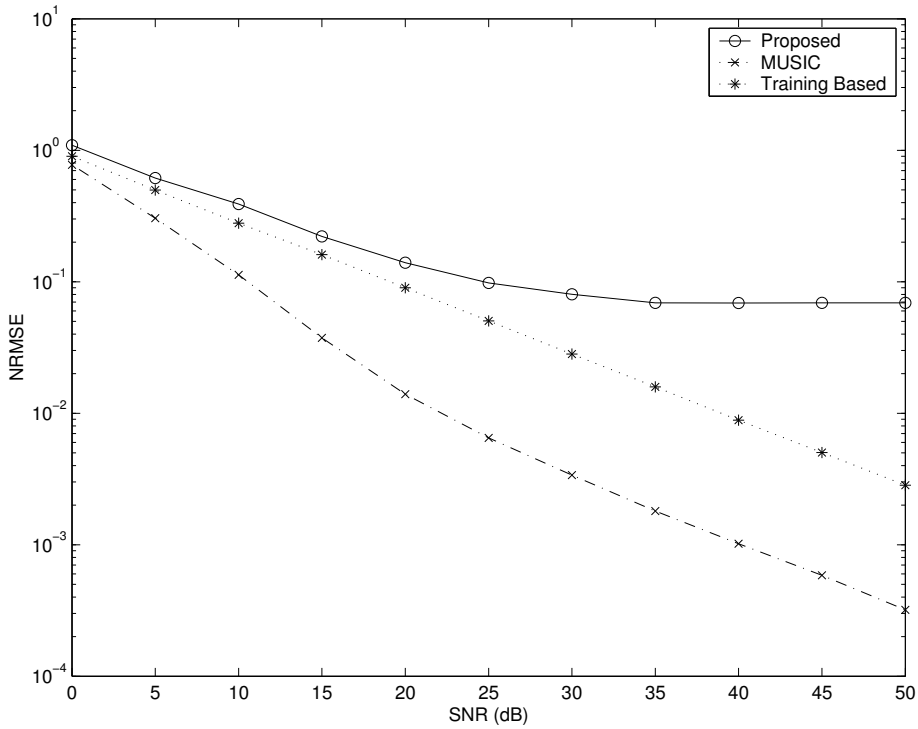

Figure 3.5: NRMSE performance as a function of SNR

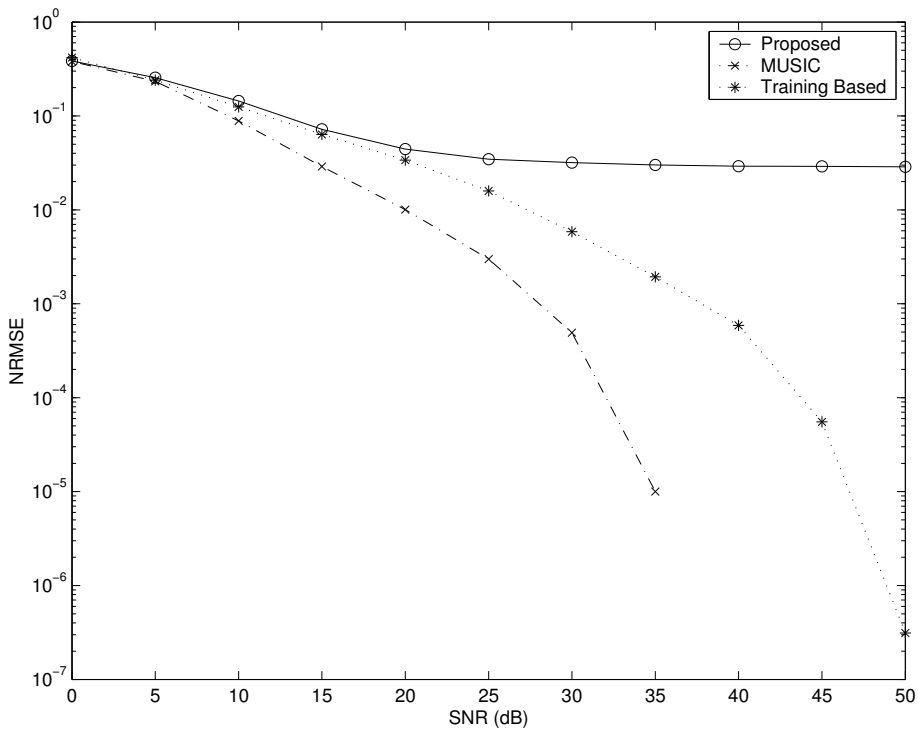

Figure 3.6: BER performance as a function of SNR

Figure 3.5 and Figure 3.6 illustrate the NRMSE and BER performance of the proposed estimation algorithm, together with the existing SS method [79] and the training based algorithm [68]. The parameter setting is the same as Example 1, except that the number of the observed OFDM symbols for proposed and existing SS methods are both 2000. Comparing these two figures with Figure 3.1 and Figure 3.2, we see that there exists performance degradation of the proposed algorithm when source signals are changed to colored. This performance degradation is due to the inaccuracy of the estimation of the target matrix $\mathbb{R}_{\mathbf{x}}$. In this example, the maximum delay lag of the nonzero correlation matrix $\mathbf{R}_{\mathbf{x}}(\kappa)$ is $\kappa_{\max} = 1$. Thus, two more correlation matrices, $\mathbf{R}_{\mathbf{x}}(\pm 2)$, need to be involved into $\mathbb{R}_{\mathbf{x}}$, which brings in extra estimation error. On the other hand, the existing MUSIC algorithm does not face such problem. Thus the performance of the proposed algorithm is worse than the existing MUSIC algorithm in this case. It must be pointed out that if the delay lag τ is bigger, then the performance of the proposed algorithm will depredate even more. However, although the proposed algorithm faces the performance degradation problem, it is still valuable because its motivation relies on its computation efficiency comparing with the existing MUSIC algorithm. Moreover, the source signals are white in many practical cases. Therefore, the proposed estimation algorithm still achieves fairly attractive performance in practice.

3.6 Summary

In this chapter, we proposed a new SS based method that admits a closed-form solution for blind channel estimation of MIMO-OFDM systems. The uniqueness of

the solution and the asymptotic performance analysis are presented. This proposed estimation algorithm does not require the number of the receive antenna to be greater than the transmit antenna. It neither requires the length of CP to be greater than the channel order. It is also more computationally efficient than the existing SS method for MIMO-OFDM systems. These properties make the proposed channel estimator to be attractive, especially in the ill conditioned environment. The analytical and the simulated NRMSE and BER performance were both presented. The existing SS algorithm and the training based LS algorithm were also simulated as comparisons. The simulations were carried out in two environments, where the systems were driven by white and colored source signals respectively. It was shown that the proposed estimation algorithm is more attractive when the source signal is white.

Chapter 4

Blind Channel Estimation For Linearly Precoded MIMO-OFDM Systems

4.1 Introduction

In the previous chapter, we proposed a subspace (SS) based method for blind channel estimation of MIMO-OFDM systems, which exploits the redundancy introduce by the CP. In general, most of the existing SS algorithms for MIMO-OFDM systems are processed in the time domain before removing the guard interval. Hence, they are either designed for CP based OFDM [70] or ZP based OFDM [74], and are not suitable for both simultaneously. Besides, some strict requirements must be satisfied for the SS methods [4].

Using the precoding based algorithm is a way to solve these problems. Fre-

quency domain linear precoding is an effective method to compensate the lack of the multipath diversity for OFDM systems so as to avoid the catastrophic effects of channel zeros at certain subcarriers [108]. It also motivates an alternative method for blind channel estimation. There are mainly two types of the precoding schemes, redundant [81, 109, 113] and nonredundant [83, 114–116]. The algorithms for the redundantly precoded OFDM systems exploit the noise subspace freedom of the signal correlation matrix, and directly apply the existing SS algorithms. In [81], a SS method was proposed for space-time coded (STC) MIMO-OFDM systems with the assistance of properly designed redundant linear precoding. In [109], a SS method was proposed for the ZP based MIMO STC-OFDM system extended from the system in [111]. In [83], a SS method was proposed by considering the existence of virtual subcarriers (VCs). In fact, using VCs can be viewed as a special case of linear precoding. On the other hand, the nonredundant precoding provides cross correlations between the signals transmitted on different subcarriers. Based on the assumption that the transmitted symbols are independent and identically distributed to each other, a new type of blind channel estimation method for SISO-OFDM systems has been proposed in [114–116], where a nonredundant linear precoder is used at the transmitter, and the CSI is possessed in all entries of the signal covariance matrix. In [117], a SS method was proposed by using the second order cyclostationary statistics induced by employing a periodic nonconstant-modulus antenna precoding.

In this chapter, we propose a novel approach for blind MIMO-OFDM channel estimation. A nonredundant linear precoder is applied to each source data block before the conventional OFDM transmission. Due to the structure introduced by

the precoding matrix, the channel can be estimated at the receiver based on general SVD operations. In many existing algorithms, the assumption that the channel transfer functions share no common zeros at subcarrier frequencies is necessary. However, our proposed algorithm can still work even when this assumption is not fulfilled. The asymptotic performance analysis of the proposed algorithm and the SNR degradation caused by the precoding scheme are also discussed.

The rest of this paper is organized as follows. In Section 4.2, we review the MIMO-OFDM system model and formulate the problem. In Section 4.3, we propose the blind estimation algorithm with the assistance of linear precoder. The identifiability, precoder designing, SNR degradation due to the precoder, and the asymptotic performance are discussed in Section 4.4. Simulations results are demonstrated in Section 4.5, and summary of this chapter is presented in the last section.

4.2 System Model

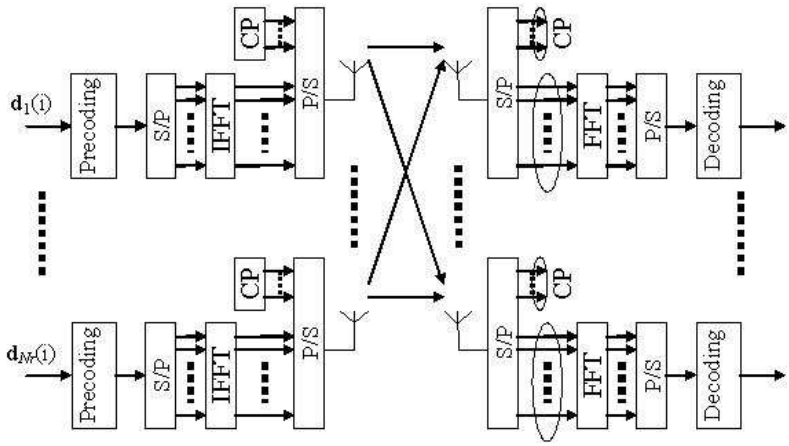

Figure 4.1: Linearly Precoded MIMO-OFDM System Block Diagram

Consider the above MIMO-OFDM system equipped with M_t and M_r transmit and receive antennae respectively. Define the kth block of data stream to be transmitted by the ith $(i = 1, \cdots, M_t)$ transmit antenna as

$$\mathbf{d}_i(k) \triangleq \begin{bmatrix} d_i(k,0) & d_i(k,1) & \cdots & d_i(k,N_c-1) \end{bmatrix}^T \tag{4.1}$$

where N_c is the number of subcarriers. We assume that the transmitted signals are independent and identically distributed (i.i.d.) with zero-mean and unit variance. A $N_c \times N_c$ precoding matrix $\mathbf{P}$ is applied to each block, mapping them as

$$\mathbf{s}_i(k) = \mathbf{P}\mathbf{d}_i(k) \tag{4.2}$$

Then the coded blocks are then transmitted through conventional MIMO-OFDM systems as shown in Figure 4.1. Let $h_{j,i}(l), (i = 1, \cdots, M_t, j = 1, \cdots, M_r, l = 0, \cdots, L)$ denote the lth tap of the time domain channel impulse response between the ith transmit antenna and the jth receive antenna, where L is the maximum channel order. We assume that the channel stays the same for a number of successive symbol blocks. Hence, the received signals after removing CP and FFT demodulation at the jth receive antenna is given by

$$\mathbf{y}_j(k) = \sum_{i=1}^{M_t} \mathcal{D}(\mathbf{H}_{ji})\mathbf{s}_i(k) + \mathbf{n}_j(k) \tag{4.3}$$

where $\mathbf{n}_j(k)$ is the zero mean white Gaussian noise vector with variance $\sigma_n^2\mathbf{I}_{N_c}$, $\mathcal{D}(\mathbf{H}_{ji})$ denotes the diagonal matrix with the elements of vector $\mathbf{H}_{ji}$ along its diagonal, and $\mathbf{H}_{ji}$ represents the frequency domain channel vector for each transmit-receive antenna pair, which is mathematically defined as

$$\mathbf{H}_{j,i} \triangleq \begin{bmatrix} H_{j,i}(0) & \cdots & H_{j,i}(N_c-1) \end{bmatrix}^T \tag{4.4}$$

where

$$H_{j,i}(n) = \sum_{l=0}^{L} h_{j,i}(l) e^{-j\frac{2\pi}{N_c}nl} \tag{4.5}$$

for $n = 0, \cdots, N_c - 1$, $i = 1, \cdots, M_t$, and $j = 1, \cdots, M_r$.

Our task is to estimate the channel matrices $\mathbf{H}_{ji}$, (for $i = 1, \cdots, M_t$, $j = 1, \cdots, M_r$), through the received OFDM symbols $\mathbf{y}_j(k)$.

4.3 Proposed Blind Channel Estimation

Without loss of generality, we focus on the wireless channels associated with the jth receive antenna. Consider the correlation matrix of the received signal

$$\begin{aligned}
\mathbf{R}_{\mathbf{y}_j} &\triangleq E\left\{\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right\} \\
&= E\left\{\left(\sum_{i_1=1}^{M_t} \mathcal{D}(\mathbf{H}_{j,i_1})\mathbf{s}_{i_1}(k) + \mathbf{n}_j(k)\right)\left(\sum_{i_2=1}^{M_t} \mathcal{D}(\mathbf{H}_{j,i_2})\mathbf{s}_{i_2}(k) + \mathbf{n}_j(k)\right)^H\right\} \\
&= \sum_{i=1}^{M_t} \mathcal{D}(\mathbf{H}_{j,i})\mathbf{P}\mathbf{P}^H\mathcal{D}(\mathbf{H}_{j,i})^H + \sigma_n^2\mathbf{I}
\end{aligned} \tag{4.6}$$

According to the properties of the diagonal matrices, it is not difficult to verify that

$$\mathcal{D}(\mathbf{H}_{ji})\mathbf{P}\mathbf{P}^H\mathcal{D}^H(\mathbf{H}_{j,i}) = \mathbf{H}_{j,i}\mathbf{H}_{j,i}^H \odot \mathbf{P}\mathbf{P}^H \tag{4.7}$$

where $\odot$ denotes the element-by-element multiplication of two matrices with identical size. Assume $\mathbf{P}$ is of full rank, and $\mathbf{\Phi}$ has unit diagonal entries and no zero entries, where we define $\mathbf{\Phi} \triangleq \mathbf{P}\mathbf{P}^H$. Hence, we can perform the element-by-element division of $\mathbf{R}_{\mathbf{y}_j}$ by $\mathbf{\Phi}$

$$\mathbb{R}_{\mathbf{y}_j} \triangleq \mathbf{R}_{\mathbf{r}_j} \oslash \mathbf{\Phi} = \sum_{i=1}^{M_t} \mathbf{H}_{j,i}\mathbf{H}_{j,i}^H + \sigma_n^2\mathbf{I} \tag{4.8}$$

Note that the noise correlation matrix $E\left\{\mathbf{n}_j(k)\mathbf{n}_j^H(k)\right\} = \sigma_n^2\mathbf{I}$ keeps unchanged after the element-by-element division since all the diagonal entries of $\mathbf{\Phi}$ are ones.

Define the $N_c \times M_t$ matrix $\mathbf{H}_j$ as

$$\mathbf{H}_j \triangleq \begin{bmatrix} \mathbf{H}_{j,1} & \cdots & \mathbf{H}_{j,M_t} \end{bmatrix} \tag{4.9}$$

Then Eqn.(4.8) can be rewritten as

$$\mathbb{R}_{\mathbf{y}_j} = \mathbf{H}_j\mathbf{H}_j^H + \sigma_n^2\mathbf{I} \tag{4.10}$$

In fact, matrix $\mathbb{R}_{\mathbf{y}_j}$ forms the outer-product of the channel matrix $\mathbf{H}_j$ [42]. Hence, the singular-value decomposition (SVD) of $\mathbb{R}_{\mathbf{y}_j}$ can be used to estimate the channel matrix [7]. Let the SVD of $\mathbb{R}_{\mathbf{y}_j}$ given in Eqn.(4.10) be denoted as

$$\mathbb{R}_{\mathbf{y}_j} = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \left(\begin{bmatrix} \mathbf{\Lambda}_s & \\ & \mathbf{0} \end{bmatrix} + \sigma_n^2\mathbf{I} \right) \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{4.11}$$

Then $\mathbf{H}_j$ can be estimated by

$$\hat{\mathbf{H}}_j = \mathbf{U}_s\mathbf{\Lambda}_s^{\frac{1}{2}} = \mathbf{H}_j\mathbf{Q}_j \tag{4.12}$$

where $\mathbf{Q}_j$ is the constant unitary ambiguity matrix. The ambiguity matrix indeed exists in all kinds of blind channel estimators and can be remedied by introducing extra constraints, e.g., using the blind source separation methods proposed in [119, 120]. Thus, we assume that the unitary ambiguity matrix is known exactly in the rest of this chapter. Later in Chapter 5, we discuss the problem of blind source separation to remove this ambiguity matrix.

By repeating the above procedures to each receive antenna, all the channels can be estimated up to different unitary ambiguity matrices, which need to be "synchronized". This is very inefficiency. To avoid this, we introduce another method

to estimate the remaining channels so that this "synchronizing" requirement can be avoided.

Consider the cross-correlation matrix of the noisy signals received by the tth and the jth receive antenna respectively, for $t = 1, \cdots, M_r$ and $t \neq j$, we have

$$\begin{aligned} \mathbf{R}_{\mathbf{y}_{t,j}} &\triangleq E\left\{\mathbf{y}_t(k)\mathbf{y}_j^H(k)\right\} \\ &= \sum_{i=1}^{M_t} \mathcal{D}(\mathbf{H}_{t,i})\mathbf{A}\mathbf{A}^H\mathcal{D}^H(\mathbf{H}_{j,i}) \\ &= \sum_{i=1}^{M_t} \mathbf{H}_{t,i}\mathbf{H}_{j,i}^H \odot \mathbf{\Phi} \end{aligned} \tag{4.13}$$

Similar to Eqn.(4.8), we perform the element-by-element division of $\mathbf{R}_{\mathbf{y}_{t,j}}$ by $\mathbf{\Phi}$, hence we get

$$\mathbb{R}_{\mathbf{y}_{t,j}} \triangleq \mathbf{R}_{\mathbf{y}_{t,j}} \oslash (\mathbf{P}\mathbf{P}^H) = \mathbf{H}_t\mathbf{H}_j^H \tag{4.14}$$

Multiply $\mathbb{R}_{\mathbf{y}_{t,j}}$ with the pseudo-inverse of $\hat{\mathbf{H}}_j^H$, where $\hat{\mathbf{H}}_j$ is the estimated channel matrix associated with the jth receive antenna, which is given by Eqn.(4.12). Thus we have

$$\hat{\mathbf{H}}_t = \mathbb{R}_{\mathbf{y}_{t,j}}(\hat{\mathbf{H}}_j^H)^\dagger = \mathbf{H}_t\mathbf{H}_j^H(\mathbf{Q}_j^H\mathbf{H}_j^H)^\dagger = \mathbf{H}_t\mathbf{Q}_j \tag{4.15}$$

By repeating the above procedure to all the $M_r - 1$ remaining receive antennae, the channel matrices can be estimated up to the same unitary matrix $\mathbf{Q}_j$. Since all of the channel matrices associated with the remaining receive antennae are estimated based on the previously estimated matrix $\mathbf{H}_j$, the jth receive antenna is therefore referred to as the "reference" antenna. Meanwhile, we name the other receive antennae as "normal" antenna.

In practice, the auto and cross correlation matrices are estimated by the time-domain average of the received symbol blocks, i.e.,

$$\hat{\mathbf{R}}_{\mathbf{y}_j} = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{y}_j(k)\mathbf{y}_j^H(k) \tag{4.16a}$$

$$\hat{\mathbf{R}}_{\mathbf{y}_{tj}} = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{y}_t(k)\mathbf{y}_j^H(k) \tag{4.16b}$$

where K is the total number of OFDM symbols collected to calculate the auto- and cross- correlation matrices. It is evident that the above estimate of the correlation matrices converges in the mean-square sense to $\mathbf{R}_{\mathbf{y}_j}$ and $\mathbf{R}_{\mathbf{y}_{tj}}$ respectively. Thus, in practice, the channel estimation is obtained by the above procedures after substituting $\mathbf{R}_{\mathbf{y}_j}$ and $\mathbf{R}_{\mathbf{y}_{tj}}$ with $\hat{\mathbf{R}}_{\mathbf{y}_j}$ and $\hat{\mathbf{R}}_{\mathbf{y}_{tj}}$ respectively.

Eqn.(4.15) also implies that for any subcarrier, the ambiguity matrices are all identical. Hence, it is not difficult to verify that the kth block of equalized signal on the nth subcarrier is $\hat{\mathbf{s}}(k,n) = \mathbf{Q}\mathbf{s}(k,n)$. Thus, the transmitted signal can be recovered by applying any source separation algorithm. Note that we only need to identify $\mathbf{Q}$ for one out of N_c subcarriers, since $\mathbf{Q}$ is identical for all subcarriers.

Therefore we can summarize our estimation algorithm as follows

1. Select a receive antenna j and calculate the auto-correlation matrix $\mathbf{R}_{\mathbf{y}_j}$ by Eqn.(4.16a).

2. Perform the element-by-element division to $\mathbf{R}_{\mathbf{y}_j}$ by $\mathbf{\Phi} \triangleq \mathbf{P}\mathbf{P}^H$ as Eqn.(4.8), and get the matrix $\mathbb{R}_{\mathbf{y}_j}$.

3. Apply SVD to $\mathbb{R}_{\mathbf{y}_j}$, and get the estimation of the channel matrix associated with the jth receive antenna, $\hat{\mathbf{H}}_j$, (up to an unitary matrix $\mathbf{Q}_j$) as Eqn.(4.12).

4. Check whether $\hat{\mathbf{H}}_j$ is of full column rank. If yes, continue; otherwise, choose another receive antenna, and go to step 1.

5. Calculate the cross correlation matrix $\mathbf{R}_{\mathbf{y}_{t,j}}$ and hence $\mathbb{R}_{\mathbf{y}_{t,j}}$ for all $t \neq j$.

6. Estimate $\mathbf{H}_t$ (up to the same unitary matrix $\mathbf{Q}_j$) by right multiplying $\mathbb{R}_{\mathbf{y}_{t,j}}$ with the pseudo inverse of $\hat{\mathbf{H}}_j^H$ as Eqn.(4.15).

4.4 Discussion

4.4.1 Identifiability

The proposed algorithm can identify any channel up to a unitary ambiguity matrix as long as $\mathbf{\Phi}$ has unit diagonal entries and no zero entries, and $\mathbf{H}_j$ is of full column rank. Note that we need only one out of M_r channel matrices to be of full column rank. Once $\mathbf{H}_j$ is estimated, the other $\mathbf{H}_t$ $(t \neq j)$ can be estimated regardless of the rank of $\mathbf{H}_t$. The uniqueness of the proposed estimator is guaranteed by the uniqueness of the SVD. Moreover, the traditional blind channel estimation algorithms for MIMO- or SIMO-OFDM usually require the channel transfer functions do not share common zeros at the subcarrier frequencies [109]. However, this assumption may not be satisfied necessarily. On the other hand, according to the uniqueness of the SVD, the identifiability of our proposed algorithm is guaranteed even when this assumption is not fulfilled.

4.4.2 Precoder Design

As discussed in Section 4.2, the precoding matrix $\mathbf{P}$ should satisfy that $\mathbf{\Phi} \triangleq \mathbf{P}\mathbf{P}^H$ has unit diagonal elements and no zero elements. Assume $\mathbf{P}$ is an arbitrary $M \times M$ full rank and symmetric matrix with unit diagonal elements. Denote the SVD of $\mathbf{\Phi}$ as

$$\mathbf{\Phi} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H \tag{4.17}$$

Since $\mathbf{\Phi}$ is of full rank, then $\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{V}^H$ is also full rank. Thus, the precoding matrix can be designed as

$$\mathbf{P} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{V}^H \tag{4.18}$$

Without loss of generality, we can enforce all of the non-diagonal elements of $\mathbf{\Phi}$ being the real number ϕ, $(\phi \neq 1)$, i.e.,

$$\mathbf{\Phi} = \begin{bmatrix} 1 & \phi & \cdots & \phi \\ \phi & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi \\ \phi & \cdots & \phi & 1 \end{bmatrix} \tag{4.19}$$

In this situation, the precoding matrix $\mathbf{P}$ is a circulant matrix. It should be noted that the precoder proposed in this section is only one of the possible precoders. The optimization of the precoder design is yet to be studied.

4.4.3 SNR Analysis

In this subsection, we focus on the SNR degradation caused by the precoder. We assume that the channel is perfectly estimated, while the signal is detected by the

zero forcing criterion. In the conventional MIMO-OFDM system without precoding, the kth detected OFDM symbol is given by

$$\hat{\mathbf{d}}(k) = \mathbf{d}(k) + [\mathcal{D}(\mathbf{H})]^{\dagger}\,\mathbf{n}(k) = \mathbf{d}(k) + \mathbf{e}(k) \tag{4.20}$$

where the channel matrix $\mathbf{H}$ and the block diagonal matrix $\mathcal{D}(\mathbf{H})$ are defined by Eqn.(2.73) and Eqn.(2.67) in Chapter 2 respectively, and $\mathbf{e}(k) = [\mathcal{D}(\mathbf{H})]^{\dagger}\,\mathbf{n}(k)$ is the detection error. According to the structure of $\mathbf{n}(k)$ and $\mathbf{P}$, $\mathbf{e}(k)$ can be written as

$$\mathbf{e}(k) = \left[\ \mathbf{e}^T(k,0),\ \ \cdots,\ \ \mathbf{e}^T(k,N_c-1)\ \right]^T \tag{4.21}$$

where

$$\mathbf{e}(k,n) = \left[\ e_1(k,n),\ \ \cdots,\ \ e_{M_t}(k,n)\ \right]^T \tag{4.22}$$

and $e_i(k,n)$ is the detection error of element the kth OFDM symbol associated with the nth subcarrier and the ith transmit antenna. Define

$$\mathbf{e}_i(k) = \left[\ e_i(k,0),\ \ \cdots,\ \ e_i(k,N_c-1)\ \right]^T \tag{4.23}$$

Substitute to Eqn.(4.20), then the detected signal of the ith user can be written as

$$\hat{\mathbf{d}}_i(k) = \mathbf{d}_i(k) + \mathbf{e}_i(k) \tag{4.24}$$

Thus, the SNR of signal transmitted on the nth subcarrier through the ith transmit antenna is given by

$$\mathrm{SNR_u}(i,n) = \frac{E\{|d_i(k,n)|^2\}}{E\{|e_i(k,n)|^2\}} = \frac{\beta_i(n)}{\gamma_i(n)} \tag{4.25}$$

On the other hand, the detected and decoded signal in the precoded MIMO-OFDM system is

$$\hat{\mathbf{d}}_i(k) = \mathbf{d}_i(k) + \mathbf{P}^{-1}\mathbf{e}_i(k) \tag{4.26}$$

Hence, the SNR is modified as

$$\begin{aligned}\mathrm{SNR_c}(i,n) &= \frac{E\{|d_{i,n}(k)|^2\}}{\sum_{m=0}^{N_c-1}\left(|p'_{n,m}|^2 E\{|e_{i,m}(k)|^2\}\right)} \\ &= \frac{\beta(i,n)}{\sum_{m=0}^{N_c-1}\left(|p'_{n,m}|^2 \gamma(i,m)\right)}\end{aligned} \tag{4.27}$$

where $p'_{n,m}$ is the $(n,m)^{th}$ element of $\mathbf{P}^{-1}$.

Lemma 4.4.1

$$\frac{\min_n \mathrm{SNR_u}(i,n)}{\varphi_n} \leqslant \mathrm{SNR_c}(i,n) \leqslant \frac{\max_n \mathrm{SNR_u}(i,n)}{\varphi_n} \tag{4.28}$$

where φ_n is the nth diagonal element of $(\mathbf{P}^H\mathbf{P})^{-1}$.

Proof: One can verify that $\sum_{m=0}^{N_c-1}|p'_{n,m}|^2 = \varphi_n$, which is the nth diagonal element of $(\mathbf{P}^H\mathbf{P})^{-1}$, thus we have

$$\varphi_n\gamma_{\min}(i,n) \leqslant \sum_{m=0}^{N_c-1}\left(|p'_{n,m}|^2\gamma(i,m)\right) \leqslant \varphi_n\gamma_{\max}(i,n) \tag{4.29}$$

Substitute Eqn.(4.29) to Eqn.(4.27), we have

$$\frac{\beta(i,n)}{\varphi_n\gamma_{\max}(i,n)} \leqslant \mathrm{SNR_c}(i,n) \leqslant \frac{\beta(i,n)}{\varphi_n\gamma_{\min}(i,n)} \tag{4.30}$$

which is equivalent to Eqn.(4.28). ■

Particularly, if the circulant precoding matrix which satisfies Eqn.(4.19) proposed in Section 4.4.2 is used, then it can be verified that

$$\varphi_n = \frac{1}{N_c}\left[\frac{1}{1+(N_c-1)\phi} + \frac{N_c-1}{1-\phi}\right] \tag{4.31}$$

for $n = 0, \cdots, N_c-1$. Since φ_n is a function of ϕ, then the SNR and hence the BER performance can be controlled by carefully selecting ϕ.

4.4.4 Asymptotic Performance Analysis

We derive the asymptotic performance of the proposed channel estimator based on the first order perturbation theory of SVD [88]. From Eqn.(4.11) we have the following equation

$$\begin{aligned} \mathbf{H}_j\mathbf{H}_j^H &= \mathbb{R}_{\mathbf{y}_j} - \sigma_n^2\mathbf{I} \\ &= \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_n \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_s & \\ & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \end{aligned} \tag{4.32}$$

Define the noisy estimates of the signal and noise subspace of $\mathbb{R}_{\mathbf{y}_j}$ as follows respectively

$$\hat{\mathbf{U}}_s \triangleq \mathbf{U}_s + \Delta\mathbf{U}_s \tag{4.33a}$$

$$\hat{\mathbf{U}}_n \triangleq \mathbf{U}_n + \Delta\mathbf{U}_n \tag{4.33b}$$

where $\Delta\mathbf{U}_s$ and $\Delta\mathbf{U}_n$ are the perturbation in the estimated signal and noise subspace respectively. Hence we obtain

$$\begin{aligned} \mathbf{H}_j\mathbf{H}_j^H + \Delta(\mathbf{H}_j\mathbf{H}_j^H) &= \hat{\mathbb{R}}_{\mathbf{y}_j} - \sigma_n^2\mathbf{I} \\ &= \begin{bmatrix} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{\Lambda}}_s & \\ & \Delta\mathbf{\Lambda}_n \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{bmatrix} \end{aligned} \tag{4.34}$$

where $\Delta(\mathbf{H}_j\mathbf{H}_j^H) = \Delta\mathbb{R}_{\mathbf{y}_j}$ denotes the perturbation of $\mathbf{H}_j\mathbf{H}_j^H$ due to the limited number of observation symbols and the AWGN noise. In the noise free case, we have $\Delta\mathbf{\Lambda}_n = 0$.

According to Lemma 3.4.2, the first order perturbation expansion of $\Delta\mathbf{U}_s$ and

$\Delta\mathbf{U}_n$ can be expressed as a linear approximation form, i.e.,

$$\hat{\mathbf{U}}_s \triangleq \mathbf{U}_s + \mathbf{U}_n\mathbf{T} \tag{4.35a}$$

$$\hat{\mathbf{U}}_n \triangleq \mathbf{U}_n + \mathbf{U}_s\mathbf{W} \tag{4.35b}$$

Left multiplying both side of Eqn.(4.34) by $\hat{\mathbf{U}}_n^H$, we have

$$\begin{aligned}\hat{\mathbf{U}}_n^H\left[\mathbf{H}_j\mathbf{H}_j^H + \Delta(\mathbf{H}_j\mathbf{H}_j^H)\right] &= \hat{\mathbf{U}}_n^H(\hat{\mathbf{U}}_s\hat{\mathbf{\Lambda}}_s\hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n\Delta\mathbf{\Lambda}_n\hat{\mathbf{U}}_n^H) \\ &= \Delta\mathbf{\Lambda}_n\hat{\mathbf{U}}_n^H \end{aligned} \tag{4.36}$$

The second equality follows from the fact the $\hat{\mathbf{U}}_s \perp \hat{\mathbf{U}}_n$ and $\hat{\mathbf{U}}_n^H\hat{\mathbf{U}}_n = \mathbf{I}$. By substituting Eqn.(4.35a) and Eqn.(4.35b) into Eqn.(4.36), we get

$$(\mathbf{U}_n + \mathbf{U}_s\mathbf{W})^H(\mathbf{\Xi} + \Delta\mathbf{\Xi}) = \Delta\mathbf{\Lambda}_n(\mathbf{U}_n + \mathbf{U}_s\mathbf{W})^H \tag{4.37}$$

Neglect the second-order terms and use the fact that $\mathbf{U}_n^H\mathbf{H}_j = \mathbf{0}$, then we get

$$\mathbf{W} \doteq -\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H\left[\Delta(\mathbf{H}_j\mathbf{H}_j^H)\right]^H\mathbf{U}_n \tag{4.38}$$

and consequently,

$$\Delta\mathbf{U}_n \approx -\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H\left[\Delta(\mathbf{H}_j\mathbf{H}_j^H)\right]^H\mathbf{U}_n \tag{4.39}$$

Using the orthogonality between the perturbed signal and noise subspace, we have

$$\hat{\mathbf{U}}_n^H\hat{\mathbf{U}}_s = (\mathbf{U}_n + \mathbf{U}_s\mathbf{W})^H(\mathbf{U}_s + \mathbf{U}_n\mathbf{T}) = \mathbf{0} \tag{4.40}$$

Since $\mathbf{U}_n^H\mathbf{U}_s = \mathbf{0}$, the above equation can be simplified as

$$\mathbf{T} = -\mathbf{W}^H \tag{4.41}$$

which directly leads to the result that

$$\Delta\mathbf{U}_s = \mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{V}_s\mathbf{\Lambda}_s^{-1} \tag{4.42}$$

Since the channel is estimated up to an unitary ambiguity matrix, then it is not meaningful to compare the estimated channel $\hat{\mathbf{H}}_j$ with the true channel $\mathbf{H}_j$ directly. Instead, we compare $\hat{\mathbf{H}}_j$ and $\tilde{\mathbf{H}}_j$, where

$$\tilde{\mathbf{H}}_j \triangleq \mathbf{H}_j \mathbf{Q}^{-1} \tag{4.43}$$

and $\mathbf{Q}$ is the ambiguity matrix, which is assumed to be known. Define the channel estimation error associated with the jth receive antenna as

$$\Delta\mathbf{H}_j \triangleq \hat{\mathbf{H}}_j - \tilde{\mathbf{H}}_j \tag{4.44}$$

Next we consider the bias and the mean-square error (MSE) of the proposed estimator, which are defined as follows respectively,

$$\text{Bias} \triangleq E\{\Delta\mathbf{H}_j\} \tag{4.45a}$$

$$\text{MSE} \triangleq E\{\|\Delta\mathbf{H}_j\|_F^2\} = \text{tr}\left(E\{\Delta\mathbf{H}_j\mathbf{H}_j^H\}\right) \tag{4.45b}$$

The following Theorem 4.4.2 and Theorem 4.4.3 settles the asymptotical performance analysis of the "reference" and "normal" channel matrices respectively.

Theorem 4.4.2 The proposed blind channel estimator of the "reference" channel matrix, $\hat{\mathbf{H}}_j$, is asymptotically unbiased (i.e. $E\{\Delta\mathbf{H}_j\} = \mathbf{0}$ and the estimated channel MSE is

$$\text{MSE} = \frac{1}{K}\sum_{i=1}^{M_t} \text{tr}\left(\mathbf{U}_n^H \mathcal{E}_{i,i} \mathbf{U}_n\right) \tag{4.46}$$

Proof: See Appendix B. ■

Theorem 4.4.3 The proposed blind channel estimator of the "normal" channel matrices, $\hat{\mathbf{H}}_t$ $(t = 1, \cdots, M_t,\ t \neq j)$, is asymptotically unbiased (i.e. $E\{\Delta\mathbf{H}_t\} = \mathbf{0}$),

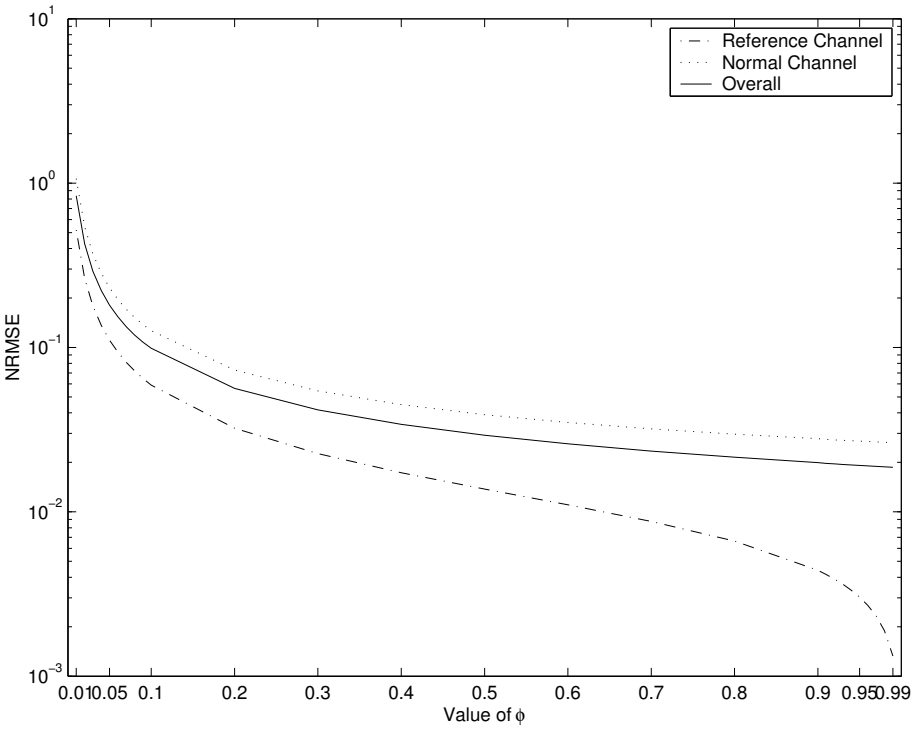

Figure 4.2: NRMSE as a function of parameter ϕ

and the estimated channel MSE is

$$\text{MSE} = \frac{1}{K}\sum_{i=1}^{M_t} \text{tr}\left\{\ddot{\boldsymbol{\Xi}}_i + \boldsymbol{\Psi}_i + \boldsymbol{\Upsilon}_i + \boldsymbol{\Omega}_i\right\} \tag{4.47}$$

where the notations are defined in Appendix C accordingly.

Proof: See Appendix C. ■

The above two theorems figure out the theoretical mean-square error (MSE) of the proposed channel estimator, which is related to the choice of the precoder. The mathematical expression of these two theorems may not be intuitive. Therefore, we present Fig. 4.2, which illustrates the theoretical normalized-mean-square-error (NMSE) [110] as the function of ϕ in a noise free 2I2O-OFDM system. It is shown that when ϕ is closer to 1, the estimation performance is better. This is not strange because ϕ is the cross correlation of source data carried by different subcarriers.

When they are more correlated, the more redundant information is brought in. However, this also causes the SNR degradation. Hence, we need to balance them.

4.5 Simulation Results

In this section, we provide some simulation results to illustrate the performance of the proposed estimator. The simulated OFDM system is modeled containing 64 subcarriers, i.e. $M = 64$. Each OFDM symbol consists of 68 elements including the guard interval of length 4, i.e. $M_g = 4$. The system is equipped with 2 transmit antennas and 2 receive antennas, and the channel model used is a 3-tap FIR filter with tap coefficients independently chosen from a white Gaussian process. As discussed, the proposed algorithm is suitable for both CP and ZP based OFDM system. Therefore, both of these two kinds of OFDM systems are simulated, and their performances are both illustrated for comparison. Additionally, we also simulate a training based LS channel estimation algorithm.

To evaluate the channel estimation error, we employed the normalized-root-mean-square-error (NRMSE) [110], which is defined as

$$\text{NRMSE} = \sqrt{\frac{1}{N_1 N_2} \sum_{t_1=1}^{N_1} \sum_{t_2=1}^{N_2} \frac{\|\hat{\mathbf{H}}^{(t_1,t_2)} - \mathbf{H}^{(t_1)}\|_F^2}{\|\mathbf{H}^{(t_1)}\|_F^2}} \tag{4.48}$$

where $\|.\|_F$ denotes the Frobenius norm, N_1 and N_2 are the number random channels and the number of Monte Carlo runs for each simulated channel respectively. $\mathbf{H}^{(t_1)}$ is the true channel matrix, and $\hat{\mathbf{H}}^{(t_1,t_2)}$ is the estimation of $\mathbf{H}^{(t_1)}$ for the t_2^{th} Monte Carlo run.

Figure 4.3 and Figure 4.4 illustrate the NRMSE and BER performance as functions of SNR respectively. We simulate 30 independent channels and each for 100

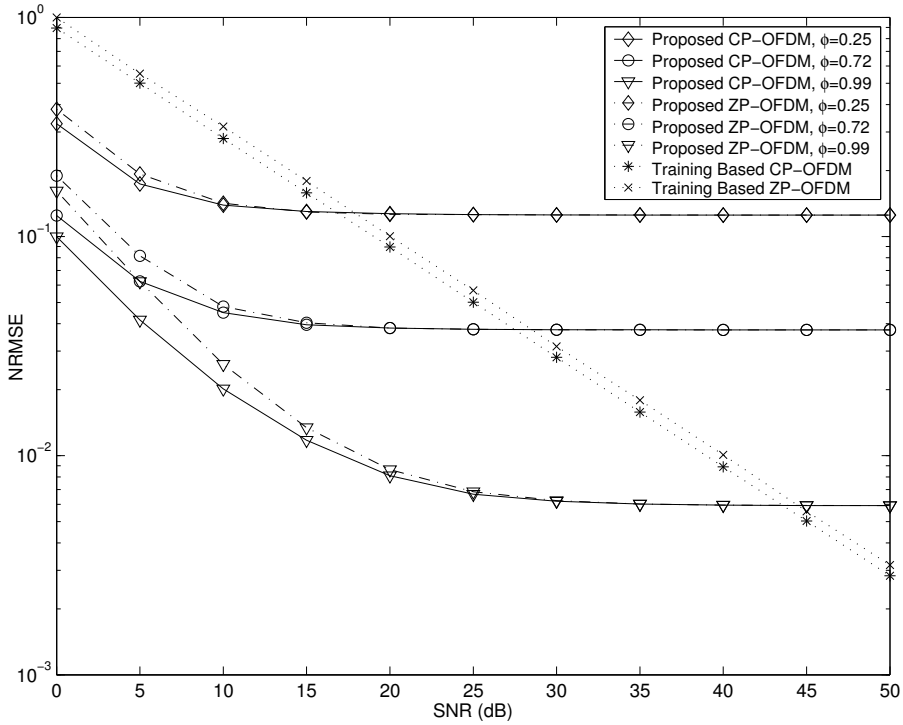

Figure 4.3: NRMSE performance as a function of SNR

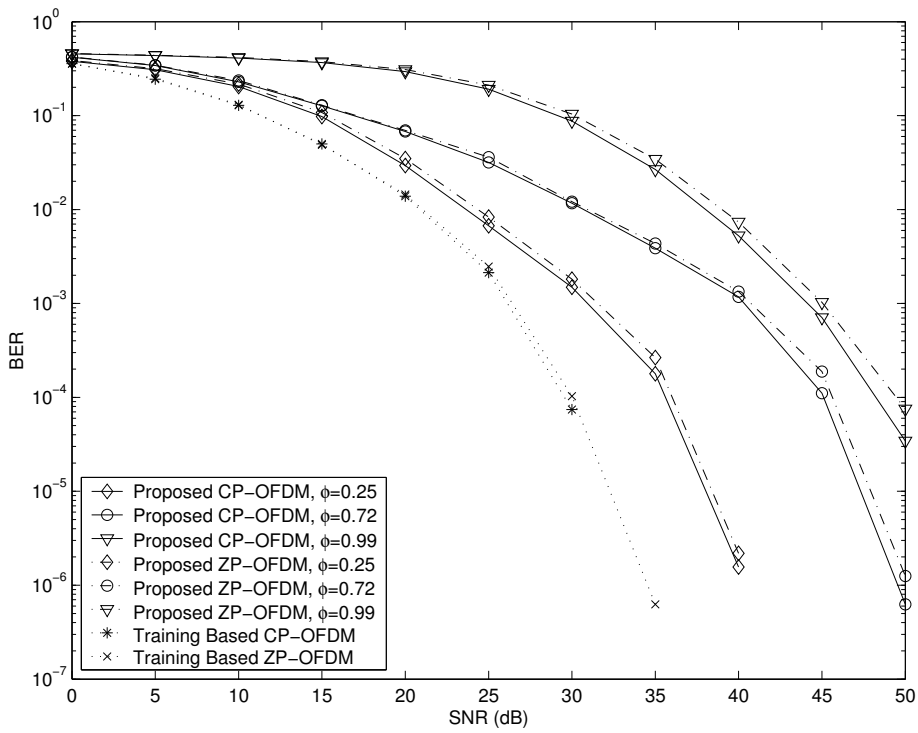

Figure 4.4: BER performance as a function of SNR

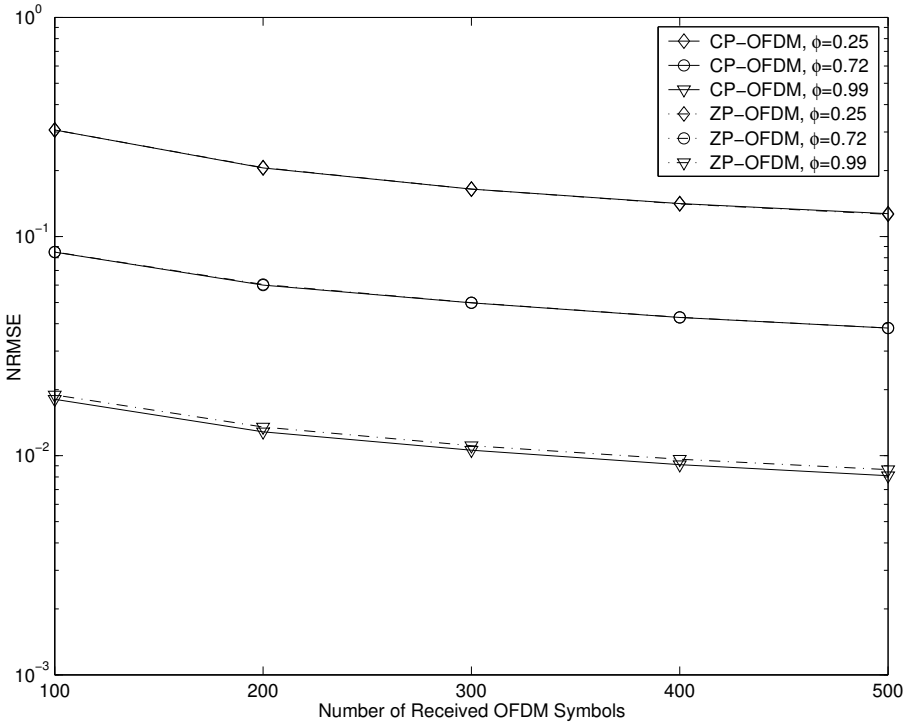

Figure 4.5: NRMSE performance as a function of NOS

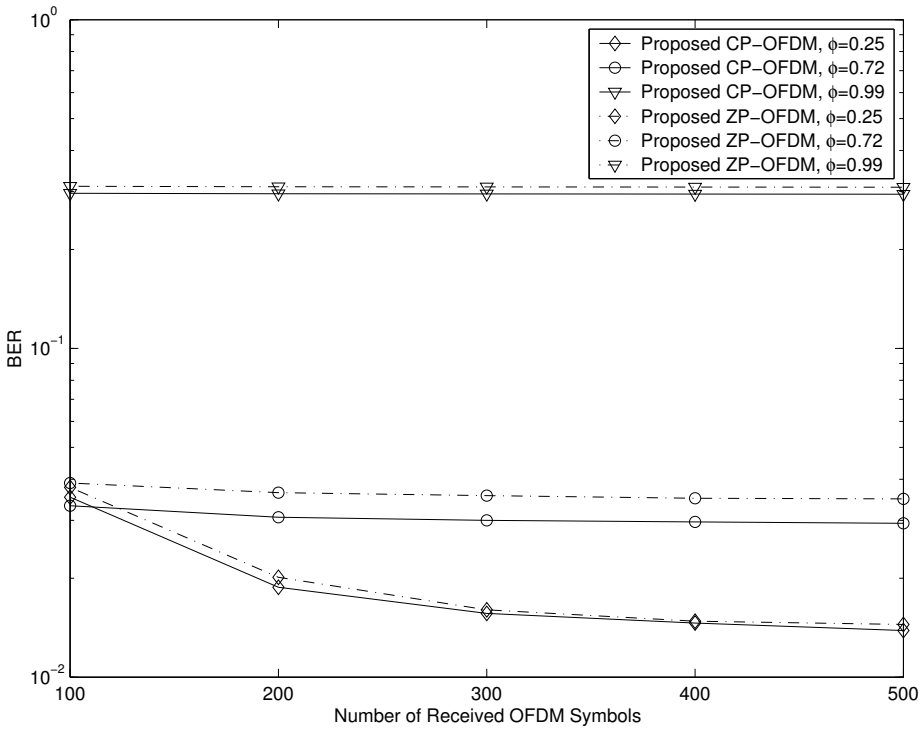

Figure 4.6: BER performance as a function of NOS

Monte Carlo runs. The number of the observed OFDM symbols (NOS) is fix to be 500. From Figure 4.3 We can see that our proposed estimator can achieve a lower NRMSE than the training-based LS algorithm at moderate or low SNR. The figure also indicates that NRMSE performance of the proposed method can be controlled by carefully selecting ϕ, as discussed in the previous section. On the other hand, as shown in Figure 4.4, the proposed estimation algorithm can achieve a BER performance close to the LS estimator if ϕ is small. As discussed in Section 4.4.4, when ϕ is small enough, the SNR degradation caused by the precoding is suppressed. Thus, the proposed BER performance can improved. However, this will cause the NRMSE performance degradation. Thus, we need to balance them.

As discussed, the auto-correlation and cross-correlation matrices, $\hat{\mathbf{R}}_{\mathbf{y}_j}$ and $\hat{\mathbf{R}}_{\mathbf{y}_{tj}}$, are estimated by Eqn.(4.16a) and Eqn.(4.16b) respectively, where the estimation accuracy depends on the number of observed OFDM symbols (NOS). Figure 4.5 and Figure 4.6 illustrate the NRMSE and BER performances as functions of NOS respectively. From these two figures, we see that the proposed estimation algorithm achieves better performance when NOS is increased. The reason is obvious. When NOS is larger, the estimation of $\hat{\mathbf{R}}_{\mathbf{y}_j}$ and $\hat{\mathbf{R}}_{\mathbf{y}_{tj}}$ are more accurate, which leads to a better performance. Meanwhile, as same as the last two figures, Eqn.(4.16a) and Eqn.(4.16b) prove again that ϕ is a key parameter which influences the NRMSE and BER performance of the proposed estimator.

In Figure 4.7, we examine the NRMSE performance of the "reference" and "normal" channel matrices for CP based MIMO-OFDM systems. The overall performance is also shown. We select the parameter of the precoding matrix to be $\phi = 0.72$,

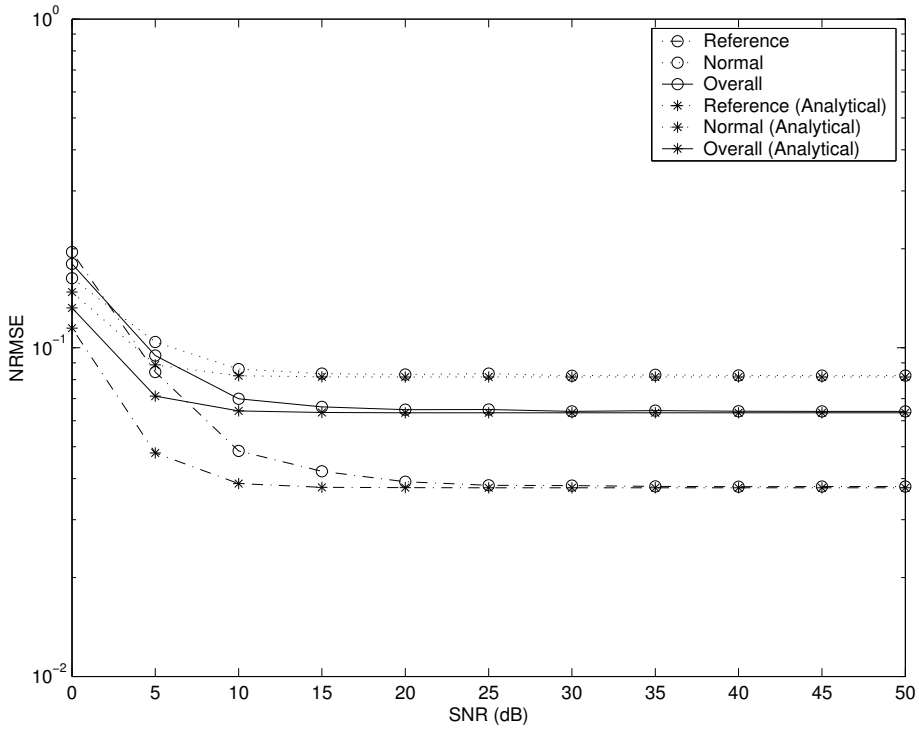

Figure 4.7: NRMSE performance of "reference" and "normal" channels as functions of SNR

and the other parameter settings are the same as those for Figure 4.2. It can be seen from the figure that when SNR is high, the computer experiment result is close to the theoretical performance proposed in Theorem 4.4.2 Theorem 4.4.3. The figure also indicates that the estimation error of "normal" channel matrix is higher than that of the "reference" channel matrix. The reason is obvious. Since the "normal" channel matrix is estimated based on the estimated "reference" channel matrix, the estimation error of the "reference" channel is thus accumulated into the error of the "normal" channel.

4.6 Summary

We presented a novel blind channel estimation method for MIMO-OFDM system where the source data is linearly precoded. With the assistance of a nonredundant linear precoder, the channel can be estimated blindly by exploiting the correlation matrix of the received signal. The proposed algorithm can identify the channel even when the channel transfer functions share zeros at subcarrier frequencies. Simulations show that the proposed algorithm compares favorably to the training based LS algorithm in both NRMSE and BER performance. The performance of the proposed algorithm could be further improved by optimizing the precoder, which is an open question.

Chapter 5

A Geometric Method for BSS of Digital Signals with Finite Alphabets

5.1 Introduction

Blind deconvolution is a problem of considerable interest in diverse fields including seismology, radio astronomy, underwater acoustics [118]. The goal of blind deconvolution is the recovery of signals transmitted through an unknown channel based solely on the channel's output without access to its input. The two most common applications of blind deconvolution in communication systems are channel identification and source separation. The former deals with the recovery of the signal with the inter-symbol interference (ISI) caused by the channel distortion. The latter deals with the simultaneous recovery of a number of signals with inter-user interference

(IUI) that are caused by the multiple antenna transmission.

In the previous chapters, the problem of blind channel identification is discussed, and three blind channel estimation methods are proposed for the MIMO-OFDM systems. It is shown that the first two methods can blindly estimate the MIMO-OFDM channels up to an invertable ambiguity matrix which need to be further removed. Unfortunately, this ambiguity matrix can not be directly removed by the proposed estimation algorithms. In fact, this is a common problem in many of the existing blind MIMO channel estimation methods.

The problem of removing the ambiguity matrix can be referred to as the problem of blind source separation (BSS), which is discussed in this chapter. This BSS problem has been of considerable interest in wireless digital communications and other fields. Past work on this problem include [119–126]. Among them, [119, 122, 123] are iterative methods and suffer from local optima. They, usually, require a good initialization in order to minimize the problem of local minima. In [121], an analytical method based on a generalized eigenvalue decomposition was developed for the constant modulus sources. Some techniques that rely directly on HOS cumulants were introduced in [124–126]. In [124], an adaptive separation algorithm which is free of undesired stationary point for an arbitrary number of users was derived from a constrained multiuser kurtosis optimization criterion. However, the HOS technique often requires a large number of observation samples for the accuracy of the numerical result, and the computational cost is comparably large. Furthermore, the source signal must be non-Gaussian, and their kurtosis must have the same sign [128]. A geometric approach for the blind separation of instantaneous mixtures

of digital signals was proposed in [120]. However, this method is specific only to the BPSK signals.

In this chapter, we propose a geometric non-iterative method that separates signals with the M-ASK or QAM digital format. We focus on the non-iterative algorithm development for the whitened real case. We compare our proposed method to the hyperplane-based algorithm [122] which has been shown to be a fast algorithm with similar performance as iterative least squares with projection (ILSP) [119]. The kurtosis-based algorithm [124] is also simulated as a comparison. It is shown that our proposed algorithm achieves a lower SER that both [122] and [124].

5.2 Problem Formulation

Consider a general MIMO FIR system which is equipped with M_t and M_r transmit and receive antennae respectively. We assume that no intersymbol interference is present, and the transmitted signals are M-ASK or QAM digital signals with are spatially independent. Thus the array output vector is an instantaneous mixture of the M_t transmitted source signals. Hence, the array output vector $\mathbf{x}(n)$ can be written as

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{w}(n) \tag{5.1}$$

where

$$\mathbf{x}(n) \triangleq \begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_{M_r}(n) \end{bmatrix}^T \tag{5.2a}$$

$$\mathbf{s}(n) \triangleq \begin{bmatrix} s_1(n) & s_2(n) & \cdots & s_{M_t}(n) \end{bmatrix}^T \tag{5.2b}$$

$$\mathbf{w}(n) \triangleq \begin{bmatrix} w_1(n) & w_2(n) & \cdots & w_{M_r}(n) \end{bmatrix}^T \tag{5.2c}$$

where $\mathbf{s}(n)$ is the vector of symbols from the alphabet $\mathcal{S}$ generated by the M_t sources, $\mathbf{w}(n)$ is a vector of M_r dimensional additive noise, $\mathbf{H}$ is an $M_r \times M_t$ unknown instantaneous mixture matrix. To recover all source signals, it is assumed that $\mathbf{H}$ is full column rank. If we concatenate N snapshots of the received data as

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}(1) & \mathbf{x}(2) & \cdots & \mathbf{x}(N) \end{bmatrix} \tag{5.2b}$$

then we have

$$\mathbf{X} = \mathbf{HS} + \mathbf{W} \tag{5.2c}$$

where we define

$$\mathbf{S} \triangleq [\mathbf{s}(1) \ \cdots \ \mathbf{s}(N)] \tag{5.4a}$$

$$\mathbf{W} \triangleq [\mathbf{w}(1) \ \cdots \ \mathbf{w}(N)] \tag{5.4b}$$

Here we assume that N is large enough to satisfy the so-called "sufficient excitation condition", i.e. every combination vector of length M_t with elements from M-ASK or QAM alphabet $\mathcal{S}$ appears at least once in $\mathbf{S}$. Our objective is to recover $\mathbf{H}$ and $\mathbf{S}$ up to a permutation matrix and a diagonal matrix from the received data $\mathbf{X}$ only.

5.3 Proposed Source Separation Algorithm

5.3.1 Real Case: M-ASK Alphabets

We first consider the M-ASK signals, i.e.

$$\mathcal{S}_{\text{M-ASK}} = \{\pm 1, \pm 3, \cdots, \pm(M-1)\}$$

Also, we assume that the channel matrix $\mathbf{H}$ and the noise matrix $\mathbf{W}$ are real since the complex equation $\mathbf{X} = \mathbf{HS} + \mathbf{W}$ can be easily converted into the following real equation

$$\begin{bmatrix} \mathbf{X}^R \\ \mathbf{X}^I \end{bmatrix} = \begin{bmatrix} \mathbf{H}^R \\ \mathbf{H}^I \end{bmatrix} \mathbf{S} + \begin{bmatrix} \mathbf{W}^R \\ \mathbf{W}^I \end{bmatrix} \tag{5.5}$$

where the superscripts $[\cdot]^R$ and $[\cdot]^I$ denote real and imaginary parts of the matrices, respectively. We now enumerate the steps for our source separation algorithm.

Clustering

In the presence of channel noise $\mathbf{w}(n)$, the observed data constellation is a union of clusters centered around the points $\tilde{\mathbf{x}}_i = \mathbf{H}\tilde{\mathbf{s}}_i; i = 1, \ldots, d$, where d is the number of clusters. Without loss of generality, the channel matrix $\mathbf{H}$ is assumed to be full column rank, and hence the number of clusters is $d = M^p$. Define $\mathbf{S}_d \triangleq [\tilde{\mathbf{s}}_1 \ \cdots \ \tilde{\mathbf{s}}_d]$ which represents the $p \times d$ matrix containing exactly d distinct column vectors with elements from the M-ASK alphabet. The cluster centers $\tilde{\mathbf{x}}_i$ can be extracted by using the unsupervised clustering algorithms such as the Neural gas algorithm [127] and the smallest distance clustering algorithm [120] (See [127] for a comprehensive treatment of unsupervised clustering methods). If we concatenate the extracted

cluster vectors $\tilde{\mathbf{x}}_i$ as

$$\mathbf{X}_d \triangleq [\tilde{\mathbf{x}}_1 \ \cdots \ \tilde{\mathbf{x}}_d] \tag{5.6}$$

then, theoretically, the matrix has the form:

$$\mathbf{X}_d = \mathbf{H}\mathbf{S}_d \tag{5.7}$$

Whitening

We now whiten the data set $\mathbf{X}_d$ by utilizing the following easily verified property: $\mathbf{S}_d\mathbf{S}_d^T = K_r\mathbf{I}$, where $K_r = 2M^{(p-1)}(1^2 + 3^2 + \cdots + (M-1)^2)$. Let $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ be the singular value decomposition (SVD) of $\mathbf{H}$, thus we have

$$\frac{\mathbf{X}_d\mathbf{X}_d^T}{K_r} = \mathbf{H}\mathbf{H}^T = \mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^T = \bar{\mathbf{U}}\Sigma\bar{\mathbf{U}}^T \tag{5.8}$$

where $\bar{\mathbf{U}}$ denotes the submatrix of $\mathbf{U}$ from 1st column to p^{th} column, $\Sigma \triangleq \text{diag}(\sigma_1^2, \cdots, \sigma_p^2)$, σ_i denotes i^{th} singular value of $\mathbf{H}$. The whitening matrix is defined as $\mathbf{W} \triangleq \Sigma^{-\frac{1}{2}}\bar{\mathbf{U}}^T$. We can then form the whitened data set as

$$\mathbf{Z}_d = \mathbf{W}\mathbf{X}_d = \mathbf{W}\mathbf{H}\mathbf{S}_d = \mathbf{Q}\mathbf{S}_d \tag{5.9}$$

where $\mathbf{Q}$ is a $p \times p$ real unitary channel matrix to be determined.

Geometric Approach for Channel Estimation and Source Recovery

The objective of this step is to estimate $\mathbf{Q}$ from the whitened data set $\mathbf{Z}_d$. Let $\text{dis}(\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_j) \triangleq \|\tilde{\mathbf{z}}_i - \tilde{\mathbf{z}}_j\|$ denotes the Euclidean distance in $\mathbb{R}^p$ between two constellation points. It is clear that we have the following

$$\begin{aligned} \text{dis}(\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_j) &= \|\tilde{\mathbf{z}}_i - \tilde{\mathbf{z}}_j\| = \|\mathbf{Q}(\tilde{\mathbf{s}}_i - \tilde{\mathbf{s}}_j)\| = \|\tilde{\mathbf{s}}_i - \tilde{\mathbf{s}}_j\| \\ &= \text{dis}(\tilde{\mathbf{s}}_i, \tilde{\mathbf{s}}_j) \end{aligned} \tag{5.10}$$

Notice that for any $i \neq j$, $\text{dis}(\tilde{\mathbf{s}}_i, \tilde{\mathbf{s}}_j)$ is minimized if and only if $\tilde{\mathbf{s}}_i$ and $\tilde{\mathbf{s}}_j$ differ only in one bit by 2, i.e. $\tilde{\mathbf{s}}_i - \tilde{\mathbf{s}}_j = \pm 2\mathbf{e}_k, k \in \{1, \ldots, p\}$, where $\mathbf{e}_k$ denotes the unit vector with its kth entry equal to one, and its other entries equal to zero. Therefore, for each pair of $\{\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_j\}$ that minimizes $\text{dis}(\tilde{\mathbf{z}}_i, \tilde{\mathbf{z}}_j)$, we have

$$\tilde{\mathbf{z}}_i - \tilde{\mathbf{z}}_j = \mathbf{Q}(\tilde{\mathbf{s}}_i - \tilde{\mathbf{s}}_j) = \pm 2\mathbf{Q}\mathbf{e}_k \;\; k \in \{1, \ldots, p\} \tag{5.11}$$

Thus some column of the unitary matrix $\mathbf{Q}$ can be determined up to a sign. It is clear that for each constellation point $\tilde{\mathbf{z}}_i$, there exist p nearest neighboring vectors $\tilde{\mathbf{z}}_{j_k}; k = 1, \ldots, p$ that allow us to recover all p distinct columns of $\mathbf{Q}$ up to a sign and a permutation of the columns. This implies that the received data constellation geometry is very rich in information pertaining to the channel. In this case, it is desirable to find a way that can extract the channel information from the constellation geometry more accurately at a moderate or low SNR. Notice that we have $\|\tilde{\mathbf{z}}_i\| = \|\tilde{\mathbf{s}}_i\|$ and usually, the constellation points with maximum vector norm contains the highest signal power, and hence achieves the highest SNR. Thus, these points are less likely to be confused. Therefore we can summarize our geometric approach as follows

1. Choose 2^p vectors $\tilde{\mathbf{z}}_{i_k}, k = 1, \ldots, 2^p$ that have maximum vector norm.

2. Choose one vector from $\tilde{\mathbf{z}}_{i_k}, k = 1, \ldots, 2^p$ as a reference vector such that $\|\tilde{\mathbf{z}}_{i_{ref}}\| - \sqrt{p}(M-1)$ is minimal.

3. Choose p nearest neighboring vectors $\tilde{\mathbf{z}}_{nb_k}, k = 1, \ldots, p$ of $\tilde{\mathbf{z}}_{i_{ref}}$ from $\tilde{\mathbf{z}}_{i_k}, k = 1, \ldots, 2^p$ by computing the Euclidean distance between $\tilde{\mathbf{z}}_{i_{ref}}$ and $\tilde{\mathbf{z}}_{i_k}$.

4. The unitary matrix $\mathbf{Q}$ is then estimated as

$$\mathbf{Q}_e = \begin{bmatrix} \tilde{\mathbf{z}}_{i_{ref}} - \tilde{\mathbf{z}}_{nb_1} & \cdots & \tilde{\mathbf{z}}_{i_{ref}} - \tilde{\mathbf{z}}_{nb_p} \end{bmatrix} \tag{5.12}$$

   The column normalized $\mathbf{Q}_e$ is an estimate of $\mathbf{Q}$ up to a sign and a permutation of the columns.

5. The input symbols are estimated as $\mathbf{S}_e = \mathbf{Q}_e^{-1}\mathbf{W}\mathbf{X}$.

5.3.2 Extension to The Complex Case: QAM Alphabets

We now discuss the extension of our source separation algorithm to the QAM alphabets, i.e. $\mathcal{S}_{\text{QAM}} = \{\alpha + j\beta : \alpha, \beta \in \mathcal{S}_{\text{M-ASK}}\}$. The complex extension is described as follows. As for the QAM alphabets, we still have $\mathbf{S}_d\mathbf{S}_d^H = K_c\mathbf{I}$, where K_c is a constant implicitly determined by the alphabets and the number of sources p. Therefore the observed data set $\mathbf{X}_d$ can be whitened by following the same way as in the real case. We have $\mathbf{Z}_d = \mathbf{Q}\mathbf{S}_d$, where $\mathbf{Q}$ is a $p \times p$ complex unitary matrix (note that $\mathbf{H}$ is allowed to be a complex matrix). In order to apply the geometric approach presented in previous subsection, we transform the complex equation $\mathbf{Z}_d = \mathbf{Q}\mathbf{S}_d$ into the following real equation

$$\begin{bmatrix} \mathbf{Z}_d^R \\ \mathbf{Z}_d^I \end{bmatrix} = \begin{bmatrix} \mathbf{Q}^R & -\mathbf{Q}^I \\ \mathbf{Q}^I & \mathbf{Q}^R \end{bmatrix} \begin{bmatrix} \mathbf{S}_d^R \\ \mathbf{S}_d^I \end{bmatrix} \triangleq \mathbf{Q}_T \begin{bmatrix} \mathbf{S}_d^R \\ \mathbf{S}_d^I \end{bmatrix} \tag{5.13}$$

where $\mathbf{Q}_T$ is a $2p \times 2p$ real unitary matrix. Thus we have successfully converted the QAM source separation problem into M-ASK source separation problem and can further estimate $\mathbf{Q}_T$ by using our proposed geometric approach. The construction of $\mathbf{Q}$ from the estimated $\mathbf{Q}_T$ is detailed in [122].

5.4 Discussion

Our work can be considered as a further development of work [120] where the latter only permits BPSK signals. Both works are clustering-based and operate on a whitened data space. However, in contrast to the work [120] that is essentially an assignment algorithm, our work focuses on extracting the rich channel information hidden in the constellation geometry, as developed in [122]. This fundamental difference accounts for the wider applicability of our geometric approach. It has been shown that, in our work, the channel information is only related to the relative difference between two constellation points while irrespective of the exact positions of the constellation points. In contrast, when assigning the received constellation points to the corresponding source vectors, the constellation points themselves as well as their relationship with other points have to be considered. In other words, we conclude that the information needed for channel estimation is less than that needed for an appropriate assignment algorithm.

Computational Complexity

We consider the computational complexity of our proposed algorithm. The proposed algorithm involves three steps: clustering, whitening and geometric approach for channel identification and signal recovery. The clustering algorithm adopted in our simulations is the smallest distance clustering algorithm [120], which is self starting, uses each received data vector only once and works very well at moderately high SNR. The complexity for clustering is NC_p, where C_p are the computations required per iteration. Treating addition and multiplication equally, i.e. counting

the flops, we find that the computations required per iteration are $\mathcal{O}(qd^2)$. Thus the complexity for clustering is $\mathcal{O}(Nqd^2)$. The complexity of step 2 and step 3 are respectively dominated by the correlation matrix $\mathbf{X}_d\mathbf{X}_d^T$ and signal recovery $\mathbf{S}_e = \mathbf{Q}_e^{-1}\mathbf{W}\mathbf{X}$, which have complexity $\mathcal{O}(q^2d)$ and $\mathcal{O}(Nqp)$, respectively. Therefore, by combining all these steps, our proposed algorithm has an overall complexity $\mathcal{O}(Nqd^2 + q^2d + Nqp)$. Since $d = M^p$, the computational complexity is exponential with respect to p. This suggests that our proposed algorithm is practical for separating a small number of discrete sources. In fact, the computational complexity is the most prohibitive issue in almost all geometric methods. On the other hand, the hyperplane-based algorithm does not need to do the clustering. The cost for signal whitening is identical to our proposed algorithms, $\mathcal{O}(q^2d)$, while the cost for all the iterations is $\mathcal{O}(Iq^2d)$, where I is the total number of iterations till the global convergence. Hence, the total computational cost of the hyperplane-based algorithm is $\mathcal{O}(q^2d + Iq^2d + Nqp)$.

Applications in MIMO-OFDM Systems

Suppose we have the MIMO-OFDM system equipped with M_t and M_r transmit and receive antennae respectively, where $M_r \geq M_t$. The transmitted signals are M-ASK or QAM digital signals which are spatially independent. At the receiver side, the channel matrices are estimated by the proposed channel estimators, and the received signals are then equalized. According to the discussions in the previous chapters, the channel estimators proposed in Chapter 3 and 4 can blindly estimate the channel matrices up to an unitary ambiguity matrices, which need to be further removed. This is essentially the BSS problem.

Recall the system model of the MIMO-OFDM systems

$$\mathbf{y}(k,n) = \mathbf{H}(n)\mathbf{s}(k,n) + \mathbf{n}(n,k) \tag{5.14}$$

where k is the OFDM symbol index and n is the subcarrier index. According to the previous chapters, the proposed estimation algorithms can blind estimate the channel up to an unitary ambiguity matrix $\mathbf{Q}$, i.e.,

$$\hat{\mathbf{H}} = \mathbf{H}\mathbf{Q} \tag{5.15}$$

Obviously, it follows that

$$\hat{\mathbf{H}}(n) = \mathbf{H}(n)\mathbf{Q} \tag{5.16}$$

where $\mathbf{H}(n)$ is the $M_r \times M_t$ channel matrix associated with the nth subcarrier, and $\hat{\mathbf{H}}(n)$ is the corresponding estimated channel matrix. In other words, the estimated channel matrix according to all the N_c subcarriers have the identical ambiguity matrix. Since the equalization of the OFDM system is performed subcarrier by subcarrier, then the equalized signals according to different subcarriers have the identical mixture matrix $\mathbf{Q}$, i.e.,

$$\hat{\mathbf{s}}(k,n) = \mathbf{Q}\mathbf{s}(k,n) \quad , \quad \text{for} \quad n = 0, \cdots, N_c - 1 \tag{5.17}$$

where $\hat{\mathbf{s}}(k,n)$ is the kth block of equalized signal block according to n^{th} subcarrier. Therefore, we apply the BSS algorithm to the equalized signals associated with one subcarrier only, and the mixture matrix $\mathbf{Q}$ can be identified blindly.

5.5 Simulation Results

We now present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the iterative hyperplane-based algorithm

proposed in [122] and the kurtosis-based algorithm proposed in [124]. For the hyperplane-based and kurtosis-based algorithms, the gradient search may converge to local minima. The magnitude of the residual $\frac{1}{Nq}\|\mathbf{X} - \mathbf{H}_e\mathbf{S}_e\|_F^2$ is a good measure to test the converged solution of the gradient search [119]. For the cases where the residual is not reduced to the noise power level, we restart the gradient search until the residual is decreased to the noise power level. In our simulations, we consider $M_t = 2$ source signals drawn from the 4-ASK alphabet $\{-3, -1, 1, 3\}$ arriving at $M_r = 2$ sensors. The entries of the tested channel matrices are independently chosen from a white Gaussian process. We totally test 100 independent channels, with 300 Monte Carlo runs for each channel realization. The symbol error rate shown in the figure is the overall average of all the 300 runs. In each run, we collect 100 data samples the perform the separation process. Figure 5.1 shows the symbol error rate (SER) of the respective algorithms as a function of SNR. We can see that our proposed algorithm achieves a slightly lower SER than the hyperplane-based algorithm and the kurtosis-based algorithms, especially at a moderately high SNR. In fact, the performance of our proposed algorithm is closely related to the adopted clustering algorithm. Hence, more accurate clustering techniques, particularly at low SNR, result in better performance of our proposed algorithm.

5.6 Summary

It has been shown that the received data constellation geometry contains rich information pertaining to the channel. Based on this observation, we develop a practical non-iterative algorithm for blind separation of digital signals with M-ASK and

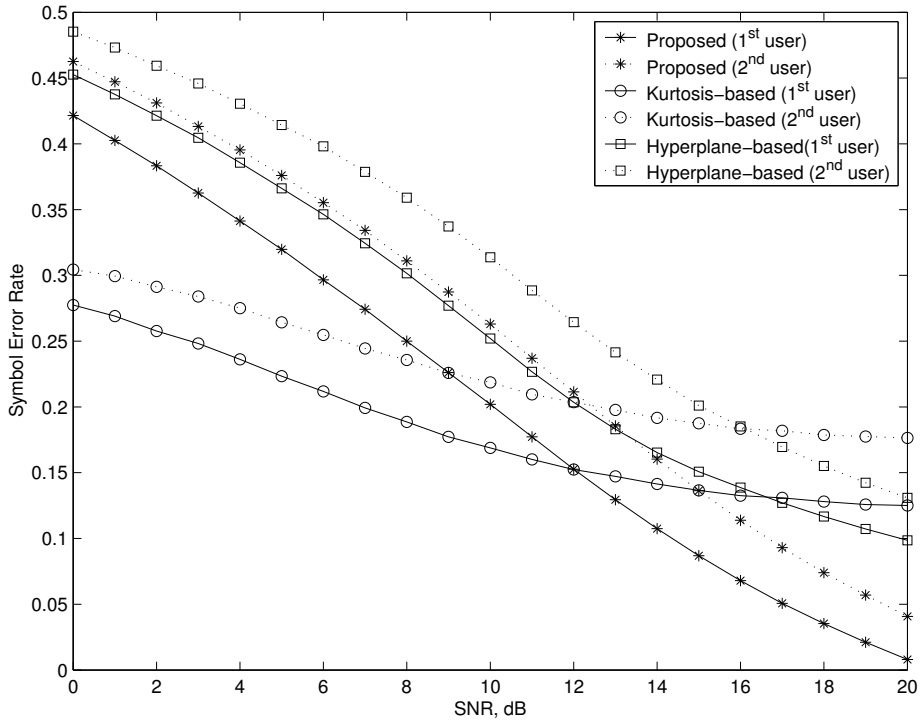

Figure 5.1: Symbol Error Rate (SER) Versus SNR

QAM alphabets. The proposed algorithm compares favorably with the existing hyperplane-based and kurtosis-based algorithms. Since only a small fraction of the constellation geometry is exploited in our geometric approach for channel estimation, it is desirable for us to devise an efficient and more accurate geometric channel estimation approach in the future by utilizing the constellation geometry to a full extent.

Chapter 6

Blind MIMO-OFDM Channel Estimation Based on Spectra Correlations

6.1 Introduction

Blind estimation of the multi-input multi-output (MIMO) finite-impulse-response (FIR) channel has been a very attractive research area for many years. For this problem, a variety of techniques has been developed based either on higher order statistics (HOS) [89, 96–98] or second order statistics (SOS) [94–96] of the observed signals. The estimation methods based on SOS are more attractive because they require far fewer samples than the traditional estimation methods based on HOS.

In Chapter 3 and 4, we proposed two blind channel estimation algorithms which can estimate the MIMO-OFDM channels up to an unknown ambiguity matrix which

must be further resolved. Therefore, a blind source separation (BSS) algorithm is proposed in Chapter 5 to resolve the ambiguity matrix. On the other hand, the MIMO FIR channel driven by colored source may provide us with advantages in developing a complete closed-form SOS-based method without an extra BSS algorithmic step. If the input sources have distinct power spectra, then the MIMO FIR channel can be blindly estimated up to a scaling by using the SOS of the channel output [106]. One successful algorithm which applies to nonstationary signals was proposed in [91]. It has been shown that the cyclostationarity induced at the receiver side by over or fractionally sampling the received waveform permits blind estimation of most FIR channels under certain conditions [99]. In [93], by inducing cyclostationarity in the input signal, a closed-form solution was obtained based on the SOS of the channel outputs. In [92], a special structure was imposed on each input and the resulting MIMO problem with colored inputs was solved using the SOS. In [107], the authors proposed a blind MIMO FIR channel identification algorithm by exploiting the second-order spectra correlations of the system outputs.

Second-order cyclostationary statistics has also been used for blind channel estimation in MIMO-OFDM systems. A subspace-based approach for blind channel estimation using cyclic correlations at the OFDM receiver was proposed in [90]. This approach is robust to the presence of stationary noise and channel order overestimation error and does not require the cyclic prefix to be longer than the channel memory [102]. However, this approach is proposed for the single-input single-output (SISO) CP-OFDM systems only. In [105], a blind MIMO-OFDM channel estimation algorithm using a periodic nonconstant modulus precoding scheme was introduced.

The basic idea of this method is to provide each transmit antenna with a different signature in the cyclostationary domain to null out the influence of all but one transmit antenna at a time, and blindly estimates the subchannels individually.

In this chapter, we exploit the second-order spectra correlations of the system output to blindly estimate the FIR channel matrix of the MIMO-OFDM systems, which are driven by stationary or cyclostationary and nonwhite inputs with known correlations. We follow the idea of [107] and generalize it to the MIMO-OFDM case. In OFDM systems, the source signals before IFFT modulation and the received signals after FFT demodulation are called frequency domain signals, while the transmitted signals after IFFT modulation and the received signals before IFFT demodulation are called time domain signals. In [107], the criterion function involves the spectra correlation of the received signal which must be calculated using the FFT. In addition, the number of the FFT points must be carefully selected to make the equality of the essential principle equation hold. On the contrary, our proposed algorithm directly uses the correlations of the frequency domain signals to construct the criterion, minimization of which yields the system impulse response within a scalar ambiguity. Therefore, our proposed algorithm is simpler and more efficient.

The rest of this chapter is organized as follows. First, the system model is reviewed in Section 6.2. Then in Section 6.3, we exploit the spectra correlation of the system output to blindly estimate the FIR channels. We propose a closed-form solution for the channel matrix. The identifiability of the proposed channel estimator and two practical casea are discussed in Section 6.4. The simulations of

the proposed algorithm are presented in Section 6.5. Finally, the summary of this chapter is given in the last section.

6.2 Problem Formulation

Consider the MIMO-OFDM system shown in Figure 2.9, which is equipped with M_t and M_r transmit and receive antennae respectively. The notations are the same as those in Subsection 2.4.2. The source information symbols to be transmitted by the ith antenna, $s_i(n)$; $(n = 0, 1, 2, \cdots)$, are divided into groups of length N_c, where N_c is the number of subcarriers, and each symbol of every group is transmitted on one subcarrier. Denotes the collection of the source symbols to be sent on the kth MIMO-OFDM symbol by

$$\mathbf{s}(k) = \begin{bmatrix} \mathbf{s}(k,0) \\ \vdots \\ \mathbf{s}(k, N_c - 1) \end{bmatrix} \tag{6.1}$$

where

$$\mathbf{s}(k,n) \triangleq \begin{bmatrix} s_1(k,n), & s_2(k,n), & \cdots, & s_{M_t}(k,n) \end{bmatrix}^T \tag{6.2}$$

and

$$s_i(k,n) \triangleq s_i(kN_c + n), \quad i = 1, \cdots, M_t, \quad n = 0, \cdots, N_c - 1 \tag{6.3}$$

represents the source symbol that is carried by the nth subcarrier in the kth MIMO-OFDM symbol, and transmitted by the ith transmit antenna.

Let $h_{j,i}(l)$ $(l = 0, \cdots, L,\ i = 1, \cdots, M_t,\ j = 1, \cdots, M_r)$ denotes the discrete time channel impulse response between the $(i,j)^{th}$ transceiver pair, where L is the

maximum channel order. Then, as discussed in Chapter 2, the kth received MIMO-OFDM symbol can be written as

$$\mathbf{y}(k) \triangleq \begin{bmatrix} \mathbf{H}(0) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{H}(N_c-1) \end{bmatrix} \mathbf{s}(k) + \mathbf{n}(k) \tag{6.4}$$

where

$$\mathbf{H}(n) = \sum_{l=0}^{L} \mathbf{h}(l) e^{-j\frac{2\pi}{N_c}nl} \quad \text{for} \quad n = 0, \cdots, N_c - 1 \tag{6.5}$$

$$\mathbf{h}(l) = \begin{bmatrix} h_{1,1}(l) & \cdots & h_{1,M_t}(l) \\ \vdots & \ddots & \vdots \\ h_{M_r,1}(l) & \cdots & h_{M_r,M_t}(l) \end{bmatrix} \text{for} \quad l = 0, 1, \cdots L \tag{6.6}$$

and $\mathbf{n}(k,n)$ represents the frequency domain noise vector whose elements are zero mean complex AWGNs with the variance σ_n^2, and are all spatially and temporally independent form each other. So, for the nth subcarrier we may write

$$\mathbf{y}(k,n) = \mathbf{H}(n)\mathbf{s}(k,n) + \mathbf{n}(k,n) \tag{6.7}$$

where

$$\mathbf{y}(k,n) = \begin{bmatrix} y_1(k,n), & y_2(k,n), & \cdots, & y_{M_r}(k,n) \end{bmatrix}^T \tag{6.8}$$

and

$$y_j(k,n) \triangleq y_j(kN_c + n), \quad i = 1, \cdots, M_r, \quad n = 0, \cdots, N_c - 1 \tag{6.9}$$

is the received symbol that is carried by the nth subcarrier in the kth MIMO-OFDM symbol, and received by the jth receive antenna.

For this system model, we adopt the following basic assumptions:

A1) The colored source signals are zero mean, wide sense stationary or cyclostationary with known statistics.

A2) The input colors are known and pairwise nonidentical.

A3) The time domain channels $h_{j,i}(l)$ are in general complex, and the maximum channel order L is known apriori.

A4) The scalar subchannel matrices $\mathbf{H}(k)$ is full column rank for $k = 0, \cdots, N_c$.

Assumptions A1) and A3) are generally required by most of the existing algorithms. It should be noted that we only require the upper bound of the channel orders to be known. Assumptions A2) and A4) are required to guarantee that the channels can be uniquely estimated up to scalar ambiguities. Assumption A4) implies that the number of outputs is greater than or equal to the number of the inputs. Unlike some existing estimation methods which require the guard intervals must be some specified type, the proposed estimation method is capable for both of the CP-based and ZP-based OFDM systems.

Our goal is to blindly estimate the channel matrices $\mathbf{H}(n)$ based on the spectra correlations of the observed channel outputs. However, direct estimation of these $\mathbf{H}(n)$ would be very inefficient since it requires the estimation of $N_c M_t M_r$ parameters, which can be a significant number in practical systems. Rather, we propose an algorithm which estimates the time domain channel taps $\mathbf{h}(l)$ $(l = 0, \cdots, L)$ instead, which greatly reduces the number of parameters need to be estimated. Thereafter, the subchannel matrices $\mathbf{H}(n)$ can be calculated from $\mathbf{h}(l)$ by using the FFT.

6.3 Proposed Blind Estimation Algorithm

Consider the autocorrelation of the source signal $s_i(n)$ which is defined as follows

$$R_{s_i}(n,\tau) \triangleq E\{s_i(n)s_i^*(n+\tau)\} \tag{6.10}$$

Note that $s_i(n)$ is the frequency domain signal in the OFDM system. Thus, $R_{s_i}(n,\tau)$ is also referred to as the spectra correlation. If $s_i(n)$ is cyclostationary, then we have

$$R_{s_i}(n,\tau) = R_{s_i}(n+\kappa N',\tau) \quad \text{for} \quad n = 0,1,\cdots,N'-1 \tag{6.11}$$

where κ is an arbitrary integer, and N' is the period of the cyclic spectra correlation. Especially, if the source signal is stationary, then

$$R_{s_i}(n,\tau) = R_{s_i}(\tau) = R_{s_i}(n+\kappa,\tau) \tag{6.12}$$

Thus, it can be looked as a special case of cyclostationary source such that the period $N' = 1$. We assume that all the users have the identical cyclic spectra correlation period N', and they are pairwise uncorrelated. Then we have

$$\begin{aligned} \mathbf{R_s}(n,\tau) &\triangleq E\{\mathbf{s}(n)\mathbf{s}^H(n+\tau)\} \\ &= \mathbf{R_s}(n+\kappa N',\tau) \\ &= \begin{bmatrix} R_{s_1}(n,\tau) & & 0 \\ & \ddots & \\ 0 & & R_{s_{M_t}}(n,\tau) \end{bmatrix} \end{aligned} \tag{6.13}$$

Thus, according to Eqn.(6.7), the auto correlation matrix of the received signals before FFT demodulation, which is referred to as the spectra correlation matrix of

the received signals, can be written as

$$\begin{aligned}\mathbf{R}_{\mathbf{y}}(n,\tau) &\triangleq E\{\mathbf{y}(n)\mathbf{y}(n+\tau)\} \\ &= \mathbf{H}\left(n \bmod N_c\right)\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)\mathbf{H}^H\left((n+\tau) \bmod N_c\right) + \sigma_n^2\mathbf{I}\delta(\tau) \\ &= \mathbf{R}_{\mathbf{y}}(n+\kappa N'',\tau) \quad \text{for} \quad n = 0,1,\cdots,N''-1 \end{aligned} \tag{6.14}$$

where $(n \bmod N_c)$ denotes the remainder of n divided by N_c, N'' equals to the least common multiple of N_c and N', and σ_n^2 is the noise power.

In Eqn.(6.5), we defined the frequency domain channel matrix $\mathbf{H}(n)$ for $n = 0,\cdots,N_c-1$. Now we extend this definition to $n = 0,\cdots,N''-1$, i.e.,

$$\mathbf{H}(n) = \sum_{l=0}^{L}\mathbf{h}(l)e^{-j\frac{2\pi}{N_c}nl} \quad \text{for} \quad n = 0,\cdots,N''-1 \tag{6.15}$$

Obviously, it always holds that

$$\mathbf{H}(n) = \mathbf{H}(n \bmod N_c) \tag{6.16}$$

Therefore, Eqn.(6.14) can be rewritten as

$$\begin{aligned}\mathbf{R}_{\mathbf{y}}(n,\tau) &= \mathbf{H}\left(n\right)\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)\mathbf{H}^H\left(n+\tau\right) + \sigma_n^2\mathbf{I}\delta(\tau) \\ &\quad \text{for} \quad n = 0,1,\cdots,N''-1 \end{aligned} \tag{6.17}$$

If we consider two different nonzero values of τ, and substitute to Eqn.(6.17), we can form the following equation which admits a closed form solution for $\mathbf{h}(l)$:

$$\mathbf{R}_{\mathbf{s}}\left(n,\tau_1\right)\mathbf{H}^H\left(n+\tau_1\right)\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_1) = \mathbf{R}_{\mathbf{s}}\left(n,\tau_2\right)\mathbf{H}^H\left(n+\tau_2\right)\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_2) \tag{6.18}$$

Note that by choosing the nonzero values of τ, the influence of the AWGN noise is completely removed. We also note that the source signals are pairwise uncorrelated. Thus, according to Eqn.(6.17), $\mathbf{R}_{\mathbf{s}}\left(n,\tau_1\right)$ and $\mathbf{R}_{\mathbf{s}}\left(n,\tau_2\right)$ are diagonal matrices. Take

every row of both the left and right hand sides of the above Eqn.(6.18), we have M_t equations associated with the M_t transmit antennae respectively, i.e.,

$$R_{s_i}(n,\tau_1)\,\mathbf{H}_i^H(n+\tau_1)\,\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_1) = R_{s_i}(n,\tau_2)\,\mathbf{H}_i^H(n+\tau_2)\,\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_2) \quad \text{for} \quad i=1,\cdots,M_t \tag{6.19}$$

where $R_{s_i}(n,\tau)$ is the ith diagonal elements of $\mathbf{R_s}(n,\tau)$ and $\mathbf{H}_i(n+\tau)$ is the ith columns of $\mathbf{H}(n+\tau)$ (for $\tau=\tau_1,\ \tau_2$). Substitute Eqn.(6.5) to Eqn.(6.19), we have

$$\sum_{l=0}^{L}\mathbf{h}_i^H(l)\left[R_{s_i}(n,\tau_1)e^{-j\frac{2\pi}{N_c}\tau_1 l}\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_1)-R_{s_i}(n,\tau_2)e^{-j\frac{2\pi}{N_c}\tau_2 l}\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_2)\right]=\mathbf{0} \tag{6.20}$$

where

$$\mathbf{h}_i(l)\triangleq\left[\ h_{1,i}(l),\ \ \cdots,h_{M_r,i}(l)\ \right]^T \tag{6.21}$$

is the ith column of the time domain channel matrix $\mathbf{h}(l)$. Evaluating Eqn.(6.20) for different ns ($n=0,\cdots,N''-1$), we can set up a system of equations with the unknown elements of $\mathbf{h}_i$, i.e.,

$$\mathcal{F}_i^H\mathbf{h}_i=\mathbf{0} \tag{6.22}$$

where

$$\mathbf{h}_i\triangleq\left[\ \mathbf{h}_i^T(0)\ \ \cdots\ \ \mathbf{h}_i^T(L)\ \right]^T \tag{6.23}$$

and $\mathcal{F}_i$ consists of the submatrices $\mathcal{F}_i(l,n)$ for $l=0,\cdots,L,\ n=0,\cdots,N''-1$, where $\mathcal{F}_i(l,n)$ denotes the $(l,n)^{th}$ submatrix, which is defined as follows

$$\mathcal{F}_i(l,n)\triangleq R_{s_i}(n,\tau_1)e^{-j\frac{2\pi}{N_c}\tau_1 l}\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_1)-R_{s_i}(n,\tau_2)e^{-j\frac{2\pi}{N_c}\tau_2 l}\mathbf{R}_{\mathbf{y}}^{\dagger}(n,\tau_2) \tag{6.24}$$

Thus, the unknown channel matrix $\mathbf{h}_i$ must satisfy the following optimization equation

$$\hat{\mathbf{h}}_i=\min_{\|\mathbf{h}_i\|=1}\|\mathcal{F}_i^H\mathbf{h}_i\|^2 \quad \text{for} \quad i=1,\cdots,M_t \tag{6.25}$$

In practice, the spectra correlations of the received signals are unknown and must be estimated as follows

$$\hat{\mathbf{R}}_{\mathbf{y}}(n,\tau) = \frac{1}{T}\sum_{n=0}^{T-1}\mathbf{y}(n)\mathbf{y}^H(n+\tau) \tag{6.26}$$

and consequently the matrix $\mathcal{F}_i(l,n)$ in the criterion function is estimated by

$$\hat{\mathcal{F}}_i(l,n) \triangleq R_{s_i}(n,\tau_1)e^{-j\frac{2\pi}{N_c}\tau_1 l}\hat{\mathbf{R}}_{\mathbf{y}}^{\dagger}(n,\tau_1) - R_{s_i}(n,\tau_2)e^{-j\frac{2\pi}{N_c}\tau_2 l}\hat{\mathbf{R}}_{\mathbf{y}}^{\dagger}(n,\tau_2) \tag{6.27}$$

Thus, the unknown channel matrix $\mathbf{h}_i$ can be found by solving the optimization problem

$$\hat{\mathbf{h}}_i = \min_{\|\mathbf{h}_i\|=1} \|\hat{\mathcal{F}}_i^H\mathbf{h}_i\|^2 \quad \text{for} \quad i = 1,\cdots,M_t \tag{6.28}$$

It is well known that the solution of the above optimization problem is the eigenvector of the matrix $\hat{\mathcal{F}}_i\hat{\mathcal{F}}_i^H$ associated with its smallest eigenvalue. We will show later that the channel can be estimated up to a complex diagonal ambiguity matrix.

6.4 Discussion

6.4.1 Identifiability

In this subsection, we discuss the identifiability of the proposed blind channel estimation method. We show that the channel vectors associated with each transmit antenna can be estimated up to a complex scalar, i.e., the ambiguity matrix for the whole channel matrix is a complex diagonal matrix.

Lemma 6.4.1 Under the assumptions A1)-A4), if there exists a matrix $\hat{\mathbf{H}}(n)$ such that Eqn.(6.17) is satisfied, then

$$\hat{\mathbf{H}}(n) = \mathbf{H}(n)\mathbf{\Lambda} \tag{6.29}$$

where $\mathbf{\Lambda}$ is a constant complex diagonal ambiguity matrix.

Proof: According to Proposition 1 of [106], if the assumptions A1)-A4) are satisfied except that the spectra correlations of the source signals are unknown, then the channel matrix $\mathbf{H}(n)$ can be reconstructed from Eqn.(6.17) up to a column permutation matrix $\mathbf{P}(n)$ and a complex diagonal matrix $\mathbf{\Lambda}(n)$, i.e.,

$$\hat{\mathbf{H}}(n) = \mathbf{H}(n)\mathbf{P}(n)\mathbf{\Lambda}(n) \tag{6.30}$$

If there exists such a solution, then according to Eqn.(6.17), for some $\tau \neq 0$ we have

$$\begin{aligned} \mathbf{R}_{\mathbf{y}}(n,\tau) &= \mathbf{H}\left(n\right)\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)\mathbf{H}^{H}\left(n+\tau\right) \\ &= \mathbf{H}\left(n\right)\mathbf{P}(n)\mathbf{\Lambda}(n)\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)\mathbf{\Lambda}^{H}(n+\tau)\mathbf{P}^{H}(n+\tau)\mathbf{H}^{H}\left(n+\tau\right) \end{aligned} \tag{6.31}$$

Since $\mathbf{H}(n)$ is assumed to be full column rank, then Eqn.(6.31) is equivalent to

$$\mathbf{P}(n)\mathbf{\Lambda}(n)\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)\mathbf{\Lambda}^{H}(n+\tau)\mathbf{P}^{H}(n+\tau) = \mathbf{R}_{\mathbf{s}}\left(n,\tau\right) \tag{6.32}$$

We note that $\mathbf{R}_{\mathbf{s}}\left(n,\tau\right)$ is a diagonal matrix of full rank, and $\mathbf{P}(n)$ and $\mathbf{P}(n+\tau)$ are both column permutation matrices for any τ and n. Hence, taking two different nonzero values of τ, say τ_1 and τ_2, yields the following equation

$$\begin{aligned} & \mathbf{R}_{\mathbf{s}}(n,\tau_1)\mathbf{\Lambda}^{H}(n+\tau_1)\mathbf{P}^{H}(n+\tau_1)\mathbf{R}_{\mathbf{s}}^{-1}(n,\tau_1) \\ = \; & \mathbf{R}_{\mathbf{s}}(n,\tau_2)\mathbf{\Lambda}^{H}(n+\tau_2)\mathbf{P}^{H}(n+\tau_2)\mathbf{R}_{\mathbf{s}}^{-1}(n,\tau_2) \end{aligned} \tag{6.33}$$

Define $\mathbf{e}_n$ as the column vector with the nth element being 1, and all other elements being 0. Let

$$\mathbf{e}_{p_1}^{T} = \mathbf{e}_1^{T}\mathbf{P}^{H}(n+\tau_1) \tag{6.34a}$$

$$\mathbf{e}_{p_2}^{T} = \mathbf{e}_1^{T}\mathbf{P}^{H}(n+\tau_2) \tag{6.34b}$$

It should be noted that since $\mathbf{P}(n+\tau_1)$ and $\mathbf{P}(n+\tau_2)$ are permutation matrices, then $\mathbf{e}_{p_1}$ and $\mathbf{e}_{p_2}$ are row vectors with only one element being nonzero, and the positions of the nonzero elements, which are indexed by p_1 and p_2, depend on $\mathbf{P}(n+\tau_1)$ and $\mathbf{P}(n+\tau_2)$. By taking the 1st row of the both side of the above Eqn.(6.33), we have

$$\lambda_1(1)\gamma_1(p_1)\mathbf{e}_{p_1}^T = \lambda_2(1)\gamma_2(p_2)\mathbf{e}_{p_2}^T \tag{6.35}$$

where $\lambda_i(1)$ is the 1st diagonal elements of $\mathbf{\Lambda}(n+\tau_i)$ for $i = 1, 2$, and

$$\gamma_1(p_1) \triangleq R_s(n,\tau_1)(1)R_s^{-1}(n,\tau_1)(p_1) \tag{6.36a}$$

$$\gamma_2(p_2) \triangleq R_s(n,\tau_2)(1)R_s^{-1}(n,\tau_2)(p_2) \tag{6.36b}$$

where $R_s(n,\tau_1)(1)$ and $R_s(n,\tau_2)(1)$ are the 1st element of the diagonal matrix $\mathbf{R_s}(n,\tau_1)$ and $\mathbf{R_s}(n,\tau_2)$ respectively, while $R_s(n,\tau_1)(p_1)$, $R_s(n,\tau_2)(p_2)$ are the p_1^{th} and p_2^{th} diagonal elements of $\mathbf{R_s}(n,\tau_1)$ and $\mathbf{R_s}(n,\tau_2)$ respectively. Since the diagonal elements of $\mathbf{\Lambda}(n+\tau_1)$, $\mathbf{\Lambda}(n+\tau_2)$, $\mathbf{R_s}(n,\tau_1)$ and $\mathbf{R_s}(n,\tau_2)$ are obviously nonzero, then Eqn.(6.35) implies that $\mathbf{e}_{p_1} = \mathbf{e}_{p_2}$. By repeating the above steps for every row of the both sides of Eqn.(6.33), we can conclude that the permutation matrices $\mathbf{P}(n+\tau)$ is constant for any n and τ. Therefore, Eqn.(6.32) can be re-expressed as

$$\mathbf{P}\mathbf{\Lambda}(n)\mathbf{R_s}\left(n,\tau\right)\mathbf{\Lambda}^H(n+\tau)\mathbf{P}^H = \mathbf{R_s}\left(n,\tau\right) \tag{6.37}$$

Note that $\mathbf{R_s}\left(n,\tau\right)$ and $\mathbf{\Lambda}(n)\mathbf{R_s}\left(n,\tau\right)\mathbf{\Lambda}^H$ are both diagonal matrices. To make the above equation hold, $\mathbf{P}$ must be an identity matrix.

Now we consider the solution of the form $\hat{\mathbf{H}} = \mathbf{H}(n)\mathbf{\Lambda}(n)$. Substitute to Eqn.(6.32), it follows that

$$\mathbf{\Lambda}(n)\mathbf{R_s}\left(n,\tau\right)\mathbf{\Lambda}^H(n+\tau) = \mathbf{R_s}\left(n,\tau\right)\mathbf{\Lambda}(n)\mathbf{\Lambda}^H(n+\tau) = \mathbf{I}_{M_t} \tag{6.38}$$

where $\mathbf{I}_{M_t}$ is the $M_t \times M_t$ identity matrix. Since this must hold for all n and τ, then $\mathbf{\Lambda}(n)$ must be independent of n, i.e., $\mathbf{\Lambda}(n) = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a complex diagonal matrix. ■

Note the difference of our proposed algorithm and the existing algorithms proposed in [106] and [107], which also exploit the spectra correlations of the system outputs. Firstly, these two existing algorithms are both proposed for common MIMO FIR systems, while our algorithm is proposed for the MIMO-OFDM systems and is capable for both CP and ZP based OFDM. Secondly, in [106] and [107], the channel estimation is performed in the frequency domain, where the FFT is employed to calculate the spectra correlation matrix. On the other hand, our proposed algorithm directly use the frequency domain OFDM symbols to compute the spectra correlation. This makes our proposed algorithm more efficient and simpler. Lastly, to make the essential principle equation hold, [106] focus on the cyclostationary sources and [107] requires the number of the FFT points satisfying certain condition. In contrast, our proposed work is capable with both the cyclostationary and stationary source. In addition, we consider two different nonzero values of τ to suppress the influence of the AWGN noise.

6.4.2 Two Practical Examples

In this subsection, we consider two examples, where the systems are driven by the source that is convolutionally or block coded respectively. These two examples are very common in practice. As will be shown, due to the inherent structure of the spectra correlations of these two sources, the channel estimation procedures are

slightly different.

Convolutional Coding

Consider the source signals which are linear processes generated as follows

$$s_i(n) = \sum_{l=0}^{L_c} c_i(l) d_i(n-l) \tag{6.39}$$

where $d_i(n)$ is a white process, $c_i(l)$, $l = 0, \cdots, L_c$, is the corresponding coding process, or in other words, the color of the corresponding transmitted signal, and L_c denotes the maximum color length in case the colors have different lengths. Obviously, the generated signal $s_i(n)$ is stationary, and according to assumption A4), $c_i(l)$ are known and pairwise distinct. Thus, the spectra correlation of the transmitted signal can be written as

$$\begin{aligned} \mathbf{R}_s(n,\tau) &= \mathbf{R}_s(\tau) = E\{\mathbf{s}(n)\mathbf{s}^H(n+\tau)\} \\ &= \sum_{l_1=0}^{L_c}\sum_{l_2=0}^{L_c} c_i(l_1)c_i^*(l_2)\delta(l_2-l_1-\tau) \end{aligned} \tag{6.40}$$

where we assume that the white processes $d_i(n)$ have unitary powers. Consequently, the spectra correlation of the received signals are cyclic with period N_c, and can be rewritten as follows

$$\begin{aligned} \mathbf{R}_{\mathbf{y}}(n,\tau) &\triangleq E\{\mathbf{y}(n)\mathbf{y}(n+\tau)\} \\ &= \mathbf{H}(n)\,\mathbf{R}_{\mathbf{s}}(\tau)\,\mathbf{H}^H((n+\tau) \bmod N_c) + \sigma_n^2\mathbf{I}\delta(\tau) \\ &= \mathbf{R}_{\mathbf{y}}(n+\kappa N_c,\tau) \quad \text{for} \quad n = 0,1,\cdots,N_c-1 \end{aligned} \tag{6.41}$$

The remaining steps of the proposed blind channel estimation algorithm for this convolutional coding case is the same as the general case expressed in Section 6.3.

It should be noted that in this case, the spectra correlation of the transmitted signal is independent of the time index n, i.e., the transmitted signal is stationary. Unlike some of the existing algorithms based on the non-stationarity of the source signals, the proposed algorithm is capable with the system driven by the stationary source.

Block Coding

In this subsection, we consider the practical application of the proposed blind channel estimation algorithm in the circumstance where the source signals are block coded. Denote the kth block of the coded source symbols as

$$\mathbf{s}_i(k) = \mathbf{P}_i \mathbf{d}_i(k) \tag{6.42}$$

where

$$\begin{aligned} \mathbf{d}_i(k) &\triangleq \left[\begin{array}{ccc} d_i(k,0), & \cdots & , d_i(k, N_c - 1) \end{array} \right]^T \\ &= \left[\begin{array}{ccc} d_i(kN_c), & \cdots & , d_i(kN_c + N_c - 1) \end{array} \right]^T \end{aligned} \tag{6.43}$$

is the kth block of the uncoded source symbols to be transmitted by the ith transmit antenna, which is assumed to be i.i.d. with unit power, $\mathbf{P}_i$ is the corresponding block coding matrix, which is distinct for different users (transmit antennae), and

$$\begin{aligned} \mathbf{s}_i(k) &\triangleq \left[\begin{array}{ccc} s_i(k,0), & \cdots & , s_i(k, N_c - 1) \end{array} \right]^T \\ &= \left[\begin{array}{ccc} s_i(kN_c), & \cdots & , s_i(kN_c + N_c - 1) \end{array} \right]^T \end{aligned} \tag{6.44}$$

is the kth block of the coded source symbols, which constitute the kth OFDM symbol corresponding to the ith user (transmit antenna). Note that we don't require the coding scheme introduce any redundancy.

Thereafter, the autocorrelation matrix of the kth OFDM symbol can be written as

$$\begin{aligned}\mathbf{R}_{\mathbf{s}_{i,j}}(k,\gamma) &\triangleq E\{\mathbf{s}_i(k)\mathbf{s}_j^H(k+\gamma)\} \\ &= \mathbf{P}_i\mathbf{P}_j^H\delta(\gamma)\delta(i-j)\end{aligned} \tag{6.45}$$

This equation indicates that the coded signals belonging to different OFDM symbols are mutually uncorrelated. More specifically, the spectra correlation of the coded source signal can be written as

$$\begin{aligned}R_{s_i}(n,\tau) &= E\{s_i(n)s_i^*(n+\tau)\} \\ &= \begin{cases}\mathbf{P}_i(n,:)\mathbf{P}_i^H(n+\tau,:) & \text{for } 0\le n\le N_c-1,\ \ -n\le\tau\le N_c-n-1 \\ 0 & \text{otherwise}\end{cases}\end{aligned} \tag{6.46}$$

where $\mathbf{P}_i(n,:)$ denotes the nth row of the precoding matrix $\mathbf{P}_i$. From Eqn.(6.46), we know that the coded source signal is cyclostationary with period N_c. Therefore, the relation between the spectra correlations of the transmitted and received signals can be reconstructed as

$$\begin{aligned}\mathbf{R}_{\mathbf{y}}(n,\tau) &= \mathbf{H}(n)\,\mathbf{R}_{\mathbf{s}}(n,\tau)\,\mathbf{H}^H(n+\tau)+\sigma_n^2\mathbf{I}\delta(\tau) \\ &\text{for}\quad n=0,1,\cdots,N_c-1,\quad -n\le\tau\le N_c-n-1\end{aligned} \tag{6.47}$$

Following the remaining steps in Section 6.3, one could construct the optimization problem described in Eqn.(6.28), settling which yields the closed-form solution for the channel matrix, i.e.,

$$\hat{\mathbf{h}}_i = \min_{\|\mathbf{h}_i\|=1} \|\hat{\mathcal{F}}_i^H\mathbf{h}_i\|^2 \tag{6.48}$$

where $\hat{\mathcal{F}}_i$ consists of the submatrices $\hat{\mathcal{F}}_i(n,l)$ for $l=0,\cdots,L$, and $n=n_{\min},\cdots,n_{\max}$,

where

$$\begin{cases} n_{\min} &= \max\{-\tau_1, \ -\tau_2, \ 0\} \\ n_{\max} &= \min\{N_c - 1 - \tau_1, \ N_c - 1 - \tau_2, \ N_c - 1\} \end{cases} \tag{6.49}$$

One should notice the slight difference between Eqn.(6.47) and Eqn.(6.17), which leads to the difference of the criterion functions displayed by Eqn.(6.48) and Eqn.(6.28). Since we need the evaluated criterion function satisfying that $\mathbf{R_s}(n, \tau) \neq 0$, then once τ is selected, the number of equations derived from Eqn.(6.47) is less than the general case. This may cause the performance degradation. Fortunately, the number of subcarriers in practical OFDM systems, N_c, is fairly large. Thus, we still have enough equations derived from Eqn.(6.47) to construct the criterion function and get an acceptable performance.

6.5 Simulation Results

In this section, we consider two simulation examples to illustrate the performance of our proposed algorithm. In both examples, the simulated MIMO-OFDM system are modeled containing 64 subcarriers, i.e. $M = 64$. Each OFDM symbol consists of 68 chips including the guard interval (CP / ZP) of length 4, i.e. $M_g = 4$. The system is equipped with 2 transmit antennae and 2 receive antennae respectively, and the channel model used is a 3-tap FIR filter with tap coefficients independently chosen from a complex white Gaussian process.

To evaluate the channel estimation error, we employed the normalized-root-mean-square-error (NRMSE), which is defined as

$$\text{NRMSE} = \sqrt{\frac{1}{N_1 N_2} \sum_{t_1=1}^{N_1} \sum_{t_2=1}^{N_2} \frac{\|\hat{\mathbf{h}}^{(t_1,t_2)} - \mathbf{h}^{(t_1)}\|_F^2}{\|\mathbf{h}^{(t_1)}\|_F^2}} \tag{6.50}$$

where $\|.\|_F$ denotes the Frobenius norm, N_1 and N_2 are the number random channel realizations and the number of Monte Carlo runs for each channel realization, respectively; $\mathbf{h}^{(t_1)}$ is true channel matrix for the t_1^{th} realization, which is randomly generated, and $\hat{\mathbf{h}}^{(t_1,t_2)}$ is the estimation of $\mathbf{h}^{(t_1)}$ for the t_2^{th} Monte Carlo run. In both of the two examples, we simulate 30 channel realizations, each for 100 Monte Carlo runs.

Example 1: Convolutionally Coded Source

In this example, we consider the system that is driven by the colored source which is generated by convolutionally coded white sources. More specifically, the colored source symbols $s(n)$ are drawn from a 4-QAM constellation according to the following rule. Let b_n be the input stream of independent and identically distributed bits, i.e., $b_n \in \{0, 1\}$. Then

$$s(n) = \begin{cases} -\frac{\sqrt{(2)}}{2} + \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (0 \; 0) \\ +\frac{\sqrt{(2)}}{2} + \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (0 \; 1) \\ -\frac{\sqrt{(2)}}{2} - \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (1 \; 0) \\ +\frac{\sqrt{(2)}}{2} - \frac{\sqrt{2}}{2}j & \text{if} \quad (b_n \; b_{n-1}) = (1 \; 1) \end{cases} \tag{6.51}$$

This generates a colored symbol sequence with autocorrelation

$$E\{s(n)s^*(n+\tau)\} = \begin{cases} 1 & \text{if} \quad \tau = 0 \\ \mp\frac{1}{2}j & \text{if} \quad \tau = \pm 1 \\ 0 & \text{else} \end{cases} \tag{6.52}$$

Figure 6.1 illustrates the NRMSE as a function of SNR. We collect 500 and 1000 received OFDM symbols to estimate the spectra correlation of the received signals respectively, and thus the criterion functions. We also simulate the training based

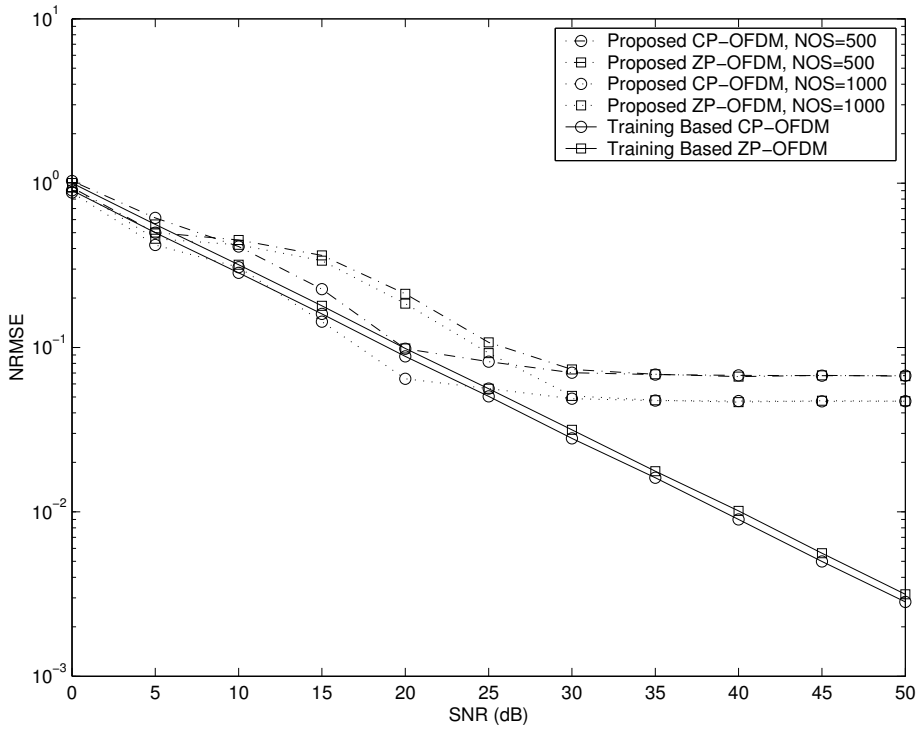

Figure 6.1: NRMSE performance as a function of SNR

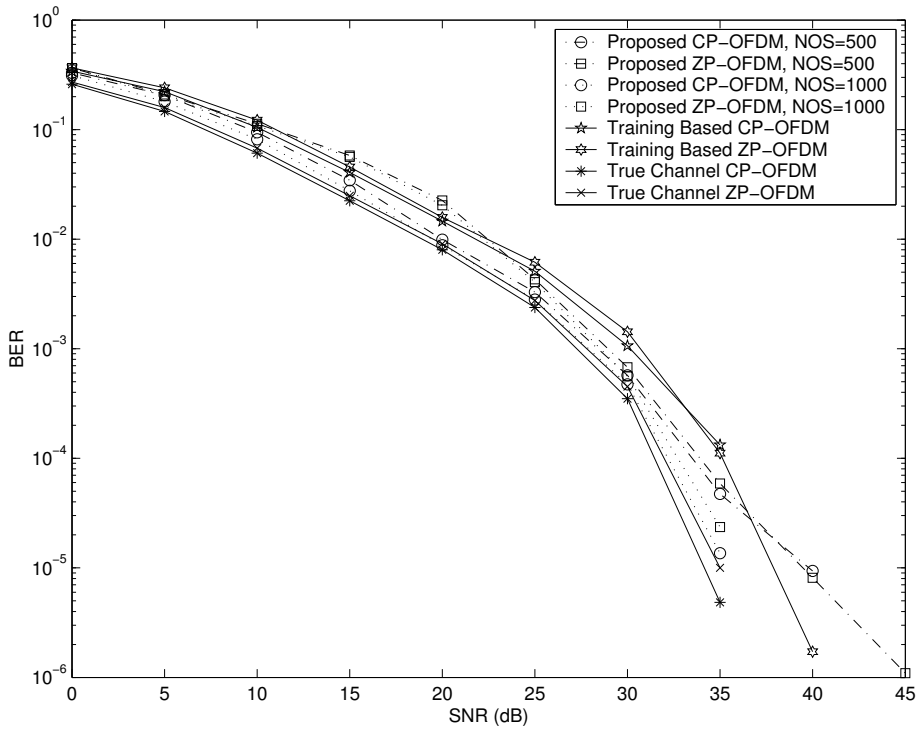

Figure 6.2: BER performance as a function of SNR

LS algorithm proposed in [68] as a comparison. Figure 6.2 illustrates the BER as a function of SNR. The detected symbols are the outputs of the Zero Forcing equalizers using the estimated channel coefficients. Moreover, the BER performance of the training-based LS algorithm and equalization output using the true channel coefficients are also simulated as comparisons. Both of the CP and ZP based OFDM systems are simulated. From the figures, we see that the CP-OFDM system achieves better performance than the ZP-OFDM system. This is not strange, because at the step removing the guard interval, the ZP-OFDM injects extra noise.

Example 2: Block Coded Source

In this example, we consider the system that is driven by the colored source which is generated by block coded white sources, i.e.,

$$\mathbf{s}_i(k) = \mathbf{P}_i \mathbf{d}_i(k) \tag{6.53}$$

where $s_i(k)$ is the kth block of white source drawn from a 4-QAM constellation and to be transmitted by the ith antenna, and $\mathbf{P}_i$ the coding matrix associated with the ith transmit antenna. In this example, we define

$$\mathbf{P}_1 = \begin{bmatrix} 1 & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & 1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & \frac{1}{2} \\ 0 & \cdots & 0 & \frac{1}{2} & 1 \end{bmatrix} \quad \mathbf{P}_2 = \begin{bmatrix} 1 & \frac{1}{2}j & 0 & \cdots & 0 \\ -\frac{1}{2}j & 1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & \frac{1}{2}j \\ 0 & \cdots & 0 & -\frac{1}{2}j & 1 \end{bmatrix} \tag{6.54}$$

Figure 6.3 illustrates the NRMSE as a function of SNR. Similar to the Example 1, we collect 500 and 1000 received OFDM symbols to estimate the spectra correlation of the received signals respectively, and thus the criterion functions. We also

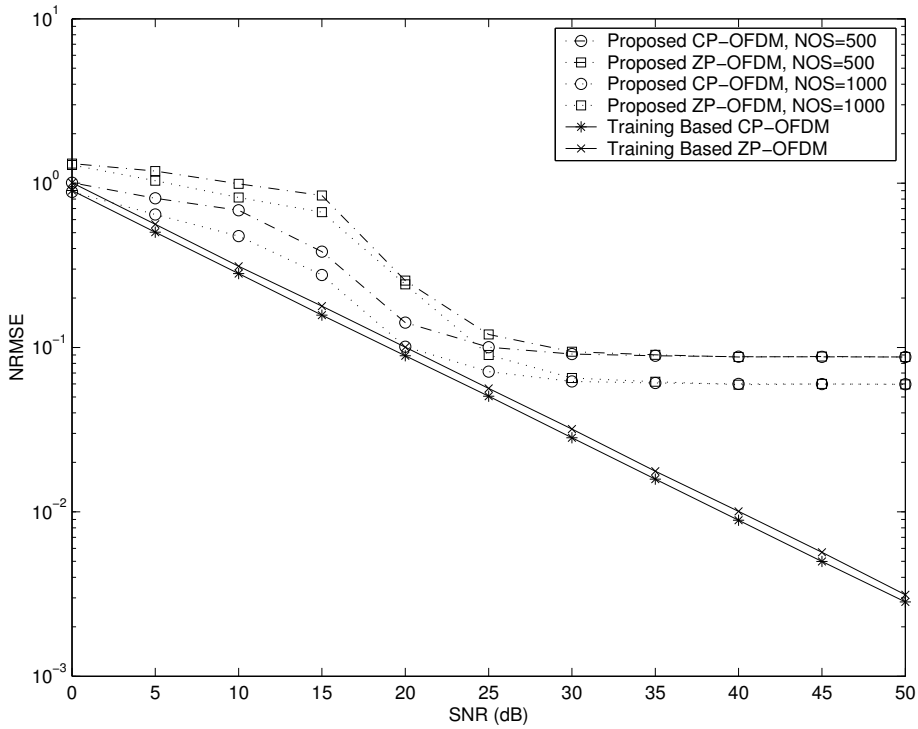

Figure 6.3: NRMSE performance as a function of SNR

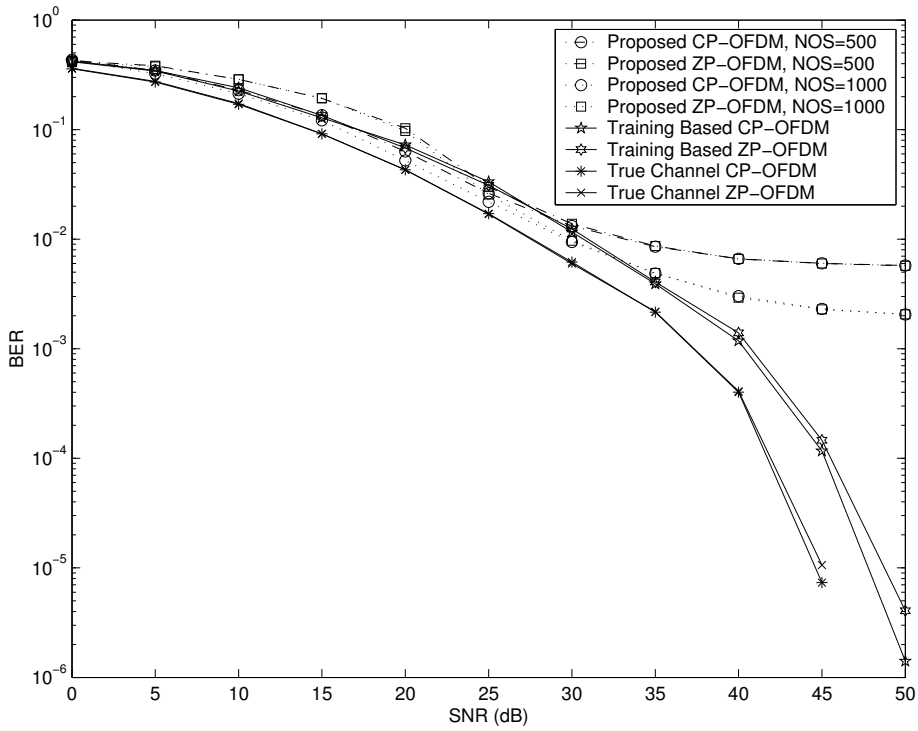

Figure 6.4: BER performance as a function of SNR

simulate the training based LS algorithm proposed in [68] as a comparison. Figure 6.4 illustrates the BER as a function of SNR. The detected symbols are the outputs of the Zero Forcing equalizers using the estimated channel coefficients. The BER performance of the training-based LS algorithm and equalization output using the true channel coefficients are also simulated as comparisons. Both of the CP and ZP based OFDM systems are simulated. We see that the performance of the proposed algorithm in this example is worse than that in Example 1. This is due to the lack of the constraints to construct the criterion function (See Section 6.4).

6.6 Summary

In this chapter, we proposed a closed-form solution for blind channel estimation of MIMO-OFDM systems by exploiting the spectra correlations of the system output. Due to the structure of the OFDM symbols, the channel vectors associated with each transmit antenna can be uniquely estimated up to a complex scalar. Thus, the source separation step which are usually required by some of the existing blind channel estimation algorithms can be avoided. The proposed estimation algorithm is capable with both the CP and ZP based OFDM systems. The computer simulations show that the proposed algorithm can achieve an attractive performance.

Chapter 7

Conclusion

7.1 Summary

In this thesis, we proposed three blind channel estimation algorithms for MIMO-OFDM systems, and one blind source separation algorithm which is designed to resolve the ambiguity matrices remaining in the estimated channel matrices. Each chapter is briefly summarized below.

In Chapter 2, mathematical models are presented for both CP and ZP based MIMO-OFDM systems. It is shown that, by properly designing the guard intervals to combat the inter-symbol interference (ISI), either using cyclic prefix (CP) or zero padding (ZP), the OFDM system can successfully transfer the frequency-selective channel into a set of flat fading subchannels [6]. Therefore the equalization process in OFDM systems is greatly simplified.

In Chapter 3, we present a novel subspace (SS) based algorithm for blind channel estimation of CP based MIMO-OFDM systems. The proposed estimation method is capable for MIMO-OFDM systems driven by either white or colored sources. It

does not requires the system to be equipped with strictly more receive antennae than transmit antennae either. Furthermore, it even doesn't require the length of the CP to be greater than the maximum channel order. This property makes the proposed channel estimator to be attractive in the ill conditioned environments. The computer experiments show that the proposed algorithm achieve attractive performance for systems driven by white source. Although there exists performance degradation when colored source presents, it is still attractive, considering its efficiency and identifiability.

The estimation algorithm presented in Chapter 4 is in fact an extension of [114] which was designed for SISO-OFDM. We design a nonredundant linear precoder for MIMO-OFDM systems which enables blind channel estimation. The identifiability of the proposed algorithm is guaranteed even when the channel matrices share common zeros at subcarrier frequencies. The theoretical performance analysis is also derived.

In Chapter 5, we proposed a geometric based blind source separation method to resolve the ambiguity matrix which is yet to be removed by using the blind channel estimation methods proposed in Chapter 6 and 3. Moreover, this proposed separation method is also a general blind separation method for all flat fading channels. It has been shown that the received data constellation geometry contains rich information pertaining to the channel. Based on this observation, we develop a practical non-iterative algorithm for blind separation of digital signals with M-ASK and QAM alphabets. The proposed algorithm compares favorably with the existing hyperplane-based and kurtosis-based algorithms. Since only a small fraction of the

constellation geometry is exploited in our geometric approach for channel estimation, it is desirable for us to devise an efficient and more accurate geometric channel estimation approach in the future by utilizing the constellation geometry to a full extent.

In Chapter 6, we proposed a closed-form solution for blind channel estimation of MIMO-OFDM systems by exploiting the spectra correlations of the system output. The system is assumed to be driven by stationary or cyclostationary and nonwhite inputs with distinct but known correlations. Under this assumption, the channel can be uniquely estimated up to a complex scalar. Therefore, the problem of the ambiguity matrix which indeed exists in many existing channel estimation algorithms is avoided without any further process.

7.2 Future Research

In this section, a list of possible extension of the research work presented in this thesis is provided.

- The estimation error of the algorithm presented in Chapter 3 is raised when the system is driven by colored source. The main contribution of the estimation error comes from the estimation of the target matrix $\mathbb{R}_{\mathbf{x}}$, which is the summation of channel output correlation matrices with different delay lags. There may be a possibility to develop some solution to improve the performance under such situation.

- In Chapter 6, we proposed a blind channel estimation algorithm with the assis-

tance of non-redundant linear precoder. It has been shown that the estimation performance is closely related with the designing of the precoder. Generally, when the precoder delivers more correlations between the subcarriers, the estimator achieves better NRMSE performance. However, this may decrease the SNR which leads to a BER performance degradation. How to balance them and find out an optimized precoder is a possible research direction.

- The algorithms presented in Chapter 3 requires large number of observed symbols. This may cause problem in fast fading environments. Therefore, it is possible to design an adaptive method to solve this problem.

- Combination of the idea proposed in Chapter 4 and 6 may leed to a new research direction. In Chapter 4, we proposed a precoded MIMO-OFDM system where the precoder is identical for every user. If each user is assigned a distinct precoder which induces distinct source color, then the ambiguity matrix may be removed without any further process following the idea in Chapter 6.

Appendix A

The Proof of Theorem 3.4.3

A.1 Preliminaries

Before we proceed to prove the theorem, we propose the following two lemmas first.

Lemma A.1.1 Let $\mathbf{d}(k)$ be the i.i.d. source signal with unit signal power such that $\mathbf{R}_{\mathbf{d}}(\kappa) \triangleq E\{\mathbf{d}(k)\mathbf{d}^H(k+\kappa)\} = \mathbf{I}\delta(\kappa)$, and

$$\mathbf{x}_1(k) = \mathbf{A}_1\mathbf{d}(k) \tag{A.1a}$$

$$\mathbf{y}_1(k) = \mathbf{B}_1\mathbf{d}(k) \tag{A.1b}$$

$$\mathbf{x}_2(k) = \mathbf{A}_2\mathbf{d}(k) \tag{A.1c}$$

$$\mathbf{y}_2(k) = \mathbf{B}_2\mathbf{d}(k) \tag{A.1d}$$

where $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_1$ and $\mathbf{B}_2$ are arbitrary coding matrices with proper sizes. Define the cross correlation matrices $\mathbf{R}_{\mathbf{x}_1\mathbf{y}_1}(\kappa)$, $\mathbf{R}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)$ and their estimates by time averaging

as the following Eqn.(A.2a) to Eqn.(A.2d) respectively,

$$\mathbf{R}_{\mathbf{x}_1\mathbf{y}_1}(\kappa) \triangleq E\left\{\mathbf{x}(k)_1\mathbf{y}_1^H(k+\kappa)\right\} \tag{A.2a}$$

$$\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa) \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_1(k)\mathbf{y}_1^H(k+\kappa) \tag{A.2b}$$

$$\mathbf{R}_{\mathbf{x}_2\mathbf{y}_2}(\kappa) \triangleq E\left\{\mathbf{x}_2(k)\mathbf{y}_2^H(k+\kappa)\right\} \tag{A.2c}$$

$$\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa) \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_2(k)\mathbf{y}_2^H(k+\kappa) \tag{A.2d}$$

Then the estimation of the cross correlation matrices are unbias, and given any matrix $\mathbf{G}$ with proper size, we have

$$E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa)\mathbf{G}\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)\right\} \tag{A.3}$$

$$= \begin{cases} \frac{1}{K}\mathrm{tr}(\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2)\mathbf{A}_1\mathbf{A}_2^H & (\kappa \neq 0) \\ \mathbf{A}_1\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2\mathbf{A}_2^H + \frac{1}{K}\left[\mathrm{tr}(\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2)\mathbf{A}_1\mathbf{A}_2^H - \mathbf{A}_1\mathrm{diag}(\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2)\mathbf{A}_2^H\right] & (\kappa = 0) \end{cases}$$

Proof: The proof of this lemma is straightforward. Let's consider

$$E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa)\right\} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_1(k)\mathbf{y}_1^H(k+\kappa) = \mathbf{A}_1\mathbf{B}_1^H\delta(\kappa) = \mathbf{R}_{\mathbf{x}_1\mathbf{y}_1}(\kappa) \tag{A.4a}$$

$$E\left\{\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)\right\} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_2(k)\mathbf{y}_2^H(k+\kappa) = \mathbf{A}_2\mathbf{B}_2^H\delta(\kappa) = \mathbf{R}_{\mathbf{x}_2\mathbf{y}_2}(\kappa) \tag{A.4b}$$

which implies that the estimation of the correlation matrices are unbias.

On the other hand,

$$\begin{aligned} & E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa))\mathbf{G}\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}^H(\kappa)\right\} \\ = & \frac{1}{K^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K}E\left\{\mathbf{A}_1\mathbf{d}(k_1)\mathbf{d}^H(k_1+\kappa)\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2\mathbf{d}(k_2+\kappa)\mathbf{d}^H(k_2)\mathbf{A}_2^H\right\} \end{aligned} \tag{A.5}$$

Without loss of generality, we define

$$\mathbf{\Phi}(k_1,k_2,\kappa) \triangleq E\left\{\mathbf{A}_1\mathbf{d}(k_1)\mathbf{d}^H(k_1+\kappa)\mathbf{B}_1^H\mathbf{G}\mathbf{B}_2\mathbf{d}(k_2+\kappa)\mathbf{d}^H(k_2)\mathbf{A}_2^H\right\} \tag{A.6}$$

Consider the following two situations:

- $\kappa \neq 0$:

It can be verified that if $\kappa \neq 0$, $\mathbf{\Phi}(k_1, k_2, \kappa) \neq 0$ only when $k_1 = k_2$, i.e.,

$$\mathbf{\Phi}(k_1, k_2, \kappa) = \mathbf{A}_1 \mathrm{tr}\left(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2\right) \mathbf{A}_2^H \delta(k_1 - k_2) \tag{A.7}$$

Thus

$$\begin{aligned} E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa)\mathbf{G}\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)\right\} &= \frac{1}{K^2}\sum_{k_1=1}^{K}\sum_{k_2=1}^{K} \mathbf{A}_1 \mathrm{tr}\left(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2\right) \mathbf{A}_2^H \delta(k_1 - k_2) \\ &= \frac{1}{K}\mathrm{tr}\left(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2\right) \mathbf{A}_1 \mathbf{A}_2^H \end{aligned} \tag{A.8}$$

- $\kappa = 0$:

When $\kappa = 0$, we have

$$\begin{aligned} E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa)\mathbf{G}\hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)\right\} &= \frac{1}{K^2}\sum_{k_1,k_2=1}^{K} E\left\{\mathbf{A}_1 \mathbf{d}(k_1)\mathbf{d}^H(k_1)\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 \mathbf{d}(k_2)\mathbf{d}^H(k_2)\mathbf{A}_2^H\right\} \\ &= \frac{1}{K^2}\sum_{k_1,k_2=1,k_1\neq k_2}^{K} \mathbf{\Phi}(k_1, k_2, 0) + \frac{1}{K^2}\sum_{k=1}^{K}\mathbf{\Phi}(k, k, 0) \end{aligned} \tag{A.9}$$

The two terms on the right hand side of Eqn.(A.9) can be derived as follows respectively.

When $k_1 \neq k_2$

$$\begin{aligned} \mathbf{\Phi}(k_1, k_2, 0) &= \mathbf{A}_1 E\left\{\mathbf{d}(k_1)\mathbf{d}^H(k_1)\right\} \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 E\left\{\mathbf{d}(k_2)\mathbf{d}^H(k_2)\right\} \mathbf{A}_2^H \\ &= \mathbf{A}_1 \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 \mathbf{A}_2^H \end{aligned} \tag{A.10}$$

Thus

$$\frac{1}{K^2}\sum_{k_1,k_2=1,k_1\neq k_2}^{K} \mathbf{\Phi}(k_1, k_2, 0) = \frac{K-1}{K}\mathbf{A}_1 \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 \mathbf{A}_2^H \tag{A.11}$$

When $k_1 = k_2$, we consider the $(m, n)^{th}$ element of $\mathbf{\Phi}(k, k, 0)$, which is denoted as $\phi_{m,n}$. We also define $\omega_{i,j}$ as the $(i, j)^{th}$ element of the matrix $\mathbf{\Omega} \triangleq \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2$, and

write $\mathbf{d}(k)$ as

$$\mathbf{d}(k) \triangleq [d_0(k), \cdots, d_{N-1}(k)] \tag{A.12}$$

Thus it can be verified that

$$\begin{aligned}\phi_{m,n} &= \sum_{i,j,p,q=0}^{N-1} \omega_{i,j} a_1(m,p) a_2^*(n,q) E\{d_i^*(k) d_j(k) d_p(k) d_q^*(k)\} \qquad (A.13)\\ &= \sum_{i,p=0}^{N-1} \underbrace{\omega_{i,i} a_1(m,p) a_2^*(n,p)}_{i=j,\ p=q} + \sum_{i,j=0}^{N-1} \underbrace{\omega_{i,j} a_1(m,i) a_2^*(n,j)}_{i=p,\ j=q} - \sum_{i=0}^{N-1} \underbrace{\omega_{i,i} a_1(m,i) a_2^*(n,i)}_{i=j=p=q}\end{aligned}$$

where $a_1(m,p)$, $a_2(n,q)$ and $\omega_{i,j}$ respectively denote the $(m,p)^{th}$, $(n,q)^{th}$ and $(i,j)^{th}$ element of $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{\Omega}$. It can be verified that Eqn.(A.13) is equivalent to the following equation, which is written in the matrix form

$$\mathbf{\Phi}(k,k,0) = \mathrm{tr}(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2) \mathbf{A}_1 \mathbf{A}_2^H + \mathbf{A}_1 \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 \mathbf{A}_2^H - \mathbf{A}_1 \mathrm{diag}\left(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2\right) \mathbf{A}_2 \tag{A.14}$$

where $\mathrm{diag}(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2)$ denotes the diagonal matrix which contains the diagonal elements of the matrix $\mathbf{\Omega} = \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2$.

Substitute Eqn.(A.10) and Eqn.(A.14) into Eqn.(A.9), we can conclude that, when $\kappa = 0$

$$\begin{aligned}E\left\{\hat{\mathbf{R}}_{\mathbf{x}_1\mathbf{y}_1}(\kappa) \mathbf{G} \hat{\mathbf{R}}_{\mathbf{x}_2\mathbf{y}_2}(\kappa)\right\} &= \frac{1}{K}\left[\mathrm{tr}(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2) \mathbf{A}_1 \mathbf{A}_2^H - \mathbf{A}_1 \mathrm{diag}\left(\mathbf{B}_1^H \mathbf{G} \mathbf{B}_2\right) \mathbf{A}_2^H\right]\\ &\quad + \mathbf{A}_1 \mathbf{B}_1^H \mathbf{G} \mathbf{B}_2 \mathbf{A}_2^H \qquad (A.15)\end{aligned}$$

The proof ends here. ■

Lemma A.1.2 Let the length N column vector $\mathbf{x}(k) = \mathbf{A}\mathbf{d}(k)$, where $\mathbf{A}$ is an arbitrary coding matrices with proper sizes, and $\mathbf{d}(k)$ is i.i.d. source signal with unit power such that $\mathbf{R}_{\mathbf{d}}(\kappa) = \mathbf{I}\delta(\kappa)$. Let $\mathbf{v}(k)$ be the AWGN noise vector which is uncorrelated with $\mathbf{x}(k)$, and the noise power is σ_v^2. Define the cross correlation

matrices $\mathbf{R}_{\mathbf{xv}}(\kappa)$ and $\mathbf{R}_{\mathbf{vx}}(\kappa)$ and their estimation by time averaging as the following Eqn.(A.16a) to Eqn.(A.16d) respectively,

$$\mathbf{R}_{\mathbf{xv}}(\kappa) \triangleq E\left\{\mathbf{x}(k)\mathbf{v}^H(k+\kappa)\right\} \tag{A.16a}$$

$$\hat{\mathbf{R}}_{\mathbf{xv}}(\kappa) \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(k)\mathbf{v}^H(k+\kappa) \tag{A.16b}$$

$$\mathbf{R}_{\mathbf{vx}}(\kappa) \triangleq E\left\{\mathbf{v}(k)\mathbf{x}^H(k+\kappa)\right\} \tag{A.16c}$$

$$\hat{\mathbf{R}}_{\mathbf{vx}}(\kappa) \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbf{v}(k)\mathbf{x}^H(k+\kappa) \tag{A.16d}$$

Then the estimation is unbias, and given any matrix $\mathbf{G}$ with proper size, we have

$$E\left\{\hat{\mathbf{R}}_{\mathbf{xv}}(\kappa)\mathbf{G}\hat{\mathbf{R}}^H_{\mathbf{xv}}(\kappa)\right\} = \frac{\sigma_v^2}{K}\mathrm{tr}(\mathbf{G})\mathbf{A}\mathbf{A}^H \tag{A.17}$$

$$E\left\{\hat{\mathbf{R}}_{\mathbf{vx}}(\kappa)\mathbf{G}\hat{\mathbf{R}}^H_{\mathbf{vx}}(\kappa)\right\} = \frac{\sigma_v^2}{K}\mathrm{tr}(\mathbf{A}^H\mathbf{G}\mathbf{A})\mathbf{I} \tag{A.18}$$

Proof: Firstly, the unbias property of the estimation can be proved exactly by the same method for Lemma A.1.1, and is omitted for brevity.

Secondly, since $\mathbf{x}(k)$ and $\mathbf{v}(k)$ are uncorrelated, then Eqn.(A.17) and Eqn.(A.18) can be rewritten as follows respectively

$$\begin{aligned} E\left\{\hat{\mathbf{R}}_{\mathbf{xv}}(\kappa)\mathbf{G}\hat{\mathbf{R}}^H_{\mathbf{xv}}(\kappa)\right\} &= \frac{1}{K^2}\sum_{k=1}^{K}E\left\{\mathbf{A}\mathbf{d}(k)\mathbf{v}^H(k+\kappa)\mathbf{G}\mathbf{v}(k+\kappa)\mathbf{d}^H(k)\mathbf{A}^H\right\} \\ &= \frac{1}{K^2}\sum_{k=1}^{K}E\left\{\mathbf{A}\mathbf{d}(k)\mathrm{tr}(\mathbf{I})\sigma_v^2\mathbf{d}^H(k)\mathbf{A}^H\right\} \\ &= \frac{\sigma_v^2}{K}\mathrm{tr}(\mathbf{G})\mathbf{A}\mathbf{A}^H \end{aligned} \tag{A.19}$$

$$\begin{aligned} E\left\{\hat{\mathbf{R}}_{\mathbf{vx}}(\kappa)\mathbf{G}\hat{\mathbf{R}}^H_{\mathbf{vx}}(\kappa)\right\} &= \frac{1}{K^2}\sum_{k=1}^{K}E\left\{\mathbf{v}(k)\mathbf{d}^H(k+\kappa)\mathbf{A}^H\mathbf{G}\mathbf{A}\mathbf{d}(k+\kappa)\mathbf{v}^H(k)\mathbf{A}^H\right\} \\ &= \frac{1}{K^2}\sum_{k=1}^{K}E\left\{\mathbf{v}(k)\mathrm{tr}(\mathbf{A}^H\mathbf{G}\mathbf{A})\mathbf{v}^H(k)\right\} \\ &= \frac{\sigma_v^2}{K}\mathrm{tr}(\mathbf{A}^H\mathbf{G}\mathbf{A})\mathbf{I} \end{aligned} \tag{A.20}$$

The proof ends here. ■

A.2 Proof of Theorem 3.4.3

Now, we derive the proof of Theorem 3.4.3.

Theorem 3.4.3 Assume that the noise is zero-mean i.i.d with covariance σ_v^2, and the transmitted signal is modeled as $\mathbf{s}(k) = \mathbf{P}\mathbf{d}(k)$, where $\mathbf{d}(k)$ is the i.i.d. source signal with unit power, and $\mathbf{P}$ is the non-redundant precoding matrix. Then the covariance matrix of the channel estimation error is approximated by

$$E\left\{\Delta\vec{\mathbf{h}}\Delta\vec{\mathbf{h}}^H\right\} = \left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] \mathcal{E} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right] \tag{A.21}$$

and the channel estimation MSE is

$$E\left\{\|\Delta\vec{\mathbf{h}}\|^2\right\} = \mathrm{tr}\{\left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] \mathcal{E} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right]\} \tag{A.22}$$

Proof: According to the definition of $\boldsymbol{\Xi}$ (see Eqn.(3.40)), we can express the perturbation of $\boldsymbol{\Xi}$ as

$$\begin{aligned}
\Delta\boldsymbol{\Xi} &= \hat{\boldsymbol{\Xi}} - \boldsymbol{\Xi} \\
&= \hat{\mathbb{R}}_{\mathbf{x}} - \sigma_v^2\mathbf{I} \\
&= \frac{1}{K}\sum_{\kappa=-1}^{1}\sum_{k=1}^{K+1}\mathbf{x}(k)\mathbf{x}^H(k+\kappa) - \boldsymbol{\Xi} - \sigma_v^2\mathbf{I} \\
&= \hat{\mathbf{R}} + \hat{\boldsymbol{\Phi}} + \hat{\boldsymbol{\Upsilon}} - (\boldsymbol{\Xi} + \sigma_v^2\mathbf{I})
\end{aligned} \tag{A.23}$$

where

$$\hat{\mathbf{R}} \approx \frac{1}{K}\sum_{k=1}^{K}\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k)\mathbf{s}^H(k)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) \tag{A.24}$$

where the approximation in the above Eqn.(A.24) is based on the fact that when K is large,

$$\frac{1}{K}\sum_{k=1}^{K}\mathbf{s}(k)\mathbf{s}^H(k+\kappa) \approx \frac{1}{K}\sum_{k=1}^{K}\mathbf{s}(k-1)\mathbf{s}^H(k-1+\kappa) \tag{A.25}$$

and

$$\begin{aligned}\hat{\Phi} = \frac{1}{K}\sum_{k=1}^{K}\Big\{&\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k)\mathbf{s}^H(k-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h}) + \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k-1)\mathbf{s}^H(k)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h})\\ &+\ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k-1)\mathbf{s}^H(k+1)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h}) + \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k+1)\mathbf{s}^H(k-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h})\\ &+\dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\left[\mathbf{s}(k+1)\mathbf{s}^H(k)+\mathbf{s}(k)\mathbf{s}^H(k+1)\right]\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h})\Big\}\end{aligned} \tag{A.26}$$

$$\begin{aligned}\hat{\boldsymbol{\Upsilon}} = \frac{1}{K}\sum_{k=1}^{K+1}\Big\{&\mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k)\mathbf{v}^H(k) + \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k-1)\mathbf{v}^H(k) + \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k+1)\mathbf{v}^H(k)\\ &+\mathbf{v}(k)\mathbf{s}^H(k)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) + \mathbf{v}(k)\mathbf{s}^H(k-1)\tilde{\mathbf{F}}_{cp}^H\mathcal{C}_N^H(\mathbf{h}) + \mathbf{v}(k)\mathbf{s}^H(k+1)\tilde{\mathbf{F}}_{cp}^H\dot{\mathcal{T}}_N^H(\mathbf{h})\\ &+\ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{s}(k-1)\mathbf{v}^H(k+1) + \mathbf{v}(k+1)\mathbf{s}^H(k-1)\tilde{\mathbf{F}}_{cp}^H\ddot{\mathcal{T}}_N^H(\mathbf{h})\\ &+\mathbf{v}(k)\mathbf{v}^H(k) + \mathbf{v}(k)\mathbf{v}^H(k+1) + \mathbf{v}(k+1)\mathbf{v}^H(k)\Big\}\end{aligned} \tag{A.27}$$

It is assumed that the transmitted signal $\mathbf{s}(k) = \mathbf{P}\mathbf{d}(k)$, where $\mathbf{d}(k)$ is the i.i.d. source signal, and $\mathbf{s}(k)$ is uncorrelated with the noise vector $\mathbf{v}(k)$. Then it is obvious that $E\{\mathbf{s}(k)\mathbf{s}^H(k+\kappa)\} = \mathbf{0}$ for $\kappa > 0$, and $E\{\mathbf{s}(k)\mathbf{v}^H(k+\kappa)\} = \mathbf{0}$ for any κ. Thus, we have

$$E\left\{\hat{\mathbf{R}} + \hat{\boldsymbol{\Phi}} + \hat{\boldsymbol{\Upsilon}}\right\} = \boldsymbol{\Xi} + \sigma_v^2\mathbf{I} \tag{A.28}$$

and consequently, the expectation of the perturbation of $\boldsymbol{\Xi}$ can be expressed as

$$E\left\{\Delta\boldsymbol{\Xi}\right\} = E\left\{\hat{\mathbf{R}} + \hat{\boldsymbol{\Phi}} + \hat{\boldsymbol{\Upsilon}}\right\} - (\boldsymbol{\Xi} + \sigma_v^2\mathbf{I}) = \mathbf{0} \tag{A.29}$$

which indicates that the estimation of $\boldsymbol{\Xi}$ is unbias.

Next, we consider the covariance matrix of the channel estimation error. Define:

$$\begin{aligned}
\mathcal{E}_{i_1,i_2}(m,n) &\triangleq E\left\{\Delta\Xi^H(\Xi^H)^\dagger\mathbb{C}_m\mathbf{h}(:,i_1)\mathbf{h}^H(:,i_2)\mathbb{C}_n^H\Xi^\dagger\Delta\Xi\right\} \\
&= E\left\{\Delta\Xi^H\left(\mathcal{C}_N^H(\mathbf{h})\right)^\dagger\left(\tilde{\mathbf{F}}_{cp}^H\right)^\dagger\mathbb{R}_{\mathbf{s}}^{-1}\tilde{\mathbf{F}}_{N_c}\mathbf{e}_{mM_t+i_1}\mathbf{e}_{nM_t+i_2}^H\tilde{\mathbf{F}}_{N_c}^H\mathbb{R}_{\mathbf{s}}^{-1}\tilde{\mathbf{F}}_{cp}^\dagger\mathcal{C}_N^\dagger(\mathbf{h})\Delta\Xi\right\} \\
&= E\left\{(\hat{\mathbf{R}}+\hat{\mathbf{\Phi}}+\hat{\mathbf{\Upsilon}})\mathcal{G}_{i_1,i_2}(m,n)(\hat{\mathbf{R}}+\hat{\mathbf{\Phi}}+\hat{\mathbf{\Upsilon}})^H\right\} \\
&\quad -(\Xi+\sigma_v^2\mathbf{I})\mathcal{G}_{i_1,i_2}(m,n)(\Xi+\sigma_v^2\mathbf{I})^H
\end{aligned} \tag{A.30}$$

where $i_1, i_2 = 1, \cdots, M_t$, $m, n = 0, \cdots, N_c - 1$, and the matrix $\mathcal{G}_{i_1,i_2}(m,n)$ is defined as

$$\mathcal{G}_{i_1,i_2}(m,n) \triangleq \left(\mathcal{C}_N^H(\mathbf{h})\right)^\dagger\left(\tilde{\mathbf{F}}_{cp}^H\right)^\dagger\mathbb{R}_{\mathbf{s}}^{-1}\tilde{\mathbf{F}}_{N_c}\mathbf{e}_{mM_t+i_1}\mathbf{e}_{nM_t+i_2}^H\tilde{\mathbf{F}}_{N_c}^H\mathbb{R}_{\mathbf{s}}^{-1}\tilde{\mathbf{F}}_{cp}^\dagger\mathcal{C}_N^\dagger(\mathbf{h}) \tag{A.31}$$

Besides, we also define the following notations to simplify the discussion

$$\mathcal{A} \triangleq \dot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{P} \tag{A.32a}$$

$$\mathcal{B} \triangleq \ddot{\mathcal{T}}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{P} \tag{A.32b}$$

$$\mathcal{C} \triangleq \mathcal{A}+\mathcal{B} = \mathcal{C}_N(\mathbf{h})\tilde{\mathbf{F}}_{cp}\mathbf{P} \tag{A.32c}$$

Thus, according to Lemma A.1.1 and Lemma A.1.2, we have

$$\begin{aligned}
&E\left\{\hat{\mathbf{R}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{R}}^H\right\} \\
=\ &\Xi\mathcal{G}_{i_1,i_2}(m,n)\Xi^H+\frac{1}{K}\left[\mathrm{tr}\left(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C}\right)\Xi-\mathcal{C}\mathrm{diag}\left(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C}\right)\mathcal{C}^H\right] \\
=\ &\Xi\mathcal{G}_{i_1,i_2}(m,n)\Xi^H+\frac{1}{K}\left[\mathrm{tr}\left(\mathbf{P}^{-1}\tilde{\mathbf{F}}_{N_c}\mathbf{e}_{mM_t+i_1}\mathbf{e}_{nM_t+i_2}^H\tilde{\mathbf{F}}_{N_c}^H(\mathbf{P}^{-1})^H\right)\Xi\right. \\
&\left.-\mathcal{C}\mathrm{diag}\left(\mathbf{P}^{-1}\tilde{\mathbf{F}}_{N_c}\mathbf{e}_{mM_t+i_1}\mathbf{e}_{nM_t+i_2}^H\tilde{\mathbf{F}}_{N_c}^H(\mathbf{P}^{-1})^H\right)\mathcal{C}^H\right]
\end{aligned} \tag{A.33}$$

$$\begin{aligned}
&E\left\{\hat{\mathbf{\Phi}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Phi}}^H\right\} \\
=\ &\frac{1}{K}\left[\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{C}\mathcal{C}^H+\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})\mathcal{B}\mathcal{B}^H+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{B}\mathcal{B}^H\right. \\
&+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{A}\mathcal{A}^H+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})\mathcal{A}\mathcal{C}^H+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{A}\mathcal{B}^H \\
&\left.+\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{C}\mathcal{A}^H+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{B}\mathcal{A}^H\right]
\end{aligned} \tag{A.34}$$

$$
\begin{aligned}
& E\left\{\hat{\mathbf{\Upsilon}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Upsilon}}^H\right\} \\
= \; & \frac{\sigma_v^2}{K}\mathrm{tr}(\mathcal{G}_{i_1,i_2}(m,n))\left(2\mathcal{C}\mathcal{C}^H+\mathcal{A}\mathcal{A}^H+\mathcal{B}\mathcal{B}^H\right)+\frac{2\sigma_v^4}{K}\mathrm{tr}(\mathcal{G}_{i_1,i_2}(m,n))\mathbf{I} \\
& +\frac{\sigma_v^2}{K}\left[2\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\right]\mathbf{I} \\
& +\sigma_v^4\mathcal{G}_{i_1,i_2}(m,n)+\frac{\sigma_v^4}{K}\left[\mathrm{tr}(\mathcal{G}_{i_1,i_2}(m,n))\mathbf{I}-\mathrm{diag}(\mathcal{G}_{i_1,i_2}(m,n))\right] \qquad \text{(A.35)}
\end{aligned}
$$

$$
E\left\{\hat{\mathbf{R}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Upsilon}}^H\right\} = \sigma_v^2\mathbf{\Xi}\mathcal{G}_{i_1,i_2}(m,n) \qquad \text{(A.36)}
$$

$$
E\left\{\hat{\mathbf{\Upsilon}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{R}}^H\right\} = \sigma_v^2\mathcal{G}_{i_1,i_2}(m,n)\mathbf{\Xi}^H \qquad \text{(A.37)}
$$

and

$$
\begin{aligned}
& E\left\{\hat{\mathbf{R}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Phi}}^H\right\} = E\left\{\hat{\mathbf{\Phi}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{R}}^H\right\} \\
= \; & E\left\{\hat{\mathbf{\Upsilon}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Phi}}^H\right\} = E\left\{\hat{\mathbf{\Phi}}\mathcal{G}_{i_1,i_2}(m,n)\hat{\mathbf{\Upsilon}}^H\right\} = \mathbf{0} \qquad \text{(A.38)}
\end{aligned}
$$

Thus, Eqn.(A.30) can be written as

$$
\begin{aligned}
\mathcal{E}_{i_1,i_2}(m,n) &= E\left\{\Delta\mathbf{\Xi}^H(\mathbf{\Xi}^H)^\dagger\mathbb{C}_m\mathbf{h}(:,i_1)\mathbf{h}^H(:,i_2)\mathbb{C}_n^H\mathbf{\Xi}^\dagger\Delta\mathbf{\Xi}\right\} \\
&= \frac{1}{K}\left[\mathcal{X}_{i_1,i_2}^{(1)}(m,n)+\sigma_v^2\mathcal{X}_{i_1,i_2}^{(2)}(m,n)+\sigma_v^4\mathcal{X}_{i_1,i_2}^{(3)}(m,n)\right] \qquad \text{(A.39)}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{X}_{i_1,i_2}^{(1)}(m,n) &= \mathrm{tr}\left(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C}\right)\mathbf{\Xi}-\mathcal{C}\mathrm{diag}\left(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C}\right)\mathcal{C}^H \\
& +\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{C}\mathcal{C}^H+\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})\mathcal{B}\mathcal{B}^H \\
& +\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{B}\mathcal{B}^H+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{A}\mathcal{A}^H \\
& +\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})\mathcal{A}\mathcal{C}^H+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\mathcal{A}\mathcal{B}^H \\
& +\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{C}\mathcal{A}^H+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})\mathcal{B}\mathcal{A}^H \qquad \text{(A.40a)}
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{X}_{i_1,i_2}^{(2)}(m,n) &= \left[2\mathrm{tr}(\mathcal{C}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{C})+\mathrm{tr}(\mathcal{A}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{A})+\mathrm{tr}(\mathcal{B}^H\mathcal{G}_{i_1,i_2}(m,n)\mathcal{B})\right]\mathbf{I} \\
& +\mathrm{tr}(\mathcal{G}_{i_1,i_2}(m,n))\left(2\mathcal{C}\mathcal{C}^H+\mathcal{A}\mathcal{A}^H+\mathcal{B}\mathcal{B}^H\right) \qquad \text{(A.40b)}
\end{aligned}
$$

$$\mathcal{X}^{(3)}_{i_1,i_2}(m,n) \quad = \quad \mathrm{tr}(\mathcal{G}_{i_1,i_2}(m,n))\mathbf{I} - \mathrm{diag}(\mathcal{G}_{i_1,i_2}(m,n)) \tag{A.40c}$$

Hence, the covariance matrix of the channel estimation error can be expressed as

$$E\left\{\Delta\vec{\mathbf{h}}\Delta\vec{\mathbf{h}}^H\right\} \quad = \quad \left[\mathbf{I}_{M_t} \otimes (\boldsymbol{\mathcal{K}}^H)^\dagger\right] \mathcal{E} \left[\mathbf{I}_{M_t} \otimes \boldsymbol{\mathcal{K}}^\dagger\right] \tag{A.41}$$

and the channel estimation MSE can be expressed as

$$E\left\{\|\Delta\vec{\mathbf{h}}\Delta\vec{\mathbf{h}}^H\|^2\right\} \quad = \quad \sum_{i=1}^{M_t} \mathrm{tr}\left\{(\boldsymbol{\mathcal{K}}^H)^\dagger \mathcal{E}_{i,i} \boldsymbol{\mathcal{K}}^\dagger\right\} \tag{A.42}$$

where

$$\mathcal{E} \quad = \quad \begin{bmatrix} \mathcal{E}_{1,1} & \cdots & \mathcal{E}_{1,M_t} \\ \vdots & \ddots & \vdots \\ \mathcal{E}_{M_t,1} & \cdots & \mathcal{E}_{M_t,M_t} \end{bmatrix} \tag{A.43}$$

and

$$\mathcal{E}_{i_1,i_2} \quad = \quad \begin{bmatrix} \mathbf{U}_n^H \mathcal{E}_{i_1,i_2}(0,0)\mathbf{U}_n & \cdots & \mathbf{U}_n^H \mathcal{E}_{i_1,i_2}(0,N_c-1)\mathbf{U}_n \\ \vdots & \ddots & \vdots \\ \mathbf{U}_n^H \mathcal{E}_{i_1,i_2}(N_c-1,0)\mathbf{U_n} & \cdots & \mathbf{U}_n^H \mathcal{E}_{i_1,i_2}(N_c-1,N_c-1)\mathbf{U}_n \end{bmatrix} \tag{A.44}$$

The proof ends here. ■

Appendix B

The Proof of Theorem 4.4.2

Theorem 4.4.2 The proposed blind channel estimator $\hat{\mathbf{H}}_j$ is asymptotically unbiased (i.e. $E\{\Delta\mathbf{H}_j\} = \mathbf{0}$), and the estimated channel MSE is

$$\text{MSE} \quad = \quad \frac{1}{K}\sum_{i=1}^{M_t} \text{tr}\left(\mathbf{U}_n^H \mathcal{E}_{i,i} \mathbf{U}_n\right) \tag{B.1}$$

Proof: According to the discussion in Section 4.4.4, the first order perturbation of $\mathbf{U}_s$ can be written as

$$\Delta\mathbf{U}_s = \mathbf{U}_n \mathbf{U}_n^H \Delta\mathbb{R}_{\mathbf{y}_j} \mathbf{V}_s \mathbf{\Lambda}_s^{-1} \tag{B.2}$$

where $\Delta\mathbb{R}_{\mathbf{y}_j} = \Delta\mathbf{R}_{\mathbf{y}_j} \oslash \mathbf{\Phi}$, and $\Delta\mathbb{R}_{\mathbf{y}_j} \triangleq \hat{\mathbf{R}}_{\mathbf{y}_j} - \mathbf{R}_{\mathbf{y}_j}$ is the estimation error of the autocorrelation matrix of the received signal. Note that $\mathbb{R}_{\mathbf{y}_j}$ is Hermitian, which indicates that $\mathbf{U}_s = \mathbf{V}_s$. Substitute to Eqn.(4.44), the channel estimation error can be written as

$$\Delta\mathbf{H}_j \quad \triangleq \quad \hat{\mathbf{H}}_j - \tilde{\mathbf{H}}_j = \mathbf{U}_n \mathbf{U}_n^H \Delta\mathbb{R}_{\mathbf{y}_j} \mathbf{U}_s \mathbf{\Lambda}_s^{-\frac{1}{2}} \tag{B.3}$$

and

$$\Delta\mathbf{H}_{ji} \quad \triangleq \quad \hat{\mathbf{H}}_{ji} - \tilde{\mathbf{H}}_{ji} = \mathbf{U}_n \mathbf{U}_n^H \Delta\mathbb{R}_{\mathbf{y}_j} \mathbf{u}_{s,i} \lambda_{s,i}^{-\frac{1}{2}} \tag{B.4}$$

where $\lambda_{s,i}$ is the ith diagonal entry of $\mathbf{\Lambda}_s$, and $\mathbf{u}_{s,i}$ is the corresponding eigenvector.

Thus, when $\mathbf{R}_{\mathbf{y}_j}$ is estimated by Eqn.(4.16a), we have

$$E\{\Delta\mathbf{R}_{\mathbf{y}_j}\} = \mathbf{0} \Rightarrow E\{\Delta\mathbb{R}_{\mathbf{y}_j}\} = \mathbf{0} \tag{B.5}$$

and consequently,

$$\text{Bias} = \mathbf{U}_n\mathbf{U}_n^H E\{\Delta\mathbb{R}_{\mathbf{y}_j} \oslash \mathbf{P}\}\mathbf{U}_s\mathbf{\Sigma}_s^{-\frac{1}{2}} = \mathbf{0} \tag{B.6}$$

Next, we consider the covariance matrix

$$\begin{aligned} E\left\{\Delta\mathbf{H}_{ji_1}\Delta\mathbf{H}_{ji_2}^H\right\} &= \frac{1}{\sqrt{\lambda_{s,i_1}\lambda_{s,i_2}}}\mathbf{U}_n\mathbf{U}_n^H E\left\{\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{u}_{s,i}\mathbf{u}_{s,k}^H\Delta\mathbb{R}_{\mathbf{y}_j}^H\right\}\mathbf{U}_n\mathbf{U}_n^H \\ &= \mathbf{U}_n\mathbf{U}_n^H\left[E\left\{\hat{\mathbb{R}}_{\mathbf{y}_j}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\hat{\mathbb{R}}_{\mathbf{y}_j}^H\right\} - \mathbb{R}_{\mathbf{y}_j}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbb{R}_{\mathbf{y}_j}^H\right]\mathbf{U}_n\mathbf{U}_n^H \\ &= \mathbf{U}_n\mathbf{U}_n^H E\left\{\hat{\mathbb{R}}_{\mathbf{y}_j}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\hat{\mathbb{R}}_{\mathbf{y}_j}^H\right\}\mathbf{U}_n\mathbf{U}_n^H \end{aligned} \tag{B.7}$$

where $i_1, i_2 = 1, \cdots, M_t$, $\boldsymbol{\mu}_{i_1}$ and $\boldsymbol{\mu}_{i_2}$ are the $N_c \times 1$ vectors defined as

$$\boldsymbol{\mu}_i = \lambda_{s,i}^{-\frac{1}{2}}\mathbf{u}_{s,i} \quad \text{for} \quad i = i_1,\ i_2 \tag{B.8}$$

Note that the last step of the above Eqn.(B.7) is based on the fact that

$$\begin{aligned} & \mathbf{U}_n\mathbf{U}_n^H\mathbb{R}_{\mathbf{y}_j}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbb{R}_{\mathbf{y}_j}^H\mathbf{U}_n\mathbf{U}_n^H \\ = \ & \mathbf{U}_n\mathbf{U}_n^H\mathbf{H}_j\mathbf{H}_j^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{H}_j\mathbf{H}_j^H\mathbf{U}_n\mathbf{U}_n^H + \sigma_n^2\mathbf{U}_n\mathbf{U}_n^H\mathbf{H}_j\mathbf{H}_j^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{U}_n\mathbf{U}_n^H \\ & + \sigma_n^2\mathbf{U}_n\mathbf{U}_n^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{H}_j\mathbf{H}_j^H\mathbf{U}_n\mathbf{U}_n^H + \sigma_n^4\mathbf{U}_n\mathbf{U}_n^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{U}_n\mathbf{U}_n^H \\ = \ & \mathbf{0} \end{aligned} \tag{B.9}$$

According to Eqn.(4.16a), we have

$$\begin{aligned}
& E\left\{\hat{\mathbb{R}}_{\mathbf{y}_j}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\hat{\mathbb{R}}_{\mathbf{y}_j}^H\right\} \\
= & \frac{1}{K^2}\sum_{k_1,k_2=1}^{K} E\left\{\left[\left(\mathbf{y}_j(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{y}_j(k_2)\mathbf{y}_j^H(k_2)\right)\oslash\boldsymbol{\Phi}\right]\right\} \\
= & \frac{1}{K^2}\left(\sum_{k_1\neq k_2} E\left\{\left(\mathbf{y}_j(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\boldsymbol{\Phi}\right\}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H E\left\{\left(\mathbf{y}_j(k_2)\mathbf{y}_j^H(k_2)\right)\oslash\boldsymbol{\Phi}\right\}\right. \\
& \left.+\sum_{k=1}^{K} E\left\{\left[\left(\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\}\right) \qquad \text{(B.10)}
\end{aligned}$$

where $\mathbf{y}_j(k)$ is the kth OFDM symbol received by the jth receive antenna, which is defined in Eqn.(4.3). Note that

$$E\left\{\left(\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right\} = \mathbf{H}_j\mathbf{H}_j^H+\sigma_n^2\mathbf{I} \qquad \text{(B.11)}$$

Hence, the first term of the right hand side of Eqn.(B.10) can be derived as

$$\begin{aligned}
& \frac{1}{K^2}\sum_{k_1\neq k_2}\left(E\left\{\left(\mathbf{y}_j(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\boldsymbol{\Phi}\right\}\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H E\left\{\left(\mathbf{y}_j(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\boldsymbol{\Phi}\right\}\right) \\
= & \frac{K-1}{K}\left(\mathbf{H}_j\mathbf{H}_j^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{H}_j\mathbf{H}_j^H+\sigma_n^2\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\right) \qquad \text{(B.12)}
\end{aligned}$$

According to Eqn.(B.9), we have

$$\mathbf{U}_n\mathbf{U}_n^H\left(\mathbf{H}_j\mathbf{H}_j^H\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\mathbf{H}_j\mathbf{H}_j^H+\sigma_n^2\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\right)\mathbf{U}_n\mathbf{U}_n^H=\mathbf{0} \qquad \text{(B.13)}$$

Thus, substitute Eqn.(B.12) to Eqn.(B.7), and we get

$$\begin{aligned}
& E\left\{\Delta\mathbf{H}_{ji_1}\Delta\mathbf{H}_{ji_2}^H\right\} \\
= & \frac{1}{K}\mathbf{U}_n\mathbf{U}_n^H E\left\{\left[\left(\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{y}_j(k)\mathbf{y}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\}\mathbf{U}_n\mathbf{U}_n^H \\
= & \frac{1}{K}\mathbf{U}_n\mathbf{U}_n^H\left(E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\}\right. \\
& +E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\} \\
& +E\left\{\left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\} \\
& \left.+E\left\{\left[\left(\mathbf{n}_j(k)\mathbf{n}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{n}_j(k)\mathbf{n}_j^H(k)\right)\oslash\boldsymbol{\Phi}\right]\right\}\right)\mathbf{U}_n\mathbf{U}_n^H \qquad \text{(B.14)}
\end{aligned}$$

where $\tilde{\mathbf{y}}_j(k)$ is defined as the k noise free OFDM symbol received by the jth receive antenna, i.e.,

$$\tilde{\mathbf{y}}_j(k) = \sum_{t=1}^{M_t} \mathcal{D}(\mathbf{H}_{jt})\mathbf{s}_t(k) \tag{B.15}$$

Define

$$\begin{aligned}
\boldsymbol{\Xi}_{i_1,i_2} &\triangleq E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right] \boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right]\right\} && \text{(B.16a)}\\
\boldsymbol{\Psi}_{i_1,i_2} &\triangleq E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right] \boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right]\right\} && \text{(B.16b)}\\
\boldsymbol{\Upsilon}_{i_1,i_2} &\triangleq E\left\{\left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right] \boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right]\right\} && \text{(B.16c)}\\
\boldsymbol{\Omega}_{i_1,i_2} &\triangleq E\left\{\left[\left(\mathbf{n}_j(k)\mathbf{n}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right] \boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\mathbf{n}_j(k)\mathbf{n}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right]\right\} && \text{(B.16d)}
\end{aligned}$$

Then we consider the $(m,n)^{th}$ elements of $\boldsymbol{\Xi}_{i_1,i_2}$, $\boldsymbol{\Psi}_{i_1,i_2}$, $\boldsymbol{\Upsilon}_{i_1,i_2}$, and $\boldsymbol{\Omega}_{i_1,i_2}$ which are denoted as $\xi_{i_1,i_2}(m,n)$, $\psi_{i_1,i_2}(m,n)$, $\upsilon_{i_1,i_2}(m,n)$ and $\omega_{i_1,i_2}(m,n)$ respectively.

Calculation of $\boldsymbol{\Xi}$

According to the definition, $\boldsymbol{\Xi}_{i_1,i_2}$ can be expressed as

$$\begin{aligned}
\boldsymbol{\Xi}_{i_1,i_2} &\triangleq E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right] \boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \boldsymbol{\Phi}\right]\right\}\\
&= \sum_{t_1,t_2,t_3,t_4=1}^{M_t} E\left\{\left[\left(\mathcal{D}(\mathbf{H}_{jt_1})\mathbf{s}_{t_1}(k)\mathbf{s}_{t_2}^H(k)\mathcal{D}^H(\mathbf{H}_{j,t_2})\right) \oslash \boldsymbol{\Phi}\right]\right.\\
&\quad \left.\cdot\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H \left[\left(\mathcal{D}(\mathbf{H}_{jt_4})\mathbf{s}_{t_4}(k)\mathbf{s}_{t_3}^H(k)\mathcal{D}^H(\mathbf{H}_{jt_3})\right) \oslash \boldsymbol{\Phi}\right]\right\} && \text{(B.17)}
\end{aligned}$$

Then, the $(m,n)^{th}$ element of $\mathbf{\Xi}_{i_1,i_2}$, $\xi_{i_1,i_2}(m,n)$ can be written as

$$\begin{aligned}
\xi_{i_1,i_2}(m,n) &= \sum_{t_1,t_2,t_3,t_4=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\varrho_{i_1,i_2}(p,q)}{\phi_{m,p}\phi_{q,n}}H_{jt_1}[m]H^*_{jt_2}[p]H_{jt_4}[q]H^*_{jt_3}[n] \\
&\quad \cdot E\left\{s_{t_1}(k,m)s^*_{t_2}(k,p)s_{t_4}(k,q)s^*_{t_3}(k,n)\right\} \\
&\doteq \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{m,n}\phi_{q,p}}{\phi_{m,p}\phi_{q,n}}H_{jt_1}[m]H^*_{jt_2}[p]H_{jt_2}[q]H^*_{jt_1}[n]\varrho_{i_1,i_2}(p,q) \\
&\quad + \sum_{t_1,t_3=1}^{M_t}\sum_{p,q=0}^{N_c-1}H_{jt_1}[m]H^*_{jt_1}[p]H_{jt_3}[q]H^*_{jt_3}[n]\varrho_{i_1,i_2}(p,q) \\
&\quad - \sum_{t=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\beta(m,n,p,q)}{\phi_{m,p}\phi_{q,n}}H_{jt}[m]H^*_{jt}[p]H_{jt}[q]H^*_{jt}[n]\varrho_{i_1,i_2}(p,q) \qquad \text{(B.18)}
\end{aligned}$$

where $\varrho_{i_1,i_2}(p,q)$ is defined as the $(p,q)^{th}$ element of the matrix $(\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}^H_{i_2})$, which can be expressed as

$$\varrho_{i_1,i_2}(p,q) = \frac{1}{\sqrt{\lambda_{s,i_1}\lambda_{s,i_1}}}u_{s,i_1}[p]u^*_{s,i_2}[q] \qquad \text{(B.19)}$$

and

$$\beta(m,n,p,q) \triangleq \sum_{i=0}^{N_c-1}\mathrm{p}_{m,i}\mathrm{p}^*_{p,i}\mathrm{p}_{q,i}\mathrm{p}^*_{n,i} \qquad \text{(B.20)}$$

where $\mathrm{p}_{m,n}$ is the $(m,n)^{th}$ element of the precoding matrix $\mathbf{P}$.

Eqn.(B.18) also implies that

$$\mathbf{\Xi}_{i_1,i_2} = \dot{\mathbf{\Xi}}_{i_1,i_2} + \ddot{\mathbf{\Xi}}_{i_1,i_2} \qquad \text{(B.21)}$$

where

$$\dot{\mathbf{\Xi}}_{i_1,i_2} = \mathbf{H}_j\mathbf{H}^H_j\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}^H_{i_2}\mathbf{H}_j\mathbf{H}^H_j \qquad \text{(B.22)}$$

and $\ddot{\mathbf{\Xi}}_{i_1,i_2}$ is the matrix with the $(m,n)^{th}$ element, $\ddot{\xi}_{i_1,i_2}(m,n)$, being

$$\begin{aligned}
\ddot{\xi}_{i_1,i_2}(m,n) &= \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{m,n}\phi_{q,p}}{\phi_{m,p}\phi_{q,n}}H_{jt_1}[m]H^*_{jt_2}[p]H_{jt_2}[q]H^*_{jt_1}[n]\varrho_{i_1,i_2}(p,q) \\
&\quad - \sum_{t=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\beta(m,n,p,q)}{\phi_{m,p}\phi_{q,n}}H_{jt}[m]H^*_{jt}[p]H_{jt}[q]H^*_{jt}[n]\varrho_{i_1,i_2}(p,q) \qquad \text{(B.23)}
\end{aligned}$$

where $u_{s,i}(p)$ $(i = i_1,\ i_2)$ denotes the ith element of the eigenvector $\mathbf{u}_{s,i}$. Meanwhile, it should be noted that according to Eqn.(B.9)

$$\mathbf{U}_n\mathbf{U}_n^H\dot{\mathbf{\Xi}}\mathbf{U}_n\mathbf{U}_n^H = \mathbf{0} \tag{B.24}$$

Calculation of $\mathbf{\Psi}$

According to the definition, $\mathbf{\Psi}_{i_1,i_2}$ can be expressed as

$$\begin{aligned}\mathbf{\Psi}_{i_1,i_2} &\triangleq E\left\{\left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\right\}\\ &= \sum_{t_1,t_2=1}^{M_t} E\left\{\left[\left(\mathcal{D}(\mathbf{H}_{jt_1})\mathbf{s}_{t_1}(k)\mathbf{n}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\right.\\ &\quad\left.\cdot\,\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathbf{n}_j(k)\mathbf{s}_{t_2}^H(k)\mathcal{D}^H(\mathbf{H}_{jt_2})\right)\oslash\mathbf{\Phi}\right]\right\}\end{aligned} \tag{B.25}$$

Thus, the $(m,n)^{th}$ element of Ψ_{i_1,i_2}, $\psi_{i_1,i_2}(m,n)$ can be written as

$$\begin{aligned}\psi_{i_1,i_2}(m,n) &= \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\varrho_{p,q}}{\phi_{m,p}\phi_{q,n}}H_{jt_1}[m]H_{jt_2}^*[n]\\ &\quad\cdot E\left\{s_{t_1}(k,m)s_{t_2}^*(k,n)\right\}E\left\{n_j^*(k,p)n_j(k,q)\right\}\\ &= \sigma_n^2\sum_{t=1}^{M_t}\sum_{p=0}^{N_c-1}\frac{\phi_{m,n}}{\phi_{m,p}\phi_{n,p}}H_{jt}[m]H_{jt}^*[n]\varrho_{i_1,i_2}(p,q)\end{aligned} \tag{B.26}$$

Calculation of $\mathbf{\Upsilon}$

According to the definition, $\mathbf{\Upsilon}$ can be expressed as

$$\begin{aligned}\mathbf{\Psi}_{i_1,i_2} &\triangleq E\left\{\left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\right\}\\ &= \sum_{t_1,t_2=1}^{M_t} E\left\{\left[\left(\mathbf{n}_j(k)\mathbf{s}_{t_1}^H(k)\mathcal{D}^H(\mathbf{H}_{jt_1})\right)\oslash\mathbf{\Phi}\right]\right.\\ &\quad\left.\cdot\,\boldsymbol{\mu}_{i_1}\boldsymbol{\mu}_{i_2}^H\left[\left(\mathcal{D}(\mathbf{H}_{jt_2})\mathbf{s}_{t_2}(k)\mathbf{n}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\right\}\end{aligned} \tag{B.27}$$

Thus, the $(m,n)^{th}$ element of Υ_{i_1,i_2}, $\upsilon_{i_1,i_2}(m,n)$ can be written as

$$\begin{aligned}\upsilon_{i_1,i_2}(m,n) &= \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\varrho_{i_1,i_2}(p,q)}{\phi_{m,p}\phi_{q,n}}H^*_{jt_1}[p]H_{jt_2}[q] \\ &\quad \cdot E\left\{s_{t_1}(k,p)s^*_{t_2}(k,q)\right\}E\left\{n^*_j(k,m)n_j(k,n)\right\} \\ &= \delta(m-n)\sigma_n^2\sum_{t=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{p,q}}{\phi_{m,p}\phi_{q,m}}H^*_{jt}[p]H_{jt}[q]\varrho_{i_1,i_2}(p,q) \quad \text{(B.28)}\end{aligned}$$

which indicates that $\boldsymbol{\Upsilon}_{i_1,i_2}$ is a diagonal matrix.

Calculation of $\boldsymbol{\Omega}$

According to the definition, the $(m,n)^{th}$ element of Ω_{i_1,i_2}, $\omega_{i_1,i_2}(m,n)$ can be written as

$$\omega_i(m,n) = \begin{cases}\sigma_n^4\sum_{p=0,p\neq m}^{N_c-1}\varrho_{i_1,i_2}(p,q)/(\phi_{m,p}\phi_{p,m}) & (m=n) \\ \sigma_n^4\varrho_{i_1,i_2}(p,q) & (m\neq n)\end{cases} \quad \text{(B.29)}$$

Thus, according to all above derivations, the covariance matrix of the channel estimation error can be expressed as

$$E\left\{\Delta\vec{\mathbf{H}}_j\Delta\vec{\mathbf{H}}_j^H\right\} = \frac{1}{K}[\mathbf{I}_{M_t}\otimes(\mathbf{U}_n\mathbf{U}_n^H)]\mathcal{E}[\mathbf{I}_{M_t}\otimes(\mathbf{U}_n\mathbf{U}_n^H)] \quad \text{(B.30)}$$

where $\Delta\vec{\mathbf{H}}_j$ denotes the column vector constructed by concatenating the columns of the matrix $\Delta\mathbf{H}_j$.

$$\mathcal{E} \triangleq \begin{bmatrix}\mathcal{E}_{1,1} & \cdots & \mathcal{E}_{1,M_t} \\ \vdots & \ddots & \vdots \\ \mathcal{E}_{M_t,1} & \cdots & \mathcal{E}_{M_t,M_t}\end{bmatrix} \quad \text{(B.31)}$$

and the $(m,n)^{th}$ element of $\mathcal{E}_{i_1,i_2}$ is defined as

$$e_{i_1,i_2}(m,n) \triangleq \ddot{\xi}_{i_1,i_2}(m,n) + \phi_{i_1,i_2}(m,n) + \upsilon_{i_1,i_2}(m,n) + \omega_{i_1,i_2}(m,n) \quad \text{(B.32)}$$

where $\ddot{\xi}_{i_1,i_2}(m,n)$, $\psi_{i_1,i_2}(m,n)$, $\upsilon_{i_1,i_2}(m,n)$ and $\omega_{i_1,i_2}(m,n)$ are defined by Eqn.(B.23), Eqn.(B.26), Eqn.(B.28), and Eqn.(B.29) respectively.

Consequently, the MSE of the channel estimator is

$$\begin{aligned} \text{MSE} &= E\left\{\text{tr}\left(\Delta\vec{\mathbf{H}}_j\Delta\vec{\mathbf{H}}_j^H\right)\right\} = \sum_{i=1}^{M_t} E\left\{\text{tr}\left(\Delta\mathbf{H}_{ji}\Delta\mathbf{H}_{ji}^H\right)\right\} \\ &= \frac{1}{K}\sum_{i=1}^{M_t}\text{tr}\left(\mathbf{U}_n^H\mathcal{E}_{i,i}\mathbf{U}_n\right) \end{aligned} \tag{B.17}$$

The proof ends here. ■

Appendix C

The Proof of Theorem 4.4.3

Theorem 4.4.3 The proposed blind channel estimator of the normal channel matrices, $\hat{\mathbf{H}}_t$ $(t = 1, \cdots, M_t,\ t \neq j)$, is asymptotically unbiased (i.e. $E\{\Delta\mathbf{H}_t\} = \mathbf{0}$), and the estimated channel MSE is

$$\text{MSE} = \frac{1}{K}\sum_{i=1}^{M_t} \text{tr}\left\{\ddot{\mathbf{\Xi}}_i + \mathbf{\Psi}_i + \mathbf{\Upsilon}_i + \mathbf{\Omega}_i\right\} \tag{C.1}$$

Proof: According to Eqn.(4.15), the channel matrix associated with the tth received antenna is estimated based on the reference channel matrix $\hat{\mathbf{H}}_j$ as follows

$$\hat{\mathbf{H}}_t = \hat{\mathbb{R}}_{\mathbf{y}_{t,j}}(\hat{\mathbf{H}}_j^H)^\dagger \tag{C.2}$$

Obliviously, $E\{\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}\} = \mathbf{H}_t\mathbf{H}_j^H$ and according to Theorem 4.4.2, the estimator $\hat{\mathbf{H}}_j$ is unbias, i.e., $E\{\hat{\mathbf{H}}_j\} = \mathbf{H}_j$, then

$$\begin{aligned} E\{\hat{\mathbf{H}}_t\} &= E\{\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}\}E\left\{(\hat{\mathbf{H}}_j^H)^\dagger\right\} \\ &= \mathbf{H}_t\mathbf{H}_j^H(\mathbf{H}_j^H)^\dagger = \mathbf{H}_t \end{aligned} \tag{C.3}$$

which indicates that the estimator is unbias. Next, we consider the covariance matrix of the estimated channel matrix $\hat{\mathbf{H}}_t$.

$$\begin{aligned}\text{Cov} &\triangleq E\{\Delta\mathbf{H}_t\Delta\mathbf{H}_t^H\} = E\{\hat{\mathbf{H}}_t\hat{\mathbf{H}}_t^H\} - \mathbf{H}_t\mathbf{H}_t^H \\ &= E\left\{\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger}\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}^H\right\} - \mathbf{H}_t\mathbf{H}_t^H \end{aligned} \tag{C.4}$$

where similar to Eqn.(4.43) and Eqn.(4.44), $\Delta\mathbf{H}_t$ is the estimation error defined as

$$\Delta\mathbf{H}_t \triangleq \hat{\mathbf{H}}_t - \tilde{\mathbf{H}}_t = \hat{\mathbf{H}}_t - \mathbf{H}_t\mathbf{Q}^{-1} \tag{C.5}$$

where $\mathbf{Q}$ is the ambiguity matrix, which is assumed to be known.

According to Eqn.(4.16b), we have

$$\begin{aligned} & E\left\{\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger}\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}^H\right\} \\ = & \frac{1}{K^2}\sum_{k_1,k_2=1}^{K} E\left\{\left[\left(\mathbf{y}_t(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\mathbf{\Phi}\right]\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger}\left[\left(\mathbf{y}_j(k_2)\mathbf{y}_t^H(k_2)\right)\oslash\mathbf{\Phi}\right]\right\} \\ = & \frac{1}{K^2}\left(\sum_{k_1\neq k_2} E\left\{\left(\mathbf{y}_t(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\mathbf{\Phi}\right\} E\left\{\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger}\right\} E\left\{\left(\mathbf{y}_j(k_2)\mathbf{y}_t^H(k_2)\right)\oslash\mathbf{\Phi}\right\}\right. \\ & \left. +\sum_{k=1}^{K} E\left\{\left[\left(\mathbf{y}_t(k)\mathbf{y}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger}\left[\left(\mathbf{y}_j(k)\mathbf{y}_t^H(k)\right)\oslash\mathbf{\Phi}\right]\right\}\right) \end{aligned} \tag{C.6}$$

where $\mathbf{y}_j(k)$ and $\mathbf{y}_t(k)$ are the kth OFDM symbol received by the jth and t^{th} receive antenna respectively. On the other hand, according the Theorem 4.4.2 again, the perturbed channel matrix $\hat{\mathbf{H}}_j$ can be written as

$$\hat{\mathbf{H}}_j \doteq \mathbf{H}_j + \mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{\frac{1}{2}} \tag{C.7}$$

where $\mathbf{U}_s$ and $\mathbf{U}_n$ are the signal and noise subspace of the matrix $\mathbb{R}_{\mathbf{y}_j} \triangleq \mathbf{H}_j\mathbf{H}_j^H$ respectively, and the diagonal matrix $\mathbf{\Lambda}_s$ contains the singular-values of $\mathbb{R}_{\mathbf{y}_j}$ on its diagonal. Thus, it follows that

$$\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^{\dagger} = \left(\mathbf{U}_s + \mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\right)\mathbf{\Lambda}_s^{-1}\left(\mathbf{U}_s + \mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\right)^H \tag{C.8}$$

and consequently

$$
\begin{aligned}
& \mathbf{H}_j^H E\left\{\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^\dagger\right\}\mathbf{H}_j \\
= \; & \mathbf{H}_j^H E\left\{\left(\mathbf{U}_s+\mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\right)\mathbf{\Lambda}_s^{-1}\left(\mathbf{U}_s+\mathbf{U}_n\mathbf{U}_n^H\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\right)^H\right\}\mathbf{H}_j \\
= \; & \mathbf{H}_j^H\mathbf{U}_n\mathbf{U}_n^H E\left\{\Delta\mathbb{R}_{\mathbf{y}_j}\right\}\mathbf{U}_s\mathbf{\Lambda}_s^{-2}\mathbf{U}_s^H\mathbf{H}_j+\mathbf{H}_j^H\mathbf{U}_s\mathbf{\Lambda}_s^{-2}\mathbf{U}_s^H E\left\{\Delta\mathbb{R}_{\mathbf{y}_j}^H\right\}\mathbf{U}_n\mathbf{U}_n^H\mathbf{H}_j \\
& \mathbf{H}_j^H\mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H\mathbf{H}_j+\mathbf{H}_j^H\mathbf{U}_n\mathbf{U}_n^H E\left\{\Delta\mathbb{R}_{\mathbf{y}_j}\mathbf{U}_s\mathbf{\Lambda}_s^{-3}\Delta\mathbb{R}_{\mathbf{y}_j}^H\mathbf{U}_n\mathbf{U}_n^H\mathbf{H}_j\right\} \\
= \; & \mathbf{I}_{M_t}
\end{aligned}
\tag{C.9}
$$

Note that

$$
E\left\{\left(\mathbf{y}_t(k)\mathbf{y}_j^H(k)\right)\oslash\mathbf{\Phi}\right\}=\mathbf{H}_t\mathbf{H}_j^H \tag{C.10}
$$

Hence, the first term of the right hand side of Eqn.(C.6) can be derived as

$$
\begin{aligned}
& \frac{1}{K^2}\sum_{k_1\neq k_2}E\left\{\left(\mathbf{y}_t(k_1)\mathbf{y}_j^H(k_1)\right)\oslash\mathbf{\Phi}\right\}E\left\{\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^\dagger\right\}E\left\{\left(\mathbf{y}_j(k_2)\mathbf{y}_t^H(k_2)\right)\oslash\mathbf{\Phi}\right\} \\
= \; & \frac{K-1}{K}\mathbf{H}_t\mathbf{H}_t^H
\end{aligned}
\tag{C.11}
$$

Meanwhile, according to Eqn.(C.8), $\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^\dagger$ can be written as

$$
\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^\dagger=\sum_{i=1}^{M_t}\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \tag{C.12}
$$

where $\hat{\boldsymbol{\mu}}_i$ is defined as

$$
\hat{\boldsymbol{\mu}}_i\triangleq\lambda_{s,i}^{-\frac{1}{2}}(\mathbf{u}_{s,i}+\Delta\mathbf{u}_{s,i}) \tag{C.13}
$$

Substitute Eqn.(C.11) and Eqn.(C.12) into Eqn.(C.6), we get

$$
\begin{aligned}
& E\left\{\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}\left(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H\right)^\dagger\hat{\mathbb{R}}_{\mathbf{y}_{t,j}}^H\right\} \\
= \; & \frac{1}{K}\sum_{i=1}^{M_t}E\left\{\left[\left(\mathbf{y}_t(k)\mathbf{y}_j^H(k)\right)\oslash\mathbf{\Phi}\right]\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H\left[\left(\mathbf{y}_j(k)\mathbf{y}_t^H(k)\right)\oslash\mathbf{\Phi}\right]\right\} \\
& +\frac{K-1}{K}\mathbf{H}_t\mathbf{H}_t^H
\end{aligned}
\tag{C.14}
$$

Consider

$$E\left\{ \left[\left(\mathbf{y}_t(k)\mathbf{y}_j^H(k)\right) \oslash \mathbf{\Phi}\right] \hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \left[\left(\mathbf{y}_j(k)\mathbf{y}_t^H(k)\right) \oslash \mathbf{\Phi}\right] \right\} = \mathbf{\Xi}_i + \mathbf{\Psi}_i + \mathbf{\Upsilon}_i + \mathbf{\Omega}_i \quad \text{(C.15)}$$

where we define

$$\mathbf{\Xi}_i = E\left\{\left[\left(\tilde{\mathbf{y}}_t(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \mathbf{\Phi}\right] \hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \left[\left(\tilde{\mathbf{y}}_j(k)\tilde{\mathbf{y}}_t^H(k)\right) \oslash \mathbf{\Phi}\right]\right\} \quad \text{(C.16a)}$$

$$\mathbf{\Psi}_i = E\left\{\left[\left(\tilde{\mathbf{y}}_t(k)\mathbf{n}_j^H(k)\right) \oslash \mathbf{\Phi}\right] \hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \left[\left(\mathbf{n}_j(k)\tilde{\mathbf{y}}_t^H(k)\right) \oslash \mathbf{\Phi}\right]\right\} \quad \text{(C.16b)}$$

$$\mathbf{\Upsilon}_i = E\left\{\left[\left(\mathbf{n}_t(k)\tilde{\mathbf{y}}_j^H(k)\right) \oslash \mathbf{\Phi}\right] \hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \left[\left(\tilde{\mathbf{y}}_j(k)\mathbf{n}_t^H(k)\right) \oslash \mathbf{\Phi}\right]\right\} \quad \text{(C.16c)}$$

$$\mathbf{\Omega}_i = E\left\{\left[\left(\mathbf{n}_t(k)\mathbf{n}_j^H(k)\right) \oslash \mathbf{\Phi}\right] \hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H \left[\left(\mathbf{n}_j(k)\mathbf{n}_t^H(k)\right) \oslash \mathbf{\Phi}\right]\right\} \quad \text{(C.16d)}$$

where $\tilde{\mathbf{y}}_j(k)$ and $\tilde{\mathbf{y}}_t(k)$ are defined as the kth noise free OFDM symbol received by the jth and t^{th} receive antenna respectively.

Then we consider the $(m,n)^{th}$ elements of $\mathbf{\Xi}_i$, $\mathbf{\Psi}_i$, $\mathbf{\Upsilon}_i$, and $\mathbf{\Omega}_i$, which are denoted as $\xi_i(m,n)$, $\psi_i(m,n)$, $\upsilon_i(m,n)$ and $\omega_i(m,n)$ respectively. The derivation of these parameters is very similar to those appear in Theorem 4.4.2, we hereby only give out the final results without unnecessary repeating of the derivation.

$$\begin{aligned}\xi_i(m,n) = & \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{m,n}\phi_{q,p}}{\phi_{m,p}\phi_{q,n}}H_{tt_1}[m]H_{jt_2}^*[p]H_{jt_2}[q]H_{tt_1}^*[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^*[q]\} \\ & + \sum_{t_1,t_3=1}^{M_t}\sum_{p,q=0}^{N_c-1}H_{tt_1}[m]H_{jt_1}^*[p]H_{jt_3}[q]H_{tt_3}^*[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^*[q]\} \\ & - \sum_{t_1=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\beta(m,n,p,q)}{\phi_{m,p}\phi_{q,n}}H_{tt_1}[m]H_{jt_1}^*[p]H_{jt_1}[q]H_{tt_1}^*[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^*[q]\} \quad \text{(C.17a)}\end{aligned}$$

$$\psi_i(m,n) = \sigma_n^2\sum_{t_1=1}^{M_t}\sum_{p=0}^{N_c-1}\frac{\phi_{m,n}}{\phi_{m,p}\phi_{n,p}}H_{tt_1}[m]H_{tt_1}^*[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^*[p]\} \quad \text{(C.17b)}$$

$$\upsilon_i(m,n) = \delta(m-n)\sigma_n^2\sum_{t_1=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{p,q}}{\phi_{m,p}\phi_{q,m}}H_{jt_1}^*[p]H_{jt_1}[q]E\{\hat{\mu}_i[p]\hat{\mu}_i^*[q]\} \quad \text{(C.17c)}$$

$$\omega_i(m,n) = \delta(m-n)\sigma_n^4\sum_{p=0}^{N_c-1}\frac{1}{\phi_{m,p}\phi_{p,m}}E\{\hat{\mu}_i[p]\hat{\mu}_i^*[p]\} \quad \text{(C.17d)}$$

where

$$\beta(m,n,p,q) \triangleq \sum_{i=0}^{N_c-1} \mathrm{p}_{m,i}\mathrm{p}_{p,i}^{*}\mathrm{p}_{q,i}\mathrm{p}_{n,i}^{*} \tag{C.17e}$$

and $\mathrm{p}_{m,n}$ is the $(m,n)^{th}$ element of the precoding matrix $\mathbf{P}$. Note that Eqn.(C.17a) suggests that

$$\mathbf{\Xi}_i = \dot{\mathbf{\Xi}}_i + \ddot{\mathbf{\Xi}}_i = \mathbf{H}_{ti}\mathbf{H}_{ti}^H + \ddot{\mathbf{\Xi}}_i \tag{C.18}$$

where $\dot{\mathbf{\Xi}}_i = \mathbf{H}_{ti}\mathbf{H}_{ti}^H$, and the $(m,n)^{th}$ element of the matrix $\ddot{\mathbf{\Xi}}_i$ is

$$\begin{aligned}\ddot{\xi}_i(m,n) &= \sum_{t_1,t_2=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\phi_{m,n}\phi_{q,p}}{\phi_{m,p}\phi_{q,n}}H_{tt_1}[m]H_{jt_2}^{*}[p]H_{jt_2}[q]H_{tt1}^{*}[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^{*}[q]\} \\ &\quad -\sum_{t_1=1}^{M_t}\sum_{p,q=0}^{N_c-1}\frac{\beta(m,n,p,q)}{\phi_{m,p}\phi_{q,n}}H_{tt_1}[m]H_{tt_1}^{*}[p]H_{jt_1}[q]H_{jt_1}^{*}[n]E\{\hat{\mu}_i[p]\hat{\mu}_i^{*}[q]\}\end{aligned} \tag{C.19}$$

In the above equations, $E\{\hat{\mu}_i[p]\hat{\mu}_i^{*}[q]\}$ is the $(p,q)^{th}$ element of the matrix $E\left\{\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H\right\}$, where

$$\begin{aligned}E\left\{\hat{\boldsymbol{\mu}}_i\hat{\boldsymbol{\mu}}_i^H\right\} &= \boldsymbol{\mu}_i\boldsymbol{\mu}_i^H + E\left\{\Delta\boldsymbol{\mu}_i\Delta\boldsymbol{\mu}_i^H\right\} \\ &= \lambda_{s,i}^{-1}\mathbf{u}_{s,i}\mathbf{u}_{s,i}^H + \frac{1}{K'}\lambda_{s,i}^{-2}\mathbf{U}_n\mathbf{U}_n^H\mathcal{E}_{i,i}\mathbf{U}_n\mathbf{U}_n^H\end{aligned} \tag{C.20}$$

where $\mathbf{u}_{s,i}$ is the ith column of left singular matrix of $\mathbb{R}_{\mathbf{y}_j}$, $\lambda_{s,i}$ is the corresponding singular-value (see Eqn.(B.8)), $\mathcal{E}_{i,i}$ is defined by Eqn.(B.32), and K' is the NOS for estimating the "reference" channel matrix $\mathbf{H}_j$.

Thus, by substituting Eqn.(C.16a) - Eqn.(C.17d) into Eqn.(C.4), we can derive the covariance matrix as

$$\begin{aligned}\mathrm{Cor} &= \frac{K-1}{K}\mathbf{H}_t\mathbf{H}_t^H + \frac{1}{K}\sum_{i=1}^{M_t}\left(\ddot{\mathbf{\Xi}}_i + \mathbf{\Psi}_i + \mathbf{\Upsilon}_i + \mathbf{\Omega}_i\right) + \frac{1}{K}\mathbf{H}_t\mathbf{H}_t^H - \mathbf{H}_t\mathbf{H}_t^H \\ &= \frac{1}{K}\sum_{i=1}^{M_t}\left(\ddot{\mathbf{\Xi}}_i + \mathbf{\Psi}_i + \mathbf{\Upsilon}_i + \mathbf{\Omega}_i\right)\end{aligned} \tag{C.21}$$

which directly leads to the channel estimation MSE

$$\mathrm{MSE} = \frac{1}{K}\sum_{i=1}^{M_t} \mathrm{tr}\left\{\ddot{\boldsymbol{\Xi}}_i + \boldsymbol{\Psi}_i + \boldsymbol{\Upsilon}_i + \boldsymbol{\Omega}_i\right\} \tag{C.22}$$

The proof ends here. ■

Bibliography

[1] D. M. Balston, and R. C. V. Macario, *Cellular Radio Systems*, Artech House, 1993.

[2] T. S. Rappaport, *Wireless Communications*, 2nd Edition, Prentice Hall, 2002.

[3] J. G. Proakis, *Digital Communications*, 4th Edition, McGraw-Hill, 2001.

[4] Z. Ding, Y. Li, *Blind Equalization and Identification*, 2nd Edition, Marcel Dekker, 2001.

[5] A. Sikora, *Wireless personal and local area networks*, Wiley, 2003.

[6] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Boston: Artech House, 2000.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third Edition, The Johhs Hopkins University Press, 1996.

[8] M. Rahnema, "Overview of the GSM system and protocol architecture," *IEEE Commun. Mag.*, vol. 31, no. 4, pp. 92-100, Apr. 1993.

[9] D. Terasawa, and E. G. Jr. Tiedemann, "CDMAOne(R) (IS-95) technology overview and evolution," in *Proc. Radio Frequency Integrated Circuits (RFIC) Symposium*, Anaheim, CA, USA, 13-15 Jun. 1999, pp.213-216.

[10] J. Chuang, and N. Sollenberger, "Beyond 3G: Wideband wireless data access based on OFDM and dynamic packet assignment," *IEEE Commun. Mag.*, vol.38, no. 7, pp. 78-87, Jul. 2000.

[11] "Third Generation Mobile Phone Licensing in Europe," *TIA on line*, Online: http://www.tiaonline.org/international/regional/nis/licensing.cfm Oct. 2000.

[12] "Auction of Third Generation Mobile Telecommunications Licences in the UK, Frequently Asked Questions," *Radiocommunications Agency of UK*, Online: http://www.spectrumauctions.gov.uk/documents/faq2.htm.

[13] "Spectrum Auctions," *Radiocommunications Agency of UK*, Online: http://www.spectrumauctions.gov.uk/auction/auction_index.htm.

[14] L. M. Correia and R. Prasad, "An overview of wireless broadband communications," *IEEE Communications Magazine*, vol. 35, no. 1, pp. 2833, Jan. 1997.

[15] M. Dinis and J. Fernandes, "Provision of sufficient transmission capacity for broadband mobile multimedia: A step toward 4G," *IEEE Communications Magazine*, vol. 39, no. 8, pp. 4654, Aug. 2001.

[16] M. Progler, C. Evci, and M. Umehira, "Air interface access schemes for broadband mobile systems," *IEEE Communications Magazine*, vol. 37, no. 9, pp. 106115, Sept. 1999.

[17] H. Rohling, R. Grünheid, and D. Galda, "OFDM air interface for the 4th generation of mobile communication systems," in *Proc. 6th international OFDM-Workshop (InOWo01)*, Hamburg, Sept. 2001, pp. 0-10-28.

[18] L. M. Pereira, "Fourth generation: now , it is personal!," in *Proc. 11th International Symposium on Personal Indoor Mobile Radio Communications PRMRC 2000*, London, Sep. 2000, pp. 1009-1016.

[19] S. Nanda, K. Balachandran, and S. Kumar, "Adaptation Techniques in Wireless Packet Data Services," *IEEE Communications Magazine*, pp. 54-64, Jan. 2000.

[20] R. van Nee, G. Awater, M. Morikura, H. Takeshi, M. Webster, and K. W. Halford, "New high-rate wireless LAN standards," *IEEE Communications Magazine*, vol. 37, no. 12, pp. 82-88, Dec. 1999.

[21] L. M. Pereira, "Fourth generation: now, it is personal!," in *Proc. PIMRC 2000 (11th International Symposium on Personal Indoor Mobile Radio Communications)*, London, Sept. 2000, pp. 1009-1016.

[22] J. Kalliokulju, P. Meche, M. J. Rinne, J. Vallstrom, P. Varshney, and S. -G. Haggman, "Radio access selection for multistandard terminals," *IEEE Communications Magazine*, vol. 39, no. 10, pp. 116124, Oct. 2001.

[23] A. Jamalipour, and S. Tekinay, "Fourth generation wireless networks and interconnecting standards," *IEEE Personal Commun. Mag.*, vol.8, No. 6, Dec. 2001.

[24] R. W. Chang, "Synthesis of band-limited orghogonal signals for mulltichannel data transmission" *Bell System Tech. J.*, No.45, pp. 1775-1796, Dec. 1966.

[25] S. B. Weinstain and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, Vol.43, No. 2-4, pp.887-895.Feb.-Apr. 1995.

[26] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational comlexity algorithms," *IEEE. Int. Conf. Acoust., Speech, Signal Processing*, Denver, CO, 1980 pp. 964-967.

[27] H. Nakagawa, "Vision in WP8F and activity in Japan for future mobile communicaiton systems," *Strategic Workshop 2002 Unified Global Infrastructrure,* Prague, Czech Republic, Sep. 6-7, 2002.

[28] G. J. Foschini, "Layered space-time architecture for wireless communicaiton in a fading environment when using multiple antennas," *Bell Laboratories Technical Journal*, vol. 1, no.2, pp.41-59, autumn 1996.

[29] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Comm*, vol. 46, no.3, pp. 357-366, Mar. 1998.

[30] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451 - 1458, Oct. 1998.

[31] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Cal derbank, "A space-time coding modem for high-data-rate wireless communications," *IEEE Journal on Selected Areas in Communicaitons*, vol. 16, no. 8, pp. 1459-1478, Oct. 1998.

[32] V. Tarokh, N. Seshadri, and R. r. Calderbank, "Spactime-codes for high data rate wireless communicaiton: performance criterion and code construction," *IEEE Trans. Info. Theory*, vol. 44, no. 3, pp. 744-756, Mar. 1998.

[33] A. van Zelst, "Space division multiplexing algorithms," in *Proc. of the 10th Mediterranean Electrotechnical Conference (MELECON) 2000*, Lemesos, Cyprus, May 2000, vol.3, pp. 1218-1221.

[34] IEEE 802.11, "IEEE P802.11, the working group for wireless LANs," Aug. 2003, Online: http://grouper.ieee.org/groups/802/11/.

[35] ETSI, "ETSI-Telecom standars," Online: http://www.etsi.org, Nov. 2003.

[36] K.S. Li and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-gaussian linear processes," *The Annals of Statistics*, vol. 10, pp. 1195-1208, 1982.

[37] J. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Proc. Mag.*, vol. 17, no. 3, pp. 17-28, May 2000.

[38] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 4-114, Jan. 1975.

[39] H. Liu, G. Xu, L. Tong, and T. Kailath. "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Processing*, vol. 50, pp. 82-99, 1996.

[40] E. de Carvalho and D.T.M. Slock, "Semi-blind methods for FIR multichannel estimation," in *Signal processing advances in wireless and mobile communications*, vol. 1, chapter 7, pp. 211-254. Prentice Hall, 2001.

[41] Z. Ding, "An outer-product decomposition algorithm for multichannel blind identification," *Proc. 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Corfu, Greece, pp. 132-135, June 1996.

[42] Z. Ding, "Matrix outer-product decomposition method for blind multichannel identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 3054-3061, Dec. 1997,

[43] A. Mansour, and C. Jutten, "Fourth-order criteria for blind sources separation," *IEEE Trans. Signal Processing*, vol. 43, no. 8, pp. 2022-2025, Aug. 1995.

[44] A. K. Nandi, and V. Zarzoso, "Fourth-order cumulant based blind source separation," *IEEE Signal Processing Letters*, vol. 3, no. 12, pp. 312-314, Dec. 1996.

[45] A. R. Leyman, Z. M. Kamran, and K. Abed-Meraim, "Higher-order time frequency-based blind source separation technique," *IEEE Signal Processing Letters*, vol. 7, no. 7, pp.193-196 Jul. 2000.

[46] L. Molgedey, and H.G. Schuster, "Separation of a mixture of independent signals using time delayed correlations," *Physical Review Letters*, vol. 72 no. 23, pp. 3634-637, 1994.

[47] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics", *IEEE Trans. Signal Processing*, vol. 45 no. 2, pp. 434-444, Feb. 1997.

[48] A. Ziehe, K.-R.Muller, G. Nolte, B.-M. Mackert, and G. Curio, "Artifact reduction in biomagnetic recordings based on time-delayed second order correlations," *IEEE Trans. Biomedical Engineering*, vol. 47 pp. 75-87, 2000.

[49] S. Choi, A. Cichocki, and A. Belouchrani, "Second order nonstationary source separation," *Journal of VLSI Signal Processing*, vol. 32 no. 1-2, pp. 93-104, Aug. 2002.

[50] A. Cichocki and A. Belouchrani, "Sources separation of temporally correlated sources from noisy data using bank of band-pass filters," in *Proc. Third International Conference on Independent Component Analysis and Signal Separation (ICA-2001)*, pp. 173-178, San Diego, USA, Dec. 9-13 2001.

[51] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Neural Networks*, vol. 8, no. 3, pp. 411-419, 1995.

[52] S. Choi, A. Cichocki, and S. Amari, "Equivariant nonstationary source separation," *Neural Networks*, vol. 15 pp. 121-130, 2002.

[53] S. Choi, A. Cichocki, and A. Belouchrani, "Second order nonstationary source separation," *Journal of VLSI Signal Processing*, vol. 32 vol. 1-2, pp. 93-104, Aug. 2002.

[54] A. Belouchrani, and M.G. Amin, "A new approach for blind source separation using time-frequency distributions," in *Proc. SPIE*, 2846: pp. 193-203, 1996.

[55] L. Tong, R. W. Liu, V. C. Soon, and Y. F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits and Systems*, vol. 38, no. 5, pp. 499-509, May 1991.

[56] J. A. C. Bingham, "Multi carrier modulation for data transmission: an idea whose time has come", *IEEE Commun. Mag.*, vol. 28, no.5, pp. 5-14, 1990.

[57] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *AT&T-Bell Labs Internal Tech. Memo*, Sept. 1995. Published in *Wireless Personal Communications*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[58] ETSI, "Broadband radio access networks (BRAN), HIPERLAN Type 2; Physical (PHY) layer," *Tech. Rep. ETSI TS 101 475*, v1.1.1, April 2000.

[59] IEEE, "Supplement to standard for telecommunications and information exchange between systems - LAN/MAN specific requirements - part 11: wireless MAC and PHY specificaitons: high speed physical layer in the 5 GHz band," *ISO/IEC 8802-11: 1999, 1999/Amd 1:2000(E)*, 2000.

[60] DVB, "The standard for the digital world," Online: http://www.dvb.org, Aug. 2003.

[61] DAB, "The world DAB fourm," Online: http://www.worlddab.org, Aug. 2003.

[62] Z. Wang, G.B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon", *IEEE Signal Process. Mag.*, Vol. 17, No. 3 2000, pp. 29-48.

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

[64] G. L. Stuber, S. W. Mclaughlin, and M. A. Ingram, "Broadband MIMO-OFDM wireless communicaitons", *Proc. of the IEEE*, vol. 92, no. 2, pp. 271-294, 2004.

[65] W. Bai, C. He, L. Jiang, and H. Zhu, "Blind channel estimation in MIMO-OFDM systems," *Proc. GLOBECOM*, Taipei, Taiwan, Nov. 2002, vol.1, pp. 317-321.

[66] W. Bai, and Z. Bu, "Subspace based channel estimation in MIMO-OFDM system," *Proc. VTC* , May 2004, vol. 2, pp. 598 - 602.

[67] C. Shin, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems using virtual carriers," *Proc. GLOBECOM*, 29 Nov.-3 Dec. 2004 , vol. 4, pp. 2465 - 2469.

[68] S. Sun, I. Wiemer, C. K. Ho, and T. T. Tjhung, "Training sequence assisted channel estimation for MIMO OFDM", *IEEE WCNC 2003*, Vol. 1, pp. 38 - 43.

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

[70] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MMIMO-OFDM systems," *IEEE Trans. Vehicular Tech.*, vol. 56, no. 2, pp. 670-685, Mar. 2007.

[71] K. Rahbar, J. P. Reilly, and J. H. Manton, "Blind identification of MIMO FIR systems driven by quasistationary sources using second-order statistics: a frequency domain approach," *IEEE Trans. Signal Proc.*, vol. 52, no. 2, Feb. 2004, pp. 406-417.

[72] B. Muquet, M. D. Courville, and P. Duhamel, "Subpsec-based blidn and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, pp. 1699-1712, Jul. 2002.

[73] S. Roy and C. Y. Li, "A subspace blind channel estimation method for OFDM systems withoutcyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 572-579, Jul. 2002.

[74] Y. Zeng, and T. S. Ng, "A semi-blind chenel estimation method for multiuser multiantenna OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1419-1429, May 2004.

[75] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences-Applicability to blind equalization," *IEEE Trans. Signal Proc.*, vol. 43, no. 12, pp. 3501-3506, Dec. 2000.

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

[77] J. Xavier, V. Barroso, and J. Mount, "Closed-form correlative coding blidn identification of MIMO channels: Isometry fitting to second order statistics," *IEEE Trans. Signal Proc.*, vol. 49, no. 5, pp. 1073-1086, May 2001.

[78] A. Gorokhov, and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits & Systems-I: Fundamental Theory & Applications*, vol. 44, no. 9, pp. 813-820, Sep. 1997

[79] W. Bai, C. He, L. Jiang, and H. Zhu, "Blind channel estimation in MIMO-OFDM systems," *Proc. GLOBECOM*, Taipei, Taiwan, vol.1, pp. 317-321, Nov. 2002.

[80] C. Shin, R. W. Heath, Jr., and E. J. Powers, "Blind channel estimation in MIMO-OFDM systems," *IEEE Trans. Vehicular Tech.*, vol. 56, no. 2, pp. 670-685, Mar. 2007.

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

[82] W. Bai, and Z. Bu, "Subspace based channel estimation in MIMO-OFDM system," *Proc. VTC* , vol. 2, pp. 598 - 602, Milan Italy, May 2004.

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

[84] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communicatin systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867-1875, Nov. 1980.

[85] M. Ghosh, "Blind decision feedback equalzation for terrestrial television receivers," *Proc. IEEE*, vol. 86, no. 10, pp. 2070-2081, Oct. 1998.

[86] S. Talwar, M. Biberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array - part I: Algorithms," *IEEE Trans. Signal Processing*, vol. 44, May 1996, pp. 1184 - 1197.

[87] K. Anand, G. Mathew, and V. U. Reddy, "Blind separation of multiple co-channel BPSK signals arriving at an antenna array," *IEEE Signal Processing Lett.,* vol. 2, Sep. 1995, pp. 176-177.

[88] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA, estimation algorthms: Futher unification, simplification, and observation," *IEEE Trans. Aerosp., Electron. Syst.,* vol. ASE-29, no. 3, Oct. 1993, pp. 1170-1184.

[89] Y. Inouye and K. Hirano, "Cumulant-based blind identification of linear multi-input-multi-output system driven by colored inputs," *IEEE Trans. Signal Proc.*, vol. 45, pp. 1543-1552, Jun. 1997.

[90] R. W. Heath Jr., and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems", *IEEE Trans. Signal Proc.*, vol. 47, no. 3, Mar. 1999, pp. 848-856.

[91] L. Parra, C. Spence, and B. De Vries, "Convolutive blind source separation based on multiple decorrelation," presented at the 8th IEEE workshop Neural Networks Signal Process., Cambridge, MA, Sept. 1998.

[92] Z. Ding, "Blind Wiener filter estimation for multi-channel sytems based on partial information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.,* vol. 5, pp. 3609-3612.

[93] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Proc.*, vol. .47, pp. 572-578, Feb. 1999.

[94] Z. Ding, "Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics," *IEEE Signal Proc. Lett.*, vol. 3, May 1996, pp. 150-152.

[95] A. Gorokhov and P. Loubaton,"Subspace-based technique for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits Syst. I*, vol. 44. pp. 813-820, Sep. 1997.

[96] S. Van Gerven and D. Van Compernolle, "Signal separation by symmetric adaptive decorrelation: Stability, convergence, and uniqueness," *IEEE Trans. Signal Proc.*, vol. 43, pp. 1602-1612, Jul. 1995.

[97] S. Shamsunder and G. B. Giannakis, "Multichannel blind signal separation and reconstruction," *IEEE Trans. Speech Audio Proc.*, vol. 5, pp. 515-527, Nov. 1997.

[98] D. Yellin and E. Weinstein, "Multichannel signal separation: Methods and analysis," *IEEE Trans. Signal Proc.*, vol. 44, pp. 106-118, Jan. 1996.

[99] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation-induced cyclostationarity," *IEEE Trans. Signal Processing*, vol. 46, no. 7, pp. 1930-1944, Jul. 1998.

[100] G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Lett.*, vol. 4, pp. 184-187, Jun. 1997.

[101] M. K. Tsatsanis, and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing*, vol. 45, pp. 1785-1794, Jul. 1997.

[102] R. W. Heath, Jr., "Mitigating channel distortions in wireless orthogonal frequency division multiplexing communicaiton systems," MS Thesis, Dept. Elect. Eng., Univ. Virginia, Charlottesville, Aug. 1997.

[103] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Trans. Info. Theory*, vol. 40, no. 2, pp. 340-349, Mar. 1994.

[104] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: a frequency-domain approach," *IEEE Trans. Info. Theory*, vol. 41, no. 1, pp. 329-334, Jan. 1995.

[105] H. Bolcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna systems," *IEEE Trans. Signal Proc.* vol. 50, no. 1, pp. 96-109, Jan. 2002.

[106] I. Bradaric, A. P. Petropulu, and K. I. Diamantaras, "On blind identifiability of FIR-MIMO systems with cyclostationary inputs using second order statistics," *IEEE Trans. Signal Proc.*, vol. 51, no. 2, pp. 434-441, Feb. 2003.

[107] I. Bradaric, A. P. Petropulu, and K. I. Diamantaras, "Blind MIMO FIR channel identification based on second-order spectra correlations," *IEEE Trans. Signal Proc.* vol. 51, no. 6, pp. 1668-1674, Jun. 2003.

[108] Z. Wang and G. B. Giannakis, "Linearly precoded or coded OFDM against wireless fades?" in *Proc. SPAWC'01*, Taoyuan, Taiwan, March 2001.

[109] Y. Zeng, W. H. Lam, and T. S. Ng, "Semiblind channel estimation and equalization for MIMO space-time coded OFDM," *IEEE Trans. Circuits and Systems I: Regular Papers*, vol. 53, no. 2, pp. 463-474, Feb. 2006.

[110] H. Ali, A. Doucet, and Y. Hua, "Blind SOS Subspace channel estimation and equalizatoin techniques exploiting spatial diversity in OFDM systems", *Digital Signal Processing*, Vol. 14, No. 2, Mar. 2004.

[111] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Trainsmit antennae space-time block coding for generalized OFDM in the presence of unkonw multipath," *IEEE J. Select. Areas Commun.*, vol. 19, no. 7, pp. 1352-1364, 2001.

[112] G. B. Giannakis, "Filterbanks for blind channel identificaiton and equalization," *IEEE Signal Process. Lett.*, vol. 4, Jun. 1997, pp. 184 -187.

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

[114] R. Zhang, "Blind channel estimation for precoded OFDM system," in *Proc. ICASSP 2005*, vol. 3, pp.469 - 472, Philadelphia, PA, USA, 18-23 March 2005.

[115] R. Lin, and A. P. Petropulu, "Linear precoding assisted blind channel estimation for OFDM systems," *IEEE Trans. Veh. Tech,* vol. 54, no. 3, pp. 983 - 995, May 2005.

[116] A. P. Petropulu, R. Zhang, and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. Wireless Comm.*, vol. 3, no. 2, pp. 647 -655, Mar. 2004.

[117] H. Bolcskei, R. W. Heath Jr., and A. J. Paulraj, "Blind equalization in OFDM-based multi-antenna systems," *Proc. ASSPCC 2000*, Lake Louise, Alberta, CA, 2000, pp.58-63.

[118] O. Shalvi, and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Info. Theory*, vol. 36, no. 2, Mar. 1990, pp. 312-321.

[119] S. Talwar, M. Biberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array - part I: Algorithms," *IEEE Trans. Signal Processing*, vol. 44, May 1996, pp. 1184-1197.

[120] K. Anand, G. Mathew, and V. U. Reddy, "Blind separation of multiple co-channel BPSK signals arriving at an antenna array," *IEEE Signal Processing Lett.,* vol. 2, Sep. 1995, pp. 176-177.

[121] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithms," *IEEE Trans. Signal Processing,* vol.44, May 1996, pp. 1136-1155.

[122] L. K. Hansen and G. Xu, "A hyperplane-based algorithm for the digital co-channel communications problem," *IEEE Trans. Information Theory*, vol. 43, Sep. 1997, pp. 1536-1548.

[123] Q. Li, E.-W. Bai, and Z. Ding, "Blind source separation of signals with known alphabets using ϵ-approximation algorithms," *IEEE Trans. Signal Processing*, vol. 51, Jan. 2003, pp. 1-10.

[124] G. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Trans. Signal Processing,* vol. 48, Dec. 2000, pp. 3508-3519.

[125] D. Yellin and E. Weinstein, "Criteria for multichannel signal separation," *IEEE Trans. Signal Processing*, vol. 42, Aug. 1994, pp. 2158-2168.

[126] A. Swami, G. Giannakis, and S. Shamsunder, "Multichannel ARMA processes," *IEEE Trans. Signal Processing*, vol. 42, Apr. 1994, pp. 898-913.

[127] T. M. Martinetz, S. G. Berkovich, and K. J. Schulten, ""Neural-gas" network for vector equantization and its application to time-series prediction," *IEEE Trans. Neural Network,* vol. 4, Jul. 1993, pp. 558-569.

[128] P. Comon and E. Moreau, "Improved contrast dedicated to blind separation in communications," *IEEE International Conference on Acoustics, Speach and Signal Processing,* vol. V, (Munich, Germany), 1997, pp. 3453-3456.

[129] R. L. Causey, *On closest normal matrix*, Tech. Rep. CS-10, Dept. Comp. Sc. Stanford 1964.