BLIND CHANNEL IDENTIFICATION/EQUALIZATION WITH APPLICATIONS IN WIRELESS COMMUNICATIONS

JUN FANG

NATIONAL UNIVERSITY OF SINGAPORE

2006

BLIND CHANNEL IDENTIFICATION/EQUALIZATION WITH APPLICATIONS IN WIRELESS COMMUNICATIONS

JUN FANG

(*B.Sc. and M.Sc., Xidian University, China*)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2006

Acknowledgment

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. A. Rahim Leyman, for his guidance, support and his liberal attitude to my research. He demonstrated great freedom and patience on my research. This enabled me to develop my own interests and enjoy the process of intellectual experience during the course of my research. Many thanks go to my co-supervisor Dr. Chew Yong Huat for his helps and discussions on my research.

I also wish to thank all my colleagues and friends who have given me so much helps and encouragements throughout these three years' studies. I would like to acknowledge the Institute for Infocomm Research and National University of Singapore for their generous financial support and facilities. Besides, my work was partially supported by the Singapore's Agency for Science, Technology and Research (A*STAR) under Research Grant Number 022-106-0041.

Finally, heartfelt thanks go to my beloved. I am deeply indebted to my parents and my wife for their constant love and untiring support. It is them who encourage and accompany me during the hardest days of my research.

Summary

The rapid growth in demand for cellular communications services has encouraged research into the design of wireless communications to improve spectrum efficiency and link quality. As opposed to their wireline counterpart, wireless communication systems pose some unique challenges. One of the main problems faced in wireless communications is the intersymbol interference (ISI) caused by channel dispersion and the multiuser interference (MUI) resulting from frequency reuse. In order to recover the desired transmitted user signals accurately, advanced space-time signal processing techniques need to be developed to simultaneously suppress the ISI and MUI. A key aspect of these is the estimation of the channel. Traditional methods for channel estimation usually resort to training sequences to enable channel identification. These periodically transmitted training sequences consume considerable bandwidth and thus reduce the bandwidth usage efficiency. Over the past decade, a promising approach called as "blind method" has received significant attention. As compared to the traditional techniques, blind channel estimation methods identify the unknown wireless channels based only on the received signals and some *a priori* statistical information or properties of the input signals, without direct access to the transmitted signals.

This dissertation focuses on the blind estimation of the wireless channels by

exploiting the statistical information of the received data. We have developed a variety of statistics-based blind channel estimation methods for different data models, i.e., single-input single-output (SISO), single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) models. The proposed algorithms can be directly applied or tailored to diverse wireless communication systems, such as TDMA and CDMA, to combat the ISI and MUI which constitute a major impediment to the system performance. In this dissertation, we, firstly, introduce the background, review, mathematical preliminaries and basic models for blind channel identification. Next, in Chapter 3, we present a higher order statistics-based linear method to estimate the SISO wireless channels. In Chapters 4 and 5, by utilizing the properties of the companion matrices, a new second-order statistics-based method for blind estimation of SIMO and MIMO channels driven by colored sources is proposed. In Chapter 6, we derive a new method to directly estimate the zero-forcing (ZF) or minimum mean-square-error (MMSE) equalizers of the SIMO channel driven by colored sources with unknown statistics. We also studied the problem of blind identification of MIMO channels driven by spatially correlated sources with *a priori* known statistics. The results are presented in Chapter 7. Finally, in Chapter 8, we conclude with a summary of contributions and directions for future research.

Contents

List of Figures

List of Tables

Abbreviations

CCI	Co-Channel Interference
CDMA	Code Division Multiple Access
FDMA	Frequency Division Multiple Access
FIR	Finite Impulse Response
GSM	Global System for Mobile Communications
HOS	Higher Order Statistics
i.i.d.	Independent and identically distributed
ISI	Intersymbol Interference
IS-54	Interim Standard 54
LOS	Line of Sight
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MUI	Multiuser Interference
NMSE	Normalized Mean Square Error

pdf	Probability density function
RF	Radio Frequency
SER	Symbol Error Rate
SIMO	Single Input Multiple Output
SIR	Signal-to-Interference Ratio
SISO	Single Input Single Output
SNR	Signal-to-Noise Ratio
SOS	Second-Order Statistics
TDMA	Time Division Multiple Access
ZF	Zero-Forcing

Notations

$\circledast$	Convolution operator
$(\cdot)^T$	Matrix transpose
$(\cdot)^*$	Complex conjugate
$(\cdot)^H$	Hermitian transpose
$(\cdot)^{-1}$	Generalized inverse
$(\cdot)^\dagger$	Moore-Penrose pseudo-inverse
$E[\cdot]$	Mathematical expectation
$\lVert \cdot \rVert$	Euclidean norm
$\lVert \cdot \rVert_{\mathrm{F}}$	Frobenius norm
$\otimes$	Kronecker product
$\mathrm{rank}(\mathbf{A})$	The rank of matrix $\mathbf{A}$
$\mathcal{R}(\mathbf{A})$	The range (column) space of matrix $\mathbf{A}$
$\mathrm{vec}(\mathbf{A})$	The column vector obtained from matrix $\mathbf{A}$ by stacking the column vectors of $\mathbf{A}$ from left to right.
$\mathbb{C}^n$	The set of $n \times 1$ column vectors with complex entries

$\mathbb{C}^{n\times m}$	The set of $n \times m$ matrices with complex entries
$\mathbf{J}$ ($\mathbf{J}^T$)	The one-lag down (up) shift square matrix whose first sub-diagonal entries below (above) the main diagonal are unity, whereas all remaining entries are zero
$\mathbf{e}_i$	The unit column vector with its ith entry equal to one, and its other entries equal to zero
$\mathbf{A}[r_1 : r_2, c_1 : c_2]$	The sub-matrix of $\mathbf{A}$ from r_1^{th} row to r_2^{th} row and from c_1^{th} column to c_2^{th} column
$\mathbf{I}$	The identity matrix

Chapter 1

Introduction

Wireless communication has become one of the fastest growing technologies during the past century. The wireless era began around 1895 when Guglielmo Marconi demonstrated the use of radio waves to communicate over large distances. After over one hundred years advancement, now the wireless systems have evolved to become a technology capable of providing instantaneous high data rates links to mobile users. Nevertheless, the rapid growth in demand for wireless communications services still encourages research into the design of wireless systems which can render a higher data rate to more mobile users.

Wireless communications systems, as opposed to their wireline counterpart, pose some unique challenges: (*i*) the limited allocated spectrum results in a limit on capacity, (*ii*) the radio propagation environment and the mobility of users give rise to signal fading and spreading in time, space and frequency, and (*iii*) multiuser interference arises from frequency reuse in cellular wireless communications systems. The search for effective technologies to mitigate these effects has been going on in the past two decades, as wireless communication

is experiencing rapid growth. Among these technologies are multiple access schemes, channel coding, and space-time signal processing techniques such as beamforming and blind methods. This thesis is focused on working out a variety of space-time signal processing algorithms addressing the above problems.

Understanding the physics of radio frequency (RF) wave propagation is crucial to the development of good models for space-time wireless signal processing. Radio wave propagation is a very complex phenomenon. In the following section, we attempt to characterize some key issues involved in this phenomenon and proceed to develop a discrete channel model.

1.1 Radio Propagation Model

A signal propagating through the wireless channel usually arrives at the destination along a number of different paths, referred to as multipath. These paths arise from scattering, reflection, refraction, or diffraction of the radiated energy off the objects that lie in the environment. Moreover, the received signal is much weaker than the transmitted signal due to phenomenon such as path loss and fading.

1.1.1 Path Loss and Fading

An important measure in mobile communications is the *path loss*. It is defined as the ratio between the received and transmitted power. The mean received signal level varies with distance d as d^{-n}, where n is a parameter in the range of $2-5$, depending on the type of environment [3, chapter 3]. The more build-up/obstructed the environment, the larger the n. $n = 2$ is realistic only for free-space propagation. In ideal free-space propagation, we have inverse square-law

spreading phenomenon and the received signal power is given by [3, chapter 3]

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2} \tag{1.1}$$

where P_r and P_t are the received and transmitted powers, respectively, λ is the wavelength, G_t, G_r are the power gains of the transmit and receive antennas, respectively, d is the transmitter-receiver (T-R) separation distance in meters. From the above equation, we can see that path loss increases not only with increasing transmitter-receiver distance d, but also with increasing operating frequency.

In addition to path loss, the signal exhibits fluctuations in power level. The fluctuations in signal level is called *fading*. There are two types of fading: *slow* (or *long-term*) fading and *fast* (or *short-term*) fading.

A signal experiences slow fading when it is shadowed by obstructions in the environment such as hills, buildings, etc. Thus this type of fading is mainly caused by terrain configuration and man-made structures between the transmitter and receiver. The envelope of the slow-fading signal is determined by the local mean of the fast-fading signal, i.e., the average signal level for areas of a few tens of wavelengths. Experiments have shown that, for paths of length of a few hundred meters or more, the received local mean power fluctuates as a log-normal distribution about the mean of the local power, that is, the local mean power expressed in logarithmic values (e.g. dB) has a Gaussian distribution [4]. Such a distribution is described by the following probability-density function:

$$p(x) = \begin{cases} \frac{1}{\sqrt{\pi}\sigma} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) & x > 0 \\ 0 & x < 0 \end{cases} \tag{1.2}$$

where x is a random variable representing the slow signal level fluctuation, μ and σ are the mean and standard deviation of x, respectively, expressed in decibels. The mean of this distribution is distance dependent, and the standard deviation is typically in the range of $5-12$ dB, with 8 dB being typical for macrocellular applications [5].

Fast fading is caused by the multiple reflections of the transmitted wave by objects around the mobile such as houses, trees, etc. The waves scattered by these objects have different attenuations and phases, and thus may add up constructively or destructively, causing fast fluctuations in the signal level. The received signal power may change by a few orders of magnitude (e.g., $20-40$ dB) within just a few wavelengths. When the mobile is completely obstructed from the base-station, i.e., there is no direct *line-of-sight* (LOS), then the envelope of the received signal is best modeled statistically as Rayleigh distribution which is given as follows [3, chapter 4]

$$p(x) = \begin{cases} \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} & x > 0 \\ 0 & x < 0. \end{cases} \tag{1.3}$$

If there is a direct path present, then it will no longer be a Rayleigh distribution, but it becomes Rician. The corresponding probability density function (pdf) is given by [3, chapter 4]

$$p(x) = \begin{cases} \frac{x}{\sigma^2} e^{-(x^2+A^2)/2\sigma^2} I_0(\frac{Ax}{\sigma^2}) & x \geq 0, A \geq 0 \\ 0 & x < 0 \end{cases} \tag{1.4}$$

where the parameter A denotes the peak amplitude of the dominant signal and $I_0(\cdot)$ is the modified Bessel function of the first kind and zero-order. See Figure 1.1 for a summary of all these fading phenomena.

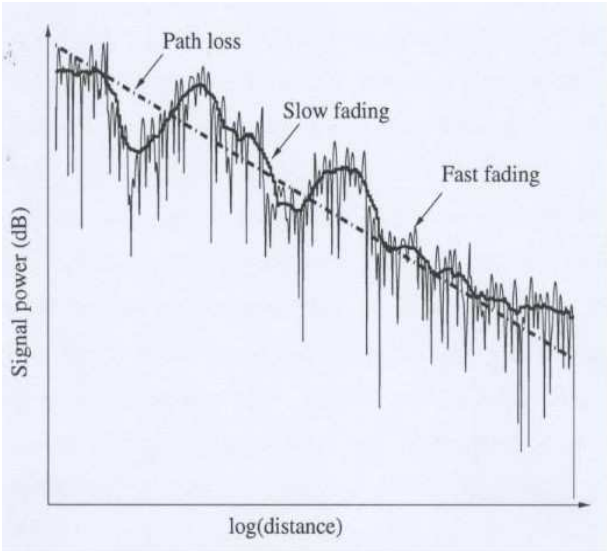


Figure 1.1: The fading phenomena (this figure is adopted from [2])

1.1.2 Multipath

The multipath phenomenon is caused by objects (scatterers) lying in the environment a radio signal is propagating in. Multipath causes the spread of signals in time and space (and also in frequency if the source is moving), i.e., the received signal consists of multiple time-delayed replicas of the transmitted signal, arriving from various directions. The cause lies in the three basic mechanisms that govern wave propagation: reflection, diffraction, and scattering [5]. Reflection occurs when a propagating wave impinges upon an obstruction with dimensions very large compared with its wavelength. Examples are the earth surface, buildings, etc. Refraction is a related phenomenon by which a component of the radio wave travels into the obstruction medium. Most buildings are made of materials that absorb a lot of the energy of the wave, such that the refracted wave is not significant in strength, compared to the reflected one.

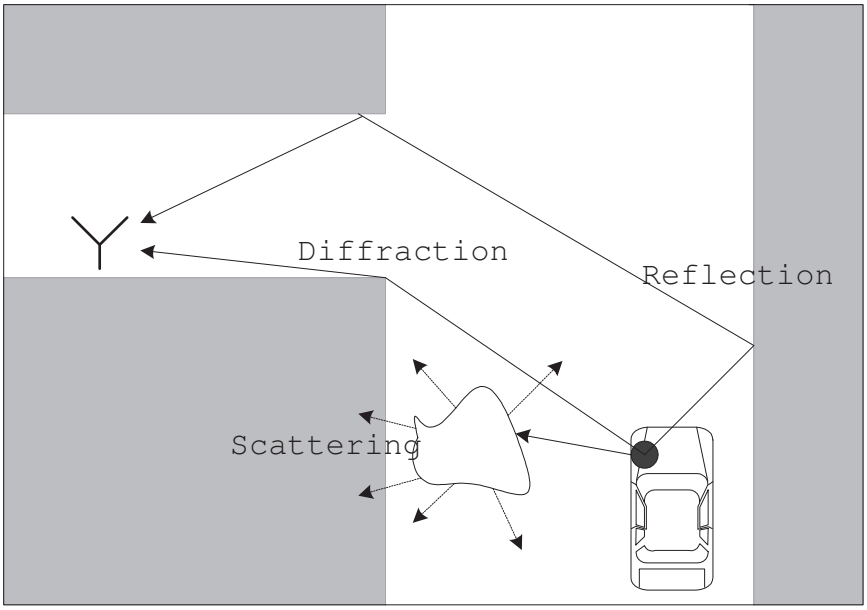


Figure 1.2: The propagation mechanisms

Reflection and refraction occur according to Snell's laws. Diffraction occurs when the radio path between the transmitter and receiver is obstructed by an impenetrable object; then, according to Huyghen's principle, secondary waves form behind this object. This phenomenon explains how radio waves arrive at the receiver even though there is no direct line-of-sight, as is the case in most urban environments. Lastly, scattering occurs when the wave impinges upon objects of dimensions that are on the order of the wavelength or less. In urban environments, such scattering objects are street signs or lamp posts. Scattering causes the energy of the wave to be radiated in many directions. See Figure 1.2 for a sketch of these propagation mechanisms.

The relative importance of these propagation mechanisms depends on the particular environment. Thus, if there is a direct line-of-sight between the mobile and base, then reflection dominates the propagation, while if the mobile is in a

heavily build-up area with no line-of-sight to the base, diffraction and scattering will play the major role.

To summarize, multipath propagation results in signal spreading in time (*delay spread*), space (*angle spread*), and frequency (*Doppler spread*). These three parameters describe the nature of the wireless communication channels. In a multipath propagation environment, several time-shifted and scaled versions of the transmitted signal arrive at the receiver through different paths. The spread of path delays is called delay spread. Delay spread causes frequency-selective fading, which implies that fading now depends on the frequency. It can be characterized in terms of *coherence bandwidth*, which represents the maximum frequency separation for which the frequency-domain channel responses at two frequency shifts remain strongly correlated. The coherence bandwidth is inversely proportional to the delay spread [3, chapter 4] and is a measure of the channel's frequency selectivity. A small ratio of coherence bandwidth to signal bandwidth indicates a frequency-selective channel. While delay spread and coherence bandwidth are parameters which describe the time dispersive nature of the channel in a local area, however, they do not offer information about the time-varying nature of the channel caused by either relative motion between the mobile and base station, or by movement of objects in the channel. Doppler spread and *coherence time* are parameters which describe the time varying nature of the channel in a small-scale region. Doppler spread is a measure of the spectral broadening caused by the time rate of change of the mobile radio channel and is defined as the range of frequencies over which the received Doppler spectrum is essentially non-zero. Coherence time is the time domain dual of Doppler spread. It is actually a statistical measure of the time duration over which the channel impulse response is essentially invariant, and quantifies the

similarity of the channel response at different times. The larger the coherence time, the slower the channel changes. A popular rule of thumb for modern digital communications is to define the coherence time as follows [3, chapter 4]

$$T_c = \sqrt{\frac{9}{16\pi f_m^2}} = \frac{0.423}{f_m} \tag{1.5}$$

where f_m is the maximum Doppler shift given by $f_m = v/\lambda$, λ is the wavelength, v is the velocity of the mobile station. It can be seen that the coherence time is inversely proportional to the Doppler spread. Time-selective fading is characterized by the coherence time of the channel. Angle spread at the receivers refers to the spread of angles of arrival of the multipath at the antenna array. Angle spread causes space-selective fading, which means that signal amplitude depends on the spatial location of the antenna. Space-selective fading is characterized by the *coherence distance*. The larger the angle spread, the shorter the coherence distance. Coherence distance represents the maximum spatial separation for which the channel responses at two antennas remain strongly correlated.

1.1.3 Space-Time Channel Model

Given all the considerations of the channel characteristics so far, we now proceed to derive a signal model for the space-time processing applications. From the above discussions, we know that the multipath propagation induces delay, angle and Doppler spreads. These spreads may have distinct effects on the channel modeling under different wireless communication systems. For example, consider a typical example of a global system for mobile communications (GSM) macrocell channel in a hilly terrain. GSM is a time division multiple

access (TDMA) cellular standard first developed in Europe and now extensively deployed around the world. It is characterized by a very short symbol period (3.7μs), a short time slot (0.577ms), and a high channel bandwidth (200kHz). Since the delay spread in hilly terrain and urban areas can be much larger ($10 \sim 15\mu$s, see [3, chapter 4]) than the symbol period, severe intersymbol interference (ISI) will be present and hence, the channel is highly frequency selective. On the other hand, as the time slot is short, the channel variation introduced from the Doppler spread is negligible for several or more time slots, despite the high velocity of the mobile stations. In contrast, the situation is reversed in the Interim Standard 54 (IS-54 – an American TDMA standard for mobile communications) system, where the symbol period is 41.6μs, the time slot is 6.66ms, and the bandwidth is much smaller (30kHz). We therefore have negligible ISI as the symbol period is large compared to the delay spread, and frequency selectivity of the channel is low. For high Doppler spread, the coherence time (say 5ms) is smaller than the time-slot duration, indicating significant channel variation within the slot.

In this thesis, we focus our study on the high rate dispersive communication systems such as GSM and DS-CDMA systems where the symbol period is short in comparison with the delay spread and thus the ISI constitutes a major impediment to the system performance. The channel is usually assumed to be time-invariant for our space-time processing in these communication systems. This is because there, the data packets used for space-time processing are relatively shorter in duration relative to the coherence time of the channel.

Consider a multipath channel illustrated in Figure 1.3. The signal from the mobile travels through a number of paths, each with its own fading and delay. The fading can be Rayleigh or Rician, with a Doppler spectrum that is flat or

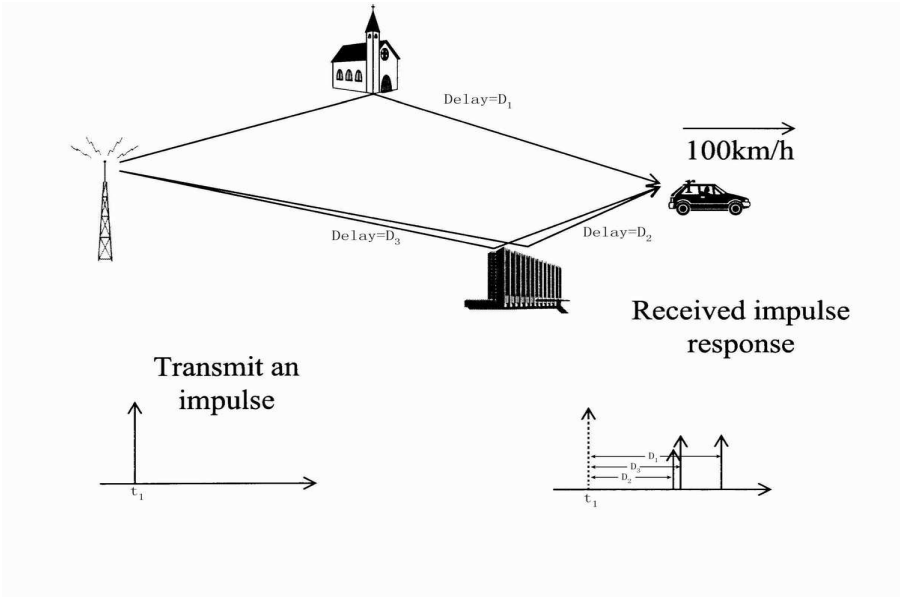


Figure 1.3: The multipath propagation

classical. These paths arrive at the receiver with different angles of arrival. Let the transmitted signal be

$$\tilde{s}(t) = s(t)e^{i2\pi f_c t} \tag{1.6}$$

The actual broadcast signal is the real part of $\tilde{s}(t)$. Here $s(t)$ is the complex baseband signal and f_c is the carrier frequency. The noiseless received signal $x(t)$ in this multipath environment, is then a superposition[1] of multiple replicas of the transmitted signal, scaled in amplitude and shifted in time, which is written as follows

$$x(t) = A \sum_n \alpha_n s(t - \tau_n) e^{j2\pi f_c (t - \tau_n)} \tag{1.7}$$

where A denotes the antenna gain, α_n denotes the amplitude scaling resulted

[1]The superposition principle applies because the medium (air) is linear.

from reflection/refraction along the nth path, and τ_n denotes the time shift resulted from the propagation delays. Clearly, if d_n is the propagation distance of the nth path, then $\tau_n = d_n/c$, where c is the speed of light.

If the transmitter (or receiver) antenna is moving with velocity v, then the received signal is also shifted in frequency. This phenomenon is known as the *Doppler effect*. The Doppler frequency shift can be shown to be $f_{D,n} = \frac{v}{\lambda}\cos(\theta_n)$, where θ_n is the direction of the nth wave with respect to the velocity vector v. Thus the signal model becomes [3, chapter 4]

$$x(t) = A\sum_n \alpha_n s(t-\tau_n)e^{j2\pi[(f_c+f_{D,n})t-f_c\tau_n]} \tag{1.8}$$

and can be further written as

$$x(t) = A\left[\sum_n \alpha_n s(t-\tau_n)e^{j2\pi(f_{D,n}t-f_c\tau_n)}\right]e^{j2\pi f_c t}. \tag{1.9}$$

We note that the parameters A, α_n, τ_n, and $f_{D,n}$ vary with time since the source and/or other objects in the environment are moving. However, this variation is usually negligible for short time intervals. Therefore we can assume that they are constant for short intervals. Moreover, for the short time intervals and small frequency shift $f_{D,n}$, the term $e^{j2\pi f_{D,n}t}$ can also be treated as a constant scalar, which is denoted as β_n. Thus the equivalent lowpass (baseband) noiseless received signal is

$$x(t) = A\sum_n \alpha_n\beta_n e^{-j2\pi f_c\tau_n}s(t-\tau_n). \tag{1.10}$$

Since the *multipath intensity profile* (or *power delay profile*), as a function of time, is always a continuum of peaks, which implies that there exist a large number of multipaths, the received baseband signal can be modeled by the

integral

$$x(t) = \int_{-\infty}^{+\infty} h(\tau)s(t-\tau)d\tau \tag{1.11}$$

where $h(\tau) = A\alpha(\tau)\beta(\tau)e^{-j2\pi f_c \tau}$, $\alpha(\tau)$ and $\beta(\tau)$ are the continuous-time forms of α_n and β_n, respectively. Eqn.(1.11) reveals that the channel operates as a linear filter with the impulse response of $h(\tau)$. Also, we note that $h(\tau)$ is time-invariant under the assumption that the parameters A, α_n, τ_n, and $f_{D,n}$ are constant for short time intervals. In conclusion, we can assume that the channel $h(\tau)$ is time-invariant for our space-time processing because, as indicated earlier, in high rate dispersive wireless communication systems, the data packets used are relatively short in duration as compared with the coherence time of the channel. For the digital wireless communication systems, the received signal at single antenna, $x(t)$, is the convolution of the transmitted sequence $\{s(k)\}$ with the channel $h(t)$

$$x(t) = \sum_k s(k)h(t-kT) + w(t) \tag{1.12}$$

where $w(t)$ is the additive noise with limited bandwidth. If we sample the received data $x(t)$ at the symbol rate $\frac{1}{T}$, then the sampled signal output is

$$\begin{aligned} x(nT) &= \sum_{k=n-L+1}^{n} s(k)h((n-k)T) + w(nT) \\ &= \sum_{l=0}^{L} h(lT)s(n-l) + w(nT). \end{aligned} \tag{1.13}$$

Here the channel $h(t)$ has a finite impulse response (FIR) of $L+1$ symbols. The assumption of a finite channel length of $h(t)$ has been verified by practical measurements [6]. These experiments show that the bulk of the energy of a

received symbol is concentrated in a finite time frame from the reception of the first ray. With a slight abuse of notation, Eqn.(1.13) can be written in a simpler form

$$x(n) = \sum_{l=0}^{L} h(l)s(n-l) + w(n). \tag{1.14}$$

The above discussed model is for single-input single-output (SISO) case. We can easily extend this model to single-input multiple-output (SIMO) by employing multiple antennas or by oversampling the received data $x(t)$ (we will discuss the oversampling model in detail in Chapter 2). This multichannel model (SIMO) arising from multiple sensors or oversampling the received data provides rich multichannel structures that can be exploited to facilitate algorithm design, which will be shown in the later part of our thesis. Let the received signal at the antenna array (q antennas) be arranged into an $q \times 1$ vector $\mathbf{x}(t) \triangleq [x_1(t)\ x_2(t)\ \ldots\ x_q(t)]^T$. Thus the received signal can be modeled as

$$\mathbf{x}(n) = \sum_{l=0}^{L} \mathbf{h}(l)s(n-l) + \mathbf{w}(n) \tag{1.15}$$

where $\mathbf{h}(l) \triangleq [h_1(l)\ h_2(l)\ \ldots\ h_q(l)]^T$, and $\mathbf{w}(n) \triangleq [w_1(n)\ w_2(n)\ \ldots\ w_q(n)]^T$, $\{h_i(l)\}$ denotes the subchannel i from the user to the ith antenna. The channel model can be written in vector form as

$$\mathbf{x}(n) = \bar{\mathbf{H}}\mathbf{s}(n) + \mathbf{w}(n) \tag{1.16}$$

where $\bar{\mathbf{H}} \triangleq [\mathbf{h}(0)\ \mathbf{h}(1)\ \cdots\ \mathbf{h}(L)]$ and $\mathbf{s}(n) \triangleq [s(n)\ s(n-1)\ \cdots\ s(n-L)]^T$. The signal model in Eqn.(1.16) is simple but rich and allows the application of many techniques developed in other signal processing contexts. $\bar{\mathbf{H}}$ is the

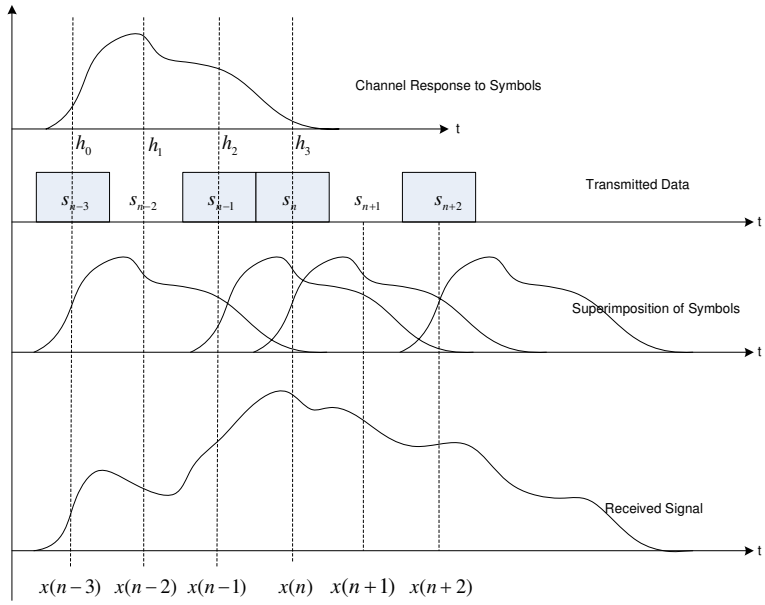


Figure 1.4: Smearing of received signal by ISI

symbol response channel that captures the effects of the array response, symbol waveform or pulse shaping function (note that although we did not consider this effect in our above derivation, it can be easily included into our model), and the path fading.

From the studied signal models, we can see that what impinges on the receiver is not only the transmitted symbol, but a superposition of all the delayed and scaled transmitted signals. This has the effect of smearing the symbols in time, which is shown in Figure 1.4. Time dispersion of the channel causes received symbols to trail for more than its allocated time period. Thus, components of one symbol begin to affect the received signal of adjacent symbols. This effect is known as *intersymbol interference*. It corrupts the received signal, thereby preventing the accurate reconstruction of the transmitted symbols. Figure 1.4 illustrates how time dispersion ultimately results in a received signal that has

little or no resemblance to the transmitted symbols. In such cases, accurate reconstruction of the transmitted symbol sequence is almost impossible without additional processing.

Besides the intersymbol interference introduced from the channel dispersion, another kind of interference arises from cellular frequency reuse in cellular mobile communication systems, which is called as *multiuser interference* (MUI) or *co-channel interference* (CCI). In wireless networks, a cellular layout with frequency reuse is exploited to support a large number of geographically dispersed users. In TDMA and frequency division multiple access (FDMA) systems, when a co-channel mobile operates in a neighboring cell, MUI will be present. The average signal-to-interference power ratio (SIR), also called as the protection ratio, depends on the reuse factor (K). The frequency reuse factor is $K = 1$ in CDMA networks, that is, the frequency is reused in every cell and, in fact, in every sector. A user signal is interfered by other users within its own cell and from neighboring sectors and cells. This leads to higher MUI. However, the MUI can be tolerated in CDMA because of the *processing gain*. The overall signal plus multiuser interference model at the base-station antenna array can be extended from Eqn.(1.16) and written as

$$\mathbf{x}(n) = \bar{\mathbf{H}}_1 \mathbf{s}_1(n) + \sum_{i=2}^{p} \bar{\mathbf{H}}_i \mathbf{s}_i(n) + \mathbf{w}(n) \tag{1.17}$$

where $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_i$ for $i \in \{2, \ldots, p\}$ are channels corresponding to signal and MUI, respectively, while $\mathbf{s}_1(n)$ and $\mathbf{s}_i(n)$ are the corresponding data sequences. Eqn.(1.17) appears to suggest that the signal and interference are baud synchronous. However, this can be relaxed and the time offsets can be absorbed into channel $\bar{\mathbf{H}}_i$ for $i \in \{2, \ldots, p\}$. Note that in multiuser cases, all the signals are desired. The above equation turns into a multiple-input multiple-output

(MIMO) model.

1.2 Motivation for Blind Channel Estimation

As analyzed in previous section, in high rate dispersive wireless communication systems, ISI arises from channel dispersion and becomes the major impediment to reliable wireless communications. We begin with the single-user case where we are only interested in demodulating the signal of interest. In this case, the interference from other users, i.e., MUI, can be treated as unknown additive noise and suppressed by an interference-suppression approach (see, e.g., [7]). Thus here what we are concerned most is how to cancel the effect of ISI; in order to do this, we need to estimate the wireless dispersive channel. Once the channel is estimated, various equalization techniques such as maximum likelihood (ML) and minimum mean-square-error (MMSE) can be used to compensate the ISI and recover the transmitted symbols accurately.

Traditional methods for channel estimation require the transmitter to periodically send signals that are known to the receiver (also called as "training sequences") in order to enable channel identification. Although the use of training sequences is probably the most robust way to estimate the channel, however, these periodically transmitted training sequences consume a considerable bandwidth and thus reduce the bandwidth usage efficiency. In fact, almost all of the current cellular systems embed training signals in the transmitted data, for example, in GSM, about 20% of the bandwidth is devoted to training. Moreover, in rapidly time-varying wireless channels, we may have to retrain frequently, resulting in poor spectral efficiency. There is, hence, an increased interest in the so-called "blind methods" that can estimate the channel without an explicit

training signal. Starting from the seminal work of Sato [8] in 1975, blind channel estimation methods have received considerable attention over the past decades. As compared to the traditional techniques, blind channel estimation methods identify the unknown wireless channels based only on the received signals and some *a priori* statistical information or properties of the input signals, without direct access to the transmitted signals. Therefore, blind methods can be used to eliminate or reduce the training sequences, thus saving precious bandwidth and improve the network capacity.

In the multiuser scenarios, our task is to jointly detect or extract all impinging signals rather than only the signal of interest. Such problems occur in channel reuse-within-cell (RWC) applications or in situations where we attempt to demodulate the interference signals in order to improve interference suppression. In this case, the multiuser interference which comes from other users is not negligible and can no longer be treated as additive noise. On the contrary, they now become the desired signals to be demodulated in order to achieve a better interference suppression effect. Obviously, to jointly demodulate all the user data sequences, the channels for all the arriving signals have to be estimated. Multiuser techniques need either training signals or blind methods to estimate the channels for all users. However, the use of training signals to estimate the channels becomes much complicated in this case. This is because the multiple training signals should be designed to have low cross-correlation properties so as to minimize cross coupling in the channel estimates. Moreover, training requires synchronization, which may not be feasible in multiuser scenarios. Thus, blind methods which do not need training and synchronization become a desirable alternative.

Outside the communications area, the need for blind channel estimation also

arises from other applications such as speech recognition and reverberation cancelation [9], image restoration [10, 11], and seismic signal processing [12, 13].

Although blind methods present some advantages, they also suffer from certain drawbacks compared to non-blind techniques. In general, blind algorithms tend to be computationally more expensive. Some blind methods converge to a local rather than global minimum due to their nonlinear nature. Also, blind methods, as opposed to the non-blind methods, introduce some inherent ambiguity in the channel estimation, e.g., an unknown phase ambiguity. The latter two problems can be resolved by using a short set of training signals. Although the algorithms are then no longer blind, they retain many of advantages associated with blind algorithms. Hence, purely blind versus training correspond to two extremes of a whole spectrum of system identification algorithms. In practice, system designers may combine ideas from both approaches to minimize the training signal requirements of non-blind methods, and yet obtain the robustness of blind methods at a lower computational cost. This *semi-blind* approach which can combine the advantages of blind and training-based (non-blind) techniques is discussed in [14, 15].

1.3 Review of Blind Channel Estimation Techniques

As indicated earlier, the term "blind" refers to methods that do not need training signals and rather exploit some a *priori* statistical information or properties of the input signals. These properties include non-Gaussianity, constant modulus (CM), finite alphabet (FA), cyclostationarity, etc. It is also noted that there is another kind of blind methods that exploit the spatial structure such as array manifold to estimate the direction-of-arrival (DOA) of the impinging

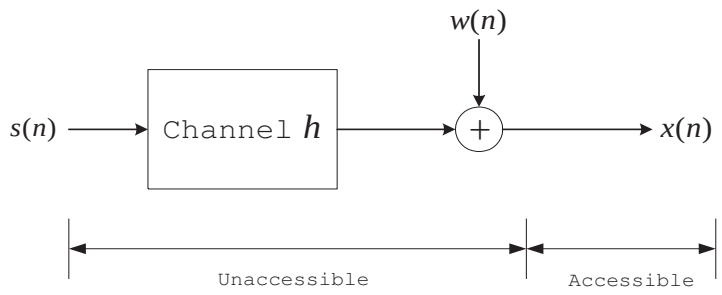


Figure 1.5: Schematic of blind channel identification

signals. They then use DOA estimates as a basis for determining the optimum beamformer. These methods were developed vigorously in the 1980s in military applications for reception of unknown or noise-like signals. See [16] for a survey. DOA-based methods, however, suffer from several drawbacks in cellular applications. First, DOA estimation requires an accurate knowledge of the array manifold. This needs expensive calibration support. Next, the number of antennas at cellular base stations varies from four to eight per sector, which might be insufficient for cellular environments with rich multipath and interference. Finally, while these methods can be quite effective against co-channel interference, their effectiveness against ISI depends upon the angle spread of multipath. In fact, when multipath and delay spread are present, they may have a poor or even disastrous performance. In this thesis, we focus our study on the blind methods which exploit the statistical information or properties of the input signals. The spatial structure, such as array manifold, is not assumed and exploited in our work.

At first glance, the blind channel estimation/identification problem illustrated in Figure 1.5 may not seem tractable. How is it possible to distinguish the signal from the channel when neither is known? The essence of blind channel

identification rests on the exploitation of structures of the channel (note that although the array manifold is not assumed, by stacking and arranging the received data, the channel matrix may still possess some certain structure like block Toeplitz structure, this will be detailed in Chapter 2) and properties of the input. A familiar case is when the input has known probabilistic description, such as probability distributions and moments. In such a case, the problem of estimating the channel using the output statistics is related to time series analysis. In communications applications, for example, the input signals may have the finite alphabet property, or sometimes exhibit cyclostationarity. This latter property was exploited in [17] to demonstrate the possibility of estimating a nonminimum phase channel using only second order statistics, which led to the development of many subspace-based blind channel identification algorithms.

The earliest blind techniques were primarily based on higher order statistics. They [8, 18–21] exploited higher order statistics (HOS) implicitly or explicitly to directly estimate the transmitted signal or estimate the single-input single-output channel/equalizer to combat the intersymbol interference. Since the phase information of the SISO channel only exists in higher order statistics, the second-order statistics (SOS) alone cannot recover the unknown SISO channel. The major breakthrough came in the 1990s. In the pioneering work [17], it was shown that under multichannel model, direct blind identification/equalization becomes possible using only the second-order statistics of the received data under quite general assumptions. This multichannel model (SIMO) arises from resorting to multiple sensors at the receiver or oversampling the received data by exploiting the receiver-induced cyclostationarity. Following [17], numerous SOS-based blind identification/equalization methods [1, 22–25] have been proposed. These methods include the matrix pair method [17], channel subspace

method [22], linear prediction method [23, 26], outer-product method [24], etc. As compared to HOS-based methods, SOS-based methods usually require much less received data samples to converge or to generate an accurate statistical estimation. Also, they are more computationally efficient, which is in contrast to the HOS-based methods that suffer from high computational costs in computing the higher order cumulants. Another advantage is that under the multichannel model, they can provide very elegant closed-form solutions to the channel estimation, while most previously HOS-based methods are iterative and suffer from the problem of local convergence. Due to the above mentioned reasons, the SOS-based methods have attracted significant attention over the past decade. Moreover, the study on blind channel estimation using SOS is not only confined to single-user's scenarios, there is also an increasing interest in blind channel estimation of MIMO systems because of its wide applications. For the MIMO systems, the multiple inputs may represent communication signals from multiple users or speech signals from multiple speakers, and the received signals are the convolutive mixtures of the multiple input signals. Many SOS-based methods have been proposed for blind MIMO channel estimation for the past decade, which include the channel subspace method for multiuser's case [27–29], the linear prediction method [30–32], the outer-product method [33,34], the whitening approach [35] and the frequency-domain approaches [36–39].

As in classical system identification problems, certain conditions about the channel and the source must be satisfied to ensure identifiability. These conditions are called as channel identifiability conditions. Channel identifiability has always been the issue closely related to various blind channel estimation problems. For the SIMO case, it is well known [17, 40] that unknown channel $\mathbf{h}$ can be blindly identified up to a constant factor from SOS of the received data if and

only if all SIMO sub-channels share no common zeros, i.e., the channel is *irreducible* [27, 41]. This condition permits a full-rank channel convolution matrix. The unknown constant factor is an inherent ambiguity for blind multichannel identification and can be determined by further knowledge available about the model. As for the MIMO case, for the independent and identically distributed (i.i.d.) inputs, it is noted that SOS-based algorithms can only estimate MIMO channels up to an unknown unitary mixing matrix. To further identify this unknown matrix, we have to resort to higher order statistics or other pertinent properties of the impinging source signals. Most SOS-based methods require the MIMO channel $\mathbf{H}(z)$ to be irreducible and *column-reduced* [27, 41], which guarantees the existence of a finite impulse response inverse to $\mathbf{H}(z)$, as shown in [27, 29]. However, it is shown that the column-reduced condition can be removed in some SOS-based algorithms [32, 33, 35]. In particular, SOS-based methods for blind system identification depend on the availability of channel diversity. In other words, the number of output signals must exceed the number of source signals in the MIMO system. The identifiability conditions for blind MIMO identification driven by the colored signals are investigated in [42]. It is shown that for the colored inputs, the sufficient conditions for the MIMO FIR channel to be identifiable (up to a scaling and permutation) using second-order statistics of the channel output are (i) the input colored signals should be of distinct power spectra; (ii) the channel is irreducible; and (iii) the number of channel outputs is strictly greater than the number of inputs. To design less restrictive algorithms, HOS can provide some distinct advantages, which include providing system phase information without requiring channel diversity, the ability to resolve matrix ambiguity to pure scaling and permutation indeterminacies, and asymptotic insensitivity to additive Gaussian noise. It is shown in [43], by exploiting HOS, the proposed algorithms only require a

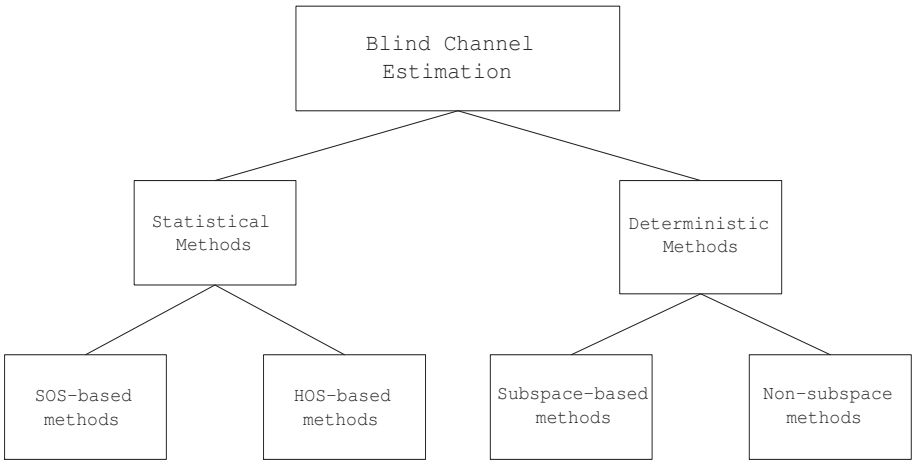


Figure 1.6: Classification of blind channel estimation methods

weaker identifiability condition, i.e., the number of antennas can be equal to the number of source signals, and the channel even need not be irreducible.

The various blind channel estimation methods can be classified based on the modeling of the input signals (see Figure 1.6 for the classification of blind methods). If inputs are assumed to be random with prescribed statistics (or distributions), the corresponding blind channel estimation schemes are considered to be *statistical*. The methods discussed up to now are all statistical and clearly, they can be divided into SOS-based or HOS-based methods. On the other hand, if the sources do not have a statistical description, or although the sources are random but the statistical properties of the sources are not exploited, the corresponding estimation algorithms are classified as *deterministic*.

Deterministic methods do not assume that the input sources have a specific statistical structure. This class of algorithms can be categorized into subspace and non-subspace algorithms. The subspace-based deterministic methods were inspired by the introduction of the multichannel SIMO platform. It has been

shown that this model provides rich multichannel structures that can be exploited for channel identification. In [44], a cross-relation (CR) approach was proposed by utilizing the cross relation between any two subchannels. Later, the classical channel subspace (CS) method [22] was put forward by Moulines *et al.*. The channel subspace method exploits the structure of the channel convolution matrix. It forces the signal subspace to have a block Toeplitz structure, which is orthogonal to the noise subspace. The channel subspace algorithm has a strong connection with the cross-relation algorithm. As pointed out in [27], they only differ in their parameterizations of noise and signal subspaces. Though, the CS method is relatively more complex than the CR method, it appears to provide better estimates under most conditions. The dual of the CS approach is to force the Toeplitz structure of the constructed input symbol matrix. This kind of approaches is presented in [45–47]. Other deterministic methods include the two-step maximum likelihood (TSML) algorithm [48], and the mutually referenced equalizers (MRE) method [49]. Perhaps a more striking property of the subspace-based deterministic methods is the so-called finite sample convergence property. Namely, when there is no noise, the estimator produces the exact channel using only a finite number of samples, provided that, of course, the identifiability conditions (the identifiability conditions for this kind of methods can be found in [40]) are satisfied. Also, deterministic methods can be applied to a much wider range of source signals since they do not rely on the specific assumptions concerning the input statistics. However, not using the source statistics affects their asymptotic performance. They are not so robust to noise as the SOS-based algorithms, especially when the identifiability condition is close to be violated. Therefore, these methods are most effective at high signal-to-noise ratio (SNR) and for small data samples scenarios. Almost all of the existing subspace-based deterministic algorithms are derived under a common

assumption that the channel order must be known *a priori*; while some statistical methods, e.g. [23, 24], require only the upper bound of the channel order L. The assumption that L is known may not be possible in practice. To address this problem, there are two approaches. First, there are well-known order detection schemes that can be used in practice, e.g. [50, 51]. Second, channel order detection and parameter estimation can be performed jointly. This joint estimation deterministic method was proposed by Tong *et al.* in [52, 53]. The method utilizes the so-called isomorphic relationship between the output and input subspaces to develop a least squares smoothing algorithm. In contrast to the subspace-based deterministic methods, the non-subspace deterministic methods are not dependent on the multichannel structures, instead, they exploit the properties of the input signals such as constant-modulus (CM) or finite alphabet (FA). This class of methods includes [54–62], etc. As compared to the subspace-based deterministic methods which can provide elegant closed-form identification solution, the non-subspace methods are usually iterative and suffer from local convergence. However, using the CM or FA properties of the input signals makes these algorithms not only confined to the multichannel models (SIMO), i.e. they are also capable of handling the model without channel diversity (SISO) and the model of multiuser scenarios (MIMO). It has been shown in [46, 47] that blind MIMO channel identification can be made more effective by exploiting the multichannel structures and the finite alphabet properties of the input signals simultaneously.

As mentioned before, statistical methods can be divided into SOS-based methods and HOS-based methods. The first SOS-based method was introduced by Tong *et al.* in [17] under the multichannel model. Since then, a variety of SOS-based methods have been proposed [1,22–25,63,64]. The channel subspace (CS)

method [22], the linear prediction (LP) approaches [23] and the outer-product decomposition (OPD) algorithm [24] are some of the most popular among them. The channel subspace method has its deterministic and stochastic versions, respectively. It is developed by exploiting the block Toeplitz structure of the channel convolution matrix (also called as channel filtering matrix). It has a simple structure and achieves good performance for SIMO system, but it requires precise knowledge of the channel order *a priori*, which may not be possible in practice. Also, its extension to MIMO channels is not successful because it generally can only estimate the channels subject to a polynomial matrix ambiguity [27]. In contrast, the extension of the LP and OPD algorithms to MIMO systems are quite straightforward, and they are valid even when the channel order is overestimated. Linear prediction-based approach was first presented by Slock [26], and was later generalized and refined by Abed-Meriam *et al.* [23]. The key idea comes from the recognition that multichannel moving average (MA) process is also autoregressive (AR). The main disadvantage of this algorithm is that it is a two-step approach whose performance depends on the accuracy of the estimated $\mathbf{h}(0)$. When noise is present and $\|\mathbf{h}(0)\|$ is small, performance degradation may be significant. Ding [24] proposed the OPD algorithm that obtains the channel directly, hence avoiding the problem of small $\|\mathbf{h}(0)\|$. Although OPD algorithm was not derived from the linear prediction view point, it has the same identification equation as the multistep linear prediction approach derived by Gesbert and Duhamel [65]. OPD algorithm is similar in properties and performance to the LP algorithm. There is another kind of SOS-based algorithms [1, 25] that directly compute the equalizers, unlike most blind algorithms available to date which first identify the channel and then use it to estimate the equalizer coefficients. Most existing SOS-based algorithms including the above mentioned algorithms assume the input signals to be i.i.d..

The works that consider the correlated (colored) input signals are much less, see [66–68] for the SIMO case. In fact, colored sources indeed occur in practice. For example, colored sources arise as a result of channel encoding [69]. Also, for the i.i.d. input sources, it is well known that the MIMO channel can only be determined up to an unknown unitary matrix that cannot otherwise be resolved using the second-order statistics. To resolve this residual static mixtures, one of several blind source separation (BSS) techniques should be resorted to. However, in contrast to the channels driven by white signals, the MIMO FIR channels driven by colored signals may provide us advantages in developing a complete closed-form SOS-based method without an extra BSS algorithmic step. The input colored signals should be of distinct power spectra, which is a sufficient condition for the MIMO FIR channel to be identifiable up to a scaling and permutation using second-order statistics of the channel output. There are some works [28,36,37,70,71] which studied blind identification/equalization of MIMO FIR channel driven by colored signals but with unknown statistics. For the case where the sources are colored with *a priori* known statistics, the proposed algorithms [72–74] may achieve a better performance by utilizing the knowledge of the input statistics. While most SOS-based methods for blind channel identification depend on the availability of channel diversity resulting from either oversampling (by exploiting the receiver-induced cyclostationarity) or multiple antennas, another class of SOS-based methods termed *transmitter induced cyclostationarity* (TIC) algorithms can identify SISO system only from the output second order statistics by exploiting the transmitter induced cyclostationarity. There are many different schemes proposed to induce cyclostationarity at the transmitter, which include periodic modulation [75, 76], repetition coding [77], combinations of repetition and modulation and filterbank precoding [78]. This class of algorithms do not assume any restrictive assumption on

the channel zeros, color of additive noise, and do not increase the transmission rate of the data stream. The blind identification of MIMO systems by utilizing the transmitter induced cyclostationarity is discussed in [79].

HOS-based methods can usually be classified as either linear or nonlinear. Most nonlinear methods involve the optimization of a nonlinear cost function constructed based on inverse filter criteria [18] or model fitting criteria [80–82]. These nonlinear methods tend to be more accurate than linear methods but usually require a good initialization to prevent local convergence and to speed up the searching process. In [83], Boss *et al.* proposed a novel two-step approach. It first iteratively ameliorates a blind linear equalizer and then estimates the channel parameters using the reference signal generated by the equalizer. Meanwhile, a large number of linear HOS approaches also exist in the literature. Linear approaches normally admit a closed-form solution, which can be used either as the final result or as an initialization for nonlinear methods. Some original contributions include [84–86]. Later in [87–90], to improve estimation accuracy, more advanced approaches were proposed to exploit a larger set of cumulants as well as certain inherent linear algebraic structure. HOS-based methods for blind MIMO system identification also have received much attention in the past. A broad class of HOS algorithms fall into the category of linear approaches [91–94]. Among them, Giannakis *et al.* generalized their earlier "GM method", which is designed for SISO systems, to MIMO systems under a somewhat strict condition imposed on the channel impulse response [91]. Later on, Swami and his colleagues presented a unified Kronecker product formulation to define cumulants of vector processes of arbitrary orders [93]. Parameter estimation algorithms for causal and noncausal multichannel AR, MA, ARMA models were developed under the similar identifiability condition required in [91]. In [94], Tong

proposed a new channel estimation algorithm for FIR MIMO systems, which can handle MIMO systems with fewer outputs than inputs. It presents a very general identifiability condition. Nevertheless, as indicated in the conclusion of [94], the algorithm implementation still requires further studies to avoid a number of matrix rank tests that may not be practical for estimated cumulant matrices. Recently, a new linear MIMO cumulant subspace method was put forward by Liang *et al.* [95]. On the other hand, a number of HOS algorithms were developed based on statistics matching or inverse filtering. Some inverse filtering-based algorithms [96,97] adopt an iterative procedure that successively recovers each active source signal using a MIMO equalizer and then estimates the corresponding subchannels based on the recovered source signal and the observed channel outputs. Clearly, both algorithms [96,97] rely on the premise of accurately converging MIMO equalizers. The channel estimation performance tends to suffer from error propagation when the number of users increases. The main limitation of all HOS-based methods, however, consists of their slow convergence rate due to the large estimation variance of HOS and thus the need of a large sample size for accurate time-average approximations of HOS. Consequently, HOS-based methods can hardly be applied in applications where fast channel variations and rapid adaptivity are essential.

In summary, we have seen that there is a wide range of possible blind techniques, which can be well suited for specific applications. The choice of a particular technique depends on the characteristics of the signals and the channel. Convergence and computational complexity are also important issues. Although we have not yet reached the final solution in blind signal processing, previous research activities have led to many fast converging and good performance techniques.

1.4 Motivations and Contributions of the Thesis

Our research interests are mostly inspired by the elegant closed-form solutions to blind channel identification offered by the SOS-based methods. Following the ground-breaking paper by Tong *et al.* [17] that presented a blind channel identification method that relies only on second-order statistics under the multichannel model, a number of SOS-based techniques have since been proposed and dominated the blind identification literature. As compared to HOS-based methods, the SOS-based methods have a much faster convergence rate and thus need a smaller data sample size for an accurate channel estimation. Also, by using the source statistics, the SOS-based techniques gain a stronger robustness to noise than the deterministic methods. Finally, unlike many iterative algorithms suffering from the problem of local minima, the elegant closed-form solutions rendered by the SOS-based methods can be used either as the final result or as an initialization for nonlinear methods. Given so many advantages, it is no wonder that they have attracted significant attention over the past decade.

Although numerous SOS-based algorithms have been proposed, however, most existing SOS-based algorithms [1, 17, 23–25] assume the input signals to be i.i.d.. The work that consider the case of correlated input signals are much sparser. In fact, colored sources indeed occur in practice. For example, colored sources arise as a result of channel encoding [69]. Also, in spatial division multiple access (SDMA) systems, the information bearing signals are filtered through correlative filters as a means of introducing redundancy for the purpose of better signal recovery at the receiver [72]. The methods for blind channel identification driven by colored sources include [66, 67, 70–73]. Usually, they

can be classified according to whether the input statistics of the sources are known *a priori* or unknown. It is clear that using the input statistics can help improve the performance of the proposed algorithms. On the other hand, the algorithms without requiring the knowledge of the input statistics may have wider applications. The previously developed work that address the case of input colored signals with *a priori* known statistics include [66, 67] for the SIMO systems and [72,73] for the MIMO systems. Among them, the work [66] imposes a somewhat restrictive condition on the source correlation, where an exponentially decaying autocorrelation function is assumed. The work [67] constitutes a direct extension of the TXK method [17] by exploiting the inherent structural relationship between the source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[1]$. Both [66] and [67] consider the blind channel estimation/equalization of SIMO models, and the extension of these algorithms to the MIMO systems is not straightforward because some of the relationships and properties in these works are no longer valid in multiuser case. The work [72] provides an elegant closed-form solution to blind MIMO channel identification. However, it is derived and presented under a correlative framework which is obtained by utilizing linear correlative filters at the transmitters, thus assigning distinct spectral patterns to the sources. Recently, a frequency-domain nonlinear iterative method [73] was proposed for blind MIMO channel estimation driven by colored sources. Due to its nonlinear nature, the method requires a good initialization in order to minimize the problem of local minima.

As we can see, although there already exist some papers addressing the problem of blind channel identification driven by colored sources with *a priori* known statistics, however, the rich information of the input statistics and the multichannel structures has not yet been fully exploited. In this thesis, by further

exploiting the inherent structural relationship between the source autocorrelation matrices and utilizing the new derived properties of the companion matrices, we propose a new closed-form solution for blind channel identification driven by colored sources. Our proposed method, unlike the works [66, 67], is applicable to both SIMO and MIMO cases. Also, as compared with [72], our proposed method achieves better performance with much less computational complexity and a less restrictive identifiability condition. The contribution of our work can be summarized as the following three aspects. First, the inherent structural relationship between the source autocorrelation matrices is further exploited as compared to the work [67]. Second, we derive some new properties of the constructed companion matrices. These properties play a key role in devising and validating our proposed algorithm. Third, the proposed algorithm compares favorably with other existing methods in many aspects. We will present this part of our work in Chapter 4 (SIMO) and Chapter 5 (MIMO) of this thesis. The results for SIMO case have been published in *IEEE Signal Processing Letters* [68] and the results for MIMO case have been published by *IEEE Trans. Signal Processing* [74].

Encouraged by the achieved results for blind channel identification driven by colored sources with known input statistics, we become interested in investigating the problem of blind channel identification/equalization driven by colored sources with unknown input statistics. For the case where the input statistics are *colored but unknown*, it seems much more difficult to devise a SOS-based algorithm since no prior statistical information of the transmitted signals can be utilized. One solution to this problem is given in [22], which proposed a subspace-based method by exploiting the block Toeplitz structure of the channel convolution matrix, and thus required no knowledge of input statistics what-

soever. The extension of [22] to MIMO systems was studied in [28]. There are some other work [70, 71] which studied blind identification/equalization of MIMO FIR channel driven by colored signals with unknown statistics. However, both work [70, 71] constitute a two-step approach that is based on [22]. They, firstly, determine source separating vectors or decorrelators to separate the sources. Once the sources are separated, the second step utilizes the subspace method [22] to estimate the resulted SIMO systems and the original MIMO systems. It is also noticed that some deterministic approaches that can handle arbitrarily correlated source signals have been discussed in [44,48,52,98] for blind SIMO channel identification/equalization. However, they are most effective at high SNR and for small data samples scenarios. In this thesis, we propose a new SOS-based method for blind equalization of SIMO FIR channel when the input signals are colored but the source statistics are unknown. It is shown that although the statistical information of the transmitted signals is not available, we can still estimate the zero-forcing (ZF) and minimum-mean-square-error (MMSE) equalizers of desired delays from the second-order statistics of the received data by exploiting the inherent structural relationship between source autocorrelation matrices of different delay lags. Our proposed method outperforms the existing methods [22, 49] significantly for the colored sources that have a weak autocorrelation. This part of our work will be presented in Chapter 6. The results have been published by *IEEE Trans. Signal Processing* [99].

After we obtained results for temporally correlated sources, our research interest then turns to blind identification of MIMO channel driven by spatially correlated sources (note that in our work [74], we assume the sources are temporally correlated but spatially uncorrelated). Thus far, although numerous SOS-

based techniques have been introduced for blind MIMO channel identification, however, they usually assume that the input sources are spatially mutually independent or, at least, uncorrelated. The work on blind identification of MIMO channel driven by spatially correlated sources are scarce. In this thesis, such a problem is also studied. It is shown that under certain specified conditions, the MIMO FIR channel can be completely identified using the second-order statistics of the channel output. A SOS-based method is proposed and the proof for the uniqueness of the system solution is provided. Our method can be successfully employed for blind nonlinear SIMO channel equalization. As compared to other existing methods [100, 101], our method renders a wider applicability for the input sources and exhibits better performance. The part of our work will be presented in Chapter 7 and also, the results have been submitted to *Signal Processing* [102].

Another thread of our research work is to study the problem of blind channel estimation by utilizing higher order statistics. As indicated earlier, HOS can provide some distinct advantages to design less restrictive algorithms. HOS-based methods could be applicable in such a scenario where channel diversity is not available, i.e. channel is SISO, or blind identifiability condition for SOS-based methods is close to be violated [103]. The HOS-based approaches have now evolved into a subspace era. To improve estimation accuracy, many advanced approaches were proposed to exploit a large set of cumulants as well as certain inherent linear algebraic structure. Compared with other methods, these subspace methods [87–90] achieve a better performance with less data samples. In this thesis, we propose a new linear subspace method that extracts the channel information by utilizing the so-called interference subspace cancellation vectors. Precisely, we devise our algorithm by exploiting the *partial column*

space overlapping relationship between a concatenated cumulant matrix and a target cumulant matrix to obtain these information-extraction vectors. With a similar computational complexity, our proposed algorithm compares favorably with other existing linear HOS-based methods. We will present this part of our work in Chapter 3. The results have been published by *IEEE Trans. Signal Processing* [104].

1.5 Thesis Outline

This thesis is organized as follows.

In Chapter 2, we introduce some mathematical preliminaries about higher order statistics and the data models for blind system identification.

Next, in Chapter 3, we propose a HOS-based linear method to identify the SISO channel. The results of this chapter have been published by *IEEE Trans. Signal Processing* [104].

In Chapter 4, we present a SOS-based method for blind identification of SIMO channel driven by colored sources. The input statistics of the source are assumed known *a priori*. The results of this chapter have been published in *IEEE Signal Processing Letters* [68].

In Chapter 5, the proposed method in Chapter 4 is extended to the MIMO systems. The properties on companion matrices are further exploited to prove the uniqueness of the system solution. The results of this chapter have been published by *IEEE Trans. Signal Processing* [74].

Chapter 6 introduces a SOS-based method for blind equalization of SIMO channel driven by colored sources with unknown statistics. The results of this chap-

ter have been published by *IEEE Trans. Signal Processing* [99].

Chapter 7 presents a SOS-based method for blind identification of MIMO FIR channel driven by spatially correlated sources. The proof of the uniqueness of the system solution is provided and the identifiability conditions are also investigated. The results of this chapter have been submitted to *Signal Processing* [102].

Finally, in Chapter 8, we conclude with a summary of our contributions and directions for future research.

Chapter 2

Background – Mathematical Preliminaries

Our primary goal in this chapter is to introduce all the important definitions and properties associated with moments and cumulants which are useful in the following chapters. We also discuss the data model for blind system identification that will be used in our work.

2.1 Moments and Cumulants

When signals are non-Gaussian the first two moments are not sufficient to define their probability density function (pdf) and consequently HOS, namely of order greater than two, can reveal other information about them than SOS alone can. Ideally, the entire pdf is needed to characterize a non-Gaussian signal. In practice this is not available but the pdf may be characterized by its moments. It should be noted that some distributions do not possess finite moments of all orders [105]. Moreover, the moments, even when they exist for

all orders, do not necessarily determine the pdf completely. Only under certain conditions will a set of moments determine a pdf uniquely. It is rather fortunate that these conditions are satisfied by most of the distributions arisen commonly. For practical purpose, the knowledge of moments may be considered equivalent to the knowledge of the pdf. Thus distributions that have a finite number of the lower moments in common will, in a sense, be close approximations to each other. In practice, approximations of this kind often turn out to be remarkably good, even when only the first three or four moments are equated [106].

2.1.1 Definitions and Properties

We focus our presentation on the real random variables. For the complex case, the readers can refer to [107] and the references therein. Let $\{y(n)\}, n = 0, \pm 1, \pm 2, \pm 3, \ldots$ be a random process, stationary up to order q; then, the pth order moment, $M_{p,y}(\tau_1, \tau_2, \ldots, \tau_{p-1})$ is defined as the joint p^{th} order moment of the random variables, $y(n), y(n+\tau_1), \ldots, y(n+\tau_{p-1})$. Because of the assumed stationarity, the pth order moment is a function only of the $(p-1)$ lags, $\{\tau_i\}_{i=1}^{p-1}$. We now write the moments of a stationary random process as

$$\begin{aligned} M_{p,y}(\tau_1, \tau_2, \ldots, \tau_{p-1}) &\triangleq Mom[y(n), y(n+\tau_1), \ldots, y(n+\tau_{p-1})] \\ &= E[y(n)y(n+\tau_1)\ldots y(n+\tau_{p-1})] \end{aligned} \tag{2.1}$$

where $E[\cdot]$ is the statistical expectation operator. The pth order cumulant exists, if all absolute moments of orders $q \leq p$ exist (and are bounded). Similarly, the pth order cumulants of $\{y(n)\}$ are $(p-1)$-dimensional functions which we now write in the form

$$C_{py}(\tau_1, \tau_2, \ldots, \tau_{p-1}) \triangleq \mathrm{Cum}[y(n), y(n+\tau_1), \ldots, y(n+\tau_{p-1})]. \tag{2.2}$$

The general relationship between moments and cumulants of any order can be found in [108]. Cumulants of orders greater than one are invariant to shift of mean. We will assume that the processes of interest are all zero-mean.

Hence, the second-order moment sequence (autocorrelation) of the zero-mean random process $\{y(n)\}$ is defined as

$$M_{2,y}(\tau) \triangleq E[y(n)y(n+\tau)]. \tag{2.3}$$

In this case, the second-order cumulants $C_{2,y}(\tau)$ are the same as $M_{2,y}(\tau)$, i.e. $C_{2,y}(\tau) = M_{2,y}(\tau)\ \forall\tau$. The third-order moment sequence is defined as

$$M_{3,y}(\tau_1,\tau_2) \triangleq E[y(n)y(n+\tau_1)y(n+\tau_2)] \tag{2.4}$$

and again $C_{3,y}(\tau_1,\tau_2) = M_{3,y}(\tau_1,\tau_2)\ \forall\tau_1,\tau_2$, where $C_{3,y}(\tau_1,\tau_2)$ is the third-order cumulant sequence. The fourth-order moment sequence is defined as

$$M_{4,y}(\tau_1,\tau_2,\tau_3) \triangleq E[y(n)y(n+\tau_1)y(n+\tau_2)y(n+\tau_3)] \tag{2.5}$$

and the fourth-order cumulants are

$$\begin{aligned} C_{4,y}(\tau_1,\tau_2,\tau_3) = {} & M_{4,y}(\tau_1,\tau_2,\tau_3) \\ & -C_{2,y}(\tau_1)C_{2,y}(\tau_2-\tau_3) - C_{2,y}(\tau_2)C_{2,y}(\tau_3-\tau_1) \\ & -C_{2,y}(\tau_3)C_{2,y}(\tau_1-\tau_2). \end{aligned} \tag{2.6}$$

As can been seen despite the fact that the second- and third-order cumulants (of zero-mean processes) are identical with autocorrelation and the third-order moment respectively, fourth-order moments are different from the fourth-order cumulants. The third- and higher-order cumulants of Gaussian processes are

identically zero (proof can be found in [109]). Since cumulants of order $p > 2$ of a Gaussian process are zero, cumulants provide a quantitative measure of the deviation from Gaussianity.

The properties of moments and cumulants can be summarized as follows [108]:

P1 If $\lambda_i, i = 1, \ldots, p$ are constants, and $y_i, i = 1, \ldots, p$ are random variables, then

$$\mathrm{Mom}(\lambda_1 y_1, \ldots, \lambda_p y_p) = (\prod_{i=1}^{p} \lambda_i)\mathrm{Mom}(y_1, \ldots, y_p)$$

and

$$\mathrm{Cum}(\lambda_1 y_1, \ldots, \lambda_p y_p) = (\prod_{i=1}^{p} \lambda_i)\mathrm{Cum}(y_1, \ldots, y_p).$$

P2 Moments and cumulants are symmetric functions in their arguments, i.e.

$$\mathrm{Mom}(y_1, \ldots, y_p) = \mathrm{Mom}(y_{j+1}, \ldots, y_p, y_1, \ldots, y_j)$$

and

$$\mathrm{Cum}(y_1, \ldots, y_p) = \mathrm{Cum}(y_{j+1}, \ldots, y_p, y_1, \ldots, y_j).$$

P3 If the random variables $\{y_i\}_{i=1}^{p}$ are independent of the random variables $\{z_i\}_{i=1}^{p}$, then

$$\mathrm{Cum}(y_1 + z_1, \ldots, y_p + z_p) = \mathrm{Cum}(y_1, \ldots, y_p) + \mathrm{Cum}(z_1, \ldots, z_p)$$

whereas in general

$$\begin{aligned}\mathrm{Mom}(y_1 + z_1, \ldots, y_p + z_p) &= E[(y_1 + z_1)(y_2 + z_2) \cdots (y_p + z_p)] \\ &\neq \mathrm{Mom}(y_1, \ldots, y_p) + \mathrm{Mom}(z_1, \ldots, z_p).\end{aligned}$$

However, for random variables $\{z_1, y_1, y_2, \ldots, y_p\}$, we have that

$$\mathrm{Cum}(y_1 + z_1, y_2, \ldots, y_p) = \mathrm{Cum}(y_1, y_2, \ldots, y_p) + \mathrm{Cum}(z_1, y_2, \ldots, y_p)$$

and

$$\mathrm{Mom}(y_1 + z_1, y_2, \ldots, y_p) = \mathrm{Mom}(y_1, y_2, \ldots, y_p) + \mathrm{Mom}(z_1, y_2, \ldots, y_p).$$

P4 If a subset of the random variables $\{y_i\}_{i=1}^{p}$ is independent of the rest, then

$$\mathrm{Cum}(y_1, y_2, \ldots, y_p) = 0$$

whereas in general

$$\mathrm{Mom}(y_1, y_2, \ldots, y_p) \neq 0.$$

2.1.2 Ergodicity and Moments

Ergodicity deals with the relationship between statistical averages and sample averages. A process $\{y(n)\}$ is ergodic in the most general form if, with probability one, all its moments can be determined from a single realization [108]. In other words, the expected value $E[\cdot]$ (or ensemble averages) can be replaced

by time averages, i.e.,

$$\begin{aligned} E[y(n)y(n+\tau_1)\cdots y(n+\tau_{p-1})] &= \langle y(n)y(n+\tau_1)\cdots y(n+\tau_{p-1})\rangle \\ &= \lim_{T\to\infty}\frac{1}{2T+1}\sum_{n=-T}^{+T} y(n)y(n+\tau_1)\cdots y(n+\tau_{p-1}) \end{aligned} \tag{2.7}$$

where $\langle\cdot\rangle$ is the time-average operator which has the same properties as the ensemble average operation $E[\cdot]$ if the process is ergodic.

We see from Eqn.(2.7) that time-averages of higher-order moments are functions of infinitely many random variables and, therefore, can be viewed as random variables themselves. What ergodicity implies is that the time averages of all possible sample sequences are equal to the same constant which, in turn, equals the ensemble average. Clearly, a process might be ergodic for certain higher-order moments and not for others [109]. Throughout this thesis we assume that if the process is ergodic, then Eqn.(2.7) holds for all orders up to p.

In practice, when we are given a finite length single realization of an ergodic process, i.e., $\{y(n)\}, n = -T, \ldots, +T$, we cannot compute the limits of Eqn.(2.7) but the *estimates*

$$\langle y(n)y(n+\tau_1)\cdots y(n+\tau_{p-1})\rangle = \frac{1}{2T+1}\sum_{n=-T}^{+T} y(n)y(n+\tau_1)\cdots y(n+\tau_{p-1}) \tag{2.8}$$

The estimation of higher-order moments and thus of a stochastic process can be found in details in [108].

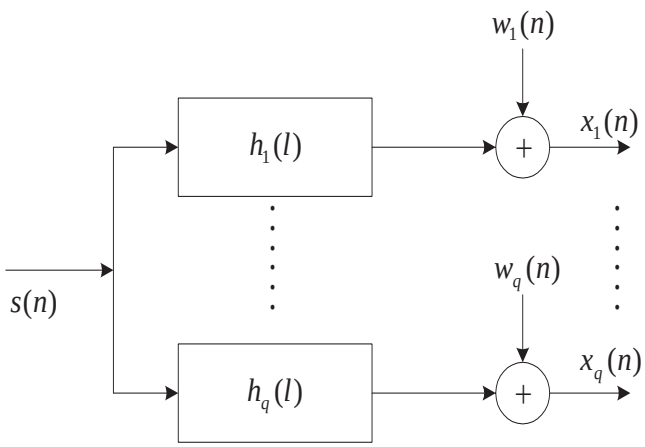


Figure 2.1: Single-input multiple-output model

2.2 Data Model

The data model for blind channel identification has been discussed briefly in Chapter 1. In this section, we elaborate on the SIMO and MIMO systems.

The SIMO model (see Figure 2.1) arises from employing multiple antennas or from oversampling the received data. For single user and multiple antennas, if we sample the received data at symbol rate, as indicated in Chapter 1, the channel model can be written as

$$\mathbf{x}(n) = \sum_{l=0}^{L} \mathbf{h}(l)s(n-l) + \mathbf{w}(n) \tag{2.9}$$

where $\mathbf{x}(n) \triangleq [x_1(n)\ x_2(n)\ ...\ x_q(n)]^T$, $\mathbf{h}(l) \triangleq [h_1(l)\ h_2(l)\ ...\ h_q(l)]^T$, and $\mathbf{w}(n) \triangleq [w_1(n)\ w_2(n)\ ...\ w_q(n)]^T$. Here q denotes the number of antennas, $\{h_i(l)\}$ denotes the subchannel i from the user to the ith antenna. For single user and single antenna, if we oversample the received data at a rate $\frac{1}{T_{\text{os}}}$, and we assume $T/T_{\text{os}} = q$, where T is the symbol period, then the received data $x(n)$

is oversampled to obtain q subsequences $x_i(n) = x(n + i/q)$, $i = 0, 1, \ldots, q-1$:

$$x_i(n) = \sum_{l=0}^{L} h_i(l)s(n-l) + w_i(n) \tag{2.10}$$

where $h_i(l) = h(l + i/q)$ and $w_i(n) = w(n + i/q)$. Thus we can still arrive at the model Eqn.(2.9) by stacking these q subsequences. We now derive the most popular data model widely used in blind system identification from Eqn.(2.9). By stacking the received data (channel output) vector $\{\mathbf{x}(n)\}$ and defining: $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^T(n)\ \ \mathbf{x}^T(n-1)\ \ \ldots\ \ \mathbf{x}^T(n-N)]^T$, $\vec{\mathbf{s}}(n) \triangleq [s(n)\ \ s(n-1)\ \ \ldots\ \ s(n-N-L)]^T$, and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^T(n)\ \ \mathbf{w}^T(n-1)\ \ \ldots\ \ \mathbf{w}^T(n-N)]^T$, where N is called the stack number or smoothed factor, we can therefore re-express Eqn.(2.9) as

$$\vec{\mathbf{x}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{2.11}$$

where the channel convolution matrix $\mathcal{H} \in \mathbb{C}^{(N+1)q \times d}$ is a block Toeplitz matrix written as follows with $d \triangleq N + L + 1$

$$\mathcal{H} \triangleq \begin{bmatrix} \mathbf{h}(0) & \ldots & \mathbf{h}(L) & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}(0) & \ldots & \mathbf{h}(L) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{h}(0) & \ldots & \mathbf{h}(L) \end{bmatrix}.$$

The above model Eqn.(2.11) is widely used in blind channel identification and most time-domain-based methods developed their algorithms based on this. An important assumption imposed on this model is that $\mathcal{H}$ is full column rank: a condition equivalent to requiring that the channel $\mathbf{h}(z)$ is irreducible, i.e., the subchannels have no common zero: $\text{rank}(\mathbf{h}(z)) = 1\ \forall z \neq 0$, where $\mathbf{h}(z)$ denote the Z-transform of the channel impulse response $\{\mathbf{h}(l)\}$. Under this

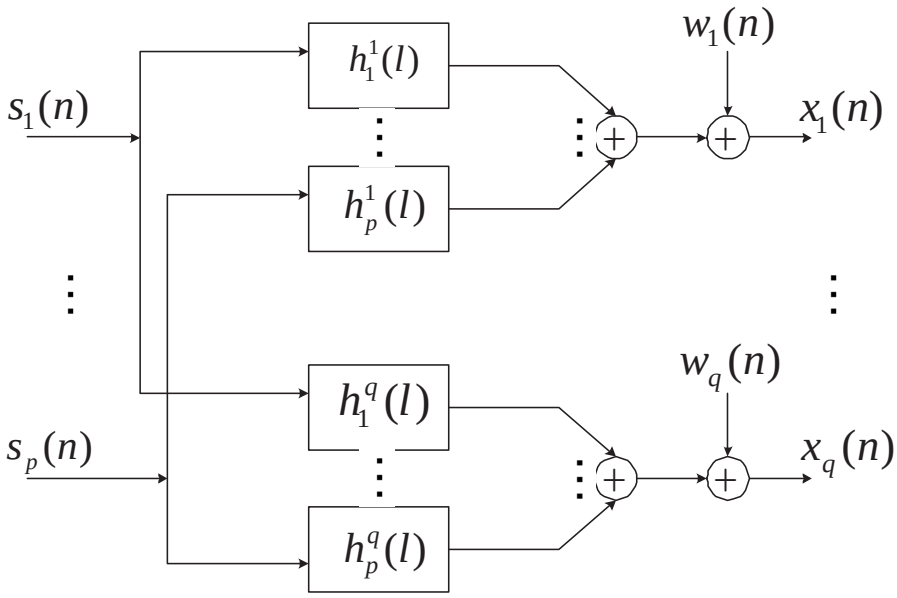


Figure 2.2: Multiple-input multiple-output model

assumption, exploiting the structure of the channel convolution matrix alone enables the channel identification [22].

For the multiuser scenarios, assume that p source signals $\{s_i(n), i = 1, 2, \ldots, p\}$ impinging at an array of q ($q > p$) antenna elements generate the set of q observations or measurements $\{x_i(n), i = 1, 2, \ldots, q\}$ at the sensor output (see Figure 2.2), we thus have

$$\begin{aligned}\mathbf{x}(n) &= \sum_{i=1}^{p} \mathbf{h}_i(n) \circledast s_i(n) + \mathbf{w}(n) \\ &= \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n) \end{aligned} \tag{2.12}$$

where $\circledast$ denotes the convolution operator, $\{\mathbf{h}_i(l) \triangleq [h_i^1(l) \;\; h_i^2(l) \;\; \ldots \;\; h_i^q(l)]^T\}$ denotes the multichannel filter corresponding to the ith user, $\{h_i^j(l)\}$ denotes the subchannel from the ith user to the jth antenna, and L_i represents the

channel order corresponding to the ith user. Similarly, by stacking the channel output vector $\mathbf{x}(n)$ and defining $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^T(n)\ \mathbf{x}^T(n-1)\ \ldots\ \mathbf{x}^T(n-N)]^T$, $\vec{\mathbf{s}}_i(n) \triangleq [s_i(n)\ s_i(n-1)\ \cdots\ s_i(n-N-L_i)]^T$ and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^T(n)\ \mathbf{w}^T(n-1)\ \ldots\ \mathbf{w}^T(n-N)]^T$, we can rewrite Eqn.(2.12) as

$$\vec{\mathbf{x}}(n) = \sum_{i=1}^{p} \mathcal{H}_i \vec{\mathbf{s}}_i(n) + \vec{\mathbf{w}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{2.13}$$

where $\mathcal{H}_i \in \mathbb{C}^{(N+1)q \times d_i}$ is a block Toeplitz matrix written as follows with $d_i \triangleq N + L_i + 1$

$$\mathcal{H}_i \triangleq \begin{bmatrix} \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{h}_i(0) & \ldots & \mathbf{h}_i(L_i) \end{bmatrix}$$

$$\mathcal{H} \triangleq \begin{bmatrix} \mathcal{H}_1 & \mathcal{H}_2 & \cdots & \mathcal{H}_p \end{bmatrix}$$

$$\vec{\mathbf{s}}(n) \triangleq \begin{bmatrix} \vec{\mathbf{s}}_1^T(n) & \vec{\mathbf{s}}_2^T(n) & \cdots & \vec{\mathbf{s}}_p^T(n) \end{bmatrix}^T.$$

Most time-domain SOS-based methods for blind MIMO FIR channel identification are based on the model Eqn.(2.13). As its counterpart for SIMO case, $\mathcal{H}$ is also required to be full column rank for most SOS-based methods. However, for the multiuser case, the condition "irreducible" is not sufficient to guarantee that $\mathcal{H}$ is full column rank. To ensure a full column rank $\mathcal{H}$, the channel is required to be irreducible and column-reduced [27]. Define $\mathbf{H}(l) \triangleq [\mathbf{h}_1(l)\ \mathbf{h}_2(l)\ \cdots\ \mathbf{h}_p(l)]$, let $\mathbf{H}(z)$ denote the Z-transform of $\{\mathbf{H}(l)\}$. Then the channel is irreducible

if and only if $\text{rank}(\mathbf{H}(z)) = p \ \forall z \neq 0$. Also, $\mathbf{H}(z)$ is said to be column-reduced if and only if its *leading column coefficient matrix* is nonsingular, i.e., $\text{rank}([H]_c) = p$, where $[H]_c$ is the leading column coefficient matrix defined as

$$[H]_c \triangleq \begin{bmatrix} \mathbf{h}_1(L_1) & \mathbf{h}_2(L_2) & \cdots & \mathbf{h}_p(L_p) \end{bmatrix}. \tag{2.14}$$

Given that the channel is irreducible and column-reduced, $\mathcal{H}$ is full column rank if the stack number N is chosen to satisfy $(N+1) \geq \sum_{i=1}^{p} L_i$ (see [27]). It is also noted that this well-known full-rank requirement for the MIMO channel convolution matrix $\mathcal{H}$ is not necessary for some existing SOS-based blind algorithms [32, 33, 35]. In these algorithms, the channel is only required to be irreducible, whereas the additional column-reduced condition can be removed. In our thesis, we still require this channel convolution matrix $\mathcal{H}$ to be full column rank. We will study how to relax this channel identifiability condition in our future work.

The additive noise $\mathbf{w}(n) \triangleq [w_1(n) \ w_2(n) \ ... \ w_q(n)]^T$ in the SIMO and MIMO models is usually characterized as spatially and temporally white with same variance, and statistically independent of the source signals. For SOS-based methods, the influence of the noise can be minimized by removing the noise contribution from the estimated autocorrelation matrices of channel output $\vec{\mathbf{x}}(n)$. From Eqn.(2.11) or Eqn.(2.13), we have

$$\mathbf{R}_x[0] = \mathcal{H}\mathbf{R}_s[0]\mathcal{H}^H + \sigma_w^2\mathbf{I}$$

where $\mathbf{R}_x[0] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^H(n)]$ and $\mathbf{R}_s[0] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n)]$. The noise variance σ_w^2 can thus be estimated as the smallest eigenvalues of $\mathbf{R}_x[0]$ and then subtracted from the estimated autocorrelation matrix to provide the proposed algorithms

with denoised autocorrelation estimates.

Up to now, we have discussed the data models for blind SIMO/MIMO channel identification and the basic assumptions imposed on the channel. The assumptions related to the source signals are not made here since they may vary from chapter to chapter. We will detail these assumptions in the respective chapters.

Chapter 3

Blind Estimation of SISO FIR Channel

In this chapter, we present a new HOS-based linear method for blind estimation of SISO FIR channel. The channel can be minimum-phase or nonminimum-phase channel. The proposed method is based on a series of fourth order cumulant matrices, where it is shown that by employing vectors chosen from the left null space of a concatenated cumulant matrix, the interference subspace of the channel convolution matrix can be cancelled and thus, channel information can be extracted. The proposed method is robust to channel order overestimation, and it has a similar computational complexity as other existing methods. Simulation results are included to validate the performance of the proposed algorithm.

3.1 Introduction

We consider the problem of blind identification of SISO FIR channel by using higher order statistics in this chapter. As indicated in Chapter 1, there have been a lot of research works [17, 22–24] on blind channel estimation by using the SOS. However, in this case, channel diversity should be obtained by oversampling the output data or by resorting to multiple sensors. Also, to identify this multichannel system, most SOS-based methods are required to meet a fundamental blind identifiability condition [17] that all subchannels do not share any common zeros, which may not be satisfied in practice [103]. In contrast, HOS methods could be applicable in such a scenario where channel diversity is not available, i.e. channel is SISO, or blind identifiability condition for SOS methods is close to be violated.

Numerous linear HOS-based methods [84–90] have been proposed over the past decade. Among them, some [87–90] can be considered as *subspace* methods because by using a set of cumulants, they exploit the special linear algebraic structure of the correspondingly constructed channel matrix. Compared to other methods, these *subspace* methods achieve better performance with lesser data samples. However, with the exception of the *Weighted Slices* (WS) algorithm proposed in [87], they usually require the precise knowledge of channel order and are sensitive to channel order overestimation. It is noted that a systematic generalization of [87] has been proposed in [110]. Compared to [87], [110] presents an enhanced way in exploiting the special linear algebraic (Toeplitz) structure. Recently, a matrix pencil technique [111] was adopted to estimate the channel by utilizing a series of fourth order cumulant matrices. As compared to [87–90], the work [111] shows an improvement in that it exploits the inherent structure

relationship between a pair of constructed cumulant matrices rather than only the linear algebraic structure of the channel matrix, thus demonstrating an enhanced performance and robustness to channel order overestimation.

In this chapter, we propose a new linear method that extracts the channel information by utilizing the so-called interference subspace cancellation vectors. It is noted that an implicit connection exists between our work and [111] because the non-trivial generalized eigenvectors derived in [111] can also be deemed as the interference subspace cancellation vectors. Compared to [111], we take a more direct approach to compute these information-extraction vectors. Precisely, we devise our algorithm by exploiting the *partial column space overlapping* relationship between a concatenated cumulant matrix and a target cumulant matrix to obtain these information-extraction vectors. This technique is essentially different from [111] that devises its algorithm by investigating the inherent structure of the constructed matrix pencil. As a consequence, our work induces the following advantages. Firstly, the algorithm in [111] requires a channel identifiability condition to make sure that the constructed matrix pencil has at least one unique nontrivial generalized eigenvalue. This identifiability condition is no longer necessary for our proposed algorithm. Secondly, the selection of interference subspace cancellation vectors can be excluded from our algorithm without detrimental effects. However, this selection procedure is necessary in [111] in order to distinguish the trivial from the non-trivial vectors.

This chapter is organized as follows. In Section 3.2, we introduce the system model and definitions of cumulant matrices. Next, in Section 3.3, we present the principle for channel identification and practical analysis of the proposed channel identification method. In the sequel, a practical algorithm based on the proposed channel identification method is developed in Section 3.4. Finally,

in Section 3.5, numerical simulation results are presented to demonstrate the performance of the proposed algorithm.

3.2 Preliminaries

3.2.1 System Model

We consider the SISO linear time-invariant FIR model derived in Chapter 1 (see Eqn.(1.14))

$$x(n) \triangleq h(n) \circledast s(n) + w(n) \triangleq \sum_{l=0}^{L} h(l)s(n-l) + w(n) \tag{3.1}$$

where $\circledast$ denotes the convolution operator; L denotes the channel order. The above model can be easily transformed as follows

$$\vec{\mathbf{x}}(n) = \mathbf{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n). \tag{3.2}$$

If we define $\vec{\mathbf{x}}(n) \triangleq [x(n)\ x(n-1)\ \ldots\ x(n-N)]^T$, $\vec{\mathbf{s}}(n) \triangleq [s(n)\ s(n-1)\ \ldots\ s(n-N-L)]^T$, $\vec{\mathbf{w}}(n) \triangleq [w(n)\ w(n-1)\ \ldots\ w(n-N)]^T$, and the channel convolution matrix $\mathbf{H}$ is an $(N+1) \times (N+L+1)$ Toeplitz matrix written as

$$\mathbf{H} \triangleq \begin{bmatrix} h(0) & \ldots & h(L) & 0 & \ldots & 0 \\ 0 & h(0) & \ldots & h(L) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & h(0) & \ldots & h(L) \end{bmatrix}. \tag{3.3}$$

We adopt the following assumptions: A1) The input signal is independent and identically distributed non-Gaussian stationary process with zero mean and

nonzero kurtosis $\gamma \triangleq \frac{\mu_4}{\mu_2^2}$, where μ_i denotes the ith central moment. A2) Additive noise is a zero-mean Gaussian process, and is statistically independent of the input signal. A3) The channel order is L. Without loss of generality, we assume $h(0) \neq 0$ and $h(L) \neq 0$. Our objective is to estimate the channel impulse response $\{h(l)\}$ by utilizing the fourth order statistics of the observed data $\vec{\mathbf{x}}(n)$.

3.2.2 Cumulant Matrices

We define a series of fourth order cumulant matrices [110, 111] of the channel output signals $\mathbf{C}[k]$ as

$$\mathbf{C}[k] \triangleq \text{cum}(\vec{\mathbf{x}}(n), \vec{\mathbf{x}}(n)^H, x(n-k), x^*(n-k)) \tag{3.4}$$

where $\mathbf{C}[k]$ is an $(N+1) \times (N+1)$ matrix with its $(i,j)^{th}$ element defined as $\mathbf{C}^{ij}[k] \triangleq \text{cum}(x(n-i+1), x^*(n-j+1), x(n-k), x^*(n-k))$. Invoking the cumulant properties P1–P4 introduced in Chapter 2 and the assumptions A1–A2, we have

$$\mathbf{C}[k] = \gamma \mathbf{H} \Lambda[k] \mathbf{H}^H \tag{3.5}$$

$$\Lambda[k] \triangleq \text{diag}(\underbrace{0, \ldots, 0}_{k}, |h(0)|^2, \ldots, |h(L)|^2, \underbrace{0, \ldots, 0}_{(N-k)}). \tag{3.6}$$

Eqn.(3.5) can be further rewritten as

$$\mathbf{C}[k] = \gamma \mathbf{H}_{[k+1:k+L+1]} \bar{\Lambda} \mathbf{H}^H_{[k+1:k+L+1]} \tag{3.7}$$

where $\bar{\Lambda} \triangleq \text{diag}(|h(0)|^2, \ldots, |h(L)|^2)$, $\mathbf{H}_{[k_1:k_2]} \triangleq [\mathbf{H}_{k_1} \ \cdots \ \mathbf{H}_{k_2}]$ denotes the part of $\mathbf{H}$ from its k_1^{th} column to its k_2^{th} column, in which $\mathbf{H}_i$ denotes the ith column of $\mathbf{H}$. In the following section, we will show that the channel information can be extracted based on a series of cumulant matrices $\mathbf{C}[k]$.

3.3 Channel Identification

3.3.1 Principle for Channel Identification

We consider a series of cumulant matrices $\mathbf{C}[k]$ with consecutive *delay lags* k, where $L \leq k \leq 2L$ and $N \geq 2L$. From Eqn.(3.7), we have

$$\begin{aligned}\mathbf{C}[L] &= \gamma \mathbf{H}_{[L+1:2L+1]} \bar{\Lambda} \mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma |h(0)|^2 \mathbf{H}_{[L+1]} \mathbf{H}_{[L+1]}^H + \gamma \mathbf{H}_{[L+2:2L+1]} \mathbf{D} \mathbf{H}_{[L+2:2L+1]}^H \end{aligned} \tag{3.8}$$

where $\mathbf{D} \triangleq \text{diag}(|h(1)|^2, \ldots, |h(L)|^2)$. Note that the column $\mathbf{H}_{[L+1]}$ is exactly an augmented channel vector with the desired channel vector $\mathbf{h} \triangleq [h(L) \ \ h(L-1) \ \ \cdots \ \ h(0)]^T$ padded with zero entries. Thus, the column space of $\mathbf{H}_{[L+1:2L+1]}$ is constructed by a rank-1 signal subspace spanned by $\mathbf{H}_{[L+1]}$ and a rank-L interference subspace spanned by columns of $\mathbf{H}_{[L+2:2L+1]}$. Our objective here is to find an *interference subspace cancellation vector*, $\mathbf{v}_c$, which is orthogonal to the interference subspace $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$, where $\mathcal{R}(\mathbf{A})$ denotes the range (column space) of the matrix $\mathbf{A}$. Such a vector can extract the signal subspace as follows

$$\begin{aligned}\mathbf{v}_c^H \mathbf{C}[L] &= \gamma \mathbf{v}_c^H \mathbf{H}_{[L+1:2L+1]} \bar{\Lambda} \mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma |h(0)|^2 \mathbf{v}_c^H \mathbf{H}_{[L+1]} \mathbf{H}_{[L+1]}^H \end{aligned}$$

$$= \alpha \mathbf{H}_{[L+1]}^H \tag{3.9}$$

where $\alpha \triangleq \gamma |h(0)|^2 \mathbf{v}_c^H \mathbf{H}_{[L+1]}$ is a complex scalar. Thus the channel is identified up to an unknown complex scalar α. In order to find such an interference subspace cancellation vector, the interference subspace, $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$, should be determined. Since the channel impulse response may contain zero coefficients, it is very hard for us to get an exact $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$. However, a subspace which includes the interference subspace can be obtained by the method described in the following theorem.

Theorem 3.1 *If we concatenate a series of cumulant matrices* $\mathbf{C}[k]$ *with consecutive delay lags to construct a new concatenated cumulant matrix*

$$\mathbf{S} \triangleq \begin{bmatrix} \mathbf{C}[k_1] & \mathbf{C}[k_1+1] & \cdots & \mathbf{C}[k_2] \end{bmatrix} \tag{3.10}$$

where $k_1 = L+1$, $k_2 = 2L$, *then we have*

$$\mathcal{R}(\mathbf{H}_{[L+2:3L+1]}) \supseteq \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L+2:2L+1]}). \tag{3.11}$$

Proof: Since $\mathbf{S}$ is a concatenation of a series of cumulant matrices, it is clear that

$$\mathcal{R}(\mathbf{S}) = \mathcal{R}(\mathbf{C}[k_1]) \cup \mathcal{R}(\mathbf{C}[k_1+1]) \cup \cdots \cup \mathcal{R}(\mathbf{C}[k_2]) \tag{3.12}$$

where the subspace $\mathcal{R}(\mathbf{C}[k])$ is spanned by columns of $\mathbf{H}$ whose corresponding diagonal elements in $\Lambda[k]$ are non-zero. Since $h(0) \neq 0$, it is obvious that

$$\mathcal{R}(\mathbf{H}_{[k+1:k+1+L]}) \supseteq \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+1]}). \tag{3.13}$$

We can easily get the results in Eqn.(3.11) by combining Eqn.(3.12) and Eqn.(3.13), given that $k_1 = L + 1$, $k_2 = 2L$. The proof is completed here. ■

On one hand, Theorem 3.1 shows that the interference subspace is included in $\mathcal{R}(\mathbf{S})$; on the other hand, it indicates that $\text{rank}(\mathbf{S}) \leq 2L$, which guarantees that $\mathbf{S}$ must have a non-degenerate left null space (i.e. dimension ≥ 1) when $N \geq 2L$. Here we use $\mathcal{R}(\mathbf{S}_l^{\perp})$ to denote the left null space of $\mathbf{S}$. It is clear that any vector that belongs to $\mathcal{R}(\mathbf{S}_l^{\perp})$ is orthogonal to the interference subspace. However, to be an interference subspace cancellation vector, not only the vector has to be orthogonal to the interference subspace, a hidden condition is that the vector should not be orthogonal to the signal subspace, otherwise α would be zero. Hence we are faced with a question: whether or not there exists a vector that belongs to $\mathcal{R}(\mathbf{S}_l^{\perp})$ satisfying this hidden condition. This problem is answered by the following theorem.

Theorem 3.2 *Suppose* $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_p]$ *is a basis for* $\mathcal{R}(\mathbf{S}_l^{\perp})$*, where* p *is the dimension of* $\mathcal{R}(\mathbf{S}_l^{\perp})$*, then we have* $\mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0}$*, which means that there exists at least one column vector in* $\mathbf{V}$ *that satisfies*

$$\mathbf{v}_i^H\mathbf{H}_{[L+1]} \neq 0 \qquad \exists \ 1 \leq i \leq p. \tag{3.14}$$

Proof: Since $\mathbf{V}$ is a basis for the orthogonal complement of $\mathcal{R}(\mathbf{S})$, it is obvious that, for any vector $\mathbf{g}$, $\mathbf{g}^H\mathbf{V} = 0$ if and only if $\mathbf{g}$ belongs to the left null space of $\mathbf{V}$, i.e., $\mathcal{R}(\mathbf{S})$. Therefore, to prove $\mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0}$, we only need to prove that $\mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{S})$. Note that $\mathbf{H}_{[L+1:N+L+1]}$ is a $(N+1)\times(N+1)$ lower triangular matrix with its main diagonal entries equal to $h(L)$ and i^{th} sub-diagonal entries below the main diagonal equal to $h(L-i)$. This structure guarantees that all columns except all-zero columns of $\mathbf{H}_{[L+1:N+L+1]}$ are linearly independent, i.e.,

$\mathbf{H}_{[L+1:N+L+1]}$ is full column rank after deleting all-zero columns. Therefore $\mathbf{H}_{[L+1:3L+1]}$ is also full column rank after deleting all-zero columns. Thus we have

$$\mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{H}_{[L+2:3L+1]}) \Rightarrow \mathbf{H}_{[L+1]} \nsubseteq \mathcal{R}(\mathbf{S}) \Rightarrow \mathbf{V}^H\mathbf{H}_{[L+1]} \neq \mathbf{0} \tag{3.15}$$

The proof is completed here. ■

Theorem 3.2 indicates that there exists at least one interference subspace cancellation vector from the basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$. Such an interference subspace cancellation vector obtained from $\mathcal{R}(\mathbf{S}_l^{\perp})$ can help to extract the channel information through Eqn.(3.9). In fact, it is noted that, in practice, the computed basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$ can always render us *more than one* interference subspace cancellation vector. Since the channel is estimated for each interference subspace cancellation vector, these multiple interference subspace cancellation vectors obtained from the basis for $\mathcal{R}(\mathbf{S}_l^{\perp})$ provide us with an estimation diversity that can be utilized to enhance our final estimation accuracy. For simplicity, here we assume that all these p column vectors $\mathbf{v}_i$ $(1 \leq i \leq p)$ can be taken as the interference subspace cancellation vectors. This assumption has been verified by our numerous simulations. Also, such an assumption can be justified in the following sense, that is, even if there exist some column vectors in $\mathbf{V}$ satisfying $\mathbf{v}_i^H\mathbf{H}_{[L+1]} = 0$, our proposed method can still work without any detrimental effects on the channel estimation. The reason is explained as follows. Without loss of generality, suppose that we have $\mathbf{v}_i^H\mathbf{H}_{[L+1]} \neq 0$ for $i \in \{1, \cdots, m\}$ and $\mathbf{v}_i^H\mathbf{H}_{[L+1]} = 0$ for $i \in \{m+1, \cdots, p\}$. Thus the estimated channel from the former vectors $\mathbf{v}_i$ for $i \in \{1, \cdots, m\}$ is $\mathbf{v}_i^H\mathbf{C}[L] = \lambda_i\mathbf{H}_{[L+1]}^H$ and the estimated channel from the latter vectors $\mathbf{v}_i$ for $i \in \{m+1, \cdots, p\}$ is $\mathbf{v}_i^H\mathbf{C}[L] = \mathbf{0}$. Since our final channel estimate is obtained by integrating the channel information

estimated from every vector $\mathbf{v}_i$ for $i \in \{1, \cdots, p\}$, theoretically, these all-zero vectors estimated from $\mathbf{v}_i$ for $i \in \{m+1, \cdots, p\}$ have no impact on the final estimation result.

3.3.2 Practical Analysis of Channel Identification

We study our channel identification method under the following two practical scenarios.

Channel with Small Head Taps

We study how and to what extent the proposed method is influenced when the multipath channel has small head taps. For simplicity, we assume that the multipath channel has only one small head tap, i.e., $|h(0)|$ is small and $|h(0)| \ll |h(1)|$. Recalling Eqn.(3.7), we have

$$\mathbf{C}[k] = \gamma |h_0|^2 \mathbf{H}_{[k+1]} \mathbf{H}_{[k+1]}^H + \gamma \mathbf{H}_{[k+2:k+L+1]} \mathbf{D} \mathbf{H}_{[k+2:k+L+1]}^H. \tag{3.16}$$

It is clear that a small $|h(0)|$ leads to a much smaller $\gamma |h(0)|^2 \mathbf{H}_{[k+1]} \mathbf{H}_{[k+1]}^H$, thus $\mathbf{H}_{[k+1]}$ has a negligible contribution in spanning the column space of $\mathbf{C}[k]$. Therefore Eqn.(3.13) should be modified as

$$\mathcal{R}(\mathbf{H}_{[k+2:k+1+L]}) \supseteq \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+2]}) \tag{3.17}$$

and, accordingly, Eqn.(3.11) should be modified as

$$\mathcal{R}(\mathbf{H}_{[L+3:3L+1]}) \supseteq \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L+3:2L+2]}). \tag{3.18}$$

It can be seen that here $\mathcal{R}(\mathbf{S})$ only includes partial subspace of the interference subspace. Hence the interference subspace cancellation vectors obtained from $\mathcal{R}(\mathbf{S}_l^{\perp})$ are only orthogonal to partial subspace of the interference subspace $\mathcal{R}(\mathbf{H}_{[L+2:2L+1]})$. Eqn.(3.9) should be rewritten as follows

$$\begin{aligned}\mathbf{v}_c^H\mathbf{C}[L] &= \gamma\mathbf{v}_c^H\mathbf{H}_{[L+1:2L+1]}\bar{\Lambda}\mathbf{H}_{[L+1:2L+1]}^H \\ &= \gamma|h(0)|^2\mathbf{v}_c^H\mathbf{H}_{[L+1]}\mathbf{H}_{[L+1]}^H \\ &\quad +\gamma|h(1)|^2\mathbf{v}_c^H\mathbf{H}_{[L+2]}\mathbf{H}_{[L+2]}^H \\ &\approx \alpha_1\mathbf{H}_{[L+2]}^H \end{aligned} \tag{3.19}$$

where $\alpha_1 \triangleq \gamma|h(1)|^2\mathbf{v}_c^H\mathbf{H}_{[L+2]}$, the term $\gamma|h(0)|^2\mathbf{v}_c^H\mathbf{H}_{[L+1]}\mathbf{H}_{[L+1]}^H$ can be omitted since $|h(0)|^2$ is much smaller as compared to $|h(1)|^2$. Thus the column $\mathbf{H}_{[L+2]}$ is extracted. It is noted that $\mathbf{H}_{[L+2]}$ is still an augmented channel vector surrounded by zero entries. As a generalization, we can conclude that the column $\mathbf{H}_{[L+k+1]}$ could be extracted if the channel has k small head taps, where $k < L$. Note that for every $k < L$, we can guarantee that the column $\mathbf{H}_{[L+k+1]}$ is an augmented channel vector which contains the complete channel information by choosing $N \geq 2L$. Therefore we can see that even if the multipath channel has small head taps, it will not have a detrimental effect to our proposed method, and we can still estimate the channel vector up to a delay ambiguity[1].

Channel Order Overestimated

In practice, it is almost impossible for us to obtain a precise channel order due to noise and estimation errors. Therefore it is very meaningful to investigate the

[1]Since, in practice, we do not know how many small head taps exist in the multipath channel, thus we also do not know how many zero entries are padded in the forepart of the estimated augmented channel vector. This can be considered as a delay ambiguity.

robustness of our proposed algorithm to channel order overestimation. When channel order is overestimated, Eqn.(3.13) should be revised as (It should be noted that here we do not consider small head taps)

$$\mathcal{R}(\mathbf{H}_{[k+1:k+1+L_e]}) \supset \mathcal{R}(\mathbf{C}[k]) \supseteq \mathcal{R}(\mathbf{H}_{[k+1]}) \tag{3.20}$$

where L_e is the overestimated channel order and $L_e > L$. Accordingly, Eqn.(3.11) should be revised as

$$\mathcal{R}(\mathbf{H}_{[L_e+2:3L_e+1]}) \supset \mathcal{R}(\mathbf{S}) \supseteq \mathcal{R}(\mathbf{H}_{[L_e+2:2L_e+1]}). \tag{3.21}$$

Since $\mathbf{H}_{[L_e+1:3L_e+1]}$ is also full column rank after deleting all-zero columns when channel order is overestimated, Theorem 3.2 remains true. Therefore Eqn.(3.9) still holds and can be rewritten as

$$\begin{aligned}
\mathbf{v}_c^H\mathbf{C}[L_e] &= \gamma\mathbf{v}_c^H\mathbf{H}_{[L_e+1:2L_e+1]}\bar{\Lambda}_{[L_e]}\mathbf{H}_{[L_e+1:2L_e+1]}^H \\
&= \gamma|h(0)|^2\mathbf{v}_c^H\mathbf{H}_{[L_e+1]}\mathbf{H}_{[L_e+1]}^H \\
&= \alpha_2\mathbf{H}_{[L_e+1]}^H
\end{aligned} \tag{3.22}$$

where $\alpha_2 \triangleq \gamma|h(0)|^2\mathbf{v}_c^H\mathbf{H}_{[L_e+1]}$ and $\bar{\Lambda}_{[L_e]} \triangleq \mathrm{diag}(|h(0)|^2, \ldots, |h(L_e)|^2)$. Thus the column $\mathbf{H}_{[L_e+1]}$ is extracted and the estimated channel vector is obtained by taking the 1st to $(L_e+1)^{th}$ entries out from $\mathbf{H}_{[L_e+1]}$. We can see that the overestimated channel taps $h(l)$ $(L_e \geq l > L)$ should be zero. It means that, theoretically, channel order overestimation has no effect on our proposed method.

3.4 Algorithm Development

Following the above analysis, we now develop a practical algorithm for channel identification. Theoretically, the interference subspace cancellation vectors can be chosen to be the left singular vectors associated with the p smallest singular value of $\mathbf{S}$, where $p = N + 1 - \text{rank}(\mathbf{S})$. Because of the finite sample size, the estimate $\hat{\mathbf{S}}$ would not be rank deficient in practice and in order to determine p, the rank of $\hat{\mathbf{S}}$ need to be estimated. Nevertheless, the determination of rank is always a tricky problem, especially for estimated cumulant matrix. Therefore it is better for us to find a simple way to go around this problem. Notice that we have $L \leq \text{rank}(\mathbf{S}) \leq 2L$ from Eqn.(3.11); this implies that the number of interference subspace cancellation vectors is upper-bounded and lower-bounded by $p_u = N + 1 - L$ and $p_l = N + 1 - 2L$ respectively. Since every interference subspace cancellation vector provides us with an estimated channel, we can only choose p_l interference subspace cancellation vectors from p_u candidate vectors which are the left singular vectors associated with the p_u smallest singular values of $\hat{\mathbf{S}}$. Of course, a simpler alternative is to choose the left singular vectors associated with the p_l smallest singular values of $\hat{\mathbf{S}}$ as the p_l interference subspace cancellation vectors, at the expense of mild performance degradation. For comparison purpose, the former which is more accurate is used in our work. Here we assume that the channel order L is known *a priori*. In practice, even if the channel order is overestimated, we can still determine $p_u = N + 1 - L_e$ and $p_l = N + 1 - 2L_e$ since under this case, we have $L_e \leq \text{rank}(\mathbf{S}) < 2L_e$ (see Eqn.(3.21)).

Until now, we have successfully circumvented the rank determination problem by choosing p_l qualified vectors from p_u candidate vectors. It is clear that these

p_u candidate vectors are not equivalent as they achieve different interference subspace cancellation effects. Hence, there are two problems faced by us. First, how to choose p_l qualified vectors from p_u candidate vectors, i.e. interference subspace cancellation vectors selection. Second, how to integrate the estimated channel information obtained from these p_l interference subspace cancellation vectors. We now enumerate the steps for our channel identification procedure.

1. Given the estimated channel order L, let $N \geq 2L$, compute a series of estimated fourth order cumulant matrices $\hat{\mathbf{C}}[k]$, where $L \leq k \leq 2L$, from the channel output samples.

2. Concatenate a series of $\hat{\mathbf{C}}[k]$ to construct a new cumulant matrix $\hat{\mathbf{S}}$ as given in Eqn.(3.10).

3. Compute the SVD of $\hat{\mathbf{S}}$. Choose p_u left singular vectors, $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_{p_u}$, associated with the p_u smallest singular values of $\hat{\mathbf{S}}$.

4. *Interference subspace cancellation vectors selection*: For each $\hat{\mathbf{v}}_i$, the estimated augmented channel vector can be computed as $\underline{\hat{\mathbf{h}}}_i = \hat{\mathbf{C}}[L]^H \hat{\mathbf{v}}_i$ for each $i \in \{1, \ldots, p_u\}$. And the ith estimated channel vector $\hat{\mathbf{h}}_i$ can be obtained by deleting the zero entries in $\underline{\hat{\mathbf{h}}}_i$. However, because of the possible delay ambiguity introduced by small head taps (see Section 3.3.2), $\hat{\mathbf{h}}_i$ should be chosen from $\{\hat{\mathbf{h}}_i^j\}$, where $j \in \{1, \ldots, L\}$ and $\hat{\mathbf{h}}_i^j$ is obtained by removing the jth to $(j+L)^{th}$ entries from $\underline{\hat{\mathbf{h}}}_i$. For each $\hat{\mathbf{h}}_i^j$, we compare the theoretical $\mathbf{C}[k]$ which is computed by using the estimated channel to $\hat{\mathbf{C}}[k]$, i.e. the estimated cumulant matrix. The distance between the theoretical $\mathbf{C}[k]$ and the estimated $\hat{\mathbf{C}}[k]$ is defined as

$$\text{dis} \triangleq \min_{\beta} \|\hat{\mathbf{C}}[k] - \beta \mathbf{C}[k]\|_{\text{F}}^2 \tag{3.23}$$

where $\|\cdot\|_{\mathrm{F}}$ stands for the Frobenius matrix norm, β is a scalar chosen to minimize the matrix norm. Thus we can obtain the resulted distance, denoted by $\mathrm{dis}(i,j)$, for each estimated channel vector $\hat{\mathbf{h}}_i^j$. Finally, from the computed $\mathrm{dis}(i,j)$ for $i \in \{1,\ldots,p_u\}$, $j \in \{1,\ldots,L\}$, we select the best p_l vectors from $\hat{\mathbf{v}}_i$, $i \in \{1,\cdots,p_u\}$, as the interference subspace cancellation vectors. The criterion for choosing these p_l vectors is as follows. Let $d_i \triangleq \min\{\mathrm{dis}(i,j)\}$ for $j \in \{1,\ldots,L\}$. Then we choose the vector $\hat{\mathbf{v}}_i$ as the interference subspace cancellation vectors if d_i is among the first p_l minimum values of $\{d_1,\cdots,d_{p_u}\}$.

5. *Channel information integration*: Given the selected p_l vectors from above, we have p_l corresponding estimated augmented channel vectors $\{\hat{\underline{\mathbf{h}}}_i\}$. We next integrate the channel information from these multiple estimated results. This step is similar to that in [111] and thus we describe it briefly as follows. First, we select a reference vector $\hat{\underline{\mathbf{h}}}_{i_r}$ by the following criterion

$$i_r = \arg\min_{i,j} |\mathrm{dis}(i,j)|. \tag{3.24}$$

Given the p_l estimated augmented channel vectors, estimate delay difference $\{\tau_i\}$ relative to the selected reference vector (the estimation of the relative delay difference can be found in the counterpart of [111]), and obtain the aligned vectors $\{\hat{\underline{\mathbf{h}}}_i^{(\tau_i)}\}$ with the same delay ambiguity. Concatenate all aligned vectors $\{\hat{\underline{\mathbf{h}}}_i^{(\tau_i)}\}$ and compute the SVD of the concatenated matrix. The ultimate estimation $\hat{\underline{\mathbf{h}}}$ is obtained as the left singular vector associated with the largest singular value of the concatenated matrix.

Finally, we compare the computational complexity of our proposed method to the other existing linear methods [87, 111]. We only consider the linear

algebraic operations involved in the algorithm implementation. It can be seen from previous part that our proposed algorithm requires to do SVD operation in step 3 and step 5 respectively. The dimension of the computed matrices are as follows

$$\begin{array}{ll} \text{Step 3} & (N+1)\times(N+1)L \\ \text{Step 5} & (N+1)\times p_l \end{array}$$

where we adopt $N = 2L + 1$ in our simulations. Thus in step 3, we have to compute the SVD of a $(2L+2)\times(2L+2)L$ matrix. However, it is noted that we only need to compute the left singular vectors of the matrix $\mathbf{S}$. This is equivalent to computing the right singular vectors of the tall matrix $\mathbf{S}^H$ with dimension $(2L + 2)L \times (2L + 2)$. From [112] [p. 254], we know that this computation requires $2mn^2 + 11n^3$ flops, where $m = (2L + 2)L$ and $n = (2L + 2)$. Hence the total flops required for our proposed algorithm are $(2L + 11)(2L + 2)^3$ and of order $O(L^4)$. In the case where the channel order L is not very large and smaller than order of tens, our proposed algorithm has a similar computational complexity as the algorithm [111][2]. Also, the computational complexity of our algorithm is less than that of the WS algorithm [87] since the latter involves computing the pseudo-inverse of a $(2L+1)\times(3L+1)(L+1)$ matrix and requires about $O(L^5)$ flops.

3.5 Simulation Results

Now we present simulation results to illustrate the performance of our algorithm. We compare our method, namely *Cumulant Interference Subspace Can-*

[2]For the proposed algorithm in [111], we need to compute the generalized eigenvalue decomposition of two $(3L+1)\times(3L+1)$ matrices, which requires at least $30(3L+1)^3$ flops (see p.385 of [112])

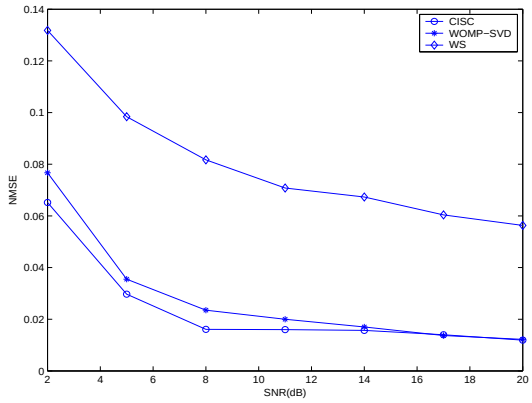
cellation (CISC) algorithm, to the other two linear methods, *Weighted Slices* algorithm (WS) proposed in [87] and *Cumulant Weighted Overlapping Matrix Pencil* algorithm (WOMP-SVD) presented in [111]. Among them, the WS method exploited the inherent linear algebraic structure, i.e. the structure of $\mathbf{H}$, of the constructed cumulant matrix (say $\mathbf{C}[k]$). The work [111] investigated the non-trivial generalized eigenvectors of the constructed matrix pencil $\{\mathbf{C}[k_1], \mathbf{C}[k_2]\}$ and showed that these vectors can be used to extract the channel information. In fact, the non-trivial generalized eigenvectors derived in [111] are exactly the so-called interference subspace cancellation vectors discussed in this chapter. For comparison purposes, we will only use the same set of fourth-order cumulants as CISC and WOMP-SVD for WS. In the implementation of CISC algorithm, we choose $N = 2L + 1$, also for simplicity, let $k = L$ in Eqn.(3.23) when computing the interference subspace cancellation vectors selection criterion. In our simulations, channel outputs are added with complex white Gaussian noise. The performance is measured by the Normalized Mean Square Error (NMSE) of the channel estimate, which is obtained by finding the complex scalar ρ that minimizes $\frac{\|\mathbf{h}-\rho\hat{\mathbf{h}}\|^2}{\|\mathbf{h}\|^2}$, and the symbol error rate (SER) of the estimated data symbols.

3.5.1 Example A

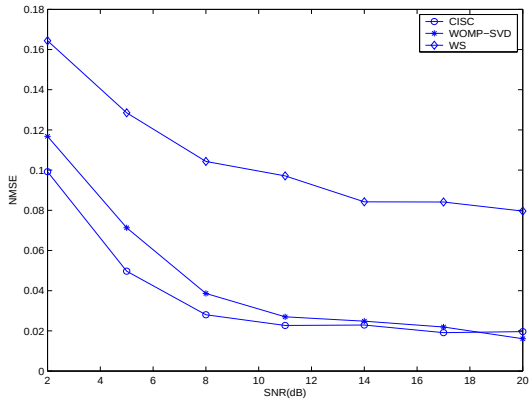
To study the robustness of the proposed algorithm to various channel conditions, we conduct simulation tests using randomly generated wireless channels, in which $\{h(l)\}$ is a complex, zero-mean Gaussian process with the channel order $L = 2$. Source signals are i.i.d QPSK signals. Results are averaged over 200 Monte Carlo runs and for each Monte Carlo run, a different FIR SISO channel is randomly generated.

In Figure 3.1, we show the NMSE of the channel estimate of these three algorithms as a function of SNR, with the number of samples used to estimate the signal statistics, T_s, varying from 400 to 1600. It can be seen that, as expected, all algorithms improve consistently as SNR or number of samples T_s increases. Also the proposed algorithm CISC presents a slightly better performance than WOMP-SVD and a significant performance advantage over WS. Once the channel is estimated, we can further detect the information sequences by adopting the Viterbi algorithm-based maximum likelihood detector. We present the SER performance of the algorithms in Table 3.1, in which the SER is a function of SNR and T_s. It can be seen that the SER performance depends on the following two parameters: SNR and NMSE of the channel estimate. On one hand, the SER performance deteriorates as SNR decreases. On the other hand, under a certain SNR, a more accurate channel estimate yields a lower SER. Also we can see that, in general, CISC shows a lower SER than WOMP-SVD since more accurate channel estimations are used in the Viterbi detector. Also CISC and WOMP-SVD outperform WS significantly in terms of SER.

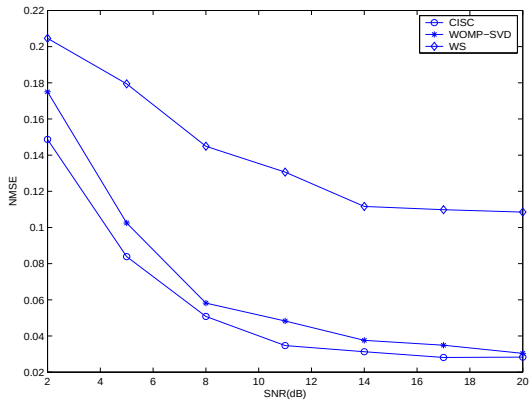
In Figure 3.2, we demonstrate the performance of CISC when channel order is overestimated with $T_s = 800$. It can be observed that the performance degrades more rapidly as SNR deteriorates. The reason, we suspect, is that the interference subspace cancellation vector selection incurs more errors when SNR becomes low and channel order becomes large. While at a moderate SNR level when SNR $\geq$ 8dB, the performance degradation is mild and acceptable, thus validating the theoretical analysis of the proposed method's robustness to channel order overestimation.



(a) $T_s = 1600$



(b) $T_s = 800$



(c) $T_s = 400$

Figure 3.1: NMSE of the channel estimate versus SNR under different number of samples used.

SNR(dB)	$T_s = 1600$		
	CISC	WOMP	WS
20	0.0002	0.0003	0.0162
17	0.0005	0.0008	0.0171
14	0.0019	0.0014	0.0254
11	0.0059	0.0082	0.0349
8	0.0415	0.0452	0.1048
5	0.1558	0.1609	0.2234
2	0.3060	0.3126	0.3546
SNR(dB)	$T_s = 800$		
	CISC	WOMP	WS
20	0.0005	0.0010	0.0299
17	0.0016	0.0015	0.0318
14	0.0061	0.0039	0.0346
11	0.0093	0.0108	0.0568
8	0.0528	0.0597	0.1324
5	0.1704	0.1794	0.2401
2	0.3226	0.3310	0.3789
SNR(dB)	$T_s = 400$		
	CISC	WOMP	WS
20	0.0038	0.0028	0.0393
17	0.0043	0.0043	0.0447
14	0.0048	0.0089	0.0560
11	0.0212	0.0280	0.0847
8	0.0708	0.0801	0.1405
5	0.2012	0.2179	0.2744
2	0.3423	0.3606	0.3950

Table 3.1: SER versus SNR of the respective algorithms under different number of data samples

3.5.2 Example B

In this example, we consider the source that employs 16-ary QAM digital format. The channel transfer function is given as $h(z) = -0.2039 + 0.4503z^{-1} + 0.7972z^{-2} - 0.3466z^{-3}$. In our simulations, results are averaged over 100 Monte Carlo runs. Figure 3.3 shows the performance of CISC and WOMP-SVD as 1600 and 800 data samples are used respectively. We can see that CISC owns a

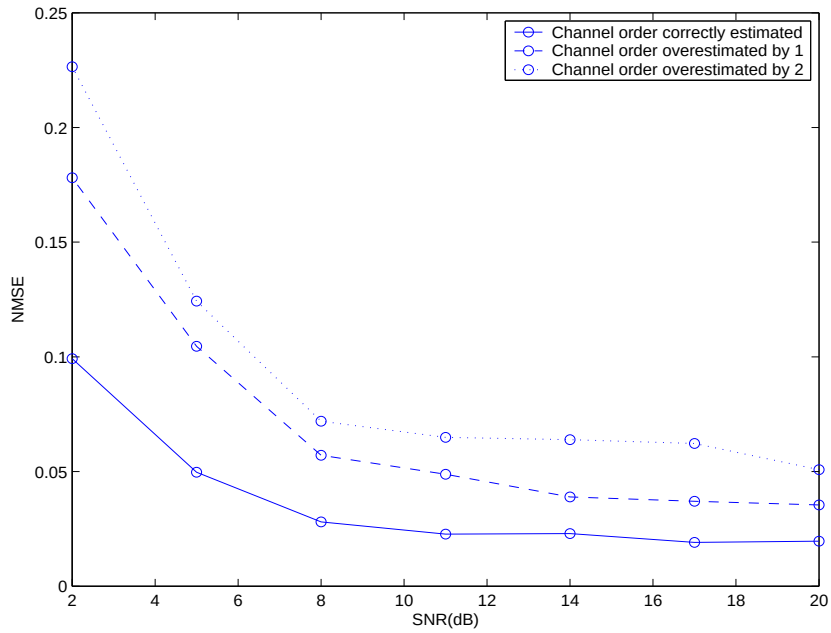


Figure 3.2: NMSE of the channel estimate versus SNR with channel order overestimated by 1 and 2 respectively

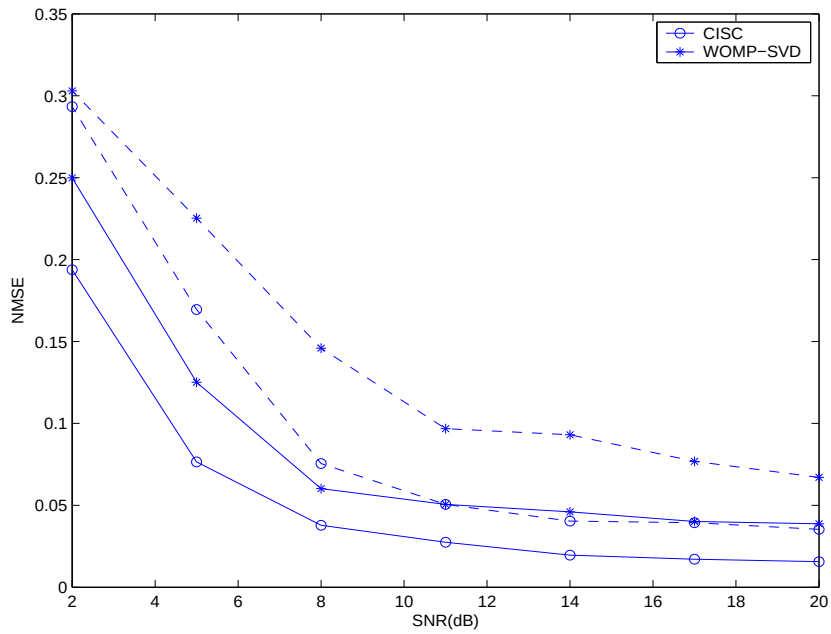


Figure 3.3: NMSE of the channel estimate versus SNR. Solid lines are for $T_s = 1600$; dashed lines for $T_s = 800$.

clear performance advantage over WOMP-SVD in both cases. It seems that, in such a channel scenario, CISC is more favorable than WOMP-SVD to obtain an accurate channel estimation. Besides, both algorithms suffer from a certain performance loss when 16-ary QAM digital modulation scheme is used. This is because, as compared to other simpler digital modulation schemes like QPSK, the source signals that employs 16-ary QAM digital modulation scheme induce a larger estimate variances between the estimated cumulants and the theoretical cumulants.

3.6 Summary

In this chapter, we presented a new linear HOS-based method for blind SISO FIR channel estimation. The proposed method exploits the partial column space overlapping relationship between a concatenated cumulant matrix and a target cumulant matrix to obtain a set of vectors which can be used to extract the channel information. The robustness of the proposed method to channel order overestimation was investigated. It was shown that, theoretically, channel order overestimation has no effect on our proposed method. This claim was also validated by our simulation examples. Simulation results showed that, with a similar computational complexity, our proposed algorithm compares favorably with existing linear HOS-based methods WS [87] and WOMP [111].

Chapter 4

Blind Identification of SIMO FIR Channel

In this chapter, we present a closed-form solution for blind estimation of SIMO FIR channel driven by colored source. The SOS of the input source are known *a priori*. The uniqueness of the system solution is proved by exploiting the property of companion matrices that are constructed from the inherent structural relationship between the source autocorrelation matrices. Numerical simulation results are presented to illustrate the performance of the proposed algorithm.

4.1 Introduction

In this chapter, we consider the problem of blind SIMO FIR channel estimation driven by colored source signals. As mentioned in Chapter 1, the multichannel model enables blind identification of the channel relying only on second-order statistics of the received data, which provides a much faster convergence rate and a more accurate channel estimation as compared with HOS-based methods.

This is the main reason for us to investigate the blind identification problem under the multichannel models. In particular, we are interested in the case where the input sources are colored and the second-order statistics of the input signals are known *a priori*. Colored sources with known statistics indeed occur in practice. For example, colored sources arise as a result of channel encoding [69], and the knowledge of the encoding scheme alone provides the required source statistics to the receiver. There are existing methods [22, 66, 67] that address the same problem as in this chapter. Among them, the work [22] proposed a subspace-based method by exploiting the block Toeplitz structure of the channel convolution matrix, and thus required no knowledge of input statistics whatsoever. Another work [66] imposes a somewhat restrictive condition on the source correlation, where an exponentially decaying autocorrelation function is assumed. The work [67] constitutes a direct extension of TXK method [17] by exploiting the inherent structural relationship between the source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[1]$. In this chapter, we propose a new closed-form solution for blind channel estimation driven by colored source. The contribution of our work consists of the following three aspects. Firstly, the inherent structural relationship between source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ is further exploited. Secondly, we derive a new property of a pair of constructed companion matrices, which plays a key role in devising and validating our algorithm. Thirdly, unlike other methods [22, 66, 67] which have difficulties in extending to the multiuser scenarios, the proposed algorithm has the potential to extend to the MIMO systems (we will show this point in next chapter). We include computer simulations to study the performance of the proposed algorithm.

4.2 System Model and Basic Assumptions

We begin by considering the SIMO FIR channel model derived in Chapter 2 (see Eqn.(2.9))

$$\mathbf{x}(n) \triangleq \mathbf{h}(n) \circledast s(n) + \mathbf{w}(n) \triangleq \sum_{l=0}^{L} \mathbf{h}(l)s(n-l) + \mathbf{w}(n) \tag{4.1}$$

where $\{s(n)\}$ is the zero mean, wide sense stationary sequence of transmitted symbols, $\{\mathbf{x}(n)\}$ is the $q \times 1$ channel output vector, $\{\mathbf{w}(n)\}$ is the $q \times 1$ white noise vector, and $\{\mathbf{h}(n)\}$ represents the multichannel impulse response. By stacking the channel output vector $\{\mathbf{x}(n)\}$ and defining: $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^T(n) \;\; \mathbf{x}^T(n-1) \;\; \ldots \;\; \mathbf{x}^T(n-N)]^T$, $\vec{\mathbf{s}}(n) \triangleq [s(n) \;\; s(n-1) \;\; \ldots \;\; s(n-N-L)]^T$ and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^T(n) \;\; \mathbf{w}^T(n-1) \;\; \ldots \;\; \mathbf{w}^T(n-N)]^T$, we can re-express Eqn.(4.1) as the following matrix form

$$\vec{\mathbf{x}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{4.2}$$

where the channel convolution matrix $\mathcal{H} \in \mathbb{C}^{(N+1)q \times d}$ is a block Toeplitz matrix defined in Chapter 2.

The following assumptions are adopted in this chapter: A1) $\mathcal{H}$ is full column rank: a condition equivalent to requiring that the channel $\mathbf{h}(z)$ is irreducible. A2) Channel order L is assumed to be known *a priori*. A3) Source signal is a zero mean, wide sense stationary colored signal whose input statistics are available. Its autocorrelation matrix with lag k is defined as $\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n-k)]$. A4) Additive noises are spatially and temporally white, and they are statistically independent of the source. Here A1 and A4 are basic assumptions for blind channel identification problem and have been elaborated in Chapter 2. Knowledge of channel order is critical to blind channel identifi-

cation methods and assumed in A2. In practice, some channel order detection schemes based on the minimum description length (MDL) principle [50, 51] or Akaike's information criterion (AIC) [113] can be employed to estimate the channel order. In this chapter, our objective is to estimate the channel impulse response by utilizing the second-order statistics of the observed output data and the knowledge of the source statistics.

4.3 Proposed Channel Identification Method

In order to simplify the presentation of the proposed channel identification method, we assume the noiseless case. Thus the autocorrelation matrix of the received data $\vec{\mathbf{x}}(n)$ with lag k can be expressed as

$$\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H. \tag{4.3}$$

Our goal is to find an estimate of $\mathcal{H}$ from Eqn.(4.3) by using the knowledge of $\mathbf{R}_s[k]$. We commence by introducing the following lemma.

Lemma 4.1 *Given* $\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H$, $\mathcal{H}$ *is full column rank and* $\mathbf{R}_s[0]$ *is invertible, we have*

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger \tag{4.4}$$

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]\mathcal{H} = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]. \tag{4.5}$$

Proof: To justify Lemma 4.1, we need to prove that

$$\mathbf{R}_x^\dagger[0] = (\mathcal{H}^H)^\dagger\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger. \tag{4.6}$$

Typically, $\mathbf{A}^\dagger$ is defined to be the unique matrix $\mathbf{T}$ that satisfies the four *Moore-Penrose conditions*: [112]

$$\begin{array}{ll} \text{(i)}\ \mathbf{ATA} = \mathbf{A} & \text{(iii)}\ (\mathbf{AT})^H = \mathbf{AT} \\ \text{(ii)}\ \mathbf{TAT} = \mathbf{T} & \text{(iv)}\ (\mathbf{TA})^H = \mathbf{TA}. \end{array}$$

Therefore we only need to prove that $\mathbf{R}_x^\dagger[0]$ defined in Eqn.(4.6) satisfies the above four *Moore-Penrose conditions*. This can be easily done and thus omitted here. ■

For convenience, let

$$\begin{array}{ll} \Upsilon_{2k-1} \triangleq \mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] & \Upsilon_{2k} \triangleq \mathbf{R}_x[-k]\mathbf{R}_x^\dagger[0] \\ \Theta_{2k-1} \triangleq \mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] & \Theta_{2k} \triangleq \mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]. \end{array}$$

We can therefore re-express Eqn.(4.5) (choose $K \geq k \geq 1$) as

$$\Upsilon_i \mathcal{H} = \mathcal{H}\Theta_i \qquad 2K \geq i \geq 1. \tag{4.7}$$

The above set of equations can be used to identify the channel $\mathcal{H}$ since the knowledge of Θ_i is known *a priori* and the information of Υ_i can be obtained from the second-order statistics of the observed data. By exploiting the block Toeplitz structure of $\mathcal{H}$, we can rewrite Eqn.(4.7) as

$$\mathcal{T}_1[\Upsilon_i]\mathbf{h} = \mathcal{T}_2[\Theta_i]\mathbf{h} \qquad 2K \geq i \geq 1 \tag{4.8}$$

where $\mathbf{h} \triangleq \begin{bmatrix}\mathbf{h}(0)^T & \cdots & \mathbf{h}(L)^T\end{bmatrix}^T$, $\mathcal{T}_1[\cdot]$ and $\mathcal{T}_2[\cdot]$ respectively represent a certain transformation on the bracketed matrix. The transformed matrices $\mathcal{T}_1[\Upsilon_i]$ and $\mathcal{T}_2[\Theta_i]$ are all of the same dimension $\mathbb{C}^{(N+L+1)(N+1)q\times(L+1)q}$. Therefore we may

estimate $\mathbf{h}$ by the following criterion

$$\hat{\mathbf{h}} = \arg \min_{\|\mathbf{u}\|=1} \sum_{k=1}^{2K} \| \left[\mathcal{T}_1[\Upsilon_k] - \mathcal{T}_2[\Theta_k] \right] \mathbf{u}\|^2. \tag{4.9}$$

The above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. However, this criterion fails to provide the true channel estimation if the solution of Eqn.(4.8) is not unique, i.e. there exist other non-zero vectors that are *linearly independent of* $\mathbf{h}$ and also satisfy Eqn.(4.8). Hence we are faced with the following problem, that is, whether or not the solution of Eqn.(4.8) is unique (up to a scalar factor) and under what conditions the solution of Eqn.(4.8) will be unique. This problem is studied in the following and we will establish the uniqueness of the solution to Eqn.(4.8) by using only the autocorrelation matrices $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm 1]$, i.e. the uniqueness of the solution can be guaranteed by choosing $K = 1$ in Eqn.(4.8).

We begin by observing the structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$. It can be seen that the last $d-1$ rows of $\mathbf{R}_s[1]$ are the first $d-1$ rows of $\mathbf{R}_s[0]$, and the first $d-1$ rows of $\mathbf{R}_s[-1]$ are the last $d-1$ rows of $\mathbf{R}_s[0]$. Hence we can establish the following relationship

$$\mathbf{R}_s[1] = \mathbf{J}\mathbf{R}_s[0] + \mathbf{e}_1\mathbf{r}_1^H \tag{4.10}$$

$$\mathbf{R}_s[-1] = \mathbf{J}^T\mathbf{R}_s[0] + \mathbf{e}_d\mathbf{r}_2^H \tag{4.11}$$

where $\mathbf{J}$ ($\mathbf{J}^T$) stands for the one-lag down (up) shift square matrix whose first sub-diagonal entries below (above) the main diagonal are unity, whereas all remaining entries are zero; $\mathbf{e}_i$ denotes the unit column vector with its ith entry

equal to one, and its other entries equal to zero; and we have

$$\mathbf{r}_1^H \triangleq \mathbf{e}_1^H \mathbf{R}_s[1] = E[s(n)\vec{\mathbf{s}}^H(n-1)] \tag{4.12}$$

$$\mathbf{r}_2^H \triangleq \mathbf{e}_d^H \mathbf{R}_s[-1] = E[s(n-d+1)\vec{\mathbf{s}}^H(n+1)]. \tag{4.13}$$

In addition, if we define $r_{1,i}$ and $r_{2,i}$ as the ith entries of the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively, then the entries in these vectors are related as follows

$$r_{1,i} = r^*_{2,d+1-i} \qquad \forall\ i \in \{1, \ldots, d\}. \tag{4.14}$$

Using Eqn.(4.10–4.11), we can express $\Theta_i, i = 1, 2$ as follows

$$(a)\ \ \Theta_1 \triangleq \mathbf{J} - \mathbf{e}_1 \vec{\alpha}_1^H \quad (b)\ \ \Theta_2 \triangleq \mathbf{J}^T - \mathbf{e}_d \vec{\alpha}_2^H \tag{4.15}$$

where

$$\vec{\alpha}_1 = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,d} \end{bmatrix}^T = -\mathbf{R}_s^{-1}[0]\mathbf{r}_1 \tag{4.16}$$

$$\vec{\alpha}_2 = \begin{bmatrix} \alpha_{2,1} & \cdots & \alpha_{2,d} \end{bmatrix}^T = -\mathbf{R}_s^{-1}[0]\mathbf{r}_2. \tag{4.17}$$

It is clear that the entries in $\vec{\alpha}_1$ are exactly the coefficients of the dth-order optimum forward prediction error filter for the process $\{s(n)\}$ and the entries in $\vec{\alpha}_2$ are exactly the coefficients of the dth-order optimum backward prediction error filter for the process $\{s(n)\}$ [114]. Moreover, the relationship of $\vec{\alpha}_1$ and $\vec{\alpha}_2$ can be formalized in the following lemma.

Lemma 4.2 *Given that $\vec{\alpha}_1$ and $\vec{\alpha}_2$ are defined in Eqn.(4.16) and Eqn.(4.17)*

respectively, there holds

$$\alpha_{1,i} = \alpha^*_{2,d+1-i} \ \ \forall \ \ i \in \{1, \ldots, d\}. \tag{4.18}$$

Proof: We rewrite Eqn.(4.16) as $\mathbf{R}_s[0]\vec{\alpha}_1 = -\mathbf{r}_1$. Let c_i denote $E[s(n)s^*(n-i)]$, thus we have

$$\sum_{k=1}^{d} c_{k-i}\alpha_{1,k} = -r_{1,i} \ \ \forall \ \ i \in \{1, \ldots, d\}. \tag{4.19}$$

Taking conjugate operation on both sides of the above equation, we have $\sum_{k=1}^{d} c_{i-k}\alpha^*_{1,k} = -r^*_{1,i} \ \ \forall \ \ i \in \{1, \ldots, d\}$. If we let $k = d+1-\hat{k}$ and $i = d+1-\hat{i}$, it is clear that

$$\sum_{\hat{k}=1}^{d} c_{\hat{k}-\hat{i}}\alpha^*_{1,d+1-\hat{k}} = -r^*_{1,d+1-\hat{i}} \ \ \forall \ \ i \in \{1, \ldots, d\}. \tag{4.20}$$

Hence we have $\mathbf{R}_s[0][\alpha^*_{1,d} \ \cdots \ \alpha^*_{1,1}]^T = -\mathbf{r}_2$. We can conclude that $\vec{\alpha}_2 = [\alpha^*_{1,d} \ \cdots \ \alpha^*_{1,1}]^T$. The proof is completed. ■

Observe that both $\Theta_1 = \mathbf{J} - \mathbf{e}_1\vec{\alpha}_1^H$ and $\Theta_2 = \mathbf{J}^T - \mathbf{e}_d\vec{\alpha}_2^H$ are companion matrices that have some important properties to be investigated and exploited. We highlight one of the exploited properties as follows

Lemma 4.3 *If matrix* $\mathbf{Y}$ *commutes with* Θ_1 *and* Θ_2 *respectively, i.e.*

$$(a) \ \ \Theta_1\mathbf{Y} = \mathbf{Y}\Theta_1 \quad (b) \ \ \Theta_2\mathbf{Y} = \mathbf{Y}\Theta_2 \tag{4.21}$$

where $\mathbf{Y} \in \mathbb{C}^{d\times d}$ *and the modulus of* $\alpha_{1,d}$ *in* Θ_1 *is not equal to one, i.e.* $|\alpha_{1,d}| \neq 1$*, then* $\mathbf{Y} = \lambda\mathbf{I}$*, where* λ *could be any complex scalar including zero.*

Proof: See Appendix A. ■

We now prove the uniqueness of the system solution to Eqn.(4.8) by using the above lemma. Notice that Eqn.(4.8) and Eqn.(4.7) can be derived from each other. Therefore we only need to prove that the solution of Eqn.(4.7) is unique (up to a scalar factor). Thus the problem can be formulated as follows: Given that the following two equations hold

$$(a)\ \ \Upsilon_1 = \mathcal{H}\Theta_1\mathcal{H}^\dagger \quad (b)\ \ \Upsilon_2 = \mathcal{H}\Theta_2\mathcal{H}^\dagger \tag{4.22}$$

and $\mathcal{H}$ is full column rank, we need to prove that $\mathcal{H}$ can be uniquely determined up to a complex scalar by the following two equations

$$(a)\ \ \Upsilon_1\mathcal{H} = \mathcal{H}\Theta_1 \quad (b)\ \ \Upsilon_2\mathcal{H} = \mathcal{H}\Theta_2 \quad . \tag{4.23}$$

It implies that, if any non-zero matrix $\mathcal{G}$ which has the same structure as $\mathcal{H}$ also satisfies Eqn.(4.23a–b), then $\mathcal{G} = \lambda\mathcal{H}$, where λ is a non-zero complex scalar.

Proof: Suppose a non-zero matrix $\mathcal{G}$ which has the same Toeplitz structure as $\mathcal{H}$ also satisfies Eqn.(4.23a–b), then we have

$$\Upsilon_1\mathcal{G} = \mathcal{G}\Theta_1 \Rightarrow \mathcal{H}\Theta_1\mathcal{H}^\dagger\mathcal{G} = \mathcal{G}\Theta_1 \Rightarrow \Theta_1\mathcal{H}^\dagger\mathcal{G} = \mathcal{H}^\dagger\mathcal{G}\Theta_1 \tag{4.24}$$

$$\Upsilon_2\mathcal{G} = \mathcal{G}\Theta_2 \Rightarrow \mathcal{H}\Theta_2\mathcal{H}^\dagger\mathcal{G} = \mathcal{G}\Theta_2 \Rightarrow \Theta_2\mathcal{H}^\dagger\mathcal{G} = \mathcal{H}^\dagger\mathcal{G}\Theta_2. \tag{4.25}$$

By invoking Lemma 4.3, we know that $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$. Therefore we only need to prove that the solution of $\mathcal{G}$ that satisfies $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$ is unique and $\mathcal{G} = \lambda\mathcal{H}$. Note that $\mathcal{G}$ has the same block Toeplitz structure as $\mathcal{H}$. If we write $\mathcal{H}^\dagger \triangleq$

$[\mathbf{V}_0 \ \cdots \ \mathbf{V}_N]$, we can rewrite $\mathcal{H}^\dagger\mathcal{G} = \lambda\mathbf{I}$ as

$$\mathcal{V}\begin{bmatrix} \mathbf{g}(0) \\ \vdots \\ \mathbf{g}(L) \end{bmatrix} = \mathrm{vec}(\lambda\mathbf{I}) \tag{4.26}$$

where $\mathbf{g}(0), \cdots, \mathbf{g}(L)$ are the corresponding column vectors used to construct the block Toeplize matrix $\mathcal{G}$ in the way as we define $\mathcal{H}$ using $\mathbf{h}(0), \cdots, \mathbf{h}(L)$, $\mathcal{V} \in \mathbb{C}^{d^2\times(L+1)q}$ is a block Toeplitz matrix written as

$$\mathcal{V} \triangleq \begin{bmatrix} \mathbf{V}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \mathbf{V}_0 & \ddots & \vdots \\ \mathbf{V}_N & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_N & \ddots & \mathbf{V}_0 \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_N \end{bmatrix}.$$

Obviously, from Eqn.(4.26) we know that $\mathcal{G}$ can be uniquely determined if $\mathcal{V}$ has full column rank. Recalling Theorem 1 in [95], $\mathcal{V}$ has full column rank if there exists a non-zero z_0 (including ∞) such that the polynomial matrix $\mathbf{V}(z_0)$ has full column rank, where $\mathbf{V}(z) \triangleq \mathbf{V}_0 + \mathbf{V}_1 z^{-1} + \cdots + \mathbf{V}_N z^{-N}$. This mild condition can be assured with probability one since generally, when $L \geq 1$, the entries of matrix $\mathcal{H}^\dagger$ can be considered as randomly generated. Thus we can conclude that the solution of $\mathcal{G}$ is unique and $\mathcal{G} = \lambda\mathcal{H}$. Note that λ cannot be zero because $\mathcal{G}$ would be zero under the condition $\lambda = 0$, which contradicts our previously made assumption $\mathcal{G} \neq \mathbf{0}$. The proof is completed here. ■

4.4 Simulation Results

We now present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the other two methods, namely, the *subspace* (SS) method proposed in [22] and the so-called *linear prediction* (LP) approach presented in [67]. In our simulations, as an approximation of a two-ray multipath environment, the channel impulse response is obtained from the two delayed raised cosine pulses with its coefficients given by

$$[\mathbf{h}(0) \ \cdots \ \mathbf{h}(3)] = \begin{bmatrix} -0.1470 & 0.4461 & 0.1126 & -0.2233 \\ 0.0213 & 0.5356 & -0.2911 & 0.0660 \end{bmatrix}.$$

The colored source is induced in the same way as the simulation example in [67]. The channel order is assumed known *a priori* and the stack number (smoothed factor) N is chosen to be 3. For our proposed method, we only use the autocorrelation matrices $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm 1]$, i.e. $K = 1$ in criterion Eqn.(4.9). Once the channel has been estimated, the MMSE equalizers can be computed. The equalizer with equalization delay, d_e, equal to 3 is used in our simulations. We present the equalization performance of the respective algorithms in Table 4.1. The results are averaged over 500 Monte Carlo runs. In the first part of Table 4.1, we show the SER as a function of SNR with the number of samples used to estimate signal statistics $T_s = 400$. Next, in the latter part of Table 4.1, the SER is shown to be a function of T_s for SNR = 10dB. From the following table, we can see that the three algorithms perform similarly with the performance of LP slightly better than that of the other two algorithms. And our proposed method seems to lie somewhere between LP and SS.

SNR(dB)	Proposed Method	LP	SS
20	0.0000	0.0000	0.0000
17.5	0.0000	0.0000	0.0000
15	0.0006	0.0006	0.0023
12.5	0.0268	0.0101	0.0366
10	0.1368	0.0744	0.1417
T_s	Proposed Method	LP	SS
2000	0.0087	0.0055	0.0121
1600	0.0099	0.0077	0.0186
1200	0.0224	0.0129	0.0292
800	0.0598	0.0277	0.0605
400	0.1368	0.0744	0.1417

Table 4.1: SER versus SNR and number of data samples respectively

Our proposed method estimates the channel matrix $\mathcal{H}$ by matching $\mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger$ and $\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]$ for $k \in \pm 1$. The accuracy of our estimated channel is subject to the estimation errors of $\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]$. This accounts for the lack of performance improvement of our proposed algorithm as compared to LP. Despite of the slightly degraded performance, our algorithm shows an advantage over [22] and [67] since its extension to MIMO systems is straightforward. For the multiuser scenarios, Eqn.(4.7) still holds and under the assumption that all sources are uncorrelated with each other, we can further decompose Eqn.(4.7) into $\Upsilon_i\mathcal{H}_l = \mathcal{H}_l\Theta_{i,l}$, where $\mathcal{H}_l$ denotes the channel convolution matrix corresponding to the lth user, $\Theta_{i,l} \triangleq \mathbf{R}_{s_l}[\bar{k}]\mathbf{R}_{s_l}^{-1}[0]$, s_l represents the lth source, $\bar{k} = (i+1)/2$ if i is odd and $\bar{k} = -i/2$ if i is even. Each user's channel convolution matrix, $\mathcal{H}_l$, can then be identified according to the described algorithm. However, the proof for the uniqueness of Eqn.(4.7) has to further exploit additional properties on companion matrices, and also to impose a spectral diversity identifiability condition on the input colored sources (this will be discussed in the next chapter).

4.5 Summary

In this chapter, we presented a new SOS-based method that admits a closed-form solution for blind estimation of SIMO FIR channel driven by colored source signals. The uniqueness of the closed-form solution was proved by exploiting the inherent structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ and the derived property of one pair of companion matrices. The proposed method is valid under very mild condition on the source correlation. In fact, our method still works even if the source signals are white (this can be easily proved by following the procedure in this chapter). Simulation results showed that our proposed algorithm achieves a better performance than the classical subspace method [22]. Also, unlike other methods [22, 66, 67] which have difficulties in extending to the multiuser scenarios, our proposed algorithm has the potential to extend to the MIMO systems, which will be shown in the next chapter.

Chapter 5

Blind Identification of MIMO FIR Channel

In this chapter, we extend the proposed method in Chapter 4 to blind estimation of MIMO FIR channel driven by colored sources. We assume that the SOS of the input sources are known *a priori*. By further exploiting the properties of the companion matrices, we provide an original proof for the uniqueness of the system solution, which serves as a theoretical basis for our new method that admits a closed-form solution. The corresponding identifiability conditions and the computational complexity of the proposed method are discussed and compared to other existing method. Numerical simulation results are presented to illustrate the performance of the proposed algorithm.

5.1 Introduction

Blind identification of MIMO FIR channel arises in a wide variety of communication and signal processing applications, which include speech enhancement,

wireless mobile communications and brain signal analysis. In particular, potential use of blind MIMO channel identification in wireless systems is of strong interest. Consequently, the study on blind channel identification of MIMO systems has attracted increasing attention. Thus far, there have been a lot of research works [17, 23, 24, 35] on blind channel identification driven by white input signals. In this case, it is well known that the MIMO channel can only be determined up to an unknown unitary matrix that cannot otherwise be resolved using the second-order statistics. To resolve this residual static mixtures, one of several blind source separation (BSS) techniques should be resorted to. However, in contrast to the channels driven by white signals, the MIMO FIR channels driven by colored signals may provide us advantages in developing a complete closed-form SOS-based method without an extra BSS algorithmic step. It is also noted that the input colored signals should be of distinct power spectra, which is a sufficient condition for the MIMO FIR channel to be identifiable up to a scaling and permutation using second-order statistics of the channel output.

In this chapter, we consider the problem of blind estimation of MIMO FIR channel driven by colored signals. Of specific interest, we focus on the case where the second order statistics of the input signals are known *a priori*. As mentioned in previous chapter, colored sources with known statistics may indeed occur in practice. For example, colored sources may arise as a result of channel encoding [69], and the knowledge of the encoding scheme alone provides the required source statistics at the receiver. Moreover, correlative filters can be utilized at the transmitters to induce distinct spectral patterns to the source signals [72]. Interestingly, there have been some works on blind SIMO FIR channel identification [22] and blind MIMO FIR channel identification [28,36,37,70,71] dealing

with the input signals that are colored but with unknown statistics. However, failing to utilize the information of input signals statistics affects the estimator's performance to some extent. The previously developed works that address the case of input colored signals with *a priori* known statistics include [66, 67] for the single-input multiple-output (SIMO) systems and [72, 73] for the MIMO systems. Among them, the work [66] imposes a somewhat restrictive condition on the source correlation, where an exponentially decaying autocorrelation function is assumed. The work [67] constitutes a direct extension of the TXK method [17] by exploiting the inherent structural relationship between the source autocorrelation matrices $\mathbf{R}_s[0]$ and $\mathbf{R}_s[1]$. Both [66] and [67] consider the blind channel estimation/equalization of SIMO models, and the extension of these algorithms to the MIMO systems is not straightforward because when extended to the multiuser's case, some of the relationships and properties in these works are no longer valid. The work [72] provides an elegant closed-form solution to blind MIMO channel identification. It is presented under a correlative framework which is obtained by utilizing linear correlative filters at the transmitters, thus assigning distinct spectral patterns to the sources. In comparison with [72], our work in this chapter addresses a more general case, i.e. the colored sources need not necessarily be generated from linear correlative filters, and they can be induced by other nonlinear methods such as channel encoding. A more detailed comparison of our work to [72] is discussed in later part of this chapter. Recently, a frequency-domain nonlinear iterative method [73] was proposed for blind MIMO channel estimation driven by colored sources. Due to its nonlinear nature, the method requires a good initialization in order to minimize the problem of local minima. [101] is another important work on blind identification of channel driven by sources with *a priori* known statistics. Although the paper [101] considers the blind equalization of SIMO nonlinear

channels, some of the results in [101] can be reformulated into the MIMO setting (the nonlinear functions of the signal of interest can be seen as the additional inputs or sources) and induce a different channel identifiability condition from that of [72] and our work.

In this chapter, we extend the method in Chapter 4 and propose a closed-form solution for blind MIMO FIR system identification by utilizing the estimated channel output autocorrelation matrices and the knowledge of the source autocorrelation matrices. The properties of companion matrices are further exploited to prove the uniqueness of the system solution. The contribution of this chapter consists of the following three aspects. First, as in Chapter 4, the inherent structural relationship between the source autocorrelation matrices is further exploited as compared to the work [5]. Second, we derive some new properties of the constructed companion matrices. These properties play a key role in devising and validating our proposed algorithm. Third, our proposed algorithm compares favorably with other existing methods in many aspects.

This chapter is organized as follows. In Section 5.2, we introduce the MIMO system model and some basic assumptions. Next, in Section 5.3, we present our blind channel identification method and provide an original proof for the uniqueness of the system solution. We compare our method to other existing method in Section 5.4. Computational complexity and identifiability conditions are mainly considered. Finally, in Section 5.5, numerical simulation results are presented to demonstrate the performance of the proposed algorithm.

5.2 System Model and Basic Assumptions

Consider a noisy linear MIMO channel with p inputs, $s_i(n), i \in \{1, 2, \cdots, p\}$, and q outputs $\mathbf{x}(n) \triangleq [x_1(n) \ \cdots \ x_q(n)]$. The MIMO channel model can be written as follows (also see Eqn.(2.12) in Chapter 2)

$$\mathbf{x}(n) = \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n) \tag{5.1}$$

where $\{\mathbf{h}_i(l)\}$ denotes the multichannel filter corresponding to the ith user, L_i represents the channel order corresponding to the ith user. Let $\mathbf{h}_i(z)$ denote the Z-transform of $\{\mathbf{h}_i(l)\}$.

As discussed in Chapter 2, this channel model can be written in the following matrix form by stacking the channel output vector $\mathbf{x}(n)$ and defining $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^T(n) \ \mathbf{x}^T(n-1) \ \ldots \ \mathbf{x}^T(n-N)]^T$, $\vec{\mathbf{s}}_i(n) \triangleq [s_i(n) \ s_i(n-1) \ \cdots \ s_i(n-N-L_i)]^T$ and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^T(n) \ \mathbf{w}^T(n-1) \ \ldots \ \mathbf{w}^T(n-N)]^T$:

$$\vec{\mathbf{x}}(n) = \sum_{i=1}^{p} \mathcal{H}_i \vec{\mathbf{s}}_i(n) + \vec{\mathbf{w}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n). \tag{5.2}$$

The readers can refer to Chapter 2 for the details and the definitions of the symbols $\mathcal{H}_i$, $\mathcal{H}$ and $\vec{\mathbf{s}}(n)$.

We adopt the following basic assumptions:

A1 The number of sources is known *a priori*, and there are more outputs than inputs, i.e. $q > p$.

A2 Channel is irreducible and column-reduced (Please refer to Chapter 2 for the definitions "irreducible" and "column-reduced").

A3 The channel order of each source is assumed to be known *a priori*.

A4 The sources are zero-mean wide-sense stationary colored signals with their input statistics being available. The sources are uncorrelated with each other.

A5 Additive noises are spatially and temporally white noises, and they are statistically independent of the sources.

As a consequence of A1–A2, the MIMO channel matrix $\mathcal{H}$ is full column rank if the stack number N is chosen to satisfy $(N+1) \geq \sum_{i=1}^{p} L_i$ (see [27]). Our objective is to estimate the channel impulse response by utilizing the second-order statistics of the observed data $\vec{\mathbf{x}}(n)$ and the knowledge of the sources' statistics.

5.3 Proposed Channel Identification Method

We begin by defining the source autocorrelation matrices as follows

$$\mathbf{R}_{s_i}[k] \triangleq E[\vec{\mathbf{s}}_i(n)\vec{\mathbf{s}}_i^H(n-k)] \tag{5.3}$$

$$\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n-k)]. \tag{5.4}$$

By invoking the assumption A4, we know that $\mathbf{R}_s[k]$ is a block diagonal matrix written as

$$\mathbf{R}_s[k] = \text{diag}\left(\mathbf{R}_{s_1}[k], \mathbf{R}_{s_2}[k], \cdots, \mathbf{R}_{s_p}[k]\right) \tag{5.5}$$

where diag$(\cdot)$ denotes block diagonal. Also, in order to simplify the presentation of the proposed channel identification method, we assume the noiseless case.

Thus the autocorrelation matrices of the received data $\vec{\mathbf{x}}(n)$ can be written as

$$\mathbf{R}_x[k] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^H(n-k)] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H. \tag{5.6}$$

In the following, we will show that the channel convolution matrix $\mathcal{H}$ can be identified up to a block diagonal matrix $\mathcal{D} \triangleq \mathrm{diag}(\lambda_1\mathbf{I}_1, \cdots, \lambda_p\mathbf{I}_p)$ by utilizing the estimated channel output autocorrelation matrices $\mathbf{R}_x[k], k \in \{0, \pm 1\}$ and the knowledge of $\mathbf{R}_s[k], k \in \{0, \pm 1\}$, where $\mathbf{I}_i$ denotes an $d_i \times d_i$ identity matrix. We commence by introducing the following lemma.

Lemma 5.1 *Given* $\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H$, $\mathcal{H}$ *is full column rank and* $\mathbf{R}_s[0]$ *is invertible, then we have*

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger \tag{5.7}$$

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]\mathcal{H} = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]. \tag{5.8}$$

Proof: The proof follows the same way as that of Lemma 4.1. ■

For convenience, let

$$\Upsilon_{2k-1} \triangleq \mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] \qquad \Upsilon_{2k} \triangleq \mathbf{R}_x[-k]\mathbf{R}_x^\dagger[0]$$
$$\Theta_{2k-1} \triangleq \mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] \quad \Theta_{2k} \triangleq \mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0].$$

We can therefore re-express Eqn.(5.8) (choose $K \geq k \geq 1$) as

$$\Upsilon_{\bar{k}}\mathcal{H} = \mathcal{H}\Theta_{\bar{k}} \qquad \forall\ \bar{k} \in \{1, \ldots, 2K\} \tag{5.9}$$

and further, for every $\bar{k} \in \{1, \ldots, 2K\}$, we have the following by exploiting the block diagonal structure of $\Theta_{\bar{k}} \triangleq \text{diag}(\Theta_{\bar{k},1}, \Theta_{\bar{k},2}, \cdots, \Theta_{\bar{k},p})$

$$\Upsilon_{\bar{k}} \mathcal{H}_i = \mathcal{H}_i \Theta_{\bar{k},i} \qquad \forall \ i \in \{1, \ldots, p\} \tag{5.10}$$

where $\Theta_{\bar{k},i} \triangleq \mathbf{R}_{s_i}[k]\mathbf{R}_{s_i}^{-1}[0]$, $k = (\bar{k}+1)/2$ if $\bar{k}$ is odd and $k = -\bar{k}/2$ if $\bar{k}$ is even. For each $i \in \{1, \ldots, p\}$, the above equation can be used to identify the channel convolution matrix of user i, i.e. $\mathcal{H}_i$, since the knowledge of $\Theta_{\bar{k},i}$ is known *a priori* and the information of $\Upsilon_{\bar{k}}$ can be obtained from the second-order statistics of the observed data. By exploiting the block Toeplitz structure of $\mathcal{H}_i$, we can rewrite Eqn.(5.10) as

$$\mathcal{T}_1[\Upsilon_{\bar{k}}]\mathbf{h}_i = \mathcal{T}_2[\Theta_{\bar{k},i}]\mathbf{h}_i \tag{5.11}$$

where $\mathbf{h}_i \triangleq \left[\mathbf{h}_i^T(0) \ \ \ldots \ \ \mathbf{h}_i^T(L_i)\right]^T$, $\mathcal{T}_1[\cdot]$ and $\mathcal{T}_2[\cdot]$ respectively represent a certain transformation on the bracketed matrix. Therefore we may estimate $\mathbf{h}_i$ by the following criterion

$$\hat{\mathbf{h}}_i = \arg \min_{\|\mathbf{h}_i\|=1} \sum_{\bar{k}=1}^{2K} \left\| \left[\ \mathcal{T}_1[\Upsilon_{\bar{k}}] - \mathcal{T}_2[\Theta_{\bar{k},i}] \ \right] \mathbf{h}_i \right\|^2 . \tag{5.12}$$

The above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. However, this criterion fails to provide the true channel estimation if the solution to Eqn.(5.11) is not unique, i.e. there exist other non-zero vectors, $\mathbf{g}_i$, that are *linearly independent of* $\mathbf{h}_i$ and also satisfy $\mathcal{T}_1[\Upsilon_{\bar{k}}]\mathbf{g}_i = \mathcal{T}_2[\Theta_{\bar{k},i}]\mathbf{g}_i$ for any $\bar{k} \in \{1, \ldots, 2K\}$. Hence we are faced with the following problem, that is, whether or not the solution to Eqn.(5.11) is unique (up to an unknown scalar factor) and under what conditions the solution of Eqn.(5.11) will be unique. This

problem is studied in the following and we will establish the uniqueness of the solution to Eqn.(5.11) by utilizing only $\mathbf{R}_x[0]$ and $\mathbf{R}_x[\pm 1]$ and the knowledge of $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$, i.e. the uniqueness of the solution can be guaranteed by choosing $\bar{k} = 1, 2$ in Eqn.(5.11). We begin by exploiting the inherent structural relationship between $\mathbf{R}_{s_i}[0]$ and $\mathbf{R}_{s_i}[\pm 1]$ for any $i \in \{1, \ldots, p\}$.

5.3.1 Inherent Structural Relationship of Source Autocorrelation Matrices

It can be readily seen that for each source s_i, the last $d_i - 1$ rows of $\mathbf{R}_{s_i}[1]$ are the first $d_i - 1$ rows of $\mathbf{R}_{s_i}[0]$, and the first $d_i - 1$ rows of $\mathbf{R}_{s_i}[-1]$ are the last $d_i - 1$ rows of $\mathbf{R}_{s_i}[0]$. Hence we can establish the following relationship

$$\mathbf{R}_{s_i}[1] = \mathbf{J}\mathbf{R}_{s_i}[0] + \mathbf{e}_1 \mathbf{r}_{i1}^H \tag{5.13}$$

$$\mathbf{R}_{s_i}[-1] = \mathbf{J}^T \mathbf{R}_{s_i}[0] + \mathbf{e}_{d_i} \mathbf{r}_{i2}^H \tag{5.14}$$

where

$$\mathbf{r}_{i1}^H \triangleq \mathbf{e}_1^H \mathbf{R}_{s_i}[1] = E[s_i(n)\vec{\mathbf{s}}_i^H(n-1)] \tag{5.15}$$

$$\mathbf{r}_{i2}^H \triangleq \mathbf{e}_{d_i}^H \mathbf{R}_{s_i}[-1] = E[s_i(n-d_i+1)\vec{\mathbf{s}}_i^H(n+1)]. \tag{5.16}$$

In addition, if we define $\mathbf{r}_{i1}(k)$ and $\mathbf{r}_{i2}(k)$ as the kth entries of the vectors $\mathbf{r}_{i1}$ and $\mathbf{r}_{i2}$, respectively, then the entries in these vectors are related as follows

$$\mathbf{r}_{i1}(k) = \mathbf{r}_{i2}^*(d_i + 1 - k) \qquad \forall\ k \in \{1, \ldots, d_i\}. \tag{5.17}$$

Using Eqn.(5.13–5.14), we can re-express $\Theta_{1,i}$ and $\Theta_{2,i}$ as follows

$$\Theta_{1,i} = \mathbf{R}_{s_i}[1]\mathbf{R}_{s_i}^{-1}[0] = \mathbf{J} - \mathbf{e}_1\vec{\alpha}_i^H \tag{5.18}$$

$$\Theta_{2,i} = \mathbf{R}_{s_i}[-1]\mathbf{R}_{s_i}^{-1}[0] = \mathbf{J}^T - \mathbf{e}_{d_i}\vec{\beta}_i^H \tag{5.19}$$

where $\Theta_{1,i}, \Theta_{2,i} \in \mathbb{C}^{d_i \times d_i}$ and $\vec{\alpha}_i$ and $\vec{\beta}_i$ can be obtained as

$$\vec{\alpha}_i = \begin{bmatrix} \alpha_{i,1} & \cdots & \alpha_{i,d_i} \end{bmatrix}^T \triangleq -\mathbf{R}_{s_i}^{-1}[0]\mathbf{r}_{i1} \tag{5.20}$$

$$\vec{\beta}_i = \begin{bmatrix} \beta_{i,1} & \cdots & \beta_{i,d_i} \end{bmatrix}^T \triangleq -\mathbf{R}_{s_i}^{-1}[0]\mathbf{r}_{i2}. \tag{5.21}$$

It can be seen that the entries in $\vec{\alpha}_i$ are exactly the coefficients of the d_i^{th}-order optimum forward prediction error filter for the process $\{s_i(n)\}$ and the entries in $\vec{\beta}_i$ are exactly the coefficients of the d_i^{th}-order optimum backward prediction error filter for the process $\{s_i(n)\}$ [114]. Also it is well known that the relationship between $\vec{\alpha}_i$ and $\vec{\beta}_i$ can be formulated in the following lemma.

Lemma 5.2 *Given that $\vec{\alpha}_i$ and $\vec{\beta}_i$ are defined as in Eqn.(5.20) and Eqn.(5.21) respectively, there holds*

$$\alpha_{i,k} = \beta_{i,d_i+1-k}^* \ \forall \ k \in \{1, \ldots, d_i\}. \tag{5.22}$$

Proof: The proof follows the same way as that of Lemma 4.2. ■

Observe that both $\Theta_{1,i}$ and $\Theta_{2,i}$ are companion matrices and they are related by the following relationship $\Theta_{2,i} = \mathbf{M}\Theta_{1,i}^*\mathbf{M}$, where $\mathbf{M}$ represents the exchange matrix with ones on the antidiagonal and zeros elsewhere. Due to their special

structures, these companion matrices have some important properties we shall investigate in the following.

5.3.2 Properties of Companion Matrices and The Identifiability Conditions

The properties of the companion matrices are highlighted as follows

Lemma 5.3 *Given that* $\mathbf{Y} \in \mathbb{C}^{d_i \times d_j}$ *satisfies the following two equations*

$$(a)\ \ \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j} \quad (b)\ \ \Theta_{2,i}\mathbf{Y} = \mathbf{Y}\Theta_{2,j} \tag{5.23}$$

and the modulus of the last entry in $\vec{\alpha}_j$ *is not equal to one, i.e.* $|\alpha_{j,d_j}| \neq 1$*, we have*

- *If* $d_i = d_j$*,* $\Theta_{1,i} = \Theta_{1,j}$ *and* $\Theta_{2,i} = \Theta_{2,j}$*, then* $\mathbf{Y} = \lambda\mathbf{I}$*, where* λ *could be any complex scalar including zero.*
- *If* $d_i = d_j$*,* $\Theta_{1,i} \neq \Theta_{1,j}$ *and* $\Theta_{2,i} \neq \Theta_{2,j}$*, then* $\mathbf{Y} = \mathbf{0}$*.*
- *If* $d_i > d_j$*, then* $\mathbf{Y} = \mathbf{0}$*.*
- *If* $d_i < d_j$*, and* $|\alpha_{i,m_i}| \neq |\alpha_{j,d_j-t_i}|$*, where* $t_i \triangleq d_i - m_i$*,* α_{i,m_i} *denotes the last non-zero entry in* $\vec{\alpha}_i$*, then* $\mathbf{Y} = \mathbf{0}$*. Such a condition* $|\alpha_{i,m_i}| \neq |\alpha_{j,d_j-t_i}|$ *can be removed if there exists a non-zero entry for* $\alpha_{j,k}, k \in \{d_j - t_i + 1, \ldots, d_j\}$*.*

Proof: See Appendix B. ■

The significance of Lemma 5.3 not only lies in the fact that it provides a theoretical basis for Theorem 5.1, but also it establishes the identifiability conditions imposed on the input colored sources. We describe these identifiability conditions as follows

IC1 The modulus of the last entry in each $\vec{\alpha}_i$ is not equal to one, i.e.

$$|\alpha_{i,d_i}| \neq 1 \qquad \forall\ i \in \{1, \ldots, p\}. \tag{5.24}$$

This condition can be guaranteed if for every user $s_i, i \in \{1, \ldots, p\}$, the source autocorrelation matrix $\mathbf{R}_{s_i}[0]$ is positive definite. This is because $|\alpha_{i,d_i}|$ will be strictly less than one under the assumption that $\mathbf{R}_{s_i}[0]$ is positive definite (see Theorem 1 in [67]).

IC2 For each pair of sources $\{s_i, s_j\}$, we have

$$\Theta_{1,i} \neq \Theta_{1,j} \qquad \Theta_{2,i} \neq \Theta_{2,j} \tag{5.25}$$

i.e. $\vec{\alpha}_i \neq \vec{\alpha}_j$ and $\vec{\beta}_i \neq \vec{\beta}_j$. Because of the relationship between $\vec{\alpha}_i$ and $\vec{\beta}_i$, this condition can be reduced as $\vec{\alpha}_i \neq \vec{\alpha}_j$ for each pair of $\{i, j\}$. This condition is satisfied if all sources have distinct second order statistics (power spectra).

IC3 Sources have distinct second order statistics (power spectra) such that for each pair of sources $\{s_i, s_j\}$, where $d_j \geq d_i$, the corresponding $\{\vec{\alpha}_i, \vec{\alpha}_j\}$ *does not* satisfy the following two conditions simultaneously

(i) $|\alpha_{i,m_i}| = |\alpha_{j,d_j-t_i}|$

(ii) $\alpha_{j,k} = 0 \qquad \forall\ k \in \{d_j - t_i + 1, \ldots, d_j\}$

where α_{i,m_i} is the last non-zero entry in $\vec{\alpha}_i$. This condition is very mild such that it is less restrictive than the following more comprehensible condition: *for each pair of sources* $\{s_i, s_j\}$*, the moduli of the last non-*

zero entries of $\vec{\alpha}_i$ *and* $\vec{\alpha}_j$ *are not equal*, i.e.

$$|\alpha_{i,m_i}| \neq |\alpha_{j,m_j}| \ \ \forall \ \ i,j \in \{1,\ldots,p\} \tag{5.26}$$

where α_{j,m_j} denotes the last non-zero entry in $\vec{\alpha}_j$. Considering that the channel order of each source may possibly change in practice, it means that every pair of $\{\vec{\alpha}_i, \vec{\alpha}_j\}$ from $\{\vec{\alpha}_1, \ldots, \vec{\alpha}_p\}$ should satisfy the condition in Eqn.(5.26) for every possible $\{d_i, d_j\}$, where $d_i = N + L_i + 1$, N is the stack number chosen at the receiver.

Remark: First, we emphasize that IC1–IC3 are *sufficient* identifiability conditions in order to distinguish them from *necessary* conditions. Among the identifiability conditions, IC2 can be considered as a redundant condition since IC3 alone guarantees that all sources have distinct power spectra. IC3 is the so-called *spectral diversity condition* required to ensure that the MIMO channel $\mathcal{H}$ can be unambiguously determined from the second-order statistics of received data. Otherwise the channel can only be determined up to an unknown unitary matrix due to the spectral symmetry.

5.3.3 Proof of The Solution Uniqueness and The Proposed Algorithm

We now prove the uniqueness of the system solution to Eqn.(5.11) by utilizing the above lemma. We, firstly, prove that the solution to Eqn.(5.10) is unique (up to a scalar factor). The problem is formulated as the following theorem.

Theorem 5.1 *Given that (note that the following two equations are directly*

from Eqn.(5.7))

$$(a)\ \ \Upsilon_1 = \mathcal{H}\Theta_1\mathcal{H}^\dagger \quad (b)\ \ \Upsilon_2 = \mathcal{H}\Theta_2\mathcal{H}^\dagger \tag{5.27}$$

if $\mathcal{H}$ is full column rank and the input colored sources satisfy the identifiability conditions IC1–IC3, then any non-zero matrix $\mathcal{G}_i, i \in \{1, \ldots, p\}$, that has the same block Toeplitz structure as $\mathcal{H}_i$ and also satisfies Eqn.(5.10) for $\bar{k} = 1, 2$, i.e. $\Upsilon_1\mathcal{G}_i = \mathcal{G}_i\Theta_{1,i}$ and $\Upsilon_2\mathcal{G}_i = \mathcal{G}_i\Theta_{2,i}$, can be written as $\mathcal{G}_i = \lambda_i\mathcal{H}_i$, where λ_i is a non-zero complex scalar.

Proof: Suppose a non-zero matrix $\mathcal{G}_i \in \mathbb{C}^{(N+1)q\times d_i}$ with the same block Toeplitz structure as $\mathcal{H}_i$ also satisfies Eqn.(5.10) for $\bar{k} = 1, 2$, then we have

$$\Upsilon_1\mathcal{G}_i = \mathcal{G}_i\Theta_{1,i} \Rightarrow \mathcal{H}\Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{1,i} \Rightarrow \Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{1,i} \tag{5.28}$$

$$\Upsilon_2\mathcal{G}_i = \mathcal{G}_i\Theta_{2,i} \Rightarrow \mathcal{H}\Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{2,i} \Rightarrow \Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{2,i}. \tag{5.29}$$

Let $\mathbf{X} \triangleq \mathcal{H}^\dagger\mathcal{G}_i \triangleq \begin{bmatrix} \mathbf{X}_1^T & \cdots & \mathbf{X}_p^T \end{bmatrix}^T$, where $\mathbf{X}_k \in \mathbb{C}^{d_k\times d_i}$, then we have

$$\Theta_{1,k}\mathbf{X}_k = \mathbf{X}_k\Theta_{1,i} \ \ \forall\ \ k \in \{1, \ldots, p\} \tag{5.30}$$

$$\Theta_{2,k}\mathbf{X}_k = \mathbf{X}_k\Theta_{2,i} \ \ \forall\ \ k \in \{1, \ldots, p\}. \tag{5.31}$$

Since the input sources satisfy the identifiability conditions IC1–IC3, by applying the results in Lemma 5.3, we know that $\mathbf{X}_k = \mathbf{0}$ for any $k \neq i$ and $\mathbf{X}_k = \lambda_i\mathbf{I}_i$ for $k = i$, i.e.

$$\mathcal{H}^\dagger\mathcal{G}_i = \begin{bmatrix} \mathbf{0} & \cdots & \lambda_i\mathbf{I}_i & \cdots & \mathbf{0} \end{bmatrix}^T \triangleq \lambda_i\mathbf{E}_i. \tag{5.32}$$

Therefore we only need to prove that the solution of $\mathcal{G}_i$ that satisfies Eqn.(5.32) is unique and $\mathcal{G}_i = \lambda_i \mathcal{H}_i$. Notice that $\mathcal{G}_i$ has the same block Toeplitz structure as $\mathcal{H}_i$. If we write $\mathcal{H}^\dagger \triangleq [\mathbf{V}_0 \ \cdots \ \mathbf{V}_N]$, we can transform $\mathcal{H}^\dagger \mathcal{G}_i = \lambda_i \mathbf{E}_i$ as

$$\mathcal{V} \begin{bmatrix} \mathbf{g}_i(0) \\ \vdots \\ \mathbf{g}_i(L_i) \end{bmatrix} = \mathrm{vec}(\lambda_i \mathbf{E}_i) \tag{5.33}$$

where $\mathbf{g}_i(0), \cdots, \mathbf{g}_i(L_i)$ are the corresponding column vectors used to construct the block Toeplize matrix $\mathcal{G}_i$ in the way as we define $\mathcal{H}_i$ using $\mathbf{h}_i(0), \cdots, \mathbf{h}_i(L_i)$, $\mathcal{V} \in \mathbb{C}^{d_i(d_1+\cdots+d_p)\times(L_i+1)q}$ is a block Toeplitz matrix written as

$$\mathcal{V} = \begin{bmatrix} \mathbf{V}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \mathbf{V}_0 & \ddots & \vdots \\ \mathbf{V}_N & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_N & \ddots & \mathbf{V}_0 \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_N \end{bmatrix}.$$

Obviously, from Eqn.(5.33) we know that $\mathcal{G}_i$ can be uniquely determined if $\mathcal{V}$ has full column rank. Recalling Theorem 1 in [95], $\mathcal{V}$ has full column rank if the following condition holds, i.e. there exists a nonzero z_0 (including ∞) such that the polynomial matrix $\mathbf{V}(z_0)$ has full column rank, where $\mathbf{V}(z) \triangleq \mathbf{V}_0 + \mathbf{V}_1 z^{-1} + \cdots + \mathbf{V}_N z^{-N}$. This mild condition can be satisfied with probability one since generally, when $L \geq 1$, the entries of matrix $\mathcal{H}^\dagger$ can be considered as randomly generated. Thus we can conclude that the solution of $\mathcal{G}_i$ is unique and $\mathcal{G}_i = \lambda_i \mathcal{H}_i$. Note that λ_i can not be zero here because $\mathcal{G}_i$ would be zero under the condition $\lambda_i = 0$, which contradicts our previously made assumption

$\mathcal{G}_i \neq \mathbf{0}$. The proof is completed here. ■

Since Eqn.(5.10) and Eqn.(5.11) can be derived from each other, it implies that the solution to Eqn.(5.11) is unique up to a scaling constant of the "true" channel $\mathbf{h}_i$. Therefore $\mathbf{h}_i$ can be estimated by the criterion in Eqn.(5.12) with $K = 1$, i.e.

$$\hat{\mathbf{h}}_i = \arg \min_{\|\mathbf{h}_i\|=1} \left\| \begin{bmatrix} \mathcal{T}_1[\Upsilon_1] - \mathcal{T}_2[\Theta_{1,i}] \\ \mathcal{T}_1[\Upsilon_2] - \mathcal{T}_2[\Theta_{2,i}] \end{bmatrix} \mathbf{h}_i \right\|^2 . \tag{5.34}$$

As mentioned before, the above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. The matrix involved in singular value decomposition (SVD) operation is of dimension $\mathbb{C}^{2(N+L_i+1)(N+1)q \times (L_i+1)q}$. Now the channel convolution matrix $\mathcal{H}$ has been identified up to a block diagonal matrix $\mathcal{D} = \text{diag}(\lambda_1 \mathbf{I}_1, \cdots, \lambda_p \mathbf{I}_p)$, where λ_i for each $i \in \{1, \ldots, p\}$ is an unknown nonzero complex scalar. It is noted that the amplitude ambiguity of this unknown complex scalar can be removed if we insert the estimated channel $\hat{\mathcal{H}}$ back into Eqn.(5.6). Thus the channel convolution matrix $\mathcal{H}$ can be further identified up to the block diagonal matrix $\mathcal{D} = \text{diag}(e^{i\theta_1} \mathbf{I}_1, \cdots, e^{i\theta_p} \mathbf{I}_p)$.

5.3.4 Joint Order Detection and Channel Estimation

In this subsection, we consider the problem of joint order detection and channel estimation. Our previous work assumed that the channel order of each user is known *a priori* or can be correctly estimated. However, in practice, channel order determination of each user from second order statistics of received data may not be possible since only the total number of channel order, i.e. $L_{total} =$

$L_1 + \cdots + L_p$, can be estimated from $\mathbf{R}_x[0]$ by applying MDL criterion [51]. Therefore a joint order detection and channel estimation algorithm is desirable. We present the following theorem as a theoretical basis for our joint estimation algorithm.

Theorem 5.2 *Let $\mathcal{L} = \{[M_1, \ldots, M_p] \in \mathbb{N}^p\}$ denote the finite set of p-tuple points, where $M_1 + \cdots + M_p = L_{total}$, $\mathbb{N}^p$ denotes the set of p-tuple points with natural number entries. Then $\vec{l_e} = [L_1, \ldots, L_p]$ is the unique p-tuple point in $\mathcal{L}$ that can render us a non-zero solution $\{\mathcal{X}_1, \ldots, \mathcal{X}_p\}$ satisfying the following sets of equations, where for each $i \in \{1, \ldots, p\}$, $\mathcal{X}_i \in \mathbb{C}^{(N+1)q \times (N+M_i+1)}$ is a non-zero block Toeplitz matrix.*

$$\Upsilon_1 \mathcal{X}_i = \mathcal{X}_i \Theta_{1,i}(M_i) \ \ \forall \ i \in \{1, \ldots, p\} \tag{5.35}$$

$$\Upsilon_2 \mathcal{X}_i = \mathcal{X}_i \Theta_{2,i}(M_i) \ \ \forall \ i \in \{1, \ldots, p\} \tag{5.36}$$

where $\Theta_{1,i}(M_i)$ and $\Theta_{2,i}(M_i)$ are the corresponding companion matrices constructed using the estimated channel order M_i. It implies that for any other p-tuple point $\vec{l} = [M_1, \ldots, M_p]$ in $\mathcal{L}$, $\vec{l} \neq \vec{l_e}$, there does not exist a non-zero solution $\{\mathcal{X}_1, \ldots, \mathcal{X}_p\}$ that satisfies Eqn.(5.35–5.36).

Proof: See Appendix C. ■

From Theorem 5.1 and Theorem 5.2, it is easy to ascertain that $\varphi(\vec{l_e}; \frac{\mathbf{h}_1}{\|\mathbf{h}_1\|}, \ldots, \frac{\mathbf{h}_p}{\|\mathbf{h}_p\|})$ is the unique zero of the function φ subject to $\|\mathbf{u}_i\| = 1 \ \forall \ i \in \{1, \ldots, p\}$.

$$\varphi(\vec{l}; \mathbf{u}_1, \ldots, \mathbf{u}_p) = \sum_{i=1}^{p} \left\| \mathcal{T}(\vec{l}; i) \mathbf{u}_i \right\|^2 \tag{5.37}$$

where

$$\mathcal{T}(\vec{l};i) \triangleq \begin{bmatrix} \mathcal{T}_1[\Upsilon_1] - \mathcal{T}_2[\Theta_{1,i}(M_i)] \\ \mathcal{T}_1[\Upsilon_2] - \mathcal{T}_2[\Theta_{2,i}(M_i)] \end{bmatrix}. \tag{5.38}$$

Therefore we can jointly determine the channel order and estimate the channel parameters by the following criterion

$$\{\hat{\vec{l}}_e; \hat{\mathbf{h}}_1, \ldots, \hat{\mathbf{h}}_p\} = \arg \min_{\vec{l}\in\mathcal{L}, \|\mathbf{u}_i\|=1} \sum_{i=1}^{p} \left\| \mathcal{T}(\vec{l};i)\mathbf{u}_i \right\|^2. \tag{5.39}$$

For each $\vec{l} \in \mathcal{L}$, the above optimization admits a closed-form solution of $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ involving a series of right singular vectors associated with the smallest singular values. To search for the optimal p-tuple point $\hat{\vec{l}}_e$ that minimizes the above criterion, we need to let $\vec{l}$ run over $\mathcal{L}$.

5.3.5 Noise Compensation

In our previous presentation, we have ignored the noise effect in order to simplify our presentation. In practice, the influence of the noise can be minimized by removing the noise contribution from the estimated autocorrelation matrices of channel output. Since the additive noises are assumed spatially and temporally white with same variance, we have

$$\mathbf{R}_x[0] = \mathcal{H}\mathbf{R}_s[0]\mathcal{H}^H + \sigma_w^2\mathbf{I}.$$

The noise variance σ_w^2 can thus be estimated as the smallest eigenvalues of $\mathbf{R}_x[0]$ and then subtracted from any estimated autocorrelation matrix $\mathbf{R}_x[k]$ to provide our proposed algorithm with denoised autocorrelation estimates. In fact, even doing in this way, the noise effect cannot yet be completely canceled.

This results in estimation errors of $\Upsilon_{\bar{k}}$ and thus directly affects the accuracy of the estimated channel, which can be easily observed through Eqn.(5.34). A full study of the theoretical asymptotic performance analysis of our proposed method is still under investigation and thus not included in this chapter.

5.4 Discussions

In this section, we compare our method to another existing method [72] that addresses the same problem. The work in [72] also provides a closed-form solution to blind estimation of MIMO channel driven by colored sources. The method [72] and our proposed method may share a certain similarity in that both methods are developed by matching theoretical and observed time-domain second-order statistics of the observations. However, the work in [72] involves a two-step estimation algorithm. In the first step, the channel convolution matrix $\mathcal{H}$ is determined up to a unitary matrix $\mathbf{Q}$ from $\mathbf{R}_x[0]$, i.e.

$$\mathbf{R}_x[0] = \mathcal{H}\mathbf{R}_s[0]\mathcal{H}^H = (\mathcal{H}\mathbf{R}_s^{1/2}[0])(\mathcal{H}\mathbf{R}_s^{1/2}[0])^H \tag{5.40}$$

$$\mathbf{G}_0 = \mathcal{H}\mathbf{R}_s^{1/2}[0]\mathbf{Q}^H. \tag{5.41}$$

Next, the residual unknown unitary matrix $\mathbf{Q}$ is resolved and estimated as an isometry fitter which matches the observed second-order statistics $\mathbf{G}_0^\dagger\mathbf{R}_x[k](\mathbf{G}_0^\dagger)^H$ and the theoretical second-order statistics $\mathbf{Q}\mathbf{R}_s^{-1/2}[0]\mathbf{R}_s[k]\mathbf{R}_s^{-1/2}[0]\mathbf{Q}^H$. On the contrary, our proposed method directly estimates the channel convolution matrix $\mathcal{H}$ by matching the estimated second-order statistics $\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]$ and the theoretical second-order statistics $\mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger$. We now consider and dis-

cuss the computational complexity and the channel identifiability conditions of the respective methods.

5.4.1 Computational Complexity

One key advantage of our method is our ability in exploiting the block Toeplitz structure of $\mathcal{H}$, which results in great reduction in computational complexity. This can be illustrated as follows. We consider the linear algebraic operations involved in the respective algorithms. In [72], the unknown unitary matrix $\mathbf{Q} = [\mathbf{Q}_1 \ \cdots \ \mathbf{Q}_p]$ is computed in p parallel threads, the ith thread leading to $\mathbf{Q}_i$. $\mathbf{Q}_i$ is estimated as a closed-form minimizer that can be obtained by computing the singular value decomposition (SVD) of a matrix with the following dimension:

$$\begin{aligned} \text{rows}: &\quad \bar{K} \times ((N+1)p + \textstyle\sum_{i=1}^{p} L_i) \times (N+1+L_i) \\ \text{columns}: &\quad ((N+1)p + \textstyle\sum_{i=1}^{p} L_i) \times (N+1+L_i) \end{aligned}$$

where $\bar{K} = 2K$ is the number of autocorrelation matrices used for matching purposes. Comparatively, our method also estimates $\mathcal{H} = [\mathcal{H}_1 \ \cdots \ \mathcal{H}_p]$ in p parallel threads, the ith thread leading to $\mathcal{H}_i$. $\mathcal{H}_i$ is obtained by computing the SVD of a matrix with the following dimension:

$$\begin{aligned} \text{rows}: &\quad \bar{K} \times (N+1)q \times (N+L_i+1) \\ \text{columns}: &\quad (L_i+1)q. \end{aligned}$$

It can be seen that the dimension of the matrix involved in SVD operation in our method is much less than that of [72], except when $q \gg p$. As a simple illustrative example, if $p = 2$, $q = 3$, $(L_1, L_2) = (3, 3)$, $\bar{K} = 2$ for both methods and the stack number N is chosen to be 6 to ensure that $\mathcal{H}$ is full column rank, then, the matrix involved in SVD operation in [72] is of size 400×200 while

in our method it is of size 420×12. Based on [112], for an $m \times n$ matrix with $m \geq n$, it needs about $2mn^2 + 11n^3$ flops to compute singular values and right singular vectors. Thus the flops needed by [72] is about $10^3 \sim 20^3$ times the flops required by our proposed method.

5.4.2 Channel Identifiability Condition

We now consider the identifiability conditions induced by respective methods. As compared to [72], our work shows an improvement in that it relaxes the identifiability conditions further by specifying and minimizing the exact number of autocorrelation matrices required for channel estimation. For our method, it has been proved that only $\mathbf{R}_x[\pm 1]$ and $\mathbf{R}_x[0]$, i.e. $\bar{K} = 2$, suffice to provide us with a unique closed-form system solution. On the other hand, it can be seen that both methods have established respective spectral diversity conditions imposed on the input colored sources. In our work, the coefficients of the optimum forward prediction error filters for the input colored processes are used to characterize the spectral diversity, and the spectral diversity condition (IC3) requires that for each pair of sources $\{s_i, s_j\}$, the moduli of the last non-zero entries of $\vec{\alpha}_i$ and $\vec{\alpha}_j$ are not equal (Note that here we adopt the more restrictive but more comprehensible condition), i.e.

$$|\alpha_{i,m_i}| \neq |\alpha_{j,m_j}| \ \ \forall \ \ i, j \in \{1, \ldots, p\}. \tag{5.42}$$

In [72], the sources autocorrelation matrices' eigenvalues are used to characterize the spectral diversity, and the spectral diversity condition requires that for each pair of sources $\{s_i, s_j\}$, there is a correlation lag k such that

$$\sigma(\mathbf{A}_i(k)) \cap \sigma(\mathbf{A}_j(k)) = \varnothing \tag{5.43}$$

where $\mathbf{A}_i(k) \triangleq \mathbf{R}_{s_i}^{-1/2}[0]\mathbf{R}_{s_i}[k]\mathbf{R}_{s_i}^{-1/2}[0]$, $\sigma(\mathbf{A}) = \{\sigma_1, \ldots, \sigma_n\}$ denotes the set of eigenvalues of matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$. From probability perspective, if the correlation lag k is confined to $\{\pm 1\}$, it seems that our spectral diversity condition is more likely to be satisfied since our spectral diversity condition requires only one pair of random variables to be unequal for each pair of sources, while the spectral diversity condition of [72] requires that every pair of entries selected respectively from the two set of eigenvalues must be unequal.

5.5 Simulation Results

In this section, we present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the *second-order statistics isometry fitting* (SIF) method proposed in [72] and the subspace (SS) method developed in [27, 28]. The subspace method yields an estimate of $\mathbf{H}(z) \triangleq [\mathbf{h}_1(z)\ \mathbf{h}_2(z)\ \cdots\ \mathbf{h}_p(z)]$ up to a constant matrix factor if $\mathbf{H}(z)$ is irreducible, column-reduced and of equal column degrees. To retrieve the unknown constant matrix, the method [115] is shown to be a good choice for incorporation into the subspace method for blind estimation of MIMO FIR channel driven by colored signals. Such a combined scheme does not require the knowledge of the input statistics, and it is free of local minima and has demonstrated a superior performance as compared to other methods [70, 71]. In our simulations, the performance is measured by the normalized root-mean-square error (NRMSE) of the channel estimate and the SER of the estimated data symbols. The NRMSE of each user's channel estimate is defined as

$$\text{NRMSE(i)} = \sqrt{\frac{1}{N_{mc}} \sum_{t=1}^{N_{mc}} \frac{\|\rho(t)\hat{\mathbf{h}}_i^{(t)} - \mathbf{h}_i\|^2}{\|\mathbf{h}_i\|^2}}$$

where N_{mc} is the number of Monte Carlo runs, $\rho(t)$ is a complex scalar that minimizes $\|\rho(t)\hat{\mathbf{h}}_i^{(t)} - \mathbf{h}_i\|^2$. Also, in our simulations, the additive noise $\mathbf{w}(n)$ is taken as spatial-temporal white Gaussian noise with variance σ_w^2. The SNR is defined as

$$\text{SNR} = 10 \cdot \log \frac{E[\|\mathcal{H}\vec{\mathbf{s}}(n)\|^2]}{E[\|\vec{\mathbf{w}}(n)\|^2]}.$$

In the following examples, the noise variance is estimated as the smallest eigenvalues of the matrix $\mathbf{R}_x[0]$ and thus subtracted to provide the respective algorithms with denoised estimated autocorrelation matrices.

5.5.1 Scenario with Multiple Sources – Channel Estimation

We consider $p = 2$ sources which are i.i.d. information sequences with 4-QPSK digital modulation format, i.e. $\mathcal{S} = \{1, -1, i, -i\}$. To generate the colored sources, we pass these two i.i.d. information sequences through correlative filters prior to transmission. The correlative filters are chosen to be $f_1(z) = k_1(1 + \frac{1}{4}e^{-i\pi/2}z^{-1})$ and $f_2(z) = k_2(1 + \frac{1}{2}e^{i\pi/4}z^{-1})$ respectively for user 1 and user 2, where k_1 and k_2 are normalizing constants to ensure unit-power outputs. We consider a wireless communication scenario with these two colored user signals arriving at a single sensor via a multipath channel. The channel impulse responses of the users are respectively given as

$$h_1(t) = (c(t, 0.1) - 0.7c(t - 0.5T, 0.1))W_{4T}(t)$$
$$h_2(t) = (c(t - 0.2T, 0.1) - 0.5c(t - 0.6T, 0.1))W_{4T}(t)$$

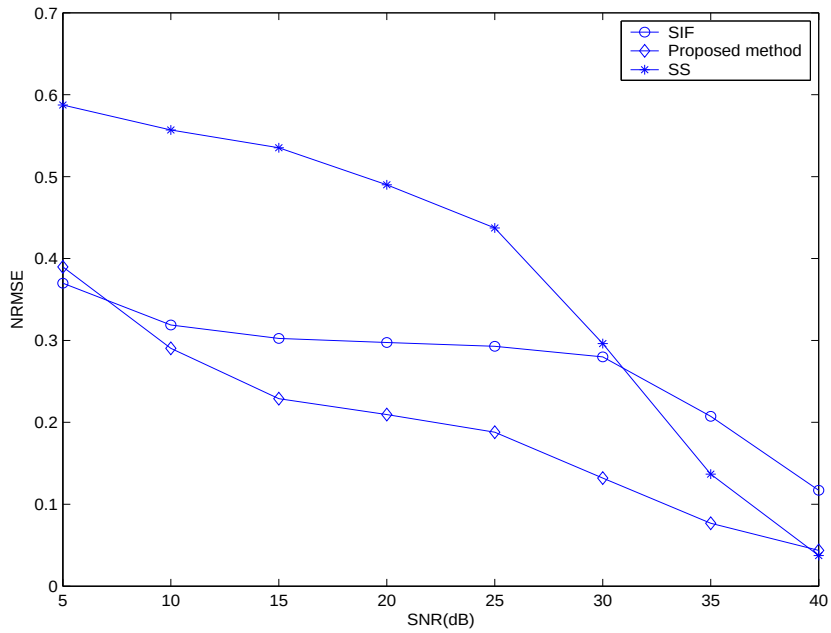
where $c(t - t_0, \theta)$ denotes a raised cosine pulse with roll-off factor θ and delay t_0, $W_{4T}(t)$ is a square window of duration 4 symbols intervals. We sample

the received signal three times the symbol rate, thus generating a two-input three-output linear system with channel order $L_1 = L_2 = 3$. For simplicity, knowledge of the channel order of each user is assumed. The rank of $\mathbf{R}_x[0]$ is estimated as $r = \sum_{i=1}^{p}(N + L_i + 1)$. Results are averaged over 100 Monte Carlo runs.

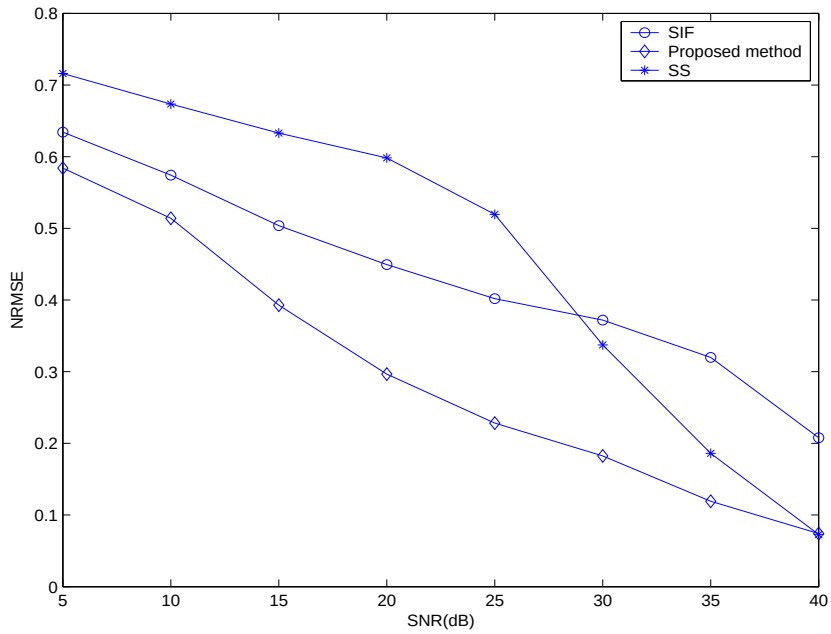
In the following, we illustrate the channel estimate performance of the respective algorithms, and show how it depends on the SNR of the received data and the number of samples used to estimate signal statistics, T_s. We choose stack number $N = 8$. For the SIF identification technique, six autocorrelation matrices are used for matching purposes, i.e. $k = \{\pm 1, \pm 2, \pm 3\}$ in Eqn.(5.6). Figure 5.1 shows the performance of the three algorithms as SNR is varied. $T_s = 2000$ samples are used to estimate signal statistics in each Monte Carlo run. Next, in Figure 5.2, the NRMSE of the channel estimate is shown as a function of the number of samples, with SNR = 30dB. Clearly, we can see that all three algorithms improve consistently as SNR or number of samples T_s increases. Also it seems that, for this specified channel, the performance of the subspace method is more liable to be affected by the number of samples used for estimation as compared to the other two algorithms. And our proposed algorithm presents a clear performance advantage over SIF and SS methods. Notice that due to the variation between the channel impulse responses of user 1 and user 2, the SNR of received signals of user 1 is higher, and as a consequence, the channel estimate of user 1 is better.

5.5.2 Scenario with Multiple Sources – Channel Equalization

In this example, we are interested in the equalization performance of the respective algorithms. We consider two colored sources arriving at $q = 3$ antennas.

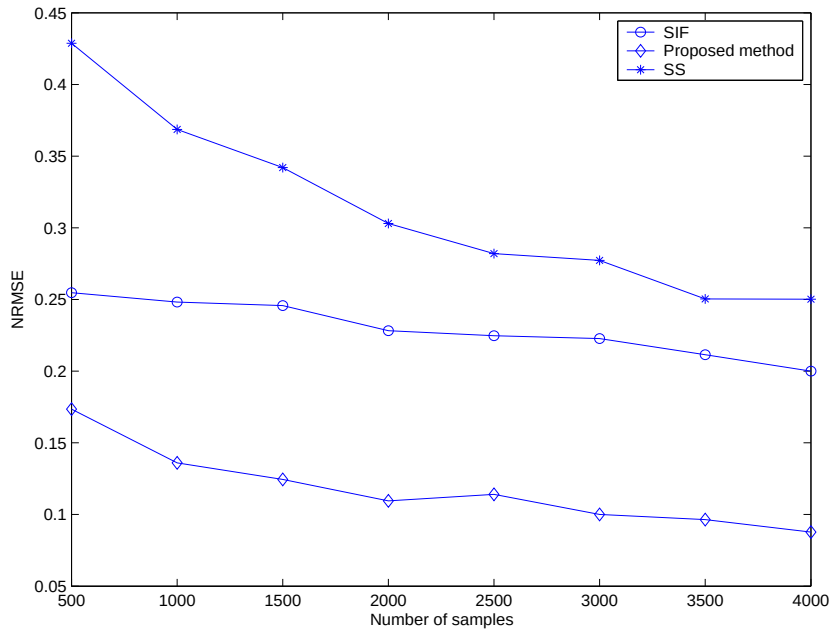


(a) channel of user 1

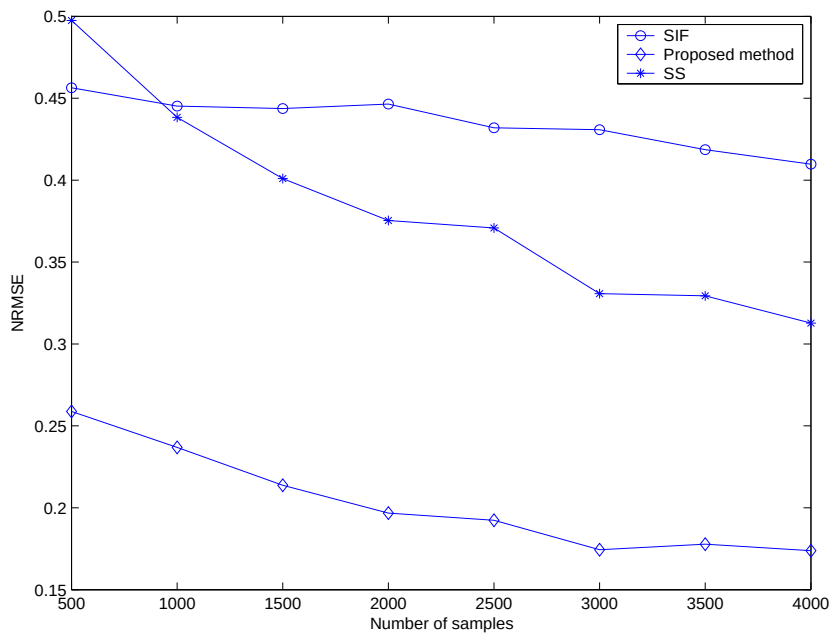


(b) channel of user 2

Figure 5.1: NRMSE of the estimated channel versus SNR, $T_s = 2000$



(a) channel of user 1



(b) channel of user 2

Figure 5.2: NRMSE of the estimated channel versus number of samples T_s, SNR = 30dB

The two colored sources are induced in the same way as in the example above. The polynomial channel matrix $\mathbf{H}(z)$ thus has dimension 3×2 and its degree is chosen to be 2, i.e. $L_1 = L_2 = 2$. In our simulations, all parameters of $\mathbf{H}(z)$ are randomly chosen from $\mathcal{N}(0,1) + i\mathcal{N}(0,1)$ (i.e. complex Gaussian variable with independent real and imaginary parts) at each run. The stack number N is chosen to be 4 and knowledge of the channel order is assumed.

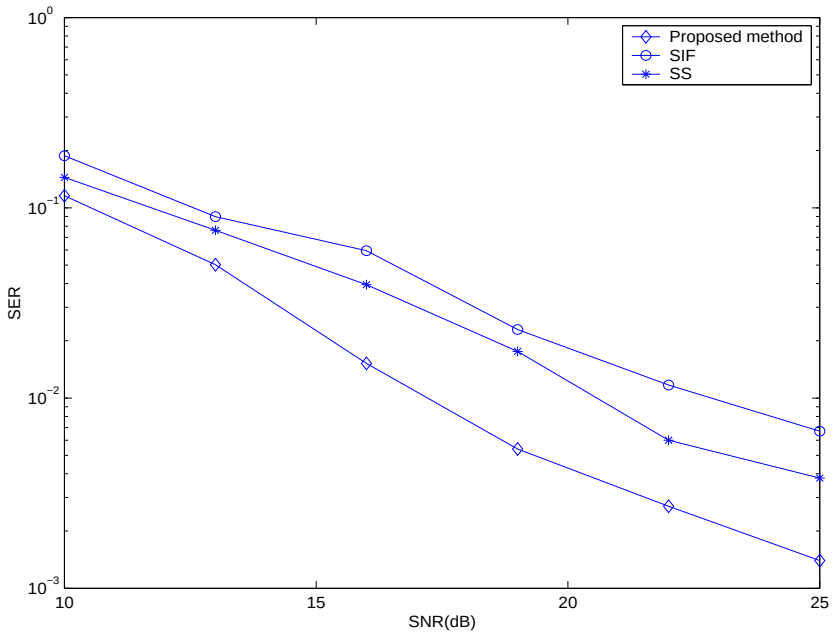
Once the channel matrix has been estimated by the respective algorithms, we can compute the ZF equalizers and the MMSE equalizers respectively as

$$\mathcal{E}_{\text{ZF}} = \hat{\mathcal{H}}^{\dagger}$$
$$\mathcal{E}_{\text{MMSE}} = \mathcal{E}_{\text{ZF}}(\mathbf{I} - \sigma_w^2 \hat{\mathbf{R}}_x^{-1}[0])$$

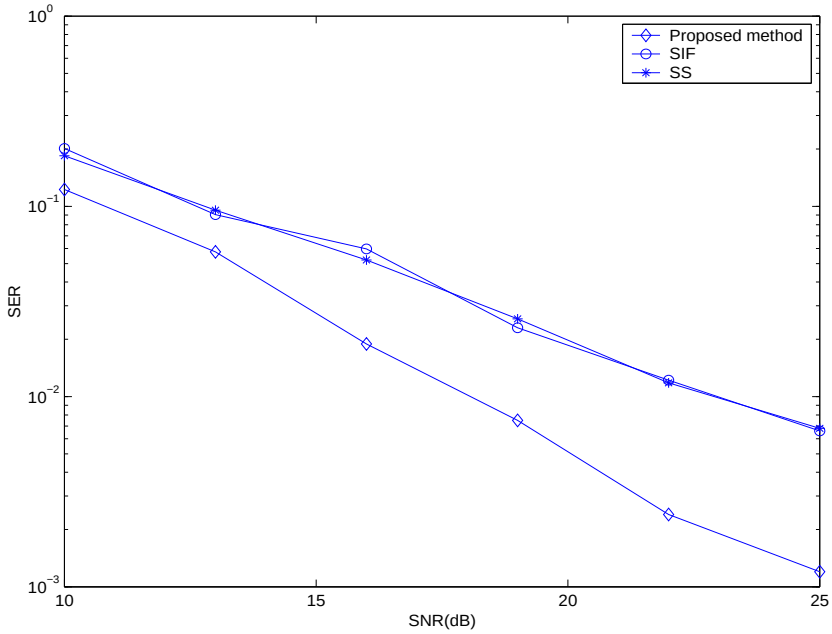
where $\hat{\mathbf{R}}_x[0]$ is the estimated autocorrelation matrix before denoised. The above expression for the MMSE equalizers was first derived in [116] for white inputs and then readily extended in [67] for colored inputs. The different rows of $\mathcal{E}_{\text{MMSE}}[1 : N + L_1 + 1, :]$ correspond to equalizers with different equalization delays of user 1; the different rows of $\mathcal{E}_{\text{MMSE}}[N + L_1 + 2 : 2(N+1) + L_1 + L_2, :]$ correspond to equalizers with different equalization delays of user 2. The scalar ambiguity of equalizers per user is removed before we perform the equalization. After channel equalization, the *filtered information sequences* (the outputs of the information sequences passing through the correlative filters) of each source are recovered and we can further detect the information sequences by adopting the Viterbi algorithm-based maximum likelihood detector. We present the equalization performance of the algorithms in the following three tables. The results are averaged over 500 Monte Carlo runs. In Figure 5.3, we show the SER associated with the two sources as a function of SNR with $T_s = 2000$ and equalization delay $d_e = 2$. It can be seen that, as expected, all the three algorithms

improve with increasing SNR. And our proposed algorithm shows a consistently lower SER than SIF and SS. Next, in Figure 5.4, the SER associated to the two sources are shown as a function of T_s for SNR = 19dB and $d_e = 2$. Once again our proposed algorithm presents a clear advantage over SIF and SS in terms of SER. Finally, keeping the SNR constant at 19dB and $T_s = 2000$, we investigate the relationship between SER performance and the equalization delays d_e. Figure 5.5 shows that for all these three algorithms, the equalizers associated with intermediate delays yield lower SER than those associated with extremal delays. And at almost all equalization delays, our proposed algorithm achieves better performance than the other two algorithms.

It is also interesting to consider the scenario where the spectral diversity condition of the input colored sources is 'weakly' satisfied. Although, theoretically, this spectral diversity condition can be satisfied with probability one for all three algorithms, however, a weakly satisfied spectral diversity condition does not stand strong against noise and estimation errors. Hence, this scenario tends to be a non-identifiable one in practice. Consider the two colored sources which are induced in the same way as in the previous example with the correlative filters replaced by $f_1(z) = k_1(1-0.25iz^{-1})$ and $f_2(z) = k_2(1-0.5iz^{-1})$. The zeros of these two filters are near and thus, the induced colored source signals have similar power spectra and only 'weakly' satisfy the spectral diversity condition. Figure 5.6 shows the equalization performance of the respective algorithms as a function of SNR with $T_s = 2000$ and $d_e = 2$. It can be seen that, as compared to Figure 5.3, all these three algorithms suffer from a certain performance loss in this case, and the performance degrades more rapidly as SNR deteriorates. This numerical example shows that a sufficiently diverse power spectra of the input colored sources is required to enhance the algorithm's robustness to the

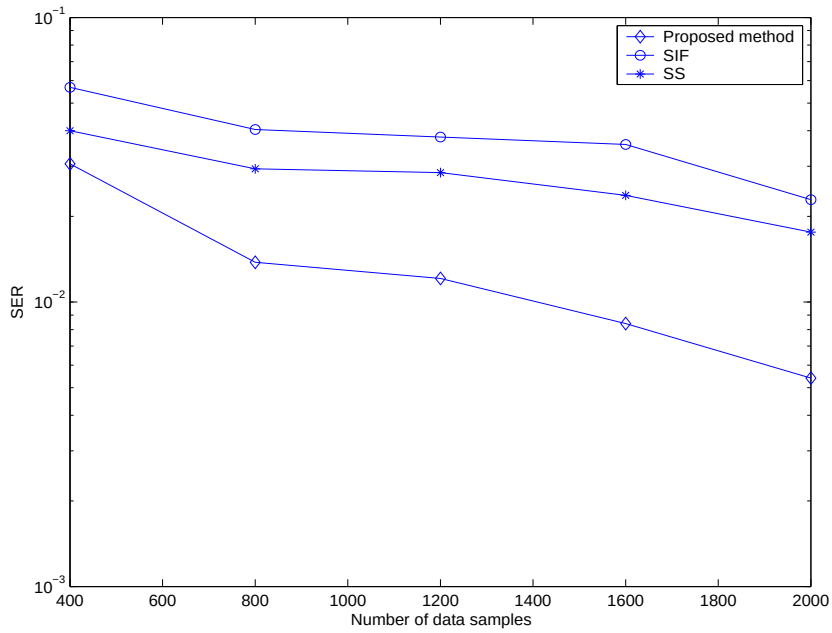


(a) user 1

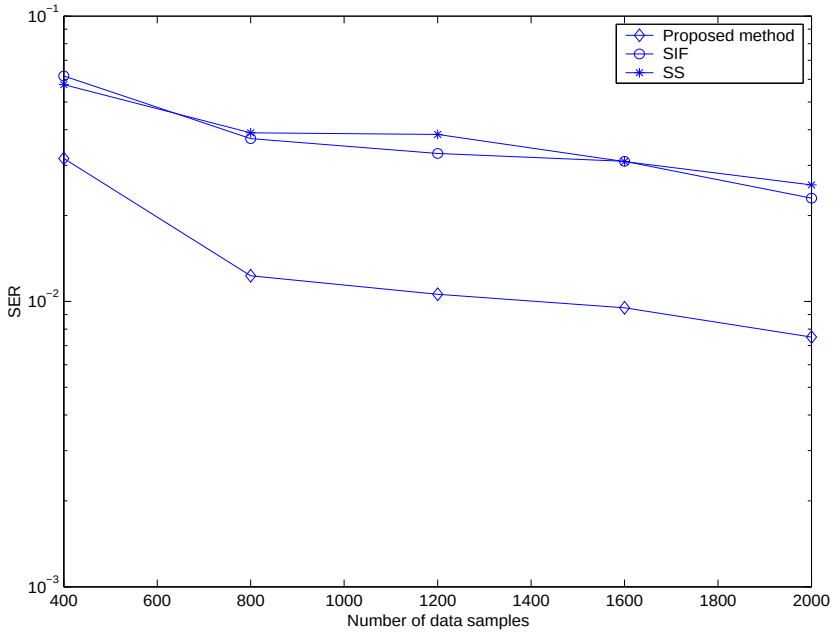


(b) user 2

Figure 5.3: SER versus SNR, $T_s = 2000$

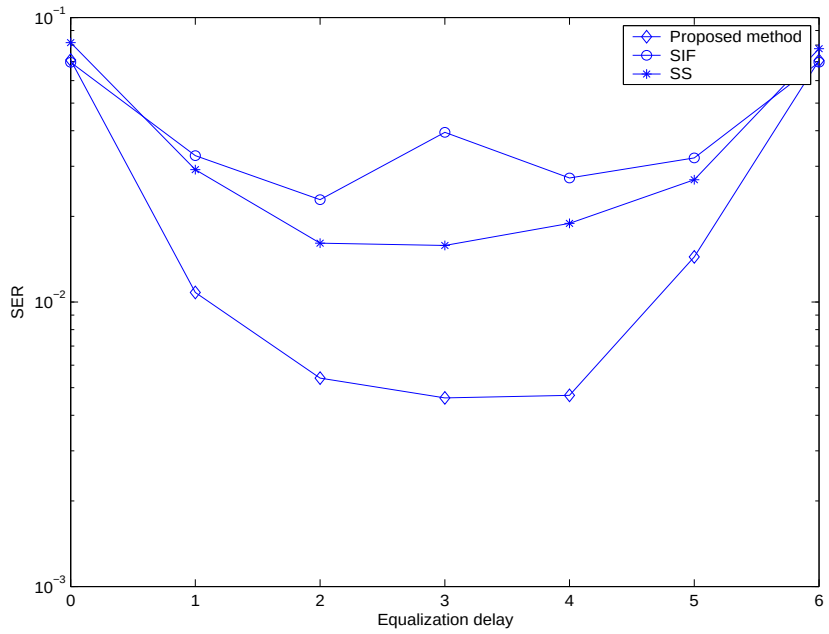


(a) user 1

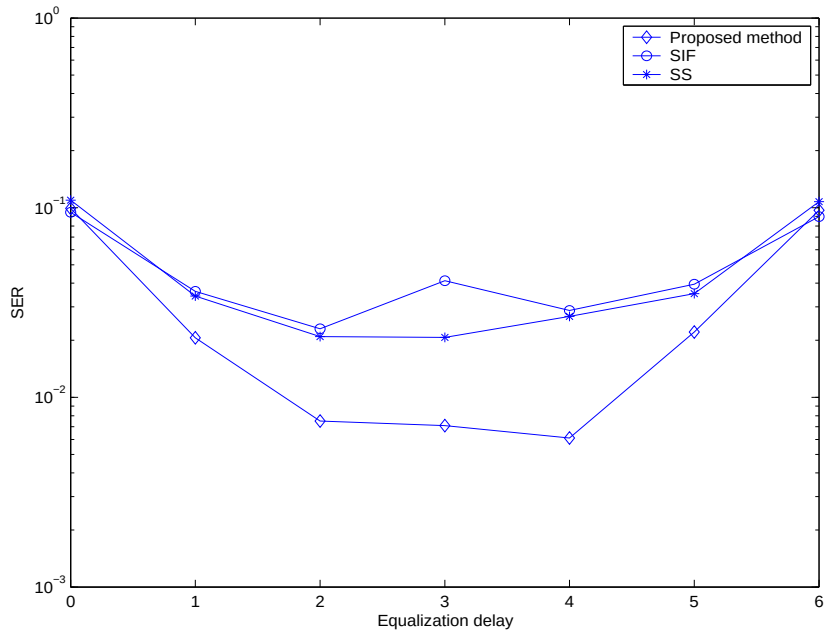


(b) user 2

Figure 5.4: SER versus number of data samples T_s, SNR = 19dB

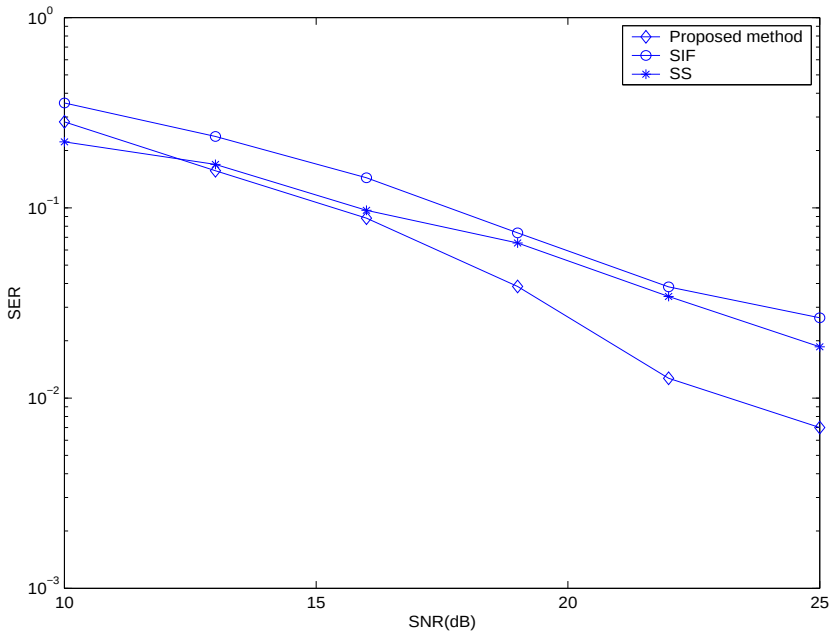


(a) user 1

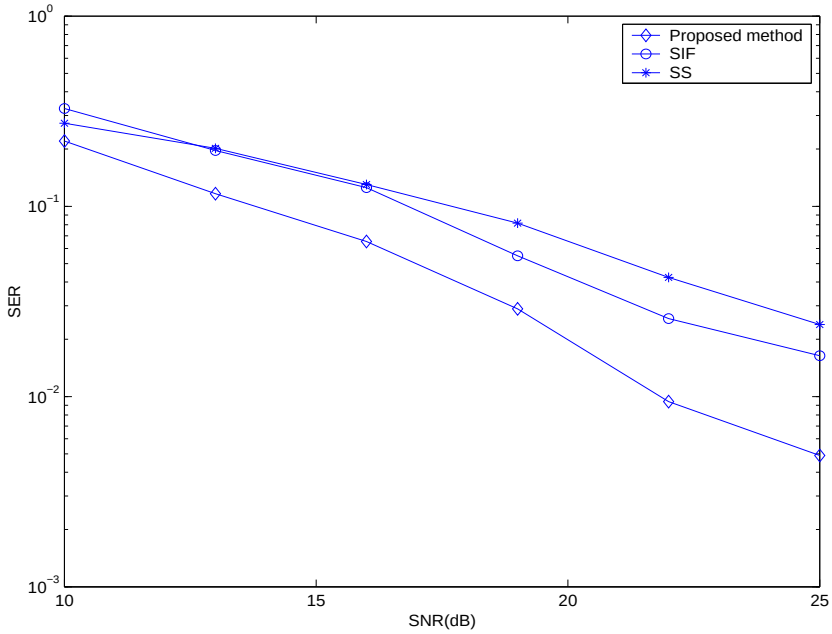


(b) user 2

Figure 5.5: SER versus equalization delays, SNR = 19dB, $T_s = 2000$



(a) user 1



(b) user 2

Figure 5.6: SER versus SNR for the 'weakly' satisfied spectral diversity condition, $T_s = 2000$

noise and estimation errors.

5.6 Summary

In this chapter, we extended our proposed method in Chapter 4 to MIMO systems and presented a closed-form solution for blind estimation of MIMO FIR channel driven by colored sources. The original proof for the uniqueness of the closed-form system solution was provided by exploiting the inherent structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$ and the derived properties of the companion matrices. It is noted that the properties of the companion matrices were further exploited in this chapter to prove the uniqueness of the system solution. Theoretical analysis showed that our method requires a much less computational complexity and induces a less restrictive identifiability condition as compared to the method [72]. Simulation results showed that the new method compares favorably with the SOS-based methods [28, 72]. Also, the numerical example demonstrated that a sufficiently diverse power spectra of the input colored sources is required to enhance the algorithms' robustness to the noise and estimation errors.

Chapter 6

Blind Equalization of SIMO FIR Channel

In this chapter, we consider the blind equalization of SIMO FIR channel driven by colored signals. The statistics of the input colored signals are unknown *a priori*. By exploiting the inherent structural relationship between the source autocorrelation matrices of different delay lags, the closed-form ZF and MMSE equalizers of desired delays are estimated from the SOS of the received data. The blind equalizability conditions of the proposed method are investigated. Numerical simulation results are presented to illustrate the performance of the proposed algorithm.

6.1 Introduction

Encouraged by the achieved results (Chapter 4) for blind identification of channel driven by colored sources with known input statistics, in this chapter, we are interested in investigating the problem of blind identification/equalization of

channel driven by colored sources with unknown input statistics. It is shown in the pioneering work [17] that the SIMO FIR channel can be perfectly identified/equalized from the second-order statistics of the received data under quite general assumptions. Following [17], numerous SOS-based blind identification/equalization methods [1, 23–25] have been proposed. Nevertheless, most existing SOS-based methods are applied to the i.i.d. input signals. The work that consider the case of correlated input signals are much less, see [66–68]. Specifically, the work [66–68] treated the case where the input signal statistics are *colored and known*. As for the case where the input statistics are *colored but unknown*, it seems much more difficult to devise a SOS-based algorithm since no prior statistical information of the transmitted signals can be utilized. One solution to this problem is given in [22], which proposed a subspace-based method by exploiting the block Toeplitz structure of the channel convolution matrix, and thus required no knowledge of input statistics whatsoever. The extension of [22] to MIMO systems was studied in [28]. There are some other work [70, 71] which studied blind identification/equalization of MIMO FIR channel driven by colored signals with unknown statistics. However, both work [70, 71] constitute a two-step approach that is based on [22]. They, firstly, determine source separating vectors or decorrelators to separate the sources. Once the sources are separated, the second step utilizes the subspace method [22] to estimate the resulting SIMO systems and the original MIMO systems. Some deterministic approaches that can handle arbitrarily correlated source signals have been discussed in [44, 48, 52, 98] for blind SIMO channel identification/equalization. They are most effective at high SNR and for small data samples scenarios. Note that the deterministic method [44] has its statistical version whose performance is similar as the subspace method [22]. Besides the above mentioned methods, the mutually referenced equalizers (MRE) method proposed in [49] and the

constrained minimum output energy (MOE) algorithm presented in [117] are also important work in blind SIMO channel identification/equalization. The work [49] was developed on the concept of the mutually referenced equalizers, i.e., the outputs of the set of filters (equalizers) act as training signals for each other. The method does not rely on the specific assumptions concerning the input statistics, and several variations of the MRE criterion including a stochastic criterion using the second-order statistics have been derived. Another interesting work [117] explored the popular constrained minimum output energy approach to derive the optimal blind equalizers. As indicated in [117], the method is also insensitive to the color of the input signals.

In this chapter, we study the blind equalization of SIMO FIR channel when the input signals are colored but the source statistics are unknown. It is shown that although the statistical information of the transmitted signals is not available, we can still estimate the equalizers of desired delays from the second-order statistics of the received data by exploiting the inherent structural relationship between source autocorrelation matrices of different delay lags. The proposed method is essentially different from the subspace method [22] since it does not exploit the structure of the channel convolution matrix. As a consequence, our proposed method can be applied to the special case where the receiver diversity is of high dimension such that the observations are not necessary to be stacked (or smoothed) in time, and thus the resulting channel convolution matrix will not have a block Toeplitz structure. Another important property of our proposed method is that it directly computes the equalizers of desired delays, which is different from most SOS-based methods available to date that identify the channel first and then use it to estimate the equalizer coefficients. Direct blind equalization algorithms have been proposed in [1, 25] for white

input signals and in [49, 117] for colored input signals. In [1], the equalizers are estimated from the eigenvectors of certain rank-one matrices constructed from the autocorrelation matrices of the received data. However, when the input signals are colored, the correspondingly constructed matrices will not be rank-one and the described properties about the matrices in [1] will not hold. The contribution of this chapter consists of the following two aspects. First, the inherent structural relationship between the source autocorrelation matrices of different delay lags is exploited. Second, we generalize the proposed algorithm in [1] to the colored sources. As will be evident in this chapter, such an extension is not so straightforward and is highly nontrivial.

The chapter is organized as follows. In Section 6.2, the SIMO system model and some basic assumptions are introduced. Next, in Section 6.3, we present our blind channel equalization method and investigate the corresponding equalizability conditions. The closed-form ZF and MMSE equalizers are derived from the matrices constructed by the autocorrelation matrices of received data. Finally, in Section 6.4, numerical simulation results are presented to demonstrate the performance of the proposed algorithm.

6.2 System Model and Basic Assumptions

The SIMO FIR channel model considered in this chapter is the same as that discussed in Chapter 4:

$$\mathbf{x}(n) \triangleq \mathbf{h}(n) \circledast s(n) + \mathbf{w}(n) \triangleq \sum_{l=0}^{L} \mathbf{h}(l) s(n-l) + \mathbf{w}(n) \tag{6.1}$$

where $\{s(n)\}$ is the zero mean, wide sense stationary sequence of transmitted symbols; $\{\mathbf{x}(n)\}$ is the $q \times 1$ channel output vector; $\{\mathbf{w}(n)\}$ is the $q \times 1$ white noise vector, and $\{\mathbf{h}(n)\}$ represents the multichannel impulse response; the number of subchannels is q. This multichannel model arises by deploying multiple sensors or by fractionally sampling the channel output.

As we did before, we stack the channel output vector $\{\mathbf{x}(n)\}$ and define: $\vec{\mathbf{x}}(n) \triangleq [\mathbf{x}^T(n)\ \mathbf{x}^T(n-1)\ \ldots\ \mathbf{x}^T(n-N)]^T$, $\vec{\mathbf{s}}(n) \triangleq [s(n)\ s(n-1)\ \ldots\ s(n-N-L)]^T$, and $\vec{\mathbf{w}}(n) \triangleq [\mathbf{w}^T(n)\ \mathbf{w}^T(n-1)\ \ldots\ \mathbf{w}^T(n-N)]^T$, where N is called the stack number or smoothed factor. Therefore we can re-express Eqn.(6.1) as the following matrix form

$$\vec{\mathbf{x}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n) \tag{6.2}$$

where the channel convolution matrix $\mathcal{H} \in \mathbb{C}^{(N+1)q \times d}$ is a block Toeplitz matrix defined in the previous chapter.

We adopt the following basic assumptions: A1) $\mathcal{H}$ is full column rank: a condition equivalent to requiring that the channel $\mathbf{h}(z)$ is irreducible. A2) Channel order L is assumed to be known *a priori*. A3) The input signals $\{s(n)\}$ are zero-mean wide-sense stationary colored signals whose input statistics are unavailable. A4) Additive noises $\{\mathbf{w}(n)\}$ are spatially and temporally white noises, and they are statistically independent of the source. Note that the assumptions made in this chapter are the same as Chapter 4, with the exception of A3. Our objective is to estimate the channel impulse response by utilizing the second-order statistics of the observed output data.

6.3 Proposed Channel Equalization Method

We begin by defining the source autocorrelation matrix with delay lag k as follows

$$\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n-k)]. \tag{6.3}$$

In order to simplify the presentation of the proposed channel equalization method, we assume the noiseless case. Thus, from Eqn.(6.2), the autocorrelation matrix of the received data $\vec{\mathbf{x}}(n)$ with delay lag k can be expressed as

$$\mathbf{R}_x[k] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^H(n-k)] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H. \tag{6.4}$$

If the input signals are i.i.d., then the source autocorrelation matrix $\mathbf{R}_s[k]$ is given by

$$\begin{cases} \mathbf{R}_s[k] = \mathbf{J}_k & k > 0 \\ \mathbf{R}_s[k] = \mathbf{I} & k = 0 \\ \mathbf{R}_s[-k] = \mathbf{J}_k^T & k > 0 \end{cases}$$

where the symbol $\mathbf{J}_k$ ($\mathbf{J}_k^T$) stands for the k-lag down (up) shift square matrix whose k^{th} sub-diagonal entries below (above) the main diagonal are unity, whereas all remaining entries are zero. This special structure possessed by the source autocorrelation matrix has facilitated the design of numerous SOS-based algorithms, e.g. [17]. However, when the input signals are colored, the source autocorrelation matrix degenerates into a more general form that is hard to be exploited directly. Moreover, for the case where the knowledge of the input statistics is not available, the exact values of the entries of the source autocorrelation matrix are unknown. To overcome all these difficulties, we need to

exploit the inherent structural relationship between the source autocorrelation matrices. Such a structural relationship is first observed in [67] between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[1]$, and later developed in [68] for $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm 1]$. In this chapter, we will further explore this structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm k]$ for $1 \leq k \leq d-1$.

6.3.1 Inherent Structural Relationship Between Source Autocorrelation Matrices

We begin by observing the structural relationship between $\mathbf{R}_s[0]$ and $\mathbf{R}_s[\pm k]$, where $1 \leq k \leq d-1$. It can be seen that the last $d-k$ rows of $\mathbf{R}_s[k]$ are the first $d-k$ rows of $\mathbf{R}_s[0]$, and the first $d-k$ rows of $\mathbf{R}_s[-k]$ are the last $d-k$ rows of $\mathbf{R}_s[0]$. Hence we can establish the following relationship

$$\mathbf{R}_s[k] = \mathbf{J}_k \mathbf{R}_s[0] + \sum_{i=1}^{k} \mathbf{e}_i \mathbf{r}_{k,i}^H \tag{6.5}$$

$$\mathbf{R}_s[-k] = \mathbf{J}_k^T \mathbf{R}_s[0] + \sum_{i=d-k+1}^{d} \mathbf{e}_i \mathbf{t}_{k,i}^H \tag{6.6}$$

where $\mathbf{r}_{k,i}$ and $\mathbf{t}_{k,i}$ are defined as follows with $1 \leq i \leq k$ for Eqn.(6.7) and $d-k+1 \leq i \leq d$ for Eqn.(6.8) respectively

$$\mathbf{r}_{k,i}^H \triangleq \mathbf{e}_i^H \mathbf{R}_s[k] = E[s(n-i+1)\vec{\mathbf{s}}^H(n-k)] \tag{6.7}$$

$$\mathbf{t}_{k,i}^H \triangleq \mathbf{e}_i^H \mathbf{R}_s[-k] = E[s(n-i+1)\vec{\mathbf{s}}^H(n+k)] \tag{6.8}$$

In addition, for any pairs (i, j) which satisfy $i+j = d+1$, we have the following relationship

$$\mathbf{r}_{k,i}[m] = \mathbf{t}^*_{k,j}[d+1-m] \qquad \forall\ m \in \{1, \ldots, d\} \tag{6.9}$$

where $\mathbf{r}_{k,i}[m]$ denotes the mth entry of $\mathbf{r}_{k,i}$; $\mathbf{t}_{k,i}[m]$ denotes the mth entry of $\mathbf{t}_{k,i}$.

6.3.2 Channel Equalization

In order to utilize the above structural relationship between source autocorrelation matrices, we introduce the following lemma that has been used in previous chapters.

Lemma 6.1 *Given* $\mathbf{R}_x[k] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H$, $\mathcal{H}$ *is full column rank and* $\mathbf{R}_s[0]$ *is invertible, we have*

$$\mathbf{R}_x[k]\mathbf{R}_x^\dagger[0] = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger. \tag{6.10}$$

Proof: See the proof of Lemma 4.1. ■

Using Eqn.(6.5–6.6), we can express $\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]$ as follows

$$\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0] = \mathbf{J}_k + \sum_{i=1}^{k} \mathbf{e}_i \vec{\alpha}^H_{k,i} \qquad 0 < k < d \tag{6.11}$$

$$\mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0] = \mathbf{J}_k^T + \sum_{i=d-k+1}^{d} \mathbf{e}_i \vec{\beta}^H_{k,i} \qquad 0 < k < d \tag{6.12}$$

where

$$\vec{\alpha}_{k,i} \triangleq \mathbf{R}_s^{-1}[0]\mathbf{r}_{k,i} \tag{6.13}$$

$$\vec{\beta}_{k,i} \triangleq \mathbf{R}_s^{-1}[0]\mathbf{t}_{k,i}. \tag{6.14}$$

It can be seen that the entries in $\vec{\alpha}_{k,i}$ are exactly the coefficients of the multistep d^{th}-order optimum forward prediction error filter for the process $\{s(n)\}$ and the entries in $\vec{\beta}_{k,i}$ are the coefficients of the multistep d^{th}-order optimum backward prediction error filter for the process $\{s(n)\}$ [114]. $\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]$ is constructed by $\mathbf{J}_k$ with its ith row of the first all-zero k rows replaced by the row vector $\vec{\alpha}_{k,i}^H$; likewise, $\mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]$ is constructed by $\mathbf{J}_k^T$ with its ith row of the last all-zero k rows replaced by the row vector $\vec{\beta}_{k,i}^H$. It is still hard to directly utilize the structure of $\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]$ and $\mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]$. Intuitively, it may be better if the ones of the constructed matrices are on the diagonal instead of on the sub-diagonal. Hence we have the following further transformation. Define

$$\Gamma_k \triangleq \mathbf{R}_x[k]\mathbf{R}_x^\dagger[0]\mathbf{R}_x[-k]\mathbf{R}_x^\dagger[0].$$

Using Eqn.(6.10), we can express Γ_k as

$$\Gamma_k = \mathcal{H}\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]\mathcal{H}^\dagger \triangleq \mathcal{H}\Omega_k\mathcal{H}^\dagger \tag{6.15}$$

where we let $\Omega_k \triangleq \mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]\mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]$. Using Eqn.(6.11–6.12), we get the following

$$\Omega_k = \left(\mathbf{J}_k + \sum_{i=1}^{k}\mathbf{e}_i\vec{\alpha}_{k,i}^H\right)\left(\mathbf{J}_k^T + \sum_{j=d-k+1}^{d}\mathbf{e}_j\vec{\beta}_{k,j}^H\right)$$

$$
\begin{aligned}
&= \hat{\mathbf{I}}_{d-k} + \sum_{i=1}^{k} \mathbf{e}_i \vec{\alpha}_{k,i}^H \left(\mathbf{J}_k^T + \sum_{j=d-k+1}^{d} \mathbf{e}_j \vec{\beta}_{k,j}^H \right) \\
&= \hat{\mathbf{I}}_{d-k} + \sum_{i=1}^{k} \mathbf{e}_i \mathbf{a}_{k,i}^H \qquad (6.16)
\end{aligned}
$$

where

$$
\hat{\mathbf{I}}_{d-k} \triangleq \mathbf{J}_k \mathbf{J}_k^T = \mathrm{diag}([\mathbf{0}_{1\times k}, \mathbf{1}_{1\times(d-k)}])
$$

$$
\mathbf{a}_{k,i}^H \triangleq \vec{\alpha}_{k,i}^H \left(\mathbf{J}_k^T + \sum_{j=d-k+1}^{d} \mathbf{e}_j \vec{\beta}_{k,j}^H \right)
$$

and the term $\mathbf{J}_k \sum_{j=d-k+1}^{d} \mathbf{e}_j \vec{\beta}_{k,j}^H$ is equal to zero. Similarly, $\Omega_{-k} = \mathbf{R}_s[-k]\mathbf{R}_s^{-1}[0]\mathbf{R}_s[k]\mathbf{R}_s^{-1}[0]$ is given as

$$
\begin{aligned}
\Omega_{-k} &= \left(\mathbf{J}_k^T + \sum_{i=d-k+1}^{d} \mathbf{e}_i \vec{\beta}_{k,i}^H \right) \left(\mathbf{J}_k + \sum_{j=1}^{k} \mathbf{e}_j \vec{\alpha}_{k,j}^H \right) \\
&= \breve{\mathbf{I}}_{d-k} + \sum_{i=d-k+1}^{d} \mathbf{e}_i \vec{\beta}_{k,i}^H \left(\mathbf{J}_k + \sum_{j=1}^{k} \mathbf{e}_j \vec{\alpha}_{k,j}^H \right) \\
&= \breve{\mathbf{I}}_{d-k} + \sum_{i=d-k+1}^{d} \mathbf{e}_i \mathbf{b}_{k,i}^H \qquad (6.17)
\end{aligned}
$$

where

$$
\breve{\mathbf{I}}_{d-k} \triangleq \mathbf{J}_k^T \mathbf{J}_k = \mathrm{diag}([\mathbf{1}_{1\times(d-k)}, \mathbf{0}_{1\times k}])
$$

$$
\mathbf{b}_{k,i}^H \triangleq \vec{\beta}_{k,i}^H \left(\mathbf{J}_k + \sum_{j=1}^{k} \mathbf{e}_j \vec{\alpha}_{k,j}^H \right)
$$

and the term $\mathbf{J}_k^T \sum_{j=1}^{k} \mathbf{e}_j \vec{\alpha}_{k,j}^H$ is equal to zero. Substituting Ω_k and Ω_{-k} with the obtained relations in Eqn.(6.16–6.17), we thus have

$$\Gamma_k = \mathcal{H}\left(\hat{\mathbf{I}}_{d-k} + \sum_{i=1}^{k} \mathbf{e}_i \mathbf{a}_{k,i}^H\right)\mathcal{H}^\dagger \tag{6.18}$$

$$\Gamma_{-k} = \mathcal{H}\left(\breve{\mathbf{I}}_{d-k} + \sum_{i=d-k+1}^{d} \mathbf{e}_i \mathbf{b}_{k,i}^H\right)\mathcal{H}^\dagger. \tag{6.19}$$

In the later part of this chapter, we will show that the closed-form channel equalizers of any desired delays can be estimated from these two matrices Γ_k, Γ_{-k} and their matrix product. We, first, define the following notation.

$$\Upsilon \triangleq \mathcal{H}\left(\sum_{i\in S} \mathbf{e}_i \mathbf{c}_i^H + \sum_{i\in S^c} \mathbf{e}_i \mathbf{c}_i^H\right)\mathcal{H}^\dagger \tag{6.20}$$

where S is a subset of $\{1, 2, \cdots, d\}$, S^c denotes the complement of S, i.e. $S \cup S^c = \{1, 2, \cdots, d\}$. Also for any $i \in S$, we have $\mathbf{c}_i = \mathbf{e}_i$. Clearly, the above defined matrix Υ is a generalized form of the matrices Γ_k, Γ_{-k} and the matrix product $\Gamma_{k_1}\Gamma_{-k_2}$ (It can be easily verified that $\Gamma_{k_1}\Gamma_{-k_2}$ can also be written as the form in Eqn.(6.20)). We would like to investigate this generalized form and exploit some property of this matrix Υ in the following theorem. Before we proceed, we define

$$\mathbf{C} \triangleq [\mathbf{c}_1 \;\; \mathbf{c}_2 \;\; \cdots \;\; \mathbf{c}_d].$$

Let $c_{i,j}$ denote the $(i,j)^{\text{th}}$ element of $\mathbf{C}$. $\mathbf{C}_1$ is defined as the matrix formed by taking the entries $c_{i,j}$ out from $\mathbf{C}$ for any $i \in S$ or $j \in S$. $\mathbf{C}_1$ is a square matrix of dimension $|S^c| \times |S^c|$, where $|S^c|$ represents the number of elements in the set S^c. In a similar way, $\mathbf{C}_2$ is defined as the matrix formed by taking

the entries $c_{i,j}$ out from $\mathbf{C}$ for any $i \in S^c$ or $j \in S$. $\mathbf{C}_2$ is a matrix of dimension $|S| \times |S^c|$. As we can see, the matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ are all constructed from the vectors $\mathbf{c}_i$ for $i \in S^c$.

Theorem 6.1 *Given that the matrix Υ defined in Eqn.(6.20), for any $i \in S$, we have $\mathbf{c}_i = \mathbf{e}_i$; also, for $i \in S^c$, the vectors $\mathbf{c}_i$ satisfy the following condition, that is, the constructed matrix by the vectors $\mathbf{c}_i$ for $i \in S^c$, $\bar{\mathbf{C}} \triangleq \begin{bmatrix} \mathbf{C}_1 - \mathbf{I} \\ \mathbf{C}_2 \end{bmatrix}$, is full column rank, then for any non-zero vector $\mathbf{g}$ that satisfies $\mathbf{g}^H \Upsilon = \mathbf{g}^H$, $\mathbf{g}^H$ is a linear combination of the rows $\mathcal{H}^\dagger[i,:]$ for $i \in S$, that is,*

$$\mathbf{g}^H = \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:] \tag{6.21}$$

where $\mathcal{H}^\dagger[i,:]$ denotes the ith row of the matrix $\mathcal{H}^\dagger$.

Proof: It is easy to see that if $\mathbf{g}^H = \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:]$, then

$$\begin{aligned}
\mathbf{g}^H \Upsilon &= \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:] \mathcal{H} \left(\sum_{j \in S} \mathbf{e}_j \mathbf{c}_j^H + \sum_{j \in S^c} \mathbf{e}_j \mathbf{c}_j^H \right) \mathcal{H}^\dagger \\
&= \sum_{i \in S} w_i \mathbf{e}_i^H \left(\sum_{j \in S} \mathbf{e}_j \mathbf{e}_j^H + \sum_{j \in S^c} \mathbf{e}_j \mathbf{c}_j^H \right) \mathcal{H}^\dagger \\
&= \sum_{i \in S} w_i \mathbf{e}_i^H \mathcal{H}^\dagger \\
&= \mathbf{g}^H.
\end{aligned} \tag{6.22}$$

Now we prove that for any vector $\mathbf{g}$ that satisfies $\mathbf{g}^H \Upsilon = \mathbf{g}^H$, we have $\mathbf{g}^H = \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:]$. Since $\mathbf{g}^H = \mathbf{g}^H \Upsilon = \mathbf{g}^H \mathcal{H} (\sum_{i \in S} \mathbf{e}_i \mathbf{c}_i^H + \sum_{i \in S^c} \mathbf{e}_i \mathbf{c}_i^H) \mathcal{H}^\dagger$, it indicates that $\mathbf{g}^H$ is a linear combination of the rows of $\mathcal{H}^\dagger$. Hence we can

write

$$\mathbf{g}^H = \sum_{i=1}^{d} w_i \mathcal{H}^\dagger[i,:] \tag{6.23}$$

therefore we have

$$\begin{aligned}
\mathbf{g}^H \Upsilon &= \sum_{j=1}^{d} w_j \mathcal{H}^\dagger[j,:] \mathcal{H} \left(\sum_{i \in S} \mathbf{e}_i \mathbf{c}_i^H + \sum_{i \in S^c} \mathbf{e}_i \mathbf{c}_i^H \right) \mathcal{H}^\dagger \\
&= \sum_{j=1}^{d} w_j \mathbf{e}_j^H \left(\sum_{i \in S} \mathbf{e}_i \mathbf{e}_i^H + \sum_{i \in S^c} \mathbf{e}_i \mathbf{c}_i^H \right) \mathcal{H}^\dagger \\
&= \sum_{i \in S} w_i \mathbf{e}_i^H \mathcal{H}^\dagger + \sum_{i \in S^c} w_i \mathbf{c}_i^H \mathcal{H}^\dagger \\
&= \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:] + \sum_{i \in S^c} w_i \mathbf{c}_i^H \mathcal{H}^\dagger. \qquad (6.24)
\end{aligned}$$

By combining Eqn.(6.23–6.24), we have

$$\begin{aligned}
\mathbf{g}^H \Upsilon - \mathbf{g}^H &= \sum_{i \in S^c} w_i \mathbf{c}_i^H \mathcal{H}^\dagger - \sum_{i \in S^c} w_i \mathcal{H}^\dagger[i,:] \\
&= \sum_{j=1}^{d} \left(\sum_{i \in S^c} w_i c_{j,i}^* \right) \mathcal{H}^\dagger[j,:] - \sum_{j \in S^c} w_j \mathcal{H}^\dagger[j,:]. \qquad (6.25)
\end{aligned}$$

Since $\mathcal{H}^\dagger$ is full row rank and the rows of $\mathcal{H}^\dagger$ are independent, in order to make $\mathbf{g}^H \Upsilon = \mathbf{g}^H$, it is required that

$$\begin{aligned}
&\text{(i)} \quad \textstyle\sum_{i \in S^c} w_i c_{j,i}^* = w_j \quad \text{for } j \in S^c \\
&\text{(ii)} \quad \textstyle\sum_{i \in S^c} w_i c_{j,i}^* = 0 \quad \text{for } j \in S.
\end{aligned}$$

The above first condition is equivalent to the following condition in matrix form

$$\mathbf{w}_{s^c}^T \mathbf{C}_1^H = \mathbf{w}_{s^c}^T \tag{6.26}$$

where $\mathbf{w}_{s^c}$ is a column vector formed by taking the entries w_i out from the vector $\mathbf{w} \triangleq [w_1 \;\; w_2 \;\; \cdots \;\; w_d]^T$ for any $i \in S$. The number of entries in $\mathbf{w}_{s^c}$ is thus the same as the number of elements in the set S^c, which is denoted by $|S^c|$. Also, the above second condition can be written as

$$\mathbf{w}_{s^c}^T \mathbf{C}_2^H = \mathbf{0}. \tag{6.27}$$

These two conditions are combined and transformed as follows

$$\begin{bmatrix} \mathbf{C}_1 - \mathbf{I} \\ \mathbf{C}_2 \end{bmatrix} \mathbf{w}_{s^c}^* = \mathbf{0}. \tag{6.28}$$

Since the matrix $\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_1 - \mathbf{I} \\ \mathbf{C}_2 \end{bmatrix}$ is full column rank, we can conclude that $\mathbf{w}_{s^c} = \mathbf{0}$. Clearly, if there exists only an all-zero solution of $\mathbf{w}_{s^c}$ to satisfy the above two conditions, then the vector $\mathbf{g}$ that satisfies $\mathbf{g}^H \Upsilon = \mathbf{g}^H$ has the form $\mathbf{g}^H = \sum_{i \in S} w_i \mathcal{H}^\dagger[i,:]$. Therefore our theorem is valid *under the condition that* $\bar{\mathbf{C}}$ *is full column rank*. Such a condition is only related to the vectors $\mathbf{c}_i$ for $i \in S^c$. It can be easily verified that this condition, in fact, has excluded the possibility that any vector $\mathbf{c}_i$, for $i \in S^c$, is equal to $\mathbf{e}_i$. The proof is completed here. ■

From Theorem 6.1, it is easy to know that if S includes only one element m, i.e. $S = \{m\}$, then the non-zero vector $\mathbf{g}^H$ that satisfies $\mathbf{g}^H \Upsilon = \mathbf{g}^H$ will be equal to $w_m \mathcal{H}^\dagger[m,:]$, which is exactly the desired zero-forcing equalizer of delay $m-1$. Inspired by this, we have the following two lemmas that provide the key to estimate the zero-forcing equalizers of desired delays.

Lemma 6.2 *If* $\mathbf{g}^H \Gamma_k = \mathbf{g}^H$ *for* $k = d-1$, *then* $\mathbf{g}^H$ *is the zero-forcing channel*

equalizer of delay $d-1$*; also if* $\mathbf{g}^H\Gamma_{-k} = \mathbf{g}^H$ *for* $k = d-1$*, then* $\mathbf{g}^H$ *is the zero-forcing channel equalizer of delay* 0*.*

Proof: From Eqn.(6.18–6.19), we have

$$\Gamma_{d-1} = \mathcal{H}\left(\mathbf{e}_d\mathbf{e}_d^H + \sum_{i=1}^{d-1}\mathbf{e}_i\mathbf{a}_{d-1,i}^H\right)\mathcal{H}^\dagger \tag{6.29}$$

$$\Gamma_{-(d-1)} = \mathcal{H}\left(\mathbf{e}_1\mathbf{e}_1^H + \sum_{i=2}^{d}\mathbf{e}_i\mathbf{b}_{d-1,i}^H\right)\mathcal{H}^\dagger. \tag{6.30}$$

The proof is obvious from Theorem 6.1 if we let $\mathbf{c}_i = \mathbf{a}_{d-1,i}$ and $\mathbf{c}_i = \mathbf{b}_{d-1,i}$ for $i \in S^c$, respectively. ■

Lemma 6.3 *Let* $\mathbf{F}_{k_1,k_2} \triangleq \Gamma_{k_1}\Gamma_{-k_2}$*. If* $\mathbf{g}^H\mathbf{F}_{k_1,k_2} = \mathbf{g}^H$ *for any* $\{k_1, k_2\}$ *satisfying* $k_1 + k_2 = d-1$*, where* $1 \le k_1 \le d-2$*, then* $\mathbf{g}^H$ *is the zero-forcing channel equalizer of delay* k_1*.*

Proof: For $k_1 + k_2 = d-1$, we have

$$\begin{aligned}
\mathbf{F}_{k_1,k_2} &\triangleq \Gamma_{k_1}\Gamma_{-k_2} \\
&= \mathcal{H}\left(\hat{\mathbf{I}}_{d-k_1} + \sum_{i=1}^{k_1}\mathbf{e}_i\mathbf{a}_{k_1,i}^H\right) \times \left(\breve{\mathbf{I}}_{d-k_2} + \sum_{j=d-k_2+1}^{d}\mathbf{e}_j\mathbf{b}_{k_2,j}^H\right)\mathcal{H}^\dagger \\
&= \mathcal{H}\left(\mathbf{e}_{k_1+1}\mathbf{e}_{k_1+1}^H + \sum_{i=k_1+2}^{d}\mathbf{e}_i\mathbf{b}_{k_2,i}^H + \sum_{i=1}^{k_1}\mathbf{e}_i\mathbf{f}_{k_1,i}^H\right)\mathcal{H}^\dagger
\end{aligned} \tag{6.31}$$

where

$$\mathbf{f}_{k_1,i}^H \triangleq \mathbf{a}_{k_1,i}^H\left(\breve{\mathbf{I}}_{d-k_2} + \sum_{j=d-k_2+1}^{d}\mathbf{e}_j\mathbf{b}_{k_2,j}^H\right).$$

Again, the proof is evident from Theorem 6.1 if we let $\mathbf{c}_i = \mathbf{f}_{k_1,i}$ for $i \in \{1, \cdots, k_1\}$ and $\mathbf{c}_i = \mathbf{b}_{k_2,i}$ for $i \in \{k_1 + 2, \cdots, d\}$. ■

Up to now, we have shown that the zero-forcing equalizers with extreme delays can be estimated directly from Γ_k and Γ_{-k} for $k = d - 1$, and the zero-forcing equalizers with intermediate delays can be estimated from $\mathbf{F}_{k_1,k_2} = \Gamma_{k_1}\Gamma_{-k_2}$ for $k_1 + k_2 = d - 1$. As an important note, from the above, all equalizers of desired delays are estimated independently of those of other delays. The solution to $\mathbf{g}^H\Gamma_k = \mathbf{g}^H$ or $\mathbf{g}^H\mathbf{F}_{k_1,k_2} = \mathbf{g}^H$ admits a closed-form and can be solved as the left singular vector associated with the smallest singular value of matrix $(\Gamma_k - \mathbf{I})$ or $(\mathbf{F}_{k_1,k_2} - \mathbf{I})$. Of course, all these results are valid under the assumption that the so-called *equalizability condition* is satisfied, i.e. $\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_1 - \mathbf{I} \\ \mathbf{C}_2 \end{bmatrix}$ is full column rank, where $\mathbf{C}_1$ and $\mathbf{C}_2$ are constructed from the vectors $\mathbf{c}_i$ for $S^c = \{1, \cdots, m-1, m+1, \cdots, d\}$. $\mathbf{C}_1$ is a $(d-1) \times (d-1)$ square matrix and $\mathbf{C}_2$ turns into a $1 \times (d-1)$ row vector. In the following, we discuss this crucial issue of equalizability condition and ascertain its implications.

6.3.3 Equalizability Condition and Relation with Other Work

As mentioned above, the equalizability condition requires that the matrix $\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_1 - \mathbf{I} \\ \mathbf{C}_2 \end{bmatrix}$ constructed by the vectors $\mathbf{c}_i$ for $i \in S^c = \{1, \cdots, m-1, m+1, \cdots, d\}$ is full column rank, where $\bar{\mathbf{C}}$ is a $d \times (d-1)$ matrix. From Lemma 6.2 and 6.3, it can be seen that the entries in $\bar{\mathbf{C}}$ are determined by the second-order statistics of the colored input signals. Hence the equalizability is only related to its "colorness" or power spectrum of the source. In fact, the equalizability condition described above is not very restrictive and can be met by most col-

ored sources. The reason can be explained as follows. The vectors $\mathbf{c}_i$ for $i \in S^c$ used to construct the matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ are all explicitly expressed by the coefficients of the multistep d^{th}-order optimum forward/backward prediction error filters for the process $\{s(n)\}$. For example, in Lemma 6.2, the vectors $\mathbf{c}_i$ for $i \in S^c$ are equal to $\mathbf{a}_{d-1,i}$ ($\mathbf{b}_{d-1,i}$), where $\mathbf{a}_{d-1,i}$ ($\mathbf{b}_{d-1,i}$) are functions of the coefficients of the multistep d^{th}-order optimum forward/backward prediction error filters. Generally, for most colored sources, the coefficients of the multistep optimum forward/backward prediction error filters are small and scarcely greater than unity. Accordingly, the entries in the vectors $\mathbf{c}_i$ for $i \in S^c$ are also of small values. Thus, $\mathbf{C}_1 - \mathbf{I}$ is a matrix whose diagonal elements are close to minus one and whose off-diagonal elements are relatively small. Such a matrix scarcely collapses into a rank-deficient matrix. Hence $\bar{\mathbf{C}}$ is full column rank and the equalizability condition can be assured for most colored sources. For an illustrative example, for the colored source in Section 6.4 Example A of this chapter, we have

$$\mathbf{C}^H = \begin{bmatrix} & & & \mathbf{0}_{5\times 7} & & & \\ 0.09i & 0.19 & -0.28i & -0.37 & 0.47i & 0.56 & 0.22i \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

for $d = 7$ and $S = \{7\}$. It can be easily verified that $\mathbf{C}_1 - \mathbf{I}$ is full rank, where

$$\mathbf{C}_1^H = \begin{bmatrix} & & & \mathbf{0}_{5\times 6} & & \\ 0.09i & 0.19 & -0.28i & -0.37 & 0.47i & 0.56 \end{bmatrix}.$$

Also, for the colored source in Section 6.4 Example B of this chapter, we have

$$\mathbf{C}^H \approx \begin{bmatrix} 0 & 0 & 0 & 0 & -0.04i \\ 0 & 0 & 0 & 0 & 0.05-0.03i \\ 0 & 0 & 0 & 0 & 0.13-0.04i \\ 0 & 0 & 0 & 0 & 0.35-0.02i \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

for $d = 5$ and $S = \{5\}$. It can also be verified that $\mathbf{C}_1 - \mathbf{I}$ is full rank, where $\mathbf{C}_1^H = \mathbf{0}_{4\times 4}$.

If the input signals are white, then we have $\mathbf{c}_i = \mathbf{0}$ for each $i \in S^c$, i.e. $i \neq m$. Hence the matrices $\mathbf{C}_1 = \mathbf{0}$ and $\mathbf{C}_2 = \mathbf{0}$. It is clear that the equalizability condition is automatically satisfied in this case. This implies that our proposed algorithm devised for colored sources also applies to the white signals. In fact, in the case where the input signals are white, the matrices Γ_k, Γ_{-k} for $k = d-1$ and $\mathbf{F}_{k_1,k_2}$ for $k_1 + k_2 = d-1$ are all reduced into rank-one matrices $\mathbf{A}_m \triangleq \mathcal{H}\mathbf{e}_m\mathbf{e}_m^H\mathcal{H}^\dagger = \mathcal{H}[:,m]\mathcal{H}^\dagger[m,:]$. Clearly, we can estimate the mth column of $\mathcal{H}$ and the mth row of $\mathcal{H}^\dagger$ as the principal left singular vector and right singular vector of these rank-one matrices respectively. Thus the channel impulse response and the zero-forcing equalizers are obtained simultaneously by computing the singular value decomposition (SVD) of $\mathbf{A}_m$. On the other hand, it is easy to prove that for any vector $\mathbf{g}$ that satisfies $\mathbf{g}^H\mathbf{A}_m = \mathbf{g}^H$, then $\mathbf{g}^H$ is the zero-forcing equalizer of delay $(m-1)$; for any vector $\mathbf{g}$ that satisfies $\mathbf{A}_m\mathbf{g} = \mathbf{g}$, then $\mathbf{g}$ is the mth column of $\mathcal{H}$. These are also the main results of work reported in [1]. As we can see, these special rank-one matrices have useful structure facilitating us to identify the channel and estimate the zero-forcing equalizers simultaneously. However, when the input signals are colored,

the matrices Γ_k, Γ_{-k} and $\mathbf{F}_{k_1,k_2}$ degenerate into a more general form and the structure does not hold any more. Nevertheless, it is shown that under certain equalizability conditions satisfied, the zero-forcing equalizers can still be estimated from $\mathbf{g}^H\mathbf{F}_{k_1,k_2} = \mathbf{g}^H$ or $\mathbf{g}^H\Gamma_k(\Gamma_{-k}) = \mathbf{g}^H$. Therefore, our work can be considered as a generalization or extension of work [1] to the colored sources. It is noted that for the colored sources, we can not estimate the channel from $\mathbf{F}_{k_1,k_2}\mathbf{g} = \mathbf{g}$ or $\Gamma_k(\Gamma_{-k})\mathbf{g} = \mathbf{g}$ although this is valid for the white input signals. The reason lies in that the special structure of rank-one matrices does not hold for the colored signals.

6.3.4 Channel Estimation

The identification of the channel for the colored sources is more complicated since the relationship $E[s(n-m)\vec{\mathbf{x}}^H(n)] = \mathcal{H}^H[m+1,:]$ which is only valid for the white inputs does not hold any more. Thus we can not expect to estimate the channel from $(\mathbf{g}_{\text{zf}}^m)^H\mathbf{R}_x[0]$, where $\mathbf{g}_{\text{zf}}^m$ denotes the zero-forcing equalizer of delay m. One possible way to identify the channel needs to use the zero-forcing equalizers of different delays from 0 to $d-1$. Let $\mathbf{G}_{\text{zf}} \triangleq [\mathbf{g}_{\text{zf}}^0 \;\; \mathbf{g}_{\text{zf}}^1 \;\; \cdots \;\; \mathbf{g}_{\text{zf}}^{d-1}]$, we have $\mathbf{G}_{\text{zf}}^H\mathcal{H} = \mathbf{D}$ and $\mathcal{H}^H\mathbf{G}_{\text{zf}} = \mathbf{D}^H$, where $\mathbf{D}$ is an unknown diagonal matrix (Note that $\mathbf{D} \neq \mathbf{I}$ because there exists a scalar ambiguity between the estimated zero-forcing equalizer and the ideal zero-forcing equalizer). We obtain

$$\mathbf{D}\mathbf{R}_s[0]\mathbf{D}^H = \mathbf{G}_{\text{zf}}^H\mathbf{R}_x[0]\mathbf{G}_{\text{zf}}. \tag{6.32}$$

On the other hand, we have

$$\mathbf{R}_x[0]\mathbf{G}_{\text{zf}} = \mathcal{H}\mathbf{R}_s[0]\mathbf{D}^H. \tag{6.33}$$

Hence the channel convolution matrix $\mathcal{H}$ can be estimated up to an unknown diagonal matrix as follows

$$\begin{aligned}\hat{\mathcal{H}} &= \mathbf{R}_x[0]\mathbf{G}_{\mathrm{zf}}(\mathbf{G}_{\mathrm{zf}}^H\mathbf{R}_x[0]\mathbf{G}_{\mathrm{zf}})^{-1} \\ &= \mathcal{H}\mathbf{R}_s[0]\mathbf{D}^H(\mathbf{D}\mathbf{R}_s[0]\mathbf{D}^H)^{-1} \\ &= \mathcal{H}\mathbf{D}^{-1}. \end{aligned} \tag{6.34}$$

Observe that for $N \geq L$, each column of $\hat{\mathcal{H}}$, except the first L and the last L columns, contains the estimated entire channel impulse response.

6.3.5 Noise Compensation and MMSE Equalizers

In our previous presentation of this chapter, we have ignored the noise effect in order to simplify our presentation. In practice, the influence of the noise can be minimized by removing the noise contribution from the autocorrelation matrices. Since the additive noises are assumed spatially and temporally white with same variance, we have

$$\mathbf{R}_x[0] = \mathcal{H}\mathbf{R}_s[0]\mathcal{H}^H + \sigma_n^2\mathbf{I}.$$

The noise variance σ_n^2 can be estimated as the smallest $(N+1)\times q-(N+L+1)$ eigenvalues of $\mathbf{R}_x[0]$, in which the channel order L is assumed known or can be estimated by applying the MDL criterion [51]. After the noise variance σ_n^2 is estimated, we can subtract the noise effect from any estimated autocorrelation matrix $\mathbf{R}_x[k]$.

We now study how to derive the MMSE equalizers from the estimated ZF equalizers. This is meaningful because equalization using zero-forcing equalizers

results in noise enhancement in some channels. The MMSE equalizers are estimated by the following MMSE criterion

$$\mathbf{g}_{\text{mmse}}^{m} = \arg\min_{\mathbf{g}} E[|s(n-m) - \mathbf{g}^H \vec{\mathbf{x}}(n)|^2] \tag{6.35}$$

where $\mathbf{g}_{\text{mmse}}^{m}$ is the sought-after MMSE equalizer of delay m. The solution to the criterion Eqn.(6.35) is given as

$$\begin{aligned}(\mathbf{g}_{\text{mmse}}^{m})^H &= E[s(n-m)\vec{\mathbf{x}}^H(n)](E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^H(n)])^{-1} \\ &= E[s(n-m)\vec{\mathbf{x}}^H(n)]\mathbf{R}_x^{-1}[0]. \end{aligned} \tag{6.36}$$

Although, when the input signals are colored, the relationship $E[s(n-m)\vec{\mathbf{x}}^H(n)] = \mathcal{H}^H[m+1,:]$ does not hold, we still have

$$E[s(n-m)\vec{\mathbf{x}}^H(n)] = (\mathbf{g}_{\text{zf}}^{m})^H(\mathbf{R}_x[0] - \sigma_n^2\mathbf{I}) \tag{6.37}$$

where $\mathbf{g}_{\text{zf}}^{m}$ is the zero-forcing equalizer with delay m. From Eqn.(6.36–6.37), we obtain

$$\begin{aligned}(\mathbf{g}_{\text{mmse}}^{m})^H &= (\mathbf{g}_{\text{zf}}^{m})^H(\mathbf{R}_x[0] - \sigma_n^2\mathbf{I})\mathbf{R}_x^{-1}[0] \\ &= (\mathbf{g}_{\text{zf}}^{m})^H(\mathbf{I} - \sigma_n^2\mathbf{R}_x^{-1}[0]). \end{aligned} \tag{6.38}$$

Thus we have established the relationship between the ZF equalizer and MMSE equalizer, and we can now obtain the MMSE equalizers from our estimated ZF equalizers. It should be noted that here $\mathbf{R}_x[0]$ is not denoised in computing the MMSE equalizers and, hence, its inverse is not a pseudo-inverse. This is different from our previous parts that use the denoised autocorrelation matrices to compute the ZF equalizers, and, consequently, the inverse of $\mathbf{R}_x[0]$ is a

pseudo-inverse.

6.3.6 Discussions

It can be seen that our proposed method does not impose/exploit any structure on the channel convolution matrix $\mathcal{H}$, which is fundamentally different from the classical subspace method [22]. This characteristic makes our method applicable to the case where the channel convolution matrix has no special structure. For example, when the environment is multipath-dominated and the receiver diversity is of high dimension, the stack number N can thus be chosen as zero and the channel convolution matrix $\mathcal{H}$ degenerates into the following form: $\mathcal{H} = [\mathbf{h}(0)\ \mathbf{h}(1)\ \cdots\ \mathbf{h}(L)]$ (Note that the receiver diversity is sufficiently high to make $\mathcal{H}$ full column rank). Our proposed method can be applied to this case without any change, whereas the classical subspace method [22] which relies on the structure of the channel convolution matrix fails.

6.4 Simulation Results

We now present simulation results to illustrate the performance of our proposed algorithm. We compare our method to the classical subspace (SS) method [22], the mutually referenced equalizers (MRE) method [49], and the linear prediction method for colored sources (LPC) proposed in [67]. It is noted that the LPC method requires the knowledge of the second-order statistics of the colored sources. In our simulations, the performance is measured by the NRMSE of the channel estimate and the SER of the estimated data symbols. The NRMSE

of the user's channel estimate is defined as

$$\text{NRMSE} = \sqrt{\frac{1}{N_{mc}} \sum_{t=1}^{N_{mc}} \frac{\|\rho(t)\hat{\mathbf{h}}^{(t)} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2}}$$

where N_{mc} is the number of Monte Carlo runs, $\mathbf{h} \triangleq \begin{bmatrix}\mathbf{h}(0)^T & \cdots & \mathbf{h}(L)^T\end{bmatrix}^T$, $\rho(t)$ is a complex scalar that minimizes $\|\rho(t)\hat{\mathbf{h}}^{(t)} - \mathbf{h}\|^2$. Also, in our simulations, the additive noise $\mathbf{w}(n)$ is taken as spatial-temporal white Gaussian noise with variance σ_n^2. The SNR is defined as

$$\text{SNR} = 10 \cdot \log \frac{E[\|\mathcal{H}\vec{\mathbf{s}}(n)\|^2]}{E[\|\vec{\mathbf{w}}(n)\|^2]}.$$

Two examples are considered to show the performance of our proposed channel equalization and identification algorithms.

6.4.1 Example One

In this example, we generate the colored source by following the same manner as in [67]. The input signal $\{s(n)\}$ draws symbols from a 4-QAM constellation $\mathcal{S} = \{-1-i, -1+i, 1-i, 1+i\}$ according to the following rule

$$s(n) = \begin{cases} -1+i & \text{if } (a(n) \;\; a(n-1)) = (0 \;\; 0) \\ +1+i & \text{if } (a(n) \;\; a(n-1)) = (0 \;\; 1) \\ -1-i & \text{if } (a(n) \;\; a(n-1)) = (1 \;\; 0) \\ +1-i & \text{if } (a(n) \;\; a(n-1)) = (1 \;\; 1) \end{cases}$$

where $\{a(n)\}$ are the digital i.i.d. sequences and $a(n) \in \{0, 1\}$. This makes the input symbols $\{s(n)\}$ to be colored with autocorrelation as

$$E[s(n)s^*(n-k)] = \begin{cases} 2 & k = 0 \\ \pm i & k = \pm 1 \\ 0 & \text{else.} \end{cases}$$

We consider a wireless communication scenario with the colored source signals arriving at a single sensor via a multipath channel. The channel impulse response is given by

$$h(t) = (0.8c(t, 0.11) - 0.5c(t - 0.5T, 0.11))W_{4T}(t)$$

where $c(t - t_0, \theta)$ denotes a raised cosine pulse with roll-off factor θ and delay t_0, $W_{4T}(t)$ is a square window of duration 4 symbols intervals. The channel is non-minimum phase and introduces significant intersymbol interference (ISI). We sample the received data two times the symbol rate, thus generating a single-input two-output linear system with channel order equal to 3.

In our simulations, the channel order is assumed known *a priori* and the stack number (smoothed factor) N is chosen to be 3. The scalar ambiguity of the estimated equalizers is removed before we perform channel equalization. Results are averaged over 500 Monte Carlo runs. We, firstly, illustrate the channel equalization performance of our proposed algorithm. Figure 6.1 shows the variation of SER with SNR for the MMSE equalizers with delays 0, 2, 4, 6 using $T_s = 1000$ data samples. Clearly, we can see that, the equalizers of intermediate delays are superior to those of extremal delays in performance. The SER of the MMSE equalizer with $d_e = 2$ using different number of data samples is displayed in Figure 6.2. As expected, the performance improves with an increasing

T_s. Figure 6.3 depicts the NRMSE of the estimated channel impulse response obtained using the identification algorithm proposed in Section 6.3.4. In Figure 6.4, we show the SER of the MMSE equalizers with best delays as the stack number N varies from 3 to 7. $T_s = 1000$ data samples are used to estimate the signal statistics and autocorrelation matrices. It can be seen that the SER decreases with an increasing N. This improved performance results from the "lengthened" length of the equalizers as N increases. We now compare our proposed method to the other three methods. The performance comparison is shown in Figure 6.5, where the MMSE equalizer with $d_e = 2$ is used for all methods, and the stack number N is chosen to be 3. We can see that, in this case, the LPC method achieves the best performance and the subspace method follows after the LPC method, while our proposed method performs similarly as the MRE method but not so well as the LPC and SS methods. The possible reason for the lack of performance improvement of our proposed algorithm in this example can be explained as follows. Note that the equalizers of intermediate delays are directly estimated from the matrix $\mathbf{F}_{k_1,k_2}$ (see Lemma 6.3), where $\mathbf{F}_{k_1,k_2}$ is constructed from the multiple matrix products of the estimated autocorrelation matrices. These multiple matrix products of the estimated autocorrelation matrices result in noise enhancement and thus deteriorate the performance, especially when the coefficients of the multistep optimum forward/backward prediction error filters are not negligible, which is exactly the case for this example, e.g. we have $\vec{\alpha}_{1,1}^H = [0.88i \;\; 0.75 \;\; -0.63i \;\; -0.50 \;\; 0.38i \;\; 0.25 \;\; -0.13i]$.

6.4.2 Example Two

We adopt another kind of colored source in this example. The digital message sequence $\{s(n)\}$ is generated from the 4-QAM constellation $\mathcal{S} = \{-1 -$

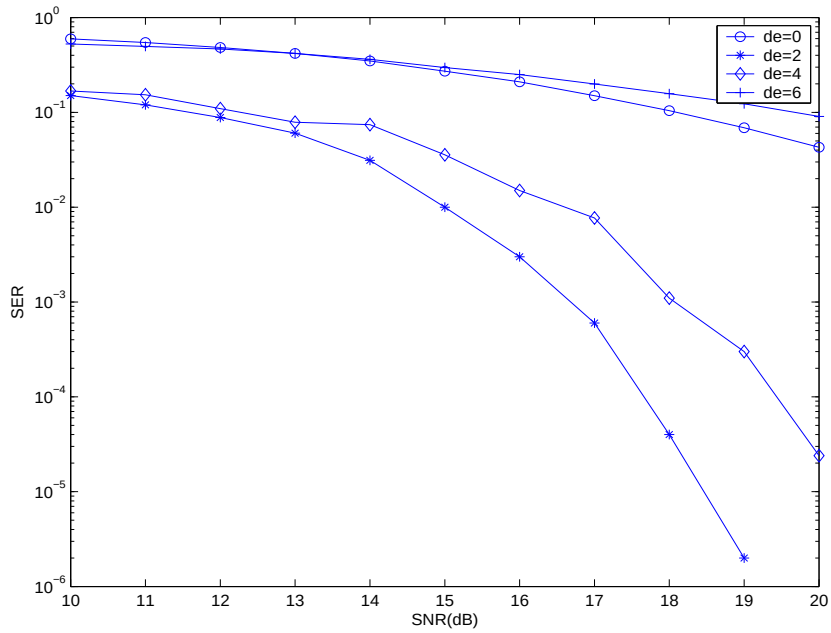


Figure 6.1: Example 1: SER versus SNR for the MMSE equalizers with different delays; $T_s = 1000$.

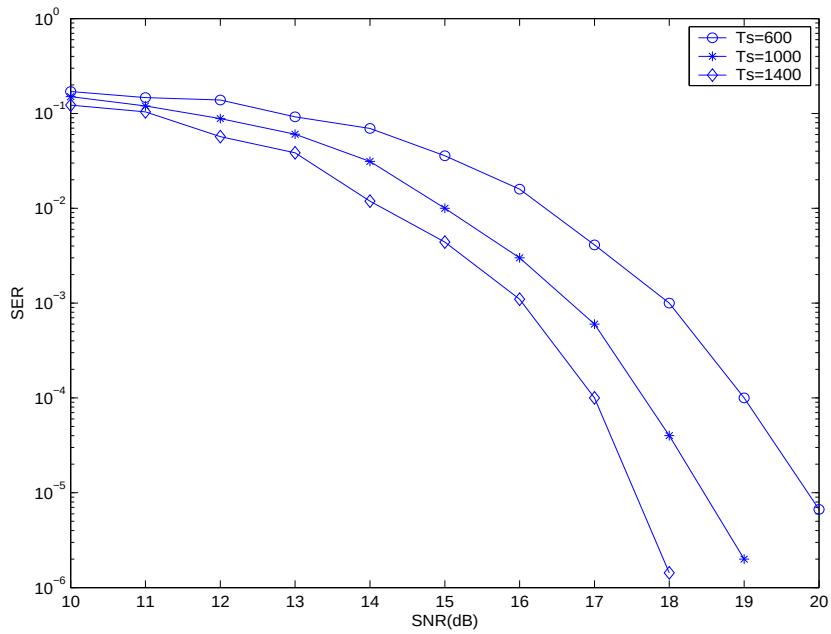


Figure 6.2: Example 1: SER of the MMSE equalizer with $d_e = 2$ versus SNR using different number of data samples.

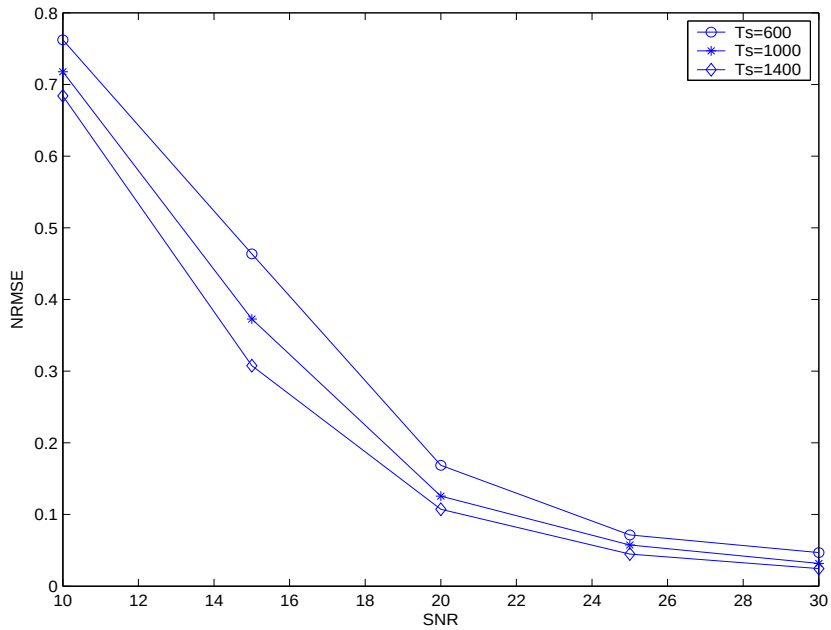


Figure 6.3: Example 1: NRMSE of the estimated channel versus SNR using different number of data samples.

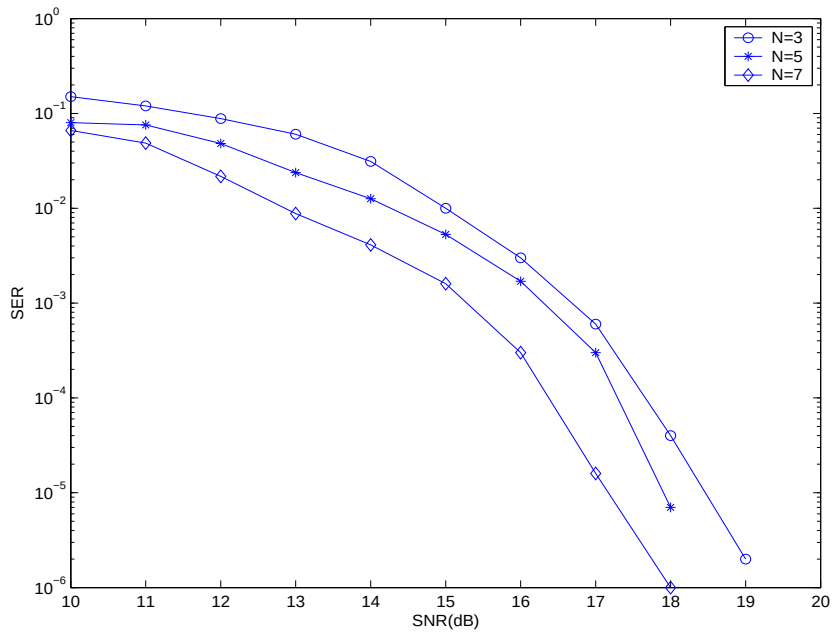


Figure 6.4: Example 1: SER of the MMSE equalizers with best delays versus SNR as the stack number varies from 3 to 7, $T_s = 1000$.

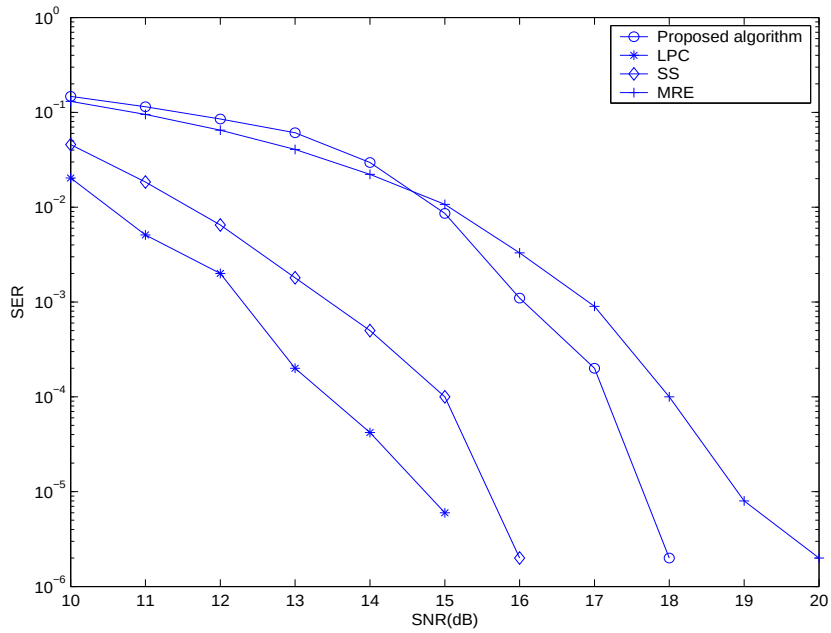


Figure 6.5: Example 1: SER of the respective algorithms versus SNR, $d_e = 2$, $T_s = 1000$.

$p(s_k \mid s_{k-1})$	$s_k = -1-i$	$s_k = -1+i$	$s_k = 1-i$	$s_k = 1+i$
$s_{k-1} = -1-i$	0.5	0.3	0.1	0.1
$s_{k-1} = -1+i$	0.2	0.4	0.2	0.2
$s_{k-1} = 1-i$	0.2	0.1	0.4	0.3
$s_{k-1} = 1+i$	0.1	0.2	0.2	0.5

Table 6.1: Transition probabilities for Markov source

$r(0)$	2.0000	$r(5)$	0.0173	$r(10)$	0.0085	$r(15)$	0.0020
$r(1)$	0.6584	$r(6)$	0.0088	$r(11)$	0.0095	$r(16)$	0.0047
$r(2)$	0.2343	$r(7)$	0.0017	$r(12)$	-0.0029	$r(17)$	0.0047
$r(3)$	0.0893	$r(8)$	0.0053	$r(13)$	-0.0068	$r(18)$	0.0027
$r(4)$	0.0384	$r(9)$	0.0082	$r(14)$	0.0031	$r(19)$	0.0039

Table 6.2: Autocorrelation function of the Markov source

$i, -1+i, 1-i, 1+i\}$ according to the following model. The model simulates a *Markov* source by implementing the transition probabilities of Table 6.1. The autocorrelation function of this source is given in Table 6.2. Consider this col-

ored source arriving at a uniform linear array of $q = 6$ sensors via a frequency selective multipath channel. The array elements are spaced half a wavelength apart. The signalling pulse shape for the user is a raised cosine pulse with a roll-off factor of 0.11, the pulse being truncated to a length of four symbols duration ($4T$). A three-ray multipath is assumed, with attenuation factors $[0.5 \ \ -0.7 \ \ 0.6]$, delays $[0.8T \ \ 1.0T \ \ 1.2T]$ and angles of arrival $[30^\circ \ \ 60^\circ \ \ 110^\circ]$. The array measurements are sampled at baud rate. In our simulations, the channel order $L = 3$ is assumed known *a priori* and the stack number N is chosen to be 1. The MMSE equalizer with equalization delay, d_e, equal to 2 is used. Results are averaged over 500 Monte Carlo runs. We present the equalization performance of the respective algorithms in Figure 6.6, in which the SER is shown as a function of SNR with $T_s = 800$. Figure 6.7 shows the variation of SER with the number of data samples T_s for SNR = 15dB. From Figures 6.6 and 6.7, we can see that our proposed algorithm presents the best performance among all these algorithms, even better than the LPC method which utilizes the knowledge of the input statistics. As compared with the previous example, our proposed algorithm shows a significant performance improvement. This is because, in this case, the coefficients of the multistep optimum forward/backward prediction error filters are relatively small, e.g. $\vec{\alpha}_{1,1}^H = [0.32-0.04i \ \ 0.01+0.013i \ \ 0.012+0.015i \ \ -0.005-0.004i \ \ 0.023+0.020i]$, and thus the noise enhancement caused by the multiple matrix products is negligible.

To study the performance of respective algorithms under various channel conditions, we conduct simulation tests using randomly generated channel where each subchannel $\{h_i(l)\}$ is a complex zero-mean Gaussian process, where $\{h_i(l)\}$ denotes the subchannel from the colored source to the ith antenna. We assume

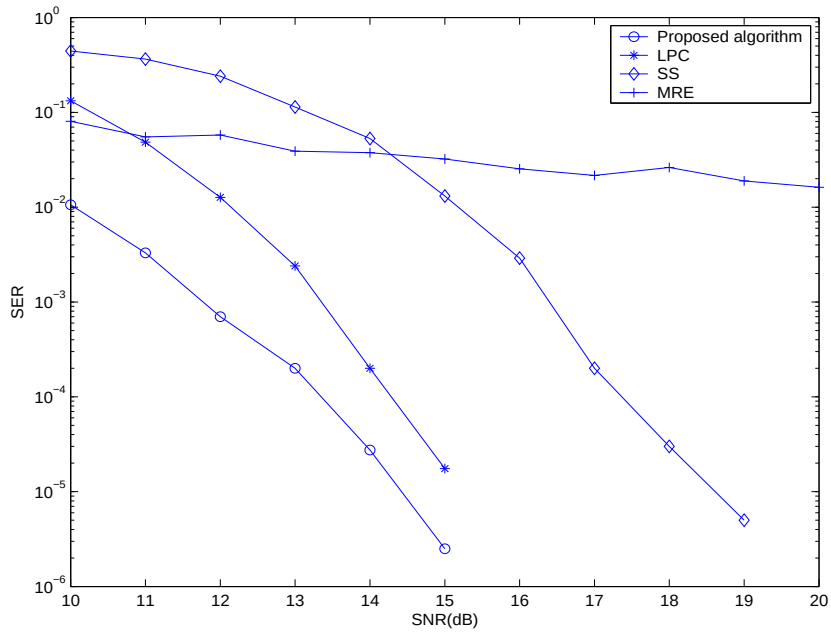


Figure 6.6: Example 2: SER of the respective algorithms versus SNR, $d_e = 2$, $T_s = 800$.

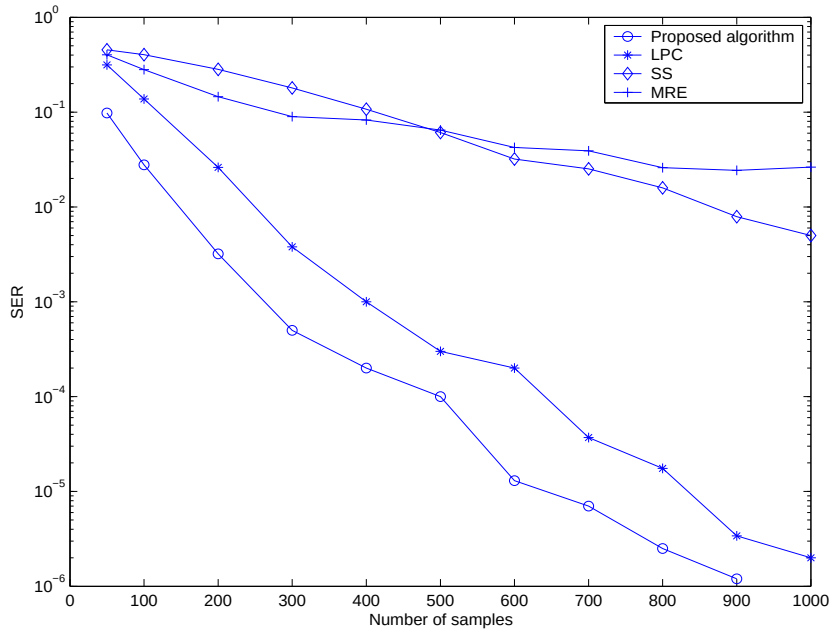
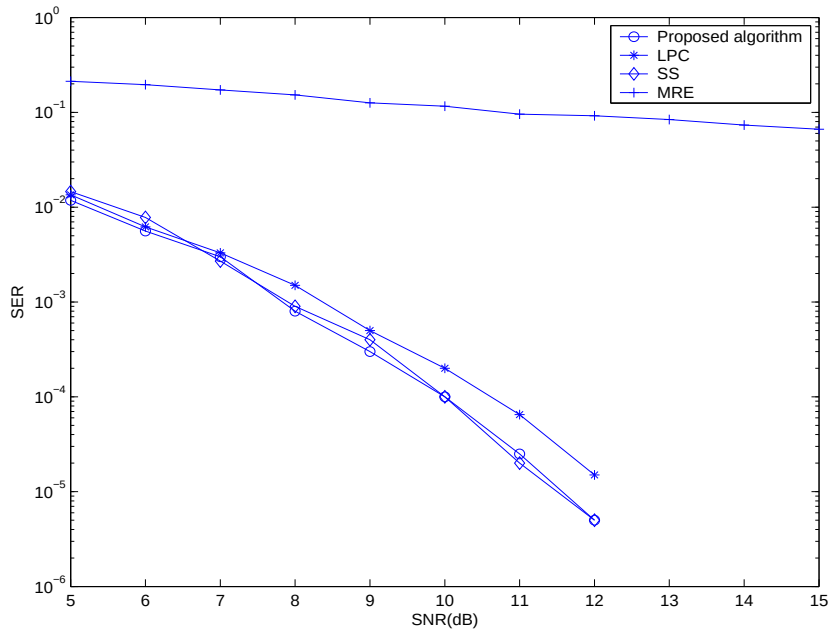
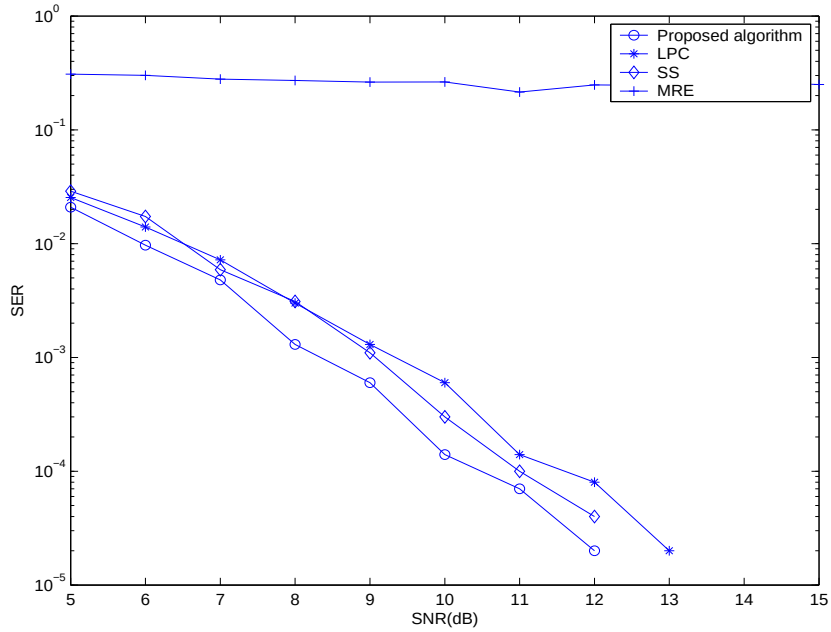


Figure 6.7: Example 2: SER of the respective algorithms versus SNR using different number of data samples, $d_e = 2$, SNR = 15dB.



(a) $T_s = 400$



(b) $T_s = 100$

Figure 6.8: Example 2: SER of the respective algorithms versus SNR, $d_e = 2$.

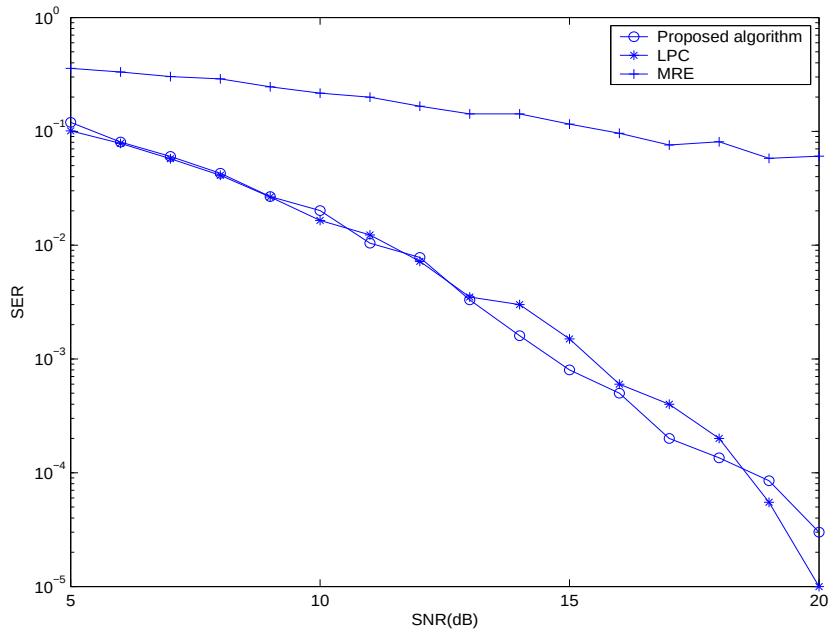


Figure 6.9: Example 2: SER of the respective algorithms versus SNR, $d_e = 1$, $T_s = 400$.

$q = 6$ antennas. Therefore there are totally six subchannels and all subchannels are generated independently of each other. The channel order L is chosen to be 3. We select $N = 1$ to compare our proposed algorithm with other methods. For the respective algorithms, the MMSE equalizer with $d_e = 2$ is used. The equalization performance of the respective algorithms is presented in Figure 6.8, in which we show the SER as a function of SNR with the number of samples used to estimate signal statistics equal to 400 and 100, respectively. We can see that our proposed algorithm performs similarly as LPC and SS methods and slightly better than these two methods at low SNR. Moreover, for our proposed algorithm, the stack number can be chosen to be zero. In this case, the proposed algorithm is used to invert the channel with baud-length equalizers. The performance of the baud-length equalizer with $d_e = 1$ is shown in Figure 6.9 (note that the subspace method no longer applies to this case). The perfor-

mance degradation of this case as compared to the case where $N = 1$ mainly results from the "shortened" length of the equalizers.

6.5 Summary

In this chapter, we presented a new SOS-based method for blind equalization of SIMO FIR channel driven by colored sources whose statistics are unknown *a priori*. It has been shown that, even if without utilizing the knowledge of the input signals' statistics and exploiting the structure of the channel convolution matrix, the closed-form ZF and MMSE equalizers of desired delays can still be estimated from the second-order statistics of the received data by exploring the inherent structure relationship between source autocorrelation matrices of different delay lags. The estimation of these equalizers is direct and does not require to identify the channel impulse response in advance. The equalizers of any desired delays are estimated independently of those of other delays. Simulation results showed that our proposed method outperforms the other existing methods [22, 49, 67] clearly for the colored sources whose coefficients of the multistep optimum forward/backward prediction error filters are small. The extension of our proposed method to multiple-input multiple-output systems is under investigation.

Chapter 7

Further Studies on MIMO FIR Channel Identification

Our previous work (Chapter 5) studies the problem of blind identification of MIMO FIR channel driven by spatially uncorrelated input signals. In this chapter, we consider the problem of blind MIMO FIR channel identification driven by spatially correlated signals. The SOS of the input sources are assumed known *a priori*. It is shown that under certain specified conditions, the MIMO FIR channel can be completely identified using the second-order statistics of the channel output. A SOS-based method is proposed and the proof for the uniqueness of the system solution is provided. As a special case, our proposed method can still entirely identify the MIMO channel even if the input source signals are spatially and temporally uncorrelated, given that the channel orders corresponding to each pair of users are different from each other. Extensive numerical simulation results are included to illustrate the performance of the proposed algorithm.

7.1 Introduction

Blind identification of MIMO FIR channel arises in a wide variety of communication and signal processing applications, which include speech enhancement, wireless mobile communications and brain signal analysis. Thus far, numerous SOS-based techniques [24, 29, 32, 34, 35, 70, 71] have been proposed within such a framework. Nevertheless, they usually assume that the input sources are mutually independent or, at least, uncorrelated (note that this assumption is also adopted in Chapter 5 of our thesis). In contrast, blind channel estimation driven by spatially correlated sources has not received much attention. Spatially correlated sources may indeed occur in practice. For example, the nonlinear SIMO channels can be reformulated into multiple-input linear systems in which the additional inputs are nonlinear functions of the signal of interest (the details of this reformulation can be referred to [100]). Clearly, in this case, the inputs of this reformulated linear MIMO channel may be correlated with each other. Specifically, this reformulated MIMO channel can be written as follows

$$\mathbf{x}(n) = \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n) \tag{7.1}$$

where $s_1(n) \triangleq a(n)$ is exactly the input signal to the nonlinear channels and also called as "linear kernel"; the terms $s_i(n) = f_i(a(n), a(n-1), \ldots)$ for $i \in \{2, \ldots, p\}$ are nonlinear functions of $a(\cdot)$ and also called as "nonlinear kernels"; $\{\mathbf{h}_i(l)\}$ are $q \times 1$ multichannel vectors; $\mathbf{x}(n)$ and $\mathbf{w}(n)$ represent the received data and the additive noise, respectively. In this case, if the statistical information of the input signal $a(n)$ and the functions $f_i(\cdot)$ generating the nonlinearities are known *a priori*, then the second-order statistics of the reformulated inputs to the MIMO channel are available. It is noted that both [100] and [101] were

presented under such a framework described by Eqn.(7.1). In [100], the authors proposed a deterministic method that exploited the channel order disparity between the linear kernel and nonlinear kernels. In fact, the techniques of [100] only resolve the kernel which has the largest channel order, irrespective of whether this kernel is a linear kernel or a nonlinear one. In the event that there are many kernels with maximum channel order, [100] has to resort to higher order methods to equalize the channel. On the other hand, the work in [101] shows that the linear kernel is resolvable under the right conditions imposed on the statistics of the signals $a(n)$ and $s_i(n)$, without the need to adhere to the particular channel order condition required by [100]. A SOS-based approach was put forward to determine the equalizers for the i.i.d. input signals $\{a(n)\}$.

In this chapter, we address the problem of blind identification of MIMO FIR channel driven by spatially correlated signals. A closed-form solution is proposed for blind MIMO FIR system identification by utilizing the estimated channel output autocorrelation matrices and the knowledge of the source autocorrelation matrices. As compared to [100] and [101], the problem we address here is more general because our goal is to identify the entire MIMO channel and recover all the source signals, which is different from the works [100, 101] that focus on extracting and equalizing only one source signals. It is noted that, in our case, the terms $s_i(n)$ for $i \in \{2, \ldots, p\}$ in Eqn.(7.1) may not necessarily be the functions of $s_1(n)$. In fact, even if we would only consider the MIMO models resulted from the nonlinear SIMO channels, our work has its own advantages over [100] and [101] in the following two aspects. Firstly, the particular channel order condition required by [100] is no longer necessary for our proposed method to identify and equalize the channel. Secondly, unlike the proposed algorithm in [101] which is only specified for the i.i.d. input signals

$\{a(n)\}$, our proposed method applies to both i.i.d. and colored input signals $\{a(n)\}$. As a special case of our work, our proposed method can entirely identify the MIMO channel even when the input source signals are spatially and temporally uncorrelated, given that the channel orders corresponding to each pair of users are different. This result is different from most existing SOS-based methods that can only identify such a channel up to an unknown unitary matrix.

This chapter is organized as follows. In Section 7.2, the MIMO system model and some basic assumptions are introduced. Next, in Section 7.3, we present our method for blind channel identification driven by spatially correlated sources. The channel identifiability conditions are investigated and an original proof for the uniqueness of the system solution is provided. In Section 7.4, we extend our method to the case of spatially and temporally uncorrelated input sources. Finally, in Section 7.5, numerical simulation results are presented to demonstrate the performance of the proposed algorithm.

7.2 System Model and Basic Assumptions

The linear MIMO channel with p inputs, $s_i(n), i \in \{1, 2, \cdots, p\}$, and q outputs $\mathbf{x}(n) \triangleq [x_1(n) \ \cdots \ x_q(n)]$ is the same as the model in Chapter 5:

$$\mathbf{x}(n) = \sum_{i=1}^{p} \sum_{l=0}^{L_i} \mathbf{h}_i(l) s_i(n-l) + \mathbf{w}(n). \tag{7.2}$$

Its corresponding matrix form is written as (see Chapter 2 or Chapter 5 for details)

$$\vec{\mathbf{x}}(n) = \sum_{i=1}^{p} \mathcal{H}_i \vec{\mathbf{s}}_i(n) + \vec{\mathbf{w}}(n) = \mathcal{H}\vec{\mathbf{s}}(n) + \vec{\mathbf{w}}(n). \tag{7.3}$$

Some basic assumptions are adopted throughout this chapter: A1) The number of sources is known *a priori*, and there are more outputs than inputs, i.e. $q > p$. A2) Channel is irreducible and column-reduced. A3) The channel order corresponding to each source is assumed to be known *a priori*. A4) The sources are zero-mean wide-sense stationary colored signals or white signals and their input statistics (include autocorrelation of the single source signals and cross-correlation between any two source signals) are available. A5) The source correlation matrix $\mathbf{R}_s[0]$ is positive definite, where $\mathbf{R}_s[0] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n)]$. A6) Additive noises $\mathbf{w}(n)$ are spatially and temporally white noises with same variance, and they are statistically independent of the sources. As a consequence of A1–A2, the MIMO channel matrix $\mathcal{H}$ is full column rank if the stack number N is chosen to satisfy $N + 1 \geq \sum_{i=1}^{p} L_i$ (see [27]). In the sequel, we assume that $\mathcal{H}$ is full column rank.

7.3 Proposed Channel Identification Method for Spatially Correlated Sources

We begin by defining the source autocorrelation matrices with delay lag k as follows

$$\mathbf{R}_s[k] \triangleq E[\vec{\mathbf{s}}(n)\vec{\mathbf{s}}^H(n-k)]. \tag{7.4}$$

Also, in order to simplify the presentation of the proposed channel identification method, we consider the noiseless case. In fact, since the additive noises are assumed spatially and temporally white with same variance, and statistically independent of the sources, the noise variance can be estimated in a standard way [51] and subtracted from any estimated autocorrelation matrix $\mathbf{R}_x[k]$. Henceforth, we can write the autocorrelation matrices of the received data $\vec{\mathbf{x}}(n)$ as follows

$$\mathbf{R}_x[k] \triangleq E[\vec{\mathbf{x}}(n)\vec{\mathbf{x}}^H(n-k)] = \mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H. \tag{7.5}$$

In the following, we will show that, given that certain identifiability conditions are satisfied, the channel convolution matrix $\mathcal{H}$ can be completely identified up to a scalar factor by utilizing the estimated channel output autocorrelation matrices $\mathbf{R}_x[k], k \in \{0, \pm 1\}$ and the knowledge of $\mathbf{R}_s[k], k \in \{0, \pm 1\}$.

It is clear that from A5, we can write the following relationship

$$\mathbf{R}_s[0] = \mathbf{P}\mathbf{P}^H \tag{7.6}$$

where $\mathbf{P}$ is an invertible matrix. Also, if we write the eigenvalue decomposition: $\mathbf{R}_s[0] \triangleq \mathbf{U}_s\mathbf{D}_s\mathbf{U}_s^H$, we have

$$\mathbf{P} = \mathbf{U}_s\mathbf{D}_s^{1/2}\mathbf{M} \tag{7.7}$$

where $\mathbf{M}$ is a unitary matrix that can be properly chosen to facilitate our algorithm design. Now we consider the eigenvalue decomposition of $\mathbf{R}_x[0]$:

$$\mathbf{R}_x[0] \triangleq \begin{bmatrix} \mathbf{U}_{x,1} & \mathbf{U}_{x,2} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{x,1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{x,1}^H \\ \mathbf{U}_{x,2}^H \end{bmatrix}. \tag{7.8}$$

Let $\mathbf{G} \triangleq \mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$. Since $\mathbf{R}_x[0] = \mathcal{H}\mathbf{P}\mathbf{P}^H\mathcal{H}^H = \mathbf{G}\mathbf{G}^H$, it is clear that we have

$$\mathbf{G} = \mathcal{H}\mathbf{P}\mathbf{Q} \tag{7.9}$$

where $\mathbf{Q}$ is an unknown unitary matrix to be determined. To resolve this unknown matrix, we have to further explore the relationship imposed on this unitary matrix $\mathbf{Q}$. Recalling Eqn.(7.5) for $k \neq 0$, we have

$$\begin{aligned}\mathbf{G}^\dagger\mathbf{R}_x[k](\mathbf{G}^\dagger)^H &= \mathbf{Q}^H\mathbf{P}^{-1}\mathcal{H}^\dagger\mathcal{H}\mathbf{R}_s[k]\mathcal{H}^H(\mathcal{H}^\dagger)^H\mathbf{P}^{-H}\mathbf{Q} \\ &= \mathbf{Q}^H\mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-H}\mathbf{Q}. \end{aligned} \tag{7.10}$$

For notational convenience, let $\bar{\mathbf{R}}_x[k] \triangleq \mathbf{G}^\dagger\mathbf{R}_x[k](\mathbf{G}^\dagger)^H$ and $\bar{\mathbf{R}}_s[k] \triangleq \mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-H}$. Thus Eqn.(7.10) can be rewritten as

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^H\bar{\mathbf{R}}_s[k]\mathbf{Q} \qquad \forall\ k \neq 0 \tag{7.11}$$

and furthermore

$$\bar{\mathbf{R}}_x[k]\mathbf{Q}^H = \mathbf{Q}^H\bar{\mathbf{R}}_s[k] \qquad \forall\ k \neq 0. \tag{7.12}$$

Eqn.(7.12) defines the relationship the unknown unitary matrix $\mathbf{Q}$ must satisfy. Since $\bar{\mathbf{R}}_x[k]$ can be estimated from the second-order statistics of the channel output, $\bar{\mathbf{R}}_s[k]$ can also be computed for a properly chosen $\mathbf{P}$, the above equation can thus be used to estimate the unknown unitary matrix $\mathbf{Q}$. By utilizing the property of Kronecker product, we rewrite Eqn.(7.12) as

$$\mathbf{I}_d \otimes \bar{\mathbf{R}}_x[k] \cdot \mathrm{vec}(\mathbf{Q}^H) = \bar{\mathbf{R}}_s^T[k] \otimes \mathbf{I}_d \cdot \mathrm{vec}(\mathbf{Q}^H) \tag{7.13}$$

where $d \triangleq \sum_{i=1}^{p} d_i$ is the dimension of the matrices $\mathbf{Q}$ and $\bar{\mathbf{R}}_s[k]$. By defining

$\mathbf{q} \triangleq \text{vec}(\mathbf{Q}^H)$, we may estimate the unknown unitary matrix $\mathbf{Q}$ by the following criterion

$$\hat{\mathbf{q}} = \arg \min_{\|\mathbf{q}\|=1} \sum_{k=-K}^{K} \left\| \left[\mathbf{I}_d \otimes \bar{\mathbf{R}}_x[k] - \bar{\mathbf{R}}_s^T[k] \otimes \mathbf{I}_d \right] \mathbf{q} \right\|^2 . \tag{7.14}$$

The above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. However, this criterion fails to provide the true channel estimation if the solution to Eqn.(7.13) is not unique, i.e. there exist other non-zero vectors, $\mathbf{g}$, that are *linearly independent of* $\mathbf{q}$ and also satisfy $\mathbf{I}_d \otimes \bar{\mathbf{R}}_x[k] \cdot \mathbf{g} = \bar{\mathbf{R}}_s^T[k] \otimes \mathbf{I}_d \cdot \mathbf{g}$ for any $k \in \{-K, \ldots, K\}$. Hence we are faced with the following problem, that is, whether or not the solution to Eqn.(7.13) is unique (up to an unknown scalar factor) and under what conditions the solution to Eqn.(7.13) will be unique. This problem is studied in the following and we will show that, under certain identifiability conditions and for a properly chosen $\mathbf{P}$, the uniqueness of the solution to Eqn.(7.13) can be established for $k \in \{\pm 1\}$.

7.3.1 Property of Triangular Matrix

Notice that we have

$$\begin{aligned} \bar{\mathbf{R}}_s[k] &= \mathbf{P}^{-1}\mathbf{R}_s[k]\mathbf{P}^{-H} \\ &= \mathbf{M}^H\mathbf{D}_s^{-1/2}\mathbf{U}_s^H\mathbf{R}_s[k]\mathbf{U}_s\mathbf{D}_s^{-1/2}\mathbf{M} \\ &= \mathbf{M}^H\ddot{\mathbf{R}}_s[k]\mathbf{M} \end{aligned} \tag{7.15}$$

where $\ddot{\mathbf{R}}_s[k] \triangleq \mathbf{D}_s^{-1/2}\mathbf{U}_s^H\mathbf{R}_s[k]\mathbf{U}_s\mathbf{D}_s^{-1/2}$ can be computed *a priori*, $\mathbf{M}$ is a unitary matrix that can be properly selected to facilitate our algorithm design and

the proof for the uniqueness of the proposed system solution. Intuitively, it may be helpful if the chosen unitary matrix can make matrix $\bar{\mathbf{R}}_s[k]$ to possess some special structure. Therefore, it is natural for us to resort to the Schur decomposition that can transform any square matrix into an upper triangular matrix using a unitary matrix, i.e. $\mathbf{M}^H\mathbf{A}\mathbf{M} = \mathbf{T}$, where $\mathbf{T}$ is an upper triangular matrix. By exploiting the special (upper triangular) structure of the upper triangular matrix, we are able to derive some important property in the following lemma. This newly derived property plays a key role in the proof of the solution uniqueness.

Lemma 7.1 *Given that* $\mathbf{T} \in \mathbb{C}^{n\times n}$ *is an upper triangular matrix,* $\mathbf{Y} \in \mathbb{C}^{n\times n}$ *and we have*

$$\mathbf{T}\mathbf{Y} = \mathbf{Y}\mathbf{T} \tag{7.16}$$

if any pair of diagonal elements in $\mathbf{T}$ *are different from each other, i.e.* $t_{i,i} \neq t_{j,j}$ *for any* $i \neq j$*, then* $\mathbf{Y}$ *is also an upper triangular matrix.*

Proof: See Appendix D. ■

7.3.2 Proof of The Solution Uniqueness and The Proposed Algorithm

From the previous discussion, it is clear that we can choose a proper matrix $\mathbf{P} = \mathbf{U}_s\mathbf{D}_s^{1/2}\mathbf{M}$ to make $\bar{\mathbf{R}}_s[1] = \mathbf{P}^{-1}\mathbf{R}_s[1]\mathbf{P}^{-H}$ an upper triangular matrix. We now proceed to prove the uniqueness of the system solution based on our properly chosen matrix $\mathbf{P}$ and the derived property of the upper triangular matrix. We, firstly, prove that the solution to Eqn.(7.12) is unique (up to a scalar factor). The problem is formulated in the following theorem.

Theorem 7.1 *Given that*

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^H \bar{\mathbf{R}}_s[k]\mathbf{Q} \qquad k \in \{\pm 1\} \tag{7.17}$$

where $\bar{\mathbf{R}}_s[1]$ *is an upper triangular matrix which satisfies the following two identifiability conditions: IC1) the diagonal entries are all unequal; IC2) for any* $i \in \{2, \ldots, d\}$*, there exists at least one non-zero entry* $r_{j_1,i}$ *for* $j_1 < i$*, or for any* $i \in \{1, \ldots, d-1\}$*, there exists at least one non-zero entry* r_{i,j_2} *for* $j_2 > i$*, where* $r_{i,j}$ *denotes the* i^{th} *row and* j^{th} *column entry in* $\bar{\mathbf{R}}_s[1]$*. If there is any non-zero matrix* $\mathbf{C}$ *that satisfies the following*

$$\bar{\mathbf{R}}_x[k]\mathbf{C} = \mathbf{C}\bar{\mathbf{R}}_s[k] \qquad k \in \{\pm 1\} \tag{7.18}$$

then we have $\mathbf{C} = \lambda \mathbf{Q}^H$*, where* λ *can be any non-zero complex scalar.*

Proof: From Eqn.(7.17–7.18), it is easy to obtain

$$\mathbf{Q}^H \bar{\mathbf{R}}_s[k]\mathbf{Q}\mathbf{C} = \mathbf{C}\bar{\mathbf{R}}_s[k] \Rightarrow \bar{\mathbf{R}}_s[k]\mathbf{Q}\mathbf{C} = \mathbf{Q}\mathbf{C}\bar{\mathbf{R}}_s[k]. \tag{7.19}$$

Let $\mathbf{Z} \triangleq \mathbf{Q}\mathbf{C}$, we can rewrite the above equation as

$$\bar{\mathbf{R}}_s[k]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[k] \qquad k \in \{\pm 1\} \tag{7.20}$$

. We now only need to prove that $\mathbf{Z} = \lambda \mathbf{I}$. Since we have $\bar{\mathbf{R}}_s[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[1]$, where $\bar{\mathbf{R}}_s[1]$ is an upper triangular matrix whose diagonal entries are all unequal, by utilizing Lemma 7.1, it is easy to conclude that $\mathbf{Z}$ is an upper triangular matrix. For $\bar{\mathbf{R}}_s[-1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[-1]$, by exploiting the symmetry $\bar{\mathbf{R}}_s[-1] = \bar{\mathbf{R}}_s^H[1]$, we have

$$\bar{\mathbf{R}}_s[-1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[-1] \Rightarrow \bar{\mathbf{R}}_s^H[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s^H[1] \Rightarrow \bar{\mathbf{R}}_s[1]\mathbf{Z}^H = \mathbf{Z}^H\bar{\mathbf{R}}_s[1]. \tag{7.21}$$

Thus $\mathbf{Z}^H$ is also proved to be an upper triangular matrix. Therefore we can conclude that $\mathbf{Z}$ must be a diagonal matrix. Obviously, under the identifiability condition IC2, the diagonal elements in $\mathbf{Z}$ must be equal in order to satisfy $\bar{\mathbf{R}}_s[1]\mathbf{Z} = \mathbf{Z}\bar{\mathbf{R}}_s[1]$. Hence we have $\mathbf{Z} = \lambda\mathbf{I}$ and $\mathbf{C} = \lambda\mathbf{Q}^H$. The proof is completed here. ■

Notice that Eqn.(7.12) and Eqn.(7.13) can be derived from each other. This implies that the solution to Eqn.(7.13) is also unique and the solution is a scaling constant of the "true" vector $\mathbf{q}$. Therefore $\mathbf{q}$ can be estimated by the criterion of Eqn.(7.14) with $k \in \{\pm 1\}$, i.e.

$$\hat{\mathbf{q}} = \arg\min_{\|\mathbf{q}\|=1} \left\| \begin{bmatrix} \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[1] - \bar{\mathbf{R}}_s^T[1] \otimes \mathbf{I}_d \\ \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[-1] - \bar{\mathbf{R}}_s^T[-1] \otimes \mathbf{I}_d \end{bmatrix} \mathbf{q} \right\|^2 . \tag{7.22}$$

As indicated earlier, the above optimization has a closed-form solution which can be obtained as the right singular vector associated with the smallest singular value. For clarity, we now enumerate the steps for our channel identification procedure.

1. Compute the eigenvalue decomposition of $\mathbf{R}_s[0] = \mathbf{U}_s\mathbf{D}_s\mathbf{U}_s^H$. Let $\ddot{\mathbf{R}}_s[1] = \mathbf{D}_s^{-1/2}\mathbf{U}_s^H\mathbf{R}_s[1]\mathbf{U}_s\mathbf{D}_s^{-1/2}$.

2. Compute the Schur decomposition of $\ddot{\mathbf{R}}_s[1] = \mathbf{M}\bar{\mathbf{R}}_s[1]\mathbf{M}^H$, where $\bar{\mathbf{R}}_s[1]$ is an upper triangular matrix, $\mathbf{M}$ is a unitary matrix. Let $\mathbf{P} = \mathbf{U}_s\mathbf{D}_s^{1/2}\mathbf{M}$.

3. Compute the eigenvalue decomposition of $\mathbf{R}_x[0]$ as Eqn.(7.8) and let $\mathbf{G} = \mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$.

4. Compute $\bar{\mathbf{R}}_x[k] = \mathbf{G}^\dagger\mathbf{R}_x[k](\mathbf{G}^\dagger)^H$ for $k \in \{\pm 1\}$.

5. Estimate the unknown unitary matrix $\mathbf{Q}$ by using the criterion in Eqn.(7.22).

And the channel is thus estimated as $\hat{\mathcal{H}} = \mathbf{G}\hat{\mathbf{Q}}^H\mathbf{P}^{-1}$.

7.3.3 Discussions

To guarantee that our proposed channel identification method works, two identifiability conditions IC1–IC2 are proposed and stated in Theorem 7.1. Since the conditions are only related to the second-order statistics of the input sources, they can be checked *a priori* to determine whether the channel identifiability conditions are satisfied. Also, we emphasize that IC1–IC2 are sufficient while not necessary identifiability conditions for the channel identification. This implies that even if these two identifiability conditions are not met, the channel may still be identified using our proposed method. The reasons are as follows. Firstly, IC1 is a sufficient but not necessary condition to determine that $\mathbf{Z}$ is a diagonal matrix. For example, for the special case where the sources are spatially and temporally uncorrelated, the source autocorrelation matrices possess some special structure other than upper triangular structure that can be utilized to arrive at the same inference. This will be shown in the later part of this chapter. Secondly, IC2 is also a sufficient but not necessary condition to ensure that the diagonal elements of $\mathbf{Z}$ are equal. Furthermore, in some special cases, IC2 can be further relaxed or removed because the diagonal elements of $\mathbf{Z}$ are not necessary to be equal. It can be elucidated as follows. From the proof for Theorem 7.1, it is evident that if IC2 is not assumed, then we can only conclude that $\mathbf{Z} = \mathbf{QC}$ is a diagonal matrix, while the diagonal elements may not be equal. Thus the unknown unitary matrix is only identified up to an unknown diagonal matrix, i.e. $\hat{\mathbf{Q}}^H = \mathbf{Q}^H\mathbf{Z}$, where $\mathbf{Z}$ is a diagonal matrix. In this case, the channel $\mathcal{H}$ is estimated as $\hat{\mathcal{H}} = \mathbf{G}\mathbf{Q}^H\mathbf{Z}\mathbf{P}^{-1}$. Obviously, only when $\mathbf{Z} = \lambda\mathbf{I}$ can the channel be estimated up to a scalar factor, otherwise the channel can

not be correctly estimated since generally, $\mathbf{P} \neq \mathbf{I}$. Nevertheless, in the special case where $\mathbf{P} = \mathbf{I}$, the diagonal elements of $\mathbf{Z}$ need not be equal. This is because in this case, we have $\hat{\mathcal{H}} = \mathcal{H}\mathbf{Z}$. Since the channel equalization can be achieved if only the channel per user $\mathcal{H}_i$ for $i \in \{1, \ldots, p\}$ rather than the MIMO channel $\mathcal{H}$ is identified up to a scalar factor, it allows that $\mathbf{Z} = \mathrm{diag}(\lambda_1 \mathbf{I}_{d_1}, \cdots, \lambda_p \mathbf{I}_{d_p})$. However, in this case, we can not directly adopt the criterion of Eqn.(7.22) to estimate the unitary matrix $\mathbf{Q}^H$ because the uniqueness of the solution to Eqn.(7.18) has been spoilt by the disparity of the diagonal elements in $\mathbf{Z}$, and now we have to estimate $\mathbf{Q}^H$ in p parallel threads. This estimation can be illustrated as follows. Specifically, if we assume that $\lambda_i = 0$ for all i except for λ_1 in $\mathbf{Z}$, then any non-zero matrix $\mathbf{C}$ that satisfies Eqn.(7.18) can be written as $\mathbf{C} = \mathbf{Q}^H \mathbf{Z} = \begin{bmatrix} \lambda_1 \bar{\mathbf{Q}}_1 & \mathbf{0} \end{bmatrix}$, where $\bar{\mathbf{Q}}_1 \triangleq \bar{\mathbf{Q}}[:, 1 : d_1]$, $\bar{\mathbf{Q}} \triangleq \mathbf{Q}^H$. Hence the solution of $\mathbf{C}$ to Eqn.(7.18) is unique up to a scalar factor and $\bar{\mathbf{Q}}_1$ can be estimated as follows by defining $\mathbf{q}_1 \triangleq \mathrm{vec}(\bar{\mathbf{Q}}_1)$

$$\hat{\mathbf{q}}_1 = \arg \min_{\|\mathbf{q}_1\|=1} \|\Omega[:, 1 : d_1 d]\mathbf{q}_1\|^2 \tag{7.23}$$

where

$$\Omega \triangleq \begin{bmatrix} \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[1] - \bar{\mathbf{R}}_s^T[1] \otimes \mathbf{I}_d \\ \mathbf{I}_d \otimes \bar{\mathbf{R}}_x[-1] - \bar{\mathbf{R}}_s^T[-1] \otimes \mathbf{I}_d \end{bmatrix}.$$

Similarly, $\bar{\mathbf{Q}}_l \triangleq \bar{\mathbf{Q}}[:, \sum_{i=1}^{l-1} d_i + 1 : \sum_{i=1}^{l} d_i]$ can be estimated as

$$\hat{\mathbf{q}}_l = \arg \min_{\|\mathbf{q}_l\|=1} \left\| \Omega \left[:, \sum_{i=1}^{l-1} d_i d + 1 : \sum_{i=1}^{l} d_i d \right] \mathbf{q}_l \right\|^2. \tag{7.24}$$

In the following, a special example where the input sources are spatially and temporally uncorrelated is considered, and we will show that for this special

case, the channel can still be identified despite the violation of IC1–IC2, given that the channel orders corresponding to each pair of users are different from each other.

7.4 Channel Identifiability Condition for Spatially and Temporally Uncorrelated Inputs

For the spatially and temporally uncorrelated input sources, we have $\mathbf{R}_s[0] = \mathbf{I}_d$ and

$$\mathbf{R}_s[1] = \text{diag}\left(\mathbf{J}_{d_1}, \mathbf{J}_{d_2}, \cdots, \mathbf{J}_{d_p}\right) \tag{7.25}$$

where diag$(\cdot)$ denotes block diagonal; $\mathbf{J}_n$ stands for the $n \times n$ one-lag down shift square matrix whose first sub-diagonal entries below the main diagonal are unity, whereas all remaining entries are zero; $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. Clearly, in this case, $\mathbf{P}$ is chosen to be the exchange matrix with ones on the antidiagonal and zeros elsewhere (see Eqn.(7.15)). And the transformed upper triangular matrix is written as

$$\bar{\mathbf{R}}_s[1] = \mathbf{P}^{-1}\mathbf{R}_s[1]\mathbf{P}^{-H} = \text{diag}\left(\mathbf{J}_{d_1}^T, \mathbf{J}_{d_2}^T, \cdots, \mathbf{J}_{d_p}^T\right). \tag{7.26}$$

We can see that the upper triangular matrix $\bar{\mathbf{R}}_s[1]$ does not satisfy IC1–IC2 of Theorem 7.1. Nevertheless, we will show that the channel can still be identified unambiguously from the second-order statistics of the channel output provided that another identifiability condition is satisfied. Notice that $\bar{\mathbf{R}}_s[1]$ is exactly equal to $\mathbf{R}_s[-1]$ and $\bar{\mathbf{R}}_s[-1]$ is exactly equal to $\mathbf{R}_s[1]$. For simplicity, we can let $\mathbf{P}$ be an identity matrix instead of the exchange matrix. And we thus have

$\bar{\mathbf{R}}_s[k] = \mathbf{R}_s[k]$. We now formulate our results in the following theorem.

Theorem 7.2 *Given that*

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^H \mathbf{R}_s[k] \mathbf{Q} \qquad k \in \{\pm 1\} \tag{7.27}$$

if for every pair of sources $\{s_i, s_j\}$, $L_i \neq L_j$, *then any non-zero matrix* $\mathbf{C}$ *that satisfies* $\bar{\mathbf{R}}_x[k]\mathbf{C} = \mathbf{C}\mathbf{R}_s[k]$ *for* $k \in \{\pm 1\}$ *can be written as* $\mathbf{C} = \mathbf{Q}^H \mathbf{D}$, *where* $\mathbf{D} \triangleq \mathrm{diag}(\lambda_1 \mathbf{I}_{d_1}, \cdots, \lambda_p \mathbf{I}_{d_p})$, λ_i *for* $i \in \{1, \ldots, p\}$ *can be any complex scalar including zero.*

Before we proceed to prove Theorem 7.2, we first introduce the following lemma that exploits the properties of the one-lag down and up shift square matrices.

Lemma 7.2 *Given that* $\mathbf{Y} \in \mathbb{C}^{m \times n}$ *satisfies the following two equations*

$$(a)\ \mathbf{J}_m \mathbf{Y} = \mathbf{Y} \mathbf{J}_n \quad (b)\ \mathbf{J}_m^T \mathbf{Y} = \mathbf{Y} \mathbf{J}_n^T \tag{7.28}$$

then we have

- *If* $m = n$, *then* $\mathbf{Y} = \lambda \mathbf{I}$, *where* λ *could be any complex scalar including zero.*
- *If* $m \neq n$, *then* $\mathbf{Y} = \mathbf{0}$.

Proof: See Appendix E. ■

Now we proceed to prove Theorem 7.2.

Proof: Similarly as the proof of Theorem 7.1, we have

$$\mathbf{Q}^H \mathbf{R}_s[k] \mathbf{Q} \mathbf{C} = \mathbf{C} \mathbf{R}_s[k] \Rightarrow \mathbf{R}_s[k] \mathbf{Q} \mathbf{C} = \mathbf{Q} \mathbf{C} \mathbf{R}_s[k] \tag{7.29}$$

Let $\mathbf{Z} \triangleq \mathbf{QC}$, we can rewrite the above equation as

$$\mathbf{R}_s[k]\mathbf{Z} = \mathbf{Z}\mathbf{R}_s[k] \qquad k \in \{\pm 1\}. \tag{7.30}$$

We now prove that $\mathbf{Z} = \mathrm{diag}(\lambda_1 \mathbf{I}_{d_1}, \cdots, \lambda_p \mathbf{I}_{d_p})$. We partition matrix $\mathbf{Z}$ as follows

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} & \cdots & \mathbf{Z}_{1p} \\ \mathbf{Z}_{21} & \mathbf{Z}_{22} & \cdots & \mathbf{Z}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{Z}_{p1} & \mathbf{Z}_{p2} & \cdots & \mathbf{Z}_{pp} \end{bmatrix}$$

where $\mathbf{Z}_{ij} \in \mathbb{C}^{d_i \times d_j}$. Since the matrices $\mathbf{R}_s[k]$ for $k \in \{\pm 1\}$ are block diagonal matrices, it is straightforward for us to obtain the following from Eqn.(7.30)

$$(a)\ \mathbf{J}_{d_i}\mathbf{Z}_{ij} = \mathbf{Z}_{ij}\mathbf{J}_{d_j} \quad (b)\ \mathbf{J}_{d_i}^T\mathbf{Z}_{ij} = \mathbf{Z}_{ij}\mathbf{J}_{d_j}^T \quad . \tag{7.31}$$

Obviously, for the case where $d_i \neq d_j$ (note that $L_i \neq L_j$ is equivalent to $d_i \neq d_j$ since $d_i = N + L_i + 1$) for each pair of $\{i, j\}$, we can conclude that $\mathbf{Z} = \mathrm{diag}(\lambda_1 \mathbf{I}_{d_1}, \cdots, \lambda_p \mathbf{I}_{d_p})$ by utilizing the results of Lemma 7.2. Hence we have $\mathbf{C} = \mathbf{Q}^H\mathbf{Z} = \mathbf{Q}^H\mathbf{D}$. The proof is completed here. ■

We now develop the corresponding algorithm for the channel identification. Let

$$\mathbf{C} \triangleq \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \cdots & \mathbf{C}_p \end{bmatrix}$$

$$\bar{\mathbf{Q}} \triangleq \mathbf{Q}^H \triangleq \begin{bmatrix} \bar{\mathbf{Q}}_1 & \bar{\mathbf{Q}}_2 & \cdots & \bar{\mathbf{Q}}_p \end{bmatrix}$$

where $\mathbf{C}_i \in \mathbb{C}^{d \times d_i}$ and $\bar{\mathbf{Q}}_i \in \mathbb{C}^{d \times d_i}$. By exploiting the block diagonal structure of $\mathbf{R}_s[k]$, the set of equations $\bar{\mathbf{R}}_x[k]\mathbf{C} = \mathbf{C}\mathbf{R}_s[k]$ for $k \in \{\pm 1\}$ can be decoupled

into the following p sets of equations

$$\bar{\mathbf{R}}_x[k]\mathbf{C}_i = \mathbf{C}_i\mathbf{R}_{s_i}[k] \qquad i \in \{1, \ldots, p\} \tag{7.32}$$

where $\mathbf{R}_{s_i}[k] \triangleq E[\vec{\mathbf{s}}_i(n)\vec{\mathbf{s}}_i^H(n-k)]$. Using Theorem 7.2, it is easy to see that any matrix $\mathbf{C}_i$ that satisfies the above equation for $k \in \{\pm 1\}$ can be written as $\mathbf{C}_i = \lambda_i \bar{\mathbf{Q}}_i$. Thus the unknown unitary matrix $\bar{\mathbf{Q}}$ can be estimated in p parallel threads with the ith thread leading to the estimation of $\bar{\mathbf{Q}}_i$, where $\bar{\mathbf{Q}}_i$ is estimated as a closed-form minimizer of the following criterion by defining $\mathbf{q}_i \triangleq \text{vec}\left(\bar{\mathbf{Q}}_i\right)$

$$\hat{\mathbf{q}}_i = \arg \min_{\|\mathbf{u}\|=1} \left\| \begin{bmatrix} \mathbf{I}_{d_i} \otimes \bar{\mathbf{R}}_x[1] - \mathbf{R}_{s_i}^T[1] \otimes \mathbf{I}_d \\ \mathbf{I}_{d_i} \otimes \bar{\mathbf{R}}_x[-1] - \bar{\mathbf{R}}_{s_i}^T[-1] \otimes \mathbf{I}_d \end{bmatrix} \mathbf{q}_i \right\|^2 . \tag{7.33}$$

It can be easily verified that, in this special case, the above criterion is equivalent to the proposed criterion in Eqn.(7.24). Clearly, the above proposed identification algorithm estimates the unknown unitary matrix $\bar{\mathbf{Q}}$ up to a diagonal matrix $\mathbf{D} = \text{diag}(\lambda_1 \mathbf{I}_{d_1}, \cdots, \lambda_p \mathbf{I}_{d_p})$, where λ_i for $i \in \{1, \cdots, p\}$ can be any *non-zero* complex scalar. Hence we have

$$\hat{\mathcal{H}} = \mathbf{G}\hat{\mathbf{Q}}^H\mathbf{P}^{-1} = \mathbf{G}\mathbf{Q}^H\mathbf{D} = \mathcal{H}\mathbf{D}. \tag{7.34}$$

Thus the channel per user $\mathcal{H}_i$ for $i \in \{1, \ldots, p\}$ is identified up to a scalar factor. This result, of course, is only valid under the condition that for each pair of users $\{s_i, s_j\}$, we have $L_i \neq L_j$. In fact, this identifiability condition can be further relaxed if only the desired user channels rather than the MIMO channels for all users are identified and equalized. This result is formulated in Theorem 7.3 and for simplicity, we consider the case where only one user

channel is desired.

Theorem 7.3 *Given that*

$$\bar{\mathbf{R}}_x[k] = \mathbf{Q}^H \mathbf{R}_s[k] \mathbf{Q} \qquad k \in \{\pm 1\} \tag{7.35}$$

if for a desired user s_l, we have $L_l \neq L_i$ for $i \in \{1, \ldots, l-1, l+1, \ldots, p\}$, then any non-zero matrix $\mathbf{C}_l$ that satisfies $\bar{\mathbf{R}}_x[k]\mathbf{C}_l = \mathbf{C}_l \mathbf{R}_{s_l}[k]$ for $k \in \{\pm 1\}$ can be written as $\mathbf{C}_l = \lambda_l \bar{\mathbf{Q}}_l$.

Proof: See Appendix F. ■

Based on Theorem 7.3, we can estimate $\bar{\mathbf{Q}}_l$ using the criterion of Eqn.(7.33) with i replaced by l and the desired user channel $\mathcal{H}_l$ can be estimated as $\hat{\mathcal{H}}_l = \mathbf{G}\hat{\bar{\mathbf{Q}}}_l$. It is noted that identifying $\mathcal{H}_l$ alone allows us to compute the MMSE equalizers to recover the transmitted signals s_l by removing the intersymbol interference and canceling the multiuser interference [118]. For clarity and comparison purpose, we also enumerate the steps for channel identification for the spatially and temporally uncorrelated sources.

1. Compute the eigenvalue decomposition of $\mathbf{R}_x[0]$ as Eqn.(7.8) and let $\mathbf{G} = \mathbf{U}_{x,1}\mathbf{D}_{x,1}^{1/2}$.

2. Compute $\bar{\mathbf{R}}_x[k] = \mathbf{G}^\dagger \mathbf{R}_x[k] (\mathbf{G}^\dagger)^H$ for $k \in \{\pm 1\}$.

3. Estimate $\bar{\mathbf{Q}}_i$ by using the criterion in Eqn.(7.33). And the desired user channel is estimated as $\hat{\mathcal{H}}_i = \mathbf{G}\hat{\bar{\mathbf{Q}}}_i$.

Remark: As we can see, for the spatially and temporally uncorrelated sources, the identifiability conditions IC1–IC2 proposed in previous section are no longer

necessary for the complete channel identification. This point can also be corroborated by [72, 74] for the spatially uncorrelated but temporally correlated sources. Since the identifiability conditions IC1–IC2 are proposed for generally correlated sources, i.e. the sources can be spatially and temporally correlated, it is no surprise that these conditions can be further relaxed for the special cases where the sources are spatially and temporally uncorrelated or spatially uncorrelated but temporally correlated because, as mentioned previously, for the special cases, the source autocorrelation matrices and their revised forms possess some special structure other than upper triangular structure that can be better utilized.

7.5 Simulation Results

We now present simulation results to validate the performance of our proposed algorithm. Four examples are studied in this chapter. In the first example, we show the equalization performance of our proposed algorithm for the case where the sources are spatially and temporally uncorrelated, and consequently we investigate how the equalization performance hinges on the following parameters: equalization delays d_e, number of samples used for statistics estimation T_s and SNR. In the rest of the examples, we consider the SIMO nonlinear channels which have been adopted by the work [100, 101], and we compare our method to the SOS-based method proposed in [101] and the deterministic method presented in [100], which are named as "RS" method (R. López-Valcarce and S. Dusgupta) and "GE" method (G. B. Giannakis and E. Serpedin), respectively. Both the cases of i.i.d. input signals and colored input signals to the nonlinear channels are investigated in our examples. Also, in our simulations, the addi-

tive noise $\mathbf{w}(n)$ is taken as spatial-temporal white complex Gaussian noise with variance σ_w^2. The SNR is defined as

$$\text{SNR} = 10 \cdot \log \frac{E[\|\mathcal{H}\vec{\mathbf{s}}(n)\|^2]}{E[\|\vec{\mathbf{w}}(n)\|^2]}.$$

7.5.1 Example One

We consider $p = 2$ sources arriving at $q = 3$ sensors via a multipath channel. The source signals are i.i.d. information sequences drawn from a 4-QPSK constellation $\mathcal{S} = \{1, -1, i, -i\}$. The channel is randomly generated as

$$\{\mathbf{h}_1(l)\} = \begin{bmatrix} 0.0572 & 0.2074 & -0.0466 & 0.1085 \\ 0.2475 & -0.1004 & 0.0213 & -0.2331 \\ 0.0968 & -0.2527 & -0.3888 & 0.2701 \end{bmatrix}$$

$$\{\mathbf{h}_2(l)\} = \begin{bmatrix} 0.2885 & 0.4926 & 0.2480 & 0 \\ 0.1714 & -0.2387 & 0.1945 & 0 \\ 0.0455 & -0.0463 & -0.0256 & 0 \end{bmatrix}.$$

It can be seen that the channel orders corresponding to these two users are different and this suffices for the complete channel identification. Once the channel has been estimated by our algorithm, we can compute the ZF equalizers and the MMSE equalizers respectively as

$$\mathcal{E}_{\text{ZF}} = \hat{\mathcal{H}}^{\dagger}$$
$$\mathcal{E}_{\text{MMSE}} = \mathcal{E}_{\text{ZF}}(\mathbf{I} - \sigma_w^2 \hat{\mathbf{R}}_x^{-1}[0])$$

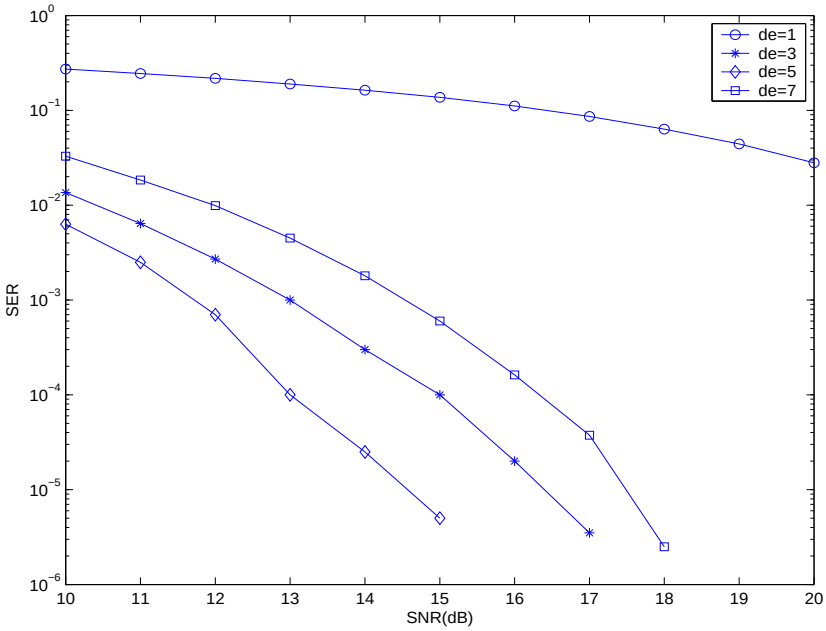
where $\hat{\mathbf{R}}_x[0]$ is the estimated autocorrelation matrix before denoised. The above expression for the MMSE equalizers was derived in [116], which is applicable for (spatially and temporally) uncorrelated sources and correlated sources. The inherent phase ambiguity of equalizers per user is removed before we perform the equalization. In the simulations, we choose stack number $N = 5$. The channel order of each user is assumed known *a priori*. Results are averaged over 500 Monte Carlo runs. Figure 7.1 shows the SER as a function of SNR for the MMSE equalizers with delays 1, 3, 5 and 7. $T_s = 2000$ data samples are used for the estimation of the autocorrelation matrices of the received data. Clearly, we can see that, the SER decreases as SNR increases. And the equalizers of intermediate delays are superior to those of extremal delays in performance. Figure 7.2 depicts the SER of the MMSE equalizer with $d_e = 5$ using different number of data samples. As expected, the performance improves with an increasing T_s. From these results, we can see that, even for the spatially and temporally uncorrelated sources, the channel can still be completely identified by exploiting the channel order disparity.

7.5.2 Example Two

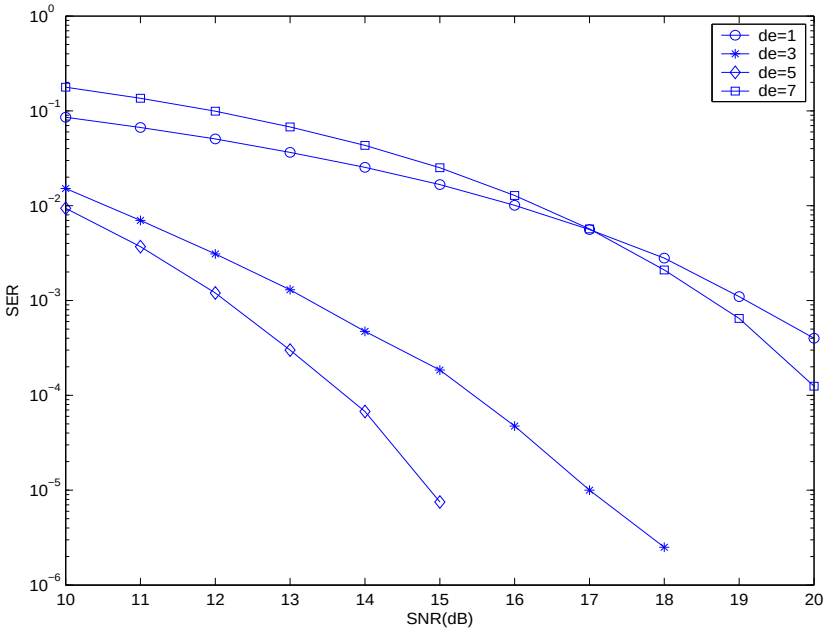
In this example, we consider the following SIMO nonlinear channel which was adopted by the third example of [100]

$$\mathbf{x}(n) = \sum_{l=0}^{3} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n)$$

where now, $s_2(n) \triangleq a(n)a(n-1)a^*(n-2)$, $a(n)$ are i.i.d. symbols drawn from the 4-QPSK constellation $\mathcal{S} = \{1, -1, i, -i\}$. It can be easily verified that $s_2(n)$ is also a temporally uncorrelated sequence, and the "two sources" $a(n)$ and

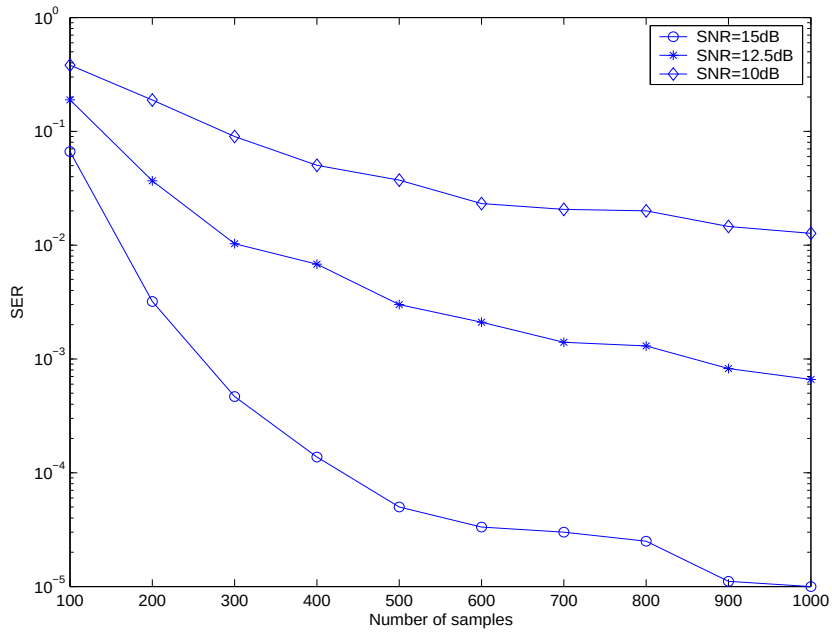


(a) user 1

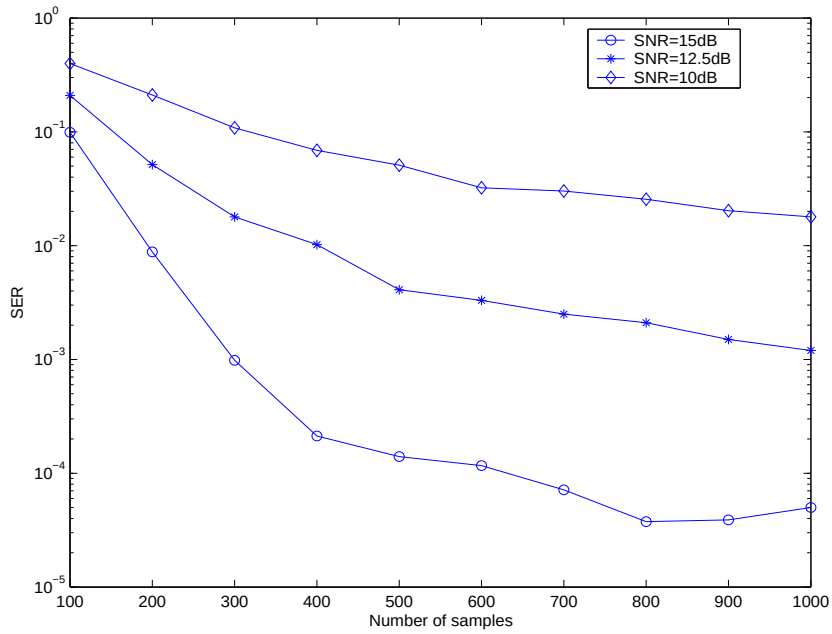


(b) user 2

Figure 7.1: Example 1: SER versus SNR for different equalization delays d_e; $T_s = 2000$



(a) user 1



(b) user 2

Figure 7.2: Example 1: SER versus T_s for different SNR; $d_e = 5$

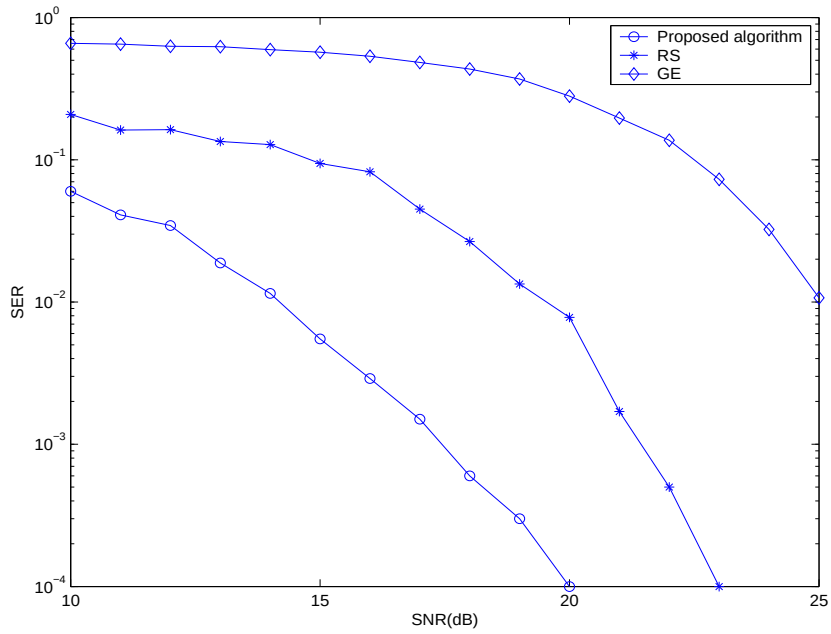
$s_2(n)$ are spatially uncorrelated with different channel order. Thus our proposed method can be applied to this nonlinear channel example. We compare our proposed algorithm with the RS method [101] and the GE method [100]. In our simulations, the stack number N is chosen to be 4 for our method and the RS method, and we use the equalizer with $d_e = 7$ which achieves the best performance. For the GE method, it only provides equalizers with minimal and maximal delays, and here we employ the maximal delay $d_e = 6$ which has a better performance. Figures 7.3 displays the equalization performance (for the nonlinear channel input $a(n)$) of the three algorithms as a function of SNR and T_s, respectively. From the figures, we can see that our proposed algorithm presents a clear performance advantage over RS and GB methods. Also, we observed that the second-order statistics methods (RS and our proposed method) are more favorable than the deterministic method (GB) to obtain an accurate symbol estimation. It seems that using the source statistics can help to gain a stronger robustness to the noise.

7.5.3 Example Three

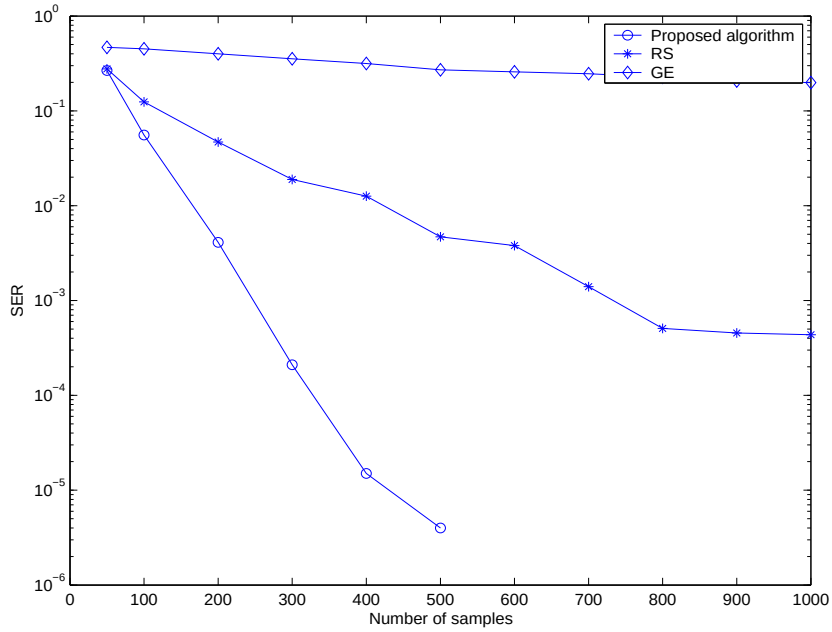
We investigate the case where the input signal $\{a(n)\}$ to the nonlinear channel is colored. The SIMO nonlinear channel used by the first example of [101] is adopted.

$$\mathbf{x}(n) = \sum_{l=0}^{2} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n).$$

where $s_2(n) \triangleq a(n)a(n-1)$. The input signals $\{a(n)\}$ are generated from the 4-QAM constellation $S = \{-1-i, -1+i, 1-i, 1+i\}$ according to the model which simulates a *Markov* source by implementing the transition probabilities of Table



(a) SER versus SNR, $T_s = 500$



(b) SER versus T_s, SNR = 20dB

Figure 7.3: Example 2: Performance of respective algorithms

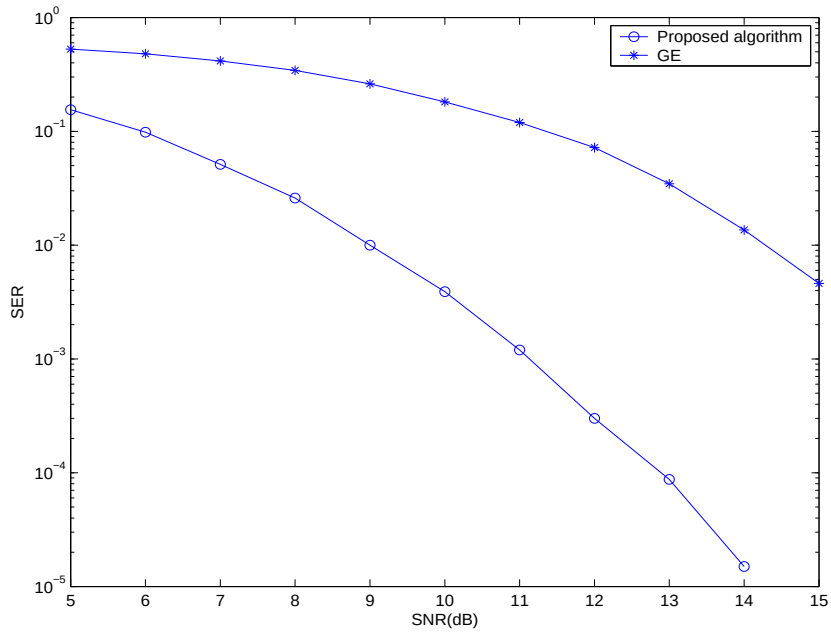
6.1. The autocorrelation function of this source is given in Table 6.2. Clearly, these two reformulated "sources" $a(n)$ and $s_2(n)$ are spatially and temporally correlated. Also, it can be verified that the identifiability conditions IC1–IC2 required by our proposed algorithm in Section 7.3 are satisfied. Since the input signals are correlated, the RS method no longer applies in this example. We compare our proposed algorithm with the GE method. In our simulations, we choose the stack number $N = 3$ for our method, and the equalizer with $d_e = 4$ is used for both methods. Figure 7.4(a) shows the SER (for the nonlinear input signals) of the respective algorithms as a function of SNR using $T_s = 500$ data samples. Figure 7.4(b) shows the variation of SER (for the nonlinear input signals) with the number of data samples T_s for SNR = 10dB. We can see that, for both cases, our proposed algorithm clearly outperforms the GE method.

7.5.4 Example Four

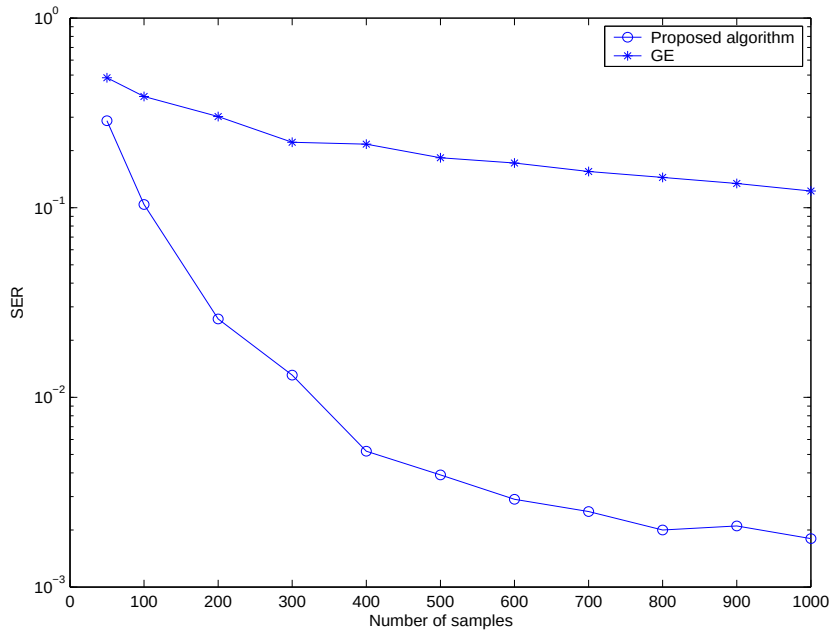
In this example, we consider the nonlinear channel which was used by the third example of [101].

$$\mathbf{x}(n) = \sum_{l=0}^{1} \mathbf{h}_1(l)a(n-l) + \sum_{l=0}^{1} \mathbf{h}_2(l)s_2(n-l) + \mathbf{w}(n)$$

The two "sources" $a(n)$ and $s_2(n)$ are spatially, temporally correlated and generated in the same way as the above example. Observe that in this case, the linear and nonlinear kernels have the same channel order. Nevertheless, given that the conditions IC1–IC2 are satisfied, our proposed method still applies. In our simulations, the stack number N is chosen to be 2. Figure 7.12(a) shows the SER (for the nonlinear input signals) versus SNR for the different equalization delays. It can be seen that the equalizer with $d_e = 1$ yields the best perfor-

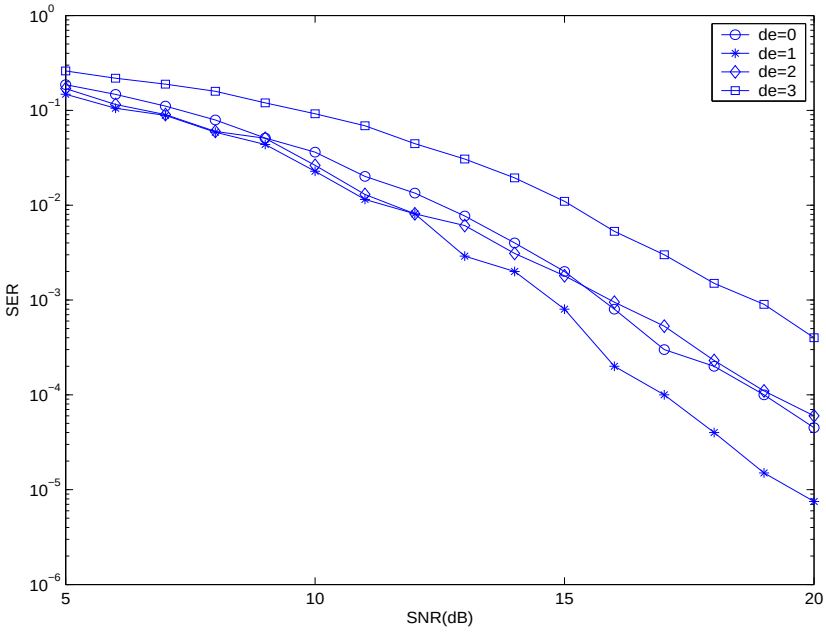


(a) SER versus SNR, $T_s = 500$

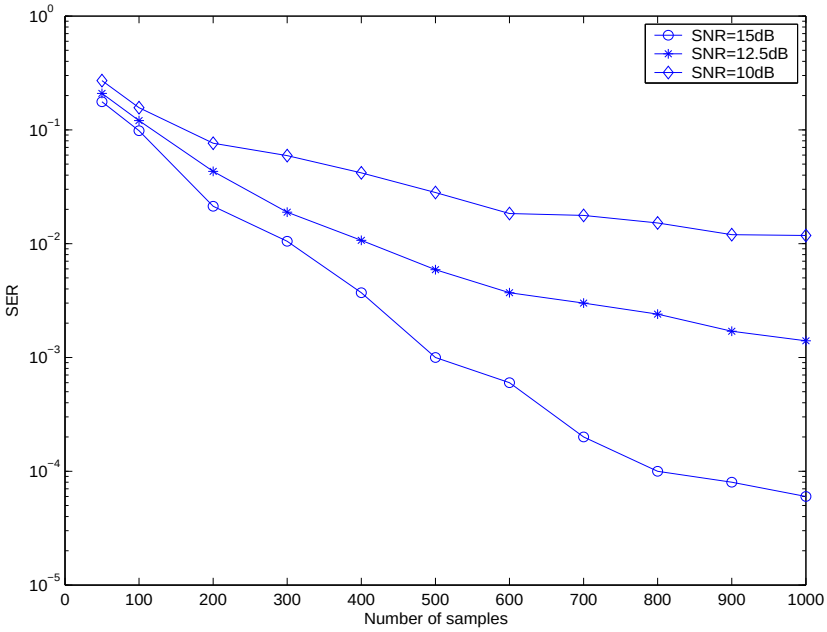


(b) SER versus T_s, SNR = 10dB

Figure 7.4: Example 3: Performance of respective algorithms



(a) SER versus SNR, $T_s = 500$



(b) SER versus T_s, $d_e = 1$

Figure 7.5: Example 4: Performance of the proposed algorithm

mance. The poorest results are obtained for $d_e = 3$. In Figure 7.12(b), the variation of the SER (for the nonlinear input signals) with the number of data samples T_s for SNR $= 10, 12.5, 15$dB is displayed. The equalizer with $d_e = 1$ is used.

7.6 Summary

In this chapter, we considered the problem of blind identification of MIMO FIR channel driven by spatially correlated sources whose second-order statistics are known *a priori*. A SOS-based method that admits a closed-form solution was proposed and its corresponding identifiability conditions were investigated. As a further result, we showed that our method still applies to the spatially and temporally uncorrelated sources given that a certain channel order disparity condition is satisfied. Simulation results showed that our method can be successfully employed for blind nonlinear SIMO channel equalization. As compared to other existing methods [100, 101] for blind nonlinear SIMO channel equalization, our method renders a wider applicability for the input sources than [101] and exhibits a better performance than [100].

Chapter 8

Conclusions and Future Work

8.1 Conclusions

In this dissertation, we have presented a variety of statistics-based methods for blind channel estimation and equalization in wireless systems. The proposed methods include the higher order statistics-based linear method for SISO channel estimation (Chapter 3), the second-order statistics-based methods for SIMO channel identification and equalization (Chapter 4 and Chapter 6, respectively), and the second-order statistics-based methods for MIMO channel identification (Chapter 5 and Chapter 7, respectively). In particular, throughout this dissertation, our work (except Chapter 3) focused on the problem of blind estimation/equalization of channels driven by colored source(s). As indicated previously, this problem has not received much attention as compared to its counterpart of blind channel estimation problem with white (or i.i.d.) input source signals. However, blind estimation/equalization of channels driven by colored source(s) never lacks its applications. In fact, this problem arises in a wide variety of communication and signal processing applications, which

include speech enhancement, wireless mobile communications and brain signal analysis.

In Chapters 4 and 5, we considered the problem of blind SIMO/MIMO channel identification driven by colored source(s) with *a priori* known statistics. By exploiting the inherent structural relationship between the source autocorrelation matrices and utilizing the derived properties of the companion matrices, we proposed a new closed-form solution for blind identification of channels driven by colored sources. The proposed algorithm compares favorably with existing methods in performance and computational complexity. Next, in Chapter 6, the problem of blind SIMO channel identification/equalization driven by colored source with unknown input statistics was investigated. It has been shown that although the statistical information of the transmitted signals is not available, we can still estimate the ZF and MMSE equalizers of desired delays from the second-order statistics of the received data by exploiting the inherent structural relationship between source autocorrelation matrices of different delay lags. The proposed method outperforms the existing methods significantly for the colored sources whose coefficients of the multistep optimum forward/backward prediction error filters are small. Finally, in Chapter 7, our research interest went to blind identification of MIMO channel driven by spatially and temporally correlated sources. We have shown that, under certain specified identifiability conditions, the MIMO FIR channel can be completely identified using the second-order statistics of the channel output. The method can be successfully employed for blind nonlinear SIMO channel equalization. As compared to other existing methods for blind nonlinear SIMO channel equalization, our method renders a wider applicability for the input sources and exhibits a better performance.

8.2 Future Work

Several future work are enumerated below.

In Chapter 6, we proposed a SOS-based method for blind equalization of SIMO channel driven by colored sources with unknown statistics. The method can directly estimate the ZF and MMSE equalizers of desired delays from the second-order statistics of the received data. The proposed method has the potential to be extended to multiuser scenarios, that is, MIMO systems. The theorem in Chapter 6 is still valid when extended to the multiuser case. For the case where the channel orders of each user are equal, the estimated equalizers cancel the intersymbol interference of all source signals and the equalized signals are the instantaneous mixtures of the source signals. We can easily recover the source signals from the equalized signals using the blind source separation techniques. If the channel orders of each user are different, then the channel order disparity alone enables us to extract the source signals successively. In this case, an effective successive extraction algorithm needs to be worked out.

It is also very meaningful for us to devise computationally efficient on-line adaptive algorithms for our proposed statistics-based methods because the existing off-line batch algorithms involve a high computation cost, and thus unsuitable for practical implementations. The high computational cost of the off-line batch algorithms mainly results from matrix decompositions such as eigenvalue decomposition (EVD) or singular value decomposition (SVD). For our proposed methods, the computation involves that of some minimal eigenvector or singular vector in the quadratic constraint or least square case. Thus the eigenvector tracking methods, e.g. [119], can be used to perform the blind estimation/equalization of our proposed method.

As mentioned before, channel identifiability conditions are closely related to blind channel estimation problems. An important issue is how to relax the channel identifiability conditions in the proposed methods. For our proposed methods [74, 102] (in Chapters 5 and 7) for blind MIMO channel identification, we require that the channel is irreducible and column-reduced to ensure a full column rank channel convolution matrix. However, in [33], the condition "column-reduced" is shown to be unnecessary for blind identification. Inspired by this, we may also design a channel identification algorithm in our scenarios with a relaxed identifiability condition.

Appendix A

Proof of Lemma 4.3

We present our proof of Lemma 4.3 in the following three steps.

Step 1: For notational convenience, let $\mathbf{G}_1 \triangleq \Theta_1\mathbf{Y} = \mathbf{Y}\Theta_1$ and $\mathbf{G}_2 \triangleq \Theta_2\mathbf{Y} = \mathbf{Y}\Theta_2$, $\mathbf{y}_i$ denotes the ith column of $\mathbf{Y}$. We consider the last column of $\mathbf{G}_1$, denoted by $\mathbf{G}_1[:,d]$, and the first column of $\mathbf{G}_2$, denoted by $\mathbf{G}_2[:,1]$. Thus we have

$$\begin{aligned}
\mathbf{G}_1[:,d] &= \begin{bmatrix} -\vec{\alpha}_1^H\mathbf{y}_d & y_{1,d} & \cdots & y_{d-1,d} \end{bmatrix}^T \\
&= -\alpha_{1,d}^* \begin{bmatrix} y_{1,1} & y_{2,1} & \cdots & y_{d,1} \end{bmatrix}^T
\end{aligned}$$

$$\begin{aligned}
\mathbf{G}_2[:,1] &= \begin{bmatrix} y_{2,1} & \cdots & y_{d,1} & -\vec{\alpha}_2^H\mathbf{y}_1 \end{bmatrix}^T \\
&= -\alpha_{1,d} \begin{bmatrix} y_{1,d} & y_{2,d} & \cdots & y_{d,d} \end{bmatrix}^T
\end{aligned}$$

and we can obtain

$$y_{k,d} = |\alpha_{1,d}|^2 y_{k,d} \ \ \forall \ k \in \{1, \ldots, d-1\} \tag{A.1}$$

$$y_{k+1,1} = |\alpha_{1,d}|^2 y_{k+1,1} \ \ \forall \ k \in \{1, \ldots, d-1\}. \tag{A.2}$$

It is known that (see Theorem 1 in [67]) $|\alpha_{1,d}|$ will be less than one under the assumption that the $(d+1) \times (d+1)$ source autocorrelation matrix $\mathbf{R}_s[0]$ is positive definite. In fact, even if this assumption does not hold, the probability of $|\alpha_{1,d}| = 1$ is still almost equal to zero. Therefore we can conclude that

$$\begin{aligned} y_{k,d} &= 0 \qquad \forall \ k \in \{1, \ldots, d-1\} \\ y_{k+1,1} &= 0 \qquad \forall \ k \in \{1, \ldots, d-1\}. \end{aligned}$$

Step 2: We consider the sub-matrix of $\mathbf{G}_1$ from second row to d^{th} row and from first column to $(d-1)^{th}$ column, denoted by $\mathbf{G}_1[2:d, 1:d-1]$. This sub-matrix can be easily computed if we write Θ_1 and $\mathbf{Y}$ as follows

$$\Theta_1 = \begin{bmatrix} -\vec{\alpha}_1^H[1:d-1] & -\alpha_{1,d}^* \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}[1:d-1, 1:d-1] & \mathbf{Y}[1:d-1, d] \\ \mathbf{Y}[d, 1:d-1] & y_{d,d} \end{bmatrix}.$$

Obviously, from $\mathbf{G}_1 = \Theta_1 \mathbf{Y}$, we have

$$\mathbf{G}_1[2:d, 1:d-1] = \mathbf{Y}[1:d-1, 1:d-1]. \tag{A.3}$$

On the other hand, if we rewrite $\mathbf{Y}$ as follows

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & \mathbf{Y}[1, 2:d] \\ \mathbf{Y}[2:d, 1] & \mathbf{Y}[2:d, 2:d] \end{bmatrix}$$

then from $\mathbf{G}_1 = \mathbf{Y}\Theta_1$, we have (note that $\mathbf{Y}[2:d, 1] = \mathbf{0}$ from step 1)

$$\mathbf{G}_1[2:d, 1:d-1] = \mathbf{Y}[2:d, 2:d]. \tag{A.4}$$

By combining Eqn.(A.3) and Eqn.(A.4), we conclude that

$$y_{i,j} = y_{i+1,j+1} \tag{A.5}$$

for any $i \in \{1, \ldots, d-1\}, j \in \{1, \ldots, d-1\}$, which shows that $\mathbf{Y}$ has a Toeplitz form.

Step 3: Based on the results of previous steps, we know that all entries of $\mathbf{Y}$ on the main diagonal are equal, and all entries of $\mathbf{Y}$ off the main diagonal are zero. Therefore we can write $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero.

Appendix B

Proof of Lemma 5.3

We present our proof of Lemma 5.3 in the following three steps.

- In the first step, we will show that the entries of $\mathbf{Y}$ in first column except $y_{1,1}$ and the entries of $\mathbf{Y}$ in the last column except y_{d_i,d_j} are zero.
- Next, we will show that the entries in $\mathbf{Y}$ satisfy the following relationship: $y_{m,n} = y_{m+1,n+1}$, for any $m \in \{1, \ldots, d_i - 1\}, n \in \{1, \ldots, d_j - 1\}$, indicating that $\mathbf{Y}$ has a Toeplitz form.
- In the last step, we prove that $\mathbf{Y} = \lambda\mathbf{I}$ or $\mathbf{Y} = \mathbf{0}$ under the following four cases
  - If $d_i = d_j$, $\Theta_{1,i} = \Theta_{1,j}$ and $\Theta_{2,i} = \Theta_{2,j}$, then $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero.
  - If $d_i = d_j$, $\Theta_{1,i} \neq \Theta_{1,j}$ and $\Theta_{2,i} \neq \Theta_{2,j}$, then $\mathbf{Y} = \mathbf{0}$.
  - If $d_i > d_j$, then $\mathbf{Y} = \mathbf{0}$.
  - If $d_i < d_j$, and $|\alpha_{i,m_i}| \neq |\alpha_{j,d_j-t_i}|$, where $t_i \triangleq d_i - m_i$, α_{i,m_i} denotes the last non-zero entry in $\vec{\alpha}_i$, then $\mathbf{Y} = \mathbf{0}$. Such a condition $|\alpha_{i,m_i}| \neq$

$|\alpha_{j,d_j-t_i}|$ can be removed if there exists a non-zero entry for $\alpha_{j,k}, k \in \{d_j - t_i + 1, \ldots, d_j\}$.

Now we provide a complete proof step by step.

Step 1: For notational convenience, let $\mathbf{G}_1 \triangleq \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j}$, $\mathbf{G}_2 \triangleq \Theta_{2,i}\mathbf{Y} = \mathbf{Y}\Theta_{2,j}$, and $\mathbf{y}_m$ denotes the mth column of $\mathbf{Y}$. Consider the last column of $\mathbf{G}_1$, denoted by $\mathbf{G}_1[:, d_j]$, and the first column of $\mathbf{G}_2$, denoted by $\mathbf{G}_2[:, 1]$. We have

$$\begin{aligned}
\mathbf{G}_1[:, d_j] &= \begin{bmatrix} -\vec{\alpha}_i^H \mathbf{y}_{d_j} & y_{1,d_j} & \cdots & y_{d_i-1,d_j} \end{bmatrix}^T \\
&= -\alpha_{j,d_j}^* \begin{bmatrix} y_{1,1} & y_{2,1} & \cdots & y_{d_i,1} \end{bmatrix}^T
\end{aligned}$$

$$\begin{aligned}
\mathbf{G}_2[:, 1] &= \begin{bmatrix} y_{2,1} & \cdots & y_{d_i,1} & -\beta_i^H \mathbf{y}_1 \end{bmatrix}^T \\
&= -\alpha_{j,d_j} \begin{bmatrix} y_{1,d_j} & y_{2,d_j} & \cdots & y_{d_i,d_j} \end{bmatrix}^T
\end{aligned}$$

thus we obtain

$$y_{k,d_j} = |\alpha_{j,d_j}|^2 y_{k,d_j} \ \ \forall \ \ k \in \{1, \ldots, d_i - 1\} \tag{B.1}$$

$$y_{k+1,1} = |\alpha_{j,d_j}|^2 y_{k+1,1} \ \ \forall \ \ k \in \{1, \ldots, d_i - 1\}. \tag{B.2}$$

It is known that (see Theorem 1 in [67]) $|\alpha_{j,d_j}|$ will be less than one under the assumption that the $(d_j + 1) \times (d_j + 1)$ source autocorrelation matrix $\mathbf{R}_{s_j}[0]$ is positive definite. In fact, even if this assumption does not hold, the probability of $|\alpha_{j,d_j}| = 1$ is still almost equal to zero because α_{j,d_j} can be considered as a continuous random variable for the randomly generated colored source s_j, and

the probability of this random variable occurring at a specified point is zero. Therefore we can conclude that

$$y_{k,d_j} = 0 \qquad \forall\ k \in \{1, \ldots, d_i - 1\}$$
$$y_{k+1,1} = 0 \qquad \forall\ k \in \{1, \ldots, d_i - 1\}.$$

Step 2: Now we consider the sub-matrix of $\mathbf{G}_1$ from second row to d_i^{th} row and from first column to $(d_j - 1)^{th}$ column, denoted by $\mathbf{G}_1[2 : d_i, 1 : d_j - 1]$. This sub-matrix can be easily computed if we write $\Theta_{1,i}$ and $\mathbf{Y}$ as follows

$$\Theta_{1,i} = \begin{bmatrix} -\vec{\alpha}_i^H[1 : d_i - 1] & -\alpha_{i,d_i}^* \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}[1 : d_i - 1, 1 : d_j - 1] & \mathbf{Y}[1 : d_i - 1, d_j] \\ \mathbf{Y}[d_i, 1 : d_j - 1] & y_{d_i,d_j} \end{bmatrix}.$$

Obviously from $\mathbf{G}_1 = \Theta_{1,i}\mathbf{Y}$ we have

$$\mathbf{G}_1[2 : d_i, 1 : d_j - 1] = \mathbf{Y}[1 : d_i - 1, 1 : d_j - 1]. \tag{B.3}$$

On the other hand, we can write $\Theta_{1,j}$ and $\mathbf{Y}$ as

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & \mathbf{Y}[1, 2 : d_j] \\ \mathbf{Y}[2 : d_i, 1] & \mathbf{Y}[2 : d_i, 2 : d_j] \end{bmatrix}$$

$$\Theta_{1,j} = \begin{bmatrix} -\vec{\alpha}_j^H[1 : d_j - 1] & -\alpha_{j,d_j}^* \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$$

then from $\mathbf{G}_1 = \mathbf{Y}\Theta_{1,j}$ we have (note that $\mathbf{Y}[2:d_i,1] = \mathbf{0}$ from step 1)

$$\mathbf{G}_1[2:d_i, 1:d_j-1] = \mathbf{Y}[2:d_i, 2:d_j]. \tag{B.4}$$

By combining Eqn.(B.3) and Eqn.(B.4), we conclude that

$$y_{m,n} = y_{m+1,n+1} \tag{B.5}$$

for any $m \in \{1,\ldots,d_i-1\}, n \in \{1,\ldots,d_j-1\}$, which shows that $\mathbf{Y}$ has a Toeplitz form.

Step 3: This step is proved by dividing into the following four cases. Before proceeding, we summarize the results derived from previous two steps as follows

$$\begin{array}{lll} \text{(i)} & y_{m,n} = 0 & \text{if } m > n \\ \text{(ii)} & y_{m,n} = 0 & \text{if } n > m - (d_i - d_j) \\ \text{(iii)} & y_{m,n} = y_{m+1,n+1}. & \end{array} \tag{B.6}$$

1. If $d_i = d_j$, from Eqn.(B.6), it is easy to know that all entries on the main diagonal are equal, and all entries off the main diagonal are zero. Thus we conclude that $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero. The proof of case 1 is completed here.

2. From the analysis of Case 1, we can still write $\mathbf{Y} = \lambda\mathbf{I}$. Since $\Theta_{1,i} \neq \Theta_{1,j}$ and $\Theta_{2,i} \neq \Theta_{2,j}$, it is clear that $\lambda = 0$ and $\mathbf{Y} = \mathbf{0}$. The proof of case 2 is completed here.

3. Also from the results Eqn.(B.6), we can conclude that

$$\begin{aligned} y_{m,n} &= 0 \quad \text{if } m > n \\ y_{m,n} &= 0 \quad \text{if } n \geq m. \end{aligned}$$

The latter part of the above equation can be easily derived since any entry with $n \geq m$ also satisfies the condition $n > m-(d_i-d_j)$ when $d_i-d_j > 0$. Thus, we can conclude that $\mathbf{Y} = \mathbf{0}$. The proof of case 3 is completed here.

4. Now we consider the case where $d_i < d_j$. We assume that α_{i,m_i} is the last non-zero entry in $\vec{\alpha}_i$ and $\alpha_{i,k} = 0, k \in \{m_i+1, \ldots, d_i\}$. Let $t_i = d_i - m_i$, where $0 \leq t_i \leq d_i - 1$. Considering $\mathbf{G}_1 = \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j}$ and by using the results in Eqn.(B.6), we have

$$\begin{aligned} \mathbf{G}_1[1, d_j - t_i : d_j] &= \begin{bmatrix} -\alpha^*_{i,m_i} y_{d_i-t_i,d_j-t_i} & 0 & \ldots & 0 \end{bmatrix} \\ &= \begin{bmatrix} -\alpha^*_{j,d_j-t_i} y_{1,1} & \ldots & -\alpha^*_{j,d_j} y_{1,1} \end{bmatrix}. \end{aligned}$$

Obviously if there exists a non-zero entry for $\alpha_{j,k}, k \in \{d_j - t_i + 1, \ldots, d_j\}$, then we can infer that $y_{1,1} = 0$ and $y_{d_i-t_i,d_j-t_i} = y_{d_i,d_j} = 0$.

On the other hand, if $\alpha_{j,k} = 0, k \in \{d_j - t_i + 1, \ldots, d_j\}$, then we only have

$$\alpha^*_{j,d_j-t_i} y_{1,1} = \alpha^*_{i,m_i} y_{d_i-t_i,d_j-t_i}. \tag{B.7}$$

In this case, we need to consider $\mathbf{G}_2$. From $\mathbf{G}_2 = \Theta_{2,i}\mathbf{Y} = \mathbf{Y}\Theta_{2,j}$, we know that

$$\begin{aligned} \mathbf{G}_2[d_i, t_i + 1] &= -\alpha_{i,m_i} y_{t_i+1,t_i+1} \\ &= -\alpha_{j,d_j-t_i} y_{d_i,d_j}. \end{aligned} \tag{B.8}$$

Since $y_{1,1} = y_{t_i+1,t_i+1}$ and $y_{d_i-t_i,d_j-t_i} = y_{d_i,d_j}$, by combining Eqn.(B.7) and Eqn.(B.8), we have

$$\frac{|\alpha_{j,d_j-t_i}|^2}{|\alpha_{i,m_i}|^2} y_{d_i,d_j} = y_{d_i,d_j}. \tag{B.9}$$

Under the condition $|\alpha_{j,d_j-t_i}| \neq |\alpha_{i,m_i}|$, it is clear that we still have $y_{d_i,d_j} = 0$ and $y_{1,1} = 0$.

In the following, we will show that all remaining entries in $\mathbf{Y}$ are zero. Based on the relationship $y_{m,n} = y_{m+1,n+1}$, we only need to show that $\{y_{1,2}, \ldots, y_{1,d_j-d_i}\}$ or $\{y_{d_i,d_i+1}, \ldots, y_{d_i,d_j-1}\}$ are zero. This can be proved in an iterative way. Considering $\mathbf{G}_1 = \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j}$, we have

$$\mathbf{G}_1[1, d_j - t_i - 1] = -\alpha^*_{i,m_i} y_{d_i-t_i,d_j-t_i-1} = y_{1,d_j-t_i}. \tag{B.10}$$

Here y_{1,d_j-t_i} can be proved to be equal to zero. If $t_i = d_i - 1$, then $y_{1,d_j-t_i} = y_{d_i,d_j} = 0$; if $t_i < d_i - 1$, then y_{1,d_j-t_i} satisfies the condition $n > m - (d_i - d_j)$ in Eqn.(B.6), also $y_{1,d_j-t_i} = 0$. Therefore from Eqn.(B.10), we have

$$y_{d_i-t_i,d_j-t_i-1} = y_{d_i,d_j-1} = 0. \tag{B.11}$$

Now suppose $\{y_{d_i,d_i+k}, \ldots, y_{d_i,d_j-1}\}$ are zero, where $d_j - d_i - 1 > k > 1$, we need to prove that $y_{d_i,d_i+k-1} = 0$. Considering $\mathbf{G}_1 = \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j}$, we have

$$\begin{aligned} \mathbf{G}_1[1, d_i + k - 1 - t_i] &= -\alpha^*_{i,m_i} y_{d_i-t_i,d_i+k-1-t_i} \\ &= y_{1,d_i+k-t_i}. \end{aligned} \tag{B.12}$$

Also y_{1,d_i+k-t_i} can be proved to be equal to zero. If $t_i = d_i - 1$, then $y_{1,d_i+k-t_i} = y_{d_i,d_i+k} = 0$; if $t_i < d_i - 1$, y_{1,d_i+k-t_i} will be equal to some entry in $\{y_{d_i,d_i+k+1}, \ldots, y_{d_i,d_j}\}$ or satisfy the condition $n > m - (d_i - d_j)$

in Eqn.(B.6), still we have $y_{1,d_i+k-t_i} = 0$. Therefore

$$y_{d_i-t_i,d_i+k-1-t_i} = y_{d_i,d_i+k-1} = 0. \tag{B.13}$$

The proof of case 4 is completed here.

Remark: In case 4, we assumed that there exists at least one non-zero entry in $\vec{\alpha}_i$ and did not consider the case where $\vec{\alpha}_i = \mathbf{0}$. This is because we assume that the input signals are all colored. Hence, generally, $\vec{\alpha}_i \neq \mathbf{0}$ for $i \in \{1, \ldots, p\}$. In the case where $\vec{\alpha}_i = \mathbf{0}$, it can be easily verified (considering only $\mathbf{G}_1 = \Theta_{1,i}\mathbf{Y} = \mathbf{Y}\Theta_{1,j}$) that a sufficient condition for $\mathbf{Y} = \mathbf{0}$ is that there exists a non-zero entry for $\alpha_{j,k}, k \in \{d_j - d_i + 1, \ldots, d_j\}$. Such a condition can be further relaxed if we consider the relationship $\mathbf{G}_2 = \Theta_{2,i}\mathbf{Y} = \mathbf{Y}\Theta_{2,j}$ in the mean time.

Appendix C

Proof of Theorem 5.2

Obviously $\vec{l_e} = [L_1, \ldots, L_p]$ is a p-tuple point that can render us a non-zero solution $\{\mathcal{X}_1, \ldots, \mathcal{X}_p\} = \{\mathcal{H}_1, \ldots, \mathcal{H}_p\}$ satisfying Eqn.(5.35–5.36). Now we prove that for any other p-tuple point, $\vec{l} \neq \vec{l_e}$, there does not exist a non-zero solution $\{\mathcal{X}_1, \ldots, \mathcal{X}_p\}$ that satisfies Eqn.(5.35–5.36), where for each $i \in \{1, \ldots, p\}$, $\mathcal{X}_i \in \mathbb{C}^{(N+1)q \times (N+M_i+1)}$ is a non-zero block Toeplitz matrix.

We prove it by contradiction. Suppose for $\vec{l_m} = [T_1, \ldots, T_p]$, $\vec{l_m} \neq \vec{l_e}$, there exists a non-zero solution $\{\mathcal{G}_1, \ldots, \mathcal{G}_p\}$ that satisfies Eqn.(5.35–5.36), where for each $i \in \{1, \ldots, p\}$, $\mathcal{G}_i \in \mathbb{C}^{(N+1)q \times (N+T_i+1)}$ is a non-zero block Toeplitz matrix. Since $\vec{l_m} \neq \vec{l_e}$, we can always find i such that $T_i < L_i$ and also

$$\Upsilon_1 \mathcal{G}_i = \mathcal{G}_i \Theta_{1,i}(T_i) \tag{C.1}$$

$$\Upsilon_2 \mathcal{G}_i = \mathcal{G}_i \Theta_{2,i}(T_i). \tag{C.2}$$

By utilizing Eqn.(5.7), we have

$$\mathcal{H}\Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{1,i}(T_i) \Rightarrow \Theta_1\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{1,i}(T_i) \tag{C.3}$$

$$\mathcal{H}\Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{G}_i\Theta_{2,i}(T_i) \Rightarrow \Theta_2\mathcal{H}^\dagger\mathcal{G}_i = \mathcal{H}^\dagger\mathcal{G}_i\Theta_{2,i}(T_i) \tag{C.4}$$

where Θ_1 and Θ_2 can be written as

$$\Theta_1 = \mathrm{diag}(\Theta_{1,1}(L_1), \Theta_{1,2}(L_2), \cdots, \Theta_{1,p}(L_p)) \tag{C.5}$$

$$\Theta_2 = \mathrm{diag}(\Theta_{2,1}(L_1), \Theta_{2,2}(L_2), \cdots, \Theta_{2,p}(L_p)). \tag{C.6}$$

Let $\mathbf{X} \triangleq \mathcal{H}^\dagger\mathcal{G}_i \triangleq \begin{bmatrix} \mathbf{X}_1^T & \cdots & \mathbf{X}_p^T \end{bmatrix}^T$, where $\mathbf{X}_k \in \mathbb{C}^{d_k \times (N+T_i+1)}$, thus we have

$$\Theta_{1,k}(L_k)\mathbf{X}_k = \mathbf{X}_k\Theta_{1,i}(T_i) \ \ \forall \ \ k \in \{1, \ldots, p\} \tag{C.7}$$

$$\Theta_{2,k}(L_k)\mathbf{X}_k = \mathbf{X}_k\Theta_{2,i}(T_i) \ \ \forall \ \ k \in \{1, \ldots, p\}. \tag{C.8}$$

If $k \neq i$, since all sources satisfy the identifiability conditions IC1–IC3, it is easy to conclude that $\mathbf{X}_k = \mathbf{0}$ for any $k \neq i$ by applying Lemma 5.3; if $k = i$, since $L_i > T_i$, the dimension of $\Theta_{1,i}(L_k)$ is strictly greater than the dimension of $\Theta_{1,i}(T_i)$, by applying the third case in Lemma 5.3, we have $\mathbf{X}_i = \mathbf{0}$. Therefore $\mathcal{H}^\dagger\mathcal{G}_i = \mathbf{0}$. From the proof of Theorem 5.1, we know that the solution of $\mathcal{H}^\dagger\mathcal{G}_i = \mathbf{0}$ is unique and $\mathcal{G}_i = \mathbf{0}$. This result contradicts the assumption $\mathcal{G}_i \neq \mathbf{0}$ we made. The proof is completed here.

Appendix D

Proof of Lemma 7.1

For notational convenience, let $\mathbf{G}_1 \triangleq \mathbf{TY}$ and $\mathbf{G}_2 \triangleq \mathbf{YT}$. Also let $g^1_{i,j}$ denote the $(i,j)^{\text{th}}$ element of $\mathbf{G}_1$ and $g^2_{i,j}$ the $(i,j)^{\text{th}}$ element of $\mathbf{G}_2$, respectively.

Step 1: We first consider the first column of $\mathbf{G}_1$ and $\mathbf{G}_2$. Clearly, we have

$$g^1_{n,1} = t_{n,n} y_{n,1} \qquad g^2_{n,1} = y_{n,1} t_{1,1}.$$

Since $g^1_{n,1} = g^2_{n,1}$ and $t_{n,n} \neq t_{1,1}$, we have $y_{n,1} = 0$. By utilizing this result, we can further derive that

$$g^1_{n-1,1} = t_{n-1,n-1} y_{n-1,1} \qquad g^2_{n-1,1} = y_{n-1,1} t_{1,1}.$$

Also, we have $y_{n-1,1} = 0$ because $g^1_{n-1,1} = g^2_{n-1,1}$ and $t_{n-1,n-1} \neq t_{1,1}$. In this iterative way, by comparing $g^1_{k,1}$ with $g^2_{k,1}$ for $k \in \{2, \ldots, n-2\}$, it is not hard to conclude that

$$y_{k,1} = 0 \qquad \forall\ k \in \{2, \ldots, n\}. \tag{D.1}$$

Step 2: Now we proceed to consider the second column of $\mathbf{G}_1$ and $\mathbf{G}_2$. We have the following by using the derived results in step 1

$$g^1_{n,2} = t_{n,n} y_{n,2} \qquad g^2_{n,2} = y_{n,2} t_{2,2}.$$

Thus we have $y_{n,2} = 0$. By a similar iterative way as in step 1, we can conclude that

$$y_{k,2} = 0 \qquad \forall\ k \in \{3, \ldots, n\}. \tag{D.2}$$

Step 3: Now we assume that $y_{i,j} = 0$ if $j \leq m$ and $i > j$, where m is a certain value between 2 and $n-2$, we need to prove that $y_{k,m+1} = 0$ for $k \in \{m+2, \ldots, n\}$. Similarly as the previous steps, it is easy to derive that

$$g^1_{n,m+1} = t_{n,n} y_{n,m+1} \qquad g^2_{n,m+1} = y_{n,m+1} t_{m+1,m+1}.$$

Since $g^1_{n,m+1} = g^2_{n,m+1}$ and $t_{n,n} \neq t_{m+1,m+1}$, we have $y_{n,m+1} = 0$. And also in an iterative way, we can conclude that

$$y_{k,m+1} = 0 \qquad \forall\ k \in \{m+2, \ldots, n\}. \tag{D.3}$$

The proof is completed here.

Appendix E

Proof of Lemma 7.2

We present our proof in the following three steps.

Step 1: For notational convenience, let $\mathbf{G}_1 \triangleq \mathbf{J}_m\mathbf{Y} = \mathbf{Y}\mathbf{J}_n$, $\mathbf{G}_2 \triangleq \mathbf{J}_m^T\mathbf{Y} = \mathbf{Y}\mathbf{J}_n^T$. Considering the relationship of $\mathbf{G}_1$, we have

$$\begin{aligned}\mathbf{G}_1[2:m,n] &= \begin{bmatrix} y_{1,n} \ y_{2,n} \ \cdots \ y_{m-1,n} \end{bmatrix}^T \\ &= \begin{bmatrix} 0 \ 0 \ \cdots \ 0 \end{bmatrix}^T \end{aligned} \tag{E.1}$$

$$\begin{aligned}\mathbf{G}_1[1,1:n-1] &= \begin{bmatrix} 0 \ 0 \ \cdots \ 0 \end{bmatrix} \\ &= \begin{bmatrix} y_{1,2} \ y_{1,3} \ \cdots \ y_{1,n} \end{bmatrix}. \end{aligned} \tag{E.2}$$

Similarly, considering the relationship of $\mathbf{G}_2$, we have

$$\begin{aligned}\mathbf{G}_2[1:m-1,1] &= \begin{bmatrix} y_{2,1} \ y_{3,1} \ \cdots \ y_{m-1,1} \end{bmatrix}^T \\ &= \begin{bmatrix} 0 \ 0 \ \cdots \ 0 \end{bmatrix}^T \end{aligned} \tag{E.3}$$

$$\begin{aligned}\mathbf{G}_2[m, 2:n] &= \begin{bmatrix} 0 \; 0 \; \cdots \; 0 \end{bmatrix} \\ &= \begin{bmatrix} y_{m,1} \; y_{m,2} \; \cdots \; y_{m,n-1} \end{bmatrix}. \end{aligned} \tag{E.4}$$

Therefore we can conclude that all entries located at the edges of the matrix $\mathbf{Y}$ are zero except the entries $y_{1,1}$ and $y_{m,n}$.

Step 2: Now we consider the sub-matrix of $\mathbf{G}_1$ from second row to m^{th} row and from first column to $(n-1)^{th}$ column, denoted by $\mathbf{G}_1[2:m, 1:n-1]$. This sub-matrix can be easily computed as if we write $\mathbf{J}_m$ and $\mathbf{Y}$ as follows

$$\mathbf{J}_m = \begin{bmatrix} \mathbf{0}_{1\times(m-1)} & 0 \\ \mathbf{I}_{(m-1)\times(m-1)} & \mathbf{0}_{(m-1)\times 1} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}[1:m-1, 1:n-1] & \mathbf{Y}[1:m-1, n] \\ \mathbf{Y}[m, 1:n-1] & y_{m,n} \end{bmatrix}.$$

Obviously from $\mathbf{G}_1 = \mathbf{J}_m\mathbf{Y}$ we have

$$\mathbf{G}_1[2:m, 1:n-1] = \mathbf{Y}[1:m-1, 1:n-1]. \tag{E.5}$$

On the other hand, we can write $\mathbf{J}_n$ and $\mathbf{Y}$ as

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & \mathbf{Y}[1, 2:n] \\ \mathbf{Y}[2:m, 1] & \mathbf{Y}[2:m, 2:n] \end{bmatrix}$$

$$\mathbf{J}_n = \begin{bmatrix} \mathbf{0}_{1\times(n-1)} & 0 \\ \mathbf{I}_{(n-1)\times(n-1)} & \mathbf{0}_{(n-1)\times 1} \end{bmatrix}.$$

Then from $\mathbf{G}_1 = \mathbf{Y}\mathbf{J}_n$ we have

$$\mathbf{G}_1[2:m, 1:n-1] = \mathbf{Y}[2:m, 2:n]. \tag{E.6}$$

By combining Eqn.(E.5) and Eqn.(E.6), we can conclude that

$$y_{i,j} = y_{i+1,j+1} \tag{E.7}$$

for $i \in \{1, \ldots, m-1\}, j \in \{1, \ldots, n-1\}$, which shows that $\mathbf{Y}$ has a Toeplitz form.

Step 3: If $m = n$, based on the above derived results, it is easy to know that all entries on the main diagonal are equal, and all entries off the main diagonal are zero. Therefore we conclude that $\mathbf{Y} = \lambda\mathbf{I}$, where λ could be any complex scalar including zero. If $m \neq n$, since $\mathbf{Y}$ has a Toeplitz form and all entries located at the edges of the matrix $\mathbf{Y}$ are zero (note that $y_{1,1}$ and $y_{m,n}$ can be easily proved to be zero by utilizing the Toeplitz form when $m \neq n$), hence $\mathbf{Y} = \mathbf{0}$. The proof is completed here.

Appendix F

Proof of Theorem 7.3

We can derive the following

$$\mathbf{Q}^H\mathbf{R}_s[k]\mathbf{Q}\mathbf{C}_l = \mathbf{C}_l\mathbf{R}_{s_l}[k] \Rightarrow \mathbf{R}_s[k]\mathbf{Q}\mathbf{C}_l = \mathbf{Q}\mathbf{C}_l\mathbf{R}_{s_l}[k]. \tag{F.1}$$

Let $\mathbf{Z}_l \triangleq \mathbf{Q}\mathbf{C}_l$, we can rewrite the above equation as

$$\mathbf{R}_s[k]\mathbf{Z}_l = \mathbf{Z}_l\mathbf{R}_{s_l}[k] \qquad k \in \{\pm 1\}. \tag{F.2}$$

If we partition matrix $\mathbf{Z}_l$ as $\mathbf{Z}_l \triangleq \begin{bmatrix} \mathbf{Z}_{1l}^T & \mathbf{Z}_{2l}^T & \cdots & \mathbf{Z}_{pl}^T \end{bmatrix}^T$, where $\mathbf{Z}_{il} \in \mathbb{C}^{d_i \times d_l}$, then, by exploiting the block diagonal structure of $\mathbf{R}_s[k]$, we have the following

$$(a)\ \ \mathbf{J}_{d_i}\mathbf{Z}_{il} = \mathbf{Z}_{il}\mathbf{J}_{d_l} \quad (b)\ \ \mathbf{J}_{d_i}^T\mathbf{Z}_{il} = \mathbf{Z}_{il}\mathbf{J}_{d_l}^T \quad . \tag{F.3}$$

By utilizing the results of Lemma 7.2, we can conclude that

$$\mathbf{Z}_l = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \lambda_l\mathbf{I}_{d_l} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}^T. \tag{F.4}$$

Hence we have $\mathbf{C}_l = \mathbf{Q}^H\mathbf{Z}_l = \lambda_l\bar{\mathbf{Q}}_l$. The proof is completed here.

Publications

Journal Articles

published

1. Jun Fang, A. Rahim Leyman, Yong Huat Chew and Huiping Duan, "Blind SIMO FIR channel estimation by utilizing property of companion matrices", *IEEE Signal Processing Letters*, vol. 12, pp. 387-390, May. 2005.

2. Jun Fang, A. Rahim Leyman, Yong Huat Chew and Ying Chang Liang, "A cumulant interference subspace cancellation method for blind SISO channel estimation", *IEEE Trans. Signal Processing*, vol. 54, pp. 784-790, Feb. 2006.

3. Jun Fang, A. Rahim Leyman, Yong Huat Chew and Huiping Duan, "Blind channel identification with colored sources by exploiting properties of companion matrices", *IEEE Trans. Signal Processing*, vol. 54, pp. 894-906, Mar. 2006.

4. Jun Fang, A. Rahim Leyman and Yong Huat Chew, "Blind equalization of SIMO FIR channel driven by colored signals with unknown statistics", *IEEE Trans. on Signal Processing*, vol. 54, pp. 1998-2008, Jun. 2006.

under review

5. Jun Fang, A. Rahim Leyman, Yong Huat Chew and Huiping Duan, "Some further results on blind MIMO FIR channel identification via second-order statistics", submitted to *Signal Processing*, May, 2006.

Conference Proceedings

1. Jun Fang, A. Rahim Leyman and Yong Huat Chew, "A cumulant subspace projection method for blind MIMO FIR identification", *2004 IEEE International Conference on Speech, Acoustics, and Signal Processing*, Montreal, Canada, May 17-21, 2004.

2. Jun Fang, A. Rahim Leyman and Yong Huat Chew, "A subspace projection method for blind identification using shifted correlation matrices", *2004 IEEE Workshop on Signal Processing Advances in Wireless Communications*, Lisbon, Portugal, July 11-14, 2004.

3. Jun Fang, A. Rahim Leyman and Yong Huat Chew, "A new closed-form solution for blind MIMO FIR channel estimation with colored sources", *2005 IEEE International Conference on Speech, Acoustics, and Signal Processing*, Philadelphia, USA, March 18-23, 2005.

4. Jun Fang, A. Rahim Leyman and Yong Huat Chew, "Blind MIMO FIR channel identification by exploiting channel order disparity", *2006 IEEE International Conference on Speech, Acoustics, and Signal Processing*, Toulouse, France, May 14-19, 2006.

Bibliography

[1] S. Prakriya, "Eigenanalysis-based blind methods for identification, equalization, and inversion of linear time-invariant channels," *IEEE Trans. Signal Processing*, vol. 50, pp. 1525–1532, July 2002.

[2] M. C. Vanderveen, "Estimation of parametric channel models in wireless communication networks," Ph.D. dissertation, Stanford University, 1997.

[3] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[4] D. C. Cox, R. Murray, and A. Norris, "800 MHz attenuation measured in and around suburban houses," *AT&T Bell Laboratory Tech. J.*, vol. 673, no. 6, July-August 1984.

[5] G. Stuber, *Principles of mobile communication.* Kluwer, Boston, 1996.

[6] T. S. Rappaport, S. Y. Seidel, and R. Singh, "900 MHz multipath propagation measurements for U.S. Digital Cellular Radiotelephone," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 132–139, May 1990.

[7] S. Wales, "Techniques for cochannel interference suppression in TDMA mobile radio systems," *IEE Proc. Communications*, vol. 142, pp. 106–114, Apr. 1995.

[8] Y. Sato, "A method of self-recovering equalization for multilevel amplitude modulation," *IEEE Trans. Commun.*, vol. 23, pp. 679–682, June 1975.

[9] B. Juang, R. J. Perdue, Jr., and D. L. Thomson, "Deployable automatic speech recognition systems: advances and challenges," *AT&T Tech. J.*, pp. 45–55, Apr. 1995.

[10] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Mag.*, vol. 13, pp. 43–64, May 1996.

[11] Y. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," *IEEE Trans. Image Processing*, vol. 5, pp. 416–428, Mar. 1996.

[12] L. C. Wood and S. Treitel, "Seismic signal processing," *Proc. IEEE*, vol. 63, pp. 649–661, Apr. 1975.

[13] S. Haykin, Ed., *Array signal processing.* Englewood Cliffs, NJ: Prentice-Hall, 1985.

[14] B. C. Ng, D. Gesbert, and A. Paulraj, "A semi-blind approach to structured channel equalization," *IEEE ICASSP, Seattle, Washington*, vol. 6, pp. 3385–3388, May 1998.

[15] E. D. Carvalho and D. T. M. Slock, "Cramer-rao bounds for semi-blind, blind and training sequence based channel estimation," *IEEE SPAWC, Paris, France*, pp. 129–132, Apr. 1997.

[16] A. Paulraj, B. Ottersten, R. Roy, L. Swindlehurst, G. Xu, and T. Kailath, *Handbook of Statistics, Vol. 10, Signal Processing and Applications.* Elsevier Press, 1993.

[17] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[18] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.

[19] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 459–472, Apr. 1983.

[20] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum-phase system using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360–377, Mar. 1989.

[21] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher-order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1307–1317, July 1990.

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

[23] K. A. Meraim, P. Loubaton, and E. Moulines, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 694–705, Mar. 1997.

[24] Z. Ding, "Matrix outer product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Processing*, vol. 45, pp. 3053–3061, Dec. 1997.

[25] G. B. Giannakis and S. D. Halford, "Blind fractionally-spaced equalizers of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. Signal Processing*, vol. 45, pp. 2277–2292, Sept. 1997.

[26] D. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," *Proc. IEEE ICASSP, Adelaide, Australia*, pp. 585–588, Apr. 1994.

[27] K. A. Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Inform. Theory*, vol. 43, pp. 499–511, Mar. 1997.

[28] A. Gorokhov and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 813–820, Sept. 1997.

[29] K. A. Meraim and Y. Hua, "Blind identification of multi-input multi-output system using minimum noise subspace," *IEEE Trans. Signal Processing*, vol. 45, pp. 254–258, Jan. 1997.

[30] A. Gorokhov, P. Loubaton, and E. Moulines, "Second-order blind equalization in multiple input multiple output FIR systems: A weighted least squares approach," *Proc. ICASSP, Atlanta, GA*, pp. 2415–2418, May 1996.

[31] A. Gorokhov and P. Loubaton, "Blind identification of MIMO FIR systems: A generalized prediction approach," *Signal Processing*, vol. 73, pp. 105–124, 1999.

[32] J. K. Tugnait and B. Huang, "Multistep linear predictors based blind identification and equalization of multiple-input multiple-output channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 26–38, Jan. 2000.

[33] Z. Ding and L. Qiu, "Blind MIMO channel identification from second order statistics using rank deficient channel convolution matrix," *IEEE Trans. Signal Processing*, vol. 51, pp. 535–544, Feb. 2003.

[34] J. Shen and Z. Ding, "Direct blind MMSE channel equalization based on second order statistics," *IEEE Trans. Signal Processing*, vol. 48, pp. 1015–1022, Apr. 2000.

[35] J. K. Tugnait and B. Huang, "On a whitening approach to partial channel estimation and blind equalization of FIR/IIR multiple-input multiple-output channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 832–845, Mar. 2000.

[36] 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 Processing*, vol. 51, pp. 434–441, Feb. 2003.

[37] K. I. Diamantaras, A. P. Petropulu, and B. Chen, "Blind two-input-two-output FIR channel identification based on frequency domain second-order statistics," *IEEE Trans. Signal Processing*, vol. 48, pp. 534–542, Feb. 2000.

[38] B. Chen and A. P. Petropulu, "Frequency domain blind MIMO system identification based on second and higher order statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 1677–1688, Aug. 2001.

[39] I. Bradaric, A. P. Petropulu, and K. I. Diamantaras, "Blind MIMO FIR channel identification based on second order spectra correlations," *IEEE Trans. Signal Processing*, vol. 51, pp. 1668–1674, June 2003.

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

[41] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice Hall, 1980.

[42] Y. Hua and J. K. Tugnait, "Blind identifiability of FIR-MIMO systems with colored inputs using second order statistics," *IEEE Signal Processing Lett.*, vol. 7, pp. 348–350, Dec. 2000.

[43] J. Liang and Z. Ding, "Blind MIMO system identification based on cumulant subspace decomposition," *IEEE Trans. Signal Processing*, vol. 51, pp. 1457–1468, June 2003.

[44] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[45] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation," *IEEE Signal Processing Lett.*, vol. 1, pp. 205–207, Dec. 1994.

[46] ——, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 45, pp. 2714–2723, Nov. 1995.

[47] A. J. V. der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 47, pp. 173–190, Jan. 1997.

[48] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, Mar. 1996.

[49] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Processing*, vol. 45, pp. 2307–2317, Sept. 1997.

[50] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.

[51] M. Wax and T. Kailath, "Detection of signals by information theoretic criterion," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 387–392, Apr. 1985.

[52] L. Tong and Q. Zhao, "Joint order detection and blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, pp. 2345–2355, Sept. 1999.

[53] Q. Zhao and L. Tong, "Adaptive blind channel estimation by least squares smoothing," *IEEE Trans. Signal Processing*, vol. 47, pp. 3000–3012, Nov. 1999.

[54] K. I. Diamantaras, "Blind channel identification based on the geometry of the received signal constellation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1133–1143, May 2002.

[55] Q. Li, E.-W. Bai, and Y. Ye, "Blind channel equalization and ϵ-approximation algorithms," *IEEE Trans. Signal Processing*, vol. 49, pp. 2823–2831, Nov. 2001.

[56] L. He, M. G. Amin, C. Reed, and R. C. Malkemes, "A hybrid adaptive blind equalization algorithm for QAM signals in wireless communications," *IEEE Trans. Signal Processing*, vol. 52, pp. 2058–2069, July 2004.

[57] I. Santamaría, C. Pantaleón, L. Vielva, and J. Ibánez, "Blind equalization of constant modulus signals using support vector machines," *IEEE Trans. Signal Processing*, vol. 52, pp. 1773–1782, June 2004.

[58] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array – part I: Algorithms," *IEEE Trans. Signal Processing*, vol. 44, pp. 1184–1197, May 1996.

[59] 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, pp. 176–178, Sept. 1995.

[60] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.

[61] L. K. Hansen and G. Xu, "A hyperplane-based algorithm for the digital co-channel communications problem," *IEEE Trans. Signal Processing*, vol. 43, pp. 1536–1548, Sept. 1997.

[62] 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, pp. 1–10, Jan. 2003.

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

[64] J. K. Tugnait, "Multistep linear predictors-based blind equalization of FIR/IIR single-input multiple-output channels with common zeros," *IEEE Trans. Signal Processing*, vol. 47, pp. 1689–1700, June 2000.

[65] D. Gesbert and P. Duhamel, "Robust blind identification and equalization based on multi-step predictors," *Proc. ICASSP, Munich, Germany*, pp. 2621–2624, 1997.

[66] K. H. Afkhamie and Z.-Q. Luo, "Blind identification of FIR systems driven by markov-like input signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 1726–1736, June 2000.

[67] R. López-Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," *IEEE Trans. Signal Processing*, vol. 49, pp. 2050–2059, Sept. 2001.

[68] J. Fang, A. R. Leyman, Y. H. Chew, and H. Duan, "Blind SIMO FIR channel estimation by utilizing property of companion matrices," *IEEE Signal Processing Lett.*, vol. 12, pp. 387–390, May 2005.

[69] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences – Applicability to blind equaliztion," *IEEE Trans. Signal Processing*, vol. 43, pp. 3501–3506, Dec. 2000.

[70] Y. Hua, S. An, and Y. Xiang, "Blind identification of FIR channels by decorrelating subchannels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1143–1155, May 2003.

[71] T. Ma, Z. Ding, and S. F. Yau, "A two-stage algorithm for MIMO blind deconvolution of colored input signals," *IEEE Trans. Signal Processing*, vol. 48, pp. 1187–1192, Apr. 2000.

[72] J. Xavier, V. Barroso, and J. Moura, "Closed-form correlative coding blind identification of MIMO channels: Isometry fitting to second order statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 1073–1086, May 2001.

[73] I. Bradaric, A. Petropulu, and K. Diamantaras, "Blind MIMO FIR channel identification based on second-order spectra correlations," *IEEE Trans. Signal Processing*, vol. 51, pp. 1668–1674, June 2003.

[74] J. Fang, A. R. Leyman, Y. H. Chew, and H. Duan, "Blind channel identification with colored sources by exploiting properties of companion matrices," *IEEE Trans. Signal Processing*, vol. 54, pp. 894–906, Mar. 2006.

[75] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Blind channel identification and equalization using periodic modulation precoders: Performance analysis," *IEEE Trans. Signal Processing*, vol. 48, pp. 1570–1586, June 2000.

[76] C. A. Lin and J. Y. Wu, "Blind identification with periodic modulation: A time-domain approach," *IEEE Trans. Signal Processing*, vol. 50, pp. 2875–2888, Nov. 2002.

[77] M. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Processing*, vol. 45, pp. 1785–1794, July 1997.

[78] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. Part I. unification and optimal designs. Part II. blind channel estimation, synchronization, and direct equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2022, July 1999.

[79] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Processing*, vol. 47, pp. 572–578, Feb. 1999.

[80] B. Porat and B. Friedlander, "FIR system identification using fourth-order cumulants with applications to channel equalization," *IEEE Trans. Signal Processing*, vol. 38, pp. 1394–1398, Sept. 1993.

[81] B. Friedlander and B. Porat, "Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 27–35, Jan. 1990.

[82] J. K. Tugnait, "Blind equalization and estimation of digital communication FIR channels using cumulant matching," *IEEE Trans. Commun.*, vol. 43, pp. 1240–1245, Apr. 1995.

[83] D. Boss, B. Jelonnek, and K. D. Kammeyer, "Eigenvector algorithm for blind MA system identification," *Signal Processing*, vol. 66, pp. 1–26, Apr. 1998.

[84] G. B. Giannakis and J. M. Mendel, "Identification of nonminimum-phase system using higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 360–377, Mar. 1989.

[85] A. Swami and J. M. Mendel, "ARMA parameter estimation using only output cumulants," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1257–1265, July 1990.

[86] J. K. Tugnait, "Approaches to FIR system identification with noisy data using higher-order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1307–1317, July 1990.

[87] J. A. R. Fonollosa and J. Vidal, "System identification using a linear combination of cumulant slices," *IEEE Trans. Signal Processing*, vol. 41, pp. 2405–2411, July 1993.

[88] L. Srinivas and K. V. S. Hari, "FIR system identification using higher-order cumulant - a generalized approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 3061–3065, Dec. 1995.

[89] ——, "FIR system identification based on subspaces of a higher order cumulant matrix," *IEEE Trans. Signal Processing*, vol. 44, pp. 1485–1491, June 1996.

[90] Z. Ding and J. Liang, "A cumulant matrix subspace algorithm for blind single FIR channel identification," *IEEE Trans. Signal Processing*, vol. 49, pp. 325–333, Feb. 2001.

[91] G. B. Giannakis, Y. Inouye, and J. M. Mendel, "Cumulant based identification of multichannel moving-average models," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 783–787, July 1989.

[92] L. Tong, Y. Inouye, and R. Liu, "A finite-step global convergence algorithm for the parameter estimation of multichannel MA processes," *IEEE Trans. Signal Processing*, vol. 40, pp. 2547–2558, Oct. 1992.

[93] A. Swami, G. B. Giannakis, and S. Shamsunder, "Multichannel ARMA processes," *IEEE Trans. Signal Processing*, vol. 42, pp. 898–914, Apr. 1994.

[94] L. Tong, "Identification of multichannel MA parameters using higher-order statistics," *Signal Processing*, vol. 53, pp. 195–209, 1996.

[95] J. Liang and Z. Ding, "Blind MIMO system identification based on cumulant subspace decomposition," *IEEE Trans. Signal Processing*, vol. 51, pp. 1457–1468, June 2003.

[96] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Processing*, vol. 45, pp. 658–672, Mar. 1997.

[97] ——, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Processing*, vol. 45, pp. 268–271, 1997.

[98] Z. Xu and B. P. Ng, "Deterministic linear prediction methods for blind channel estimation based on dual concept of zero-forcing equalization," *IEEE Trans. Signal Processing*, vol. 50, pp. 2855–2865, Nov. 2002.

[99] J. Fang, A. R. Leyman, and Y. H. Chew, "Blind equalization of SIMO FIR channel driven by colored signals with unknown statistics," *IEEE Trans. Signal Processing*, vol. 54, pp. 1998–2008, June 2006.

[100] G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear FIR volterra channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 67–81, Jan. 1997.

[101] R. López-Valcarce and S. Dasgupta, "Blind equalization of nonlinear channels from second-order statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 3084–3097, Dec. 2001.

[102] J. Fang, A. R. Leyman, and Y. H. Chew, "Some further results on blind MIMO FIR channel identification via second-order statistics," *Submmited to Signal Processing*, May 2006.

[103] J. K. Tugnait, "On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics," *IEEE Trans. Inform. Theory*, vol. 41, pp. 308–311, Jan. 1995.

[104] J. Fang, A. R. Leyman, Y. H. Chew, and Y. C. Liang, "A cumulant interference subspace cancellation method for blind SISO channel estimation," *IEEE Trans. Signal Processing*, vol. 54, pp. 784–790, Feb. 2006.

[105] A. K. Nandi, *Blind estimation using higher-order statistics.* Boston: Kluwer Academic Publishers, 1999.

[106] O. Shalvi and E. Weinstein, "New criterion for blind deconvolution of non-minimum phase systems," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, 1990.

[107] J. F. Cardoso, "Eigen-structure to the fourth-order cumulant tensor with application to the blind source separation problem," *Proc. IEEE ICASSP*, vol. 5, pp. 2655–2658, Apr. 1990.

[108] C. L. Nikias and A. P. Petropulu, *Higher-order spectra analysis: A nonlinear signal processing framework*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[109] A. Papoulis, *Probability, random variables, and stochastic processes*, 3rd ed. New York: McGraw-Hill, 1991.

[110] J. Liang and Z. Ding, "FIR channel estimation through generalized cumulant slice weighting," *IEEE Trans. Signal Processing*, vol. 52, pp. 657–667, Mar. 2004.

[111] ——, "Nonminimum-phase FIR channel estimation using cumulant matrix pencils," *IEEE Trans. Signal Processing*, vol. 51, pp. 2310–2319, Sept. 2003.

[112] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, 1996.

[113] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. on Automatic Control*, vol. 19, pp. 716–723, Dec. 1974.

[114] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[115] A. Belouchrani, K. A. Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, Feb. 1997.

[116] C. B. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction," *IEEE Trans. Signal Processing*, vol. 47, pp. 641–654, Mar. 1999.

[117] M. K. Tsatsanis and Z. Xu, "Constrained optimization methods for direct blind equalization," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 424–433, Mar. 1999.

[118] R. López-Valcarce, Z. Ding, and S. Dasgupta, "Equalization and interference cancellation in linear multiuser systems based on second-order statistics," *IEEE Trans. Signal Processing*, vol. 49, pp. 2042–2049, Sept. 2001.

[119] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.