2005

Interference suppression and parameter estimation in wireless communication systems over time-varing multipath fading channels

Jianqiang He
Louisiana State University and Agricultural and Mechanical College

INTERFERENCE SUPPRESSION AND PARAMETER ESTIMATION IN WIRELESS COMMUNICATION SYSTEMS OVER TIME-VARYING MULTIPATH FADING CHANNELS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Electrical and Computer Engineering

by
Jianqiang He
B.S., Shenyang Institute of Aeronautical Engineering, 1998
M.S., Louisiana State University, 2001
May 2005

Acknowledgements

My sincere appreciation and thanks are given to my adviser and committee chair Dr. Guoxiang Gu for his patient academic guidance and constant input throughout the research and preparation of this dissertation. His insight and expertise in this field have deeply influenced me and my work recorded herein. Without his constructive direction and invaluable advice, this work would not have been completed.

My gratitude is also extended to Dr. Morteza Naraghi-Pour and Dr. Jagannathan Ramanujam for their valuable courses and insightful discussions, from which I will continue to benefit in my career.

I would like to thank Dr. Ambar Sengupta and Dr. Rajgopal Kannan for their willingness to serve as my examining committee member and their sincere encouragement. I am grateful to Dr. Sengupta for taking his time and effort to be my minor professor. His course empowers me with the mathematical skills necessary for this dissertation and gives me confidence.

Thanks also go to my colleagues Zhongshan Wu, Bin Fu, Xinjia Chen, Xiang Gao, Phalguna Kumar Rachinayani, Nike Liu, Shuguang Hao etc. for their assistance. I will always remember the happy time we share together at LSU.

Finally, I want to express my deepest love for my wife Yang, whose endless love, understanding, and support during all these years are the greatest assets in my life.

Table of Contents

Notation and Symbols

$\mathbf{A}_{M\times N}$:	M-row N-column matrix
$\mathbf{A}^{-1}$:	Inverse of $\mathbf{A}$
Tr($\mathbf{A}$):	Trace of $\mathbf{A}$, Tr($\mathbf{A}$)=$\sum_i \mathbf{A}_{ii}$
$\mathbf{A}^T$:	Transpose of $\mathbf{A}$
$\mathbf{A}^*$:	Complex conjugate transpose of $\mathbf{A}$
$\mathbf{I}_N$:	Identity matrix of size $N \times N$
$\star$:	Linear convolution
Quan($\cdot$):	Quantization operation
E[$\cdot$]:	Expectation
Var[$\cdot$]:	Variance

List of Acronyms

AR:	Autoregressive
BER:	Bit error rate
CDMA:	Code division multiple access
CSI:	Channel state information
DFE:	Decision-feedback equalizer
FDMA:	Frequency division multiple access
FIR:	Finite impulse response
FL:	Finite length
GSM:	Global system of mobile communications
ICI:	Inter-channel interference, or Inter-carrier interference
IIR:	Infinite impulse response
ISI:	Inter-symbol interference
KF:	Kalman filter
LAN:	Local area network
LE:	Linear equalizer
LTI:	Linear time invariant
MAI:	Multiple access interference
MAN:	Metropolitan area network
MIMO:	Multi-input, multi-output
MLSE:	Maximum-likelihood sequence estimator
MMSE:	Minimum mean squared error
MUI:	Multiuser interference
OFDM:	Orthogonal frequency division multiplexing
SIC:	Successive interference cancellation
SISO:	Single-input single-output

SNR: Signal to noise ratio

SOS: Second order statistics

WSS: Wide-sense stationary

ZF: Zero-forcing

Abstract

This dissertation focuses on providing solutions to two of the most important problems in wireless communication systems design, namely, 1) the interference suppression, and 2) the channel parameter estimation in wireless communication systems over time-varying multipath fading channels. We first study the interference suppression problem in various communication systems under a unified multirate transmultiplexer model. A state-space approach that achieves the optimal realizable equalization (suppression of inter-symbol interference) is proposed, where the Kalman filter is applied to obtain the minimum mean squared error estimate of the transmitted symbols. The properties of the optimal realizable equalizer are analyzed. Its relations with the conventional equalization methods are studied. We show that, although in general a Kalman filter has an infinite impulse response, the Kalman filter based decision-feedback equalizer (Kalman DFE) is a finite length filter. We also propose a novel successive interference cancellation (SIC) scheme to suppress the inter-channel interference encountered in multi-input multi-output systems. Based on spatial filtering theory, the SIC scheme is again converted to a Kalman filtering problem. Combining the Kalman DFE and the SIC scheme in series, the resultant two-stage receiver achieves optimal realizable interference suppression. Our results are the most general ever obtained, and can be applied to any linear channels that have a state-space realization, including time-invariant, time-varying, finite impulse response, and infinite

impulse response channels. The second half of the dissertation devotes to the parameter estimation and tracking of single-input single-output time-varying multipath channels. We propose a novel method that can blindly estimate the channel second order statistics (SOS). We establish the channel SOS identifiability condition and propose novel precoder structures that guarantee the blind estimation of the channel SOS and achieve diversities. The estimated channel SOS can then be fit into a low order autoregressive (AR) model characterizing the time evolution of the channel impulse response. Based on this AR model, a new approach to time-varying multipath channel tracking is proposed.

Chapter 1
Introduction

The ultimate goal of communication systems design is to achieve four *W*'s, i.e. the exchange of *W*hatever information regardless of *W*hoever, *W*henever, and *W*herever there is need of communications. With the advance of wireless communications, we are getting closer to this goal. However, the high data rate requirement of modern communication systems and the time-varying fading nature of the wireless channels also make the task of wireless communication systems design more challenging than before.

This dissertation focuses on solving two of the most important problems in wireless communication transceiver design, namely (1) the interference suppression in wireless communication systems, and (2) the parameter estimation of time-varying multipath fading channels.

The results presented in this dissertation are expected to provide effective solutions to performance improvement in many of the contemporary and future wireless communication systems, such as TDMA based and CDMA based cellular systems, OFDM based wireless LAN and wireless MAN systems, multiple antenna systems, to name a few.

1.1 Overview

1.1.1 Unified Mathematical Description of Wireless Communication Systems

Information theory and coding theory are the cornerstones of modern communication technologies. However, many of the results obtained in information theory and coding theory have assumed ideal channel conditions. In other words, no system or channel induced interference has been considered. In reality, the performance of all the practical communication systems are limited by various interferences. For example, the TDMA based GSM cellular system is subject to inter-symbol interference (ISI) that is introduced by the multipath wireless channels; the capacity and performance of the CDMA based cellular systems are limited by the multiple-access interference (MAI, also known as multi-user interference or MUI) that is the direct result of the CDMA system design; multi-input multi-output (MIMO) systems such as multiple antenna systems and OFDM systems[1] are also affected by the inter-channel interference or inter-carrier interference (both are known as ICI).

To approach the capacity and performance limits predicted by information theorists and coding researchers, effective interference suppression in communication systems is indispensable. Although different types of interferences have quite different causes, we are fortunate enough to be able to study them within a unified framework, as will be shown in the following.

[1] An OFDM system can be regarded as a virtual MIMO systems

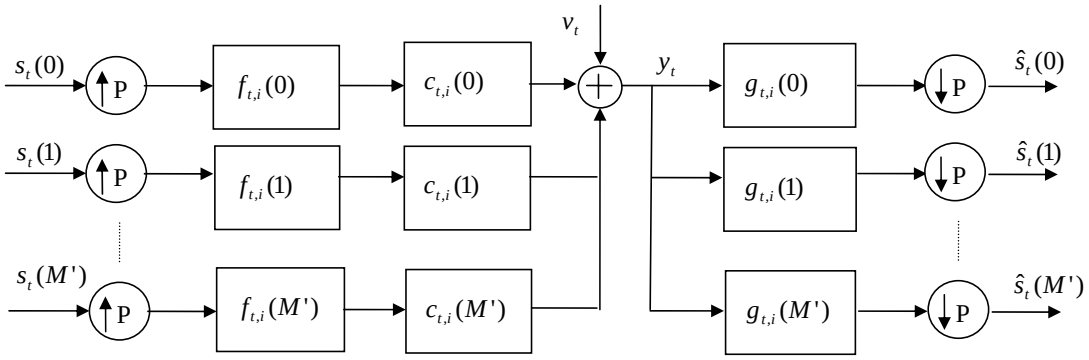

Figure 1.1: Time-varying multirate transmultiplexer model for multiple access systems in the base band ($M' = M$).

Figure 1.1 shows a time-varying multirate transmultiplexer model for multiple access systems, where M is the number of simultaneous users and $M' = M - 1$, $\uparrow P$ is the P-factor up sampler which inserts $P - 1$ zeros after each signal sample [77]. Denote $s_t(m)$ to be the mth user's signal at time t. Note that time t is discrete time and hence takes integer values. The signal $s_t(m)$ is first processed by a $(P-1)$-th order precoding filter, whose impulse response coefficients $\{f_{t,i}(m)\}_{i=0}^{P-1}$, $m = 0, 1, \cdots, M-1$, are determined by the multiplexing scheme of the system. The precoded signal is then transmitted through a possibly time-varying multipath fading channel. Let $\{c_{t,i}(m)\}_{i=0}^{L_{c,m}}$, $m = 0, 1, \cdots, M-1$, denote the impulse response coefficients of the physical channel of the mth user at time t, where $L_{c,m}$ is the mth user's physical channel order. Then the received signal at the receiver side can be written as

$$y_t = \sum_{m=0}^{M-1} \sum_{i=0}^{L_m} s_{t-i}(m) h_{t,iP}(m) + v_t, \tag{1.1}$$

where v_t is the additive white Gaussian noise with mean zero and variance σ_v^2,

$$\begin{aligned} h_{t,i}(m) &= f_{t,i}(m) \star c_{t,i}(m) \\ &= \sum_{k=0}^{P-1} c_{t,i-k}(m) f_{t,k}(m) \end{aligned} \tag{1.2}$$

is the ith coefficient of the composite system consisting of the precoding filter and the physical channel at time t, with $L_m = L_{c,m} + P$ the order of the composite system, and $\star$ denotes the linear convolution operation. Hence $\{h_{t,i}(m)\}_{i=0}^{L_m}$ can be regarded as an L_m-th order equivalent channel between the input signal $s_t(m)$ and the received signal y_t.

Here we must point out that the multirate transmultiplexer model in Figure 1.1 is so versatile that we can use this model to describe different multiplexing schemes by choosing different precoding filters [64], [85]. Obviously, a multiple antenna system is just a special case of the transmultiplexer model in Figure 1.1 with $P = 1$ and $f_{t,i}(m) = \delta_i$, where δ_i is the Kronecker delta. For TDMA systems, we can choose $P \geq M$ and $f_{t,i}(m) = \delta_{i-m}$. In this case P is the maximum number of users supported by the system. For FDMA systems, again we can choose $P \geq M$ with P the maximum number of users supported. But the precoding filter should be $f_{t,i}(m) = \exp(j\omega_m i)$, where ω_m is the mth user's carrier frequency. For CDMA systems, $P \geq M$ and $f_{t,i}(m) = \alpha_i(m)$ can be chosen, where $\alpha_i(m)$ denotes the mth user's discrete time equivalent spread spectrum code at the chip rate, and in this case P is the so-called processing gain [23], [75]. It can be shown that OFDM systems can also be described by the multirate transmultplexer model with the precoding filter

$$f_{t,i}(m) = \exp(j2\pi mi/M), \tag{1.3}$$

where $m \in [0, M-1]$, $i \in [0, P-1]$, and $P = M + \overline{L}$ with $\overline{L} \geq L$ an upper bound of the equivalent composite channel order L [64], [85]. In light of Figure 1.1, a unified mathematical description can be obtained for different communication systems such as multiple antenna systems, OFDM system, and various multiple access systems.

In Figure 1.1, because the sampling rate through the channel is P times of the symbol rate, we can further simplify the expression in (1.1) by blocking the input and output signals of the equivalent channel into vectors as below,

$$\mathbf{s}_t = [s_t(0)\ s_t(1)\ \cdots\ s_t(M-1)]^T, \tag{1.4}$$

$$\mathbf{y}_t = [y_{tP}\ y_{tP+1}\ \cdots\ y_{tP+P-1}]^T, \tag{1.5}$$

$$\mathbf{v}_t = [v_{tP}\ v_{tP+1}\ \cdots\ v_{tP+P-1}]^T. \tag{1.6}$$

Then it is straightforward to verify that (1.1) is equivalent to the following multi-input multi-output system model [8]

$$\mathbf{y}_t = \sum_{i=0}^{L} \mathbf{H}_{t,i}\mathbf{s}_{t-i} + \mathbf{v}_t, \tag{1.7}$$

where $L = \max\{L_0, L_1, \cdots, L_{M-1}\}$, and

$$\mathbf{H}_{t,i} = \begin{bmatrix} h_{tP,iP}(0) & h_{tP,iP}(1) & \cdots & h_{tP,iP}(M-1) \\ h_{tP+1,iP}(0) & h_{tP+1,iP}(1) & \ddots & h_{tP+1,iP}(M-1) \\ \vdots & \ddots & \ddots & \vdots \\ h_{t'P-1,iP}(0) & h_{t'P-1,iP}(1) & \cdots & h_{t'P-1,iP}(M-1) \end{bmatrix}, \tag{1.8}$$

with $t' = t+1$. We must note that the input-output relation in (1.7) is a simplified general description of many different wireless communication systems. Hence the channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ in (1.7) may not necessarily represent the physical channel.

The system model in (1.7) will be the basis of the subsequent chapters, based on which, we will study the problem of optimal realizable interference suppression in wireless communication systems. New interference suppression methods will be proposed.

1.1.2 Interferences and Interference Suppression in Wireless Communication Systems

In this subsection, we will introduce the concepts of different interferences in a wireless communication system.

Consider a special case of (1.7) where $P = M = 1$. In this case, the MIMO model reduces to a single-input single-output (SISO) model and the received signal can be wriiten as

$$y_t = \sum_{i=0}^{L} h_{t,i} s_{t-i} + v_t. \tag{1.9}$$

The communication receiver aims to recover the transmitted signal s_t from the received signal y_t. Suppose that the channel coefficients $\{h_{t,i}\}_{i=0}^{L}$ are perfectly known at the receiver side. Then we can rewrite (1.9) as

$$y_t = h_{t,d} s_{t-d} + \sum_{i=0, i\neq d}^{L} h_{t,i} s_{t-i} + v_t, \tag{1.10}$$

where s_{t-d} is the symbol to be detected at time t, and integer $d \in [0, L]$ is the detection delay to be specified. In (1.10), the second term on the right hand side of the equation is due to the contribution of all the other symbols that are not of interest at time t. Its existence will interfere with the successful detection of the desired symbol s_{t-d} even if there is no noise in presence. Since the interference in this case comes from other

symbols, it is known as inter-symbol interference (ISI) in communication systems. From (1.10), we can see that as long as the channel has multipath, i.e. $L \geq 1$, there exists ISI.

For MIMO channels, there is another source of interference. Consider the special case of (1.10), where $L = 0$ and $P = M \geq 2$. The input-output relation can be written as

$$\mathbf{y}_t = \mathbf{H}_{t,0}\mathbf{s}_t + \mathbf{v}_t. \tag{1.11}$$

Since the channel has only one path in this case, there is no ISI in the system. However there may exist interference between the symbols in the signal vector $\mathbf{s}_t$. Written element wise, Equation (1.11) can be expressed as

$$\begin{aligned} y_t(m) &= \textstyle\sum_{i=0}^{M-1} h_{t,0}(i,m)s_t(i) + v_t(m) \\ &= h_{t,0}(m,m)s_t(m) + \textstyle\sum_{i=0,i\neq m}^{M-1} h_{t,0}(i,m)s_t(i) + v_t(m), \end{aligned} \tag{1.12}$$

where $h_{t,0}(i,m)$ is the $(i+1, m+1)$th entry in matrix $\mathbf{H}_{t,0}$, $i, m = 0, 1, \cdots, M-1$. Due to the concurrent transmission of multiple symbols, each going through an individual channel, the elements in the received signal are the superimposition of multiple symbols in general. The undesired superimposed symbols contribute to interference, as shown in the second term of the second equation in (1.12). This type of interference is known as inter-(sub)channel interference or ICI. As long as the sub-channels are not orthogonal to each others, there always exists ICI. For some virtual MIMO systems such as OFDM system, the ICI can arise due to the time-variation of the channel during the transmission of an OFDM symbol, or due to the frequency offset between the transmitter and the receiver. Either case can cause loss

of orthogonality between each sub-carrier of the OFDM modulation. Hence this type of interference is also known as inter-carrier interference in OFDM systems [24].

For multiple access systems such as CDMA, the so-called multiple access interference or multiuser interference is in fact a combination of the inter-symbol interference and the inter-channel interference discussed previously. With the unified multirate transmultiplexer model in Figure 1.1, we will not use the terms MAI and MUI in this dissertation to simplify notation. They will be treated as ISI/ICI instead.

The purpose of interference suppression is to cancel the ISI/ICI from the received signal so that the transmitted symbols can be recovered. Any operation for this sake can be called interference suppression. However ISI suppression is also conventionally known as equalization. MAI/MUI suppression is also known as multiuser detection in CDMA systems [79].

1.1.3 Parameter Estimation of Time-varying Multipath Fading Channels

We have seen the importance of interference suppression in wireless communication systems. However, its success depends on the availability of the channel state information (CSI). In other words, we need to either explicitly or implicitly make use of the information of the equivalent channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ to effectively suppress the interferences. Otherwise there always exist ambiguities in symbol detection. As such, channel estimation is necessary. For this reason, we will also devote to this problem in this dissertation.

For time-invariant systems, channel estimation is a relatively well studied problem. There have been many results in this area. Either training based or blind methods can be used to estimate the channel. For a more comprehensive discussion on time-invariant channel estimation, see [22] and references therein.

However, the wireless communication channels are time-varying in nature due to the constant variation of the transmission medium and the relative moving between the transmitters and the mobile receivers. To achieve more mobility, high moving velocity is expected and this results in fast time-varying channels. Even when the physical channel is time-invariant, due to the frequency mismatch between the transmitter and the receiver or some other issues, the resultant equivalent channel in (1.7) can be time-varying too. Time-varying channel estimation is also known as channel tracking in communication system design. It is more challenging than time-invariant channel estimation.

In this dissertation, we will focus on the channel tracking problem only. To render the problem solvable, we need have an appropriate time evolution model to characterize the channel variation. One such model is the multichannel autoregressive (AR) model, which can be used to approximate the MIMO channel variation with satisfactory accuracy [36]. Let $\mathbf{h}_t$ be the vector formed by all the elements of $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$. Then a first order multichannel AR process of the channel variation is given by

$$\mathbf{h}_{t+1} = \mathbf{A}\mathbf{h}_t + \mathbf{B}\mathbf{w}_t \tag{1.13}$$

where $\mathbf{w}_t$ is a zero-mean unit-variance complex white Gaussian vector process, matrix

$\mathbf{A}$ is determined by the rate of variation, and matrix $\mathbf{B}$ is determined by the magnitude of variation.

Most existing channel tracking methods based on the AR model in (1.13) have assumed perfect knowledge of the matrices $\mathbf{A}$ and $\mathbf{B}$. This assumption is unrealistic in any practical systems. Instead, we have to estimate the model parameters (matrices $\mathbf{A}$ and $\mathbf{B}$) before we can use the AR model to track the time-varying channel. To achieve this, we will propose in this dissertation a novel blind method that does not require any training sequence.

1.2 Literature Review

The problem of ISI suppression in SISO communication systems has a long history. In [18], Forney derived the maximum likelihood sequence estimator (MLSE), which achieved the optimal ISI suppression in the sense of maximum likelihood symbol detection. Although it has the best performance among all the ISI suppression schemes, this method has high computational complexity, which grows exponentially with respect to the channel order. This prevents its wide applicability in many practical systems.

On the other hand, a linear equalizer employs a linear transversal filter to suppress the ISI. Although, it has very simple structure and low complexity [46], [47], it also has much poorer performance as compared to the MLSE, especially when the underline channel has deep spectral nulls [59].

As a tradeoff between complexity and performance, decision-feedback equalization was proposed and analyzed in [1], [3], [4], [6], [9], [16], [21], [58], [62], etc. Depending on the design criterion, there are two types of DFEs, namely the minimum mean squared error (MMSE) DFE and the zero forcing DFE. It can be shown that the zero forcing DFE is just a special case of the MMSE DFE as the additive noise goes to zero [3]. Because it allows easy adaptive implementation, the MMSE DFE is also more widely used than the LE in practical communication systems [21]. In the past, many people studied the DFE in the frequency domain, assuming the channel being time-invariant. The resultant DFEs are normally of infinite legnth [5] ,[10], [11], [13], [38], [61]. For the sake of realizable implementation and stability considerations, practical DFEs are normally restricted to be of finite length [1], [3]. However this length constraint may also limit the performance of the DFEs, and the resultant DFEs may not be optimal in the sense of minimum mean squared error[2] [34]. It seems that the realizability of a DFE has to compromise with its performance, or vice versa.

The results obtained in optimal estimation theory, mainly due to the work of R. E. Kalman [35], render the possibility of optimal realizable design. Several people have studied the optimal realizable equalizer design problem before [30]. To the author's knowledge, the Kalman filter based LE was first proposed by Lawrence and Kaufman in [37]. Mulgrew and Cowan generalized the results to DFE in [54]. However, they failed to go further to reveal the Kalman equalizers' properties and

[2]From now on, the words "optimality" and "optimal" refer to achieving linear minimum mean squared error, unless stated otherwise.

relations with the commonly used finite length LE and DFE. In [67], Sternad and Ahlén derived the optimal realizable DFE using a polynomial decomposition method, which was mathematically tedious and did not provide an intuitive interpretation. The most recent work on realizable LE and DFE is probably due to López-Valcarce [45], who employed Wiener and Kalman filter theories to study the properties of the optimal realizable equalizers. The results in [45], [67], however, cannot be applied to time-varying channels, which are commonly encountered in mobile communications. Furthermore, most of the existing results on optimal realizable design focus on SISO systems only. Hence they are not applicable to many new communication systems that are MIMO in nature, such as CDMA and OFDM systems.

For multiple access systems, the optimal interference suppression based on the maximum likelihood criterion was first studied by Verdu in his pioneering paper [78]. This work opened the new research arena of multiuser detection that has been very active for the past two decades [79]. Due to the rapid growth of the market of the CDMA based cellular systems, there have been many research activities in the area of multiuser detection in CDMA systems. See for example [17], [27], [28], [48], [49], [51], [53], [75], [85]. For a more comprehensive list of references, please see [79] and references therein. From the author's point of view, many of the results obtained for multiuser detection in CDMA systems are in fact parallel to those obtained for ISI suppression in SISO systems, thanks to the unified mathematical description dis-

cussed in the previous section. However, the multiple access in CDMA does introduce something unique to CDMA-like MIMO systems, i.e., the ICI problem.

An effective approach to the suppression of the ICI is the so-called successive interference cancellation (SIC) method, which utilizes the detected symbols of strong users to help the detection of weak users, an idea similar to, but different from that of the DFE. For some reason unclear to the author, the existing SIC schemes were proposed almost exclusively for CDMA systems. See [31], [33], [52], [57], [63] and references therein. All these schemes require at least the channel information and the signature sequence of the desired user to explicitly calculate the MUI caused by each user or its contribution to the noise subspace. As such, the existing SIC schemes cannot be directly applied to the ICI suppression in a general MIMO system, where it is very difficult, if not impossible, to calculate the ICI in a MIMO receiver after the ISI suppression.

MIMO data transmission has received great attention in recent years because of its many advantages, such as increased channel capacity, spatial diversity, and etc [19], [72], [76]. Many researchers have studied the interference suppression for MIMO channels. In [25], [39], [41], [70], blind methods (without training signals) were investigated. These methods are normally computationally complicated, and suffer from slow convergence. In practical systems though, a more commonly adopted approach is to use the training signals to improve the performance at the expense of the reduced channel bandwidth efficiency [26]. Such an approach employs the training

signals to either explicitly or implicitly estimate the MIMO channel [36], and then carries out the interference suppression.

Assuming perfect knowledge of the MIMO channel, Tidestav *et al* extended the results in [67] and derived an optimal realizable MIMO DFE to suppress the ISI using a polynomial decomposition method that is applicable to infinite impulse response (IIR) channels [73]. While the optimality of the method is shown mathematically, the complicated formulas are less insightful, and lack an intuitive interpretation. In addition it failed to address the ICI suppression. DFE based interference suppression methods were also studied in [3], [16], [17] and [66]. These methods did not distinguish between the ISI and the ICI, but tried to suppress both together. Furthermore, all the existing methods deal with time-invariant channels only.

As we have pointed out before, successful interference suppression depends on the accurate estimation of channel information. In wireless communications, due to multipath transmission and the moving of subscribers, the underlying channel may experience both frequency- and time-selective fading [60], [69]. For time-invariant channels, blind or training based channel estimation/equalization techniques have been studied extensively. See [12], [15] and references therein. However, when the channel is time-varying, the estimation/equalization problem becomes more challenging. In [7], [71], [83], [84], it is verified that most time-varying channels encountered in wireless communications can be sufficiently modeled by the low order auto-regressive (AR) Gaussian Markov processes discussed in Section 1.1.3. Based on this low or-

der AR Markov model, many channel estimation/tracking and equalization methods based on Kalman filtering have been proposed (e.g. [36], [44], [74]). However, most of the existing works assume that the second-order statistics (SOS) of the channel, and hence the parameters of the state equation in the Kalman filter, are exactly known or perfectly estimated. In [74], it shows that this condition holds true only when sufficiently long training sequence is used and frequent retraining is necessary. In many applications such as radio broadcasting, frequent and long training is not feasible. This may result in unreliable estimation of the channel SOS. And the estimation error in the channel SOS may cause severe performance degradation in channel estimation.

1.3 Dissertation Contributions

This dissertation provides solutions to two of the most important problems in wireless communication systems design. The first part (Chapters 2 and 3) of this dissertation focuses on the optimal realizable interference suppression in MIMO communication systems with channel information available at the receiver side. The second part (Chapter 4) devotes to the parameter estimation in time-varying channels and channel tracking, the results of which can be integrated with the interference suppression schemes proposed in the first part. The contributions of this dissertation are as follows.

- Development of optimal realizable interference suppression schemes for MIMO communication systems under a unified framework.

  This dissertation employs the optimum linear estimation theory to suppress both the ISI and the ICI in a MIMO communication system. In contrast to the existing work that seeks to balance between realizability and performance, our proposed methods achieve both with a new structure. A two-stage receiver that suppresses the ISI and the ICI in series is proposed. In this dissertation, we first adopt a state-space approach, and derive state-space solutions for the optimum realizable LE and DFE based on Kalman filtering. The state-space approach provides insights into the properties of the equalizers that are not clearly seen otherwise. We show that increasing the detection delay of an equalizer results in smaller estimation error at the expense of higher complexity. This result reveals a direct relation between performance, complexity, and detection delay. In general, the impulse response of a Kalman filter has an infinite length. For LEs, we show that the Kalman filter based optimal LE (Kalman LE) is an IIR filter. However, to our surprise, we prove in this dissertation that the Kalman filter based optiaml DFE (Kalman DFE) is equivalent to a finite length minimum mean squared error (MMSE) DFE whose feed-forward filter (FFF) order and feedback filter (FBF) order are bounded below by the detection delay and the channel length, respectively. Hence the optimal realizable DFE is an FIR filter. The only difference between a Kalman DFE and a conventional finite length MMSE

DFE of proper orders is the way how they are implemented. Our result thus clarifies the optimality of the conventional finite length MMSE DFE. Following the ISI suppression, we propose a novel successive interference cancellation (SIC) scheme to suppress the ICI based on spatial filtering. Moreover a reliability metric for the detected symbols is introduced to aid our proposed SIC scheme to successively remove the ICI from the ISI suppressed signals. It employs a state-space realization similar to that of the Kalman equalizers, and uses again Kalman filtering to suppress the ICI. Different from the Kalman equalizers, the proposed SIC scheme does not require any channel information. It only needs the output of the ISI suppression from the first stage to initialize, and thus can be combined with different equalizers in the first stage such as the Kalman DFE, the adaptive MIMO DFE, or the much simpler LE etc. In light of the optimality of the Kalman filter, optimal ISI and ICI suppression is achieved when the SIC scheme is combined with the Kalman DFE. This implementation flexibility is very appealing to MIMO systems over wireless fading channels. Since in this case the channel is constantly changing, it is desirable to adaptively adjust the receiver parameters according to the channel condition. The state-space approach adopted in this dissertation renders our method applicable to any linear channels, including time-invariant, time-varying, FIR, and IIR channels, provided that the channel model admits a state-space realization. Hence the

results reported in this paper are more general than the existing work. To the author's knowledge, our results are the most general obtained so far.

- Development of a novel semi-blind channel parameter estimation method and its associated channel tracking scheme for SISO time-varying multipath fading channels.

  Since the success of interference suppression depends on the availability of the channel state information, we also study the channel estimation problem in this dissertation. However, we focus on the more challenging problem of parameter estimation and channel tracking in time-varying multipath fading channels. To render the problem solvable, we employ the well-accepted low order autoregressive (AR) model to characterize the variation of channel impulse response. We show that the parameters of the AR model is uniquely determined by the channel second-order statistics (SOS), i.e., the covariance information of the channel impulse response. Hence the parameter problem is translated to that how to estimate the channel SOS. We establish the identifiability condition of the channel SOS. Then we convert the identifiability condition into the design constraints on the transceiver structure. Linear precoders at the transmitter side that satisfy the constraints and achieve diversities are constructed. Kalman filtering based channel tracking and equalization methods are developed. Since the frequency offset in a communication system can also be described by an first order AR

model, the proposed blind channel parameter estimation method can also be applied to blind estimation of the frequency offset in a communication system.

1.4 Organization of the Dissertation

The organization of the dissertation is outlined briefly as follows. Chapter 2 is devoted to the optimal realizable ISI suppression in MIMO systems, where Kalman filter based equalizers will be proposed. Chapter 3 focuses on the optimal realizable ICI suppression and its combination with the ISI suppression schemes. Chapter 4 studies the parameter estimation and channel tracking problem in time-varying multipath fading channels. Finally Chapter 5 concludes the dissertation by suggesting some possible future research. Throughout this dissertation, symbols for matrices are in boldface capital letters, and vectors are in boldface small letters. Other notational conventions are listed in the earlier page of this dissertation titled as "Notations and Symbols".

Chapter 2

Optimal Realizable Suppression of Inter-Symbol Interference

Minimum mean square error (MMSE) linear equalizers (LEs) and decision feedback equalizers (DFEs) are widely used in communication systems to combat the inter-symbol interference (ISI) introduced by multipath channels. An LE conventionally consists of a linear transversal filter that acts as a linear estimator of the transmitted symbols, and a quantizer. Although it has a simple structure, an LE suffers large performance loss when there exist deep channel spectral nulls. On the other hand, a DFE is formed by adding a linear feedback filter to the LE structure. Utilizing the past detected symbols with the help of the feedback filter, a DFE normally has much better performance than an LE. By the commonly adopted assumption that all the past detections are correct, the design of a DFE reduces to a linear estimator design problem too [1]. Consequently, we can study both the LE and the DFE within the framework of linear estimator design, and treat the LE as a special case of the DFE.

In practical systems, the linear filter(s) in a conventional LE or DFE are normally restricted to be of finite length for the sake of realizable implementation. The resulting equalizers will be called finite length LE and DFE hereafter. This finite length constraint may limit the equalizer performance. Motivated by the fact that Kalman

filter is the optimal linear estimator[1] and is realizable, we apply Kalman filter theory to the study of the optimal realizable MMSE LEs and DFEs in this chapter.

We will derive a state-space solution to the optimum LE and the optimum DFE based on Kalman filtering. This state-space approach yields optimal realizable LE and DFE. More importantly, it provides insights into the properties of the LE and the DFE that are not clearly seen otherwise. We will show that, for both the LE and the DFE, increasing the detection delay results in smaller estimation error at the expense of higher complexity. Although in general a Kalman filter is an IIR filter due to its recursive structure, we will show in this chapter that a Kalman filter based DFE (Kalman DFE) is equivalent to a conventional finite length MMSE DFE with feed-forward filter length and feedback filter length bounded below by the detection delay and the equivalent channel length, respectively. Hence a Kalman DFE has an FIR structure. Despite the fact that an LE (Kalman LE) is only a special case of a DFE, a Kalman filter based LE (Kalman LE) does not share the same property as the Kalman DFE. It will be shown that a Kalman LE is equivalent to a conventional MMSE LE of infinite filter length, or an IIR filter. Simulation examples will be provided to illustrate the theoretical results.

[1]The words "optimality" or "optimal" are in the sense of achieving linear minimum mean squared error [34].

2.1 System Model

Consider an M-input/P-output channel of length $L+1$. The input-output relation of the channel can be written as

$$\mathbf{y}_t = \sum_{i=0}^{L} \mathbf{H}_{t,i}\mathbf{s}_{t-i} + \mathbf{v}_t, \tag{2.1}$$

where $\mathbf{s}_t = [\, s_t(1) \;\; s_t(2) \;\; \cdots \;\; s_t(M) \,]^T$ is an $M \times 1$ vector of the transmitted signal at time t, whose elements are drawn from a finite size alphabet, $\mathbf{y}_t$ and $\mathbf{v}_t$ are $P \times 1$ vectors of the received signal and the additive noise at time t, respectively, and $\mathbf{H}_{t,i}$ is a $P \times M$ matrix of the ith channel impulse response coefficient matrix at time t with $i = 0, 1, \cdots, L$. Without loss of generality, we assume that both $\mathbf{s}_t$ and $\mathbf{v}_t$ are white and uncorrelated to each other. Hence we have

$$\mathrm{E}\left[\mathbf{s}_t\mathbf{s}_{t-k}^*\right] = \delta_k\mathbf{R}_s, \quad \mathrm{E}\left[\mathbf{v}_t\mathbf{v}_{t-k}^*\right] = \delta_k\mathbf{R}_v, \tag{2.2}$$

$$\mathrm{E}\left[\mathbf{s}_t\mathbf{v}_{t-k}^*\right] = \mathbf{0}_{M\times P}, \tag{2.3}$$

where $\mathbf{R}_s$ and $\mathbf{R}_v$ are the covariance matrices of the signal and the noise, respectively, and δ_k is the Kronecker delta. Note that we do not require $\mathbf{R}_s$ and $\mathbf{R}_v$ to be identity matrices, and no Gaussianity is assumed for $\mathbf{v}_t$.

To simplify our discussion, we assume throughout this chapter that the receiver has perfect knowledge of the channel, and focus only on the design of the receiver. As shown in Figure 2.1, the proposed receiver consists of an ISI suppression first stage followed by an ICI suppression second stage. We will focus on the optimal realizable

ISI suppression problem in this chapter. The ICI suppression problem will be studied in the next chapter.

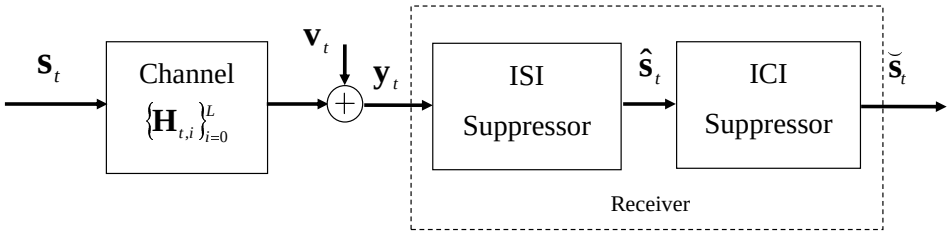

Figure 2.1: System Model

As we have pointed out in the previous chapeter, the time-varying MIMO model in (2.1) is the most general possible. Many communication systems such as the multiple antennas systems, OFDM systems, and many multiple access systems such as CDMA, FDMA, TDMA etc., can all be regarded as some special cases of the general MIMO system considered in (2.1). Thus the results presented in this chapter have wide applicability.

2.2 Preliminary Analysis: Finite Length MMSE LE and DFE

This section briefly reviews the conventional MMSE LE and DFE. Our derivation generalizes the results obtain in [1] from SISO channels to MIMO channels. The expressions we obtained are equivalent to but different from those in [3].

Figure 2.2 gives a block diagram of the conventional DFE, where the FFF and the FBF denote the feed-forward filter and the feedback filter, respectively, z^{-1} denotes the unit delay operation, and Quan$(\cdot)$ denotes the quantization operation. Both the FFF and the FBF are linear (possibly time-varying) filters whose coefficients at time t are given by the $M \times P$ matrices $\{\mathbf{M}_{t,i}\}_{i=0}^{L_g}$ and the $M \times M$ matrices $\{\mathbf{N}_{t,j}\}_{j=1}^{L_b}$, respectively. The integers L_g and L_b are fixed denoting the FFF length and the FBF length, respectively. In practical systems, for the sake of realizable implementation, both L_g and L_b are normally restricted to be of finite length. Note that the FBF of the DFE is required to be strictly causal, i.e. it can only process past detections obtained in previous time. Hence when $L_b = 0$, the FBF does not exist and no detected symbol will be fed back to the equalizer. In this case, the DFE reduces to an LE. Therefore we can treat an LE as a special case of a DFE and study both within the same framework. However, it will be shown later that a DFE and an LE have quite different properties.

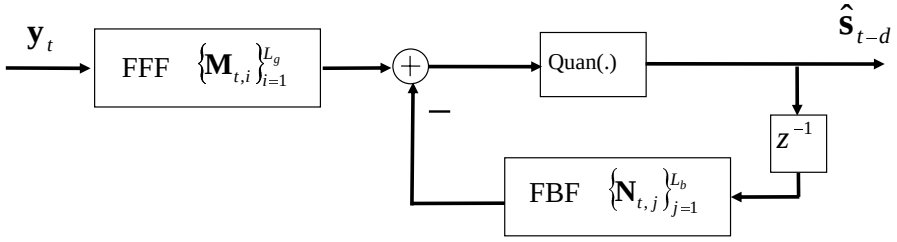

Figure 2.2: Block diagram of the conventional finite length DFE. (When $L_b = 0$, the feedback path does not exist and the above reduces to an LE.)

For the brevity of notations, in this section we will focus our discussions on DFEs. The results on LE can be obtained in a similar way by setting the FBF length $L_b = 0$. We consider an $(L+1)$-path channel in this chapter. It is well known that a multipath channel provides frequency diversity that can be exploited by the receiver to improve detection performance [59]. To achieve the frequency diversity, a detection delay $d \in [0, L + L_g]$ is desired [1, 81]. Thus at time t, the symbol to be detected is $\mathbf{s}_{t-d}$ rather than $\mathbf{s}_t$. Denote the symbol estimation error at time t (before the quantizer) to be

$$\boldsymbol{\Upsilon}_t = \mathbf{s}_{t-d} - \sum_{i=0}^{L_g} \mathbf{M}_{t,i}^* \mathbf{y}_{t-i} + \sum_{j=1}^{L_b} \mathbf{N}_{t,j}^* \hat{\mathbf{s}}_{t-d-j}, \tag{2.4}$$

where $\hat{\mathbf{s}}_t$ is the detected symbol of $\mathbf{s}_t$. A finite length MMSE DFE aims to minimize the mean squared error $\mathrm{Tr}\left[\mathbf{E}_t\right] = \mathrm{Tr}\left[\mathrm{E}\left(\boldsymbol{\Upsilon}_t \boldsymbol{\Upsilon}_t^*\right)\right]$ with properly designed FFF and FBF, where

$$\mathbf{E}_t = \mathrm{E}\left(\boldsymbol{\Upsilon}_t \boldsymbol{\Upsilon}_t^*\right), \tag{2.5}$$

is the estimation error covariance matrix. To simplify discussion, we adopt the commonly accepted assumption that the symbol detection is error free, i.e., $\hat{\mathbf{s}}_t = \mathbf{s}_t$ for all t. Then we can obtain

$$\begin{array}{rcl} \boldsymbol{\Upsilon}_t & = & \mathbf{s}_{t-d} - \sum_{i=0}^{L_g} \mathbf{M}_{t,i}^* \mathbf{y}_{t-i} + \sum_{j=1}^{L_b} \mathbf{N}_{t,j}^* \mathbf{s}_{t-d-j} \\ & = & \left[\mathbf{0}_{M \times Md}\ \mathbf{N}_t^*\ \mathbf{0}_{M \times M\gamma}\right] \underline{\mathbf{s}}_t - \mathbf{M}_t^* \underline{\mathbf{y}}_t \end{array}, \tag{2.6}$$

where $\underline{\mathbf{s}}_t = \left[\mathbf{s}_t^T\ \mathbf{s}_{t-1}^T\ \cdots\ \mathbf{s}_{t-L-L_g}^T\right]^T$, $\underline{\mathbf{y}}_t = \left[\mathbf{y}_t^T\ \mathbf{y}_{t-1}^T\ \cdots\ \mathbf{y}_{t-L_g}^T\right]^T$, $\gamma = L+L_g-L_b-d \geq 0$, and $\mathbf{M}_t^* = \left[\mathbf{M}_{t,0}^*\ \mathbf{M}_{t,1}^*\ \cdots\ \mathbf{M}_{t,L_g}^*\right]$ and $\mathbf{N}_t^* = \left[\mathbf{I}_M\ \mathbf{N}_{t,1}^*\ \mathbf{N}_{t,2}^*\ \cdots\ \mathbf{N}_{t,L_b}^*\right]$ are the matrices which consist of the coefficients of the FFF and the FBF of the finite length DFE at

time t, respectively. Denote

$$\underline{\mathbf{H}}_t = \begin{bmatrix} \mathbf{H}_{t,0} & \mathbf{H}_{t,1} & \cdots & \mathbf{H}_{t,L} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{t-1,0} & \mathbf{H}_{t-1,1} & \cdots & \mathbf{H}_{t-1,L} & \mathbf{0} & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_{t-L_g,0} & \mathbf{H}_{t-L_g,1} & \cdots & \mathbf{H}_{t-L_g,L} \end{bmatrix} \tag{2.7}$$

to be the block channel matrix. From (2.1), we have the following equivalent input-output relation

$$\underline{\mathbf{y}}_t = \underline{\mathbf{H}}_t \underline{\mathbf{s}}_t + \underline{\mathbf{v}}_t, \tag{2.8}$$

where $\underline{\mathbf{v}}_t = \begin{bmatrix} \mathbf{v}_t^T & \mathbf{v}_{t-1}^T & \cdots & \mathbf{v}_{t-L_g}^T \end{bmatrix}^T$ is the zero mean additive noise vector with covariance given by $\mathbf{R}_{\underline{vv}} = \mathrm{E}\left[\underline{\mathbf{v}}_t \underline{\mathbf{v}}_t^*\right]$. Define $\mathbf{R}_{\underline{ss}} = \mathrm{E}\left[\underline{\mathbf{s}}_t \underline{\mathbf{s}}_t^*\right]$, $\mathbf{R}_{\underline{sy}} = \mathrm{E}[\underline{\mathbf{s}}_t \underline{\mathbf{y}}_t^*] = \mathbf{R}_{\underline{ss}} \underline{\mathbf{H}}_t^*$, and $\mathbf{R}_{\underline{yy}} = \mathrm{E}[\underline{\mathbf{y}}_t \underline{\mathbf{y}}_t^*] = \underline{\mathbf{H}}_t \mathbf{R}_{\underline{ss}} \underline{\mathbf{H}}_t^* + \mathbf{R}_{\underline{vv}}$. Then it is easy to verify that [3]

$$\mathbf{E}_t = \mathbf{N}_t^* \tilde{\mathbf{R}} \mathbf{N}_t, \tag{2.9}$$

where

$$\begin{aligned} \tilde{\mathbf{R}} \;=\; & \begin{bmatrix} \mathbf{0}_{M(L_b+1)\times Md} & \mathbf{I}_{M(L_b+1)} & \mathbf{0}_{M(L_b+1)\times M\gamma} \end{bmatrix} \\ & \times \left[\mathbf{R}_{\underline{ss}}^{-1} + \underline{\mathbf{H}}_t^* \mathbf{R}_{\underline{vv}}^{-1} \underline{\mathbf{H}}_t \right]^{-1} \begin{bmatrix} \mathbf{0}_{Md\times M(L_b+1)} \\ \mathbf{I}_{M(L_b+1)} \\ \mathbf{0}_{M\gamma\times M(L_b+1)} \end{bmatrix}. \end{aligned} \tag{2.10}$$

Let $\tilde{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{A}}_{M\times M} & \tilde{\mathbf{B}}_{M\times ML_b} \\ \tilde{\mathbf{B}}_{ML_b\times M}^* & \tilde{\mathbf{C}}_{ML_b\times ML_b} \end{bmatrix}$. After some matrix manipulation, it can be shown that the optimal FBF and FFF that minimize $\mathrm{Tr}\left[\mathbf{E}_t\right]$ are given by

$$\mathbf{N}_t^{\mathrm{opt}} = \begin{bmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^* & \tilde{\mathbf{C}} \end{bmatrix}^{-1} \begin{bmatrix} \Delta_{\tilde{C}} \\ \mathbf{0}_{ML_b\times ML_b} \end{bmatrix}, \tag{2.11}$$

$$\mathbf{M}_t^{\mathrm{opt}} = \mathbf{R}_{\underline{yy}}^{-1} \mathbf{R}_{\underline{sy}}^* \left[\mathbf{0}_{M\times Md} \; \left(\mathbf{N}_t^{\mathrm{opt}}\right)^* \; \mathbf{0}_{M\times M\gamma} \right]^*, \tag{2.12}$$

where $\Delta_{\tilde{C}} = \tilde{\mathbf{A}} - \tilde{\mathbf{B}}\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{B}}^*$ is the Schur complement of $\tilde{\mathbf{C}}$ in $\tilde{\mathbf{R}}$, and

$$\Delta_{\tilde{C}}^{-1} = \tilde{\mathbf{R}}^{-1}(1:M, 1:M) \tag{2.13}$$

is an $M \times M$ sub-matrix[2] of $\tilde{\mathbf{R}}^{-1}$. The corresponding minimum possible error covariance matrix can be shown to be

$$\mathbf{E}_{\min,t} = \left(\mathbf{N}_t^{\text{opt}}\right)^* \tilde{\mathbf{R}} \mathbf{N}_t^{\text{opt}} = \Delta_{\tilde{C}}. \tag{2.14}$$

We can obtain the optimal finite length LE and its associated error covariance matrix in a similar way by setting $L_b = 0$, or equivalently by setting $\mathbf{N}_t^* = [\mathbf{I}_M \ \mathbf{0}_{M \times ML_b}]$. By the correct past detections assumption, it also can be shown that

$$\text{Tr}\left[\mathbf{E}_{\min,t}^{FLLE}\right] \leq \text{Tr}\left[\mathbf{E}_{\min,t}^{FLDFE}\right] \tag{2.15}$$

holds true for any given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ [3], where the superscripts in (2.15) denote finite length LE and finite length DFE, respectively.

2.3 Optimal Realizable MMSE LE and DFE via Kalman Filter

It is clear from Equations (2.11) to (2.14) that, given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$, the MMSE achieved by a finite length equalizer is a function of the detection delay d and its filter order(s). For fixed L_g and L_b, the optimization of the finite length DFE over the detection delay d is studied in [1]. It turns out that the performance of the finite length DFE and d, L_g, and L_b are closely related to each other. Intuitively,

[2]For any matrix $\mathbf{A}$, we denote $\mathbf{A}\,(k:l, m:n)$ to be the sub-matrix of $\mathbf{A}$ whose entries are formed by $\mathbf{A}(i,j)$, $i = k, k+1, \cdots, l$, $j = m, m+1, \cdots, n$, where $\mathbf{A}(i,j)$ is the (i,j)th entry of $\mathbf{A}$.

when d, L_g, and L_b are large, more information is used to obtain the estimate of the transmitted symbol $\mathbf{s}_{t-d}$, and smaller estimation error is expected. However, large L_g and L_b also increases the complexity of the DFE. Thus there is a tradeoff between the choices of d, L_g, L_b and the achievable MMSE. In this section, we will further reveal their relations using the Kalman filter theory. Some interesting results will be presented.

2.3.1 Structure

The key to the development of the Kalman filter based equalizer (Kalman equalizer) is the use of state-space representation for the channel model. Let $L_x \geq L$ be an integer. Choose the state vector to be $\mathbf{x}_t = \left[\mathbf{s}_t^T \ \mathbf{s}_{t-1}^T \ \cdots \ \mathbf{s}_{t-L_x}^T\right]^T$. It can be verified that the input-output relation in (2.1) is equivalent to the following state-space model

$$\begin{cases} \mathbf{x}_{t+1} &= \mathbf{F}\mathbf{x}_t + \mathbf{G}\mathbf{s}_{t+1} \\ \mathbf{y}_t &= \mathbf{\Pi}_t\mathbf{x}_t + \mathbf{v}_t \end{cases} \tag{2.16}$$

where $\{\mathbf{F}, \mathbf{G}, \mathbf{\Pi}_t\}$ is the realization given by

$$\begin{aligned} \mathbf{F} &= \begin{bmatrix} \mathbf{0}_{M\times ML_x} & \mathbf{0}_{M\times M} \\ \mathbf{I}_{ML_x} & \mathbf{0}_{ML_x\times M} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{ML_x\times M} \end{bmatrix}, \\ \mathbf{\Pi}_t &= \begin{bmatrix} \mathbf{H}_{t,0} & \cdots & \mathbf{H}_{t,L} & \mathbf{0}_{N\times M(L_x-L)} \end{bmatrix}. \end{aligned}$$

Note that the transmitted symbols are contained in the state vector.

When the detection delay satisfies $0 \leq d \leq L_x - 1$ and all the past detections are correct, we can represent the most recently detected symbol vector by

$$\hat{\mathbf{s}}_{t-d-1} = \mathbf{s}_{t-d-1} = \mathbf{\Gamma}_{d+1}\mathbf{x}_t, \tag{2.17}$$

where

$$\mathbf{\Gamma}_d = \begin{bmatrix} \mathbf{0}_{M\times Md} & \mathbf{I}_M & \mathbf{0}_{M\times M(L_x-d)} \end{bmatrix}. \tag{2.18}$$

Let $\mathbf{u}_t = \begin{bmatrix} \mathbf{y}_t^T & \hat{\mathbf{s}}_{t-d-1}^T \end{bmatrix}^T$. Then $\mathbf{u}_t$ is the observed signal vector to the DFE at time t by (2.16) and (2.17). Denote $\mathbf{w}_t = \begin{bmatrix} \mathbf{v}_t^T & \mathbf{0}_{M\times M}^T \end{bmatrix}^T$, and $\mathbf{D}_t = \begin{bmatrix} \mathbf{\Pi}_t^T & \mathbf{\Gamma}_{d+1}^T \end{bmatrix}^T$. From (2.16) and (2.17) again, we arrive at a state-space model

$$\begin{cases} \mathbf{x}_{t+1} &= \mathbf{F}\mathbf{x}_t + \mathbf{G}\mathbf{s}_{t+1} \\ \mathbf{u}_t &= \mathbf{D}_t\mathbf{x}_t + \mathbf{w}_t \end{cases} \tag{2.19}$$

whose observation vector consists of the received signals and the past detected symbols.

The design of the optimum DFE aims to synthesize a linear estimator that achieves the MMSE for the estimation of the symbol vector prior to the detection via quantization. In light of the optimum linear estimation theory, its state-space solution is the Kalman filter for the signal model in (2.19), due to the fact that all symbol vectors over the time interval from $t - L_x$ to t are contained in state vector $\mathbf{x}_t$. That is, the Kalman filter generates the optimal linear estimate of the state vector based on all the observations up to current time t. Denote $\tilde{\mathbf{x}}_t$ to be the optimal linear estimate of $\mathbf{x}_t$ based on $\{\mathbf{u}_i\}_{i=0}^t$. Applying the Kalman filter to state-space model (2.19), we can obtain $\tilde{\mathbf{x}}_t$ through the following Kalman recursions with respect to t [34]:

$$\begin{aligned} \mathbf{R}_{e,t} &= \mathbf{R} + \mathbf{D}_t\left(\mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^* + \mathbf{G}\mathbf{R}_s\mathbf{G}^*\right)\mathbf{D}_t^* \\ \mathbf{K}_t &= \left(\mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^* + \mathbf{G}\mathbf{R}_s\mathbf{G}^*\right)\mathbf{D}_t^*\mathbf{R}_{e,t}^{-1} \\ \mathbf{P}_t &= \mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^* + \mathbf{G}\mathbf{R}_s\mathbf{G}^* - \mathbf{K}_t\mathbf{R}_{e,t}\mathbf{K}_t^* \\ \tilde{\mathbf{x}}_t &= \mathbf{F}\tilde{\mathbf{x}}_{t-1} + \mathbf{K}_t\left(\mathbf{u}_t - \mathbf{D}_t\mathbf{F}\tilde{\mathbf{x}}_{t-1}\right) \end{aligned} \tag{2.20}$$

where $\mathbf{P}_t = \mathrm{E}\left[\left(\mathbf{x}_t - \tilde{\mathbf{x}}_t\right)\left(\mathbf{x}_t - \tilde{\mathbf{x}}_t\right)^*\right]$ is the error covariance matrix at time t,

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_v & \mathbf{0}_{P\times M} \\ \mathbf{0}_{M\times P} & \mathbf{0}_{M\times M} \end{bmatrix}, \tag{2.21}$$

and the initial condition is given by $\tilde{\mathbf{x}}_0 = \mathrm{E}\left[\mathbf{x}_0\right]$ with covariance $\mathbf{P}_0 = \mathrm{Var}\left[\mathbf{x}_0\right]$.

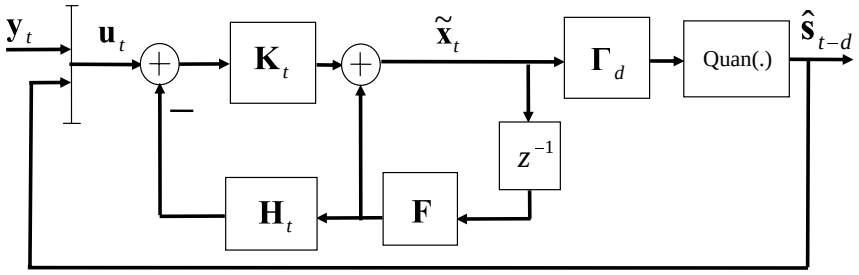

Figure 2.3: Block diagram of the Kalman DFE

After $\tilde{\mathbf{x}}_t$ is obtained, $\mathbf{s}_{t-d}$, the symbol to be detected can be retrieved by selecting the $((d-1)M+1)$th to the (dM)th elements of the estimated state vector and quantizing them to the nearest constellation points, i.e.,

$$\hat{\mathbf{s}}_{t-d} = \mathbf{\Gamma}_d \times \mathrm{Quan}\left(\tilde{\mathbf{x}}_t\right). \tag{2.22}$$

This equalization process produces the optimal linear estimates of the transmitted symbols based not only on the received signals, but also on all the past detected symbols. The resulting detected symbol $\hat{\mathbf{s}}_{t-d}$ is the quantized output of a linear filter whose input consists of both the received signals $\{\mathbf{y}_i\}_{i=0}^{t}$ and the past detected symbols $\{\mathbf{s}_j\}_{j=0}^{t-d-1}$. Thus this equalizer is a DFE and is referred to as Kalman DFE. A block diagram of the Kalman DFE is provided in Figure 2.3, where z^{-1} denotes the unit delay operation.

As in the case of the finite length equalizers, a Kalman filter based linear equalizer can be regarded as a special case of the Kalman DFE by setting its feedback signals to zero. Setting $\mathbf{u}_t = \begin{bmatrix} \mathbf{y}_t^T & \mathbf{0}_{M\times 1}^T \end{bmatrix}^T$ and performing the Kalman recursions in (2.20), we obtain the optimal linear equalizer. This LE will be referred to as Kalman LE hereafter. It is obvious that the Kalman LE is simpler in computation than the Kalman DFE. However it also has much poorer performance, especially when the underline channel has nulls in the channel frequency response.

Remark 2.1 *For time-invariant channels, the state-space models in (2.16) and (2.19) are stationary. In this case, the Kalman gain matrix* $\mathbf{K}_t = \mathbf{K}$ *can be chosen constant and can be calculated in advance by solving a discrete-time algebraic Riccati equation (DARE) [34]. After the constant Kalman gain* $\mathbf{K}$ *is obtained, only the last equation in the recursive algorithm (2.20) needs to be implemented to obtain the optimal state estimate. Hence the computational complexity for time-invariant channel can be largely reduced.*

2.3.2 Performance

Given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$, the performance of a Kalman equalizer is determined by only one parameter, the detection delay d. Let $\epsilon_{KF}^2(t;d) = \mathrm{Tr}\left(\mathbf{\Gamma}_d \mathbf{P}_t \mathbf{\Gamma}_d^T\right)$ denote the MMSE achieved by a Kalman equalizer in detecting $\mathbf{s}_{t-d}$ at time t. We have the following result.

Lemma 2.1 *Let* $1 \leq d \leq L_x - 1$. *Then*

$$\epsilon^2_{KF}(t;d) \leq \epsilon^2_{KF}(t-1;d-1) \tag{2.23}$$

holds for any given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$.

Proof: Let $\zeta_{t-1} = \left[\mathbf{y}_{t-1}^T \cdots \mathbf{y}_0^T \mathbf{s}_{t-d-1}^T \cdots \mathbf{s}_0^T\right]^T$. Let $\xi_t = \left[\mathbf{y}_t^T \zeta_{t-1}^T\right]^T$. From the Kalman filtering theory [34], it is clear that $\epsilon^2_{KF}(t;d)$ is the linear MMSE achieved in estimating $\mathbf{s}_{t-d}$ based on ξ_t, and $\epsilon^2_{KF}(t-1;,d-1)$ is the linear MMSE achieved in estimating $\mathbf{s}_{t-d}$ based on ζ_{t-1}.

Denote $\mathbf{K}_{\xi_t}$ and $\mathbf{K}_{\zeta_{t-1}}$ to be linear estimators of $\mathbf{s}_{t-d}$ based on ξ_t and ζ_{t-1}, respectively. Then the error covariance matrix corresponding to $\mathbf{K}_{\xi_t}$ can be written as

$$\begin{aligned} \mathrm{MSE}\left(\mathbf{K}_{\xi_t}\right) &= \mathrm{E}\left[\left(\mathbf{s}_{t-d} - \mathbf{K}_{\xi_t}\xi_t\right)\left(\mathbf{s}_{t-d} - \mathbf{K}_{\xi_t}\xi_t\right)^*\right] \\ &= \mathbf{R}_s - \mathbf{R}_{s\xi_t}\mathbf{K}^*_{\xi_t} - \mathbf{K}_{\xi_t}\mathbf{R}^*_{s\xi_t} + \mathbf{K}_{\xi_t}\mathbf{R}_{\xi_t}\mathbf{K}^*_{\xi_t} \\ &= \begin{bmatrix} \mathbf{I}_M & \mathbf{K}^*_{\xi_t} \end{bmatrix} \begin{bmatrix} \mathbf{R}_s & -\mathbf{R}_{s\xi_t} \\ -\mathbf{R}^*_{s\xi_t} & \mathbf{R}_{\xi_t} \end{bmatrix} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{K}^*_{\xi_t} \end{bmatrix}, \end{aligned} \tag{2.24}$$

where $\mathbf{R}_{s\xi_t} = \mathrm{E}\left[\mathbf{s}_{t-d}\xi_t^*\right]$ and $\mathbf{R}_{\xi_t} = \mathrm{E}\left[\xi_t\xi_t^*\right]$. The above equation is of standard quadratic form. It is straightforward to show that $\mathrm{MSE}\left(\mathbf{K}_{\xi_t}\right)$ is minimized when $\mathbf{K}_{\xi_t} = \mathbf{R}_{s\xi_t}\mathbf{R}^{-1}_{\xi_t}$, and

$$\mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) = \mathbf{R}_s - \mathbf{R}_{s\xi_t}\mathbf{R}^{-1}_{\xi_t}\mathbf{R}^*_{s\xi_t}. \tag{2.25}$$

Similarly, we can show that the error covariance matrix corresponding to $\mathbf{K}_{\zeta_{t-1}}$ is minimized when $\mathbf{K}_{\zeta_{t-1}} = \mathbf{R}_{s\zeta_{t-1}}\mathbf{R}^{-1}_{\zeta_{t-1}}$, and

$$\mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right) = \mathbf{R}_s - \mathbf{R}_{s\zeta_{t-1}}\mathbf{R}_{\zeta_{t-1}}^{-1}\mathbf{R}_{s\zeta_{t-1}}^{*}, \tag{2.26}$$

where $\mathbf{R}_{s\zeta_{t-1}} = \mathrm{E}\left[\mathbf{s}_{t-d}\zeta_{t-1}^{*}\right]$ and $\mathbf{R}_{\zeta_{t-1}} = \mathrm{E}\left[\zeta_{t-1}\zeta_{t-1}^{*}\right]$.

On the other hand,

$$\begin{aligned}
\mathbf{R}_{\xi_t} &= \mathrm{E}\left[\xi_t\xi_t^{*}\right] \\
&= \mathrm{E}\left(\begin{bmatrix} \mathbf{y}_t \\ \zeta_{t-1} \end{bmatrix}\begin{bmatrix} \mathbf{y}_t^{*} & \zeta_{t-1}^{*} \end{bmatrix}\right), \\
&= \begin{bmatrix} \mathbf{R}_y & \mathbf{R}_{y\zeta_{t-1}} \\ \mathbf{R}_{y\zeta_{t-1}}^{*} & \mathbf{R}_{\zeta_{t-1}} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{R}_{s\xi_t} &= \mathrm{E}\left[\mathbf{s}_{t-d}\xi_t^{*}\right] \\
&= \mathrm{E}\left(\mathbf{s}_{t-d}\begin{bmatrix} \mathbf{y}_t^{*} & \zeta_{t-1}^{*} \end{bmatrix}\right), \\
&= \begin{bmatrix} \mathbf{R}_{sy} & \mathbf{R}_{s\zeta_{t-1}} \end{bmatrix}
\end{aligned}$$

where $\mathbf{R}_y = \mathrm{E}\left[\mathbf{y}_t\mathbf{y}_t^{*}\right]$, $\mathbf{R}_{y\zeta_{t-1}} = \mathrm{E}\left[\mathbf{y}_t\zeta_{t-1}^{*}\right]$, and $\mathbf{R}_{sy} = \mathrm{E}\left[\mathbf{s}_{t-d}\mathbf{y}_t^{*}\right]$. Hence

$$\begin{aligned}
\mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) &= \mathbf{R}_s - \mathbf{R}_{s\xi_t}\mathbf{R}_{\xi_t}^{-1}\mathbf{R}_{s\xi_t}^{*} \\
&= \mathbf{R}_s - \begin{bmatrix} \mathbf{R}_{sy} & \mathbf{R}_{s\zeta_{t-1}} \end{bmatrix}\begin{bmatrix} \mathbf{R}_y & \mathbf{R}_{y\zeta_{t-1}} \\ \mathbf{R}_{y\zeta_{t-1}}^{*} & \mathbf{R}_{\zeta_{t-1}} \end{bmatrix}^{-1}\begin{bmatrix} \mathbf{R}_{sy}^{*} \\ \mathbf{R}_{s\zeta_{t-1}}^{*} \end{bmatrix} \\
&= \mathbf{R}_s - \begin{bmatrix} \mathbf{R}_{sy} & \mathbf{R}_{s\zeta_{t-1}} \end{bmatrix}\left(\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\zeta_{t-1}}^{-1} \end{bmatrix} + \mathbf{\Psi}\right)\begin{bmatrix} \mathbf{R}_{sy}^{*} \\ \mathbf{R}_{s\zeta_{t-1}}^{*} \end{bmatrix} \\
&= \mathbf{R}_s - \mathbf{R}_{s\zeta_{t-1}}\mathbf{R}_{\zeta_{t-1}}^{-1}\mathbf{R}_{s\zeta_{t-1}}^{*} - \begin{bmatrix} \mathbf{R}_{sy} & \mathbf{R}_{s\zeta_{t-1}} \end{bmatrix}\mathbf{\Psi}\begin{bmatrix} \mathbf{R}_{sy}^{*} \\ \mathbf{R}_{s\zeta_{t-1}}^{*} \end{bmatrix} \\
&= \mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right) - \begin{bmatrix} \mathbf{R}_{sy} & \mathbf{R}_{s\zeta_{t-1}} \end{bmatrix}\mathbf{\Psi}\begin{bmatrix} \mathbf{R}_{sy}^{*} \\ \mathbf{R}_{s\zeta_{t-1}}^{*} \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{\Psi} &= \begin{bmatrix} \mathbf{I}_N \\ -\mathbf{R}_{\zeta_{t-1}}^{-1}\mathbf{R}_{y\zeta_{t-1}}^{*} \end{bmatrix}\left(\mathbf{R}_y - \mathbf{R}_{y\zeta_{t-1}}\mathbf{R}_{\zeta_{t-1}}^{-1}\mathbf{R}_{y\zeta_{t-1}}^{*}\right) \\
&\quad \times \begin{bmatrix} \mathbf{I}_N & -\mathbf{R}_{y\zeta_{t-1}}\mathbf{R}_{\zeta_{t-1}}^{-1} \end{bmatrix}
\end{aligned}$$

It is clear that $\mathbf{\Psi} \geq 0$. Thus $\mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) \leq \mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right)$. By the definition of $\epsilon^2_{KF}(t;d)$ and $\epsilon^2_{KF}(t-1;d-1)$, we have

$$\begin{aligned} \epsilon^2_{KF}(t;d) &= \mathrm{Tr}\left(\mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right)\right) \\ &= \textstyle\sum_{i=1}^{M} \mathbf{e}_i \mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) \mathbf{e}_i^T \end{aligned}, \tag{2.27}$$

and

$$\begin{aligned} \epsilon^2_{KF}(t-1;d-1) &= \mathrm{Tr}\left(\mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right)\right) \\ &= \textstyle\sum_{i=1}^{M} \mathbf{e}_i \mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right) \mathbf{e}_i^T \end{aligned}, \tag{2.28}$$

where $\mathbf{e}_i = [\overbrace{0 \cdots 0}^{(i-1)0's} 1\ 0 \cdots 0]$ is an $M \times 1$ vector. Since $\mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) \leq \mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right)$, so $\mathbf{e}_i^T \mathrm{MSE}_{\min}\left(\mathbf{K}_{\xi_t}\right) \mathbf{e}_i \leq \mathbf{e}_i^T \mathrm{MSE}_{\min}\left(\mathbf{K}_{\zeta_{t-1}}\right) \mathbf{e}_i$ holds true for $i = 1, \cdots, M$. Thus we obtain $\epsilon^2_{KF}(t;d) \leq \epsilon^2_{KF}(t-1;d-1)$.

This completes the proof. ∎

Note that $\epsilon^2_{KF}(t;d)$ and $\epsilon^2_{KF}(t-1;d-1)$ are the MMSEs achieved in detecting the same symbol $\mathbf{s}_{t-d}$, but at different time and with different detection delay. For time-invariant channels, the MMSE achieved by the Kalman equalizer is independent of time t; that is, $\epsilon^2_{KF}(t;d) = \epsilon^2_{KF}(d)$ and $\epsilon^2_{KF}(t-1;d-1) = \epsilon^2_{KF}(d-1)$, and thus the inequality (2.1) reduces to $\epsilon^2_{KF}(d) \leq \epsilon^2_{KF}(d-1)$.

Lemma 2.1 suggests that increasing the detection delay improves the performance of the Kalman equalizer. Lemma 2.1 can also be explained from the Kalman smoother point of view. The larger is the detection delay d, the more received signals in the future of $\mathbf{s}_{t-d}$ are utilized to detect $\mathbf{s}_{t-d}$. Hence better performance is achieved by using large detection delay. However, large detection delay also means high dimension of the state-space model, or complexity.

2.3.3 Comparison with Finite Length MMSE Equalizers

For any given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ and the detection delay d, the corresponding Kalman LE and DFE are unique[3]. However, there exist many finite length LEs and DFEs that have the same detection delay but of different filter order(s). Due to the optimality of the Kalman filter, it is straightforward to verify the following lemma.

Lemma 2.2 *Given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$, the following inequality holds true for all finite length equalizers that have the same detection delay as the corresponding Kalman equalizers:*

$$\epsilon_{KF}^2(t;d) \leq \epsilon_{FL}^2(t;d), \tag{2.29}$$

where $\epsilon_{FL}^2(t;d) = \min\{\mathrm{Tr}\left[\mathrm{E}\left(\mathbf{\Upsilon}_t\mathbf{\Upsilon}_t^*\right)\right]\} = \mathrm{Tr}\left(\mathbf{E}_{\min,t}\right)$ denotes the MMSE achieved by a finite length equalizer in detecting $\mathbf{s}_{t-d}$ at time t.

Proof: This follows directly from the optimality of the Kalman filter. The Kalman DFE utilizes all the received signals and the past detected symbols up to current time to produce an estimate of the transmitted symbol. Let $\mathbf{z}_{KF} = \left[\mathbf{y}_t^T \cdots \mathbf{y}_0^T \mathbf{s}_{t-d-1}^T \cdots \mathbf{s}_0^T\right]^T$. Then it is clear that $\epsilon_{KF}^2(t;d)$ is the linear MMSE achieved in estimating $\mathbf{s}_{t-d}$ based on $\mathbf{z}_{ZF}$. In contrast, due to the restriction on filter length, a finite length equalizer utilizes only part of the received signal and part of the past detected symbols to produce an estimate. Let $\mathbf{z}_{FL} = \left[\mathbf{y}_t^T \cdots \mathbf{y}_{t-L_g}^T \mathbf{s}_{t-d-1}^T \cdots \mathbf{s}_{t-d-L_b}^T\right]^T$. Then $\epsilon_{FL}^2(t;d)$ is the linear MMSE achieved in estimating $\mathbf{s}_{t-d}$ based on $\mathbf{z}_{FL}$. Following the same procedure

[3]This follows from the uniqueness of the Kalman filter.

as in the proof of Lemma 2.1, (replacing ζ_{t-1} and ξ_t with $\mathbf{z}_{FL}$ and $\mathbf{z}_{KF}$, respectively,) we can show that $\epsilon^2_{KF}(t;d) \leq \epsilon^2_{FL}(t;d)$. ∎

The Kalman filter recursions in (2.20) based on the state-space realization (2.19) is in fact a linear filtering process. Different from the finite length equalizer, (the finite length DFE is also a linear filter by the assumption of correct past detections,) the Kalman equalizer utilizes *all* the received signals and *all* the past detected symbols available up to current time to produce an estimate of the symbol to be detected. However, due to the restrictions on the filter length in a finite length equalizer, it may utilize only part of the available received signals and part of the past detected symbols to produce an estimate. Consequently, a finite length equalizer can never outperform its corresponding Kalman equalizer. This also qualitatively explains the inequality in (2.29).

It is very interesting, however, that a finite length DFE does achieve the performance of a Kalman DFE under some conditions, as shown in the following theorem.

Theorem 2.1 *Denote $\epsilon^2_{FLDFE}(t;d)$ [$\epsilon^2_{KFDFE}(t;d)$] to be the MMSE achieved by a finite length DFE [Kalman DFE] in detection $\mathbf{s}_{t-d}$ at time t. Given channel $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ and detection delay d, a finite length DFE achieves the mean squared error of the Kalman DFE of the same detection delay, i.e., $\epsilon^2_{FLDFE}(t;d) = \epsilon^2_{KFDFE}(t;d)$, if $L_g \geq d$ and $L_b \geq L$.*

Proof: We need to show that the finite length DFE actually achieves the minimum possible mean squared error $\epsilon^2_{KFDFE}(t,d)$ for a given detection delay d if $Lg \geq d$ and

$L_b \geq L$. Assume $L_g \geq d$ and $L_b \geq L$. Let

$$\begin{aligned}\mathbf{R} &= \left[\mathbf{R}_{\underline{ss}}^{-1} + \underline{\mathbf{H}}_t^* \mathbf{R}_{\underline{vv}}^{-1} \underline{\mathbf{H}}_t\right] \\ &= \begin{bmatrix} \mathbf{A}_{M(d+L_b+1)\times M(d+Lb+1)} & \mathbf{B}_{M(d+Lb+1)\times M\gamma} \\ \mathbf{B}^*_{M\gamma\times M(d+Lb+1)} & \mathbf{C}_{M\gamma\times M\gamma} \end{bmatrix}.\end{aligned}$$

Then

$$\begin{aligned}\tilde{\mathbf{R}} &= \begin{bmatrix} \mathbf{0}_{M(L_b+1)\times Md} & \mathbf{I}_{M(L_b+1)} & \mathbf{0}_{M(L_b+1)\times M\gamma} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{C} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0}_{Md\times M(L_b+1)} \\ \mathbf{I}_{M(L_b+1)} \\ \mathbf{0}_{M\gamma\times M(L_b+1)} \end{bmatrix} \\ &= \left[\mathbf{0}_{M(L_b+1)\times Md} \;\; \mathbf{I}_{M(L_b+1)}\right] \Delta_C^{-1} \begin{bmatrix} \mathbf{0}_{Md\times M(L_b+1)} \\ \mathbf{I}_{M(L_b+1)} \end{bmatrix}\end{aligned} \quad , \tag{2.30}$$

where $\Delta_C = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^*$ is the Schur complement of $\mathbf{C}$ in $\mathbf{R}$. By matrix inversion theorem [34],

$$\Delta_C^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(-\mathbf{C} + \mathbf{B}^*\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^*\mathbf{A}^{-1}. \tag{2.31}$$

Further decompose $\mathbf{A}$ and $\mathbf{B}$ into sub-matrices. Let $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{bmatrix}$, where $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, and $\mathbf{A}_{22}$ have dimensions $Md\times Md$, $Md\times M(L_b+1)$, and $M(L_b+1)\times M(L_b+1)$, respectively. Let $\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}$, where $\mathbf{B}_1$, $\mathbf{B}_2$ have dimensions $Md \times M\gamma$, $M(L_b + 1) \times M\gamma$, respectively. Then

$$\begin{aligned}\mathbf{A}^{-1} &= \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^* & \mathbf{A}_{22} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\Delta_{11}^{-1}\mathbf{A}_{12}^*\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\Delta_{11}^{-1} \\ -\Delta_{11}^{-1}\mathbf{A}_{12}^*\mathbf{A}_{11}^{-1} & \Delta_{11}^{-1} \end{bmatrix}\end{aligned} \tag{2.32}$$

where $\Delta_{11} = \mathbf{A}_{22} - \mathbf{A}_{12}^*\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ is the Schur complement of $\mathbf{A}_{11}$ in $\mathbf{A}$. Note that if the input signal and the noise are both white, i.e. $\mathbf{R}_{\underline{ss}}$ and $\mathbf{R}_{\underline{vv}}$ are diagonal matrices,

then $\mathbf{R}$ is a block band-limited matrix[4] with bandwidth equal to the channel order L. In this case, if $L_b \geq L$, then $\mathbf{B}_1 = 0$. Combining (2.30), (2.31), and (2.32) and by direct calculation, we arrive at

$$\begin{aligned} \tilde{\mathbf{R}} &= \Delta_{11}^{-1} - \Delta_{11}^{-1}\mathbf{B}_2\left(-\mathbf{C} + \mathbf{B}_2^*\Delta_{11}^{-1}\mathbf{B}_2\right)^{-1}\mathbf{B}_2^*\Delta_{11}^{-1} \\ &= \left(\Delta_{11} - \mathbf{B}_2\mathbf{C}^{-1}\mathbf{B}_2^*\right)^{-1} \\ &= \left(\mathbf{A}_{22} - \mathbf{A}_{12}^*\mathbf{A}_{11}^{-1}\mathbf{A}_{12} - \mathbf{B}_2\mathbf{C}^{-1}\mathbf{B}_2^*\right)^{-1} \\ &= \left(\mathbf{A}_{22} - \mathbf{Z}_1 - \mathbf{Z}_2\right)^{-1} \end{aligned} , \tag{2.33}$$

where we have defined $\mathbf{Z}_1 = \mathbf{A}_{12}^*\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ and $\mathbf{Z}_2 = \mathbf{B}_2\mathbf{C}^{-1}\mathbf{B}_2^*$.

From (2.13) and (2.33), we have $\Delta_{\tilde{C}}^{-1} = \mathbf{A}_{22}(1:M, 1:M) - \mathbf{Z}_1(1:M, 1:M) - \mathbf{Z}_2(1:M, 1:M)$. It is clear that for any given detection delay d, $\mathbf{A}_{22}(1:M, 1:M)$ is a constant matrix. Due to the band-limited structure of $\mathbf{R}$, the first $M(L_b+1-L)$ rows of $\mathbf{B}_2$, and the last $M(L_b+1-L)$ columns of $\mathbf{A}_{12}$ are all zeros. By direct calculation, it can be verified that $\mathbf{Z}_1(1:M, 1:M)$ is a constant matrix and $\mathbf{Z}_2(1:M, 1:M)$ is a zero matrix. Thus if $L_g \geq d$ and $L_b \geq L$, then $\mathbf{E}_{\min,t}^{FLDFE} = \Delta_{\tilde{C}}$ is a constant matrix too. In this case increasing L_b and L_g will not reduce the MMSE achieved by the finite length DFE.

On the other hand, following the same procedure as in the proof of Lemma 2.1, we can show that $\epsilon_{FL}^2(t;d)$ is a decreasing function with respect to increasing L_b

[4]We define a block band-limited matrix as a Hermitian matrix

$$\mathbf{A} = \mathbf{A}^* = \begin{pmatrix} \mathbf{\Omega}_{11} & \mathbf{\Omega}_{12} & \cdots & \mathbf{\Omega}_{1M} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{\Omega}_{M1} & \mathbf{\Omega}_{M2} & \cdots & \mathbf{\Omega}_{MN} \end{pmatrix},$$

where $\mathbf{\Omega}_{ij}$ is of the same dimension for all i and j, satisfying $\mathbf{\Omega}_{ij} = 0$ if $j > i + \beta$ or $i > j + \beta$, and β is known as the bandwidth of the block band-limited matrix.

or L_g. Therefore, increasing the FFF order and the FBF order results in smaller estimation error. However, we have shown previously that when $L_b \geq L$ and $L_g \geq d$, $\epsilon^2_{FLDFE}(t;d) = \mathrm{Tr}\left[\mathbf{E}^{FLDFE}_{\min,t}\right]$ is a constant. Thus, the finite length DFE achieves the minimum possible mean squared error if $L_g \geq d$ and $L_b \geq L$. From Lemma 2.2 and by the uniqueness of Kalman filter, we have $\epsilon^2_{FL}(t;d) = \epsilon^2_{KF}(t;d)$ if $L_g \geq d$ and $L_b \geq L$.

This completes the proof. ∎

In general, a Kalman filter is an IIR filter. By Theorem 2.1 and the fact that the Kalman DFE is unique[5], we find it somehow surprising that, although it is based on the Kalman filter, the Kalman DFE is in fact equivalent to a finite length DFE with $L_b = L$ and $L_g = d$, or an FIR filter.

In fact, the Kalman gain matrix in the recursive algorithm (2.20) can be decomposed into two parts, i.e. $\mathbf{K}_t = \left[\mathbf{K}^F_t\ \mathbf{K}^B_t\right]$, where $\mathbf{K}^F_t$ and $\mathbf{K}^B_t$ are of dimension $(L_x+1)\times P$ and $(L_x+1)\times M$, respectively. As such, the FFF and the FBF of the Kalman DFE are actually realized through the following state-space recursive algorithms

$$\text{FFF}: \quad \tilde{\mathbf{x}}^F_t = \mathbf{F}\tilde{\mathbf{x}}^F_{t-1} + \mathbf{K}^F_t\left(\mathbf{y}_t - \mathbf{D}_t\mathbf{F}\tilde{\mathbf{x}}^F_{t-1}\right) \tag{2.34}$$

$$\text{FBF}: \quad \tilde{\mathbf{x}}^B_t = \mathbf{F}\tilde{\mathbf{x}}^B_{t-1} + \mathbf{K}^B_t\left(\hat{\mathbf{s}}_{t-d-1} - \mathbf{D}_t\mathbf{F}\tilde{\mathbf{x}}^B_{t-1}\right) \tag{2.35}$$

where $\tilde{\mathbf{x}}^F_t$ and $\tilde{\mathbf{x}}^F_t$ are the state vectors for the FFF and the FBF, respectively. In this sense, the Kalman DFE is just another form of DFE that implements the FFF and the FBF using recursive algorithms. It seems that the only difference between the

[5]This follows from the uniqueness of the Kalman filter.

Kalman DFE and the finite length DFE of proper filter length is the way how they are implemented. However, the Kalman DFE computes the optimal estimate at time t based on the optimal estimate at time $t-1$, which is a recursive algorithm as shown in (2.20). On the other hand, for time-varying channels, the conventional finite length DFE (*cf.* (2.11) and (2.12)) needs to update the FFF and the FBF coefficients at every time instance without utilizing any previously obtained information about the transmitted signals. As a result, the Kalman DFE is more computationally efficient than the finite length DFE for time-varying channels.

We have shown in our previous discussions that a linear equalizer can be regarded as the special case of a decision-feedback equalizer by setting the feedback signal to zero. Since the FBF length $L_b = 0$ in an LE, the conditions in Theorem 2.1 do not hold any more. As such, an LE has quite different properties than a DFE, as shown in the following Corollary.

Corollary 2.1 *Given channel* $\{\mathbf{H}_{t,i}\}_{i=0}^{L}$ *and the detection delay* d*, a Kalman LE is an IIR filter. The equality holds in (2.29) only when the filter length of the finite length LE approaches to infinity, i.e.,* $L_g = \infty$.

Proof: We will show that only an IIR filter with the received signal $\mathbf{y}_t$ as input and an estimate of $\mathbf{s}_{t-d}$ as output can achieves the performance of a Kalman LE of detection delay d. Then by the uniqueness of the Kalman LE, we have that a Kalman LE is an IIR filter.

Let $\mathbf{z}_{Lg} = \left[\mathbf{y}_t\ \mathbf{y}_{t-1}\ \cdots\ \mathbf{y}_{t-L_g}\right]^T$. Denote $\epsilon^2_{FLLE}(t,d,L_g)$ to be the MSE achieved by the finite length LE in detecting $\mathbf{s}_{t-d}$ based on $\mathbf{z}_{L_g}$. Following the same procedure as in the proof of Lemma 2.1, we obtain that $\epsilon^2_{FLLE}(t,d,L_g) \leq \epsilon^2_{FLLE}(t,d,L_g-1)$. Therefore, the MSE achieved by a finite length LE is a monotonically decreasing function with respect to L_g. From Lemma 2.2, we know that $\epsilon^2_{FLLE}(t,d,L_g) \leq \epsilon^2_{KFLE}(t,d)$ for any given L_g. Note that any lower (upper) bounded monotonically decreasing (increasing) function must have a limit [55]. Since both Kalman LE and finite length LE are linear filters, so the limit of the MSE of a finite length LE has to be the MSE of the corresponding Kalman LE, i.e., $\epsilon^2_{FLLE}(t,d,\infty) = \epsilon_{KFLE}(t,d)$.

On the other hand, since the FBF length $L_b = 0$ in an LE, we obtain from (2.11) and (2.12) the FFF coefficient vector of an optimal finite length LE as below

$$\begin{aligned} \mathbf{M}_t^{opt} &= \mathbf{R}_{\underline{yy}}^{-1}\mathbf{R}_{\underline{sy}}^{*}\left[\mathbf{0}_{M\times Md}\ \mathbf{I}_M\ \mathbf{0}_{M\times M\gamma}\right]^* \\ &= \left(\underline{\mathbf{H}}_t\mathbf{R}_{\underline{ss}}\underline{\mathbf{H}}_t^* + \mathbf{R}_{\underline{vv}}\right)^{-1}\left[\mathbf{0}_{M\times Md}\ \mathbf{I}_M\ \mathbf{0}_{M\times M\gamma}\right]^* \end{aligned} . \tag{2.36}$$

Since $\left(\underline{\mathbf{H}}_t\mathbf{R}_{\underline{ss}}\underline{\mathbf{H}}_t^* + \mathbf{R}_{\underline{vv}}\right)^{-1}$ is not a band-limited matrix[6] in general, direct calculation shows that we can always find a arbitrarily large integer L_g such that $\mathbf{M}_{t,L_g} \neq \mathbf{0}_{M\times M}$ if $L > 0$. Hence a Kalman LE is an IIR filter.

This completes the proof. ∎

[6]We define a band-limited matrix as a Hermition matrix $A = A^* = [a_{ij}]$ whose entries satisfy $a_{ij} = 0$ if $j > i+\beta$ or $i > j+\beta$, where *beta* is known as the bandwidth of the band-limited matrix.

2.3.4 Comparison with Existing Results

We highlight the connections between existing results and ours in this subsection.

- Our results in Theorem 2.1 and Corollary 2.1 are in consistent with those obtained in [45, 67, 73]. However, our derivations are carried out in the time domain. Hence our results are applicable to both time-invariant and time-varying channels. In contrast, [45] employs the Wiener filter theory, while [67] and [73] employ a frequency domain polynomial decomposition method in their derivations. Therefore the results of [45, 67, 73] cannot be applied to time-varying channels. Moreover, we have assumed a MIMO channel in our derivations, while both [45] and [67] work on single input channels only. In this sense, our results are more general than those of [45, 67, 73].

- In [1, 81], Al-Dhahir *et al* showed that the optimal detection delay $d \in [L_g, L_g + L]$. They suggested to choose $d = L_g$. From Lemma 2.1 and Theorem 2.1, we find that the optimal detection delay is indeed given by $d_{\text{opt}} = L_g$, if $L_b \geq L$. This confirms the choice in [1] and [81].

- Smee *et al* showed in [65] that simultaneously optimizing the FFF and FBF is equivalent to separately optimizing the two filters. This follows naturally from the equivalence of the Kalman DFE and the finite length DFE.

2.4 Extension to IIR Channels and Colored Signals/Noises

2.4.1 Optimal LE and DFE for IIR Channels

Wireline channels normally have very long channel impulse responses. To simplify channel modeling, many researchers have proposed to use IIR linear filters to characterize the wireline channels (for example [2], [14]). Recent results show that, even for wireless channels, IIR channel modeling is more desirable than its FIR counterpart in terms of accuracy and complexity of the model. See [88] and references therein. In this section, we will extend our results on optimal realizable equalization to IIR channels. To simplify discussion, we will focus on time-invariant SISO channels only. Generalization to MIMO channels can be done in a similar way.

As in [2], [14], and [67], a time-invariant SISO IIR channel can be represented by an auto-regressive moving average (ARMA) model, whose input-output relation can be written as

$$\begin{array}{rcl} y_t & = & \mathcal{H}(q)s_t + v_t = \frac{\mathcal{A}(q)}{\mathcal{B}(q)}s_t + v_t \\ & = & \sum_{i=0}^{M} \alpha_i s_{t-i} - \sum_{j=1}^{N} \beta_j y_{t-j} + v_t \end{array}, \tag{2.37}$$

where

$$\mathcal{H}(q) = \frac{\mathcal{A}(q)}{\mathcal{B}(q)} = \sum_{i=0}^{\infty} h_i q^{-i}, \tag{2.38}$$

is the transform domain representation of the channel with h_i the ith channel impulse response coefficient, $\mathcal{A}(q) = \alpha_0 + \alpha_1 q^{-1} + \cdots + \alpha_M q^{-M}$ and $\mathcal{B}(q) = 1 + \beta_1 q^{-1} + \cdots + \beta_N q^{-N}$ are linear filters of orders M and N, respectively, $\{\alpha_i\}_{i=0}^{M}$ and $\{\beta_j\}_{j=1}^{N}$ are the filter coefficients, q^{-1} is the unit shift operator ($q^{-1}s_t = s_{t-1}$). We call the roots of

$\mathcal{A}(q)$ and $\mathcal{B}(q)$ the zeros and the poles of the channel, respectively. For any physically stable channel, all its poles have to be strictly inside the unit circle. Note that when $N = 0$, the channel reduces to a conventional FIR channel of order M. We first assume that both s_t and v_t are white and uncorrelated to each other, with covariance given by σ_s^2 and σ_v^2, respectively. Extensions to colored signals and colored noises will be made later.

We have seen from our previous discussions that the key to the development of the optimal realizable equalization is to find an appropriate state-space model. For the IIR channel as specified in (2.37), we can choose $\mathbf{x}_t = [s_t\ s_{t-1} \cdots\ s_{t-L_s}\ y_{t-1}\ y_{t-2} \cdots\ y_{t-N}]^T$ as the state vector, where $L_s \geq M$. Then the input-output relation in (2.37) is equivalent to the following state-space realization

$$\begin{cases} \mathbf{x}_{t+1} &= \tilde{\mathbf{F}}\mathbf{x}_t + \tilde{\mathbf{G}}s_{t+1} \\ y_t &= \tilde{\mathbf{\Pi}}\mathbf{x}_t + v_t \end{cases}, \tag{2.39}$$

where $\tilde{\mathbf{F}} = \left[\tilde{\mathbf{F}}_1^T\ \tilde{\mathbf{F}}_2^T\ \tilde{\mathbf{F}}_3^T\right]^T$ with

$$\tilde{\mathbf{F}}_1 = \begin{bmatrix} 0 & \cdots & \cdots & 0 & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}_{(L_s+1)\times(N+L_s+1)}, \tag{2.40}$$

$$\tilde{\mathbf{F}}_2 = [1\ \alpha_1\ \cdots\ \alpha_M\ 0\ \cdots\ 0\ -\beta_1\ \cdots\ -\beta_N], \tag{2.41}$$

$$\tilde{\mathbf{F}}_3 = \left[\mathbf{0}_{(N-1)\times(L_s+1)}\ \mathbf{I}_{N-1}\ \mathbf{0}_{(N-1)\times 1}\right], \tag{2.42}$$

and $\tilde{\mathbf{G}} = [1\ 0\ \cdots\ 0]^T_{1\times(N+L_s+1)}$, $\tilde{\mathbf{\Pi}} = \tilde{\mathbf{F}}_2$.

Applying Kalman filter to (2.39), we can obtain the optimal linear estimate of $\mathbf{x}_t$ based on observations $\{y_i\}_{i=0}^{t}$. Repeat the process as discussed in Section 2.3.1. The symbol to be detected can be fetched by

$$\hat{s}_{t-d} = \text{Quan}\left(\tilde{\mathbf{\Gamma}}_d \hat{\mathbf{x}}_t\right), \tag{2.43}$$

where $\tilde{\mathbf{\Gamma}}_d = [\overbrace{0 \cdots 0}^{d\ zeros}\ 1\ 0\ \cdots\ 0]^T_{1\times(N+L_s+1)}$, $\hat{\mathbf{x}}_t$ is the estimate of $\mathbf{x}_t$, and $\text{Quan}(\cdot)$ represents quantization operation. Since only the received signals are used to generate the detected symbol, this process corresponds to an optimal LE (Kalman LE) for IIR channels.

The optimal DFE for IIR channels can be developed in a similar way. Define $\tilde{\mathbf{u}}_t = [y_t\ \hat{s}_{t-d-1}]^T$, $\tilde{\mathbf{H}} = \left[\tilde{\mathbf{\Pi}}^T\ \tilde{\mathbf{\Gamma}}_{d+1}^T\right]^T$, and $\tilde{\mathbf{w}}(t) = [v_t\ 0]^T$. A state-space realization for the Kalman DFE can be written as

$$\begin{cases} \mathbf{x}_{t+1} &= \tilde{\mathbf{F}}\mathbf{x}_t + \tilde{\mathbf{G}}s_{t+1} \\ \tilde{\mathbf{u}}_t &= \tilde{\mathbf{H}}\mathbf{x}_t + \tilde{\mathbf{w}}_t \end{cases} \tag{2.44}$$

Again Kalman filter can be applied to obtain the optimal estimate of $\mathbf{x}_t$, and the symbol to be detected can be fetched by (2.43). This results in the optimal DFE (Kalman DFE) for IIR channels.

Remark 2.2 *From Lemma 2.2 and Theorem 2.1, a Kalman LE and a Kalman DFE always perform strictly better than the corresponding finite length LE and DFE, respectively. This is in contrast to the case of FIR channels.*

2.4.2 Colored Signals and Noises

Colored signals/noises can be generated by passing white signals/noises through linear filters that are represented by ARMA models. We can obtain arbitrary "color", or rational power spectral of the signals/noises, by proper choices of the ARMA models. Figure 2.4 shows the input-output relation of a linear channel with colored input signal and colored additive noise, where s_t and v_t are zero-mean white sequences with variance given by σ_s^2 and σ_v^2, respectively, $\Psi(q)$, $\Phi(q)$ are both unit gain rational functions of q, representing the ARMA models of the signal and the noise, respectively, and $\mathcal{H}(q)$ is the transform domain representation of the channel as defined in (2.38). Thus, the received signal can be written as

$$y_t = \mathcal{H}(q)\Psi(q)s_t + \Phi(q)v_t. \tag{2.45}$$

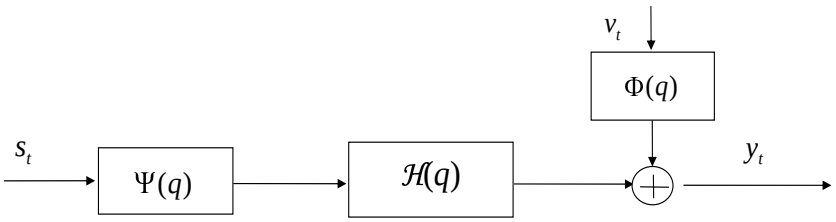

Figure 2.4: Linear channel with colored input signal and colored additive noise. (Original Channel)

To bypass the difficulty that arises due to the color of the signal and the noise, an equivalent channel model as shown in Figure 2.5 can be used [67], where $\Phi^{-1}(q)$ is

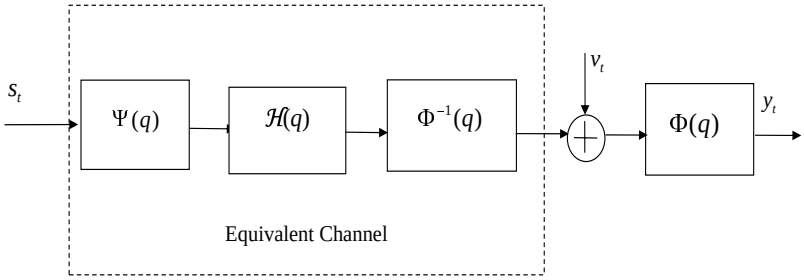

Figure 2.5: Linear channel with colored input signal and colored additive noise. (Equivalent Channel)

the inverse filter of $\Phi(q)$. Let $\tilde{H}(q) = \Psi(q)\mathcal{H}(q)\Phi(q)^{-1}$ denote the equivalent channel model. Depending on the signal filter $\Psi(q)$, the noise filter $\Phi(q)$ and the original channel $\mathcal{H}(q)$, the equivalent channel $\tilde{H}(q)$ can be either FIR or IIR. It is clear that the input signal and the additive noise to $\tilde{H}(q)$ are both white. Therefore, all the results we have obtained on optimal equalization for FIR and IIR channels are still applicable in this case. The only difference is that the LE or the FFF of the DFE need to incorporate an inverse filter of $\Phi(q)$ to compensate the effects of the noise filter. This does not affect our previous results.

2.5 Simulation Results

In this section, we will illustrate our results through simulations.

As we have pointed out before, all the derivations in this paper are carried out in the time domain. Thus our results are applicable to both time-invariant and time-varying channels. To show this, we first generate a single-input single-output time-varying multipath channel according to the Jakes' Model [32]. In practice, time-

varying channel has to be estimated online. This problem has been discussed in [29], [36], and will be studied in details in Chapter 5. For simplicity, we will assume that the receiver has perfect channel information in our simulations. In the first four examples, a SISO time-varying multipath channel with six-path ($L = 5$) will be simulated. BPSK modulation is used.

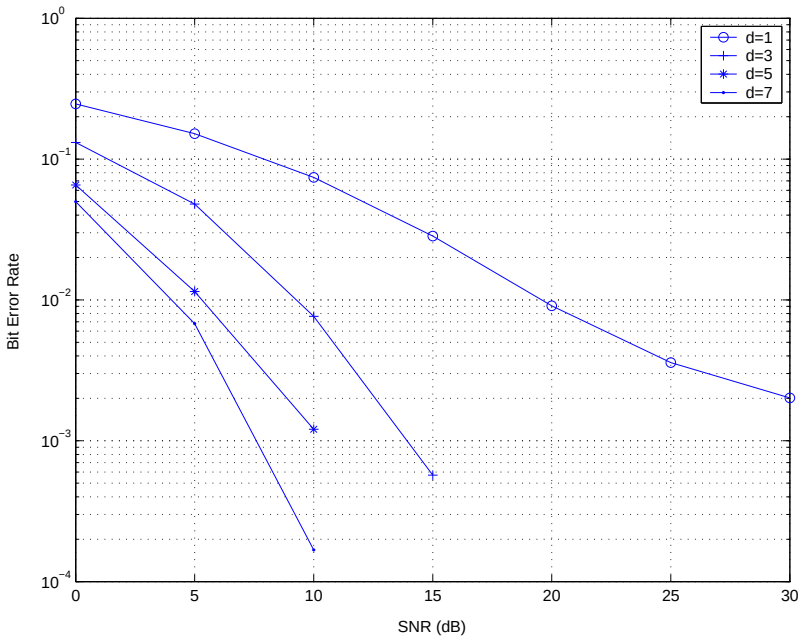

Figure 2.6: BER vs detection delay performance of Kalman DFE

Figures 2.6 and 2.7 give the bit error rate (BER) performance of the Kalman LE and Kalman DFE with respect to different detection delays, where the signal-to-noise ratio (SNR) is defined as SNR $= \frac{\sigma_s^2}{\sigma_v^2}$. Figure 2.6 shows the performance of the Kalman DFE, where the detected symbols are fed back to the FBF. Figure 2.7 shows that

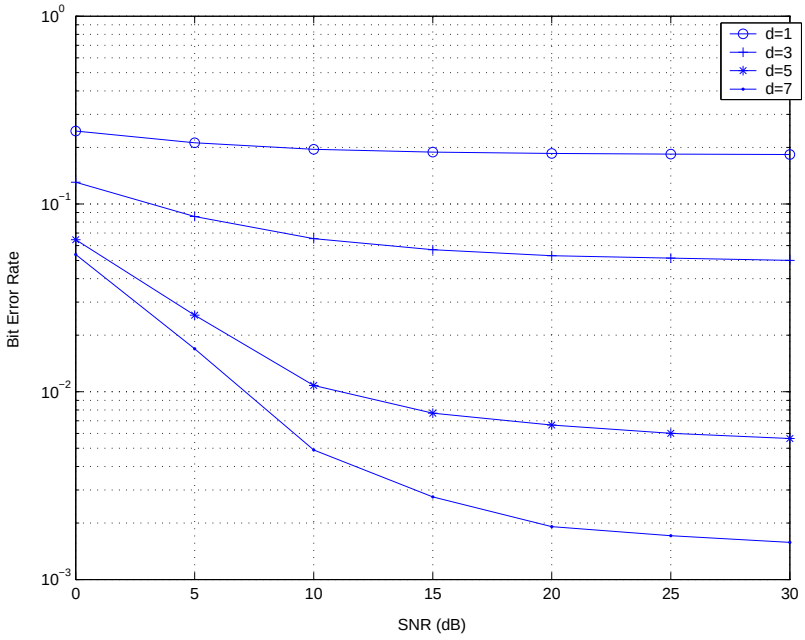

Figure 2.7: BER vs detection delay performance of Kalman LE

of the Kalman LE. Comparing Figure 2.6 and Figure 2.7, we see that the Kalman DFE has much better performance than the Kalman LE for all detection delays. It is also seen that for both the Kalman DFE and LE, their performance improve with respect to increasing the detection delay. This confirms our result in Lemma 2.1. Here we must note that there is only one parameter, the detection delay, needs to be optimized in Kalman equalizers. Since the detection delay is directly related to the dimension of the state vector, better performance (larger delay) also translates into higher complexity.

Figure 2.8 shows the relation between the Kalman DFE and the finite length DFE. In this example, the same time-varying multipath channel generated before is used, and the detection delays for both the finite length DFE and the Kalman DFE are

fixed to be $d = 7$. We find that the Kalman DFE has the best performance, which serves as the performance bound of a finite length DFE for a given detection delay. The performance of a finite length DFE improves as the FFF and the FBF lengths increase. When $L_g \geq d$ and $L_b \geq L$, it achieves the performance of the corresponding Kalman DFE. This confirms our results in Theorem 2.1.

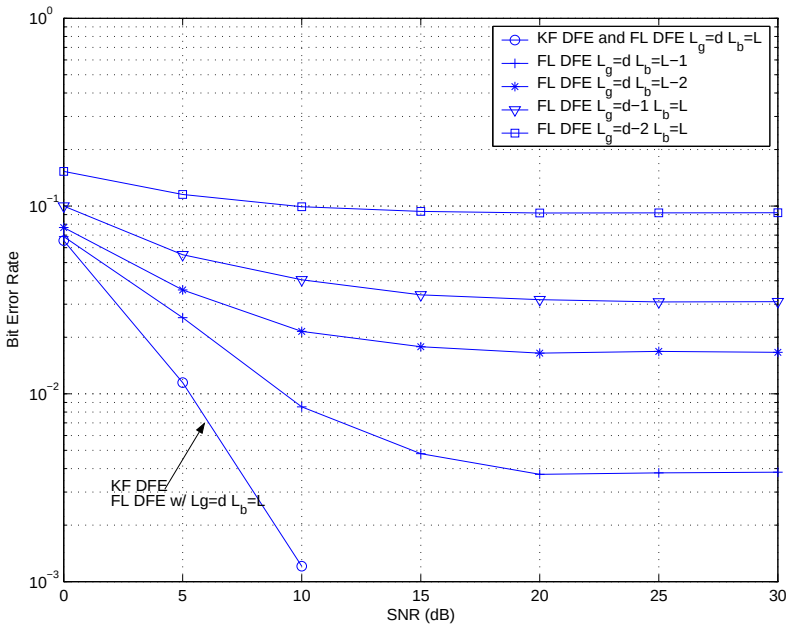

Figure 2.8: Relations between the Kalman DFE and the finite length DFE

Figure 2.9 shows the relation between the Kalman LE and the finite length LE. In this example, we have used the same time-varying multipath channel generated in the previous example. The detection delays for both the finite length LE and the Kalman LE are fixed to be $d = 7$. Again we see that Kalman LE has the best performance.

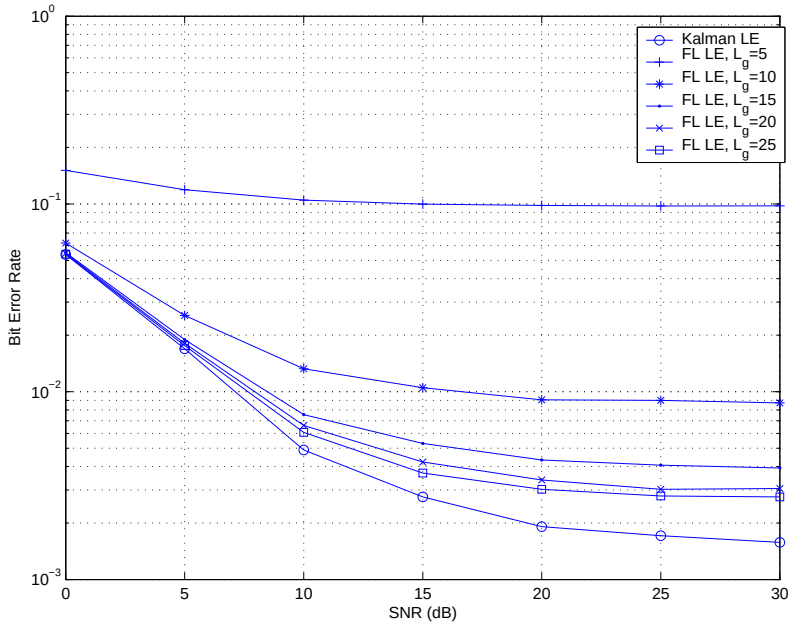

Figure 2.9: Relations between the Kalman LE and the finite length LE

It outperforms all the finite length LEs of different filter length. It is also seen that the performance of the finite length LE improves as its filter order increases. The larger the filter order of a finite length LE, the smaller is the performance gap as compared to the Kalman LE. However, different from the DFE case, there is always a performance gap between the finite length LE and the Kalman LE. This confirms Corollary 2.1.

We must note that our results are applicable to multi-input multi-output channels too, as will be shown in the following two examples, where BPSK modulation is used, and MIMO frequency selective fading channels that have zero mean and unit power

gain are randomly generated[7]. For the sake of simplicity, we simulate time-invariant MIMO channels only. However, as we have shown previously, all the results and conclusions can be carried over to time-varying channels too.

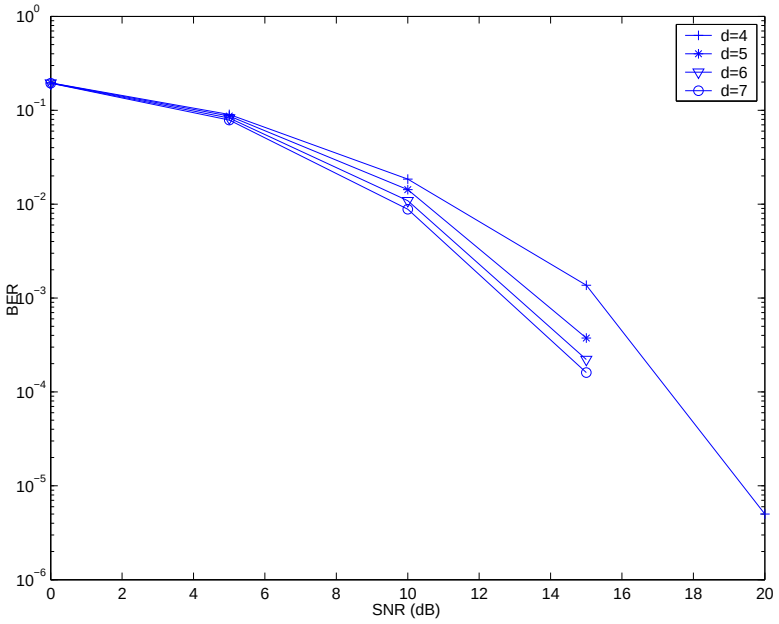

Figure 2.10: BER performance of Kalman DFE with respect to the detection delay. MIMO channel with $M = 2$, $N = 2$, $L = 4$.

We first focus on the Kalman DFE and its relation with the finite length MIMO DFE. A 2-input/2-output ($M = 2$, $P = 2$) system over a five-path ($L = 4$) frequency selective fading channels is simulated. 10000 vector signals are transmitted. Figure 2.10 shows the bit error rate (BER) performance of the Kalman DFE with respect to the detection delay. The SNR is defined as SNR $= \frac{\sigma_s^2}{\sigma_v^2}$. The curves are averaged over

[7]The elements in $\mathbf{H}_{t,i}$, $i = 0, 1, \cdots, L$, are i.i.d. zero mean complex Gaussian random variables. Unit power gain means that in the absence of noise, input signal and output signal have equal power.

30 randomly generated channels. It is clear to see that the performance improves as the detection delay increases. Hence it confirms our result in Lemma 2.1. It is also interesting to observe that as the detection delay increases, the improvement of the performance gain becomes less obvious.

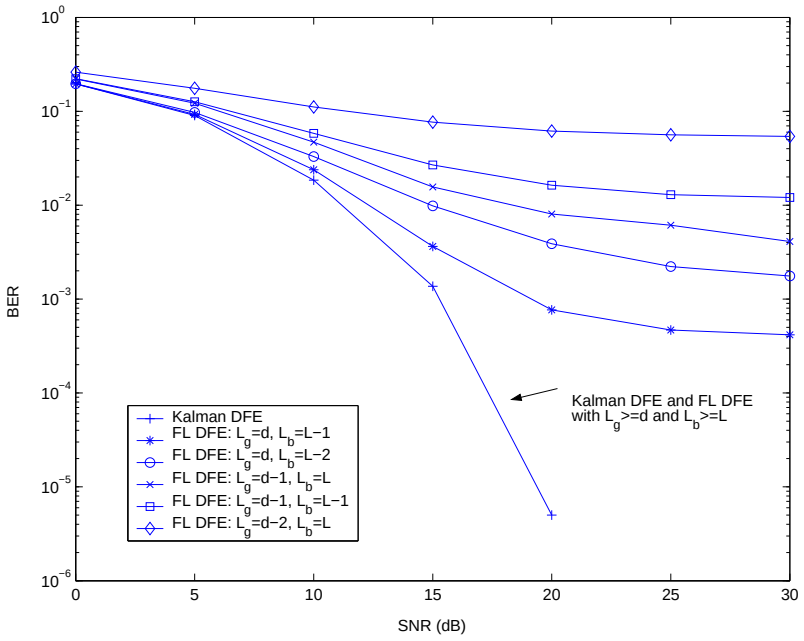

Figure 2.11: Comparison between the Kalman DFE and the finite length DFE (FL DFE). $M = 2$, $N = 2$, $L = 4$, $d = 4$.

Figure 2.11 compares the BER performance of the Kalman DFE and the finite length DFE. 10000 vector signals are transmitted and the channels used are the same as in the previous example. The curves are averaged over 30 randomly generated channels. The detection delay is fixed to be $d = 4$ for both the Kalman DFE and the

finite length DFE. We see that the finite length DFE never outperforms the Kalman DFE. The performance of the finite length DFE improves with the increase of the FFF and the FBF orders. When $L_g \geq d$ and $L_b \geq L$, the finite length DFE actually has identical performance as the Kalman DFE. This confirms Lemma 2.2 and Theorem 2.1. We also observe that reducing the FFF length causes much more performance loss than reducing the FBF length. In fact when the FBF order reduces to zero, the DFE becomes an LE that is still able to suppress the ISI to some extent.

2.6 Chapter Summary

In this chapter, we studied the realizable MMSE LE and DFE through a state-space approach. Kalman LE and Kalman DFE are proposed. Their properties and relations with the conventional finite length LEs and DFEs were investigated. It is shown that the state-space approach yields the optimal realizable equalizers. The resulting Kalman LE is an IIR filter, while the resulting Kalman DFE is an FIR filter.

As compared with those in [3], [16], [17], [66], the proposed two-stage method has an advantage that it enjoys more flexibility in implementation. More importantly, the state-space approach adopted in this paper renders our method applicable to any linear channels, including time-invariant, time-varying, FIR, and IIR channels, provided that the channel model admits a state-space realization. Hence the results reported in this paper are more general than the existing work. To our knowledge, our results are the most general results ever obtained.

Chapter 3

Optimal Realizable Suppression of Inter-Channel Interference

Different from single-input/single-output channels, a multi-input/multi-output frequency selective fading channel introduces not only inter-symbol interference (ISI), but also inter-(sub)channel interference (ICI). Even if the ISI can be completely cancelled by the Kalman DFE, due to the non-orthogonality between each pair of the sub-channels in a MIMO system, the symbols transmitted at the same time but from different transmitters may still interfere with each other. This causes the ICI. In this chapter, we assume that the ISI has been completely cancelled by the equalizer, and focus on the ICI suppression[1]. we will propose a novel successive interference cancellation (SIC) scheme based on spatial filtering to which Kalman filtering can again be applied to suppress the ICI. Combining the Kalman equalizer with the SIC, we can obtain an optimal realizable ISI and ICI cancellation with Kalman filtering a natural design tool. Simulation results show that the combined DFE and the SIC scheme based on Kalman filtering can effectively suppress both the ISI and the ICI.

[1]More accurately speaking, a DFE can suppress not only the ISI, but also part of the ICI. The ICI suppression discussed in this section is to remove the residue ICI at the output of a DFE.

3.1 ICI Suppression Via SIC

3.1.1 Basic SIC

Let us take another look at the state-space model in (2.19). As before, we denote $\tilde{\mathbf{x}}_t$ to be the state estimate (before the quantization) of the Kalman DFE at time t, based on $\{\mathbf{u}_i\}_{i=0}^{t}$. Recall that $\mathbf{u}_t$ consists of not only the received signal $\mathbf{y}_t$, but also the detected symbol $\mathbf{s}_{t-d-1}$. The optimal linear estimate of $\mathbf{s}_{t-d}$ based on $\{\mathbf{u}_i\}_{i=0}^{t}$ is thus given by

$$\tilde{\mathbf{s}}_{t-d} = \mathbf{\Gamma}_d \tilde{\mathbf{x}}_t. \tag{3.1}$$

Without further ICI suppression, a hard decision on $\mathbf{s}_{t-d}$ can be obtained by quantization of $\tilde{\mathbf{s}}_{t-d}$ to the nearest constellation point, i.e.,

$$\hat{\mathbf{s}}_{t-d} = \text{Quan}\left(\tilde{\mathbf{s}}_{t-d}\right). \tag{3.2}$$

Note that (3.2) is just another form of (2.22).

Instead of using (3.2) to directly generate a hard decision on $\mathbf{s}_t$ from $\tilde{\mathbf{s}}_t$, the SIC scheme aims to further remove the ICI from $\tilde{\mathbf{s}}_t$ before a final hard decision is made. The basic idea of the SIC is to sequentially detect each symbol in $\tilde{\mathbf{s}}_t = [\tilde{s}_t(1)\ \tilde{s}_t(2)\ \cdots\ \tilde{s}_t(M)]^T$, based on spatial filtering and decision feedback. After $s_t(m)$ is detected, calculate its contribution to the ICI and remove the ICI caused by $s_t(m)$ from $\tilde{\mathbf{s}}_t$. Repeat the process and continue to detect other undetected symbols until all the symbols in $\tilde{\mathbf{s}}_t$ are detected.

The concept of the SIC scheme sounds simple. However, when applied to the ICI suppression in $\tilde{\mathbf{s}}_t$, a major difficulty arises in calculating the ICI contributed by each $s_t(m)$ for $m = 1, 2, \cdots, M$. To bypass this difficulty, we again seek the help of the state-space approach. Without loss of generality, in the following we assume that we detect $s_t(m)$ with respect to the ascending order of m, i.e., we always detect $s_t(m)$ before we detect $s_t(m+1)$. The more general case of choosing different detection orders will be discussed in the next subsection.

Suppose that we have just detected $s_{t-d}(m)$. We want to remove the ICI caused by $s_{t-d}(m)$ so that we can have a better detection of $s_{t-d}(m+1)$. Assume that all the past detections are correct. Using $s_{t-d}(m)$ as the observation, a state-space model can be established as

$$\begin{cases} \mathbf{x}_{t|m+1} &= \mathbf{x}_{t|m} \\ s_{t-d}(m) &= \mathbf{e}_m \mathbf{x}_{t|m} \end{cases}, \tag{3.3}$$

where $\mathbf{x}_{t|m} = \mathbf{\Gamma}_d \mathbf{x}_t = \mathbf{s}_{t-d}$, and

$$\mathbf{e}_m = \left[\mathbf{0}_{1\times(m-1)}\ 1\ \mathbf{0}_{1\times(M-m)}\right] \tag{3.4}$$

for $m = 1, 2, \cdots, M$. Applying the Kalman filter to state-space realization (3.3), we obtain the optimal linear estimate of $\mathbf{x}_t$ based on not only $\{\mathbf{u}_i\}_{i=0}^{t}$ but also $\{s_{t-d}(k)\}_{k=1}^{m}$. Denote the resulting optimal estimate to be $\tilde{\mathbf{x}}_{t|m}$. It is clear that $\tilde{\mathbf{x}}_{t|m}$ is a more reliable estimate of $\mathbf{x}_t$ than $\tilde{\mathbf{x}}_t$. Then we can move on to detect $s_{t-d}(m+1)$ by

$$\check{s}_{t-d}(m+1) = \mathbf{e}_{m+1} \times \text{Quan}\left(\tilde{\mathbf{x}}_{t|m}\right), \tag{3.5}$$

where $\check{s}_t(m)$ denotes the refined detection of $s_t(m)$, $m = 1, 2, \cdots, M$. In this way, the information provided by $\{s_{t-d}(k)\}_{k=1}^{m}$ is fully utilized in generating the final hard decision $\check{s}_{t-d}(m+1)$, and the ICI due to $\{s_{t-d}(k)\}_{k=1}^{m}$ is indirectly suppressed. Here we must note that the time index in (3.3) is m instead of t. Hence it in fact corresponds to a spatial filtering process.

The procedure described above can be repeated until all the symbols in $\tilde{\mathbf{s}}_{t-d}$ are detected. We summarize the basic SIC method into the following theorem.

Theorem 3.1 *After the ISI suppression by the Kalman DFE, the following recursions give the optimal ICI suppression for $m = 1, 2, \cdots, M$,*

$$\begin{aligned} \mathbf{\Phi}_m &= \lambda_m \mathbf{\Sigma}_{m-1} \mathbf{e}_m^* \\ \mathbf{\Sigma}_m &= \mathbf{\Sigma}_{m-1} - \lambda_m \mathbf{\Phi}_m \mathbf{\Phi}_m^* \\ \tilde{\mathbf{x}}_{t|m} &= \tilde{\mathbf{x}}_{t|m-1} + \mathbf{\Phi}_m \left(\check{s}_{t-d}(m) - \mathbf{e}_m \tilde{\mathbf{x}}_{t|m-1} \right) \\ \check{s}_{t-d}(m+1) &= \mathbf{e}_{m+1} \times \mathrm{Quan}\left(\tilde{\mathbf{x}}_{t|m} \right) \end{aligned} , \tag{3.6}$$

where $\lambda_m = \mathbf{e}_m \mathbf{\Sigma}_{m-1} \mathbf{e}_m^*$, the initial conditions are given by $\tilde{\mathbf{x}}_{t|0} = \mathbf{\Gamma}_d \tilde{\mathbf{x}}_t$, $\check{s}_{t-d}(1) = \mathbf{e}_1 \times \mathrm{Quan}\left(\tilde{\mathbf{x}}_t\right)$, and $\mathbf{\Sigma}_0 = \mathbf{\Gamma}_d \mathbf{P}_t \mathbf{\Gamma}_d^* = \mathrm{E}\left[\left(\mathbf{s}_t - \tilde{\mathbf{s}}_t \right) \left(\mathbf{s}_t - \tilde{\mathbf{s}}_t \right)^* \right]$ is the estimation error covariance matrix at the output of the Kalman DFE at time t.

Proof: Applying Kalman filter to the state-space model (3.3), we can obtain the recursions in (3.6). The optimality of this ICI suppression scheme follows from the optimality of the Kalman filter. ∎

In the recursive algorithm above, each transmitted symbol in $\mathbf{s}_t$ is serially detected. Since $\mathbf{e}_m$ has only one nonzero element, the recursive algorithm in (3.6) is very simple in computation and can be implemented easily.

3.1.2 Performance Enhancement with Reliability Sorting

In the previous discussions, we have assumed that all the detections are correct. By this assumption, the SIC method presented above will definitely improve the performance of interference suppression. However, since the correct detection of each symbol depends on the previously detected symbols, an error in detection may cause trouble in all the subsequent detections, resulting in error propagation. The earlier a symbol is incorrectly detected, the more negative effect will the error propagation have. Hence we wish to first detect those symbols that can be more reliably detected. These symbols correspond to those of strong power and they also cause more ICI than symbols of weak power. Detecting them in a early stage will not only prevent error propagation, but also create a more benign condition for the detection of weak signals. We call this scheme reliability sorting before SIC.

There are different ways of determining how reliable a symbol can be detected. In this paper, we define the reliability metric as the distance between the state estimate of each symbol and its associated hard decision at the output of the Kalman DFE, i.e.,

$$\delta_t(m) = \left|\tilde{s}_t(m) - \hat{s}_t(m)\right|, \tag{3.7}$$

for $m = 1, 2, \cdots, M$, where $\delta_t(m)$ denotes the reliability metric of detecting $s_t(m)$, $\tilde{s}_t(m)$ and $\hat{s}_t(m)$ are the mth elements of $\tilde{\mathbf{s}}_t$ and $\hat{\mathbf{s}}_t$ (*cf.* (3.1) and (3.2)), respectively. From statistical detection theory, we expect that small $\delta_t(m)$ corresponds to more reliable detection.

The reliability sorting can be implemented in the following way. After the Kalman DFE, we first obtain $\hat{\mathbf{s}}_t$ by making a tentative detection of $\mathbf{s}_t$. Then calculate the reliability metric of each element in $s_t(m)$ according to (3.7). Sort the reliability metric from low to high (the smallest is the distance between $\tilde{s}_t(m)$ and $\hat{s}_t(m)$, the lower is its metric). After the reliability sorting, we perform the SIC scheme discussed in the previous subsection, detecting the signal with the lowest metric first. Continue until all the symbols in $\hat{\mathbf{s}}_t$ are detected.

Compared with the basic SIC, performing reliability sorting improves the performance of ICI suppression at the expense of higher complexity. In Section 3.3, we will compare the two schemes through simulation examples.

3.2 Combined ISI and ICI suppression

The proposed SIC scheme is based on a state-space model. It uses the same state vector as the Kalman DFE and utilizes the Kalman DFE output to initialize its implementation. Hence the SIC algorithm can be naturally concatenated with the Kalman DFE. Note that the SIC recursions in (3.6) is also a variant of the Kalman filter. By the optimality of the Kalman filter, the combination of the two results in the optimal realizable ISI and ICI suppression. A complete block diagram of the optimal combined receiver is given in Figure 3.1. In the Kalman DFE discussed in Section 2.3, the feedback signal is $\hat{\mathbf{s}}_{t-d}$. However, in the combined receiver, the feedback signal to the Kalman DFE is the SIC output $\check{\mathbf{s}}_{t-d} = [\check{s}_{t-d}(1)\ \check{s}_{t-d}(2)\ \cdots\ \check{s}_{t-d}(M)]^T$. Since $\check{\mathbf{s}}_{t-d}$

is more reliable than $\hat{\mathbf{s}}_{t-d}$, the well known error propagation problem of the DFE can be lessened too.

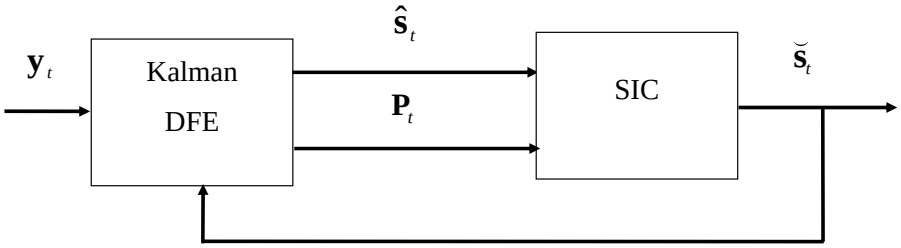

Figure 3.1: Optimal ISI and ICI suppression receiver.

A closer look at the recursive algorithm in (3.6) also reveals that the proposed SIC scheme does not require any channel information. It only needs the output of the ISI suppression in first stage to initialize, and thus can be combined with different first stages such as the Kalman DFE, the adaptive MIMO DFE, or the much simpler linear equalizer, etc. This feature gives the proposed method a lot of flexibility in implementation.

3.3 Simulation Results

In this section, we evaluate the proposed ICI suppression methods through simulations. The simulation settings are the same as those in Chapter 3. In all the simulation examples, BPSK modulation is used. MIMO frequency selective fading

channels that have zero mean and unit power gain are randomly generated. For the sake of simplicity, we simulate time-invariant channels only. However, as we have shown previously, all the results and conclusions can be carried over to time-varying channels too.

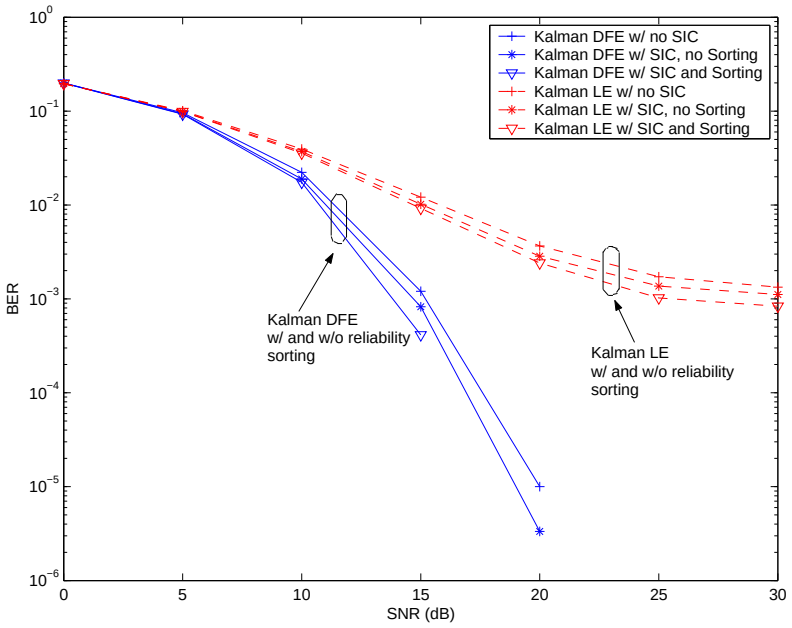

Figure 3.2: BER performance of the combined ISI and ICI suppression. 2-input/2-output system over 4th order channels.(MIMO channel with $M = 2$, $P = 2$, $L = 4$).

Figure 3.2 shows the BER performance of the combined ISI and ICI suppression, where a 2-input/2-output system over 4th order channels ($L = 4$) is simulated. The detection delay is fixed to be $d = 4$. Same as before, 10000 vector signals are transmitted and the results are averaged over 30 channel realizations that are randomly generated. The lower three curves in Figure 3.2 correspond to the performance of

the combined receiver with Kalman DFE as the first stage. We observe that the SIC scheme for ICI suppression does improve the performance. Compared with the case in which only Kalman DFE is employed and no further ICI suppression is performed, the basic SIC scheme (without reliability sorting) has about 0.5 dB SNR gain at the 10^{-3} BER level, while the SIC scheme with reliability sorting has another 1 dB gain over the basic SIC scheme.

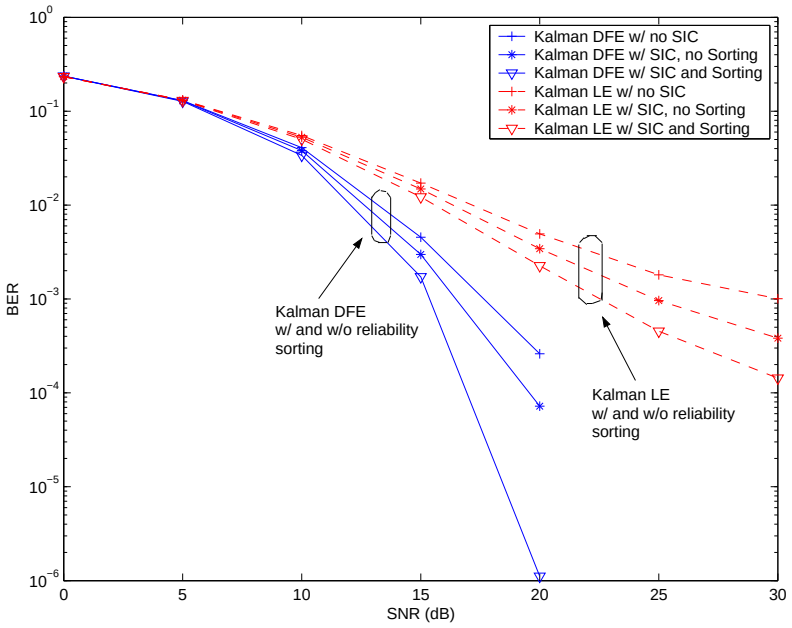

Figure 3.3: BER performance of the combined ISI and ICI suppression. 3-input/3-output system over 4th order channels.(MIMO channel with $M = 3$, $P = 3$, $L = 4$).

To show the flexibility of the proposed SIC scheme, we also combine the linear equalizer with the SIC. The performance of the combined receiver with Kalman LE as the first stage is shown by the upper three curves in Figure 6(a). We have shown

in Section 2.3 that the Kalman LE can be regarded as a special case of the Kalman DFE by setting the feedback signals to zero. While it is simpler than the Kalman DFE, the Kalman LE also has much worse performance than the Kalman DFE, as can be clearly seen in Figure 3.2. However, we can observe that the SNR gain by using the SIC scheme with a Kalman LE is larger than that with a Kalman DFE.

Figure 3.3 shows the averaged BER performance of the proposed methods in a 3-input/3-output system. Both the Kalman DFE (lower three curves) and the Kalman LE (upper three curves) are evaluated. Except the increased number of transmitters and receivers, all the other settings are the same as Figure 3.2. In this case, since more signals are transmitted at the same time, the ICI is severer than that in a 2-input/2-output system too. Thus there is more room of improvement for the SIC schemes. Compared Figure 3.2 with Figure 3.3, the advantage of using the SIC schemes is more clearly seen in this example. We can observe that the basic SIC scheme achieves about 1 dB SNR gain at the 10^{-3} BER level over the Kalman DFE without any further ICI suppression. The SIC scheme with reliability sorting achieves about 2.5 dB SNR gain as compared to the bare Kalman DFE. When the Kalman LE is used as the first stage, the SNR gain becomes even more obvious. We expect that for systems with large number of transmitter and receiver antennas, the SIC schemes can improve the performance of the DFE substantially.

3.4 Chapter Summary

This chapter investigated the inter-(sub)channel interference suppression problem in MIMO communication systems. Based on a state-space approach, a novel SIC scheme to suppress the ICI was proposed. Combining together the SIC scheme with the Kalman equalizers discussed in the previous chapter, we obtained the optimal realizable ISI and ICI suppression. Simulation results showed the effectiveness of the proposed methods.

We must point out that the MIMO model (2.1) used in Chapter 3 and Chapter 4 is general enough that the multiple antennas systems, and many multiple access systems such as CDMA, OFDM etc., can all be regarded as some special cases of the general MIMO system considered in this paper. On the other hand, our derivation places no restrictions on the channel except that it has a state-space realization. This render the proposed methods applicable to any channel that has a state-space realization, including time-invariant, time-varying, FIR, and IIR channels. Hence the methods proposed in Chapter 3 and Chapter 4 have wide applicability. For example, with some small modifications, the combined receiver can be applied to multiuser detection in uplink CDMA multipath fading channels. It can be applied to remove the Inter-Carrier Interference in OFDM systems too.

Chapter 4
Blind Parameter Estimation in Time-varying Channels

Our discussions in the previous two chapters have assume perfect channel information at the receiver side. However, in any practical communication systems, the channel state information has to be estimated online. This is a challenging problem if the channel experiences both time-selective and frequency-selective fading (i.e. time-varying multipath channel). In this chapter, we will propose a Kalman filtering based (semi-)blind channel tracking and equalization method. By using a linear precoder at the transmitter, *a priori* known correlation is introduced into the transmitted signal. The receiver can thus make use of the known correlation to blindly estimate the channel SOS and no training is necessary. Accurate estimate can be obtained with sufficiently long observation of the received signals. With only a very short training sequence (as short as the channel order) to initialize the Kalman recursions, the proposed method can track the time-varying channel and equalization can be done accordingly.

On the other hand, the time-varying multipath channel provides both temporal and frequency diversities that can be exploited to improve the performance of the system. These diversities can be achieved by using linear precoders too. In [85], it is shown that the temporal diversity of the channel can be achieved by using the so-called spread precoding, where each symbol is transmitted over a period of time that

is much longer than the original symbol duration. In [82], a zero-padding precoder is proved to be able to achieve the full frequency diversity. In [50], linear precoders that achieve both the maximum temporal and frequency diversities are designed. However, all these works have assumed perfect channel information to be available at the receiver side and the channel estimation/tracking problem is not addressed. Therefore, our paper may bridge this gap.

In this chapter, we first establish the identifiability condition of the channel SOS. Then we translate the identifiability condition into the design constraints on the linear precoder. Linear precoders that satisfy the constraints and achieve diversities are constructed. The Kalman filtering based channel tracking and equalization methods are developed. Simulation results show the effectiveness of our proposed method.

This chapter is organized as follows. Section 4.1 presents the signal and channel model and Section 4.2 establishes the identifiability condition of the channel SOS and the corresponding design constraints on the linear precoders. In Section 4.3, the Kalman-filtering based channel tracking and equalization methods are introduced. Section 4.4 presents the simulation results of the proposed method. Section 4.5 concludes this paper.

4.1 System Model

4.1.1 Signal Model

In this chapter, we focus on single-input single-output systems. Instead of directly transmitting the data sequence s_t through the channel, we include a linear precoder at the transmitter side to introduce *a priori* known correlation into the signal.

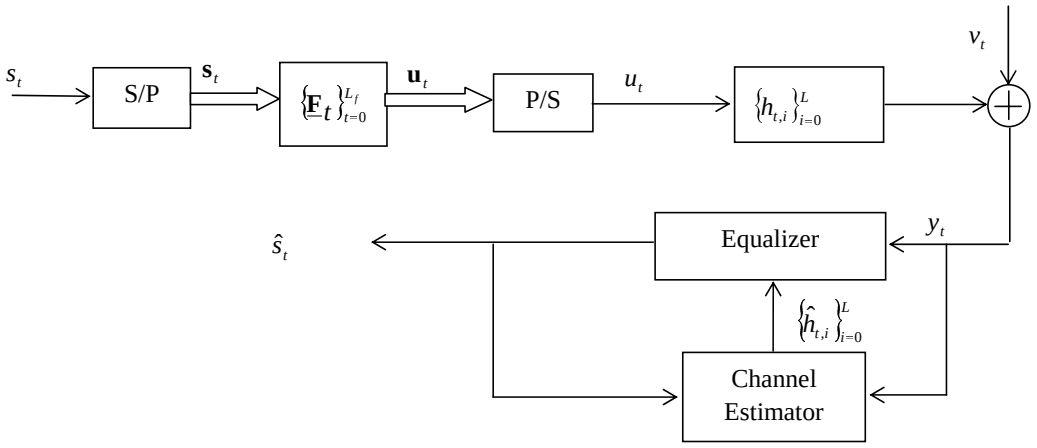

Figure 4.1: System Model.

The system model considered in this chapter is given in Figure 4.1. The data sequence s_t is blocked into an $(M+1)\times 1$ vector $\underline{\mathbf{s}}_t = \left[s_{t(M+1)}\ s_{t(M+1)+1}\ \cdots\ s_{t(M+1)+M}\right]^T$, which is then processed by an L_f-th order linear FIR precoder $\{\underline{\mathbf{F}}_t\}_{t=0}^{t=L_f}$, where $\underline{\mathbf{F}}_t$ is the $P \times M$ precoder coefficient matrix with $P \geq M$. After parallel-to-serial (P/S) conversion, the output of the precoder $\underline{\mathbf{u}}_t$ is transformed back to a scalar process u_t. The precoded signal u_t is transmitted through a linear time-varying multipath chan-

nel with $L+1$ paths. Denote $\{h_{t,i}\}_{i=0}^{L}$ to be the channel impulse response coefficients at time t. We obtain the channel input-output relation

$$\begin{aligned} y_t &= \sum_{i=t-L}^{t} u_i h_{t,t-i} + v_t \\ &= \sum_{i=0}^{L} h_{t,i} u_{t-i} + v_t \end{aligned} \tag{4.1}$$

where v_t is the sampled zero-mean AWGN with covariance given by σ_v^2. We will design equalizer (and "postcoder") to recover the original information sequence s_t based on the received signal y_t.

4.1.2 Channel Model

It is shown in [71], [83], [84] that the time-varying wireless channels can be modeled by low order Gaussian Markov processes. In most cases, even first order models are quite sufficient. Denote $\mathbf{h}_t = \begin{bmatrix} h_{t,0} & h_{t,1} & \cdots & h_{t,L} \end{bmatrix}^T$ to be the channel impulse response coefficients vector at time t. Then we can represent the channel by the following first order AR process,

$$\mathbf{h}_{t+1} = \mathbf{A}\mathbf{h}_t + \mathbf{B}\mathbf{w}_t \tag{4.2}$$

where $\mathbf{w}_t$ is a zero-mean unit-variance complex white Gaussian vector process, $\mathbf{A}$, $\mathbf{B}$ are slowly time-varying matrices that can be regarded constant for very long period of time. From (2.1), the received signal at time t can be written as

$$\begin{aligned} y_t &= \sum_{t=0}^{L} u_{t-i} h_{t,i} + v_t \\ &= \mathbf{r}_t^T \mathbf{h}_t + v_t \end{aligned} \tag{4.3}$$

where $\mathbf{r}_t = \begin{bmatrix} u_t & u_{t-1} & \cdots & u_{n-L} \end{bmatrix}^T$. Equations (4.2) and (4.3) form a state-space representation of the system. We will show later that Kalman filtering can be applied to track the time-varying channel, given $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{r}_t$.

As commonly accepted in wireless communications, we assume the channel coefficients to be wide-sense stationary and uncorrelated scattering (WSSUS)[1], i.e.,

$$\begin{aligned} R_h(k;i,j) &= \mathrm{E}\left[h_{n,i}h^*_{n-k,j}\right] \\ &= R_h(k,i)\delta_{i-j} \end{aligned} \tag{4.4}$$

and each $h_{t,i}$ is a WSS random process with respect to t. For simplicity, we also assume that $\mathrm{E}[h_{t,i}] = 0$ for all i. This corresponds to Rayleigh fading channels. The channel auto-correlation function $R_h(k;i,j)$ is closely related to the parameters $\mathbf{A}$ and $\mathbf{B}$ in the state equation (4.2). Denote $\underline{\mathcal{R}}_h(k) = \mathrm{diag}\left(R_h(k,0),\ R_h(k,1),\ \cdots\ R_h(k,L)\right)$, for $k = 0, 1$. It can be proved that

$$\mathbf{A} = \underline{\mathcal{R}}_h(1)\underline{\mathcal{R}}_h(0)^{-1}, \tag{4.5}$$

$$\mathbf{B}\mathbf{B}^* = \underline{\mathcal{R}}_h(0) - \mathbf{A}\underline{\mathcal{R}}_h^{\mathcal{H}}(1). \tag{4.6}$$

See Appendix 4.6.1 for details. Therefore, once we obtain the channel auto-correlation function, the state equation governing the channel can be uniquely determined.

Without loss of generality, we assume throughout the chapter that the input data sequence s_t is a zero-mean white process with unit variance. Perfect timing and synchronization are also assumed for simplicity. However, our derivations can also be extended to correlated scattering and non-zero-mean (Rician) channels and colored inputs.

[1]It can be shown that at least for full response systems [59], this is a valid assumption.

Remark 4.1 *The state equation in (4.2) describes a variety of channels encountered in mobile communications. When the channel is time-invariant, (4.2) reduces to* $\mathbf{h}_{t+1} = \mathbf{h}_t$*, i.e.,* $\mathbf{A} = \mathbf{I}$ *and* $\mathbf{B} = \mathbf{0}$*. There is no need to estimate the channel SOS. When the time variation of the channel is induced by the receiver carrier frequency shift, (4.2) becomes* $\mathbf{h}_{t+1} = e^{j\theta}\mathbf{h}_t$*, where* θ *is the frequency offset that needs to be estimated and compensated. In this case* $\mathbf{A} = \text{diag}\left(e^{j\theta}, e^{j\theta}, \cdots, e^{j\theta}\right)$ *and* $\mathbf{B} = \mathbf{0}$*. Therefore, our proposed methods is applicable to carrier frequency offset estimation and compensation too.*

4.2 Blind Estimation of Channel Statistics

As shown in the previous section, information of the channel SOS is indispensable for tracking time-varying channels. In this section, we will show that the channel SOS can be estimated blindly from the received signal, given that the input of the channel satisfy some conditions.

4.2.1 Identifiability Conditions

Denote $R_u(t,k) = \text{E}[u_t u_{t-k}^*]$ to be the auto-correlation of the channel input. Due to the precoder, it can be shown that $R_u(t,k)$ is generally cyclostationary[2] with period $P+1$, i.e., $R_u(t,k) = R_u(t+P+1,k)$. See Appendix 4.6.2 for details. The

[2]We regard stationary process as a special case of the cyclostationary process.

auto-correlation of the received signal is given by

$$\begin{aligned} R_y(t,k) &= \mathrm{E}[y_t y_{t-k}^*] \\ &= \mathrm{E}\left[\left\{\sum_{i=0}^{L} h_{t,i}u_{t-i}\right\}\left\{\sum_{i=0}^{L} h_{t-k,i}u_{t-k-i}\right\}^*\right] + \sigma_v^2\delta_k \\ &= \sum_{i=0}^{L} \mathrm{E}\left[h_{t,i}h_{t-k,i}^*\right]\mathrm{E}\left[u_{t-i}u_{t-k-i}^*\right] + \sigma_v^2\delta_k \\ &= \sum_{i=0}^{L} R_h(k,i)R_u(t-i,k) + \sigma_v^2\delta_k \end{aligned} . \tag{4.7}$$

Since $R_u(t,k) = R_u(t+P+1,k)$, it is clear to see that $R_y(t,k) = R_y(t+P+1,k)$ too. Thus the received sequence y_t is also a cyclostationary process with period $P+1$. Therefore we obtain a set of linear equations given by

$$\begin{aligned} R_y(t,k) &= \sum_{i=0}^{L} R_h(k,i)R_u(t-i,k) + \sigma_v^2\delta_k, \\ & \quad t = 0, 1, \cdots, P \end{aligned} . \tag{4.8}$$

Denote

$$\mathbf{R}_{\underline{y}}(t,k) = \begin{bmatrix} R_y(t,k) - \sigma_v^2\delta_k \\ R_y(t+1,k) - \sigma_v^2\delta_k \\ \vdots \\ R_y(t+P,k) - \sigma_v^2\delta_k \end{bmatrix},$$

$$\mathbf{R}_{\underline{h}}(k) = \begin{bmatrix} R_h(k,0) \\ R_h(k,1) \\ \vdots \\ R_h(k,L) \end{bmatrix},$$

$$\underline{\mathbf{R}}_u(t,k) = \begin{bmatrix} R_u(t,k) & R_u(t-1,k) & \cdots & R_u(t-L,k) \\ R_u(t+1,k) & R_u(t,k) & \cdots & R_u(t-L+1,k) \\ \vdots & \ddots & \ddots & \vdots \\ R_u(t+P,k) & R_u(t+P-1,k) & \cdots & R_u(t-L+P,k) \end{bmatrix}. \tag{4.9}$$

Then (4.8) can be conveniently represented in matrix form

$$\mathbf{R}_{\underline{y}}(t,k) = \underline{\mathbf{R}}_u(t,k)\mathbf{R}_{\underline{h}}(k). \tag{4.10}$$

Because (4.10) holds true for all integers t and k, the channel autocorrelation $\mathbf{R}_{\underline{h}}(k)$ can be uniquely determined by

$$\mathbf{R}_{\underline{h}}(k) = \underline{\mathbf{R}}_u(t,k)^+ \mathbf{R}_{\underline{y}}(t,k) \tag{4.11}$$

if and only if the following rank condition is satisfied for all integer k

$$\text{rank}\left(\underline{\mathbf{R}}_u(t,k)\right) = L+1 \tag{4.12}$$

where $^+$ denote pseudo inverse. Condition (4.12) implies that $P \geq L$. Hence $\mathbf{R}_{\underline{u}}(t,k)$ is a square or tall matrix.

In light of the fact that $R_u(t,k) = R_u(t+P+1,k)$, it is easy to prove that $\underline{\mathbf{R}}_u(t,k)$ is a column-wise circulant matrix, which can be diagonalized by the discrete Fourier transform (DFT) matrix. Denote $\mathbf{R}_{\underline{u}}(k) = [R_u(t,k) \;\; R_u(t+1,k) \;\; \cdots \;\; R_u(t+P,k)]^T$. Then the DFT-based diagonalization of $\underline{\mathbf{R}}_u(t,k)$ yields

$$\underline{\mathbf{R}}_u(t,k) = \mathcal{F}^* \text{diag}(\mathcal{F}\mathbf{R}_{\underline{u}}(k))\mathcal{F}[0:L], \tag{4.13}$$

where $\mathcal{F}$ is the $(P+1)\times(P+1)$ discrete Fourier transform (DFT) matrix with the (m,n)th entry $[\mathcal{F}]_{m,n} = \frac{1}{\sqrt{P+1}}\exp(-j2\pi mn/(P+1))$, and $\mathcal{F}[0:L]$ is formed by the first $L+1$ columns of $\mathcal{F}$ [20]. The following results establishes the sufficient and necessary condition for the rank condition as required in equation (4.12).

Lemma 4.1 *Suppose $\mathcal{F}\mathbf{R}_{\underline{u}}(k)$ has $N(k) \leq P+1$ nonzero entries. The channel second order statistics is identifiable from the received signal, irrespective to the underlying channel, if and only if*

$$N(k) \geq L + 1, \tag{4.14}$$

holds true for all integer k.

Proof: Suppose that $\text{rank}\left(\underline{\mathbf{R}}_u(t,k)\right) = L+1$. Since $\mathcal{F}$ is a unitary matrix and is always full rank, we have

$$\begin{aligned} \text{rank}\left(\underline{\mathbf{R}}_u(t,k)\right) &= L+1 \\ &= \text{rank}\left(\mathcal{F}^* \text{diag}\left(\mathcal{F}\mathbf{R}_{\underline{u}}(k)\right)\mathcal{F}[0:L]\right) \\ &= \text{rank}\left(\text{diag}\left(\mathcal{F}\mathbf{R}_{\underline{u}}(k)\right)\mathcal{F}[0:L]\right) \\ &\leq \min\left(N(k), L+1\right) \end{aligned}.$$

Thus $N(k) \geq L+1$.

Conversely, assume that $N(k) \geq L+1$. From the above derivation, we know that $\text{rank}\left(\underline{\mathbf{R}}_u(t,k)\right) \leq L+1$. We need to show that the equality holds. Let $\text{diag}(r_0, r_1, \cdots, r_P) = \text{diag}(\mathcal{F}\mathbf{R}_{\underline{u}}(k))$. Let $\mathcal{F}_i$, $i = 0, 1, \cdots, P$, denote the $(i+1)$th row of $\mathcal{F}[0:L]$. Then $\text{diag}(\mathcal{F}\mathbf{R}_{\underline{u}}(k))\mathcal{F}[0:L] = \left[r_0\mathcal{F}_0^{\mathrm{T}}, r_1\mathcal{F}_1^{\mathrm{T}}, \cdots, r_P\mathcal{F}_P^{\mathrm{T}}\right]^{\mathrm{T}}$. Since $\mathcal{F}_i$'s are the rows of a Vandermond matrix and at least $L+1$ elements of r_i's are non-zeros, there exists at least $L+1$ non-zeros rows in $\text{diag}\left(\mathcal{F}\mathbf{R}_{\underline{u}}(k)\right)\mathcal{F}[0:L]$. Any $L+1$ of these non-zero rows form an $(L+1)\times(L+1)$ submatrix that is full rank. Hence $\text{rank}\left(\underline{\mathbf{R}}_u(t,k)\right) = L+1$. ∎

Lemma 4.1 imposes some constraints on the precoder design. In the following, we will derive the structure of the precoders that can generate the desired channel input u_t.

4.2.2 Constraints on the Linear Precoders

We assume the linear precoder $\underline{\mathbf{F}}_t$ to be FIR of order L_f. From the system model in Figure 4.1, we obtain that

$$\underline{\mathbf{u}}_t = \underline{\mathbf{F}}_t \star \underline{\mathbf{s}}_t = \sum_{i=0}^{L_f} \underline{\mathbf{F}}_i \underline{\mathbf{s}}_{t-i}, \tag{4.15}$$

where $\underline{\mathbf{u}}_t = \begin{bmatrix} u_{t(P+1)} & u_{t(P+1)+1} & \cdots & u_{t(P+1)+P} \end{bmatrix}^T$ and $\underline{\mathbf{s}}_t = \begin{bmatrix} s_{t(M+1)} & s_{t(M+1)+1} & \cdots & s_{t(M+1)+M} \end{bmatrix}^T$ are both block-wise stationary. The corresponding covariance matrix of $\underline{\mathbf{u}}_t$ is given by

$$\begin{aligned} \underline{\mathbf{\Phi}}_u(k) &= \mathrm{E}\left[\underline{\mathbf{u}}_t \underline{\mathbf{u}}_{t-k}^*\right] \\ &= \mathrm{E}\left\{\left[\textstyle\sum_{i=0}^{L_f} \underline{\mathbf{F}}_i \underline{\mathbf{s}}_{t-i}\right]\left[\textstyle\sum_{j=0}^{L_f} \underline{\mathbf{F}}_j \underline{\mathbf{s}}_{t-k-j}\right]^*\right\} \\ &= \textstyle\sum_{i=0}^{L_f} \sum_{j=0}^{L_f} \underline{\mathbf{F}}_i \mathrm{E}\left[\underline{\mathbf{s}}_{t-i} \underline{\mathbf{s}}_{t-k-j}^*\right] \underline{\mathbf{F}}_j^* \\ &= \textstyle\sum_{i=0}^{L_f} \underline{\mathbf{F}}_i \underline{\mathbf{F}}_{i-k}^* \end{aligned} . \tag{4.16}$$

The last equation in (4.16) is obtained since the input signal $\underline{\mathbf{s}}_t$ is white and has unit variance. It is straightforward to prove that the elements of $\underline{\mathbf{\Phi}}_u(k)$ and $\underline{\mathbf{R}}_u(t,k)$ are related by

$$R_u(t,k) = \underline{\Phi}_u(\tau; i, j) \tag{4.17}$$

where $\underline{\Phi}_u(\tau; i, j)$ is the $(i+1, j+1)$th element of $\underline{\mathbf{\Phi}}_u(\tau)$, $i = \mathrm{mod}(t, P+1)$, $j = \mathrm{mod}(i-k, P+1)$ and $\tau = \lfloor \frac{k+j-i}{P+1} \rfloor$ is the largest integer smaller than or equal to $\frac{k+j-i}{P+1}$. Let $P' = P+1$. Thus

$$\underline{\mathbf{\Phi}}_u(k) = \begin{bmatrix} R_u(0, kP') & R_u(0, kP'-1) & \cdots & R_u(0, kP'-P) \\ R_u(1, kP'+1) & R_u(1, kP') & \cdots & R_u(1, kP'-P+1) \\ \vdots & \cdots & \cdots & \vdots \\ R_u(P, kP'+P) & R_u(P, kP'+P-1) & \cdots & R_u(P, kP') \end{bmatrix} \tag{4.18}$$

.

Define the infinite matrix

$$\underline{\mathbf{\Phi}}_u = \left[\begin{array}{cccc} \cdots & \underline{\mathbf{\Phi}}_u(1) & \underline{\mathbf{\Phi}}_u(0) & \underline{\mathbf{\Phi}}_u(-1) \quad \cdots \end{array} \right]. \tag{4.19}$$

Comparing (4.19) with (4.9), we find that $R_{\underline{u}}(t,k)$ is determined solely by the vectors formed by the diagonals of $\underline{\mathbf{\Phi}}_u$. These vectors will be called the diagonal vectors from now on. For time-invariant channels, $\mathbf{R}_{\underline{h}}(k) = \mathbf{R}_{\underline{h}}(0)$, $k = 0, \pm 1, \pm 2, \cdots$. The main diagonal vector of $\underline{\mathbf{\Phi}}_u(0)$ contains all the information necessary to determine the SOS of the channel. We have shown in Section 4.1 that even for time-varying channels, $\mathbf{R}_{\underline{h}}(0)$ along with $\mathbf{R}_{\underline{h}}(1)$ and $\mathbf{R}_{\underline{h}}(-1)$ are sufficient to characterize the dynamics of the channel. Thus we can truncate (4.19) and redefine $\underline{\mathbf{\Phi}}_u$ as

$$\underline{\mathbf{\Phi}}_u = \left[\begin{array}{ccc} \underline{\mathbf{\Phi}}_u(1) & \underline{\mathbf{\Phi}}_u(0) & \underline{\mathbf{\Phi}}_u(-1) \end{array} \right]. \tag{4.20}$$

The identifiability condition in Lemma 4.1 is therefore equivalent to that the following two conditions hold simultaneously:

C1) Denote $\mathrm{diag}\,(\underline{\mathbf{\Phi}}_u(0))$ to be the main diagonal vector of $\mathbf{\Phi}_{\underline{u}}(0)$. Then the DFT of $\mathrm{diag}\,(\underline{\mathbf{\Phi}}_u(0))$ must have at least $L+1$ nonzero entries.

C2) The DFT of the two neighboring diagonal vectors of $\mathrm{diag}\,(\underline{\mathbf{\Phi}}_u(0))$ in $\underline{\mathbf{\Phi}}_u$ must have at least $L+1$ nonzero entries.

From (4.15), we know that $\underline{\mathbf{\Phi}}_u(k)$, $k = -1, 0, +1$, is determined by the precoder $\{\underline{\mathbf{F}}_i\}_{i=0}^{L_f}$. Thus **C1** and **C2** actually serve as the design constraints on the linear precoder. Fortunately, these constraints are not strict at all, as shown in the following theorem.

Theorem 4.1 *Define* $\underline{\mathbf{F}} = \left\{\underline{\mathbf{F}}_t : \underline{\mathbf{F}}_t = \sum_{i=0}^{L_f} \underline{\mathbf{F}}_i \delta_{t-i}\right\}$ *to be the set of all block FIR precoders. Let* $\underline{\mathbf{F}}^{(1)}$ *denote the subset of* $\underline{\mathbf{F}}$ *that contains all the precoders satisfying* **C1** *and* **C2**. *Then* $\underline{\mathbf{F}}^{(1)}$ *is a dense subset of* $\underline{\mathbf{F}}$.

Proof: Let $\underline{\mathbf{F}}^{(2)}$ denote the subset of $\underline{\mathbf{F}}$ that contains all the FIR precoders *not* satisfying **C1** and **C2**. Then $\underline{\mathbf{F}} = \underline{\mathbf{F}}^{(1)} \bigcup \underline{\mathbf{F}}^{(2)}$. It is clear that $\underline{\mathbf{F}}^{(1)}$ is not empty. For any nonzero scalar ξ, if $\underline{\mathbf{F}}_t \in \underline{\mathbf{F}}^{(1)}$, then $\underline{\mathbf{F}}'_t = \xi \underline{\mathbf{F}}_t \in \mathbf{F}^{(1)}$. Now consider two arbitrary points $\mathbf{p}_0 \in \underline{\mathbf{F}}^{(2)}$ and $\mathbf{p}_1 \in \underline{\mathbf{F}}^{(1)}$. For any local neighborhood of $\mathbf{p}_0$, $B_\epsilon(\mathbf{p}_0) = \{\mathbf{p} \in \underline{\mathbf{F}}^{(2)} : \|\mathbf{p} - \mathbf{p}_0\| < \epsilon\}$, we can always find a scalar $\xi < \epsilon / \|\mathbf{p}_1\|$ such that $\mathbf{p}' = \mathbf{p}_0 + \xi \mathbf{p}_1$ falls into $B_\epsilon(\mathbf{p}_0)$, (i.e., $\mathbf{p}' \in B_\epsilon(\mathbf{p}_0)$) and $\mathbf{p}' \in \underline{\mathbf{F}}^{(1)}$. Thus $\mathbf{p}_0$ is an adherence point of $\underline{\mathbf{F}}^{(1)}$. Since $\mathbf{p}_0$ is an arbitrary point in $\underline{\mathbf{F}}^{(2)}$, so all the points in $\underline{\mathbf{F}}^{(2)}$ are adherence points of $\underline{\mathbf{F}}^{(1)}$. Hence $\underline{\mathbf{F}}^{(1)}$ is a dense subset in $\underline{\mathbf{F}}$.

This completes the proof. ∎

Theorem 4.1 assures us that even if the optimal precoder to be designed is ill-conditioned (i.e., **C1** and **C2** are not satisfied), we can always obtain a suboptimal precoder that guarantees channel SOS identifiability without sacrificing much its performance.

4.2.3 Linear Precoder Example

Following the guidelines in Section 4.2.2, we construct linear precoders that guarantee channel SOS identifiability. Theorem 4.1 shows that, in theory, there exists

infinite number of linear precoders that guarantee channel SOS identifiability. A natural question is what kind of precoder has better performance.

It is well known that the time-varying multipath channel provides both temporal and frequency diversities that can be exploited by the receiver to improve performance. In [82], it shows that full frequency diversity can be achieved by simply padding as many as L zeros at the end of each data block, where L is the channel order. In light of **C2**, we also need the adjacent symbols to be correlated. Based on these considerations, a simple linear precoder is as constructed

$$\underline{\mathbf{F}}_i^{(1)} = \begin{cases} \left[\mathbf{F}^T_{(M+1)\times(M+1)}\ \mathbf{0}^T_{L\times(M+1)}\right]^T, & i = 0; \\ \mathbf{0}_{(M+L+1)\times(M+1)}, & i \neq 0. \end{cases} \tag{4.21}$$

where

$$\mathbf{F} = \frac{1}{\lambda}\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha & \beta & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \alpha & \beta & 0 \\ 0 & \cdots & \cdots & \alpha & \beta \end{bmatrix}_{(M+1)\times(M+1)},$$

$\alpha + \beta = 1$, and $\lambda = \|\mathbf{F}\|$ is the matrix norm of $\mathbf{F}$. Division by λ normalizes the power gain of the precoder so that it always has unit power gain.

It is easy to prove that the above precoder does satisfy the design constraints **C1** and **C2**. Due to the zero-padding, frequency diversity is achieved. Moreover, its simple structure renders the corresponding receiver design simplified. This will be clear when we discuss the equalizer design in the next section.

Remark 4.2 *Based on the considerations to guarantee channel SOS identifiability and to achieve channel diversities, we intuitively construct a precoder* $\underline{\mathbf{F}}^{(1)}$*. However, We do not establish the optimality of* $\underline{\mathbf{F}}^{(1)}$*. Better precoders may exist but it is beyond the scope of this dissertation.*

4.3 (Semi-)blind Tracking/Equalization of Time-varying Channels

4.3.1 Time-varying Channels Tracking

Blind estimation of time-invariant channels is a well studied problem. In this section, we focus on the more challenging problem of tracking the time-varying channels. From equations (4.2) and (4.3), we have the following state-space representation of the system

$$\begin{cases} \mathbf{h}_{t+1} = \mathbf{A}\mathbf{h}_t + \mathbf{B}\mathbf{w}_t \\ y_t = \mathbf{r}_t^T\mathbf{h}_t + v_t \end{cases} \tag{4.22}$$

Given $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{r}_t$, Kalman filtering can be employed to obtain the minimum mean squared error (MMSE) estimate of $\mathbf{h}_t$. Due to the iterative structure of the Kalman filter, this channel estimation method is suitable to track even fast fading channels. We list the equations of the Kalman filter recursions below [34]:

$$\begin{aligned} e_t &= y_t - \mathbf{r}_t^T\hat{\mathbf{h}}_t \\ \hat{\mathbf{h}}_{t+1} &= \mathbf{A}\hat{\mathbf{h}}_t + \mathbf{K}_t e_t \\ \mathbf{K}_t &= \left(\mathbf{A}\mathbf{P}_t\mathbf{r}_t^T\right) R_e^{-1}(n) \\ \mathbf{R}_{e,t} &= \sigma_v^2\mathbf{I} + \mathbf{r}_t^T\mathbf{P}_t(\mathbf{r}_t^T)^* \\ \mathbf{P}_{t+1} &= \mathbf{A}\mathbf{P}_t\mathbf{A}^* + \sigma_w^2\mathbf{B}\mathbf{B}^* - \mathbf{K}_t\mathbf{R}_{e,t}\mathbf{K}_{p,t}^* \\ e_0 &= y_0, \quad \hat{\mathbf{h}}_t = 0 \ , \quad \mathbf{P}_0 = \mathbf{h}_t\mathbf{h}_0^* \end{aligned} \quad . \tag{4.23}$$

The above channel estimation process can be implemented in a training-aided/decision-feedback manner [36]. During the training phase, $\mathbf{r}_t$ is exactly known to the receiver. During the data transmission phase, $\mathbf{r}_t$ must be replaced by its estimated value $\hat{\mathbf{r}}_t$, which is formed by the equalizer outputs. The Kalman filter assumes that the equalizer outputs are correct and uses them to estimate the next channel value, whereas the equalizer assumes correct Kalman channel estimate and uses them in turn to equalize the channel. However, the equalizer introduces a detection delay d. To make full use of the multipath diversity, the optimal delay should be equal to or greater than the channel order, i.e., $d \geq L$. This delay causes a *time gap* in the iteration. When $\mathbf{h}_t$ needs to be estimated, what available to the measurement equation is $\mathbf{r}_{t-d}$. To bridge this gap, Kalman prediction instead of Kalman filtering has to be used in the iterations. For channels of large order (L is large), there is a tradeoff between the detection delay and estimation accuracy. Large delay results in better equalization performance but also causes larger channel prediction error, while small delay improves the channel estimation but degrades the equalization performance. A thorough analysis of this tradeoff is not pursued in this dissertation.

4.3.2 Channel Equalization

Due to the effects of the linear precoder, the transmitted signals become "colored". With the channel state information estimated online as shown in (4.23), we can use the techniques discussed in Section 2.4 to construct the optimal realizable equalizer

to obtain the information symbols before precoding. A detailed discussion is similar to that in Section 2.4 and hence is omitted here.

4.4 Simulation Results

In this section, we evaluate the performance of the proposed methods through simulations.

In all the simulations, BPSK signal constellation is used. The precoder size is chosen to be $P = 11$, $M = 9$. The channel order is fixed at $L = 2$. This introduces one step delay in the optimal equalizer output and requires the Kalman filtering to predict two steps ahead of the current time to track the channel. All the channel paths are initialized with zero-mean Gaussian random variables. Matrices $\mathbf{A}$ and $\mathbf{B}$ in (4.2) are chosen to be

$$\mathbf{A} = \begin{pmatrix} 0.996 & 0 & 0 \\ 0 & 0.996 & 0 \\ 0 & 0 & 0.996 \end{pmatrix}, \tag{4.24}$$

$$\mathbf{B} = \begin{pmatrix} 0.08935 & 0 & 0 \\ 0 & 0.08042 & 0 \\ 0 & 0 & 0.07238 \end{pmatrix}. \tag{4.25}$$

This defines an exponentially decaying WSSUS channel that corresponds to a 2.4-GHz transmission with baud rate 40 kHz and Doppler frequency $f_D = 200$ Hz (equivalent to vehicle speed at 90 kM/h or 56 mi/h) for all the three paths [36].

We first evaluate the performance of estimating the state equation parameters $\mathbf{A}$ and $\mathbf{B}$, using the precoder designed in Section 3.3. Figure 2 shows the estimates of

$\mathbf{A}(1,1)$, $\mathbf{A}(2,2)$, and $\mathbf{A}(3,3)$. We observe that the estimates converge to the desired value after about 1500 blocks (each block corresponds to $M+1$ data symbols). We must note that no training symbol is required to obtain these estimates, although the convergence speed is not fast.

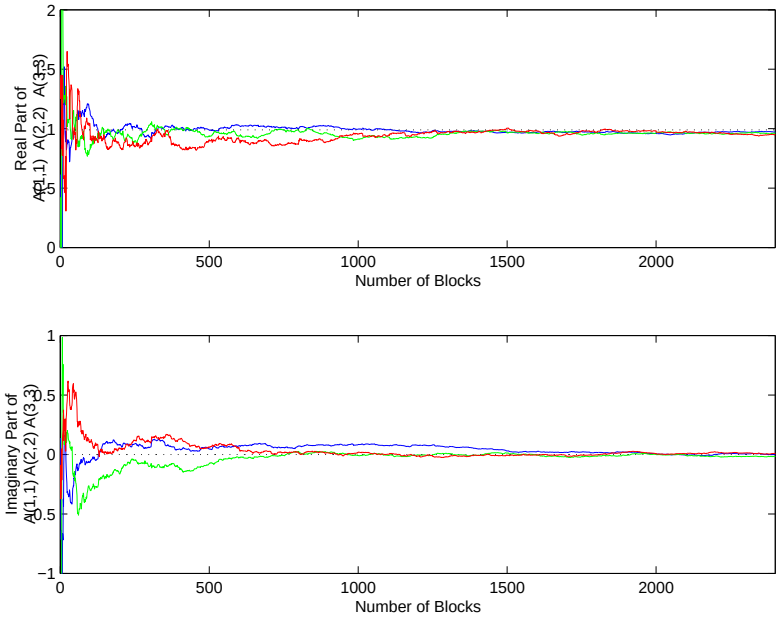

Figure 4.2: Estimates of matrix $\mathbf{A}$. Upper is the real parts of $\mathbf{A}(i,i)$, $i=1,2,3$. Lower is the imaginary parts.

The main advantage of the proposed methods is that only very short training is needed to initialize the channel tracking/equalization process, as compared to other methods where frequent and long training is necessary. The following examples illustrate the effectiveness of the proposed methods. The matrices $\mathbf{A}$ and $\mathbf{B}$ are estimated

after 3000 data blocks to guarantee convergence. (Note that these received data blocks are not training but any signal transmitted from the base station.) Then only *one* training block is inserted to initialize the Kalman channel tracking/equalizatin iterations. Figure 3 shows the channel tracking performance of the Decision-feedback Equalizer when SNR=20 dB. Only the real part of the first channel path is shown here. For the imaginary part and the other two channels, we observe similar performance that is not included here. In Figure 3, we observe that in most cases, the tracking is successful. Only occasionally the estimated channel is very different from the true one. This can be explained by the error propagation of the DFE. However, the DFE manages to recover from the errors quickly. Here, we do not plot the tracking performance of other methods for comparison, because all the other (not blind) methods cannot converge to the true state equation parameters with only one training block ($M+1$ data symbols) and the Kalman filtering iterations cannot be successful.

The overall performance of the proposed methods is evaluated through BER vs. SNR curves. Figure 4 shows BER curves of the LE and the DFE. It is clear to see that with perfect channel information, the DFE outperforms the LE largely. However, with estimated channels, the performance gain of the DFE over the LE is not so much. We also observe that both the DFE and the LE suffer error floors. This can be explained by the error propagation in the tracking/equalization iterations.

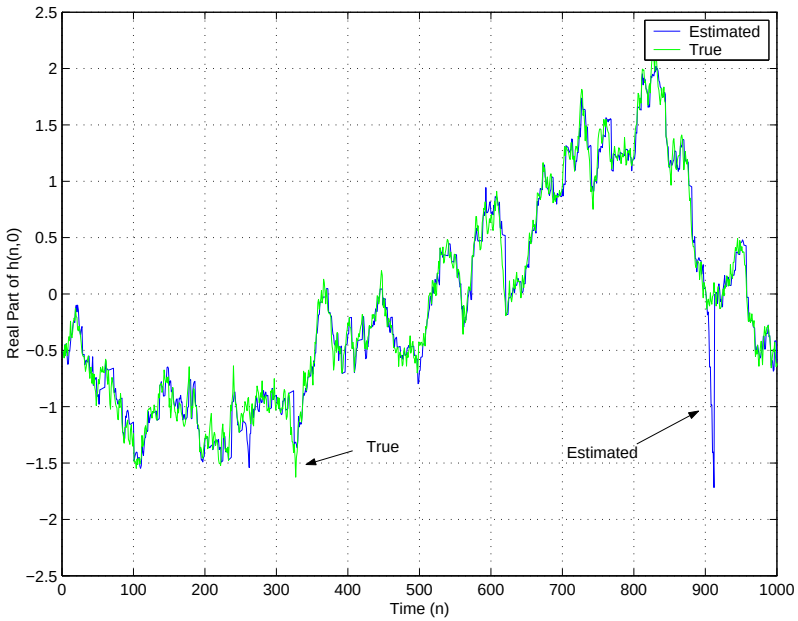

Figure 4.3: Channel tracking with DFE. SNR = 40 dB.

4.5 Chapter Summary

We proposed a semi-blind method for tracking and equalizing time-varying multipath channels. By using a linear precoder at the transmitter to introduce *a priori* known correlation into the transmitted signals, the receiver can estimate the channel second order statistics (SOS) blindly. This allows the proposed channel tracking method to be able to initialize with very short training sequences. This property is extremely useful in applications like wireless broadcasting, where frequent and long training is not possible. We established the channel SOS identifiability condition and the design constraints on the linear precoders. Linear precoders that satisfy

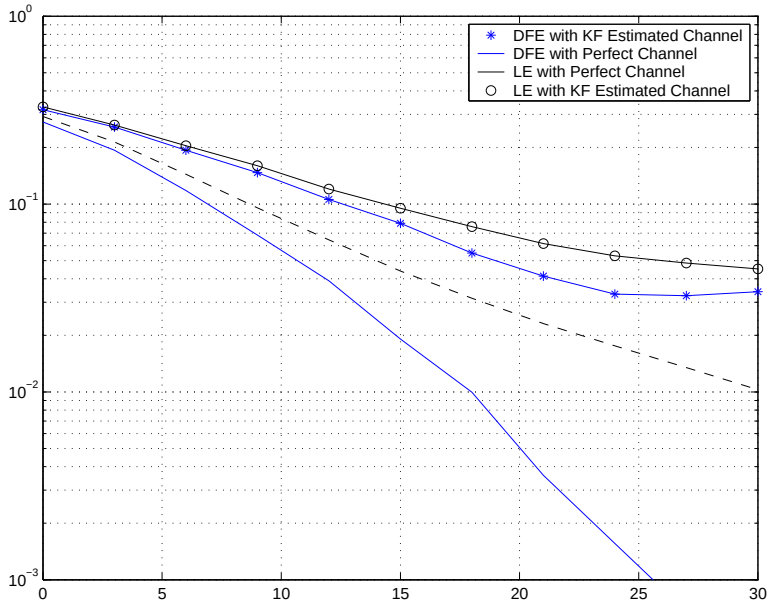

Figure 4.4: BER Performance of the LE and the DFE.

the constraints and achieve the temporal and frequency diversities are constructed. Simulation results show the effectiveness of the proposed method.

4.6 Appendices

4.6.1 Proof of Equations (4.5) and (4.6)

By multiplying both sides of the state equation in (4.2) by $\mathbf{h}_t^*$ and taking expectation, we obtain

$$\mathrm{E}\left[\mathbf{h}_{t+1}\mathbf{h}_t^*\right] = \mathbf{A}\mathrm{E}\left[\mathbf{h}_t\mathbf{h}_t^*\right] + \mathbf{B}\mathrm{E}\left[\mathbf{w}_t\mathbf{h}_t^*\right]. \tag{4.26}$$

Since $\mathbf{w}_t$ is zeros-mean, so $\mathrm{E}\left[\mathbf{w}_t\mathbf{h}_t^*\right] = \mathbf{0}$. Since $\mathbf{h}_t$ is uncorrelated scattering, so $\mathrm{E}\left[\mathbf{h}_t\mathbf{h}_t^*\right] = \underline{\mathcal{R}}_h(t)$. It follows that

$$\underline{\mathcal{R}}_h(1) = \mathbf{A}\underline{\mathcal{R}}_h(0)$$

and

$$\mathbf{A} = \underline{\mathcal{R}}_h(1)\underline{\mathcal{R}}_h^{-1}(0).$$

Similarly, multiplying both sides of the state equation in (4.2) by $\mathbf{h}_{t+1}^*$ and taking expectation, we obtain

$$\mathrm{E}\left[\mathbf{h}_{n+1}\mathbf{h}_{n+1}^*\right] \;=\; \mathbf{A}\mathrm{E}\left[\mathbf{h}_t\mathbf{h}_{t+1}^*\right] + \mathbf{B}\mathrm{E}\left[\mathbf{w}_t\mathbf{w}_t^*\right]\mathbf{B}^* \;.$$

It follows that

$$\underline{\mathcal{R}}_h(0) = \mathbf{A}\underline{\mathcal{R}}_h^*(1) + \mathbf{B}\mathbf{B}^*$$

and

$$\mathbf{B}\mathbf{B}^* = \underline{\mathcal{R}}_h(0) - \mathbf{A}\underline{\mathcal{R}}_h^*(1).$$

4.6.2 Cyclostationarity of Precoder Outputs

Let $\underline{\mathbf{s}}_t = \left[s_{t(M+1)}\ s_{t(M+1)+1}\ \cdots\ s_{t(M+1)+M}\right]^T$ be the blocked input signal. Since s_t is white with unit variance, we have

$$\mathbf{R}_s(k) = \mathrm{E}\left[\underline{\mathbf{s}}_t\underline{\mathbf{s}}_{n-k}^*\right] = \delta_k\mathbf{I}. \tag{4.27}$$

Let $\underline{\mathbf{u}}_t = \left[u_{t(P+1)}\ u_{t(P+1)+1}\ \cdots\ u_{t(P+1)+P}\right]^T$ be the blocked precoder output. From the system model in Figure 4.1, we obtain that

$$\underline{\mathbf{u}}_t = \underline{\mathbf{F}}_t \star \underline{\mathbf{s}}_t = \sum_{i=0}^{L_f} \underline{\mathbf{F}}_t \underline{\mathbf{s}}_{t-i} \tag{4.28}$$

where L_f is the order of the precoder. Let $\underline{\mathbf{f}}_t^{(m)}$, $m = 0, 1, \cdots, P$ be the $(m+1)$th row of $\underline{\mathbf{F}}_t$. Then any precoder output u_ℓ is given by $u_\ell = u_{t(P+1)+m} = \sum_{i=0}^{L_f} \underline{\mathbf{f}}_i^{(m)} \underline{\mathbf{s}}_{n-i}$, where $m = \mathrm{mod}(\ell, P+1)$ and $t = \lfloor \frac{\ell}{P+1} \rfloor$. The auto-correlation of u_ℓ can thus be calculated by

$$\begin{aligned} R_u(\ell, \ell-\ell') &= \mathrm{E}\left[u_\ell u_{\ell'}^*\right] \\ &= \mathrm{E}\left[u_{t(P+1)+m} u_{t'(P+1)+m'}^*\right] \\ &= \sum_{i,j=1}^{L_f} \underline{\mathbf{f}}_i^{(m)} \left[\underline{\mathbf{f}}_j^{(m')}\right]^* \delta_{n-n'-i-j} \\ &= \sum_{i=1}^{L_f} \underline{\mathbf{f}}_i^{(m)} \left[\underline{\mathbf{f}}_{t-t'-i}^{(m')}\right]^* \end{aligned} \tag{4.29}$$

From (4.29), it is clear to see that the auto-correlation of u_ℓ depends on t in general. Since $t = \lfloor \frac{\ell}{P+1} \rfloor$, $R_u(\ell, \ell-\ell')$ depends on ℓ too. Hence u_ℓ is in general not stationary. However, note that $\mathrm{mod}(\ell+P+1, P+1) = \mathrm{mod}(\ell, P+1)$ and $\lfloor \frac{\ell+P+1}{P+1} \rfloor = \lfloor \frac{\ell}{P+1} \rfloor$. Thus $R_u(\ell) = R_u(\ell+P+1)$. Therefore $R_u(\ell)$ is cyclostationary.

Chapter 5
Conclusions

This dissertation provided solutions to two of the most important problems in wireless communication systems design, namely, 1) the interference suppression, and 2) the channel parameter estimation in wireless communication systems over time-varying multipath fading channels.

The interference suppression problem was studied under a unified multirate transmultiplexing model. Employing the optimal estimation theory, we proposed the optimal realizable interference suppression schemes to cancel the ISI and the ICI encountered in various different communication systems. Our approach was based on the state-space description of the communication system and channel. This allowed us to use the famous Kalman filter to obtain the MMSE estimates of the transmitted symbols, and equivalently, to optimally cancel the interferences in the MMSE sense. We proposed the Kalman filter based decision-feedback equalizer and Kalman filter based linear equalizer. The latter can be regarded as a special case of the Kalman DFE. However, it was shown in this dissertation that the Kalman LE has quite different properties as compared to the Kalman DFE. Due to the recursive structure of the Kalman filter, it has an IIR structure in general. We showed that, for the Kalman LE, it does have infinite length. However, different from what normally believed, the

Kalman DFE has an FIR structure. This result thus clarifies the optimality of the conventional finite length DFE.

In addition, we also proposed a novel successive interference cancellation scheme to remove the inter-channel interference encountered in MIMO systems. This method only need a coarse estimate of the strong signal to initialize, and does not require explicit calculation of the ICI or its subspace contributed by each user as in many existing methods. Combined with the proposed optimal ISI suppression schemes, we achieve the optimal realizable interference suppression. Our results are all derived in the time domain, and can be applied to any linear channel that has a state-space realization. To our knowledge, our results are the most general ever obtained.

The problem of time-varying channel tracking is very challenging, especially when the channel also experiences multipath fading. A feasible approach to tackle this problem is to use a low order AR model to characterize the time variation of the channel impulse response. We showed that the AR model parameters can be uniquely determined by the channel second order statistics. We established the channel SOS identifiability condition for SISO communication systems, based on which, we designed linear precoders that render the blind estimation of channel SOS possible. With the AR model estimated, a new channel tracking method was also proposed in this dissertation.

There are still several issues that deserve further research. For example, at each time instance, the proposed Kalman equalizers can estimate not only the desired

symbols, but also all the other symbols contained in the state vector. However, the information of other symbols are never properly used in our interference suppression schemes or channel tracking schemes. We expect that if this information is employed, the performance of combined interference suppression and channel tracking scheme can be improved. Unfortunately, it is challenging to make use of the information of other symbols contained in the state vector because the symbols before the desired symbol are normally much unreliable than the desired one. Directly use these symbols for channel estimation may cause the error propagation problem, which will further deteriorate the symbol detection quality and cause poor channel estimation. One possible solution to this challenge may be the constant gain channel estimator that was proposed and analyzed in [42], [43], and [68]. This channel estimator is a constant filter that is not determined by the detected symbols. Hence it avoids the error propagation problem. Further study on how to combine this constant gain channel estimator with our proposed optimal realizable interference suppressor is definitely worth more exploration.

We have also mentioned that our proposed interference suppression method can be applied to many different communication systems including OFDM, while our proposed channel parameter estimation method can be applied to blindly estimate the frequency offset between the transmitter and the receiver. Note that frequency offset is also one of the causes of the ICI problem in an OFDM system. Hence both the interference suppression scheme and the channel parameter estimation scheme

propose in this dissertation can be applied to ICI cancellation in OFDM systems. However, a detailed implementation strategy still needs investigation.

References

[1] N. Al-Dhahir and J. Cioffi, "MMSE decision feedback equalizers and coding: Finite-Length Results," *IEEE Transactions on Information Theory*, pp. 961-976, July 1995.

[2] N. Al-Dhahir A. H. Sayed, and J. Cioffi, "Stable pole-zero modeling of long FIR filters with application to the MMSE-DFE," *IEEE Transactions on Communications*, pp. 508-513, Vol. 45, No. 5, May, 1997.

[3] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE DFE," *IEEE Transactions on Signal Processing*, Vol. 48, pp. 2921-2936, Oct 2000.

[4] S. A. Altekar, and N. C. Beaulieu, "Upper bounds on the error probability of decision feedback equalization," *IEEE Transactions on Information Theory*, vol. 39, pp. 145-156, Jan 1993.

[5] S. L. Ariyavisitakul, J. H. Winters, and I. Lee, "Optimum space-time processors with dispersive interference – unified analysis and required filter span," *IEEE Transactions on Communications*, Vol. 47, No. 7, pp. 1073-1083, 1999

[6] M. E. Austin, "Decision-feedback equalization for digital communication over dispersive channels," *MIT Lincoln Laboratory, Lexington, MA. Tech Report No. 437*, Aug. 1967.

[7] K. E. Baddour, and N. C. Beaulieu, "Autoregressive models for fading channel simulation," in *Proc. IEEE Global Telecommun. Conf.*, vol. 2, 2001, pp. 1187-1192.

[8] Ehab F. Badran, "Optimal Channel Equalization for filterbank transeivers in presence of white noise," *Ph.D. Dissertation*, Dept. of Electrical Engineering, Louisiana State University, 2002.

[9] C. A. Belfiore and J. H. Park, Jr, "Decision feedback equalization," *Proceedings of IEEE*, Vol. 67, pp. 1143-1156, Aug. 1979.

[10] P. Banaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio – part I: theoretical considerations," *IEEE Transactions on Communications*, Vol. 40, pp. 885-894, 1992

[11] P. Banaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio – part II: numerical results," *IEEE Transactions on Communications*, Vol. 40, pp. 895-907, 1992

[12] R. A. Casas, Thomas J. Endres, Azzedine Touzni, Richard Johnson, Jr. and John R. Treichler, "Current approaches to blind decision feedback equalization," in Ch. 11 of *Signal Processing Advances in Wireless and Mobile Communications*, Prentice Hall, New Jersey, 2001.

[13] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decisioin-feedback equalizers and coding — part I: equalization results," *IEEE Transactions on Communications*, Vol. 43, No. 10, pp. 2582-2594, October 1995.

[14] P. Crespo and M. Honig, "Pole-zero decision feedback equalization with a rapidly converging adaptive IIR algorithm," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 817C829, Aug. 1991.

[15] Zhi Ding, Ye (Geoffery) Li, *Blind Equalization and Identification*, Marcel Dekker, January 15, 2001

[16] A. Duel-Hallen, "Equalizers for multiple input / multiple output channels and PAM systems with cyclostationary input sequences," *IEEE Journal on Selected Areas in Communications*, Vol. 10, No. 3, pp. 630-639, April 1992.

[17] —— , "A family of multiuser decision-feedback detectors for asynchronous code division multiple accuess channels," *IEEE Transactions on Communications*, vol. 43, pp. 421-434, Feb/Mar/Apr 1995.

[18] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Informaiton Theory*, Vol. IT-18, No. 3, pp. 363-378, May 1972.

[19] G. J. Foschini, Jr., and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications* Vol. 6 , No. 3, pp. 311-335, March 1998.

[20] Gene H. Golub and Charles F. Van Loan, *Matrix Computations (3rd Edition)*, The John Hopkons University Press, London, 1996.

[21] D. A. George, R. R. Brown, and J. R. Storey, "An adaptive decision-feedback equalizer," *IEEE Transactions on Communications Technology*, vol. 19, 281-293, June 1971.

[22] G. B. Giannakis, Y. Hua, P. Stoica and L. Tong (Editors), *Signal Processing Advances in Wireless and Mobile Communications - Volume I, Trends in Channel Estimation and Equalization*, Prentice-Hall, September 2000.

[23] G. Gu, J. He, "MMSE multiuser detection for CDMA data networks", *WCNC 2003 - IEEE Wireless Communications and Networking Conference*, vol. 4, no. 1, Mar 2003 pp. 332-337

[24] G. Gu, J. He, X. Gao, and M. Naraghi-Pour, "An analytic approach to modeling and estimation of OFDM channels," *Proceedings of IEEE Global Telecommunications Conference, 2004. GLOBECOM '04* pp 2381-2386

[25] D. Guo, X. Wang, "Blind detection in MIMO systems via sequential Monte Carlo," *IEEE Journal on Selected Areas in Communications*, Vol. 21, No 3, pages 464- 473, Apr 2003.

[26] B. Hassibi and B. M. Hochwald, "How much training is needed in a multiple-antenna wireless link," *IEEE Transactions on Information Theory*, vol.49, no.4, pages 951-964, Apr. 2003.

[27] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Transactions on Information Theory*, Vol.41, No.4, pp.944-960, 1995.

[28] J. He, "Multiple Access Interference Suppression in CDMA Wireless Systems," *Master Thesis*, Dept. of Electrical Engineering, Louisiana State University, 2001

[29] J. He, Z. Wu, and G. Gu, " Semi-blind tracking and equalization of time-varying multipath channels," *Proceedings of IEEE Wireless Communications and Networking Conference, 2004*, WCNC 2004, Vol. 4 , pp. 2444 - 2449, 21-25 March 2004.

[30] J. He, Z. Wu, and G. Gu, "Optimal realizable linear and decision feedback equalizers: time domain results," *Proceedings of IEEE Wireless Communications and Networking Conference, 2005*, WCNC 2005, New Orleans, LA, March 2004.

[31] R. A. Iltis and S. Kim, "Geometric derivation of expectation maximization and generalized successive interference cancellation algorithms with CDMA channel estimation," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1367-1377, May 2003.

[32] W. C. Jakes, *Microwave Mobile Communications*, John Wiley and Sons, New York, 1974.

[33] A. Johansson and A. Svensson, "Successive interference cancellation in multiple data rate DS/CDMA systems," in *Proc. IEEE VTC95*, Chicago, IL, July 26-28, 1995.

[34] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, New Jersey, 2000.

[35] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME – Journal of Basic Engineering*, vol. 82, series D, pp. 35-45, 1960.

[36] C. Konminakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Transactions on Signal Processing*, Vol. 50, No. 5, pp. 1065-1076, May 2002.

[37] R. E. Lawrence and H. Kaufman, "The Kalman filter for the equalization of a digital communicaitons channel," *IEEE Transactions on Communication Technology*, Vol. 19, No. 6, pp. 1137-1141, December 1971.

[38] Y. (Geoffrey) Li and Z. Ding, "A simplified approach to optimum diversity combining and equalization in digital data transmission," *IEEE Transactions on Communications*, Vol. 43, pp. 47-53, 1995

[39] Y. (Geoffrey) Li and K. J. R. Liu, "Adaptive blind multi-channel equalization for multiple signal separation," *IEEE Transactions on Information Theory*, vol. 44, pp. 2864-2876, November 1998.

[40] Tao Li, and N.D. Sidiropoulos, "Blind digital signal separation using successive interference cancellation iterative least squares," *IEEE Transactions on Signal Processing*, Vol. 48 No. 11, pp. 3146-3152, Nov. 2000.

[41] J. Liang, Z. Ding, "Blind MIMO system identification based on cumulant subspace decomposition Signal Processing," *IEEE Transactions on Signal Processing*, Vol. 51, No. 6, pages 1457- 1468, June 2003.

[42] L. Lindbom, A. Ahlen, M. Sternad, and M. Falkenstrm, "Tracking of time-varying mobile radio channels, part II: a case study," *IEEE Transactions on Communications*, vol. 50 January 2002, pp. 156-167.

[43] L. Lindbom , M. Sternad, and A. Ahlen, "Tracking of time-varying mobile radio channels, part I: the Wiener LMS algorithm," *IEEE Transactions on Communications*, vol. 49 December 2001, pp. 2207-2217.

[44] Zhiqiang Liu, Xiaoli Ma and Georgios B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Tran. on Communications*, vol. 50, No. 2, pp. 183-186, Feb., 2002.

[45] R. López-Valcarce, "Realizable linear and decision feedback equalizers: properties and connections," *IEEE Transactions on Signal Processing*, Vol. 52, No. 3 pp. 757-773, March 2004.

[46] R. W. S. Lucky, "Automatic equalization for digital communications",*Bell Systems Tech. Journal*, vol. 44, pp. 547-588, 1966.

[47] R. W. S. Lucky, "Techniques for adative equalization of digital communications," *Bell Systems Tech. Journal*, vol. 45, pp. 255-286, 1966.

[48] R. Lupas, and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, Vol.35, No.1, pp.123-136, 1989.

[49] R. Lupas, and S. Verdu, "Near-far resistance of multiuser detectors in asynchornous channels," *IEEE Transaction on Communications*, Vol.38, No.4, pp.496-508, 1990.

[50] Xiaoli Ma and Georgios B. Giannakis, "Maximum-diversity transmisssion over doubly-selective wireless channels," *IEEE Tran. on Information Theory*, vol. 49, No. 7, pp. 1832-1840, July, 2003.

[51] Yao Ma, and Teng Joon Lim, "Linear and nonlinear chip-rate minimum mean-squared-error multiuser CDMA detection," *IEEE Transactions on Communications*, vol. 49, no. 3, March 2001 pp. 530-542

[52] M. F. Madhour, S. C. Gupta, and Y. E. Wang, "Successive interference cancellation algorithms for downlink W-CDMA communications," *IEEE Transactions on Wireless Communications*, Vol. 1. No. 1, Jan 2002.

[53] U. Madhow, and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Transactions on Communications*, Vol.42, No.12, pp.3178-3188, 1994.

[54] B. Mulgrew and C. F. N. Cowan, "An adaptive Kalman equalizer: structure and performance," *IEEE Transactions on Acoustics, Speech, and Singal Processing*, Vol. 35, No. 12, pp. 1727-1735, December 1987.

[55] A. W. Naylor, and G. R. Sell, *Linear Operator Thoery in Engineering and Science (Second Edtion)*, Springer, 1982

[56] R. Price, "Nonlinear feedback equalized PAM versus capacity for noisy filter channels," in *Proc. International Confernece on Communications*, 1972, pp. 22.12C22.17

[57] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in DS/CDMA system," *IEEE Journal on Selected Areas in Communications*, Vol. 12, pp. 796-807, June 1994.

[58] R. Price, "Nonlinearly feedback-equalized PAM vs. capacity," *Proc. IEEE International Conferences on Communications*, Philadephia, Penn, 1972.

[59] J. G. Proakia, *Digital Communications (Fourth Edition)*, Prentice Hall, New Jersey, 2000

[60] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, Inc., New Jersey, 1996

[61] S. Roy, "Optimum infinite-length MMSE multi-user decision-feedback space-time processing in broadband cellular radio," *Wireless Personal Communications*, Vol. 27, No. 1, pp. 1-32, Oct. 2003.

[62] J. Salz, "Optimum mean-square decision-feedback equalization," *Bell System Tech. Journal*, vol. 52, pp. 1341-1373, October 1973.

[63] D. Samardzija, N. Mandayam, I. Seskar, "Blind successive interference cancellation for DS-CDMA systems," *IEEE Transactions on Communications*, vol. 50, No. 2, pp. 276-290, February 2002.

[64] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers, part I: unification and optimal designs," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1988-2006, July 1999.

[65] J. E. Smee, N. C. Beaulieu, "On the equivalence of the simultaneous and separate MMSE optimizations of a DFE FFF and FBF," *IEEE Transactions on Communications*, Vol. 45, pp. 156-158, February 1997.

[66] A. Stamoulis, G. B. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 69-83, Jan 2001.

[67] M. Sternad and A. Ahlén, "The structure and design of realizable decision feedback eqalizers for IIR channels with colored noise," *IEEE Transactions on Information Theory*, Vol. 36, pp. 848-858, July 1990.

[68] M. Sternad, L. Lindbom, and A. Ahlen, "Wiener design of adaptation algorithms with time-invariant gains," *IEEE Transactions on Signal Processing*, vol. 50, August 2002, pp 1895-1907.

[69] Gordon L. Stuber, *Principles of Mobile Communication (2nd Edition)*, Kluwer Academic Publishers, February 1, 2001

[70] S. Talwar, M. Viberg, and A. Paulraj,"Blind separation of synchronous co-channel digital signals using an antenna array — Part I: Algorithms," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1184-1197, May 1996.

[71] Christopher C. Tan and Norman Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Tran. on Communications*, vol. 48, No. 12, pp. 2032-2040, December 2000.

[72] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construcation," *IEEE Transactions on Information Theory*, Vol. 44, No. 2, pp. 744-765 March 1998.

[73] C. Tidestav, A. Ahlen, and M. Sternad, "Realizable MIMO decision feedback equalizers: structure and design," *IEEE Transactions. on Signal Processing*, Vol. 49, No. 1, pp. 121 - 133, Jan. 2001.

[74] Michail K. Tsatsannis, Georgios B. Giannakis and Guotong Zhou,"Estimation and equalization of fading channels with random coefficients," *Signal Processing*, Vol. 53, pp. 211-229, 1996.

[75] Michail K. Tsatsannis, and Georgios B. Giannakis , "Optimal linear receivers for DS-CDMA systems: a signal processing approach," *IEEE Transactions on Signal Processing*, Vol. 44, pp. 3044-3055, 1996.

[76] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T-Bell Lab Internal Tech. Memo.*, June 1995.

[77] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[78] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Transactions on Information Theory*, Vol.IT-32, No.1, 1986.

[79] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998

[80] Andrew J. Viterbi, *CDMA: principles of spread spectrum communication*, Addison Wesley Longman Publishing Co., Inc. Redwood City, CA, USA, 1995

[81] P. A. Voois, I. Lee, and J. M. Cioffi, "The effect of decision delay in finite-length decision feedback equalization," *IEEE Transactions on Information Theory*, Vol. 42, No. 2, pp. 618-621, March 1996.

[82] Zhengdao Wang and Georgios B. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Tran. on Information Theory*, vol. 49, No. 3, March, 2003

[83] Hong Shen Wang and Pai-Chi Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Tran. on Vehicular Technology*, vol. 45, No. 2, pp. 353-357, May 1996

[84] Hong Shen Wang and N. Moayeri, "Finite-state Markov channel – a useful model for radio communication channels," *IEEE Trans. Vehicular Technology*, vol. 44, pp. 163-171, Feb. 1995

[85] Gregory W. Wornell, "Spread-response precoding for communications over fading channels," *IEEE Tran. on Information Theory*, vol. 42, No. 2, pp. 488-501, March, 1996.

[86] G. W. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proc. IEEE*, vol. 84, pp. 586C603, Apr. 1996

[87] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Transactions on Information Theory*, vol. 40, pp. 1334-1347, Sept. 1994.

[88] C. Zhang, and R. R. Bitmead, "State Space Modeling for MIMO Wireless Channels," *Preprint*.

Vita

Jianqiang He was born in Chengdu, Sichuan, China, on May 6, 1976. He received Bachelor of Science degree in Electrical Engineering from Shenyang Institute of Aeronautical Engineering in July 1998. From September 1998 to June 2000, he was a graduate student in the Department of Electrical Engineering at Beijing University of Aeronautics and Astronautics. He entered the graduate program in the Department of Electrical and Computer Engineering at Louisiana State University in the fall of 2000. Now he is a candidate for the degree of Doctor of Philosophy in electrical engineering. He was awarded the Excellent Graduate of Liaoning Province in 1998.