FREQUENCY-DOMAIN EQUALIZATION OF SINGLE CARRIER TRANSMISSIONS OVER DOUBLY SELECTIVE CHANNELS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the

Graduate School of The Ohio State University

By

Hong Liu, B.Sc., M.Sc.

* * * * *

The Ohio State University

2007

Dissertation Committee:

Prof. Philip Schniter, Adviser

Prof. Hesham El Gamal

Prof. Randolph L. Moses

Approved by

Adviser

Graduate Program in
Electrical and Computer
Engineering

ABSTRACT

Wireless communication systems targeting at broadband and mobile transmissions commonly face the challenge of fading channels that are both time and frequency selective. Therefore, design of effective equalization and estimation algorithms for such channels becomes a fundamental problem. Although multi-carrier transmissions demonstrate prominent potential to combat doubly selective fading, several factors may retard their applications, such as: high peak-to-average power ratio, sensitivity to phase noise, etc. Meanwhile, single-carrier transmission is a conventional approach and has important applications, such as HDTV broadcasting, underwater acoustic communication. In this dissertation, we focus on receiver design for single-carrier transmissions. Our goal is to design and develop a group of channel estimation and equalization algorithms in the frequency-domain, which enable high performance and low complexity reception of single-carrier transmissions through doubly selective channels.

For single-carrier transmissions over moderately fast fading channels with long-delay spread, we present an improved iterative frequency-domain equalization (IFDE) algorithm based on soft-interference-cancellation (SIC) and propose a novel adaptive frequency-domain channel estimation (AFDCE) based on soft-input Kalman filter, where soft information feedback from the IFDE can be exploited in the channel estimator. Simulation results show that, compared to other existing schemes, the

proposed scheme offers lower MSE in channel prediction, lower BER after decoding, and robustness to non-stationary channels.

We extend the IFDE/AFDCE scheme to accommodate the application of digital television (DTV) signal reception. Compared with the traditional joint decision feedback equalization (DFE) /decoding plus frequency-domain least-mean-square (FDLMS) channel estimation approach, the proposed scheme achieves better performance at a fraction of the implementation cost.

For very fast fading large-delay-spread channels, traditional FDE methods fail, because channel variation within a FFT block induces significant off-main-diagonal coefficients in the frequency domain. To conquer the problem, we apply Doppler channel shortening to shape the energy distribution of those coefficients and derive a pilot-aided MMSE estimator to estimate them for SIC. We also propose a novel IFDE by leveraging both the sparse structure of shortened channel and finite-alphabet property of transmitted symbols. Numerical results show that the proposed scheme has advantages over the well-known FIR-MMSE-DFE/RLS-CE scheme in both performance and complexity.

To my family

ACKNOWLEDGMENTS

First of all, I would like to thank my advisor, Prof. Phil Schniter, for his valuable advice, guidance and help throughout this academic experience. Without his critical feedback on my research and technique writing, it would not be possible for me to accomplish this work. I thank Prof. Hesham El Gamal for his valuable comments and his Wireless Communication class and paper discussion seminars motivated me to always keep a open mind towards research. I also thank Prof. Randolph L. Moses and Prof. Andrea Serrani for serving in my candidacy and dissertation committees and their constructive feedback.

I am grateful to the Department of ECE for the financial support and National Science Foundation for its funding support under Grant No. 0237037. I am also grateful to ATI Research Inc. for offering me a chance to turn my research into practice. I particularly thank Dr. Raul. A. Casas for mentoring and supporting me throughout the internship and serving in my final defense committee. Without his inspirations, enthusiasm and strong support, this dissertation couldn't have been completed. I also thank Dr. Troy Schaffer, Haosong Fu and all the other members in the system group, for their collaboration and help, which makes my internship a wonderful experience.

I thank my labmates in the IPS lab for the stimulating discussions, and for providing the fun environment in which we learn and grow. I thank Ms. Jeri McMichael,

IPS administrative assistant, for her efforts to make IPS lab a big happy family. I am thankful to all my good friends - that I met in America and in China. Wherever you are, without your moral support, I could not go through those difficult moments and achieve the goal.

Finally, I would like to thank my family for their unconditional love and support throughout my life. Especially, I want to thank my husband Jun, for his endless patience and encouragement. Both as a colleague and my dearest friend, he accompanies me throughout this journey all along.

VITA

February, 1976	Born - Wuhan, P. R. China
Sept. 1994 - June 1998	B.Sc. Electrical Engineering, Wuhan University, Wuhan, P. R. China
Sept. 1998 - June 2001	M.Sc. Electrical Engineering, Wuhan University, Wuhan, P. R. China
Sept. 2002 - Aug. 2003	Graduate Fellow, Dept. of Elec. & Computer Eng. The Ohio State University Columbus, OH.
Sept. 2003 - Mar. 2004 July 2004 - Sept. 2004	Graduate Research Associate, Dept. of Elec. & Computer Eng. The Ohio State University Columbus, OH.
Apr. 2004 - June 2004 Sept. 2004- June 2005	Graduate Teaching Associate, Dept. of Elec. & Computer Eng. The Ohio State University Columbus, OH.
July 2005 - Dec. 2005	System Engineer Intern, ATI Research, Inc. Yardley, PA.
Jan. 2006 - June 2006	Graduate Teaching Associate, Dept. of Elec. & Computer Eng. The Ohio State University Columbus, OH.

June 2006 - June 2007 ..... Graduate Fellow,
Dept. of Elec. & Computer Eng.
The Ohio State University
Columbus, OH.

PUBLICATIONS

Research Publications

1. H. Liu and P. Schniter, "Iterative Frequency-Domain Channel Estimation and Equalization for Single-Carrier Transmissions without Cyclic-Prefix," *IEEE Transactions on Wireless Communications*, submitted Apr. 2007.

2. P. Schniter and H. Liu, "Iterative Frequency-Domain Equalization of Single-Carrier Transmissions over Doubly Dispersive Channels," *IEEE Transactions on Signal Processing*, under revision.

3. H. Liu and P. Schniter, "Iterative Frequency-Domain Channel Estimation and Equalization for Single-Carrier Transmission without Cyclic Prefix," *Proc. Conference on Information Sciences and Systems*, (Baltimore, MD), Mar. 2007.

4. H. Liu, P. Schniter, H. Fu, and R. A. Casas, "Frequency Domain Turbo Equalization for Vestigial Sideband Modulation with Punctured Trellis Coding," *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, (Cannes, France), July 2006.

5. P. Schniter and H. Liu, "Iterative Frequency-Domain Equalization for Single-Carrier Systems in Doubly Dispersive Channels," *Proc. Asilomar Conf. on Signals, Systems, and Computers*, (Pacific Grove, CA), pp. 667-671, Nov. 2004.

6. P. Schniter and H. Liu, "Iterative Equalization for Single-Carrier Cyclic-Prefix in Doubly Dispersive Channels," *Proc. Asilomar Conf. on Signals, Systems, and Computers*, (Pacific Grove, CA), vol. 1, pp. 502-506, Nov. 2003.

PATENT: H. Liu, R. A. Casas, H. Fu, "A frequency-domain Turbo Equalizer for DTV signals," filed January 2006.

FIELDS OF STUDY

Major Field: Electrical and Computer Engineering

Studies in:

Comm. and Signal Proc.	Prof. Philip Schniter
Comm. and Signal Proc.	Prof. Hesham El Gamal
Comm. and Signal Proc.	Prof. Randolph L. Moses
Control Theory	Prof. Vadim Utkin

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

In mobile wireless and digital television (DTV) transmission, time-varying multipath phenomenon is generally induced by the randomly changing propagation characteristics as well as the reflection, diffraction and scattering of the transmitted signals from the buildings, large moving vehicles, mountains, etc. Such phenomenon distorts received signals and poses critical challenges in the design of communication systems for high-rate and high-mobility wireless communication applications. High rate information symbols, after transmitting through multipath channel, often spread into neighboring symbol periods, and cause serious inter-symbol interference (ISI) at the receiver side. In addition, relative mobility between the transmitter and receiver leading to fast channel variations, along with oscillator drifts and phase noise, gives rise to time selectivity. The combined time-frequency selectivity induces Doppler-delay spreading, which significantly affects communication system performance. Therefore, the design of effective equalization and estimation algorithms for such channels becomes a fundamental problem of communication systems.

In order to implement commercially competitive communication systems, low-complexity and low-cost systems are highly desirable. Among various proposed candidates for the new system design, the diversity reception with multiple transmitter and receiver antennas [3, 4] and the multi-carrier transmission [5] combined with advanced signal processing algorithms to estimate and equalize the dynamic channels are considered to be the most promising. However, the introduction of multiple antennas demands dedicated amplifiers in all configurations. Multi-carrier transmission exhibits very high peak-to-average-power ratio (PAPR) and utilizes a combination of highly linear power amplifiers, amplitude clipping and amplifier backoff to mitigate the problem [6]. Since a big portion of the cost of terminals in communication systems is due to the transmitter power amplifier, single-carrier (SC) modulation system is a favorable alternative for commercial success. In addition, in some applications such as HDTV transmission, the transmitter is standardized to adopt SC modulation, which also motivates the receiver design for SC transmission systems.

This dissertation considers receiver design for effective and efficient reception of single-carrier transmission through time-varying multipath channels. Our goal is to design and develop a group of channel estimation and equalization algorithms in the frequency domain, which enable high performance reception of SC transmission with low computational complexity.

1.2 Background

1.2.1 Doubly Selective Channels

Wireless communications operate through electromagnetic radiation from the transmitter to the receiver. The communication medium, commonly referred as the channel, usually distorts the signal based on its propagation characteristics. Two important factors which characterize the distortion effects of the channel are multipath fading and Doppler effect. Multipath fading is the phenomenon in which the transmitted signal arrives at the receiver via multiple propagation paths at different delays due to reflection, diffraction and scattering of the radio waves. It results in a wide variation of the received signal strength, since the multiple signals arriving at the receiver may add up constructively or destructively. The Doppler effect, named after Christian Doppler, is the change in frequency and wavelength of a wave that is perceived by an observer moving relative to the source of the waves [7]. In mobile wireless communication scenario, Doppler effect is attributed to the relative movement of the surrounding objects as well as the transmitter and receiver. It leads to fast phase oscillation of the received signals on multiple paths, thus accelerates the time variation of the channel distortion.

Future wireless communication services featuring high-data-rate and high-mobility can aggravate the multipath and Doppler effect. In digital communication systems, for most of the channels, the discrete information bearing symbols are modulated with a continuous pulse shape and transmitted across the channel [8]. In most cases, the pulse shapes are localized in time and frequency so that transmission of each symbol consumes a small tile in the time-frequency plane. For high data rate transmission, the duration of the pulse becomes small and comparable to the multipath delay, thus

ISI occurs and the channel distortion is called frequency-selective. In high-mobility scenarios, the channel response varies significantly in the signaling duration due to Doppler effect, thus the channel distortion becomes time-selective within a single processing block. Channels whose response are both time and frequency selective are commonly referred as doubly-selective channels.

Theoretically, the doubly selective channel can be modeled as a linear time-varying system [9]. When the surrounding objects are stationary, the input and output relationship between transmitter and receiver can be represented as a linear time-invariant system with the impulse response

$$c(\tau) = \sum_{\ell=1}^{N_\ell} c_\ell \delta(\tau - \tau_\ell), \tag{1.1}$$

where c_ℓ and τ_ℓ are the attenuation and propagation delay of the ℓ-th path respectively. This model is widely adopted for description of multipath frequency-selective channel. When there is relative movement between the surrounding objects including transmitter and receiver, the attenuation and delay of the ℓ-th path vary with time. Therefore the impulse response of the channel becomes

$$c(t, \tau) = \sum_{\ell=1}^{N_\ell} c_\ell(t) \delta(\tau - \tau_\ell(t)). \tag{1.2}$$

This is the continuous time model for a doubly selective channel.

Doppler spread and delay spread are two important quantities that measure the time selectivity and frequency selectivity of the channel respectively. The Doppler shift of the ℓ-th path is defined as $f_c \frac{d\tau_l(t)}{dt}$, where f_c is the carrier frequency. The Doppler spread f_d is defined as the largest difference between the Doppler shift of all paths.

$$f_d = \max_{i,j} f_c \left| \frac{d\tau_i(t)}{dt} - \frac{d\tau_j(t)}{dt} \right| \tag{1.3}$$

Larger f_d implies that the channel varies more rapidly in time. The delay spread (or multipath spread) is defined as the difference in the propagation time between the longest and shortest path. Thus,

$$T_d := \max_{i,j} |\tau_i(t) - \tau_j(t)| . \tag{1.4}$$

When T_d is larger, the multipath effect is more evident.

1.2.2 System Model

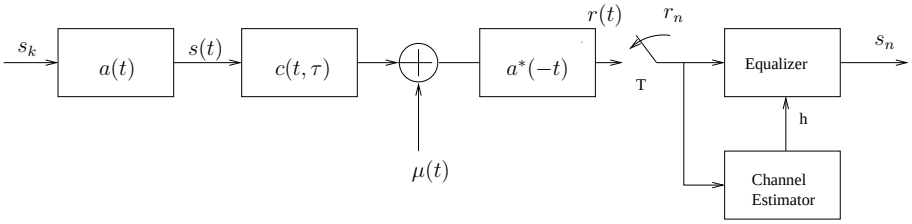

Figure 1.1: Base band transmission system model.

The continuous time transmission system model is depicted in Fig. 1.1, where the information signal is first modulated by a pulse shaping filter (PSF) $a(t)$ and then transmitted through time-varying frequency-selective channel $c(t,\tau)$, the received signal is distorted by AWGN noise $\mu(t)$ and then passes through matched PSF $a^*(t)$. The base-band transmitted symbol sequence and modulated signal waveform of data rate $1/T$ symbols/sec depicted in Fig. 1.1 are given by

$$s_T(t) = \sum_k s_k \delta(t - kT), \tag{1.5}$$

$$s(t) = s_T(t) * a(t) = \sum_k s_k a(t - kT), \tag{1.6}$$

where $\{s_k\}$ are the transmitted symbols.

To include the transmitter PSF $a(t)$ and receiver PSF $a^*(-t)$, the composite channel impulse response can be defined as

$$h(t,\tau) = a(\tau) * c(t,\tau) * a^*(-\tau) = c(t,\tau) * b(\tau) \tag{1.7}$$

$$= \sum_{\ell=1}^{N_\ell} c_\ell(t) b(\tau - \tau_\ell(t)), \tag{1.8}$$

where $b(\tau) = a(\tau) * a^*(-\tau)$. As implied by (1.3), when $f_d \ll f_c$, we can assume $\tau_\ell(t) = \tau_\ell$ for a long time period (approximately proportional to $\frac{f_c}{2f_d f_s}$, where f_s is the sampling frequency). In this case, we can rewrite (1.8) as

$$h(t,\tau) = \sum_{\ell=1}^{N_\ell} c_\ell(t) b(\tau - \tau_\ell). \tag{1.9}$$

In general, the received signal $r(t)$ is defined as

$$r(t) = h(t,\tau) * s_T(t) + \nu(t) = \int \sum_{\ell=1}^{N_\ell} c_\ell(t) b(\tau - \tau_\ell(t)) \sum_k s_k \delta(t - \tau - kT) d\tau + \nu(t) \tag{1.10}$$

$$= \sum_k s_k \sum_{\ell=1}^{N_\ell} c_\ell(t) b(t - kT - \tau_\ell(t)) + \nu(t), \tag{1.11}$$

where $\nu(t) = \mu(t) * a^*(-t)$ and $\mu(t)$ is the AWGN noise. Sampling $r(t)$ with period T, we obtain

$$r(nT) = \sum_k s_k \sum_{\ell=1}^{N_\ell} c_\ell(nT) b(nT - kT - \tau_\ell(nT)) + \nu(nT) \tag{1.12}$$

$$= \sum_l s_{n-l} \sum_{\ell=1}^{N_\ell} c_\ell(nT) b(lT - \tau_\ell(nT)) + \nu(nT) \tag{1.13}$$

Define $r_n = r(nT)$, $\nu_n = \nu(nT)$ and $h_{n,l} = \sum_{\ell=1}^{N_\ell} c_\ell(nT) b(lT - \tau_\ell(nT))$, then discrete time system model is given by:

$$r_n = \sum_{l=0}^{L} s_{n-l} h_{n,l} + v_n, \tag{1.14}$$

where we assume $h_{n,l}$ has finite support $[0\ L]$. Generally, the PSF is assumed to be a Nyquist filter of bandwidth f_s, therefore $\{v_n\}$ can be treated as AWGN noise. While for North America terrestrial digital TV transmission, the PSF is a root raised cosine filter of bandwidth $f_s/2$, therefore $\{v_n\}$ is a colored noise. However in this thesis we would still treat it as AWGN similar as in [10, 11], and extension to the colored noise case can be done with a little bit more efforts.

1.2.3 Channel Equalization

In this section, we give a brief retrospection on channel equalization schemes. First, we consider channel equalization schemes for moderately fast-fading channels, where the channel can be viewed as static within one processing block and varying across blocks, then we move on to discuss channel equalization algorithms for vary fast-fading channels, where the channel's time-variation within a single block can not be ignored.

In traditional low-mobility communication applications, the dominant factors which degrade the performance of communication systems are the multipath fading and noise. A conventional anti-multipath approach, which was pioneered in voiceband telephone modems, is to transmit a single carrier modulated by data symbols and a time-domain equalizer is applied at the receiver to compensate for ISI [12]. Various equalization methods, ranging from optimal approaches such as maximum a posteriori probability (MAP) symbol detection, maximum-likelihood (ML) sequence detection to suboptimal linear equalization such as zero-forcing (ZF), minimum mean square error (MMSE) symbol estimation, and nonlinear minimum mean square error decision feedback equalization (MMSE-DFE) have been proposed and researched in various

ways of trading of complexity for performance. However for severe multipath channels, which is more evident in wireless high-data-rate transmission, all these single carrier time-domain equalization (SC-TDE) schemes suffer from heavy computation complexity due to the long delay spread.

Multi-carrier modulation with frequency-domain equalization (FDE) techniques are proposed as alternative anti-multipath approaches for such kind of channels, and orthogonal frequency division multiplexing modulation (OFDM) with FDE system can be viewed as a successful example. OFDM transmits symbols through a large number of closely-spaced orthogonal sub-carriers, which is essentially using many slowly-modulated narrow band signals rather than one rapidly-modulated wide-band signal [13], therefore it transfers a severe frequency-selective channel into an parallel array of frequency-flat channels on each sub-carrier. As a favorable result, the channel equalization is simplified to a channel inversion operation on each sub-carrier, and the computational complexity of OFDM-FDE is approximately proportional to the logarithm of delay spread per symbol, which is much lower than the SC-TDE schemes. However, the transmitted OFDM signal is the sum of a large number of modulated sub-carriers, so OFDM suffers from high PAPR. This drawback increases the cost of power amplifiers. In addition, OFDM can be sensitive to carrier frequency offset and phase noise [6].

Single carrier FDE (SC-FDE) schemes are proposed as a promising alternatives to solve the high PAPR issue associated with OFDM [1]. SC-FDE transfers the FFT module from transmitter to receiver, thus avoids the high PAPR, but still inherits the low complexity advantage attributed to frequency-domain signal processing. In addition, it has some merits not shared by OFDM system. For example, coding,

while desirable, is not necessary for combating frequency selectivity, as it is in OFDM. Meanwhile, SC modulation is a well-proven technology in many existing wireless and wireline applications, and its radio frequency (RF) system linearity requirements are well known [1]. As shown in table 1.1, SC-FDE scheme possesses attractive features and especially fits applications with constraints on PAPR and power. Furthermore, SC-FDE shares a number of common signal processing functions with OFDE-FDE, thus SC and OFDM modems can easily be configured to coexist. In this dissertation, we conduct investigation of new equalization schemes to combat the doubly selective channels in the framework of SC-FDE.

Table 1.1: A Comparison of Anti-multipath Schemes [1].

	OFDM	SC-FDE	SC-TDE
Signal PAPR	High	Low	Low
Computational Complexity	Low	Low	High
Coding Requirement	Strict	Flexible	Flexible

With the increasing application/deployment of high-mobility and high-rate wireless communication, Doppler spread becomes an important factor in the system design. When the channel varies significantly within one OFDM symbol duration, sub-carriers are no longer perfectly orthogonal, severe ICI will degrade system performance substantially. The same dilemma also plagues SC-FDE schemes, since the resulting frequency-domain (FD) channel matrix is not diagonal any more, therefore the simple one-tap equalizer is not viable. Effective equalization for rapid time-varying frequency-selective channels is a challenging problem. In recent years, various approaches to suppress ICI for multi-carrier systems are investigated. Choi proposes

a MMSE successive detection algorithm [14] to cancel ICI, but the computation complexity is too high if the number of subcarrier is large. Assuming that some "small" ICI coefficients can be directly ignored, several ICI suppression algorithms with lower complexity are proposed in [15–17]. However, such assumption may not be valid as shown in [18], where a maximizing signal to noise plus interference ratio (SINR) window is derived to restrict ICI influence and then an iterative MMSE estimator is applied to cancel ICI as well as estimate finite-alphabet frequency-domain symbols. Rugini applies banded LDL factorization [19] to further reduce the complexity of estimation step in [18]. Besides, some parametric models are adopted to describe the doubly selective channel, various equalization algorithms based on those models are explored. Gorokhov [20] uses Taylor series expansion to linearly approximate time-domain channel variations and achieves low complexity channel equalization based on the structural property of data model. Barhumi proposes a frequency-domain per-tone equalizer based on complex exponential basis expansion model (CE-BEM) [21]. Motivated by the low-complexity ICI suppression scheme for OFDM systems in [18], we studied the FDE with analogous ICI suppression for SC systems and proposed iterative FDE schemes for both cyclic prefixed (CP) and non-cyclic prefixed (NCP) SC systems with the desired logarithmic per-symbol processing complexity.

1.2.4 Turbo Equalization

Turbo codes are first introduced by Berrou, Glavieux and Thitimajshima in [22]. They present stunning results that performance near the theoretical limits of shannon can be achieved with relatively simple code structure and decoding algorithm.

The magic comes from the decoding algorithm: iterative exchanging soft information between two simple constituent codes.

Inspired by the success of turbo decoding, researchers start investigating the application of such iterative soft information exchanging algorithms, which is termed "turbo principle", to solve other problems. Ever since then, turbo equalization becomes an active research direction. The idea of turbo equalization is first introduced in [23], where a soft-output Viterbi algorithm (SOVA) is applied for soft-in-soft-out channel equalization and decoding. A soft multi-user interference cancellation algorithm is proposed for code division multiple access (CDMA) system in [24]. Such idea is applied to turbo equalization in [25–27], and various techniques to reduce the computational cost required to compute the equalizer coefficients are discussed. Frequency-domain approaches for MMSE turbo equalization are proposed in [28–30] and [31] for single-input-single-output and multiple-input-multiple-output systems, respectively.

The key philosophy behind turbo equalization is to incorporate "soft information" into the equalization and decoding tasks. Traditionally, the equalizer estimates the symbols, makes a hard decision, and then feeds them to a decoder. This approach actually destroys information pertaining to how likely each of the possible data symbols might have be. However, this additional "soft" information can be converted into probabilities that a optimal decoding algorithm (such as BCJR algorithm [32]) can exploit for better performance.

Another key characteristic of turbo equalization is its iterative treatment. In turbo equalization, once the decoder processes the soft information it can, in turn, generate its own soft information indicating the relative likelihood of each transmitted bit. This

soft information from the decoder is fed back to the equalizer to aid symbol estimation. This process is often termed "belief propagation"or "message passing" [33,34] and has connections to methods in artificial intelligence, statistical inference, and graphical learning theory.

A closely related research topic with Turbo equalization is iterative channel estimation. For coherent detection (detecting transmitted symbols from received signal using an estimated channel impulse response (CIR)), channel estimator plays an important role. A number of researches consider exploiting the soft output information of turbo equalizer to improve the accuracy of channel estimation. Iterative CIR estimators based on least mean square (LMS), recursive least square (RLS) and Kalman filter are proposed in [35] and [36] respectively, which take soft information of data symbol estimates from equalizer as input and update filter coefficients accordingly. Application of iterative detection and channel estimation techniques in global systems for mobile communications (GSM) and enhanced data rates for global evolution (EDGE) shows a significant performance enhancement in [37]. In this dissertation, we also consider fitting soft input channel estimation into the turbo equalization framework, giving a receiver with iterative channel estimation, equalization, and decoding. The most relevant references to our work are [25, 28, 29, 36].

1.2.5 Channel Estimation

Channel estimation (CE) for doubly selective channel is a challenging and interesting problem. Generally speaking, CE schemes can be divided into two big families, one is training based CE (TB-CE) or decision directed CE (DD-CE) schemes, the other is blind CE schemes. For doubly selective channel, TB-CE and DD-CE are

more common, which can be roughly categorized into three classes: finite parametric model based CE schemes, statistical model based CE schemes and adaptive CE schemes.

Finite parametric model based CE schemes assume that the time variation of each independent channel coefficient can be captured by a linear combination of limited number of basis functions, thus CIR over a time interval can be attained by estimating those basis expansion parameters. Such models are commonly adopted for estimating very fast fading channels over a block duration. Various CE schemes have been researched for difference parametric models. Kalman filter, MMSE and LS channel estimator based on basis expansion model (BEM), Slepian basis, kernels for Rayleigh fading are investigated in [38–42]. A Taylor expansion based channel model is proposed in [43] to facilitate the design of ICI cancellation filter.

Statistical model based CE schemes assume the second-order statistics information of the channel is either known or available through estimation, thus CIR can be obtained by exploiting the correlation between received signal and priori known pilot symbols/ detected symbols. For very fast fading channels, pilot-aided channel estimation for multicarrier modulation are investigated in [44, 45]. For moderately fast fading channels, various frequency-domain channel estimation (FDCE) schemes have been proposed to track and predict wireless channels for OFDM systems, with or without pilot symbols, and with or without knowledge of channel statistics [46–48]. For SC systems, time-domain channel estimation is the typical approach [30, 36, 49], though a few pilot-aided FDCE schemes have been proposed [50–52]. A survey about linear channel estimation for systems with multiple antennas is presented in [53].

Adaptive CE schemes apply an adaptive filter to track channel variation in time, while both pilot symbols and detected data symbols can be used to update the filter coefficients. For rapid time-varying channels, adaptive filters are adopted to estimate the model parameters with low computational complexity for a BEM and polynomial basis in [54] and [55] respectively. Iterative CIR estimators based on LMS and RLS filter is investigated in [35]. For modest time-varying channels, a frequency-domain adaptive algorithm is proposed in [56] to track channels for SC transmission systems.

Blind CE schemes are adopted in communication system where training symbols are not available or not sufficient to initialize channel estimates. Various blind equalization methods have been proposed during the last ten years. These methods include higher-order statistical approaches [57], constant modulus algorithm (CMA) [58], subspace method based on second-order statistics [59], etc. In adaptive CE schemes, blind CE can serve as initialization step.

In this dissertation, we focus on FDCE algorithms and develop adaptive Kalman filter based per-tone channel estimator to track and predict channels for SC systems. the most relevant references with our work are [36, 39, 52].

1.3 Contribution and Outline

In the sequel, we give the dissertation outline and its main contributions.

In Chapter 2, we consider the receiver design for single carrier transmission systems over moderately fast-fading frequency-selective channels [60,61]. Particularly we investigate iterative frequency-domain equalization (IFDE) with explicit frequency-domain channel estimation (FDCE). First, an improved IFDE algorithm is presented

based on soft iterative interference-cancellation. Second, soft-decision-directed channel estimation algorithms are derived and analyzed both in time and frequency domain. As it turns out, frequency-domain approach is more computational efficient than time-domain approach. Therefore a new adaptive FDCE (AFDCE) algorithm based on per-tone Kalman filtering is proposed to track and predict the frequency-domain channel coefficients. The AFDCE algorithm employs across-tone noise reduction, exploits temporal correlation between successive blocks, and adaptively updates the auto-regressive model coefficients, bypassing the need for prior knowledge of channel statistics. Finally, a block overlapping idea is proposed which facilitates the joint operation of IFDE and AFDCE. Simulation results show that, compared to other existing IFDE and adaptive channel estimation schemes, the proposed schemes offer lower mean-square error (MSE) in channel prediction, lower BER after decoding, and robustness to non-stationary channels.

In Chapter 3, we consider a frequency-domain turbo equalization and adaptive frequency-domain channel estimation (FDTE/AFDCE) scheme for the reception of transmissions that employ trellis coded vestigial sideband (TCVSB) modulation, as specified by the ATSC North American terrestrial digital television (DTV) standard [62, 63]. The proposed FDTE/AFDCE scheme enables low-cost and high-performance reception of highly impaired DTV signals. Through numerical simulation, we demonstrate that our FDTE/CE scheme outperforms the traditional joint DFE/decoding plus frequency-domain least-mean-square (FDLMS) channel estimation approach at a fraction of the implementation cost.

In Chapter 4, we consider the receiver design for single carrier transmission systems over very fast fading frequency-selective channels [64–66]. In these quickly varying large-delay-spread channels, the traditional FDE methods fail when the channel response varies significantly over the FFT analysis window. Here we propose a novel FDE that is based on Doppler channel shortening, soft iterative interference cancellation, and block decision feedback. In addition, we derive a MMSE channel estimator for the pilot-aided estimation of significant channel coefficients in frequency domain, which are necessary for FDE. Numerical simulations show that the proposed scheme has advantages over the well-known FIR-MMSE-DFE plus RLS based CE scheme in both performance and complexity.

Finally in Chapter 5, we offer some concluding remarks and indicate future research possibilities.

To enhance the flow of the dissertation, we collect all detailed derivations in the appendices of each chapter.

1.4 Notation and Abbreviations

Matrices (column vectors) are denoted by upper (lower) bold face letters. Conjugate, transpose, Hermitian transpose, and inverse of $\boldsymbol{A}$ are denoted by $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, $\boldsymbol{A}^H$ and $\boldsymbol{A}^{-1}$, respectively. The Frobenius norm and l_2 norm are denoted by $\|\cdot\|_F$ and $\|\cdot\|$, respectively. The expectation, Kronecker delta, Kronecker product, modulo-N and integer ceiling operations are denoted by $\mathrm{E}[\cdot]$, $\delta(\cdot)$, $\otimes$, $<\cdot>_N$, $\lceil\cdot\rceil$, respectively. The $N \times N$ identity matrix and unitary discrete Fourier transform (DFT) matrix are denoted by $\mathbf{I}_{\mathbf{N}\times\mathbf{N}}$ and $\boldsymbol{F}_{N\times N}$, $\boldsymbol{i}_n$ for the nth column of $\mathbf{I}$. $\mathcal{C}(\boldsymbol{a})$ denotes the circulant matrix with first column $\boldsymbol{a}$, and $\mathcal{D}(\boldsymbol{a})$ is the diagonal matrix with diagonal elements

$\boldsymbol{a}$. $\mathrm{Re}(\cdot)$ denotes the real part, and $\mathrm{diag}(\boldsymbol{A})$ is the vector formed from the diagonal elements of square matrix $\boldsymbol{A}$. Finally, $CN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the multi-dimensional circular Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

Table 1.2: Abbreviations

AWGN	Additive White Gaussian Noise
AFDCE	Adaptive Frequency-Domain Channel Estimation
APPLE	Approximate Linear Estimation
AR	Auto-regressive
ATCR	Across Tone Channel Refinement
BER	Bit Error Rate
BPSK	Binary PSK (2-PSK)
CE	Channel Estimation
CIR	Channel Impulse Response
CMA	Constant Modulus Algorithm
CP	Cyclic Prefix
CPR	Cyclic Prefix Reconstruction
CSI	Channel State Information
CWGN	Circular White Gaussian Noise
DFE	Decision Feedback Equalizer
DTV	Digital Television
FDCE	Frequency-Domain Channel Estimation
FDE	Frequency-Domain Equalization
FDLMS	Frequency-Domain Least Mean Square
FDTE	Frequency-Domain Turbo Equalization
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
IBI	Interblock Interference
ICI	InterCarrier Interference
IFDE	Iterative Frequency-Domain Equalization
i.i.d.	independent and identically distributed
ISI	InterSymbol Interference
LMS	Least Mean Square
LS	Least Square
MAP	Maximum a Posterior
MF	Match filter
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MMSE-DFE	Minimum Mean Square Error Decision Feedback Equalization
MSE	Mean Squared Error
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak-to-Average-Power Ratio
QPSK	Quaternary PSK (4-PSK)
RLS	Recursive Least Squares
SC	Single Carrier
SCCP	Single Carrier Cyclic Prefix

SC-FDE	Single Carrier Frequency-Domain Equalization
SC-TDE	Single Carrier Time-Domain Equalization
SDD-CE	Soft-Decision-Directed Channel Estimation
SDD-TDCE	Soft-Decision-Directed Time-Domain Channel Estimation
SDD-FDCE	Soft-Decision-Directed Frequency-Domain Channel Estimation
SER	Symbol Error Rate
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
TE	Turbo Equalization
TCVSB	Trellis Coded Vestigial Side-band
VSB	Vestigial side-band
WSSUS	Wide-Sense-Stationary uncorrelated scattering
ZF	Zero-Forcing

CHAPTER 2

FREQUENCY-DOMAIN EQUALIZATION OF MODERATELY FAST FADING FREQUENCY-SELECTIVE CHANNELS

2.1 Introduction

Broadband wireless access systems offering high data-rates are likely to face severe multipath fading, including channel delay spreads spanning tens or hundreds of symbol intervals. While orthogonal frequency division multiplexing (OFDM) is a popular means of combating these multipath effects, its drawbacks include high PAPR and high sensitivity to carrier-frequency offset (CFO). Single carrier (SC) transmission with FDE presents an alternative to OFDM that retains robustness to channel delay spread without the disadvantages of high peak-to-average power ratio (PAPR) and CFO-sensitivity [1]. When FDE is accomplished via turbo equalization (TE) [23, 27], an iterative reception scheme whereby the equalizer and decoder iteratively exchange soft information to jointly exploit channel structure and code structure, significant performance gains result with only modest increase in demodulator complexity [28, 29, 62]. Hence, the focus of this chapter is SC transmission with turbo FDE.

When targeting practical implementation, accurate and efficient channel estimation (CE) is critical. For OFDM systems, various frequency-domain channel estimation (FDCE) schemes have been proposed to track and predict either slow-fading or fast-fading wireless channels, with or without pilot symbols, and with or without knowledge of channel statistics [39, 46, 47]. For SC systems, time-domain channel estimation is the typical approach [30, 36, 49], though a few pilot-aided FDCE schemes have been proposed [50–52]. With the decision-directed time-domain schemes, it has been observed that performance improvements result from the use of soft decoder outputs in place of hard symbol estimates [36, 49].

In this chapter, we propose a new joint channel-estimation/equalization scheme for the reception of SC transmissions over wireless channels with moderately fast fading and long delay spread. First, an improved iterative FDE (IFDE) algorithm is presented based on a frequency-domain TE idea. Second, soft-decision-directed channel estimation (SDD-CE) is studied both in time and frequency domain. Though the time-domain approach is optimal in minimizing the MSE, its heavy computational complexity prohibit practical applications. Therefore, we focus on frequency-domain approach, where a new adaptive FDCE (AFDCE) algorithm based on soft-input Kalman filtering and across-tone noise reduction is proposed to track and predict the channel in each frequency bin. Our AFDCE algorithm also exploits the temporal correlation between successive blocks and adaptively updates the channel's autoregressive (AR) model coefficients in case the channel statistics are unknown. Finally, a block-overlapping scheme is adopted to facilitate the joint operation of IFDE and AFDCE. Our approach differs from related work in the following ways.

1. Existing[1] IFDE algorithms [28, 29] are derived in the time domain and approximately translated to the frequency domain using the cyclic property of the equalizer. In contrast, our IFDE algorithm is derived in the frequency domain directly.

2. Existing soft-input CE algorithms [36, 49] work in the time domain. We focus on soft-input frequency-domain CE, hoping for low-complexity operation in the case of long channel delay spread.

3. Existing FDCEs [39, 46, 52] are pilot-aided in nature, even though practical pilots may be sparse. To better track time-varying channels, we consider (soft) decision-directed CE.

The chapter is organized as follows. Section 2.2 briefly introduces the communication system model. Section 2.3 summarizes the receiver architecture and section 2.4 describes the CPR procedure. IFDE and SDD-CE are detailed in sections 2.5 and section 2.6, respectively. Section 2.7 discusses implementation issues, and section 2.8 presents numerical results. Finally, section 2.9 concludes.

2.2 System Model

Consider coded single-carrier transmission where a bit stream $\{b_m\}$ is coded and mapped to symbols $\{s_n\}$ in a finite alphabet $\mathcal{S}$ and transmitted over a noisy linear time-varying multipath wireless channel. For simplicity, we assume $\{s_n\}$ to be uncorrelated. The complex-baseband channel can be described by the time-varying length-N_h impulse response $\{h_{n,l}\}_{l=0}^{N_h-1}$, where $h_{n,l}$ denotes the time-n response to an

[1]The IFDE we proposed in [62], appropriate for vestigial side-band (VSB) modulation, is a special case of the IFDE described here.

impulse applied at time $n-l$. The complex-valued observations $\{r_n\}$ are then given by

$$r_n = \sum_{l=0}^{N_h-1} h_{n,l} s_{n-l} + u_n, \tag{2.1}$$

where $\{u_n\}$ is zero-mean circular white Gaussian noise with variance σ_w^2. Note that (2.1) describes SC transmission without cyclic prefix (CP).

To implement IFDE and AFDCE jointly, we will eventually use overlapped block-processing with block length N and block shift interval N_d. (A detailed discussion is postponed until section 2.7.) Furthermore, we assume channel variation is slow enough to be modeled as time-invariant within a block. Thus, in terms of the block-based quantities $r_n(i) = r_{iN_d+n}$, $s_n(i) = s_{iN_d+n}$, $u_n(i) = u_{iN_d+n}$, and $h_l(i) = h_{iN_d+\frac{N}{2},l}$, the signal received during the ith block can be expressed as

$$r_n(i) = \begin{cases} u_n(i) + \sum_{l=0}^{n} h_l(i) s_{n-l}(i) + \sum_{l=n+1}^{N_h-1} h_l(i) s_{<n-l>_N}(i-1), & 0 \le n < N_h - 1, \\ u_n(i) + \sum_{l=0}^{N_h-1} h_l(i) s_{n-l}(i), & N_h - 1 \le n < N. \end{cases} \tag{2.2}$$

Note that $\{r_n(i)\}_{n=0}^{N_h-2}$ contain inter-block interference (IBI), i.e., symbol contributions from the previous block. In the sequel, we will make extensive use of the N-dimensional vectors $\boldsymbol{r}(i) := [r_0(i), \ldots, r_{N-1}(i)]^T$, $\boldsymbol{s}(i) := [s_0(i), \ldots, s_{N-1}(i)]^T$, $\boldsymbol{u}(i) := [u_0(i), \ldots, u_{N-1}(i)]^T$, and $\boldsymbol{h}(i) := [h_0(i), \ldots, h_{N_h-1}(i), 0, \cdots, 0]^T$.

2.3 Receiver Structure

The proposed receiver is illustrated in Fig. 2.1 and the corresponding processing steps are described below (for the ith block). Since steps 1-5 can be repeated several times for the same block, a superscript j is used to denote the iteration index.

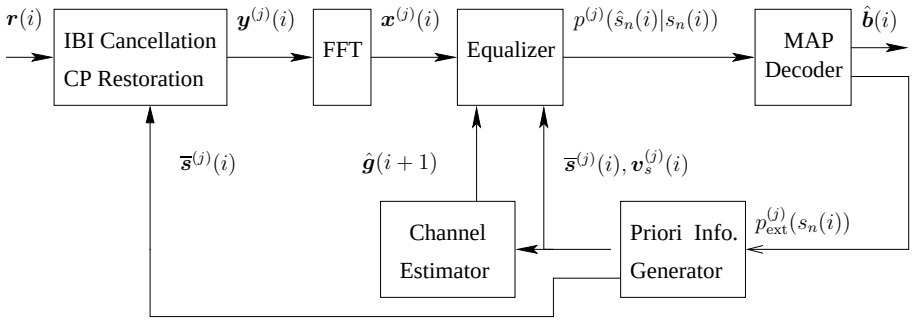

Figure 2.1: Receiver structure.

1. Perform IBI-cancellation and CP-reconstruction (CPR) on $\boldsymbol{r}(i)$ using the methods of [67, 68].

2. With the aid of FFTs, perform frequency-domain MMSE equalization assuming symbol means and variances obtained through the previous round of decoding. From the time-domain symbol estimates $\hat{\boldsymbol{s}}(i)$, extract the conditional probabilities $\left\{p^{(j)}\left(\hat{s}_n(i)|s_n(i)=s\right), \forall s \in \mathcal{S}\right\}_{n=0}^{N-1}$ for later use in decoding.

3. Perform maximum *a posteriori* (MAP) decoding, and update the extrinsic *a priori* distribution $p_{\text{ext}}^{(j)}(s_n(i))$.

4. Using $p_{\text{ext}}^{(j)}(s_n(i))$, generate symbol means $\overline{\boldsymbol{s}}^{(j)}(i)$ and variances $\boldsymbol{v}_s^{(j)}(i)$ to be used as priors in the next round of equalization.

5. Use $\overline{\boldsymbol{s}}^{(j)}(i)$ and $\boldsymbol{v}_s^{(j)}(i)$ to smooth current channel estimates and predict the channel for the next block.

For step 3, we assume that the LOGMAP algorithm [69] is used for MAP decoding and that the standard procedure is used to generate the *a priori* distribution (see, e.g., [27, 62]). In the following three sections, we describe the IBI cancellation and

CPR (step 1), IFDE algorithm (Steps 2 and 4) and the AFDCE algorithm (Step 5) in detail.

2.4 Cyclic Prefix Reconstruction

Usually, for OFDM or SC-FDE systems, a CP is added to the beginning of each transmission block to prohibit IBI as well as transform the linear convolution between channel and transmitted data block into a circular convolution, thus simplify channel equalization task. The CP is a repetition of the last data symbols in a block and for sufficient IBI cancellation, the length of CP should be larger or equal to $N_h - 1$. When CP is not available, the received signal is contaminated by IBI as shown in (2.2). In order to recover the contaminated samples, two steps, called IBI-cancellation and CPR must be taken [67, 68]. We adopted the IBI-cancellation and CPR algorithm proposed in [68] in our system, and here we briefly describe it for completeness of the dissertation.

An iterative IBI-cancellation and CPR is implemented jointly with IFDE. For the first iteration where $j = 1$, the IBI-cancellation and CPR is performed as

$$y_n^{(1)}(i) = \begin{cases} r_n(i) - \sum_{l=n+1}^{N_h-1} h_l(i)\hat{s}_{<n-l>_N}(i-1) + \sum_{l=n+1}^{N_h-1} h_l(i)\hat{s}^{(0)}_{<n-l>_N}(i), & 0 \le n < N_h - 1 \\ r_n(i) & N_h - 1 \le n < N \end{cases}. \tag{2.3}$$

where $\{\hat{s}_n(i-1)\}$ are the final estimates of previous-block symbols, and $\{\hat{s}^{(0)}_{<n-l>_N}(i)\}$ can be linearly estimated from $\{y_n^{(0)}(i)\}$, which is obtained from a linear combination between $\boldsymbol{r}(i)$ and $\boldsymbol{r}(i+1)$ as:

$$y_n^{(0)}(i) = \begin{cases} r_n(i) - \sum_{l=n+1}^{N_h-1} h_l(i)\hat{s}_{<n-l>_N}(i-1) + \xi_n r_{n+N-N_d}(i+1) \ , & 0 \le n < N_h - 1 \\ r_n(i) & N_h - 1 \le n < N \end{cases} \tag{2.4}$$

The optimum weighting coefficients $\{\xi_n\}$ to minimize the interference power in (2.4) is given by [68]

$$\xi_n = \frac{\sum_{l=n+1}^{N_h-1} |h_l(i)|^2}{\sum_{l=0}^{N_h-1} |h_l(i)|^2}. \tag{2.5}$$

For the subsequent iterations where $j > 1$, we perform CPR with $\{\bar{s}_n^{(j-1)}(i)\}$—the most recent estimates of current-block symbols after the $j-1$th iteration, according to

$$y_n^{(j)}(i) = \begin{cases} r_n(i) - \sum_{l=n+1}^{N_h-1} h_l(i)\hat{s}_{<n-l>_N}(i-1) + \sum_{l=n+1}^{N_h-1} h_l(i)\bar{s}_{<n-l>_N}^{(j-1)}(i), & 0 \le n < N_h - 1 \\ r_n(i) & N_h - 1 \le n < N \end{cases}. \tag{2.6}$$

More details on the generation of $\{\hat{s}_n(i-1)\}$ and $\{\bar{s}_n^{(j-1)}(i)\}$ will be provided in the sequel.

2.5 Iterative Frequency-Domain Equalization

Throughout this section, we will assume perfect knowledge of the channel coefficients $\{h_l(i)\}_{l=0}^{N_h-1}$. In practice, these coefficients are estimated using the AFDCE algorithm described in section 2.6.2.

Assuming that IBI-cancellation and CP-restoration are perfectly executed, $\boldsymbol{y}^{(j)}(i) := [y_0^{(j)}(i), \ldots, y_{N-1}^{(j)}(i)]$ can be considered as a noise-corrupted output of a circular convolution between the channel $\boldsymbol{h}(i)$ and the transmitted symbols $\boldsymbol{s}(i)$, i.e.,

$$\boldsymbol{y}(i) \;=\; \mathcal{C}\left(\boldsymbol{h}(i)\right)\boldsymbol{s}(i) + \boldsymbol{u}(i), \tag{2.7}$$

where, for notational simplicity, we suppress the iteration index j for the remainder of this section. Taking the discrete Fourier transform (DFT) of both sides of (2.7), we obtain the frequency-domain description

$$\boldsymbol{x}(i) \;=\; \boldsymbol{G}(i)\boldsymbol{t}(i) + \boldsymbol{w}(i), \tag{2.8}$$

where $\boldsymbol{x}(i)$, $\boldsymbol{t}(i)$ and $\boldsymbol{w}(i)$ denote the DFTs of $\boldsymbol{y}(i)$, $\boldsymbol{s}(i)$ and $\boldsymbol{u}(i)$, respectively, and where $\boldsymbol{G}(i) = \mathcal{D}\left(\boldsymbol{g}(i)\right)$, $\boldsymbol{g}(i)=\sqrt{N}\boldsymbol{F}\boldsymbol{h}(i)$, and $\boldsymbol{w}(i) \sim CN(\boldsymbol{0}, \sigma_w^2\mathbf{I})$. In the sequel, we refer to the elements in $\boldsymbol{t}(i)$ as *virtual subcarriers*.

Packing the mean and variance of each element in $\boldsymbol{s}(i)$ into the vectors $\overline{\boldsymbol{s}}(i)$ and $\boldsymbol{v}_s(i)$, respectively, it follows that

$$\begin{aligned}
\overline{\boldsymbol{t}}(i) \;&:=\; \mathrm{E}[\boldsymbol{t}(i)] \\
&=\; \boldsymbol{F}\overline{\boldsymbol{s}}(i), && (2.9)\\
\boldsymbol{R}_{\boldsymbol{tt}}(i) \;&:=\; \mathrm{E}\left[\left(\boldsymbol{t}(i) - \overline{\boldsymbol{t}}(i)\right)\left(\boldsymbol{t}(i) - \overline{\boldsymbol{t}}(i)\right)^H\right] \\
&=\; \boldsymbol{F}\,\mathcal{D}\left(\boldsymbol{v}_{\boldsymbol{s}}(i)\right)\boldsymbol{F}^H, && (2.10)\\
\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i) \;&:=\; \mathcal{D}\left(\mathrm{diag}\left(\boldsymbol{R}_{\boldsymbol{tt}}(i)\right)\right). && (2.11)
\end{aligned}$$

To simplify the equalization task, we use the approximate correlation matrix $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$ in place of the true correlation matrix $\boldsymbol{R}_{\boldsymbol{tt}}(i)$. Note that, when the elements in $\boldsymbol{s}(i)$ are i.i.d, the approximation is perfect (i.e., $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i) = \boldsymbol{R}_{\boldsymbol{tt}}(i)$). Note also that the

approximate linear estimation (APPLE) algorithm proposed in [28, 29] makes the more severe approximation $\boldsymbol{R_{tt}}(i) \approx \mathbf{I}$.

Taking $\overline{\boldsymbol{t}}(i)$ and $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$ as priors, the MMSE estimate of $\boldsymbol{t}(i)$ becomes [70]

$$\hat{\boldsymbol{t}}(i) = \overline{\boldsymbol{t}}(i) + \tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)\boldsymbol{G}^H(i)\boldsymbol{R}_{\boldsymbol{xx}}^{-1}\left(\boldsymbol{x}(i) - \boldsymbol{G}(i)\overline{\boldsymbol{t}}(i)\right), \tag{2.12}$$

$$\boldsymbol{R}_{\boldsymbol{xx}} = \boldsymbol{G}(i)\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)\boldsymbol{G}^H(i) + \sigma_w^2\mathbf{I}. \tag{2.13}$$

From a straightforward examination of $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$, it can be shown that the diagonal elements identically equal $v_t(i) = \frac{1}{N}\sum_{n=0}^{N-1} v_{s_n}(i)$, so that $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i) = v_t(i)\mathbf{I}$. Therefore, from (2.12), the kth element in $\hat{\boldsymbol{t}}(i)$ can be conveniently computed via

$$\hat{t}_k(i) = \overline{t}_k(i) + \underbrace{\frac{v_t(i)g_k^*(i)}{v_t(i)|g_k(i)|^2 + \sigma_w^2}}_{:= b_k(i)}\left(x_k(i) - g_k(i)\overline{t}_k(i)\right). \tag{2.14}$$

The time-domain symbol estimates are then obtained via inverse FFT as $\hat{\boldsymbol{s}}(i) = \boldsymbol{F}^{-1}\hat{\boldsymbol{t}}(i)$. Assuming that the symbol estimation error has a Gaussian distribution (which is reasonable for large N), we can generate priors for the MAP decoder as follows:

$$p(\hat{s}_n(i)|s_n(i) = s) = \frac{1}{\sqrt{\pi\sigma_{n,i|s}^2}}\exp\left(-\frac{(\hat{s}_n(i) - u_{n,i|s})^2}{\sigma_{n,i|s}^2}\right), \tag{2.15}$$

$$u_{n,i|s} := \mathrm{E}\{\hat{s}_n(i)|s_n(i) = s\}, \tag{2.16}$$

$$\sigma_{n,i|s}^2 := \mathrm{var}\{\hat{s}_n(i)|s_n(i) = s\}, \tag{2.17}$$

where $s \in \mathcal{S}$. In Appendix 2.A we show that $u_{n,i|s}$ and $\sigma_{n,i|s}^2$ can be written as

$$u_{n,i|s} = \overline{s}_n(i) + \frac{s - \overline{s}_n(i)}{N}\sum_{k=0}^{N-1} b_k(i)g_k(i), \tag{2.18}$$

$$\sigma_{n,i|s}^2 \approx \frac{1}{N}\sum_{k=0}^{N-1}|b_k(i)|^2\left(|g_k(i)|^2\tilde{v}_n(i) + \sigma_w^2\right), \tag{2.19}$$

where $\tilde{v}_n(i) := \frac{1}{N}\sum_{k\neq n} v_{s_k}(i)$ and where $b_k(i)$ was defined in (2.14). The approximation in (2.19) follows from the use of $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$ in place of $\boldsymbol{R}_{\boldsymbol{tt}}(i)$, as was done in (2.12)-(2.13).

Finally, we consider the update of *a priori* information for the MMSE estimator using the extrinsic information provided by the decoder. As in [27, 62], the soft feedback information can be expressed as $\{P(s_n(i) = s|\hat{\boldsymbol{s}}(i))\}_{s\in\mathcal{S}}$, which can be used to update the mean and variance of $s_n(i)$ as follows:

$$\begin{aligned} \bar{s}_n(i) &:= \mathrm{E}\{s_n(i)|\hat{\boldsymbol{s}}(i)\} \\ &= \sum_{s\in\mathcal{S}} sP(s_n(i) = s|\hat{\boldsymbol{s}}(i)) \qquad (2.20) \\ v_{s_n(i)} &:= \mathrm{var}\{s_n(i)|\hat{\boldsymbol{s}}(i)\} \\ &= \sum_{s\in\mathcal{S}} |s - \bar{s}_n(i)|^2 P(s_n(i) = s|\hat{\boldsymbol{s}}(i)). \qquad (2.21) \end{aligned}$$

2.6 Soft-Decision-Directed Channel Estimation

In this section, we investigate soft-decision-directed channel estimation (SDD-CE) algorithms that works in conjunction with the IFDE. Unlike conventional approaches for channel estimation, which rely on pilot symbols [50–52] or hard decisions [47], the soft outputs of a turbo equalizer can be exploited to improve CE performance [36, 49] and combat error propagation. First we consider a SDD time-domain CE (SDD-TDCE) in section 2.6.1, which is the optimal estimator to minimize the MMSE under perfect model match assumption. However, the SDD-TDCE is computational intensive, thus not attractive for practical applications. Motivated by a significant reduction in complexity, we propose a two-stage channel estimator, the structure of which is depicted in Fig. 2.2. In the first stage, per-tone soft-input Kalman filtering

is applied to track the channel in the frequency-domain. In the second stage, across-tone filtering is applied to refine the channel estimates. Finally, to handle the case where the channel statistics are unknown or time-varying, we propose a method to track the channel statistics inspired by [39].

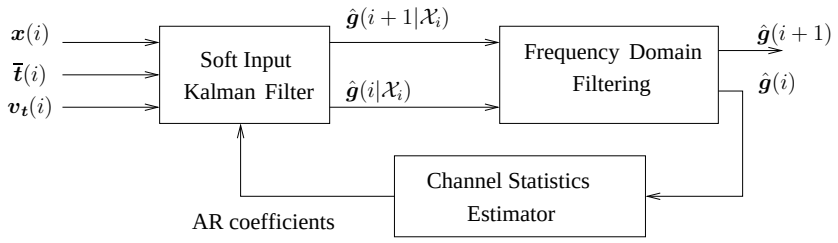

Figure 2.2: Adaptive frequency-domain channel estimator.

2.6.1 Soft-Decision-Directed Time-Domain Channel Estimation

In order to exploit the soft outputs of IFDE, we write the symbol $s_k(i)$ as

$$s_k(i) \;\; = \;\; \overline{s}_k(i) + \tilde{s}_k(i), \tag{2.22}$$

where it is assumed that the mean $\overline{s}_k(i)$ is known and the deviation $\tilde{s}_k(i)$ has zero mean and a correlation that obeys $\mathrm{E}[\tilde{s}_k(i)\tilde{s}_{k+p}(i+q)^*] = v_{s_k}(i)\delta_p\delta_q$.

In this section, we assume a wide sense stationary uncorrelated scattering (WSSUS) channel, in which case we can write $\mathrm{E}[h_l(i)h_{l+p}(i+q)] = \sigma^2_{h_l}\rho_q\delta_p$, where $\{\rho_q\}$ is the time-domain autocorrelation sequence (normalized so that $\rho_0 = 1$) and where $\{\sigma^2_{h_l}\}_{l=0}^{N_h-1}$ is the ISI-power profile. Taking the above decomposition of $s_k(i)$ into consideration, the AR model and observation equation for the lth channel tap can be

formulated as (see Appendix 2.B):

$$h_l(i) = \sum_{m=1}^{M} \alpha_l h_l(i-m) + \epsilon_l(i), \tag{2.23}$$

$$y_n(i) = \sum_{l=0}^{N_h-1} h_l(i)\bar{s}_{<n-l>_N}(i) + \underbrace{\sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i)}_{:=\mu_n(i)} \qquad n \in \{0, \cdots, N-1\} \tag{2.24}$$

where it is shown that the combined noise term $\mu_n(i) = \sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i)$ is zero-mean Gaussian with autocorrelation $\mathrm{E}[\mu_n(i)\mu^*_{n+p}(i+q)] = \sigma^2_{\mu_n(i)}\delta_p\delta_q$ and $\sigma^2_{\mu_n(i)} = \rho_q \sum_{l=0}^{N_h-1} \sigma^2_{\tilde{s}_{<n-l>_N}(i)}\sigma^2_{h_l} + \sigma^2_u$.

In the sequel, we define $\underline{\boldsymbol{h}}_l(i) = [h_l(i), h_l(i-1), \cdots, h_l(i-M+1)]^T$, $\underline{\boldsymbol{h}}(i) = [\underline{\boldsymbol{h}}_0(i)^T, \cdots, \underline{\boldsymbol{h}}_{N_h-1}(i)^T]^T$, $\underline{\boldsymbol{\epsilon}}_l(i) = [\epsilon_l(i), 0, \cdots, 0]^T$, $\underline{\boldsymbol{\epsilon}}(i) = [\underline{\boldsymbol{\epsilon}}_0(i)^T, \underline{\boldsymbol{\epsilon}}_1(i)^T, \cdots, \underline{\boldsymbol{\epsilon}}_{N_h-1}(i)^T]^T$, $\boldsymbol{y}(i) = [y_0(i), \cdots, y_{N-1}(i)]^T$, $\boldsymbol{\mu}(i) = [\mu_0(i), \cdots, \mu_{N-1}(i)]^T$, Then (2.23) and (2.24) can be rewritten in state-space model as:

$$\underline{\boldsymbol{h}}(i) = (\mathbf{I}_{N_h} \otimes \boldsymbol{A})\underline{\boldsymbol{h}}(i-1) + \underline{\boldsymbol{\epsilon}}(i) \tag{2.25}$$

$$\boldsymbol{y}(i) = \sqrt{N}\boldsymbol{F}^H_{N\times N}\,\mathcal{D}\left(\overline{\boldsymbol{t}}(i)\right)\boldsymbol{F}_{N\times N_h}(\mathbf{I}_{N_h} \otimes \boldsymbol{i}_1^H)\underline{\boldsymbol{h}}(i) + \boldsymbol{\mu}(i) \tag{2.26}$$

where $\mathrm{E}[\boldsymbol{\mu}(i)\boldsymbol{\mu}(i+q)^H] = \mathcal{D}(\boldsymbol{\sigma}^2_{\boldsymbol{\mu}(i)})\delta_q$, $\boldsymbol{\sigma}^2_{\boldsymbol{\mu}(i)} = [\sigma^2_{\mu_0(i)}, \cdots, \sigma^2_{\mu_{N-1}(i)}]^T$, and

$$\boldsymbol{A} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{M-1} & \alpha_M \\ 1 & 0 & \ddots & \ddots & 0 \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \tag{2.27}$$

using $\{\alpha_l\}_{l=1}^M$ to denote the AR model coefficients. Note $\epsilon_l(i) \sim CN(0, \sigma^2_{\epsilon_l})$, and given the channel statistics, $\{\alpha_l\}_{l=1}^M$ and $\sigma^2_{\epsilon_l}$ can be obtained via the Yule-Walker method [70].

In time domain, the Kalman filtering process with "soft" input can be carried out iteratively through [70]

$$\boldsymbol{Q}_i := \sqrt{N}\boldsymbol{F}_{N\times N}^H \mathcal{D}\left(\overline{\boldsymbol{t}}(i)\right) \boldsymbol{F}_{N\times N_h}(\mathbf{I}_{N_h}\otimes \boldsymbol{i}_1^H) \tag{2.28}$$

$$\boldsymbol{G}_i = \boldsymbol{P}_{i|i-1}\boldsymbol{Q}_i^H\left(\boldsymbol{Q}_i\boldsymbol{P}_{i|i-1}\boldsymbol{Q}_i^H + \mathcal{D}(\boldsymbol{\sigma}^2_{\boldsymbol{\mu}(i)})\right)^{-1} \tag{2.29}$$

$$\boldsymbol{e}(i) = \boldsymbol{y}(i) - \mathrm{E}[\boldsymbol{y}(i)|\mathcal{Y}_{i-1}] = \boldsymbol{y}(i) - \boldsymbol{Q}_i\underline{\hat{\boldsymbol{h}}}(i|\mathcal{Y}_{i-1}) \tag{2.30}$$

$$\underline{\hat{\boldsymbol{h}}}(i|\mathcal{Y}_i) = \mathrm{E}[\underline{\hat{\boldsymbol{h}}}(i)|\mathcal{Y}_i] = \underline{\hat{\boldsymbol{h}}}(i|\mathcal{Y}_{i-1}) + \boldsymbol{G}_i\boldsymbol{e}(i) \tag{2.31}$$

$$\underline{\hat{\boldsymbol{h}}}(i+1|\mathcal{Y}_i) = \mathrm{E}[\underline{\boldsymbol{h}}(i+1)|\mathcal{Y}_i] = \underline{\boldsymbol{A}}\underline{\hat{\boldsymbol{h}}}(i|\mathcal{Y}_i) \tag{2.32}$$

$$\boldsymbol{P}_{i+1|i} = \underline{\boldsymbol{A}}\left(\mathbf{I} - \boldsymbol{G}_i\boldsymbol{Q}_i\right)\boldsymbol{P}_{i|i-1}\underline{\boldsymbol{A}}^T + \mathcal{D}(\boldsymbol{\sigma}^2_{\underline{\boldsymbol{\epsilon}}(i)}) \tag{2.33}$$

where $\underline{\boldsymbol{A}} = \mathbf{I}_{N_h}\otimes\boldsymbol{A}$, $\boldsymbol{\sigma}^2_{\underline{\boldsymbol{\epsilon}}_l(i)} = [\sigma^2_{\epsilon_l}, 0, \cdots, 0]^T$, $\boldsymbol{\sigma}^2_{\underline{\boldsymbol{\epsilon}}(i)} = [{\boldsymbol{\sigma}^2_{\underline{\boldsymbol{\epsilon}}_0(i)}}^T, \cdots, {\boldsymbol{\sigma}^2_{\underline{\boldsymbol{\epsilon}}_{N_h-1}(i)}}^T]^T$, and $\boldsymbol{P}_{i+1|i} = \mathrm{E}\left[\left(\underline{\boldsymbol{h}}(i+1) - \underline{\hat{\boldsymbol{h}}}(i+1|\mathcal{Y}_i)\right)\left(\underline{\boldsymbol{h}}(i+1) - \underline{\hat{\boldsymbol{h}}}(i+1|\mathcal{Y}_i)\right)^H\right]$. $\mathcal{Y}_i$ denotes the set of all observations up to the ith block, namely, $\mathcal{Y}_i = \{\boldsymbol{y}(j)\}_{j=0}^i$.

In general, $\boldsymbol{Q}_i$ and $\boldsymbol{P}_{i|i-1}$ are time varying. The implementation of the TD Kalman filter would require an online inversion of a $N\times N$ matrix in each block. Since this is clearly impractical for large N, we will next develop a simplified FDCE with significantly reduced complexity.

2.6.2 Soft-Decision-Directed Frequency-Domain Channel Estimation

The motivation to search for a lower computational demanding CE algorithm leads us to an alternative approach - FDCE. First we derive the FD state-space model and compare it with the TD state-space model. The similarity between these two models stimulated us to propose a two-stage channel estimator to approximate the TDCE scheme, where per-tone soft-input Kalman filtering is applied to track each frequency bin and across-tone filtering is applied to refine the channel estimates. Finally, a

recursive filter inspired by [39] is adopted to track the channel statistics so as to adaptively update AR model.

Soft-Input Per-Tone Channel Estimation

Since the virtual subcarrier vector $\boldsymbol{t}(i)$ is the DFT of $\boldsymbol{s}(i)$, we can write

$$t_k(i) \;=\; \bar{t}_k(i) + \tilde{t}_k(i), \tag{2.34}$$

where $\bar{t}_k(i)$ is given in (2.9) and $\tilde{t}_k(i)$ is a random variable with zero mean and variance $v_t(i)$. Consistent with the use of $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$ in place of $\boldsymbol{R}_{\boldsymbol{tt}}(i)$ for MMSE equalization, we assume that $\mathrm{E}[\tilde{t}_k(i)\tilde{t}_{k+p}(i+q)^*] = v_t(i)\delta_p\delta_q$, which holds without approximation when $v_{s_k}(i)$ is invariant over k. Numerical experiments have convinced us that this is a reasonable assumption.

Taking the above decomposition of $t_k(i)$ into consideration, the state-space model for the kth frequency bin can be formulated as (see Appendix 2.C):

$$\underline{\boldsymbol{g}}_k(i) \;=\; \boldsymbol{A}\underline{\boldsymbol{g}}_k(i-1) + \underline{\boldsymbol{\eta}}_k(i), \tag{2.35}$$

$$x_k(i) \;=\; g_k(i)\bar{t}_k(i) + \underbrace{g_k(i)\tilde{t}_k(i) + w_k(i)}_{:=\, v_k(i)}, \tag{2.36}$$

where $\underline{\boldsymbol{g}}_k(i) = [g_k(i), g_k(i-1), \cdots, g_k(i-M+1)]^T$, $\underline{\boldsymbol{\eta}}_k(i) = [\eta_k(i), 0, \cdots, 0]^T$, and $\eta_k(i) \sim CN(0, \sigma_\eta^2)$. Given the channel statistics, $\{\alpha_l\}_{l=1}^M$ and σ_η^2 can be obtained via the Yule-Walker method [70]. In Appendix 2.C, it is shown that the combined noise term $v_k(i) = g_k(i)\tilde{t}_k(i) + w_k(i)$ is zero-mean Gaussian with autocorrelation $\mathrm{E}[v_k(i)v_{k+p}^*(i+q)] = \sigma_v^2(i)\delta_p\delta_q$ where $\sigma_v^2(i) = v_t(i)\sum_{l=0}^{N_h-1}\sigma_{h_l}^2 + \sigma_w^2$.

It follows naturally that the Kalman filtering process [70] can be carried out iteratively through the following steps. Assuming that $\boldsymbol{P}_k(i)$ and $\hat{\underline{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1})$ are

available from the previous block,[2]

$$\boldsymbol{q}_k(i) = \boldsymbol{P}_k(i)\bar{t}_k(i)^*\boldsymbol{i}_1\left(\bar{t}_k(i)\boldsymbol{i}_1^H\boldsymbol{P}_k(i)\boldsymbol{i}_1\bar{t}_k^*(i)+\sigma_v^2(i)\right)^{-1} \quad (2.37)$$

$$e_k(i) = x_k(i) - \bar{t}_k(i)\boldsymbol{i}_1^H\underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1}) \quad (2.38)$$

$$\underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i}) = \underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1}) + e_k(i)\boldsymbol{q}_k(i) \quad (2.39)$$

$$\underline{\hat{\boldsymbol{g}}}_k(i+1|\mathcal{X}_{k,i}) = \boldsymbol{A}\underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i}) \quad (2.40)$$

$$\boldsymbol{P}_k(i+1) = \boldsymbol{A}(\mathbf{I}-\boldsymbol{q}_k(i)\bar{t}_k(i)\boldsymbol{i}_1^H)\boldsymbol{P}_k(i)\boldsymbol{A}^T + \mathcal{D}(\boldsymbol{\sigma}_\eta^2), \quad (2.41)$$

where $\boldsymbol{\sigma}_\eta^2 = [\sigma_\eta^2, 0, \ldots, 0]^T$ and where $\mathcal{X}_{k,i}$ denotes the set of all observations up to the ith block, namely, $\mathcal{X}_{k,i} = \{\boldsymbol{x}_k(j)\}_{j=0}^i$. Recall that $\boldsymbol{i}_k$ denotes the kth column of the identity matrix, and note that $\boldsymbol{P}_k(i) := \mathrm{E}\left[\boldsymbol{\varepsilon}_k(i)\boldsymbol{\varepsilon}_k^H(i)\right]$ where $\boldsymbol{\varepsilon}_k(i) := \underline{\boldsymbol{g}}_k(i) - \underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1})$.

The channel estimator presented above is for a particular frequency bin. While channel is independently distributed for each frequency bin, we can formulate the state space equations for the channel over all frequency bins as:

$$\underline{\boldsymbol{g}}(i+1) = (\mathbf{I}_N \otimes \boldsymbol{A})\underline{\boldsymbol{g}}(i) + \underline{\boldsymbol{\eta}}_N(i) \quad (2.42)$$

$$\boldsymbol{x}(i) = \mathcal{D}(\overline{\boldsymbol{t}}(i))(\mathbf{I}_N \otimes \boldsymbol{i}_1^H)\underline{\boldsymbol{g}}(i) + \boldsymbol{v}(i) \quad (2.43)$$

where $\underline{\boldsymbol{g}}(i) = [\underline{\boldsymbol{g}}_0(i)^T, \cdots, \underline{\boldsymbol{g}}_{N-1}(i)^T]^T$, $\underline{\boldsymbol{\eta}}_N(i) = [\underline{\boldsymbol{\eta}}_0(i)^T, \cdots, \underline{\boldsymbol{\eta}}_{N-1}(i)^T]^T$, $\boldsymbol{v}(i) = [v_0(i), \cdots v_{N-1}(i)]^T$.

Compare (2.42)-(2.43) with (2.25)-(2.26), we can see that when $N_h = N$, these two models are exactly the same. Actually we can attain (2.43) from (2.26) by noting that $(\mathbf{I}_{N_h} \otimes \boldsymbol{i}_1^H)\underline{\boldsymbol{h}}(i) = \frac{1}{\sqrt{N}}F_{N\times N_h}^H(\mathbf{I}_N \otimes \boldsymbol{i}_1^H)\underline{\boldsymbol{g}}(i)$. When $N_h < N$, such relationship does not hold and TD channel model might be able to achieve better estimation

[2]For initialization (i.e., $i = 0$), we set $\underline{\hat{\boldsymbol{g}}}_k(0|\mathcal{X}_{k,0}) = \mathbf{0}$ and $\boldsymbol{P}_k(0) = \boldsymbol{R} := \mathrm{E}\{\underline{\boldsymbol{g}}_k(0)\underline{\boldsymbol{g}}_k^H(0)\}$.

performance, since we ignore the correlation between frequency bins in the state equation (2.42). However only FD state space model can enable us to attain decoupled channel estimator, which greatly simplify the computation. Therefore, we propose a two-stage channel estimator. In the first stage, we track and predict the channel for each frequency bin as in (2.37)-(2.41). In the second stage, we apply a across-tone channel refinement to exploit the correlation between frequency bins and reduce the noise contaminating the channel estimates.

Across-Tone Channel Refinement

Because the above channel tracking scheme is done on a per-tone basis, it has significantly less complexity than the optimal Kalman filtering (i.e., across all tones). However, it is suboptimal because it ignores correlation between the elements of $\boldsymbol{g}(i) = [g_0(i), g_1(i), \ldots, g_{N-1}(k)]^T$. Therefore, in this section, we propose a computationally efficient means of refining the per-tone channel estimates by leveraging the correlation structure of $\boldsymbol{g}(i)$.

Say that $\hat{\boldsymbol{g}}(i|\mathcal{X}_i) := [\hat{g}_0(i|\mathcal{X}_{0,i}), \hat{g}_1(i|\mathcal{X}_{1,i}), \cdots, \hat{g}_{N-1}(i|\mathcal{X}_{N-1,i})]^T$ denotes the per-tone estimates of $\boldsymbol{g}(i)$ generated via (2.37)-(2.41), and that $\boldsymbol{\varepsilon}(i) := \hat{\boldsymbol{g}}(i|\mathcal{X}_i) - \boldsymbol{g}(i)$ denotes the corresponding estimation error. We are interested in the linear refinement $\breve{\boldsymbol{g}}(i) = \boldsymbol{B}\hat{\boldsymbol{g}}(i|\mathcal{X}_i)$ which minimizes the MSE $\mathrm{E}\{\|\breve{\boldsymbol{g}}(i) - \boldsymbol{g}(i)\|^2\}$. Assuming that $\boldsymbol{\varepsilon}(i)$ is zero-mean with $\mathrm{E}\{\boldsymbol{\varepsilon}(i)\boldsymbol{\varepsilon}^H(i)\} = \sigma_\varepsilon^2\boldsymbol{I}$ and that $\mathrm{E}\{\boldsymbol{\varepsilon}(i)\boldsymbol{g}^H(i)\} = \boldsymbol{0}$, the orthogonality principle of MMSE estimation (i.e., $\mathrm{E}\{(\breve{\boldsymbol{g}}(i) - \boldsymbol{g}(i))\hat{\boldsymbol{g}}^H(i|\mathcal{X}_i))\} = \boldsymbol{0}$) straightforwardly implies that

$$\boldsymbol{B} \;=\; \boldsymbol{\Sigma}_g \left(\boldsymbol{\Sigma}_g + \sigma_\varepsilon^2\boldsymbol{I}\right)^{-1}, \tag{2.44}$$

where $\boldsymbol{\Sigma}_g := \mathrm{E}\{\boldsymbol{g}(i)\boldsymbol{g}^H(i)\}$. For our WSSUS channel, we show in Appendix 2.C that $\mathrm{E}[g_k(i)g_p(i)^*] = \sum_{l=0}^{N_h-1} \sigma_{h_l}^2 e^{-j\frac{2\pi}{N}(k-p)l}$, which implies that

$$\boldsymbol{\Sigma}_g \;=\; N\boldsymbol{F}\,\mathcal{D}(\boldsymbol{\sigma}_h^2)\boldsymbol{F}^H \tag{2.45}$$

for length-N vector $\boldsymbol{\sigma}_h^2 := [\sigma_{h_0}^2, \sigma_{h_1}^2, \ldots, \sigma_{h_{N_h-1}}^2, 0, \ldots, 0]^T$. Using (2.45), we can rewrite (2.44) as

$$\boldsymbol{B} \;=\; \boldsymbol{F}\,\mathcal{D}(\boldsymbol{\gamma})\boldsymbol{F}^H \tag{2.46}$$

where $\boldsymbol{\gamma} = [\gamma_0, \gamma_1, \ldots, \gamma_{N-1}]^T$ such that $\gamma_l = (1 + \sigma_\varepsilon^2/(N\sigma_{h_l}^2))^{-1}$. Equation (2.46) implies that MMSE refinement can be accomplished using a fast FFT-based algorithm. In the case that $\{\sigma_\varepsilon^2/\sigma_{h_l}^2\}_{l=0}^{N_h}$ are unknown, the high-SNR approximation $\sigma_\varepsilon^2 \to 0$ can be used, which leads to

$$\gamma_l \;=\; \begin{cases} 1 & 0 < l \le N_h - 1 \\ 0 & N_h \le l \le N-1 \end{cases}. \tag{2.47}$$

Note that this high-SNR approximation solution is the same as the constrained least-square solution $\min_{\breve{\boldsymbol{g}}(i)\in\boldsymbol{F}_{N\times N_h}} \|\breve{\boldsymbol{g}}(i) - \boldsymbol{g}(i)\|^2$, which is optimal in minimizing the square error with the assumption $\sigma_\varepsilon^2 = 0$.

While so far we have discussed across-tone refinement of a single vector $\hat{\boldsymbol{g}}(i|\mathcal{X}_i)$, merging the across-tone refinement procedure with the per-tone Kalman algorithm (2.37)-(2.41) requires that, for each i, across-tone refinement is applied to the entire M-sample block $\underline{\hat{\boldsymbol{G}}}(i|\mathcal{X}_i) := [\underline{\hat{\boldsymbol{g}}}_0(i|\mathcal{X}_{0,i}), \underline{\hat{\boldsymbol{g}}}_1(i|\mathcal{X}_{1,i}), \cdots, \underline{\hat{\boldsymbol{g}}}_{N-1}(i|\mathcal{X}_{N-1,i})]^T$, and that the refined outputs are used in the forward-prediction step (2.40). In total, this procedure consumes $2M$ FFTs at each index i.

Adaptive Tracking of AR Model Coefficients

When the Doppler spread of the channel is unknown or time-varying, we can estimate the AR model coefficients by tracking the channel statistics. As we can see from (2.35) and (2.27),

$$g_k(i) = \boldsymbol{\alpha}^H \underline{\boldsymbol{g}}_k(i-1) + \eta_k(i), \tag{2.48}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_M]^H$. For a stationary channel with $\boldsymbol{R} := \mathrm{E}\{\underline{\boldsymbol{g}}_k(i)\underline{\boldsymbol{g}}_k^H(i)\}$ and $\boldsymbol{r} := \mathrm{E}\{g_k^*(i+1)\underline{\boldsymbol{g}}_k(i)\}$, the Yule-Walker equations [70] specify that

$$\boldsymbol{\alpha} = \boldsymbol{R}^{-1}\boldsymbol{r}, \tag{2.49}$$

$$\sigma_\eta^2 = [\boldsymbol{R}]_{0,0} - \boldsymbol{\alpha}^H \boldsymbol{R}\boldsymbol{\alpha}. \tag{2.50}$$

When the statistics are slowly varying, we can use (2.49) to track the unknown AR coefficients $\boldsymbol{\alpha}(i)$ using estimates of $\boldsymbol{R}(i)$ and $\boldsymbol{r}(i)$, similar to [39]. In particular, we can use the recursive estimates

$$\hat{\boldsymbol{r}}(i) = \lambda\hat{\boldsymbol{r}}(i-1) + \frac{(1-\lambda)}{N}\sum_{k=0}^{N-1} \hat{g}_k^*(i)\hat{\underline{\boldsymbol{g}}}_k(i-1), \tag{2.51}$$

$$\hat{\boldsymbol{R}}(i) = \lambda\hat{\boldsymbol{R}}(i-1) + \frac{(1-\lambda)}{N}\sum_{k=0}^{N-1} \hat{\underline{\boldsymbol{g}}}_k(i-1)\hat{\underline{\boldsymbol{g}}}_k^H(i-1), \tag{2.52}$$

where $\lambda \in (0,1)$ is a suitably chosen forgetting factor, to generate the AR-coefficient estimate

$$\hat{\boldsymbol{\alpha}}(i) = \hat{\boldsymbol{R}}^{-1}(i)\hat{\boldsymbol{r}}(i). \tag{2.53}$$

While one might think to estimate $\sigma_\eta^2(i)$ via (2.50) with $\hat{\boldsymbol{R}}(i)$ and $\hat{\boldsymbol{\alpha}}(i)$ from (2.52)-(2.53), our experiments indicate that more robust estimates of $\sigma_\eta^2(i)$ can be obtained

via

$$\hat{\eta}_k(i) = \hat{g}_k(i) - \hat{\boldsymbol{\alpha}}^H(i)\underline{\hat{\boldsymbol{g}}}_k(i-1), \tag{2.54}$$

$$\hat{\sigma}_\eta^2(i) = \lambda\hat{\sigma}_\eta^2(i-1) + \frac{(1-\lambda)}{N}\sum_{k=0}^{N-1}|\hat{\eta}_k(i)|^2. \tag{2.55}$$

2.7 Implementation Considerations

In this section, we describe how the IFDE algorithm described in section 2.5 can be mated with the adaptive FDCE (AFDCE) algorithm described in section 2.6.2. In addition, we analyze the complexity of the joint IFDE/AFDCE algorithm and compare to existing approaches in the literature.

2.7.1 Block Overlapping

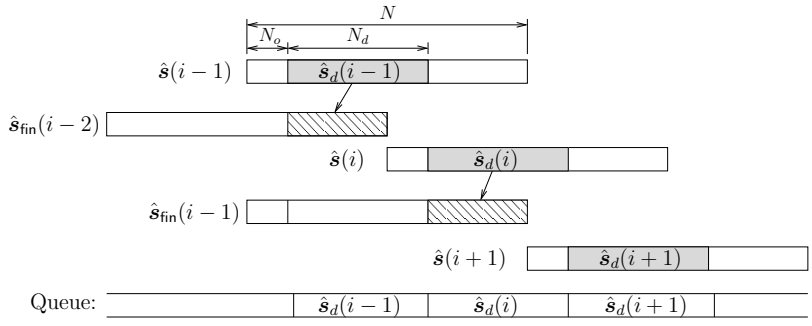

Figure 2.3: The block-overlapping scheme.

As mentioned in [62], due to causal channel dispersion and lack of CP, the symbols near the end of the block contribute little energy to the observation. As a result, these symbols are prone to detection errors. Though the CP-restoration procedure attempts to mitigate this problem, the procedure itself makes use of these end-of-block symbol estimates, which are inherently unreliable, and thus ultimately fails. Note that the

symbol estimates near the beginning of the block are contaminated by errors in IBI-cancellation as well as CP-restoration, due to both imperfect channel and symbol estimates. For these reasons, we propose the block-overlapping technique shown in Fig. 2.3. The key idea is to retain only the $N_d < N$ symbol estimates from the middle of the block (shown in grey), since only they are reliable; the symbols at the edges of a given block will be better estimated when they land near the center of a different block.

In particular, we propose to perform the equalization and channel estimation tasks as follows. Say that $\hat{\boldsymbol{s}}_d(i-1)$, estimates of the beginning symbols in $\boldsymbol{s}(i-1)$, have been computed and saved in a queue of "final symbol estimates." Then, as shown in Fig. 2.3, $\hat{\boldsymbol{s}}_d(i-1)$ is enough to complete the estimation of $\boldsymbol{s}(i-2)$. With final estimates of $\boldsymbol{s}(i-2)$, we can compute our final estimate of channel $\boldsymbol{g}(i-2)$ and forward-predict a tentative estimate of channel $\boldsymbol{g}(i)$. Using this tentative estimate of $\hat{\boldsymbol{g}}(i)$, in conjunction with IBI-cancellation based on the final estimates of $\boldsymbol{s}(i-2)$ and $\boldsymbol{g}(i-2)$, we can estimate $\boldsymbol{s}(i)$ and output the middle symbols $\hat{\boldsymbol{s}}_d(i)$ to the queue. Fig. 2.3 shows that we now have enough reliable symbol estimates to complete $\hat{\boldsymbol{s}}(i-1)$. With final estimates of $\boldsymbol{s}(i-1)$, we can compute our final estimate of channel $\boldsymbol{g}(i-1)$ and forward-predict a tentative estimate of channel $\boldsymbol{g}(i+1)$. These three latter quantities can be used to estimate $\boldsymbol{s}(i+1)$, after which the middle symbols $\boldsymbol{s}_d(i+1)$ are added to the queue, and so on.

2.7.2 Complexity Analysis

The computational complexity of the IFDE and AFDCE algorithms is reported in Table 2.1[3]. The complexity of similar algorithms from the literature is also reported

[3] $|\mathcal{S}|$ denotes the size of symbol constellation $\mathcal{S}$.

for comparison. For the symbol detection algorithms, Table 2.1 reports the number of real multiplications/divisions required per-iteration[4] to yield N time-domain symbol estimates, while for the CE algorithms, Table 2.1 reports the number of real multiplications/divisions required to yield N frequency-domain channel coefficient estimates.

Focusing first on symbol detection, we compare IFDE to the APPLE/MF algorithm from [28, 29]. As described in [28, 29], APPLE/MF alternates between the APPLE and MF tasks depending on the current system state, making an exact complexity count impossible. However, since we know that APPLE/MF complexity falls somewhere in-between the APPLE and MF complexities, we anticipate from Table 2.1 that IFDE is slightly cheaper than APPLE/MF. We note that Table 2.1 includes the cost of generating priors for MAP-decoding, but does not include the cost of computing priors for MMSE-equalization, since this latter cost is identical for APPLE/MF and IFDE.

Focusing next on channel estimation, we compare AFDCE to the least-mean-square structured channel estimation (LMS-SCE) algorithm from [56]. Notice that, in reporting AFDCE complexity, we have isolated the costs of across-tone channel refinement (ATCR) and adaptive tracking of AR model coefficients (ATARMC). Assuming the typical case of large block-length N and small AR-model order M (e.g., $M = 2$), the dominant complexity terms[5] in Table 2.1 indicate that the complexity

[4]We find it appropriate to report *per-iteration* complexity in Table 2.1 since we have observed that APPLE/MF and IFDE require approximately the same number of iterations before saturating.

[5]In deriving Table 2.1, we assumed radix-2 FFTs that cost $2N\log_2(N)$ real multiplications and real additions per real N-vector and $4N\log_2(N)$ real multiplications and real additions per complex N-vector.

Table 2.1: Computational Complexity

Task	Real $\times$	Real $\div$
IFDE	$8N\log_2(N) + 18N + \lvert\mathcal{S}\rvert N$	$2N$
APPLE	$8N\log_2(N) + 11N + \lvert\mathcal{S}\rvert N$	$2N$
MF	$14N\log_2(N) + 12N + \lvert\mathcal{S}\rvert N + 24N_h\log_2(2N_h)$	N
AFDCE	$N(4M^2 + 8M + 11) + 8N\log_2(N)$	N
(from ATCR)	$+ 8MN\log_2(N)$	
(from ATARMC)	$+ M^2N + 7MN + \frac{2}{3}M^3 + 3M^2$	1
LMS-SCE	$16N\log_2(N) + 6N$	

of AFDCE is about M times that of LMS-SCE. In other words, the complexities of the two algorithms are of the same order.

2.8 Numerical Results

2.8.1 Simulation Setup

We consider a single-carrier non-CP system, where the information bit sequence is encoded with the code generator $G(D) = (1 + D^2, 1 + D + D^2)$ and mapped to QPSK symbols via Gray mapping. The time-varying channel is simulated using Jakes' model with delay spread $N_h = 128$, uniform power profile $\sigma^2_{N_l} = \frac{1}{N_h}$, and autocorrelation $\rho_q = \mathcal{J}_0(2\pi f_d T_s N q)$, where $\mathcal{J}_0(\cdot)$ denotes the 0th-order Bessel function of the first kind. Note that the factor "N" appears in ρ_q because "q" denotes time-lag in blocks. Our experiments focused on (single-sided) normalized Doppler spread $f_d T_s \in \{0.00001, 0.00005\}$, which, e.g., corresponds to Doppler spreads of $f_d \in \{100, 500\}$Hz at sampling rate $T_s^{-1} = 10$ MHz. Our receiver used a block with length $N = 512$, offset $N_o = 50$, and reliable-symbol duration $N_d = 256$. The AR model order was $M = 2$, and $N_{iter} = 5$ iterations were used for both APPLE/MF

and IFDE (since more iterations did not appreciably improve performance). In the AFDCE across-tone channel-refinement step, we used the "high-SNR" approximation (2.47). The reported numerical results represent the average of 100 independent experiments of 51200 consecutive data symbols. Each length-51200 data-symbol sequence was preceded by a length-N pilot-symbol sequence (containing randomly chosen QPSK) that was used to initialize the channel estimator.

2.8.2 Performance Assessment

AR Model Based Time-Varying Channel

The time-varying channel is simulated as an AR process for each channel tap, while the second order AR model coefficients and the driving noise's variance is obtained through solving the Yule-Walker equations for the autocorrelation described above. In this set of simulations, we assume perfect IBI cancellation and CPR, which is equivalent to a cyclic-prefixed transmission, and assume perfect knowledge of the AR model. We evaluate the BER and MSE performance of IFDE/FDCE algorithm through simulations.

In Fig. 2.4, we compare the BER performance of IFDE/FDCE versus IFDE with perfect CSI, it can be seen that channel estimation error is negligible when $f_d T_s = 0.00001$ and for faster fading where $f_d T_s = 0.00005$, there is less than 0.25dB loss at high SNR region. In Fig. 2.5, we compare the MSE performance of IFDE/FDCE, FDCE with training versus the theoretical lower bound derived in Appendix 2.D. For the training based FDCE, all the transmitted symbols are pilot symbols and are selected from a particular class of polyphase, constant-magnitude sequences known as Chu sequences [71], which has constant frequency-domain amplitude [1,50]. It can be seen from Fig. 2.5 that the training based FDCE algorithm achieves the lower bound

at all SNR region, and IFDE/FDCE approaches the lower bound as SNR increase when $f_d T_s = 0.00001$, while has a almost constant gap when $f_d T_s = 0.00005$.

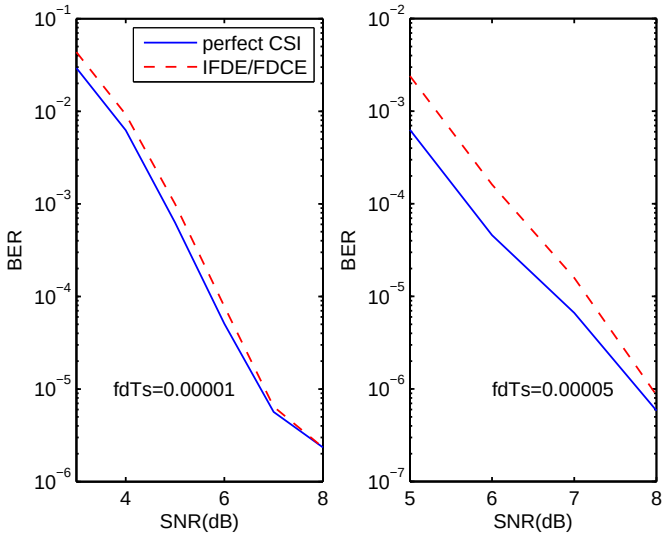

Figure 2.4: BER versus SNR for AR channels.

Rayleigh Fading Time-Varying Channel

We evaluate the performance of the proposed IFDE/AFDCE algorithm in Rayleigh fading time-varying channel. First, we compare the proposed IFDE/AFDCE algorithm with the LMS-SCE algorithm proposed by Morelli, Sanguinetti and Mengali in [56]. Three variants of our AFDCE algorithm were tested: (a) adaptive soft-input Kalman CE (ASKCE), where the inputs to the Kalman filter are the mean and variance of the virtual subcarrier symbols $\boldsymbol{t}$; (b) adaptive hard-input Kalman CE (AHKCE), where the inputs to the Kalman filter are hard-decisions on the virtual subcarrier symbols with the variance set at zero; and (c) adaptive soft/hard Kalman CE (ASHKCE), which uses AHKCE when the estimation error variance v_t is above

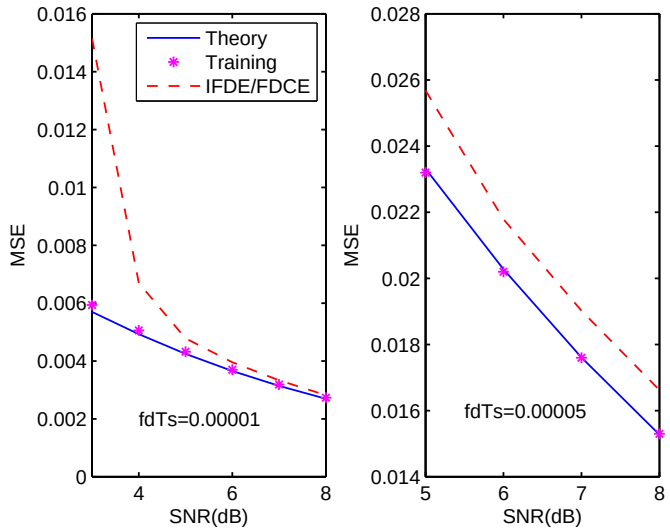

Figure 2.5: MSE versus SNR for AR channels.

a threshold (e.g., 0.1 in our simulations) and ASKCE otherwise. For LMS-SCE, we empirically chose step-sizes of $\mu = 0.1$ when $f_d T_s = 0.00001$, and $\mu = 0.5$ when $f_d T_s = 0.00005$, since no optimal choice of μ was specified in [56]. Steady-state BER and channel-estimation-MSE are reported in Fig. 2.6 and Fig. 2.7, respectively. There it can be seen that AHKCE and ASHKCE yield better performance than LMS-SCE throughout the SNR range, with significant improvements at higher Doppler. In Fig. 2.6, we also plot the BER of IFDE with perfectly known channel. There we see that both AHKCE and ASHKCE perform within 1dB of this genie-aided case. Though ASKCE does not perform well at low SNR, it slightly outperforms AHKCE at high SNR. ASHKCE combines the best features of the ASKCE and AHKCE algorithms, as seen in Fig. 2.6.

In Fig. 2.8, we compare the performance of ASKCE to adaptive-step-size LMS-SCE in a non-stationary channel. In particular, we use a channel for which $f_d T_s =$

0.00001 for the first 51200 symbols, and $f_d T_s = 0.00005$ for the last 51200 symbols. During the intermediate phase (i.e., the middle 51200 symbols), the channel smoothly transitioned between those two states. Both BER and channel-estimate-MSE are reported in Fig. 2.8. There we see that ASKCE achieves lower MSE than adaptive-stepsize LMS-SCE and that ASKCE demonstrates the ability to adapt to changing channel statistics while maintaining excellent BER performance.

Finally, we compare the channel equalization performance of our proposed IFDE algorithm with Tüchler and Hagenauer's APPLE/MF algorithm from [28,29]. Fig. 2.9 shows that our proposed IFDE/ASHKCE scheme outperforms APPLE/MF when the latter is used with either AHKCE or ASKCE.

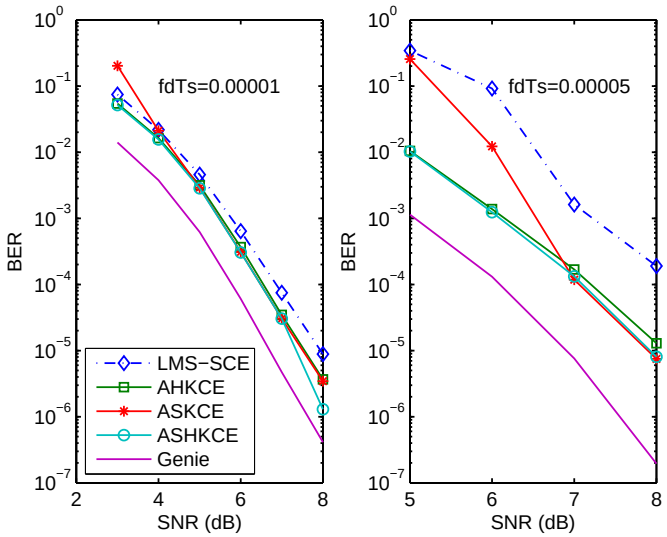

Figure 2.6: BER versus SNR for WSSUS Rayleigh channels.

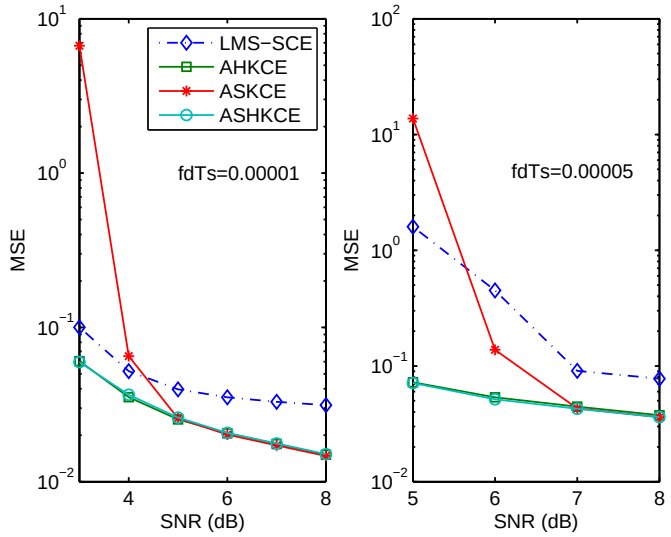

Figure 2.7: Channel-estimate-MSE versus SNR for WSSUS Rayleigh channels.

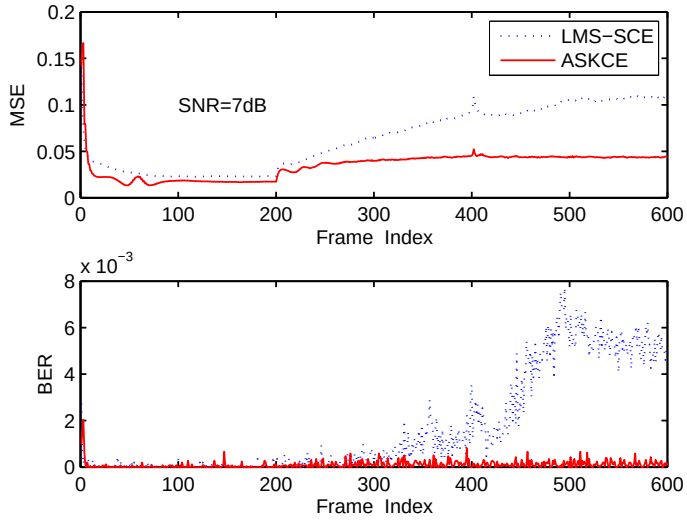

Figure 2.8: Channel-estimate-MSE and BER versus block index at SNR= 7dB for a non-stationary Rayleigh channel which transitions from $f_dT_s = 0.00001$ to $f_dT_s = 0.00005$.

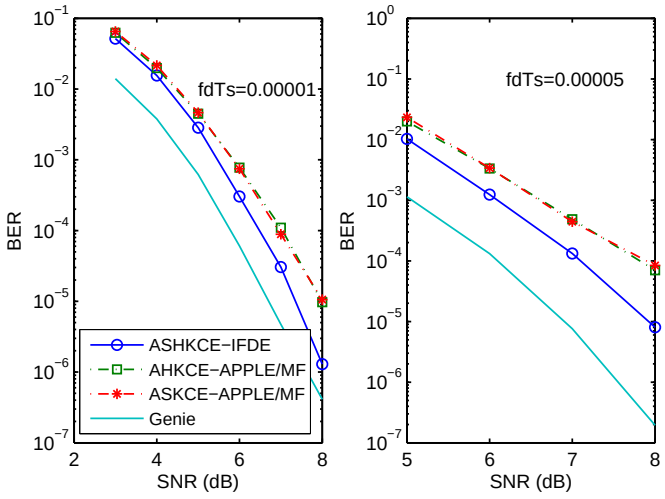

Figure 2.9: BER versus SNR for WSSUS Rayleigh channels.

2.9 Conclusion

In this chapter, we present an algorithm for joint frequency-domain equalization and channel estimation appropriate for the reception of single-carrier non-CP transmissions over time-varying long-delay-spread channels. In particular, we detail an improved IFDE algorithm based on frequency-domain turbo equalization, and propose a novel AFDCE with robustness to fast-fading. Numerical results show that the proposed IFDE-plus-AFDCE scheme demonstrates state-of-the-art performance in both stationary and non-stationary channels and maintains low complexity as a result of its frequency-domain operation. Deeper analytical insights into the convergence behavior of AFDCE will be the subject of future work.

Appendix

2.A Derivation of Conditional Mean and Variance

Combining (2.12) with $\hat{\boldsymbol{s}}(i) = \boldsymbol{F}^{-1}\hat{\boldsymbol{t}}(i)$, we can write

$$\begin{aligned}
\hat{s}_n(i) &= \boldsymbol{i}_n^H \boldsymbol{F}^H \hat{\boldsymbol{t}}(i) \\
&= \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \hat{t}_k(i) \\
&= \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \overline{t}_k(i) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k b_k(i) \left(x_k(i) - g_k(i)\overline{t}_k(i)\right) \\
&= \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \boldsymbol{i}_k^H \boldsymbol{F}\overline{\boldsymbol{s}}(i) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k b_k(i) \left(g_k(i)\boldsymbol{i}_k^H \boldsymbol{F}\boldsymbol{s}(i) - g_k(i)\boldsymbol{i}_k^H \boldsymbol{F}\overline{\boldsymbol{s}}(i) + w_k(i)\right) \\
&= \overline{s}_n(i) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \boldsymbol{i}_k^H b_k(i) g_k(i) \boldsymbol{F}\left(\boldsymbol{s}(i) - \overline{\boldsymbol{s}}(i)\right) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k b_k(i) w_k(i). \quad (2.56)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathrm{E}\left[\hat{s}_n(i)|s_n(i) = s\right] &= \overline{s}_n(i) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \boldsymbol{i}_k^H b_k(i) g_k(i) \boldsymbol{F}\, \mathrm{E}\left[\boldsymbol{s}(i) - \overline{\boldsymbol{s}}(i)|s_n(i) = s\right] \\
&= \overline{s}_n(i) + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \boldsymbol{i}_k^H b_k(i) g_k(i) \boldsymbol{F}\boldsymbol{i}_n\left(s - \overline{s}_n(i)\right) \\
&= \overline{s}_n(i) + \frac{s - \overline{s}_n(i)}{N} \sum_{k=0}^{N-1} b_k(i) g_k(i), \quad (2.57)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{var}\left[\hat{s}_n(i)|s_n(i) = s\right] &= \mathrm{E}\left[\left|\hat{s}_n(i) - \mathrm{E}\left[\hat{s}_n(i)|s_n(i) = s\right]\right|^2\right] \\
&= \mathrm{E}\left[\left| \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k \boldsymbol{i}_k^H b_k(i) g_k(i) \boldsymbol{F}\left(\boldsymbol{s}(i) - \overline{\boldsymbol{s}}(i) - \boldsymbol{i}_n(s - \overline{s}_n(i)\right)\right.\right. \\
&\qquad \left.\left. + \boldsymbol{i}_n^H \boldsymbol{F}^H \sum_{k=0}^{N-1} \boldsymbol{i}_k b_k(i) w_k(i) \right|^2\right] \\
&\approx \frac{1}{N} \sum_{k=0}^{N-1} |b_k(i)|^2 \left(|g_k(i)|^2 \tilde{v}_n(i) + \sigma_w^2\right), \quad (2.58)
\end{aligned}$$

where $\tilde{v}_n(i) := \frac{1}{N}\sum_{k\neq n} v_{s_k}(i)$. In arriving at (2.58), we ignored the off-diagonal elements of

$$\mathrm{E}\left[\boldsymbol{F}\big(\boldsymbol{s}(i)-\overline{\boldsymbol{s}}(i)-\boldsymbol{i}_n(s-\overline{s}_n(i)\big)\big(\boldsymbol{s}(i)-\overline{\boldsymbol{s}}(i)-\boldsymbol{i}_n(s-\overline{s}_n(i)\big)^H\boldsymbol{F}^H\right]$$

in a manner consistent with our previous approximation that used $\tilde{\boldsymbol{R}}_{\boldsymbol{tt}}(i)$ in place of $\boldsymbol{R}_{\boldsymbol{tt}}(i)$ for MMSE estimation.

2.B State-Space Model for Time-Domain Kalman Filter

In this appendix, we formulate the state-space model of Kalman filter for the case that the channel is WSSUS and an M-order Kalman filter is used to track the channel variation in each channel tap. The auto-regressive (AR) model for the lth channel tap is given by

$$h_l(i) = \sum_{m=1}^{M} \alpha_{m,l} h_l(i-m) + \epsilon_l(i), \tag{2.59}$$

where $\{\alpha_{m,l}\}_{m=1}^{M}$ are AR model coefficients and $\epsilon_l(i)$ is white Gaussian noise with zero mean and variance $\sigma^2_{\epsilon_l}$. Here, the parameters $\{\alpha_{m,l}\}_{m=1}^{M}$ and $\sigma^2_{\epsilon_l}$ are selected to match the desired autocorrelation function $\mathrm{E}[h_l(i)h_l(i+q)^*] = \sigma^2_{h_l}\rho_q$, implying that $\{\alpha_{m,l}\}_{m=1}^{M}$ are invariant w.r.t. tap index l, hence we omit the l-dependence from these parameters in the state-space formulation (2.23)-(2.24). Though $\sigma^2_{h_l}$ might be variant across l.

Assume perfect IBI cancellation and CPR, combine (2.6) and (2.22), the observation equation is defined as:

$$y_n(i) = \sum_{l=0}^{N_h-1} h_l(i)\bar{s}_{<n-l>_N}(i) + \sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i) \qquad n\in\{0,\cdots,N-1\} \tag{2.60}$$

Defining the combined noise term $\mu_n(i) = \sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i)$, we note that, since both $\tilde{s}_{<n-l>_N}(i)$ and $u_n(i)$ are Gaussian, $\mu_n(i)$ must also be Gaussian. Since $\tilde{s}_{<n-l>_N}(i)$ and $u_n(i)$ are independent zero-mean Gaussian random variables,

$$\mathrm{E}[\mu_n(i)] = \mathrm{E}\left[\sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i)\right] = 0, \tag{2.61}$$

$$\begin{aligned}
&\mathrm{E}[\mu_n(i)\mu^*_{n+p}(i+q)] \\
&= \mathrm{E}\left[\left(\sum_{l=0}^{N_h-1} h_l(i)\tilde{s}_{<n-l>_N}(i) + u_n(i)\right)\left(\sum_{l'=0}^{N_h-1} h^*_{l'}(i+q)\tilde{s}^*_{<n+p-l'>_N}(i+q) + u^*_{n+p}(i+q)\right)\right] \\
&= \sum_{l=0}^{N_h-1}\sum_{l'=0}^{N_h-1} \mathrm{E}\left[\tilde{s}_{<n-l>_N}(i)\tilde{s}^*_{<n+p-l'>_N}(i+q)\right]\mathrm{E}\left[h_l(i)h^*_{l'}(i+q)\right] + \sigma_u^2\delta_p\delta_q \\
&= \sum_{l=0}^{N_h-1} \mathrm{E}\left[\tilde{s}_{<n-l>_N}(i)\tilde{s}^*_{<n+p-l>_N}(i)\right]\sigma^2_{h_l(i)}\rho_q\delta_q + \sigma_u^2\delta_p\delta_q \\
&= \big(\rho_q \sum_{l=0}^{N_h-1} \sigma^2_{\tilde{s}_{<n-l>_N}(i)}\sigma^2_{h_l} + \sigma_u^2\big)\delta_p\delta_q.
\end{aligned} \tag{2.62}$$

2.C State-Space Model for Frequency-Domain Kalman Filter

In this appendix, we formulate the state-space model of Kalman filter used to track the channel variation in each frequency bin. The auto-regressive (AR) model for the kth frequency bin is given by

$$g_k(i) = \sum_{m=1}^{M} \alpha_{m,k} g_k(i-m) + \eta_k(i), \tag{2.63}$$

where $\{\alpha_{m,k}\}_{m=1}^{M}$ are AR model coefficients and $\eta_k(i)$ is white Gaussian noise with zero mean and variance $\sigma^2_{\eta_k}$. Here, the parameters $\{\alpha_{m,k}\}_{m=1}^{M}$ and $\sigma^2_{\eta_k}$ are selected to match the desired autocorrelation function $\mathrm{E}[g_k(i)g_k(i+q)^*]$. Since, for a WSSUS

channel, $\mathrm{E}[h_l(i)h_{l+p}(i+q)] = \sigma^2_{h_l}\rho_q\delta_p$ with $\rho_0 = 1$, we can see that

$$\begin{aligned}\mathrm{E}[g_k(i)g_k(i+q)^*] &= \mathrm{E}\left[\sum_{l=0}^{N_h-1} h_l(i)e^{-j\frac{2\pi}{N}kl}\sum_{m=0}^{N_h-1} h_m(i+q)^* e^{j\frac{2\pi}{N}km}\right], \\ &= \rho_q\sum_{l=0}^{N_h-1}\sigma^2_{h_l}, \end{aligned} \tag{2.64}$$

implying that $\{\alpha_{m,k}\}_{k=1}^M$ and $\sigma^2_{\eta_k}$ are invariant w.r.t. bin index k. Hence, we omit the k-dependence from these parameters in the state-space formulation (2.35)-(2.36). We can also see the across-tone correlation that is ignored in per-tone channel tracking:

$$\begin{aligned}\mathrm{E}[g_k(i)g_p(i)^*] &= \mathrm{E}\left[\sum_{l=0}^{N_h-1} h_l(i)e^{-j\frac{2\pi}{N}kl}\sum_{m=0}^{N_h-1} h_m(i)^* e^{j\frac{2\pi}{N}pm}\right], \\ &= \sum_{l=0}^{N_h-1}\sigma^2_{h_l}e^{-j\frac{2\pi}{N}(k-p)l}, \end{aligned} \tag{2.65}$$

Combining (2.8) and (2.34), the observation equation for per-tone filtering becomes

$$x_k(i) = g_k(i)\overline{t}_k(i) + g_k(i)\tilde{t}_k(i) + w_k(i). \tag{2.66}$$

Defining the combined noise term $v_k(i) := g_k(i)\tilde{t}_k(i) + w_k(i)$, we note that, since both $\tilde{t}_k(i)$ and $w_k(i)$ are Gaussian, $v_k(i)$ must also be Gaussian. Since $w_k(i)$ and $\tilde{t}_k(i)$ are independent zero-mean Gaussian random variables,

$$\mathrm{E}[v_k(i)] = 0, \tag{2.67}$$

$$\begin{aligned}\mathrm{E}[v_k(i)v^*_{k+p}(i+q)] &= \mathrm{E}\left[\left(g_k(i)\tilde{t}_k(i)+w_k(i)\right)\left(g^*_{k+p}(i+q)\tilde{t}^*_{k+p}(i+q)+w^*_{k+p}(i+q)\right)\right], \\ &= \mathrm{E}\left[g_k(i)g^*_{k+p}(i+q)\right]\mathrm{E}\left[\tilde{t}_k(i)\tilde{t}^*_{k+p}(i+q)\right] + \sigma^2_w\delta_p\delta_q, \\ &= \mathrm{E}\left[\sum_{l=0}^{N_h-1} h_l(i)e^{-j\frac{2\pi}{N}kl}\sum_{m=0}^{N_h-1} h^*_m(i+q)e^{j\frac{2\pi}{N}(k+p)m}\right] v_t(i)\delta_p\delta_q \\ &\qquad + \sigma^2_w\delta_p\delta_q, \end{aligned} \tag{2.68}$$

$$= \left(v_t(i)\sum_{l=0}^{N_h-1}\sigma^2_{h_l} + \sigma^2_w\right)\delta_p\delta_q. \tag{2.69}$$

In (2.68), we utilize the assumption that $\mathrm{E}[\tilde{t}_k(i)\tilde{t}_{k+p}(i+q)^*] = v_t(i)\delta_p\delta_q$ mentioned in section 2.6.2.

2.D Performance Bound of Channel Estimator

Assume block fading channel and perfect IBI cancellation and CP restoration, we evaluate the performance of the proposed CE scheme in a genie aided mode, where N constant pilot symbols $\{p_k\}$ are transmitted over the channel repeatedly. Furthermore, the pilot symbols are selected from a particular class of polyphase, constant-magnitude sequences known as Chu sequences [71], which satisfied the desired property that the training sequence has constant frequency-domain amplitude [1,50]. Then (2.35)-(2.36) changes to

$$\underline{\boldsymbol{g}}_k(i) = \boldsymbol{A}\underline{\boldsymbol{g}}_k(i-1) + \underline{\boldsymbol{\eta}}_k(i), \tag{2.70}$$

$$x_k(i) = \boldsymbol{i}_1^H\underline{\boldsymbol{g}}_k(i)p_k + w_k(i). \tag{2.71}$$

First, without the consideration of the frequency-domain filtering process, for such time invariant system, the discrete algebraic Riccati equation (DARE) is defined as

$$\boldsymbol{P}_k = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{A}^T - \frac{|p_k|^2\boldsymbol{A}\boldsymbol{P}_k\boldsymbol{i}_1\boldsymbol{i}_1^H\boldsymbol{P}_k\boldsymbol{A}^T}{|p_k|^2\boldsymbol{i}_1^H\boldsymbol{P}_k\boldsymbol{i}_1 + \sigma_w^2} + \sigma_\eta^2\boldsymbol{i}_1\boldsymbol{i}_1^H. \tag{2.72}$$

The unique stabilizing solution of (2.72) can be attained through applying the invariant subspace method as in [72] Theorem E.7.1.

$$\boldsymbol{M}_k := \begin{bmatrix} \boldsymbol{A}^{-1} & -\sigma_\eta^2\boldsymbol{A}^{-1}\boldsymbol{i}_1^H\boldsymbol{i}_1 \\ -\frac{1}{\sigma_w^2}\boldsymbol{i}_1^H\boldsymbol{i}_1\boldsymbol{A}^{-1} & \boldsymbol{A}^H - \frac{\sigma_\eta^2}{\sigma_w^2}\boldsymbol{i}_1^H\boldsymbol{i}_1\boldsymbol{A}^{-1}\boldsymbol{i}_1^H\boldsymbol{i}_1 \end{bmatrix}, \tag{2.73}$$

$$\boldsymbol{M}_k\begin{bmatrix}\boldsymbol{U}_k \\ \boldsymbol{V}_k\end{bmatrix} = \begin{bmatrix}\boldsymbol{U}_k \\ \boldsymbol{V}_k\end{bmatrix}\Lambda_k, \tag{2.74}$$

$$\boldsymbol{P}_k = \boldsymbol{U}_k\boldsymbol{V}_k^{-1}, \tag{2.75}$$

where $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$ can be any $M \times M$ matrices that form a basis for the stable eigenspace of the symplectic matrix $\boldsymbol{M}_k$ in (2.73) and Λ_k is an $M \times M$ matrix with all its eigenvalues inside the unit disc. For channel with i.i.d frequency bins, we can omit the subscript k from (2.72)-(2.75). In addition, since $\boldsymbol{i}_1^H \boldsymbol{P}_k(i)\boldsymbol{i}_1 = \mathbf{Var}\left(g_k(i) - g_k(i|\mathcal{X}_{k,i-1})\right)$, the variance of prediction error at steady state can be predicted through $\boldsymbol{i}_1^H \boldsymbol{P}\boldsymbol{i}_1$.

Second, we examine the influence of the channel refinement process on the lower bound when high SNR approximation is applied. To facilitate our analysis, we reformulate equation (2.37)-(2.41) as:

$$\tilde{\boldsymbol{q}}_k(i) = \boldsymbol{P}_k(i)\boldsymbol{i}_1\big(|p_k|^2 \boldsymbol{i}_1^H \boldsymbol{P}_k(i)\boldsymbol{i}_1 + \sigma_w^2\big)^{-1}, \tag{2.76}$$

$$\tilde{e}_k(i) = p_k^* x_k(i) - |p_k|^2 \boldsymbol{i}_1^H \underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1}), \tag{2.77}$$

$$\underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i}) = \underline{\hat{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1}) + \tilde{e}_k(i)\tilde{\boldsymbol{q}}_k(i), \tag{2.78}$$

$$\boldsymbol{P}_k(i+1) = \boldsymbol{A}(\mathbf{I} - \tilde{\boldsymbol{q}}_{k,i}|p_k|^2 \boldsymbol{i}_1^H)\boldsymbol{P}_k(i)\boldsymbol{A}^T + \mathcal{D}(\boldsymbol{\sigma}_{\boldsymbol{\eta}}^2). \tag{2.79}$$

Without lose of generality, we can assume $|p_k|^2 = 1$, since they are of constant amplitude. It follows that $\tilde{\boldsymbol{q}}_k(i)$ is k-independent with the same initialization of $\boldsymbol{P}_k(0)$, thus we omit the index k. Furthermore, $\tilde{e}_k$ can be decomposed as:

$$\begin{aligned}
\tilde{e}_k(i) &= p_k^*\big(p_k g_k(i) + w_k(i)\big) - |p_k|^2 \hat{g}_k(i|\mathcal{X}_{k,i-1}), \\
&= |p_k|^2\big(g_k(i) - \hat{g}_k(i|\mathcal{X}_{k,i-1})\big) + p_k^* w_k(i), \\
&= g_k(i) - \hat{g}_k(i|\mathcal{X}_{k,i-1}) + p_k^* w_k(i).
\end{aligned} \tag{2.80}$$

Now denote $\underline{\hat{\boldsymbol{G}}}(i|\mathcal{X}_{i-1}) = [\underline{\hat{\boldsymbol{g}}}_0(i|\mathcal{X}_{0,i-1}), \underline{\hat{\boldsymbol{g}}}_1(i|\mathcal{X}_{1,i-1}), \cdots, \underline{\hat{\boldsymbol{g}}}_{N-1}(i|\mathcal{X}_{N-1,i-1})]^T$, $\hat{\boldsymbol{g}}(i|\mathcal{X}_{i-1}) = [\hat{g}_0(i|\mathcal{X}_{0,i-1}), \hat{g}_1(i|\mathcal{X}_{1,i-1}), \cdots, \hat{g}_{N-1}(i|\mathcal{X}_{N-1,i-1})]^T$, $\tilde{\boldsymbol{e}}(i) = [\tilde{e}_0(i), \tilde{e}_1(i), \cdots, \tilde{e}_{N-1}(i)]^T$, and

$\boldsymbol{p}^* = [p_0^*, p_1^*, \cdots, p_{N-1}^*]^T$ it follows that

$$\begin{aligned}\hat{\underline{\boldsymbol{G}}}(i|\mathcal{X}_i) &= \hat{\underline{\boldsymbol{G}}}(i|\mathcal{X}_{i-1}) + \tilde{\boldsymbol{e}}(i)\tilde{\boldsymbol{q}}(i)^T, \\ &= \hat{\boldsymbol{G}}(i|\mathcal{X}_{i-1}) + \big(\boldsymbol{g}(i) - \hat{\boldsymbol{g}}(i|\mathcal{X}_{i-1}) + \boldsymbol{p}^* \cdot \boldsymbol{w}(i)\big)\tilde{\boldsymbol{q}}(i)^T. \end{aligned} \tag{2.81}$$

Assume $\{\hat{\underline{\boldsymbol{G}}}(i|\mathcal{X}_{i-1}), \hat{\boldsymbol{G}}(i|\mathcal{X}_{i-1})\} \in \text{span}(F_{N\times N_h})$, then

$$\begin{aligned}\breve{\boldsymbol{G}}(i) &= \boldsymbol{F}_{N\times N_h}\boldsymbol{F}_{N\times N_h}^H \hat{\underline{\boldsymbol{G}}}(i|\mathcal{X}_i), \\ &= \hat{\boldsymbol{G}}(i|\mathcal{X}_{i-1}) + \big(\boldsymbol{g}(i) - \hat{\boldsymbol{g}}(i|\mathcal{X}_{i-1})\big)\tilde{\boldsymbol{q}}(i)^T + \boldsymbol{F}_{N\times N_h}\boldsymbol{F}_{N\times N_h}^H\big(\boldsymbol{p}^* \cdot \boldsymbol{w}(i)\big)\tilde{\boldsymbol{q}}(i)^T \end{aligned} \tag{2.82}$$

Furthermore, denote $\tilde{\boldsymbol{w}}(i) = \boldsymbol{F}_{N\times N_h}\boldsymbol{F}_{N\times N_h}^H\big(\boldsymbol{p}^* \cdot \boldsymbol{w}(i)\big)$, it is easy to show that

$$\breve{\underline{\boldsymbol{g}}}_k(i) = \hat{\underline{\boldsymbol{g}}}_k(i|\mathcal{X}_{k,i-1}) + \big(g_k(i) - \hat{g}_k(i|\mathcal{X}_{k,i-1})\big)\tilde{\boldsymbol{q}}(i) + \tilde{w}_k(i)\tilde{\boldsymbol{q}}(i), \tag{2.83}$$

and $\text{E}[\tilde{w}_k(i)\tilde{w}_k(i)^*] = \frac{N_h}{N}\sigma_w^2$. Therefore, the channel refinement process reduces the noise energy by a factor $\frac{N_h}{N}$. As a result, in order to calculate the new lower bound, we only need to scale σ_w^2 in (2.72) and (2.73) by the factor $\frac{N_h}{N}$. Meanwhile, we need to update (2.76) to comply with this result.

CHAPTER 3

FREQUENCY-DOMAIN TURBO EQUALIZATION FOR DIGITAL TV TRANSMISSION SYSTEMS

3.1 Introduction

The performance of ATSC [2] digital television (DTV) receivers has been steadily increasing over the last decade [73, 74]. Receivers have become increasingly reliable in difficult channel conditions, such as indoor reception in urban settings, where dense multipath can heavily impair the transmitted signal. The current state-of-the-art ATSC reception scheme employs decision feedback equalization (DFE) [75, 76]. To handle difficult channels, receiver complexity—in particular, the DFE filter length—has increased significantly. While first-generation ATSC receivers typically employed DFEs with 100 forward and 400 feedback taps, current-generation receivers typically employ DFEs with 500 forward and 600 feedback taps. If broadcasters adopt the use of repeaters and distributed transmitters to increase coverage [77], then further increases in filter lengths can be expected.

For channels with large delay spreads, which are commonly encountered in DTV applications, even linear equalization can be quite costly when implemented in the

time domain, as suggested by DFE filter lengths mentioned above. For such channels, it might be more effective to consider frequency-domain equalization (FDE), which leverages fast circular convolution via the FFT to drastically reduce the cost of implementing long filters [1]. Motivated by the success of IFDE/AFDCE algorithm proposed in Chapter 2, here we attempt to design the receiver in the same framework. As it turns out, for DTV applications, both the FDTE and AFDCE algorithms are further complicated by the ATSC's use of 8-ary TCVSB modulation. Thus, we propose a novel FDTE/AFDCE scheme suitable for non-CP TCVSB modulated transmission system. Through numerical comparisons, we find that the proposed scheme gives simultaneous performance and complexity gains over time-domain DFE reception plus FDLMS CE.

The chapter is organized as follows. Section 3.2 and section 3.3 briefly describe TCVSB modulation and the communication system model. Section 3.4 and section 3.5 extends the proposed IFDE and AFDCE algorithms in Chapter 2 to adapt to TCVSB modulation. Section 3.6 presents numerical results. Finally, section 3.7 concludes.

3.2 Trellis Coded Vestigial Sideband Modulation

North American DTV transmission systems adopt TCVSB modulation for both power and spectrum efficiency. Here we briefly describe its properties for better understanding of the challenges for receiver design.

The VSB spectrum in a 6-MHz channel is depicted in Fig. 3.1 [2]. The digital VSB spectrum is flat throughout most of the channel due to the noise-like feature of randomized data. The RF peaks can only be described in terms of a statistical

cumulative distribution function [78] and remains within 6.3dB of the total in-band average power for 99.9% of time. As can be seen from Fig. 3.1, the half-power frequencies of the VSB signal are 5.38MHz apart with two steep transition bands (each of 0.62MHz wide). The 8-VSB modulation is implemented by a root raised cosine (RRC) pulse shaping filter (PSF) with roll-off factor $\beta = 0.115$ and the composite pulse waveform $b(t)$ is raised cosine (RC) waveform. These waveforms are illustrated in Fig. 3.2 and Fig. 3.3. Detailed time and frequency domain representation of these filters are provided in Appendix 3.B. As implied in Fig. 3.4, about half spectrum of the received signal will be suppressed by the PSFs, therefore special treatment is needed to restore the transmitted symbols as illustrated in section 3.4.

The transmission system employs a two-thirds rate (R=2/3) trellis code [2]. One input bit is encoded into two output bits using a one-half rate convolutional code while the other input bit is precoded. Standard 4-state optimal Ungerboeck code is used for the encoding. The signaling waveform used with the trellis code is an 8-level (3 bit) real valued constellation. The transmitted signal is referred to as 8-VSB. In ATSC standard, trellis code intrasegment interleaving shall be used, where twelve identical trellis encoders and precoders operating on interleaved data symbols. Fig. 3.5 shows the trellis code utilizes the 4-state feedback encoder as well as the precoder and the symbol mapper. The trellis coded modulation (TCM) complicates the soft information generation process for the Turbo equalization, which is discussed in details in section 3.4.2.

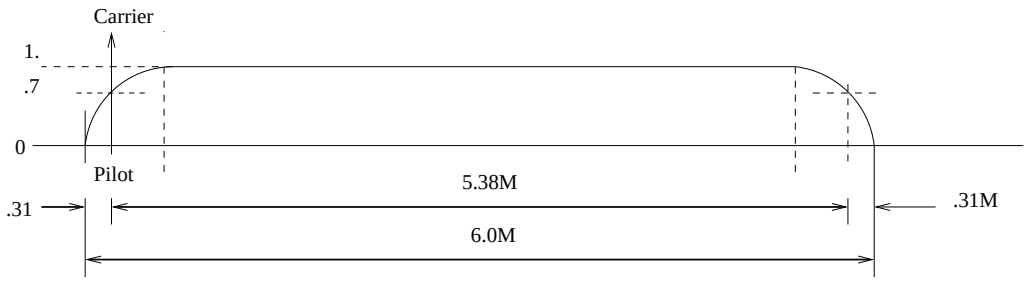

Figure 3.1: The spectrum of VSB modulated signal [2].

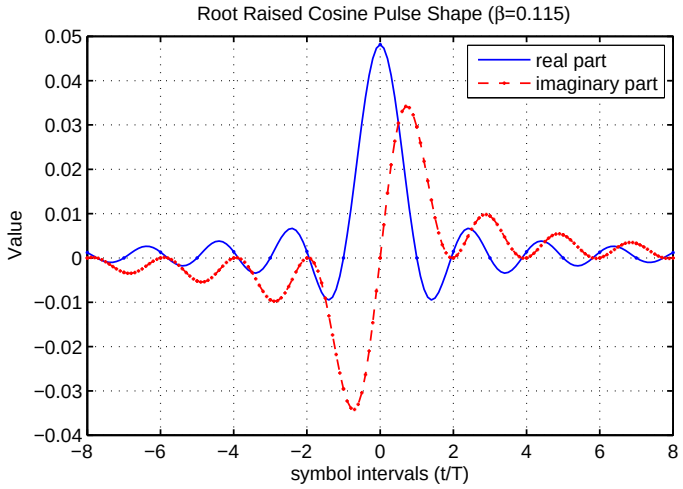

Figure 3.2: 8-VSB root raised cosine pulse shape .

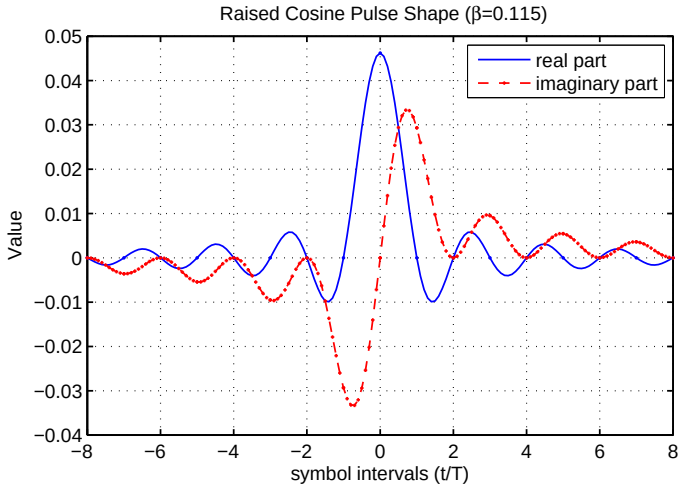

Figure 3.3: 8-VSB raised cosine pulse shape .

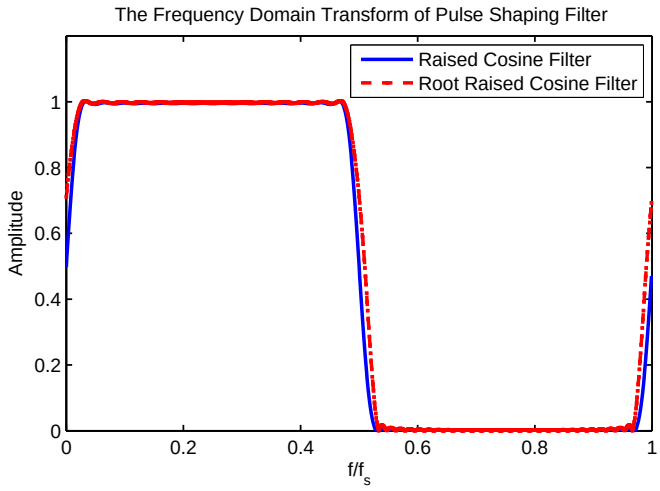

Figure 3.4: frequency-domain transform of root raised cosine and raised cosine pulse shapes.

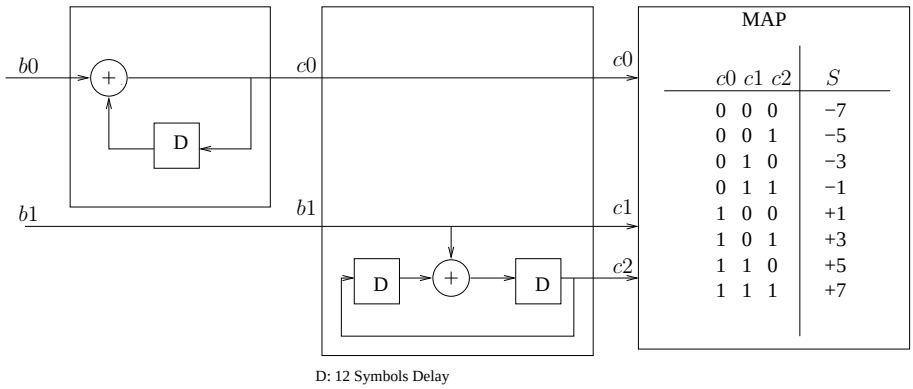

Figure 3.5: Trellis encoder, precoder, and symbol mapper [2].

3.3 System Model

For terrestrial DTV transmission, the normalized Doppler spread is relatively small, therefore, the composite channel impulse response with pulse shaping filters included can be defined as in (1.9)

$$h(t,\tau) = \sum_{\ell=1}^{N_\ell} c_\ell(t) b(\tau - \tau_\ell), \tag{3.1}$$

where the PSF $a(t)$ is the Hermitian symmetric RRC pulse such that $a(t) = a^*(-t)$, and $b(t)$ is the RC pulse. Similar as in (1.11), the received signal $r(t)$ is defined as

$$r(t) = h(t,\tau) * s_T(t) + \nu(t) = \int \sum_{\ell=1}^{N_\ell} c_\ell(t) b(\tau - \tau_\ell) \sum_k s_k \delta(t - \tau - kT) d\tau + \nu(t) \tag{3.2}$$

$$= \sum_{\ell=1}^{N_\ell} \sum_k s_k c_\ell(t) b(t - kT - \tau_\ell) + \nu(t), \tag{3.3}$$

where $\nu(t) = \mu(t) * a^*(t)$ and $\mu(t)$ is the AWGN noise. Due to the impact of PSF $a(t)$, $\nu(t)$ is a colored additive Gaussian noise, however in this chapter we would still treat it as AWGN similar as in [10, 11], and extension to the colored noise case can be done with a little bit more efforts.

The equivalent discrete time base-band transmission model can be obtained through sampling with period T. Define $r_n = r(nT)$, $c_{n,\ell} = c_\ell(nT)$, $b_{l,\ell} = b(lT - T_\ell)$ and $\nu_n = \nu(nT)$, the discrete time model is given by

$$h_{n,l} = \sum_{\ell=1}^{N_\ell} c_{n,\ell} b_{l,\ell}, \tag{3.4}$$

$$r_n = \sum_{l=0}^{L} h_{n,l} s_{n-l} + \nu_n, \tag{3.5}$$

where ν_n is assumed to be circular white Gaussian noise (CWGN) with mean zero and variance σ^2. Since two real-valued observations are made for every real-valued symbol s_n, one might consider (3.5) to be oversampled by a factor of two.

For our FDTE scheme, we assume block-wise processing with block length N. In fact, we focus on overlapping blocks, with block interval $N_d < N$. Furthermore, we assume that the channel is time-invariant over the duration of a single block. For convenience, we define the block-based quantities $r_n(i) = r_{iN_D+n}$, $h_l(i) = h_{iN_d+N/2,l}$, $s_n(i) = s_{iN_d+n}$, and $\nu_n(i) = \nu_{iN_d+n}$, and their vector counterparts $\boldsymbol{r}(i) := [r_0(i), r_1(i), \ldots, r_{N-1}(i)]$, $\boldsymbol{s}(i) := [s_0(i), s_1(i), \ldots, s_{N-1}(i)]$, $\boldsymbol{h}(i) := [h_0(i), \ldots, h_L(i), 0, \cdots, 0]$, and $\boldsymbol{\nu}(i) := [\nu_0(i), \nu_1(i), \ldots, \nu_{N-1}(i)]$. Thus, the signal received during the i-th block can be expressed as

$$r_n(i) = \begin{cases} \nu_n(i) + \sum_{l=0}^{n} h_l(i) s_{n-l}(i) + \sum_{l=n+1}^{L} h_l(i) s_{<n-l>_N}(i-1), & 0 \le n < L, \\ \nu_n(i) + \sum_{l=0}^{L} h_l(i) s_{n-l}(i), & L \le n < N, \end{cases} \tag{3.6}$$

where $< n >_N$ denotes n modulo N. Note that the samples $\{r_n(i)\}_{n=0}^{L-1}$ contain IBI from $\boldsymbol{s}(i-1)$.

The receiver structure and major processing steps for the DTV transmission system are exactly the same as those described in Section 2.3, while the implementation of FDTE and AFDCE modules are different due to the TCVSB modulation of DTV signal. For simplicity, here we only describe the special aspect of proposed FDTE and AFDCE algorithms, while directly refer to Section 2.3 for details of other similar modules, such as IBI cancellation and CPR, etc.

We assume that CP restoration has been perfectly executed, so that $\boldsymbol{y}(i)$ can be considered as a noise-corrupted output of a circular convolution between the channel $\boldsymbol{h}(i)$ and the transmitted symbols $\boldsymbol{s}(i)$. Assuming perfect CPR, the time-domain system model can be rewritten in matrix form as

$$\boldsymbol{y}(i) \;=\; \mathcal{C}(\boldsymbol{h}(i))\boldsymbol{s}(i) + \boldsymbol{\nu}(i), \tag{3.7}$$

where $\mathcal{C}(\boldsymbol{h}(i))$ denotes the circulant matrix with first column $\boldsymbol{h}(i)$. Taking the discrete Fourier transform (DFT) of both sides of (3.7), we obtain

$$\boldsymbol{x}(i) \;=\; \mathcal{D}(\boldsymbol{g}(i))\boldsymbol{t}(i) + \boldsymbol{w}(i), \tag{3.8}$$

where $\boldsymbol{x}(i)$, $\boldsymbol{t}(i)$ and $\boldsymbol{w}(i)$ denote DFTs of $\boldsymbol{y}(i)$, $\boldsymbol{s}(i)$ and $\boldsymbol{\nu}(i)$, respectively, and $\boldsymbol{g}(i) = \sqrt{N}\boldsymbol{F}\boldsymbol{h}(i)$. We refer to the elements in $\boldsymbol{t}(i)$ as *virtual subcarriers*.

3.4 Frequency-Domain Turbo Equalization

3.4.1 MMSE Estimation of Virtual Subcarriers

For notational brevity, the symbol index (i) will be suppressed for the remainder of this section. In our VSB model, the time-domain symbols $\boldsymbol{s}$ are real valued, so that the virtual subcarriers exhibit conjugate symmetry, i.e.,

$$t_n \;=\; \begin{cases} t_n^*, & n \in \{0, \frac{N}{2}\} \\ t_{N-n}^*, & n \in \{1, 2, \ldots, \frac{N}{2} - 1\} \end{cases}. \tag{3.9}$$

Using this fact, (3.8) can be rewritten with $\underline{\boldsymbol{t}} \in \mathbb{C}^{N/2+1}$:

$$\underline{\boldsymbol{x}} \;=\; \boldsymbol{\mathcal{H}}\underline{\boldsymbol{t}} + \underline{\boldsymbol{w}}, \tag{3.10}$$

$$\underline{t}_n \;:=\; t_n \quad n \in \{0, \ldots, \frac{N}{2}\} \tag{3.11}$$

$$\boldsymbol{\mathcal{H}} \;=\; \begin{bmatrix} g_0 & 0 & 0 & \cdots & 0 \\ 0 & g_1 & 0 & \cdots & 0 \\ 0 & g_{N-1}^* & 0 & \cdots & 0 \\ 0 & \vdots & \ddots & \ddots & 0 \\ \vdots & \cdots & 0 & g_{N/2-1} & \vdots \\ 0 & \cdots & 0 & g_{N/2+1}^* & 0 \\ 0 & \cdots & 0 & 0 & g_{N/2} \end{bmatrix} \tag{3.12}$$

where

$$\underline{x}_n := \begin{cases} x_n & n = 0 \\ x^*_{N-k} & n = 2k \quad k \in \{1, \ldots, \frac{N}{2} - 1\} \\ x_{k+1} & n = 2k+1 \quad k \in \{0, \ldots, \frac{N}{2} - 2\} \\ x_{\frac{N}{2}} & n = N-1 \end{cases} \tag{3.13}$$

$$\underline{w}_n := \begin{cases} w_n & n = 0 \\ w^*_{N-k} & n = 2k \quad k \in \{1, \ldots, \frac{N}{2} - 1\} \\ w_{k+1} & n = 2k+1 \quad k \in \{0, \ldots, \frac{N}{2} - 2\} \\ w_{\frac{N}{2}} & n = N-1. \end{cases} \tag{3.14}$$

Essentially, (3.10) removes the redundancy inherent in the VSB system model (3.8). We note that, with bandlimited VSB pulse shapes, some entries in $\underline{\boldsymbol{x}}$ may contain little signal energy. To reduce complexity, these elements could be ignored when estimating $\underline{\boldsymbol{t}}$. Doing so would require only that certain rows be omitted from $\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{w}}$, and $\boldsymbol{\mathcal{H}}$.

We use a linear MMSE technique to estimate the virtual subcarrier vector $\underline{\boldsymbol{t}}$, where we incorporate prior statistics on $\underline{\boldsymbol{t}}$ (i.e., mean and covariance) calculated from the MAP decoder outputs during the previous iteration. To simplify equalization task, similar as in (2.11), we approximate $\mathrm{E}\{(\underline{\boldsymbol{t}} - \underline{\bar{\boldsymbol{t}}})(\underline{\boldsymbol{t}} - \underline{\bar{\boldsymbol{t}}})^H\}$ by the diagonal matrix $\mathcal{D}(\boldsymbol{v}_{\underline{t}})$, where $\boldsymbol{v}_{\underline{t}} := \mathrm{diag}\big(\mathrm{E}\left[(\underline{\boldsymbol{t}} - \underline{\bar{\boldsymbol{t}}})(\underline{\boldsymbol{t}} - \underline{\bar{\boldsymbol{t}}})^H\right]\big)$. Therefore, the MMSE estimate of $\underline{\boldsymbol{t}} \in \mathbb{C}^{N/2+1}$ can be computed as

$$\underline{\hat{\boldsymbol{t}}} = \underline{\bar{\boldsymbol{t}}} + \mathcal{D}(\boldsymbol{v}_{\underline{t}}) \underbrace{\boldsymbol{\mathcal{H}}^H \big(\boldsymbol{\mathcal{H}}\, \mathcal{D}(\boldsymbol{v}_{\underline{t}}) \boldsymbol{\mathcal{H}}^H + \sigma^2 \mathbf{I_N}\big)^{-1}}_{:=\boldsymbol{G}} (\underline{\boldsymbol{x}} - \boldsymbol{\mathcal{H}} \underline{\bar{\boldsymbol{t}}}) \tag{3.15}$$

where $\underline{\bar{\boldsymbol{t}}} := \mathrm{E}\{\underline{\boldsymbol{t}}\}$. For the first iteration, we set $\underline{\bar{\boldsymbol{t}}}(i) = \mathbf{0}$ and $\boldsymbol{v}_{\underline{t}}(i) = \mathbf{1}$. Due to the sparse structure of $\boldsymbol{\mathcal{H}}$, (3.15) can be computed via

$$\hat{\underline{t}}_k = \begin{cases} \dfrac{g^*_k x_k + \frac{\sigma^2}{v_{\underline{t}_k}} \bar{\underline{t}}_k}{|g_k|^2 + \frac{\sigma^2}{v_{\underline{t}_k}}}, & k = 0,\ \frac{N}{2} \\ \dfrac{g^*_k x_k + g_{N-k} x^*_{N-k} + \frac{\sigma^2}{v_{\underline{t}_k}} \bar{\underline{t}}_k}{|g_k|^2 + |g_{N-k}|^2 + \frac{\sigma^2}{v_{\underline{t}_k}}}, & 1 \le k \le \frac{N-1}{2} \end{cases} \tag{3.16}$$

where $v_{\underline{t}_k} := [\boldsymbol{v}_{\underline{t}}]_k$. From $\hat{\underline{\boldsymbol{t}}}$, we reconstruct $\hat{\boldsymbol{t}}$ via (3.9) and (3.11).

3.4.2 Generation of MAP Inputs

The soft information that is passed to the MAP decoder is computed from the conditional probabilities $\{p(\hat{s}_n|s_n = s)\}_{s\in\mathcal{S}}$, where $\mathcal{S}$ denotes the symbol alphabet. Here we describe how these conditional probabilities are generated from the equalizer outputs, and how they are passed to the decoder.

Assuming Gaussian-distributed symbol estimation error,

$$p(\hat{s}_n|s_n = s) = \frac{1}{\sqrt{2\pi\sigma_{n,s}^2}} \exp\left(-\frac{(\hat{s}_n - u_{n,s})^2}{2\sigma_{n,s}^2}\right) \tag{3.17}$$

$$u_{n,s} := \mathrm{E}\left[\hat{s}_n|s_n = s\right] \tag{3.18}$$

$$\sigma_{n,s}^2 := \mathrm{var}\left[\hat{s}_n|s_n = s\right]. \tag{3.19}$$

It is shown in Appendix 3.A that $u_{n,s}$ and $\sigma_{n,s}^2$ can be calculated as

$$u_{n,s} = \bar{s}_n + \frac{(s - \bar{s}_n)}{N} \sum_{k=0}^{N-1} d_k \tag{3.20}$$

$$\sigma_{n,s}^2 \approx \frac{1}{N} \sum_{k=0}^{N-1} (|d_k|^2 \tilde{v}_n + \sigma^2 b_k) \tag{3.21}$$

where $\bar{s}_n := \mathrm{E}\{s_n\}$, and d_k, b_k, $\tilde{v}_n$ are defined in Appendix 3.A.

With TCM, a subset of the bits that determine each symbol are left uncoded, and this influences the way that soft information is passed to and from the MAP decoder. Let us represent $s_n \in \mathcal{S}$ via $s_n \equiv [\boldsymbol{c}_n, \boldsymbol{b}_n]$, where $\boldsymbol{c}_n \in \{0,1\}^m$ and $\boldsymbol{b}_n \in \{0,1\}^{\tilde{m}}$ are vectors of coded and uncoded bits, respectively. Since the ATSC standard does not employ bit-level random interleaving, we will not assume that the bits in $\boldsymbol{c}_n$ are

independent. As a result, we pass $\{p(\hat{s}_n|\boldsymbol{c}_n = \boldsymbol{c}), \forall \boldsymbol{c}\}$ to the MAP decoder, where

$$\begin{aligned} & p(\hat{s}_n|\boldsymbol{c}_n = \boldsymbol{c}) \\ &= \sum_{\boldsymbol{b}} p(\hat{s}_n|\boldsymbol{c}_n = \boldsymbol{c}, \boldsymbol{b}_n = \boldsymbol{b}) P(\boldsymbol{b}_n = \boldsymbol{b}|\boldsymbol{c}_n = \boldsymbol{c}) && (3.22) \\ &= \sum_{\boldsymbol{b}} p(\hat{s}_n|\boldsymbol{c}_n = \boldsymbol{c}, \boldsymbol{b}_n = \boldsymbol{b}) 2^{-\tilde{m}} && (3.23) \\ &= 2^{-\tilde{m}} \sum_{s \in \mathcal{S}(\boldsymbol{c})} p(\hat{s}_n|s_n = s), && (3.24) \end{aligned}$$

and where $\mathcal{S}(\boldsymbol{c})$ denotes the subset of $\mathcal{S}$ corresponding to bits $\boldsymbol{c}$.

3.4.3 Update of Virtual Subcarrier Statistics

MAP decoding yields the posterior probabilities $\{P(\boldsymbol{c}_n = \boldsymbol{c}|\hat{\boldsymbol{s}}), \forall \boldsymbol{c}\}$, which can be combined with information on the uncoded bits to update the symbol means and variances for use in the next turbo iteration. With the correspondence $s \equiv [\boldsymbol{c}, \boldsymbol{b}]$, we have

$$\begin{aligned} P(s|\hat{\boldsymbol{s}}) &= P(\boldsymbol{b}|\boldsymbol{c}, \hat{\boldsymbol{s}}) P(\boldsymbol{c}|\hat{\boldsymbol{s}}) && (3.25) \\ &= \frac{p(\hat{\boldsymbol{s}}|\boldsymbol{b}, \boldsymbol{c}) P(\boldsymbol{b}|\boldsymbol{c})}{p(\hat{\boldsymbol{s}}|\boldsymbol{c})} P(\boldsymbol{c}|\hat{\boldsymbol{s}}) && (3.26) \\ &= \frac{p(\hat{\boldsymbol{s}}|\boldsymbol{b}, \boldsymbol{c}) P(\boldsymbol{b}|\boldsymbol{c})}{\sum_{\boldsymbol{b}'} p(\hat{\boldsymbol{s}}|\boldsymbol{b}', \boldsymbol{c}) P(\boldsymbol{b}'|\boldsymbol{c})} P(\boldsymbol{c}|\hat{\boldsymbol{s}}) && (3.27) \\ &= \frac{p(\hat{\boldsymbol{s}}|s)}{\sum_{s' \in \mathcal{S}(\boldsymbol{c})} p(\hat{\boldsymbol{s}}|s')} P(\boldsymbol{c}|\hat{\boldsymbol{s}}), && (3.28) \end{aligned}$$

where we used the shorthand $P(\boldsymbol{b}|\boldsymbol{c}, \hat{\boldsymbol{s}}) = P(\boldsymbol{b}_n = \boldsymbol{b}|\boldsymbol{c}_n = \boldsymbol{c}, \hat{\boldsymbol{s}})$. For (3.28), we assumed that $P(\boldsymbol{b}_n = \boldsymbol{b}|\boldsymbol{c}_n = \boldsymbol{c})$ is uniform over $\boldsymbol{b}$.

The posteriors $\{P(s_n = s|\hat{\boldsymbol{s}})\}_{s \in \mathcal{S}}$ from (3.28) are then used to update the mean and variance of s_n as follows.

$$\begin{aligned} \bar{s}_n &:= \mathrm{E}\left[s_n|\hat{\boldsymbol{s}}\right] = \sum_{s \in \mathcal{S}} s P(s_n = s|\hat{\boldsymbol{s}}) && (3.29) \\ v_{s_n} &:= \mathrm{var}\left[s_n|\hat{\boldsymbol{s}}\right] = \sum_{s \in \mathcal{S}} |s - \bar{s}_n|^2 P(s_n = s|\hat{\boldsymbol{s}}). && (3.30) \end{aligned}$$

Assuming that $\{s_n\}$ are uncorrelated, the mean and variance of the virtual subcarriers $\{t_n\}$ become

$$\begin{aligned} \bar{\boldsymbol{t}} &:= \mathrm{E}\{\boldsymbol{t}\} = \boldsymbol{F}\bar{\boldsymbol{s}} && (3.31)\\ \boldsymbol{v}_t &:= \mathrm{diag}(\mathrm{E}\left[(\boldsymbol{t}-\bar{\boldsymbol{t}})(\boldsymbol{t}-\bar{\boldsymbol{t}})^H\right]) && (3.32)\\ &= \mathrm{diag}(\boldsymbol{F}\,\mathcal{D}(\boldsymbol{v}_s)\boldsymbol{F}^H). && (3.33) \end{aligned}$$

Since (3.33) implies that all the elements in $\boldsymbol{v}_t$ are identical, the variance calculation can be simplified to

$$v_{t_k} = \frac{1}{N}\sum_{n=0}^{N-1} v_{s_n}, \quad \forall k. \tag{3.34}$$

3.5 Modified Adaptive Frequency-Domain Channel Estimation

In this section, we modify the AFDCE algorithm proposed in Chapter 2 to exploit the property of VSB modulation. Though (2.8) and (3.8) appear the same at the first glance, there is a big difference between them due to VSB modulation. In both cases, $\boldsymbol{g}(i)$ is the FFT of $\boldsymbol{h}(i)$, however, in (2.8), we assume that $\mathrm{E}[h_l(i)h_{l+p}(i+q)] = \sigma^2_{h_l}\rho_q\delta_p$, while in (3.8), $h_{n,l} = \sum_{\ell=0}^{N_\ell} c_{n,\ell}b_{l,\ell}$ and we assume $\mathrm{E}[c_\ell(i)c_{\ell+p}(i+q)] = \sigma^2_{c_\ell}\rho_q\delta_p$. Therefore, the influence of the composite PSF $\{b_{l,\ell}\}$ should be taken into account in designing the channel estimator.

Consider the CIR $h(t,\tau)$ given by equation (3.1). Taking the Fourier transform on both sides w.r.t the τ variable, we obtain

$$g(t,f) = C(t,f)B(f), \tag{3.35}$$

where $g(t,f)$ and $B(f)$ are continuous Fourier transform of $g(t,\tau)$ and $b(\tau)$ respectively, and $C(t,f) = \sum_\ell c_\ell(t)e^{-j2\pi f\tau_\ell}$. Evaluating $g(t,f)$ at discrete time and frequency points and defining $g_k(i) = g(iNT_s, \frac{k}{N}f_s)$, $\mathsf{c}_k(i) = \mathsf{c}(iNT_s, \frac{k}{N}f_s)$, $\mathsf{b}_k = \mathsf{b}(\frac{k}{N}f_s)$, where f_s is the sampling frequency, we obtain

$$g_k(i) = \mathsf{b}_k\mathsf{c}_k(i) \tag{3.36}$$

Substitute (3.36) into (3.8), we can set up the state space model for VSB modulated system as

$$\underline{\mathbf{c}}_k(i) = \boldsymbol{A}\underline{\mathbf{c}}_k(i-1) + \eta_k(i) \tag{3.37}$$

$$x_k(i) = \mathsf{b}_k\hat{t}_k(i)\mathsf{c}_k(i) + w_k(i) \tag{3.38}$$

where $\underline{\mathbf{c}}_k(i) = [\mathsf{c}_k(i), \mathsf{c}_k(i-1), \cdots, \mathsf{c}_k(i-M+1)]^T$, $\underline{\boldsymbol{\eta}}_k(i) = [\eta_k(i), 0, \cdots, 0]^T$, $\eta_k(i) \sim CN(0, \sigma_\eta^2)$. Given the channel statistics, $\boldsymbol{A}$ and σ_η^2 can be obtained via the Yule-Walker method [70].

Then it follows that the Kalman filtering process is carried out through:

$$\boldsymbol{q}_k(i) = \boldsymbol{P}_k(i)\mathsf{b}_k^*\hat{t}_k(i)^*\boldsymbol{i}_1\left(\mathsf{b}_k\hat{t}_k(i)\boldsymbol{i}_1^H\boldsymbol{P}_k(i)\boldsymbol{i}_1\hat{t}_k^*(i)\mathsf{b}_k^* + \sigma_v^2(i)\right)^{-1} \tag{3.39}$$

$$e_k(i) = x_k(i) - \mathsf{b}_k\hat{t}_k(i)\boldsymbol{i}_1^H\hat{\underline{\mathbf{c}}}_k(i|\mathcal{X}_{k,i-1}) \tag{3.40}$$

$$\hat{\underline{\mathbf{c}}}_k(i|\mathcal{X}_{k,i}) = \hat{\underline{\mathbf{c}}}_k(i|\mathcal{X}_{k,i-1}) + e_k(i)\boldsymbol{q}_k(i) \tag{3.41}$$

$$\hat{\underline{\mathbf{c}}}_k(i+1|\mathcal{X}_{k,i}) = \boldsymbol{A}\hat{\underline{\mathbf{c}}}_k(i|\mathcal{X}_{k,i}) \tag{3.42}$$

$$\boldsymbol{P}_k(i+1) = \boldsymbol{A}(\mathbf{I} - \boldsymbol{q}_k(i)\mathsf{b}_k\hat{t}_k(i)\boldsymbol{i}_1^H)\boldsymbol{P}_k(i)\boldsymbol{A}^T + \mathcal{D}(\boldsymbol{\sigma}_\eta^2) \tag{3.43}$$

$$\hat{g}_k(i+1) = \mathsf{b}_k\boldsymbol{i}_1^H\hat{\underline{\mathbf{c}}}_k(i+1|\mathcal{X}_{k,i}), \tag{3.44}$$

Finally across-tone CE refinement and AR model coefficients adaptation can be implemented on $\{\hat{\underline{\mathbf{c}}}_k(i|\mathcal{X}_{k,i})\}$ similar as in Chapter 2.

In the following, we describe a technique to reduce the computational cost of AFDCE by leveraging the property of VSB modulation. As can be seen from Fig. 3.4, almost half spectrum of the channel is suppressed by the PSF, therefore we can ignore those channel coefficients with small energy to simplify computation without sacrificing much performance. Here we examine the relationship between performance loss and the threshold which is used to decide whether the kth tap $\mathsf{c}_k(i)$ is set to zero directly or need to be estimated by AFDCE. Particulary, we focus on analyzing the initial state of Kalman filtering, which is equivalent to analyze a MMSE estimator, since a good start is the key to guarantee the success of joint IFDE/AFDCE for the subsequent symbols.

Given (3.38), the MMSE estimate $\hat{\mathsf{c}}_k(i)$ and estimation errors are defined as

$$\hat{\mathsf{c}}_k(i) = \frac{\mathrm{E}\{|\mathsf{c}_k(i)|^2\}\hat{t}_k^*(i)\mathsf{b}_k^*}{\mathsf{b}_k\hat{t}_k(i)\,\mathrm{E}\{|\mathsf{c}_k(i)|^2\}\hat{t}_k^*(i)\mathsf{b}_k^* + \sigma^2} x_k(i), \tag{3.45}$$

$$\mathrm{E}\{|\hat{\mathsf{c}}_k(i) - \mathsf{c}_k(i)|^2\} = \frac{\mathrm{E}\{|\mathsf{c}_k(i)|^2\}\sigma^2}{|\mathsf{b}_k|^2|\hat{t}_k(i)|^2\,\mathrm{E}\{|\mathsf{c}_k(i)|^2\} + \sigma^2}. \tag{3.46}$$

Note if $\hat{\mathsf{c}}_k(i)$ is set to 0, $\mathrm{E}\{|\hat{\mathsf{c}}_k(i) - \mathsf{c}_k(i)|^2\} = \mathrm{E}\{|\mathsf{c}_k(i)|^2\}$. Denote the extra MMSE ratio as κ, then it is clear that

$$\frac{\mathrm{E}\{|\mathsf{c}_k(i)|^2\}}{\frac{\mathrm{E}\{|\mathsf{c}_k(i)|^2\}\sigma^2}{|\mathsf{b}_k|^2|\hat{t}_k(i)|^2\,\mathrm{E}\{|\mathsf{c}_k(i)|^2\}+\sigma^2}} = 1 + \underbrace{\frac{|\mathsf{b}_k|^2|\hat{t}_k(i)|^2\,\mathrm{E}\{|\mathsf{c}_k(i)|^2\}}{\sigma^2}}_{:=\kappa} \tag{3.47}$$

Since we assume $\mathrm{E}\{|\mathsf{c}_k(i)|^2\} = 1$ and $\mathrm{E}\{|t_k(i)|^2\} = 1$ while $\sigma_s^2 = 1$, therefore $\kappa \propto \frac{|\mathsf{b}_k|^2}{\sigma^2}$. Fix κ, the threshold Γ is determined by $\Gamma = \kappa\sigma^2$. Therefore, only those channel coefficients $\{\mathsf{c}_k(i), k \in \mathbf{K}\}$ will be tracked through AFDCE, where $\mathbf{K} = \{k, |\mathsf{b}_k|^2 > \Gamma\}$, while other channel coefficients will be set to 0 directly. This technique reduces the computational cost roughly by half.

3.6 Numerical Results

In this section, we compare the performance and complexity of the proposed FDTE with that of the DFE-plus-Viterbi-decoding (DFE-VD) method proposed by Ariyavisitakul and Li [79] using the fast DFE filter update proposed by Al-Dhahir and Cioffi [80]. In the DFE-VD scheme, the (delayed) Viterbi estimates are fed to adequately-delayed DFE feedback taps, while sub-optimal symbol-by-symbol decisions are fed to the DFE feedback taps corresponding to shorter delays. In addition, we compare the performance and complexity of the modified AFDCE with FDLMS algorithm [70], which is widely used in tracking slow fading channels.

3.6.1 Simulation Setup

For performance comparison, we used the three propagation models summarized in Table 3.1. These were chosen similar to the ATSC R2.2 ensembles from [81]. Six paths were employed, each with a different delay, and with either a constant phase offset or a single-sided Doppler frequency spread of $f_d = 50$Hz. With the ATSC sampling rate $T_s^{-1} = 10.76$MHz, this corresponds to a normalized Doppler spread $f_d T_s = 0.000005$. The relative attenuations of the reflected paths vary among the three propagation models in Table 3.1; channel #1 is the least selective channel, #2 is the most time-selective, and #3 is the most frequency selective. To create the $\{h_{n,l}\}_{l=0}^{L}$, we generated propagation responses using Jakes method [82] and convolved them with the PSF, using an overall channel order of $L = 511$.

We assumed an 8-VSB modulated single-carrier system (i.e., no CP) that used rate-2/3 Ungerboeck coding with constraint length 3 [83]. For perfect CSI case, the receiver was assumed to have perfect knowledge of the CIR during the middle of each

N-length block. Otherwise, estimated CIR was used for symbol detection purpose. For FDTE/AFDCE, we used $N = 2048$ and $N_d = N/2$, and we reconstructed a CP of length L. Meanwhile we set the maximum iteration number to be 5. For DFE-VD/FDLMS, we updated the filter coefficients once every N_d symbols, and we used a feedforward filter of length $N_f = 2(L + 1)$ and a feedback filter of length L. The feedback filter length allows perfect post-cursor ISI cancellation, and the feedforward filter length was chosen so that further increases yielded little improvement in BER performance. The DFE-VD decoding delay was 30. For FDLMS, the step size was set to be 0.0005. For all the simulations, we averaged 20 realizations of 100 contiguous data blocks preceded by a pilot block (to initialize channel estimates).

3.6.2 Performance Assessment

Figure 3.6 shows the BER performance of FDTE and DFE-VD with perfect CSI (denoted by Perfect FDTE and Perfect DFE), FDTE/AHKCE and DFE-VD/FDLMS. From Fig. 3.6, we can see that, after 5 iterations, FDTE outperforms DFE-VD by 1dB (in SNR) approximately when perfect CSI is available. At BER= 10^{-4}, there is less than 1dB loss between FDTE/AHKCE and FDTE with perfect CSI, and FDTE/AHKCE achieves about 1dB gain over DFE-VD/FDLMS .

Figure 3.7 and Fig. 3.8 show the BER and steady state symbol estimation MSE performance of FDTE/AHKCE versus FDTE/FDLMS. From Fig. 3.7, we can see that FDTE/AHKCE performs better than FDTE/FDLMS in all the scenario, especially in the most time selective case. Figure 3.8 demonstrates that FDTE/AHKCE achieves lower steady state MSE than FDTE/FDLMS.

Figure 3.9 shows the symbol estimation MSE after CPR. Here CPR_H denotes CPR with perfect known CSI, CPR_Hhat denotes CPR with predicted channel output from AHKCE. Note both MSE measurements of CPR_H and CPR_Hhat are took right before the first iteration of FDTE algorithm, where the symbols located at the end of block are estimated through a simple FDE (it is equivalent to MMSE estimator described in section 3.4.1 without any priors). Since those symbol estimates are noisy, therefore the MSE of CPR are higher than 0.01 (noise variance at 20dB) at pass band of VSB filter. CPR_CE denotes CPR with channel estimates and symbol estimates right before the AFDCE. In this case, symbol estimates output from FDTE are much more reliable, therefore, the MSE of CPR are close to 0.01. As illustrated in Fig. 3.9, symbol estimation errors take a more important role in influencing the CPR performance than channel estimation error. As a result, we combat the CPR error before FDTE through block overlapping idea, while ignore the CPR error in AFDCE.

Table 3.2 specifies the cost to generate N_d symbol estimates for fast DFE-VD (with feedback filter length L) and for FDTE (per iteration). Meanwhile, it also specifies the cost to update channel estimates once per N_d symbols for AFDCE and FDLMS algorithm, where N_c denotes the number of active sub-carriers tracked by AFDCE. Figure 3.10 plots DFE-VD and FDTE complexity for the same design choices used in Fig. 3.6[6], i.e., FDTE with CP length L, $N = 4(L+1)$, $N_d = N/2$, and 5 iterations; and DFE-VD with $N_f = 2(L+1)$. We see that, when the channel order $L \geq 64$, the FDTE is cheaper to implement than the fast DFE-VD. Practical DTV receivers need to handle channels of order $L \approx 511$, in which case the FDTE is

[6]We assume that 1 division is equivalent to 10 multiplications when we generate those plots, which is a reasonable assumption when finite precision is applied in practical implementations.

Table 3.1: DTV Propagation Models.

path delay		-1.8μs	0μs	0.15μs	1.8μs	5.7μs	39.8μs
chan #1	gain	-8 dB	0 dB	-3 dB	-4 dB	-3 dB	-12 dB
	Doppler	125°	0°	80°	45°	50 Hz	90°
chan #2	gain	-8 dB	0 dB	-3 dB	-4 dB	-3 dB	-12 dB
	Doppler	50 Hz	0°	50 Hz	50 Hz	50 Hz	50 Hz
chan #3	gain	-3 dB	0 dB	-1 dB	-1 dB	-3 dB	-9 dB
	Doppler	125°	0°	80°	45°	50 Hz	90°

an order of magnitude cheaper than DFE-VD. Figure 3.11 compares the computational complexity of AFDCE versus FDLMS, where we assumed $N_c = N/2$ approximately. It can be seen that both curves demonstrate increasing slope of order $\log_2(N)$, while AFDCE is about 2 times more complex than FDLMS as N increase. However, combine Figure 3.10 with Figure 3.11, we can see that the computation complexity of FDTE/AFDCE scheme is an order-of-magnitude lower than DFE-VD/FDLMS scheme as $N \geq 512$, since the computation complexity of channel equalization is the dominating factor.

3.7 Conclusions

We presented a FDTE/AFDCE scheme suitable for TCVSB modulation, as is used in the ATSC DTV standard. Simulations show that it outperforms the fast DFE-VD/FDLMS approach while maintaining up to an order-of-magnitude lower complexity.

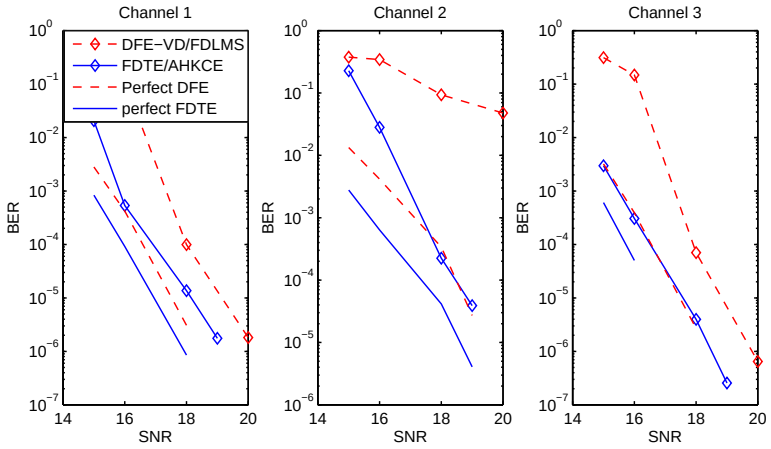

Figure 3.6: BER performance comparison for DFE-VD/FDLMS versus FDTE/AHKCE.

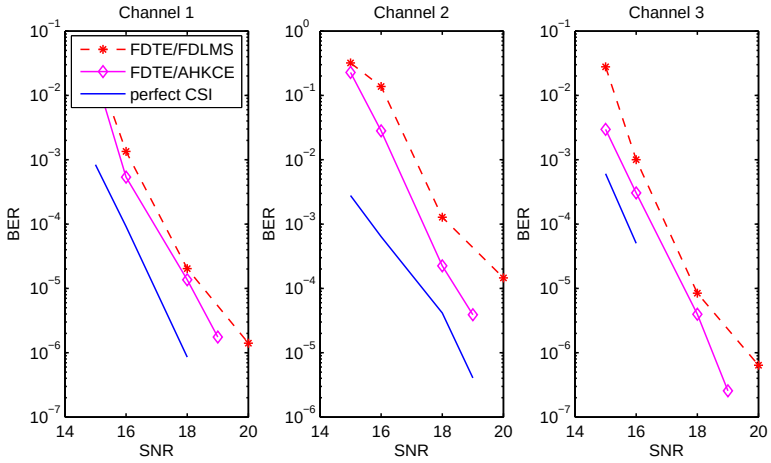

Figure 3.7: BER performance comparison for FDTE/AHKCE versus FDTE/FDLMS.

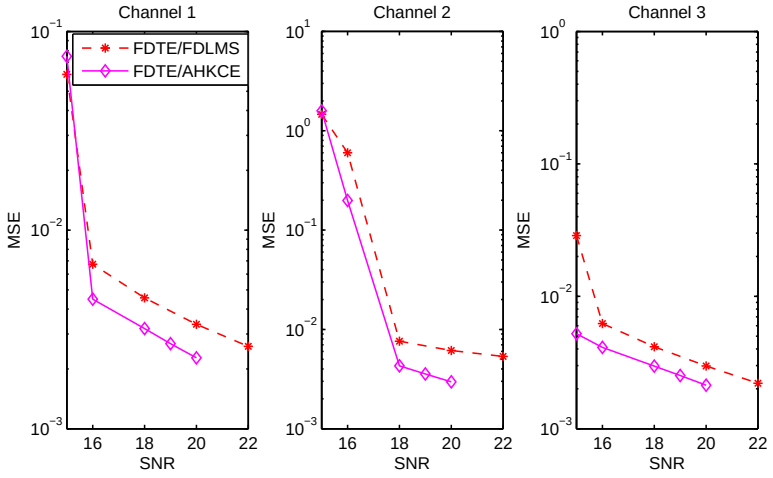

Figure 3.8: MSE performance comparison for FDTE/AHKCE versus FDTE/FDLMS.

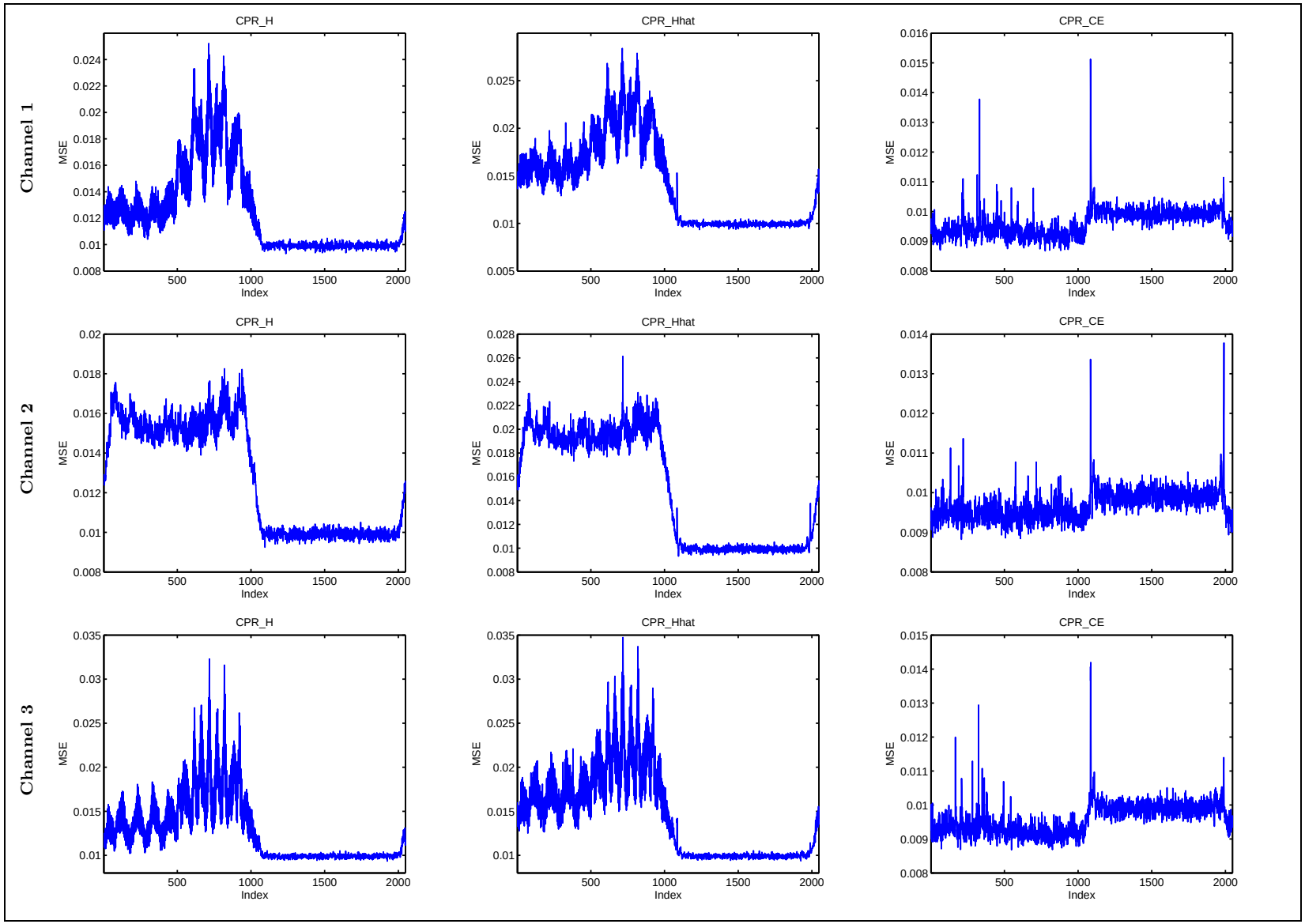

Figure 3.9: MSE of CPR with perfect CSI, estimated channels at SNR=20dB.

Table 3.2: Computational Complexity (per N_d symbols).

algorithm	real $\times$	real $\div$	exp	log
FDTE	$54.5N + 6N\log_2(N) + 11$	$6.5N + 3$	$12N$	$4N$
DFE-VD	$14N_fL + 30.5N_f + 0.5N_f^2$ $+2N_f + 2N_fN_D + LN_D$ $+8N_d - L^2 - 12L - 21$	$2N_f + N_d$	0	0
AFDCE (from ATCR) (from ATARMC)	$N_c(4M^2 + 12M + 19) + 6N\log_2(N)$ $+ 8MN\log_2(N)$ $+ M^2N_c + 7MN_c + \frac{2}{3}M^3 + 3M^2$	N_c 0 1	0	0
FDLMS	$14N\log_2(N) + 10N$	0	0	0

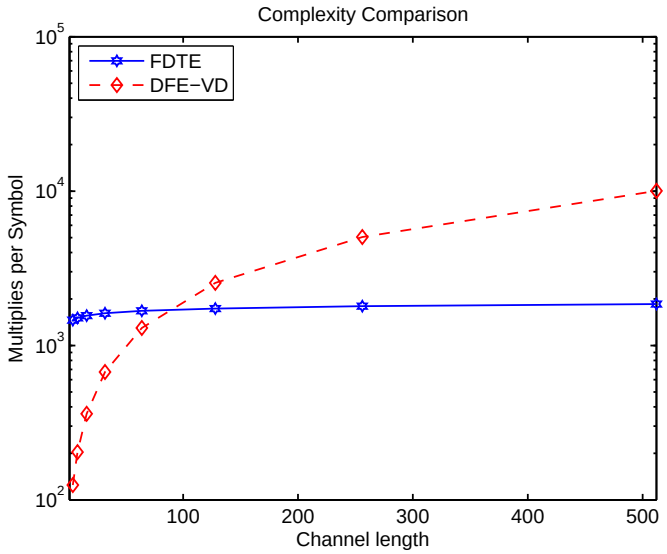

Figure 3.10: Computational complexity (per symbol).

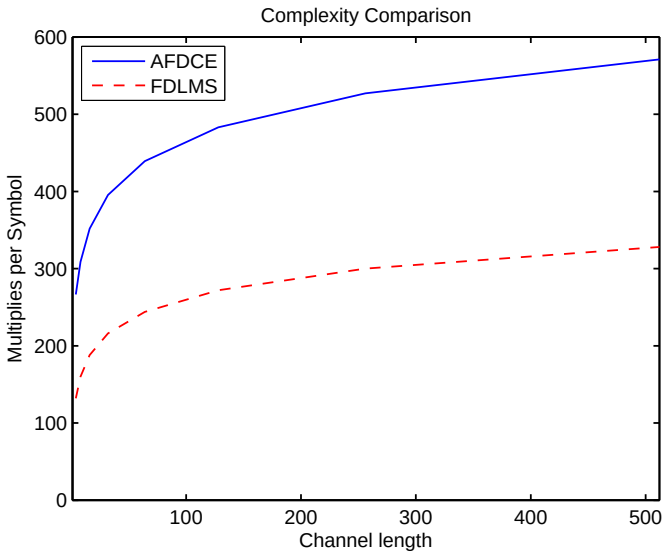

Figure 3.11: Computational complexity (per symbol).

Appendix

3.A Derivation of Conditional Mean and Variance

Symbol estimates $\boldsymbol{s}$ can be calculated by taking the inverse FFT of the virtual subcarriers,

$$\hat{\boldsymbol{s}} = \boldsymbol{F}^H \hat{\boldsymbol{t}}.$$

where by the conjugate symmetric property (3.9), $\hat{\boldsymbol{t}}$ is

$$\hat{\boldsymbol{t}} = \begin{bmatrix} \hat{\underline{\boldsymbol{t}}} \\ \boldsymbol{J}\hat{\underline{\boldsymbol{t}}}^* \end{bmatrix}, \quad \boldsymbol{J} := \left[\mathbf{0}_{(\frac{N}{2}-1)\times 1}\ \boldsymbol{\Pi}_{\frac{N}{2}-1}\ \mathbf{0}_{(\frac{N}{2}-1)\times 1}\right], \tag{3.48}$$

$$\boldsymbol{\Pi}_{\frac{N}{2}-1} := \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \cdots & \cdots & 0 \end{bmatrix}. \tag{3.49}$$

Observe the special structure of $\boldsymbol{G}$ and $\boldsymbol{\mathcal{H}}$, we can rewrite $\boldsymbol{J}\hat{\underline{\boldsymbol{t}}}^*$ as

$$\boldsymbol{J}\hat{\underline{\boldsymbol{t}}}^* = \boldsymbol{J}\bar{\underline{\boldsymbol{t}}}^* + \mathcal{D}(\boldsymbol{J}\boldsymbol{v}_{\underline{\boldsymbol{t}}})(\boldsymbol{J}\boldsymbol{G}^*)(\boldsymbol{x}^* - \boldsymbol{\mathcal{H}}^*\boldsymbol{J}^H\boldsymbol{J}\bar{\underline{\boldsymbol{t}}}^*). \tag{3.50}$$

Combine (3.48), (3.15) and (3.50), we rewrite $\hat{\boldsymbol{t}}$ as

$$\hat{\boldsymbol{t}} = \bar{\boldsymbol{t}} + \mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{\mathcal{H}}}\bar{\boldsymbol{t}}), \tag{3.51}$$

where

$$\tilde{\boldsymbol{G}} := \begin{bmatrix} \boldsymbol{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{J}\boldsymbol{G}^* \end{bmatrix}, \quad \tilde{\boldsymbol{\mathcal{H}}} := \begin{bmatrix} \boldsymbol{\mathcal{H}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{H}}^*\boldsymbol{J}^H \end{bmatrix}, \tag{3.52}$$

$$\bar{\boldsymbol{t}} = \begin{bmatrix} \bar{\underline{\boldsymbol{t}}} \\ \boldsymbol{J}\bar{\underline{\boldsymbol{t}}}^* \end{bmatrix} = \boldsymbol{F}\bar{\boldsymbol{s}}, \quad \tilde{\boldsymbol{x}} := \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{x}^* \end{bmatrix}. \tag{3.53}$$

It follows that

$$\hat{\boldsymbol{s}} = \bar{\boldsymbol{s}} + \boldsymbol{F}^H\,\mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{\mathcal{H}}}\bar{\boldsymbol{t}}), \tag{3.54}$$

$$\begin{aligned} \hat{s}_n &= \bar{s}_n + \boldsymbol{i}_n^H\boldsymbol{F}^H\,\mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}(\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{\mathcal{H}}}\bar{\boldsymbol{t}}) \\ &= \bar{s}_n + \boldsymbol{i}_n^H\boldsymbol{F}^H\,\mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}\tilde{\boldsymbol{\mathcal{H}}}\boldsymbol{F}(\boldsymbol{s} - \bar{\boldsymbol{s}}) + \boldsymbol{i}_n^H\boldsymbol{F}^H\,\mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}\tilde{\boldsymbol{w}}, \end{aligned} \tag{3.55}$$

where $\tilde{\boldsymbol{w}} = \begin{bmatrix} \boldsymbol{w}^T & \boldsymbol{w}^H \end{bmatrix}^T$. Therefore,

$$\begin{aligned}
\mathrm{E}\left[\hat{s}_n | s_n = s\right] &= \bar{s}_n + \boldsymbol{i}_n^H \boldsymbol{F}^H \underbrace{\mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}\tilde{\boldsymbol{\mathcal{H}}}}_{\mathcal{G}} \boldsymbol{F}\boldsymbol{i}_n (s - \bar{s}_n) \\
&= \bar{s}_n + \boldsymbol{i}_n^H \boldsymbol{F}^H \mathcal{G}\boldsymbol{F}\boldsymbol{i}_n (s - \bar{s}_n) \\
&= \bar{s}_n + \frac{s - \bar{s}_n}{N} \sum_{k=0}^{N-1} d_k, \qquad (3.56)
\end{aligned}$$

where

$$d_k = \begin{cases} \dfrac{|g_k|^2}{|g_k|^2 + \frac{\sigma^2}{v_{\underline{t}_k}}} & k \in \{0, N/2\}, \\ \dfrac{|g_k|^2 + |g_{N-k}|^2}{|g_k|^2 + |g_{N-k}|^2 + \frac{\sigma^2}{v_{\underline{t}_k}}}, & k \notin \{0, N/2\}. \end{cases} \qquad (3.57)$$

$$\mathrm{var}\left[\hat{s}_n | s_n = s\right] = \mathrm{E}\left[\left|\hat{s}_n - \mathrm{E}\left[\hat{s}_n | s_n = s\right]\right|^2\right] \qquad (3.58)$$

$$= \mathrm{E}\left[\left|\boldsymbol{i}_n^H \boldsymbol{F}^H \mathcal{G}\boldsymbol{F}\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s - \bar{s}_n)\big) + \boldsymbol{i}_n^H \boldsymbol{F}^H \mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}\tilde{\boldsymbol{w}}\right|^2\right] \qquad (3.59)$$

$$= \boldsymbol{i}_n^H \boldsymbol{F}^H \mathcal{G}\, \mathrm{E}\left[\boldsymbol{F}\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)^H \boldsymbol{F}^H\right] \mathcal{G}^H \boldsymbol{F}\boldsymbol{i}_n$$

$$+ \boldsymbol{i}_n^H \boldsymbol{F}^H \mathcal{D}(\boldsymbol{v_t})\tilde{\boldsymbol{G}}\, \mathcal{D}(\boldsymbol{\sigma}^2_{\tilde{\boldsymbol{w}}})\tilde{\boldsymbol{G}}^H \mathcal{D}(\boldsymbol{v_t})\boldsymbol{F}\boldsymbol{i}_n \qquad (3.60)$$

$$\approx \frac{1}{N} \sum_{k=0}^{N-1} (|d_k|^2 \tilde{v}_n + \sigma^2 b_k), \qquad (3.61)$$

where

$$b_k = \begin{cases} \dfrac{|g_k|^2}{(|g_k|^2 + \frac{\sigma^2}{v_{\underline{t}_k}})^2}, & k \in \{0, N/2\}, \\ \dfrac{|g_k|^2 + |g_{N-k}|^2}{(|g_k|^2 + |g_{N-k}|^2 + \frac{\sigma^2}{v_{\underline{t}_k}})^2}, & k \notin \{0, N/2\}, \end{cases} \qquad (3.62)$$

$$\tilde{v}_n = \frac{1}{N} \sum_{k \neq n} v_{s_k}, \qquad (3.63)$$

and $v_{s_k} := \mathrm{var}\{s_k\}$. Here we have applied the approximation that

$$E\left[\boldsymbol{F}\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)^H \boldsymbol{F}^H\right] \qquad (3.64)$$

$$\approx \mathcal{D}\left(\mathrm{diag}\Big(E\left[\boldsymbol{F}\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)\big(\boldsymbol{s} - \overline{\boldsymbol{s}} - \boldsymbol{i}_n(s_n - \bar{s}_n)\big)^H \boldsymbol{F}^H\right]\Big)\right). \qquad (3.65)$$

3.B 8-VSB Pule Shape

In this section we review the pulse shapes used by the 8-VSB TV systems [84]. The 8-VSB DTV transmitter uses a transmit pulse shaping filter with spectrum shifted root-raised cosine response. The transmit filter is denoted by $a(t)$. The receiver uses $a^*(-t)$ as the matched filter, which is optimal for AWGN channel.

Let T denote the symbol period and $1/T$ is the symbol rate, define

$$F_s = \frac{1}{2T}. \tag{3.66}$$

The roll-off factor for 8-VSB standard is $\beta = 0.115$.

3.B.1 Root-Raised Cosine Pulse

The time domain response of general RRC is defined as

$$a_{RRC}(t) = \frac{F_s}{1-(4\beta F_s t)^2}\left[(1-\beta)\frac{\sin(\pi(1-\beta)F_s t)}{\pi(1-\beta)F_s t} + \frac{4\beta}{\pi}\cos(\pi(1+\beta)F_s t)\right], \tag{3.67}$$

which has a symmetric spectrum around frequency 0Hz. The PSF $a(t)$ is $\frac{F_s}{2}$ shifted version of $a_{RRC}(t)$, which is defined as

$$a(t) = e^{j\pi F_s t} a_{RRC}(t). \tag{3.68}$$

It is straightforward to show that

$$a(t) = a^*(-t), \tag{3.69}$$

(3.69) means that the pulse shape $a(t)$ is Hermitian symmetric, therefore the receiver filter is the same as the transmit filer. The Fourier transform of $a(t)$ is:

$$A(f) = \begin{cases} 1 & \frac{\beta}{2}F_s \le f \le (1-\frac{\beta}{2})F_s \\ \cos(\frac{\pi}{2\beta F_s}[f-(1-\frac{\beta}{2})F_s]) & (1-\frac{\beta}{2})F_s \le f \le (1+\frac{\beta}{2})F_s \\ \cos(\frac{\pi}{2\beta F_s}[f-\frac{\beta}{2}F_s]) & -\frac{\beta}{2}F_s \le f \le \frac{\beta}{2}F_s \end{cases} \tag{3.70}$$

which implies that $a(t)$ has a roughly flat spectrum of width F_s and centers at $\frac{F_s}{2}$.

3.B.2 Raised Cosine Pulse

The convolution of the transmit and receive filter is denoted by $b(t)$ and the composite pulse shape is given by

$$b(t) = a(t) * a^*(-t) = a(t) * a(t). \tag{3.71}$$

The general RC filter is defined as

$$b_{RRC}(t) = F_s \frac{\sin(\pi F_s t)\cos(\pi\beta F_s t)}{\pi F_s t\big(1-(2\beta F_s t)^2\big)}. \tag{3.72}$$

$b(t)$ is $\frac{F_s}{2}$ shifted version of $b_{RC}(t)$, which is defined as

$$b(t) = e^{j\pi F_s t} b_{RRC}(t). \tag{3.73}$$

$b(t)$ has the Fourier transform:

$$\mathsf{b}(f) = \begin{cases} 1 & \frac{\beta}{2}F_s \le f \le (1-\frac{\beta}{2})F_s \\ 0.5\left[1+\cos(\frac{\pi}{\beta F_s}[f-(1-\frac{\beta}{2})F_s])\right] & (1-\frac{\beta}{2})F_s \le f \le (1+\frac{\beta}{2})F_s \\ 0.5\left[1+\sin(\frac{\pi}{\beta F_s}f)\right] & -\frac{\beta}{2}F_s \le f \le \frac{\beta}{2}F_s \end{cases} \tag{3.74}$$

CHAPTER 4

FREQUENCY-DOMAIN EQUALIZATION OF VERY FAST FADING FREQUENCY-SELECTIVE CHANNELS

4.1 Introduction

The FDE techniques discussed in previous chapters assume a delay-spread channel whose impulse response varies only moderately quickly. Some applications, however, have channels with more significant time variation, i.e., significant Doppler spread. For such doubly dispersive channels, the standard approach to FDE (i.e., pointwise multiplication in the frequency-domain) fails when the channel varies significantly over the FFT block duration. Essentially, the channel variation induces significant off-main-diagonal coefficients in the frequency domain channel matrix, which leads to ICI (virtual ICI) in multi-carrier (single-carrier) transmission systems. In response, several equalization schemes for doubly dispersed CP-OFDM have been proposed (see, e.g., [85–89] and the references within). While most of these schemes are computationally intensive, [88] maintains per-symbol processing complexity logarithmic in the block length, in keeping with the spirit of FDE. In addition, [88] exploits the finite-alphabet property of frequency-domain symbols, allowing its performance to surpass that of MMSE linear equalization. The CP-OFDM FDE scheme [88] was

extended to SCCP in [90]. Though the SCCP FDE scheme [90] is complicated by the fact that the finite-alphabet property resides in the time domain, it nevertheless maintains the desired logarithmic per-symbol processing complexity.

Though capable of FDE on doubly dispersive channels, the algorithms [85–88] and [90] require block-based transmissions with an adequate inter-block guard interval. While [89] does not require a guard, its complexity scaling properties restrict its application to channels with mild spreading. Thus, one might wonder: Is it possible to build a FDE algorithm for single-carrier *continuous-stream* modulation over doubly dispersive channels that exhibits logarithmic complexity scaling? If so, such an algorithm would present an efficient frequency-domain alternative to the time-domain equalization approaches that are commonly used in doubly selective single-carrier receivers (e.g., for North American terrestrial digital television [91] and underwater acoustic communication [92]).

In this chapter, we propose joint pilot-aided CE (PACE) and IFDE for the reception of a continuous finite-alphabet stream corrupted by a noisy and doubly dispersive channel. First, motivated by [44], we derive a MMSE channel estimator, which directly utilize the pilot symbols to estimate the significant frequency domain channel coefficients during the data symbols' transmission in between and then feeds those estimates to enable IFDE. Then, we present the IFDE. In brief, the algorithm first parses the received time-domain signals into blocks which are first windowed and then transformed into the frequency domain by an FFT. The window is designed so that both channel variations and the lack-of-CP manifest as a sparse frequency-domain response. A low-complexity serial technique is then applied to equalize the channel response in the presence of off-main-diagonal channel coefficients, and the output is

transformed back to the time domain, yielding soft symbol estimates. Using the finite-alphabet symbol property, reliability information on these soft estimates is computed for use in another round of (frequency-domain) equalization and interference cancellation. The time- and frequency-domain steps are alternated until the soft symbols estimates converge.

Simulation results demonstrate superb performance of IFDE algorithm with or without the influence of channel estimation error. When perfect CSI is available, our IFDE algorithm's performance is compared to that of the FIR-MMSE-DFE [93], a well-known benchmark, as well as to that of the matched filter bound (MFB) [94]—the "holy grail" of uncoded equalization. We find that our IFDE performs 1 dB worse than the MFB, and several dB better than the FIR-MMSE-DFE, over the SNR range of interest. When CSI is not available, the joint IFDE/CE scheme is compared with FIR-MMSE-DFE plus RLS based CE (RLS-CE) scheme, which is the state-of-art technique in tracking and equalizing fast time-varying channels [95]. It turns out that our IFDE/CE scheme can handle faster fading and perform better than FIR-MMSE-DFE/RLS-CE. In addition, we analyze the number of multiplications required by our IFDE and compare it to that of the FIR-MMSE-DFE updated using a fast algorithm. We find that our IFDE has complexity advantages over the FIR-MMSE-DFE for channels of a reasonable length.

The chapter is organized as follows. Section 4.2 gives the system model (in time- and frequency- domains), Section 4.3, 4.4 and 4.5 describe our IFDE scheme and the PACE algorithm. Section 4.6 presents our fast IFDE implementation. Section 4.7 reports the results of numerical studies, and section 4.8 concludes.

4.2 System Model

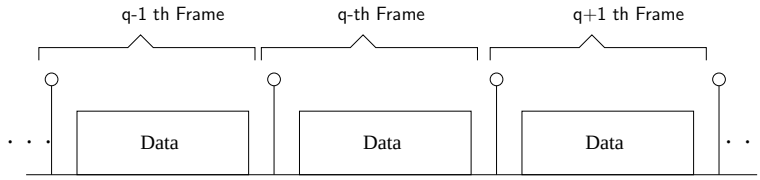

Figure 4.1: Frame structure of transmitted signal.

We consider identification and equalization of a doubly-selective channel when the transmitted signal is an infinite stream of frames, where each frame consists of a pilot block followed by a data block as shown in Fig. 4.1. Channel identification of current frame is accomplished using pilot blocks of current and the next consecutive frame. The length-N_p pilot block is assumed to have a Kronecker delta structure with $N_p = 2N_h$, where N_h is the channel delay spread. The length-QN data block consists of a stream of finite-alphabet symbols $\{s_n\}$, since we adopts single-carrier modulation. For simplicity of presentation, we assume $N = N_p$, therefore the frame length $N_f = (Q+1)N$. We use $\mathcal{S}_{\mathsf{p}}^{(q)} := \{s_n^{(q)}\}_{n=0}^{N_p-1}$ to denote the pilot portion of the q-th transmission frame and $\mathcal{S}_{\mathsf{d}}^{(q)} := \{s_n^{(q)}\}_{n=N_p}^{N_f-1}$ to denote the corresponding data portion. The Kronecker pattern implies that $s_n^{(q)} = 0$ for $0 \le n \le N_h - 1$ and $N_h+1 \le n \le N_p-1$, and that $s_{N_h}^{(q)} = \sqrt{2N_h}$. The complete set of samples transmitted during the p-th frame is denoted by $\mathcal{S}^{(q)} := \mathcal{S}_{\mathsf{p}}^{(q)} \cup \mathcal{S}_{\mathsf{d}}^{(q)} = \{s_n^{(q)}\}_{n=0}^{N_f-1}$, and the multi-frame transmitted signal $\{s_n\}$ is defined by $s_n := s_{\langle n \rangle_{N_f}}^{\lfloor \frac{n}{N_f} \rfloor}$.

The channel is described by its time-variant discrete impulse response $h_{n,l}$, defined as the time-n response to an impulse applied at time $n-l$. We assume a causal impulse

response of length N_h. The signal observed by the receiver is

$$r_n = \nu_n + \sum_{l=0}^{N_h-1} h_{n,l} s_{n-l} \tag{4.1}$$

where ν_n denotes samples of zero-mean CWGN with variance σ^2. We assume WSSUS [94] so that $\mathrm{E}\{h_{n,l} h^*_{n-q,l-m}\} = \rho_q \sigma^2_{h_l} \delta_m$. Here, ρ_q denotes the normalized autocorrelation (i.e., $\rho_0 = 1$) and $\sigma^2_{h_l}$ denotes the variance of the channel at delay l^{th}.

To facilitate the block-based FDE processing, we first segment the qth received frame into overlapping blocks of length N_b and then window each block by a pulse shaping filter $\{b_n\}_{n=0}^{N_b-1}$ so as to mitigate ICI. Those windowed N_b samples are transformed into frequency domain with spacing $\frac{2\pi}{PN}$, where IFDE is applied and N most reliable symbol estimates are detected and output. Therefore, the consecutive processing blocks are of shift N. The remainder of this section establishes the block-based frequency-domain equivalent of (4.1). At each block index $i \in \mathbb{Z}$, the receiver windows an iN-shifted version of the time-domain observation $\{r_n\}$ and applies a DFT with frequency spacing $\frac{2\pi}{PN}$, yielding the i^{th}-block frequency domain observation $\{x_d(i)\}_{d=0}^{PN-1}$:

$$x_d^{(q)}(i) = \frac{1}{\sqrt{PN}} \sum_n r^{(q)}_{iN+N_o+n} b_n e^{-j\frac{2\pi}{PN}dn}. \tag{4.2}$$

Note that the window length N_b is arbitrary. The offset N_o in (4.2) is set to satisfy the condition that $\{x_d^{(q)}(0)\}_{d=0}^{PN-1}$ yields output $\{\hat{s}_n^{(q)}\}_{n=N_p}^{N_p+N-1}$ (the estimates of the first N data symbols of q-th frame) after channel equalization. For convenience, we define

$$\begin{aligned} s_n^{(q)}(i) &:= s^{(q)}_{iN+N_o+n}, \quad n \in \{0, \ldots, PN-1\} \\ a_n &:= \begin{cases} 1 & n \in \{0, \ldots, PN-1\}, \\ 0 & \text{else}, \end{cases} \end{aligned} \tag{4.3}$$

noting that $\{a_n\}$ specifies a PN-length rectangular window and that

$$s^{(q)}_{iN+N_o+n} \;=\; \sum_{\ell=-\infty}^{\infty} s^{(q)}_{\langle n\rangle_{PN}}(i-P\ell)\, a_{\ell PN+n}. \tag{4.4}$$

Equation (4.4) says that, for a particular i, the transmitted sequence $\{s_{iN+N_o+n}\}$ can be constructed using PN-sample shifts of the disjoint subsequences $\{s^{(q)}_n(i-P\ell)\}_{n=0}^{PN-1}$ for $\ell \in \mathbb{Z}$. Combining (4.1)-(4.4), we find

$$x^{(q)}_d(i) = w^{(q)}_d(i) + \frac{1}{\sqrt{PN}} \sum_{n=0}^{PN-1} b_n \sum_{l=0}^{N_h-1} h^{(q)}_{iN+N_o+n,l} \sum_{\ell=-\infty}^{\infty} s^{(q)}_{\langle n-l\rangle_{PN}}(i-P\ell) a_{\ell PN+n-l} e^{-j\frac{2\pi}{PN}nd} \tag{4.5}$$

$$w^{(q)}_d(i) := \frac{1}{\sqrt{PN}} \sum_{n=0}^{PN-1} b_n \nu^{(q)}_{iN+N_o+n} e^{-j\frac{2\pi}{PN}nd}. \tag{4.6}$$

Frequency-domain equalization involves the i^{th}-block *virtual subcarriers* $\{t_k(i)\}_{k=0}^{PN-1}$, where

$$t^{(q)}_k(i) \;:=\; \frac{1}{\sqrt{PN}} \sum_{n=0}^{PN-1} s^{(q)}_n(i) e^{-j\frac{2\pi}{PN}nk}. \tag{4.7}$$

Equation (4.7) implies that $s^{(q)}_n(i) = \frac{1}{\sqrt{PN}} \sum_{k=0}^{PN-1} t^{(q)}_k(i) e^{j\frac{2\pi}{PN}nk}$ for $n \in \{0,\ldots,PN-1\}$. Using this in (4.5) gives

$$x^{(q)}_d(i) \;=\; w^{(q)}_d(i) + \sum_{\ell=-\infty}^{\infty} \sum_{k=0}^{PN-1} t^{(q)}_k(i-P\ell) H^{(q)}_{d-k,k}(i,\ell) \tag{4.8}$$

$$H^{(q)}_{d,k}(i,\ell) \;:=\; \frac{1}{PN} \sum_{n} \sum_{l=0}^{N_h-1} h^{(q)}_{iN+N_o+n,l} b_n a_{\ell PN+n-l}\, e^{-j\frac{2\pi}{PN}(kl+nd)}. \tag{4.9}$$

Equation (4.8) indicates that $H^{(q)}_{d,k}(i,\ell)$ can be interpreted as the response, at DFT output $k+d$ in block i, to a frequency-domain impulse applied at virtual subcarrier k in block $i-\ell$.

In practice we implement a causal length-N_b window $\{b_n\}$ implying that, for any i, only a finite number of terms in the set $\{H^{(q)}_{d,k}(i,\ell), \ell \in \mathbb{Z}\}$ will be non-zero.

Specifically, (4.9) implies that non-zero terms result from indices ℓ which satisfy $0 \leq \ell NP + n - l \leq PN - 1$ for some $n \in \{0, \ldots, N_b - 1\}$ and some $l \in \{0, \ldots, N_h - 1\}$. It is straightforward to show that $H_{d,k}^{(q)}(i, \ell)$ is non-zero for $\ell \in \{-L_{\mathsf{pre}}, \ldots, L_{\mathsf{pst}}\}$ where $L_{\mathsf{pre}} = -\lfloor \frac{N_b - 1}{PN} \rfloor$ and $L_{\mathsf{pst}} = \lfloor \frac{PN + N_h - 2}{PN} \rfloor$.

With the definitions $\boldsymbol{x}^{(q)}(i) := [x_0^{(q)}(i), \ldots, x_{PN-1}^{(q)}(i)]^T$, $\boldsymbol{w}^{(q)}(i) := [w_0^{(q)}(i), \ldots, w_{PN-1}^{(q)}(i)]^T$, $\boldsymbol{t}^{(q)}(i) := [t_0^{(q)}(i), \ldots, t_{PN-1}^{(q)}(i)]^T$, $\boldsymbol{s}^{(q)}(i) := [s_0^{(q)}(i), \ldots, s_{PN-1}^{(q)}(i)]^T$, and $[\boldsymbol{\mathcal{H}}(i, \ell)]_{d,k}^{(q)} := H_{d-k,k}^{(q)}(i, \ell)$, (4.8) implies the LTV multiple-input multiple-output (MIMO) system

$$\boldsymbol{x}^{(q)}(i) \;=\; \boldsymbol{w}^{(q)}(i) + \sum_{\ell=-L_{\mathsf{pre}}}^{L_{\mathsf{pst}}} \boldsymbol{\mathcal{H}}^{(q)}(i, \ell) \boldsymbol{t}^{(q)}(i - P\ell). \tag{4.10}$$

For any i, nonzero $\{\boldsymbol{\mathcal{H}}^{(q)}(i, \ell)\}_{\ell \neq 0}$ cause inter-block interference (IBI) and nonzero off-diagonal elements of $\{\boldsymbol{\mathcal{H}}(i, 0)\}$ cause ICI among the virtual subcarriers. In the sequel, we refer to $\{\boldsymbol{\mathcal{H}}^{(q)}(i, \ell)\}_{\ell < 0}$ as pre-cursor IBI and $\{\boldsymbol{\mathcal{H}}^{(q)}(i, \ell)\}_{\ell > 0}$ as post-cursor IBI.

It will sometimes be convenient to write the windowed frequency-domain noise vector $\boldsymbol{w}^{(q)}(i)$ as

$$\boldsymbol{w}^{(q)}(i) \;=\; \underbrace{\boldsymbol{F}\boldsymbol{J}\,\mathcal{D}(\boldsymbol{b})}_{:=\boldsymbol{C}} \boldsymbol{\nu}^{(q)}(i) \tag{4.11}$$

$$\boldsymbol{J} \;:=\; \left[\begin{array}{ccc|c} \boldsymbol{I}_{PN} & \cdots & \boldsymbol{I}_{PN} & \begin{matrix} \boldsymbol{I}_{\bar{N}_o} \\ \boldsymbol{0}_{PN - \bar{N}_o \times \bar{N}_o} \end{matrix} \end{array}\right] \tag{4.12}$$

where $\boldsymbol{F}$ denotes the PN-point unitary DFT matrix, $\bar{N}_o = \langle N_b \rangle_{PN}$, and the number of $\boldsymbol{I}_{PN}$ matrices in $\boldsymbol{J}$ is $\lfloor \frac{N_b}{PN} \rfloor$.

For notational brevity, we omit the frame index q in the rest of the paper except section 4.4, where a MMSE estimator is derived to leverage the pilot blocks of successive frames to estimate the channel in between.

4.3 Max-SINR Window Design

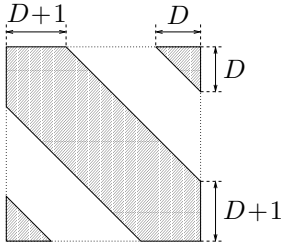

Figure 4.2: Desired "banded" structure of matrix $\boldsymbol{\mathcal{H}}(i,0)$.

The choice of window $\{b_n\}$ affects the IBI/ICI patterns of the MIMO system (4.10). Motivated by the low-pass nature of typical Doppler spectra, we aim to find $\{b_n\}$ such that the "cursor" coefficient $\boldsymbol{\mathcal{H}}(i,0)$ has the banded structure illustrated in Fig. 4.2 and the IBI coefficients $\{\boldsymbol{\mathcal{H}}(i,\ell)\}_{\ell\neq 0}$ vanish. This approach can be viewed as the frequency-domain dual of inter-symbol interference (ISI) response shortening used to reduce the complexity of maximum likelihood sequence detection (MLSD) [96]. For our purposes, the goal of time-domain windowing is to give the channel a sparse structure that leads to low-complexity estimation of $\boldsymbol{t}(i)$, and hence, low-complexity detection of $\boldsymbol{s}(i)$. We choose time-domain windowing, rather than a general matrix operation on the received signal, due to its low complexity. Since complete cancellation of out-of-target ICI/IBI is, in general, not possible with time-domain windowing, we choose to maximize signal to interference-plus-noise ratio (SINR) as a means of suppressing residual IBI/ICI.

We define SINR by $\mathcal{E}_{\mathsf{s}}/\mathcal{E}_{\mathsf{ni}}$, where $\mathcal{E}_{\mathsf{s}} := \sum_d \mathcal{E}_{\mathsf{s},d}$ and $\mathcal{E}_{\mathsf{ni}} := \sum_d \mathcal{E}_{\mathsf{ni},d}$. For each $x_d(i)$, $\mathcal{E}_{\mathsf{s},d}$ is defined as the signal energy contributed by neighboring carriers $\{t_k(i)\}_{k=d-D}^{d+D}$, and $\mathcal{E}_{\mathsf{ni},d}$ is defined as the interference-plus-noise energy contributed by non-neighboring

carriers $\{t_k(i)\}_{k=0}^{d-D-1} \cup \{t_k(i)\}_{k=d+D+1}^{PN-1}$, non-cursor carriers $\{t_k(j)\}_{j\neq i}$, and additive noise $\boldsymbol{w}(i)$. Note that indices here are taken modulo-PN. The ICI radius D is typically chosen as $D = \lceil f_{\mathsf{d}} T_{\mathsf{s}} PN \rceil$, where $f_{\mathsf{d}} T_{\mathsf{s}}$ is the maximum Doppler frequency normalized to the symbol rate. Using the approach outlined in [97], we find that the SINR-maximizing window $\boldsymbol{b}_\star$ is given by

$$\begin{aligned} \boldsymbol{b}_\star &= \underset{\boldsymbol{b}:\|\boldsymbol{b}\|^2=PN}{\arg\max} \frac{\boldsymbol{b}^H\big(\boldsymbol{R}_b \odot \boldsymbol{D}_b \odot \boldsymbol{A}_{\mathsf{s}}\big)\boldsymbol{b}}{\boldsymbol{b}^H\big(\sigma^2\boldsymbol{I} + \boldsymbol{R}_b \odot \boldsymbol{C}_b \odot \boldsymbol{A}_{\mathsf{t}} - \boldsymbol{R}_b \odot \boldsymbol{D}_b \odot \boldsymbol{A}_{\mathsf{s}}\big)\boldsymbol{b}} \\ &= \boldsymbol{v}_\star\big(\boldsymbol{R}_b \odot \boldsymbol{D}_b \odot \boldsymbol{A}_{\mathsf{s}},\ \sigma^2\boldsymbol{I} + \boldsymbol{R}_b \odot \boldsymbol{C}_b \odot \boldsymbol{A}_{\mathsf{t}} - \boldsymbol{R}_b \odot \boldsymbol{D}_b \odot \boldsymbol{A}_{\mathsf{s}}\big) \cdot \sqrt{PN} \end{aligned} \quad (4.13)$$

where $\boldsymbol{R}_b$, $\boldsymbol{A}_{\mathsf{s}}$, $\boldsymbol{C}_b$, $\boldsymbol{D}_b$ and $\boldsymbol{A}_{\mathsf{t}}$ are $N_b \times N_b$ matrices defined element-wise as $[\boldsymbol{R}_b]_{m,n} := \rho_{n-m}$, $[\boldsymbol{A}_{\mathsf{s}}]_{m,n} := \sum_{l=0}^{N_h-1} \sigma_l^2 a_{n-l} a_{m-l}^*$, $[\boldsymbol{C}_b]_{m,n} := \delta_{\langle n-m\rangle_{PN}}$, $[\boldsymbol{D}_b]_{m,n} := \frac{1}{PN}\sin(\frac{\pi}{PN}(2D+1)(n-m))/\sin(\frac{\pi}{PN}(n-m))$ and $[\boldsymbol{A}_{\mathsf{t}}]_{m,n} := \sum_{\ell=-L_{\mathsf{pre}}}^{L_{\mathsf{pst}}} \sum_{l=0}^{N_h-1} \sigma_l^2 a_{\ell PN+n-l} a_{\ell PN+m-l}^*$. In (4.13), $\boldsymbol{v}_\star(\boldsymbol{B},\boldsymbol{C})$ denotes the principle generalized eigenvector [98] of the matrix pair $(\boldsymbol{B},\boldsymbol{C})$. Through max-SINR windowing and proper selection of other design parameters, the IBI and non-neighboring ICI can be made small enough to base the symbol detection procedure on the following approximate system model.

$$\boldsymbol{x}(i) \approx \boldsymbol{\mathcal{H}}(i,0)\boldsymbol{t}(i) + \boldsymbol{C}\boldsymbol{\nu}(i). \quad (4.14)$$

As an alternative, the design parameters (e.g., block length PN) could be chosen to yield non-negligible post-cursor IBI, which could then be canceled using block decision-feedback equalization (BDFE). In this case, the window should be designed to suppress *only* ICI and pre-cursor IBI, implying $[\boldsymbol{A}_{\mathsf{t}}]_{m,n} := \sum_{\ell=-L_{\mathsf{pre}}}^{0} \sum_{l=0}^{N_h-1} \sigma_l^2 a_{\ell PN+n-l} a_{\ell PN+m-l}^*$. Figure 4.3 plots windows for both BDFE and non-BDFE cases at $f_{\mathsf{d}} T_{\mathsf{s}} \in \{0.001, 0.0075\}$. In generating Fig. 4.3, we assumed $N_h = 64$, $PN = 256$, $N_b = PN + N_h - 1$, SNR= 10dB, and $\sigma_l^2 = N_h^{-1}$, which are typical choices for the numerical results in

section 4.7. Assuming reliable post-cursor IBI cancellation, (4.14) would still describe the model used for detection of symbols in the current block.

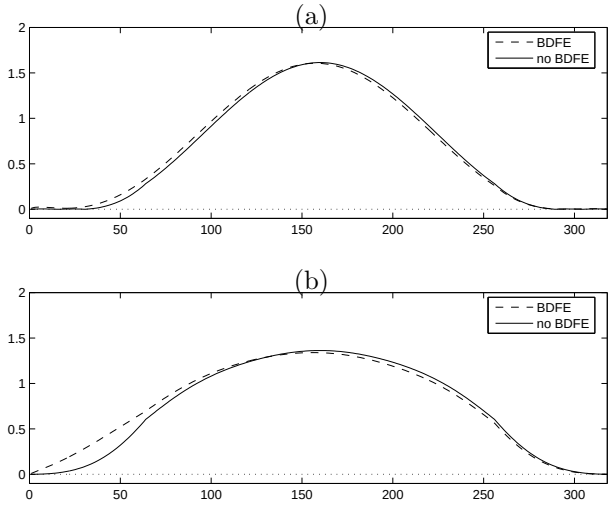

Figure 4.3: Example window shapes for $PN = 256$, $N_h = 64$, SNR=10dB and (a) $f_{\mathsf{d}}T_{\mathsf{s}} = 0.001$, (b) $f_{\mathsf{d}}T_{\mathsf{s}} = 0.0075$.

While windowing gives a sparse channel response that enables a reduced complexity symbol detection procedure, it can lead to a non-uniform collection of energy from symbols in the current block. Specifically, it is shown in Appendix 4.A that the energy in $\boldsymbol{x}(i)$ contributed by $s_n(i)$ is

$$\begin{aligned}\mathcal{E}_{\mathsf{ss},n} &:= \sum_{d=0}^{PN-1} \mathrm{E}\left\{\left|\sum_{k=0}^{PN-1} H_{d-k,k}(i,0)\cdot\frac{1}{\sqrt{PN}}e^{-j\frac{2\pi}{PN}nk}s_n(i)\right|^2\right\} \\ &= \sum_{l=0}^{N_h-1}\sigma_l^2|b_{l+n}|^2 \end{aligned} \tag{4.15}$$

which is clearly dependent on n, the symbol position within the block. This implies that, for typical max-SINR window shapes, we will collect less energy from symbols near the block edges. This phenomenon motivates the block-overlapping procedure proposed in section 4.5.

4.4 Pilot-Aided Channel Estimation

In this section, we will assume that the pulse $\{b_n\}$ is designed properly such that (4.14) is a valid approximate system model for symbol detection purpose. In addition, we are only interested in estimating $PN(2D+1)$ coefficients out of $\boldsymbol{\mathcal{H}}(i,0)$, which are within the shaded region of Fig. 4.2, since only them will be exploited by our proposed IFDE algorithm in section 4.5. In [44], several methods for the pilot-aided estimation of significant ICI coefficients resulting from pulse-shaped multicarrier modulation over doubly selective channels are proposed. Inspired by [44], we will derive the MMSE estimator which fits into the single-carrier transmission systems discussed here.

First we define some quantities that follow from our pilot pattern. Say that we are interested in estimating $\{\boldsymbol{\mathcal{H}}^{(q)}(i,0)\}_{i=0}^{Q-1}$ through the pilot block $\mathcal{S}_{\mathsf{p}}^{(q)}$ and $\mathcal{S}_{\mathsf{p}}^{(q+1)}$, for these i and q we can write

$$\boldsymbol{r}_{\mathsf{p}}^{(q)} := [r_{N_h}^{(q)},\, r_{N_h+1}^{(q)},\, \cdots,\, r_{2N_h-1}^{(q)}]^T \tag{4.16}$$

$$\boldsymbol{h}_{\mathsf{p}}^{(q)} := [h_{N_h,0}^{(q)},\, h_{N_h+1,1}^{(q)},\, \cdots,\, h_{2N_h-1,N_h-1}^{(q)}]^T \tag{4.17}$$

$$\boldsymbol{\nu}_{\mathsf{p}}^{(q)} := [\nu_{N_h}^{(q)},\, \nu_{N_h+1}^{(q)},\, \cdots,\, \nu_{2N_h-1}^{(q)}]^T \tag{4.18}$$

$$\boldsymbol{r}_{\mathsf{p}}^{(q)} = \sqrt{2N_h}\boldsymbol{h}_{\mathsf{p}}^{(q)} + \boldsymbol{\nu}_{\mathsf{p}}^{(q)} \tag{4.19}$$

where (4.19) follows from the Kronecker pattern of the pilot block. Define $\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)} := [\boldsymbol{r}_{\mathsf{p}}^{(q)^T},\, \boldsymbol{r}_{\mathsf{p}}^{(q+1)^T}]^T$, $\underline{\boldsymbol{h}}_{\mathsf{p}}^{(q)} := [\boldsymbol{h}_{\mathsf{p}}^{(q)^T},\, \boldsymbol{h}_{\mathsf{p}}^{(q+1)^T}]^T$ and $\underline{\boldsymbol{\nu}}_{\mathsf{p}}^{(q)} := [\boldsymbol{\nu}_{\mathsf{p}}^{(q)^T},\, \boldsymbol{\nu}_{\mathsf{p}}^{(q+1)^T}]^T$, then it is clear that

$$\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)} = \sqrt{2N_h}\underline{\boldsymbol{h}}_{\mathsf{p}}^{(q)} + \underline{\boldsymbol{\nu}}_{\mathsf{p}}^{(q)} \tag{4.20}$$

Same as in [44], for convenience, we collect those $PN(2D+1)$ significant ICI coefficients into $\boldsymbol{g}_D^{(q)}(i) \in \mathbb{C}^{PN(2D+1)}$:

$$\boldsymbol{g}_D^{(q)}(i) := [\mathrm{diag}_{-D}(\boldsymbol{\mathcal{H}}^{(q)}(i,0))^T, \cdots, \mathrm{diag}_D(\boldsymbol{\mathcal{H}}^{(q)}(i,0))^T]^T \tag{4.21}$$

where $\mathrm{diag}_k(\cdot)$ extracts the kth sub-diagonal of matrix, i.e.,
$\mathrm{diag}_k(\boldsymbol{\mathcal{H}}) := [\boldsymbol{\mathcal{H}}_{k,0}, \boldsymbol{\mathcal{H}}_{k+1,1}, \cdots, \boldsymbol{\mathcal{H}}_{k+PN-1,PN-1}]^T$ with modulo-N indexing assumed. By setting $\ell = 0$ in (4.9), we can show that

$$\boldsymbol{g}_D^{(q)}(i) = \boldsymbol{C}_D \boldsymbol{h}^{(q)}(i), \tag{4.22}$$

where $\boldsymbol{C}_D \in \mathbb{C}^{(PN(2D+1)\times N_b N_h)}$ and $\boldsymbol{h}^{(q)}(i) \in \mathbb{C}^{N_b N_h}$ are defined element-wise as

$$[\boldsymbol{C}_D]_{n,m} := \frac{1}{PN} b_{\langle m \rangle_{N_b}} a_{\langle m \rangle_{N_b} - \left\lfloor \frac{m}{N_b} \right\rfloor} e^{-j\frac{2\pi}{PN}\left(\left\lfloor \frac{m}{N_b} \right\rfloor n + \langle m \rangle_{N_b}\left(\left\lfloor \frac{n}{PN} \right\rfloor - D\right)\right)} \tag{4.23}$$

$$[\boldsymbol{h}^{(q)}(i)]_m := h^{(q)}_{\langle m \rangle_{N_b} + N_o + iN, \left\lfloor \frac{m}{N_b} \right\rfloor} \tag{4.24}$$

Here we have utilized the trick that $\langle n \rangle_{PN} = k$. Note that $\boldsymbol{h}^{(q)}(i)$ contains all time-domain impulse response coefficients affecting $\boldsymbol{\mathcal{H}}^{(q)}(i,0)$. Our goal is to estimate $\underline{\boldsymbol{g}}_D^{(q)} := [\boldsymbol{g}_D^{(q)}(0)^T, \boldsymbol{g}_D^{(q)}(1)^T, \cdots, \boldsymbol{g}_D^{(q)}(Q-1)^T]^T$, which are required for coherent symbol detection, from $\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)}$, the pilot observations.

The linear MMSE estimation of $\underline{\boldsymbol{g}}_D^{(q)}$ from $\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)}$ is [70]

$$\hat{\underline{\boldsymbol{g}}}_D^{(q)} = \boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}} \boldsymbol{R}^{-1}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}} \underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)} \tag{4.25}$$

where $\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}} := \mathrm{E}[\underline{\boldsymbol{g}}_D^{(q)} \underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)H}]$ and $\boldsymbol{R}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}} := \mathrm{E}[\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)} \underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)H}]$. Note that for stationary channel considered here, $\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}}$ and $\boldsymbol{R}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}}$ are unrelated with frame index q, therefore we only need to compute $\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}} \boldsymbol{R}^{-1}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}}$ once and store them for future use, which greatly

reduces the computational cost of PACE. From (4.19) and (4.22), we know that

$$\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(0,0)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(0,1)} \\ \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(1,0)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(1,1)} \\ \vdots & \vdots \\ \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(Q-1,0)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(Q-1,1)} \end{bmatrix} \quad \boldsymbol{R}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{r}_{\mathsf{p}}\boldsymbol{r}_{\mathsf{p}}}^{(0)} & \boldsymbol{R}_{\boldsymbol{r}_{\mathsf{p}}\boldsymbol{r}_{\mathsf{p}}}^{(-1)} \\ \boldsymbol{R}_{\boldsymbol{r}_{\mathsf{p}}\boldsymbol{r}_{\mathsf{p}}}^{(1)} & \boldsymbol{R}_{\boldsymbol{r}_{\mathsf{p}}\boldsymbol{r}_{\mathsf{p}}}^{(0)} \end{bmatrix} \tag{4.26}$$

where

$$\boldsymbol{R}_{\boldsymbol{g}\boldsymbol{r}_{\mathsf{p}}}^{(i,k)} = \boldsymbol{C}_D \boldsymbol{R}_{\boldsymbol{h}\boldsymbol{r}_{\mathsf{p}}}^{(i,k)} \tag{4.27}$$

$$[\boldsymbol{R}_{\boldsymbol{h}\boldsymbol{r}_{\mathsf{p}}}^{(i,k)}]_{m,n} = \sqrt{2N_h}\rho_{\langle m \rangle_{N_b}+N_o+iN-N_h-kN_f-n}\sigma_{h_n}^2 \delta(n - \left\lfloor \frac{m}{N_b} \right\rfloor) \tag{4.28}$$

$$\boldsymbol{R}_{\boldsymbol{r}_{\mathsf{p}}\boldsymbol{r}_{\mathsf{p}}}^{(k)} = 2N_h \boldsymbol{R}_{\boldsymbol{h}_{\mathsf{p}}\boldsymbol{h}_{\mathsf{p}}}^{(k)} + \delta_k \sigma_\nu^2 \mathbf{I} \tag{4.29}$$

$$\boldsymbol{R}_{\boldsymbol{h}_{\mathsf{p}}\boldsymbol{h}_{\mathsf{p}}}^{(k)} = \rho_{|k|N_f}\, \mathcal{D}(\boldsymbol{\sigma}_{\boldsymbol{h}}^2) \tag{4.30}$$

and $\boldsymbol{\sigma}_{\boldsymbol{h}}^2 := [\sigma_{h_0}^2,\, \sigma_{h_1}^2,\, \cdots,\, \sigma_{h_{N_h-1}}^2]^T$.

The estimation error $\underline{\tilde{\boldsymbol{g}}}_D = \underline{\hat{\boldsymbol{g}}}_D - \underline{\boldsymbol{g}}_D$ has covariance

$$\mathrm{E}[\underline{\tilde{\boldsymbol{g}}}_D \underline{\tilde{\boldsymbol{g}}}_D^H] = \boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{g}}} - \boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}} \boldsymbol{R}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}}^{-1} \boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}}^H \tag{4.31}$$

where same as in [44], $\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{g}}} := \mathrm{E}[\underline{\boldsymbol{g}}_D \underline{\boldsymbol{g}}_D^H]$ is given by

$$\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{g}}} = \begin{bmatrix} \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(0)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(-1)} & \cdots & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(1-Q)} \\ \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(1)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(0)} & \cdots & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(2-Q)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(Q-1)} & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(Q-2)} & \cdots & \boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(0)} \end{bmatrix} \tag{4.32}$$

$$\boldsymbol{R}_{\boldsymbol{g}\boldsymbol{g}}^{(i)} = \boldsymbol{C}_D \boldsymbol{R}_{\boldsymbol{h}\boldsymbol{h}}^{(i)} \boldsymbol{C}_D^H \tag{4.33}$$

$$\boldsymbol{R}_{\boldsymbol{h}\boldsymbol{h}}^{(i)} = \mathcal{D}(\boldsymbol{\sigma}_{\boldsymbol{h}}^2) \otimes \boldsymbol{R}_{\rho}^{(i)} \tag{4.34}$$

$$[\boldsymbol{R}_{\rho}^{(i)}]_{m,n} = \rho_{m-n+iN} \quad m,n \in \{0,\, \cdots,\, N_b - 1\} \tag{4.35}$$

4.5 Symbol Detection

In section 4.5.1, we propose an iterative method for the detection of the finite-alphabet symbol vector $\boldsymbol{s}(i) = \boldsymbol{F}^H \boldsymbol{t}(i)$ assuming the observation model (4.14). For

simplicity, we assume perfect CSI here, when it is not available, we use channel estimates obtained through (4.25) instead. We are careful to leverage the banded structure of $\boldsymbol{\mathcal{H}}(i,0)$ and the existence of fast algorithms for the transformation $\boldsymbol{F}$. It was previously observed that the max-SINR windowing described in section 4.3 collects less energy from symbols near the block edges, which, if unaccounted for, could lead to high block-averaged error rates. Hence, section 4.5.2 proposes a scheme whereby block overlap (i.e., $P > 1$) is exploited, in conjunction with the algorithm of section 4.5.1, to circumvent this problem.

4.5.1 Intrablock Processing

Here we propose an iterative method for the detection of the finite-alphabet symbol vector $\boldsymbol{s}(i)$ from the windowed frequency-domain observation $\boldsymbol{x}(i)$ in (4.14). Note that the focus of this section is *intrablock* processing, whereas the focus of section 4.5.2 is *interblock* processing. Since here we focus exclusively on the i^{th} symbol and on the cursor IBI coefficient, we can omit the symbol and lag indices, abbreviating, e.g., $\boldsymbol{s}(i)$ by $\boldsymbol{s}$ and $\boldsymbol{\mathcal{H}}(i,0)$ by $\boldsymbol{\mathcal{H}}$. We now give a brief summary of the intrablock detection algorithm illustrated in Fig. 4.4; a more detailed description will be given in Sections 4.5.1–4.5.1.

Figure 4.4: Intrablock interference cancellation.

Given current guesses of the log-likelihood ratios (LLRs) of the symbols $\{s_k\}$ (which, on the first iteration, are set to zero), the means and variances of the elements in $\boldsymbol{s}$ are calculated as $\overline{\boldsymbol{s}}$ and $\boldsymbol{v}$, respectively. These are then transformed into the mean and covariance of $\boldsymbol{t}$. Using linear MMSE estimation and incorporating these mean/variance priors, the elements $\{t_k\}$ are estimated one-at-a-time, leveraging the banded structure of $\boldsymbol{\mathcal{H}}$ for complexity reduction. The resulting estimates $\hat{\boldsymbol{t}}$ are then transformed back into the $\boldsymbol{s}$-domain, from which the LLRs are updated. To accomplish this last step we assume a conditionally-Gaussian model for the estimates $\{\hat{s}_k\}$. The procedure then repeats, starting with the most recent LLRs. In the detailed description below, we use the superscript $^{(n)}$ to denote the n^{th} iteration.

Linear Estimation with Priors

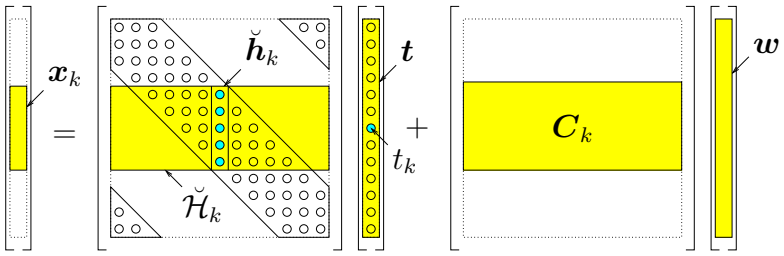

Figure 4.5: Truncated observation model.

The banded structure of $\boldsymbol{\mathcal{H}}$ suggests that linear estimation of a particular element t_k might be accomplished with reasonable accuracy from the truncated observation $\boldsymbol{x}_k := [x_{k-D}, \ldots, x_{k+D}]^t$, with indices taken modulo-PN, as opposed to the full observation $\boldsymbol{x}$. (See Fig. 4.5.) We hope to realize substantial complexity reduction in

doing so. The truncated model becomes

$$\boldsymbol{x}_k \;=\; \boldsymbol{\mathcal{H}}_k\boldsymbol{t} + \boldsymbol{C}_k\boldsymbol{\nu}, \tag{4.36}$$

where $\boldsymbol{\mathcal{H}}_k$ contains rows $\{k-D,\ldots,k+D\}$ of $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{C}_k$ contains rows $\{k-D,\ldots,k+D\}$ of $\boldsymbol{C}$. The MMSE linear estimate of t_k given $\boldsymbol{x}_k$ is [99]

$$\hat{t}_k \;=\; \mathrm{E}\{t_k\} + \mathrm{cov}(t_k,\boldsymbol{x}_k)\,\mathrm{cov}(\boldsymbol{x}_k,\boldsymbol{x}_k)^{-1}(\boldsymbol{x}_k - \mathrm{E}\{\boldsymbol{x}_k\}). \tag{4.37}$$

We assume $\mathrm{E}\{\boldsymbol{\nu}\}=\boldsymbol{0}$, $\mathrm{cov}(\boldsymbol{\nu},\boldsymbol{\nu})=\sigma^2\boldsymbol{I}$, and $\mathrm{cov}(\boldsymbol{s},\boldsymbol{\nu})=\boldsymbol{0}$, and we model the elements in $\boldsymbol{s}$ as uncorrelated with means $\overline{\boldsymbol{s}}^{(n)}$ and variances $\boldsymbol{v}^{(n)}$ during the n^{th} iteration. Then, defining $\overline{\boldsymbol{t}}^{(n)} := \boldsymbol{F}\overline{\boldsymbol{s}}^{(n)}$, (4.37) becomes

$$\hat{t}_k^{(n)} \;=\; \bar{t}_k^{(n)} + \boldsymbol{g}_k^{(n)H}(\boldsymbol{x}_k - \boldsymbol{\mathcal{H}}_k\overline{\boldsymbol{t}}^{(n)}) \tag{4.38}$$

$$\boldsymbol{g}_k^{(n)} \;:=\; \left(\boldsymbol{\mathcal{H}}_k\boldsymbol{F}\,\mathcal{D}(\boldsymbol{v}^{(n)})\boldsymbol{F}^H\boldsymbol{\mathcal{H}}_k^H + \sigma^2\boldsymbol{C}_k\boldsymbol{C}_k^H\right)^{-1}\boldsymbol{\mathcal{H}}_k\boldsymbol{F}\,\mathcal{D}(\boldsymbol{v}^{(n)})\boldsymbol{F}^H\boldsymbol{i}_k \tag{4.39}$$

from which estimates of $\boldsymbol{s}$ can be obtained as

$$\hat{\boldsymbol{s}}^{(n)} \;=\; \boldsymbol{F}^H\hat{\boldsymbol{t}}^{(n)} \quad\Leftrightarrow\quad \hat{s}_l^{(n)} \;=\; \boldsymbol{i}_l^H\boldsymbol{F}^H\sum_k \boldsymbol{i}_k\hat{t}_k^{(n)}. \tag{4.40}$$

A Conditionally Gaussian Model

Leveraging the finite-alphabet structure of the elements $\{s_k\}$ and assuming reasonably large PN (to invoke the Central Limit Theorem), we assume that the estimation error is Gaussian, or, equivalently, that the estimates are conditionally Gaussian:

$$p(\hat{s}_l^{(n)}|s_l=b) \;=\; \frac{1}{\sigma_l^{(n)}(b)}\phi\!\left(\frac{\hat{s}_l^{(n)}-\mu_l^{(n)}(b)}{\sigma_l^{(n)}(b)}\right), \tag{4.41}$$

where $\phi(w) := \frac{1}{\sqrt{\pi}}e^{-w^2}$, $\mu_l^{(n)}(b) := \mathrm{E}\{\hat{s}_l^{(n)}|s_l=b\}$, and $[\sigma_l^{(n)}(b)]^2 := \mathrm{cov}(\hat{s}_l^{(n)},\hat{s}_l^{(n)}|s_l=b)$.

In Appendix 4.B we show that

$$\mu_l^{(n)}(b) = \bar{s}_l^{(n)} + Q_{l,l}^{(n)*}(b - \bar{s}_l^{(n)}) \tag{4.42}$$

$$[\sigma_l^{(n)}(b)]^2 = \boldsymbol{q}_l^{(n)H}\,\mathcal{D}(\boldsymbol{v}^{(n)})\boldsymbol{q}_l^{(n)} - |Q_{l,l}^{(n)}|^2 v_l^{(n)} + \sigma^2\|\boldsymbol{p}_l^{(n)}\|^2, \tag{4.43}$$

where $\boldsymbol{q}_l^{(n)}$ denotes the l^{th} column of $\boldsymbol{Q}^{(n)}$ and where $\boldsymbol{p}_l^{(n)}$ denotes the l^{th} column of $\boldsymbol{P}^{(n)}$:

$$\boldsymbol{Q}^{(n)} = \boldsymbol{F}^H\Big(\sum_k \boldsymbol{\mathcal{H}}_k^H \boldsymbol{g}_k^{(n)} \boldsymbol{i}_k^H\Big)\boldsymbol{F} \tag{4.44}$$

$$\boldsymbol{P}^{(n)} = \Big(\sum_k \boldsymbol{C}_k^H \boldsymbol{g}_k^{(n)H} \boldsymbol{i}_k^H\Big)\boldsymbol{F}. \tag{4.45}$$

Log-Likelihood Ratio and Update of Priors

From now on, we restrict ourselves to the BPSK alphabet so that $b \in \{-1, +1\}$; QAM extensions are straightforward but tedious (see, e.g., [100, 101]). The n^{th}-iteration *a priori* and *a posteriori* LLRs are then defined as $L_l^{(n)} := \log\frac{P(s_l=+1)}{P(s_l=-1)}$ and $L_l(\hat{s}_l^{(n)}) := \log\frac{P(s_l=+1|\hat{s}_l^{(n)})}{P(s_l=-1|\hat{s}_l^{(n)})}$, respectively. Note that, after the first iteration, we expect to have partial information on s_l such that $L_l^{(n)} \neq 0$. The LLR update $\Delta(\hat{s}_l^{(n)}) := L_l(\hat{s}_l^{(n)}) - L_l^{(n)}$ can be written

$$\begin{aligned}
\Delta(\hat{s}_l^{(n)}) &= \log\frac{p(\hat{s}_l^{(n)}|s_l=+1)}{p(\hat{s}_l^{(n)}|s_l=-1)} \\
&= \frac{|\hat{s}_l^{(n)} - \mu_l^{(n)}(-1)|^2 - |\hat{s}_l^{(n)} - \mu_l^{(n)}(+1)|^2}{[\sigma_l^{(n)}(\pm 1)]^2} \\
&= \frac{4\left(\mathrm{Re}\big(Q_{l,l}^{(n)}(\hat{s}_l^{(n)} - \bar{s}_l^{(n)})\big) + |Q_{l,l}^{(n)}|^2\bar{s}_l^{(n)}\right)}{\boldsymbol{q}_l^{(n)H}\,\mathcal{D}(\boldsymbol{v}^{(n)})\boldsymbol{q}_l^{(n)} - |Q_{l,l}^{(n)}|^2 v_l^{(n)} + \sigma^2\|\boldsymbol{p}_l^{(n)}\|^2},
\end{aligned} \tag{4.46}$$

where we used the facts that $\sigma_l^{(n)}(+1) = \sigma_l^{(n)}(-1)$ and

$$\begin{aligned}
&\big|\hat{s}_l^{(n)} - \mu_l^{(n)}(-1)\big|^2 - \big|\hat{s}_l^{(n)} - \mu_l^{(n)}(+1)\big|^2 \\
&= \big|\hat{s}_l^{(n)} - (1 - Q_{l,l}^{(n)*})\bar{s}_l^{(n)} + Q_{l,l}^{(n)*}\big|^2 \\
&\quad - \big|\hat{s}_l^{(n)} - (1 - Q_{l,l}^{(n)*})\bar{s}_l^{(n)} - Q_{l,l}^{(n)*}\big|^2 \\
&= 4\,\mathrm{Re}\big\{\big(\hat{s}_l^{(n)} - (1 - Q_{l,l}^{(n)*})\bar{s}_l^{(n)}\big)Q_{l,l}^{(n)}\big\} \\
&= 4\,\mathrm{Re}\big\{Q_{l,l}^{(n)}(\hat{s}_l^{(n)} - \bar{s}_l^{(n)})\big\} + |Q_{l,l}^{(n)}|^2\bar{s}_l^{(n)}, \qquad (4.47)
\end{aligned}$$

since the use of BPSK implies $\bar{s}_l^{(n)} \in \mathbb{R}$. Updates of the symbol mean and variance can be accomplished via

$$\begin{aligned}
\bar{s}_l^{(n+1)} &= \sum_{b\in\mathcal{B}} b \cdot P(s_l = b|\hat{s}_l^{(n)}) \\
&= \tanh\left(\frac{L_l(\hat{s}_l^{(n)})}{2}\right) \qquad (4.48) \\
v_l^{(n+1)} &= \sum_{b\in\mathcal{B}} \big(b - \bar{s}_l^{(n+1)}\big)^2 P(s_l = b|\hat{s}_l^{(n)}) \\
&= 1 - (\bar{s}_l^{(n+1)})^2. \qquad (4.49)
\end{aligned}$$

To update the *a priori* LLR, we set $L_l^{(n+1)} := L_l(\hat{s}_l^{(n)})$, giving

$$L_l^{(n+1)} = L_l^{(n)} + \Delta(\hat{s}_l^{(n)}). \qquad (4.50)$$

Hard symbol estimates can be generated as $\hat{\hat{s}}_l^{(n)} := \mathrm{sign}\big(\mathrm{Re}(\hat{s}_l^{(n)})\big) = \mathrm{sign}\big(\bar{s}_l^{(n)}\big) = \mathrm{sign}\big(L(s_l|\hat{s}_l^{(n)})\big)$. An algorithm summary appears in Table 4.1. Note that a soft decoding algorithm could be easily embedded within the bottom path of Fig. 4.4, as proposed in [100] and investigated in [101].

4.5.2 Interblock Processing

As previously discussed, the use of max-SINR windowing causes less energy to be collected from symbols near the edges of block $\boldsymbol{s}(i)$ than from those near the center

of the block. As a result, the iterative detection algorithm described in section 4.5.1 is more likely to incorrectly detect symbols near the block edges. However, by overlapping the frames (i.e., choosing $P > 1$), we can exploit the fact that every symbol will be near the center of some block. Specifically, (4.3) implies that s_m maps to the block-quantities $\left\{ s_{\langle m-N_o \rangle_N}(\lfloor \frac{m-N_o}{N} \rfloor), \ldots, s_{\langle m-N_o \rangle_N + (P-1)N}(\lfloor \frac{m-N_o}{N} \rfloor - P + 1) \right\}$, i.e., s_m appears in P distinct blocks. The block index i_m for which s_m appears closest to block center is easily found to be

$$i_m = \left\lfloor \frac{m - N_o}{N} \right\rfloor - j_m \tag{4.51}$$

$$j_m := \arg \min_{j=0,\ldots,P-1} \left| \langle m - N_o \rangle_N + jN - \frac{PN}{2} \right|. \tag{4.52}$$

Thus, in exploiting block overlap, we stipulate that

1. the hard estimate of s_m is generated at block index i_m, i.e., $\hat{\hat{s}}_m = \hat{\hat{s}}_{\langle m-N_o \rangle_N + j_m N}(i_m)$, and

2. the final LLR calculated for symbol s_m during block i_m is used to initialize the LLR of that symbol in subsequent frames within which it appears, i.e., in frames with index $i \in \{i_m + 1, i_m + 2, \ldots, \lfloor \frac{m-N_o}{N} \rfloor\}$.

In the case that BDFE is employed, these hard estimates are then also used for post-cursor IBI cancellation. Figure 4.6 illustrates this process for $P = 2$.

Since every symbol s_m is estimated P times, the overall equalizer complexity increases linearly with P. Numerical simulations suggest that the performance with $P > 2$ is not significantly better than $P = 2$, while the performance with $P = 1$ is relatively poor. Hence, we focus on $P = 2$ for the remainder of the paper.

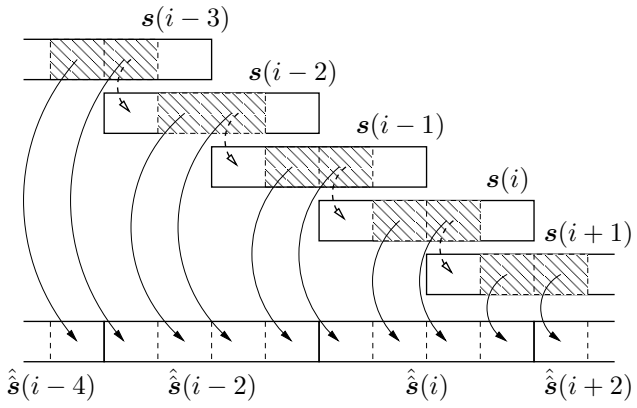

Figure 4.6: Interblock detection process for $P = 2$. Solid arrows pass final hard estimates; dashed arrows pass soft initializations.

4.6 Fast Algorithm and Complexity Analysis

In Table 4.2 we present a fast version of the detection algorithm summarized in Table 4.1. In the fast version, we avoid explicit computation of $\boldsymbol{Q}^{(n)}$ and $\boldsymbol{P}^{(n)}$, instead computing $y_k^{(n)} := \boldsymbol{q}_k^{(n)H} \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{q}_k^{(n)}$, $z_k^{(n)} := \|\boldsymbol{p}_k^{(n)}\|^2$, and $Q_{k,k}^{(n)}$ for $k \in \{0, 1, \ldots, PN-1\}$. The number of complex multiplications[7] per step is given in the right column of Table 4.2, and per-symbol averages are summarized in Table 4.3 (assuming M iterations) for both BDFE and non-BDFE cases. We include the cost of estimating frequency-domain channel coefficients $\{\boldsymbol{\mathcal{H}}(i,\ell)\}$, as well as that of post-cursor IBI cancellation in the BDFE case. Table 4.3 also includes the per-symbol cost of a fast version of the LTV-channel FIR-MMSE-DFE and RLS for CE. The details of each step are enumerated in Appendix 4.C in correspondence with the left column of Table 4.2.

[7]While the number of additions and divisions could also be counted, we feel that such an endeavor would complicate the presentation without providing significant additional insight.

4.7 Numerical Results

4.7.1 IFDE with Perfect CSI

In this section, we compare the performance and complexity of the fast IFDE algorithm summarized in Table 4.2 with the well known FIR-MMSE-DFE. While the FIR-MMSE-DFE was originally derived for LTI channels [93], it can be straightforwardly extended to the LTV channel case. and then design a recursive algorithm to update the filter coefficients at the symbol rate assuming a fixed estimation delay Δ. In all simulations, BPSK symbols are transmitted over a noisy WSSUS Rayleigh-fading channel with uniform power profile (i.e., $\sigma_l^2 = N_h^{-1}$) that is generated using Jakes method [82]. Throughout, we assume IFDE uses an ICI radius of $D = \lceil f_{\mathsf{d}} T_{\mathsf{s}} PN \rceil$ and block overlap factor of $P = 2$. Both IFDE and FIR-MMSE-DFE designs are based on known time-domain coefficients $\{h_{n,l}\}_{l=0}^{N_h-1}$.

First, we establish IFDE-BDFE design rules for block length PN and number-of-iterations M. While we will see that smaller values of PN (for fixed N_h) are advantageous from a complexity standpoint (see Fig. 4.15), Fig. 4.7 suggest the choice $PN \geq 4N_h$ for good symbol error rate (SER) performance. With radix-2 FFTs in mind, we choose $PN = 2^{\lceil \log_2 4N_h \rceil}$ in the sequel. A related set of experiments in Fig. 4.8 has shown that SER performance improves with M up to about $M = 10$, after which there is little additional improvement. Interestingly, we find that, after 2 iterations, IFDE-BDFE gives approximately the same performance as FIR-MMSE-DFE. Hence, we focus on IFDE-BDFE-2 and IFDE-BDFE-10 in the sequel, i.e., IFDE-BDFE using $M = 2$ and $M = 10$, respectively.

Next, we establish FIR-MMSE-DFE design rules for feedforward filter length N_f and estimation delay Δ, assuming that the feedback filter is just long enough to cancel

all post-cursor ISI. To investigate the effect of Δ, we fixed $N_f = N_h$ and conducted experiments measuring MSE for several values of N_f (assuming $f_{\mathsf{d}}T_{\mathsf{s}} = 0.003$ and SNR=10dB). From Fig. 4.9 we can see that the choice $\Delta = N_f - 1$ maximized performance in every case, we adopt this rule. To investigate the effect of N_f, we fixed $\Delta = N_f - 1$ and conducted experiments measuring MSE at several values of SNR (when $f_{\mathsf{d}}T_{\mathsf{s}} = 0.003$ and $N_h = 64$). As shown in Fig. 4.10, in every case, performance increased with N_f, though the gains diminished rapidly when $N_f > N_h$. With complexity in mind, we adopt the rule $N_f = N_h$.

Having established IFDE-BDFE and FIR-MMSE-DFE design rules, we are ready to compare the two approaches in performance and complexity. In Fig. 4.11, we compare SER performances when $N_h = 64$, $f_{\mathsf{d}}T_{\mathsf{s}} \in \{0.001, 0.003, 0.0075\}$ over a wide range of SNR. Note that, at all $f_{\mathsf{d}}T_{\mathsf{s}}$, IFDE-BDFE-2 performs equivalently to FIR-MMSE-DFE whereas IFDE-BDFE-10 outperforms FIR-MMSE-DFE, significantly so when SNR$>$ 5. We also plot the matched-filter bound (MFB) [94]—the ultimate in (uncoded) receiver performance—which is not far from IFDE-BDFE-10.

4.7.2 IFDE with PACE

In this section, we investigate the performance of IFDE with PACE scheme when CSI is not available. The design rules for IFDE and FIR-MMSE-DFE are the same as in section 4.7.1. For simplicity, we only consider IFDE without BDFE in order to achieve lower computational complexity for PACE. All the simulation results presented here are based on averaging 100 packets, and each packet consists of 10 consecutive data frames.

First, we study the performance of joint IFDE/CE scheme in terms of SER and MSE of CE. For simplicity, we set $Q = P = 2$, and test the joint IFDE/CE scheme when $N_h = 32$, $f_{\mathsf{d}}T_{\mathsf{s}} \in \{0.001, 0.002\}$ and $N_h = 64$, $f_{\mathsf{d}}T_{\mathsf{s}} = 0.001$ over a wide range of SNR. As shown in Fig. 4.12, when $f_{\mathsf{d}}T_{\mathsf{s}} \times N_h$ is small, the IFDE/CE scheme suffers only 1dB loss compared with perfect CSI case (Genie), though the loss increases as $f_{\mathsf{d}}T_{\mathsf{s}} \times N_h$ increase. In Fig. 4.13, we show the experimental MSE of PACE versus theoretic predicted MSE from (4.31). They demonstrate a good match. Notice the parameters setting up $f_{\mathsf{d}}T_{\mathsf{s}} = 0.002$ with $N_h = 32$ and $f_{\mathsf{d}}T_{\mathsf{s}} = 0.001$ with $N_h = 64$ are close to the limit - Nyquist frequency ($\frac{1}{12N_h}$), therefore they produce similar MSE, which is much higher than the case when $Nh = 32$, $f_{\mathsf{d}}T_{\mathsf{s}} = 0.001$. This observation justify the increased performance loss of IFDE in more dispersive channels as shown in Fig. 4.12.

Second, we compare the performance of joint IFDE/CE scheme with FIR-MMSE-DFE plus RLS-CE (denoted as DFE-RLS in Fig. 4.14) [93,102]. Experimental results show that FIR-MMSE-DFE can not work well when $f_{\mathsf{d}}T_{\mathsf{s}} > 0.0005$ and $N_h = 32$, therefore we pick $f_{\mathsf{d}}T_{\mathsf{s}} \in \{0.0001, 0.0005\}$. For simplicity, we set $Q = 5$, $P = 2$, larger Q means higher data transmission rate versus pilots. Enough pilots symbols are inserted for the initialization of RLS-CE. From Fig. 4.14, we can see that joint IFDE/CE scheme performs much better than FIR-MMSE-DFE plus RLS-CE scheme and is close to the ideal case when perfect CSI is available.

Figure 4.15 examines the multiplies-per-symbol ratio of FIR-MMSE-DFE/RLS-CE to IFDE-2/PACE using the expressions in Table 4.3. Note that values > 1 in Fig. 4.15 imply a complexity *advantage* for IFDE/PACE, and that this complexity advantage increases with N_h and decreases with $f_{\mathsf{d}}T_{\mathsf{s}}$. Since FIR-MMSE-DFE

and IFDE-2 have similar performance, Fig. 4.15 constitutes a direct *complexity* comparison. A similar comparison in Fig. 4.16 between FIR-MMSE-DFE/RLS-CE and IFDE-10/PACE shows simultaneously complexity and performance advantages.

A final comment regarding the complexity comparison in Fig. 4.15 and Fig. 4.16 is in order. One could argue that due to imperfect CE, IFDE/PACE may not perform well in some range of $(f_{\mathsf{d}}T_{\mathsf{s}}, N_h)$, therefore the complexity gain at those points are meaningless. The question is what is the region in Fig. 4.15 and Fig. 4.16 that IFDE/PACE do perform well. To answer this question, first we investigate the relationship between $(f_{\mathsf{d}}T_{\mathsf{s}}, N_h)$ and CE error of PACE through numerical results. In Fig. 4.17, we plot the contours of theoretical MSE of PACE calculated from (4.31) versus $(f_{\mathsf{d}}T_{\mathsf{s}}, N_h)$ when $Q = P = 2$. The solid lines stand for contour and the dashed lines stand for points of constant product of $f_{\mathsf{d}}T_{\mathsf{s}}N_h$. As illustrated in Fig. 4.17, almost the same MSE can be achieved by PACE for those points with the same product of $f_{\mathsf{d}}T_{\mathsf{s}}N_h$. Bearing this conclusion in mind, we superimpose the curves $f_{\mathsf{d}}T_{\mathsf{s}}N_h = \{0.064, 0.016, 0.0032\}$ on Fig. 4.15 and Fig. 4.16, which corresponds to the Doppler and delay spread setting up in Fig. 4.12 and Fig. 4.14. We can see that the IFDE/PACE algorithm can enjoy significant cost savings compared with the DFE/RLS-CE scheme in a relative wide range.

4.8 Conclusion

In this chapter, we presented an iterative frequency domain equalization (IFDE) scheme for single-carrier transmissions over noisy doubly dispersive channels. Time-domain windowing is used to make the effective ICI/IBI response sparse, after which iterative symbol estimation is performed in the frequency domain. The estimation

algorithm leverages the finite-alphabet property of symbols, the sparse ICI/IBI structure, and the low computational cost of the FFT. Simulations show that with perfect CSI, the IFDE performs significantly better than the FIR-MMSE-DFE and within about 1 dB of the MFB over the SNR range of interest. A fast version of the IFDE algorithm was also derived and its complexity compared to that of a fast FIR-MMSE-DFE for LTV channels. When CSI is not available, a pilot-aided CE (PACE) is derived to work jointly with IFDE, which demonstrates remarkable performance gain versus the conventional FIR-MMSE-DFE plus RLS-CE scheme. In addition, the IFDE/PACE algorithm was found to yield significant cost savings relative to the FIR-MMSE-DFE plus RLS-CE scheme for reasonable channel lengths.

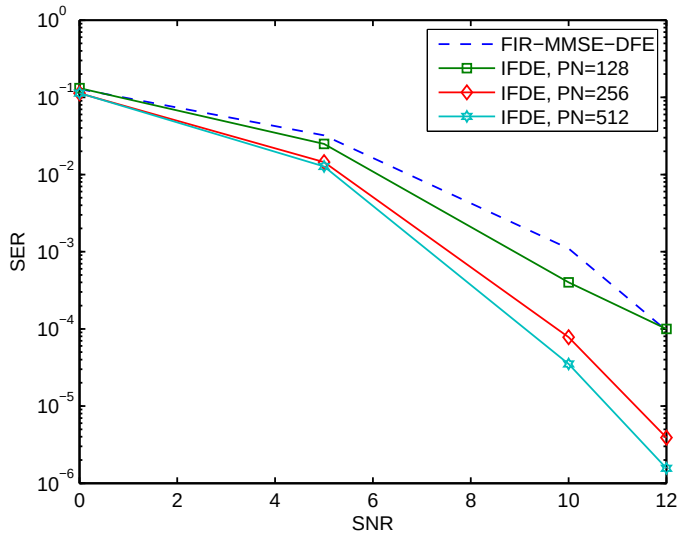

Figure 4.7: Symbol error rate for various PN when $M = 10$.

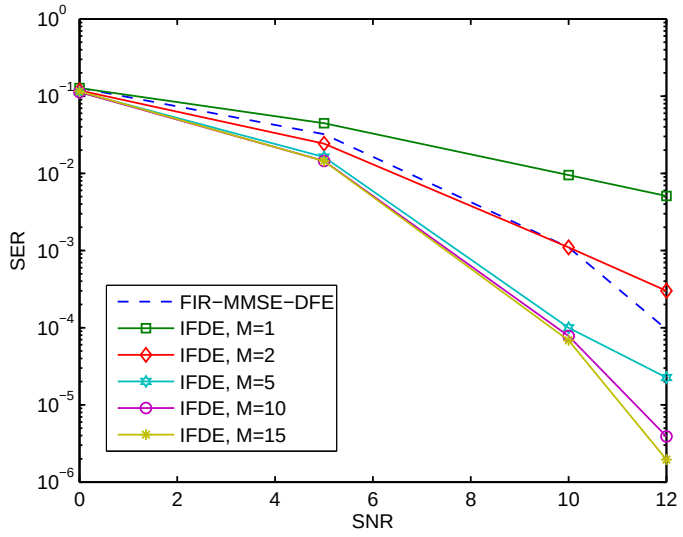

Figure 4.8: Symbol error rate for various M when $N_h = 64$ and $f_{\mathsf{d}}T_{\mathsf{s}} = 0.003$.

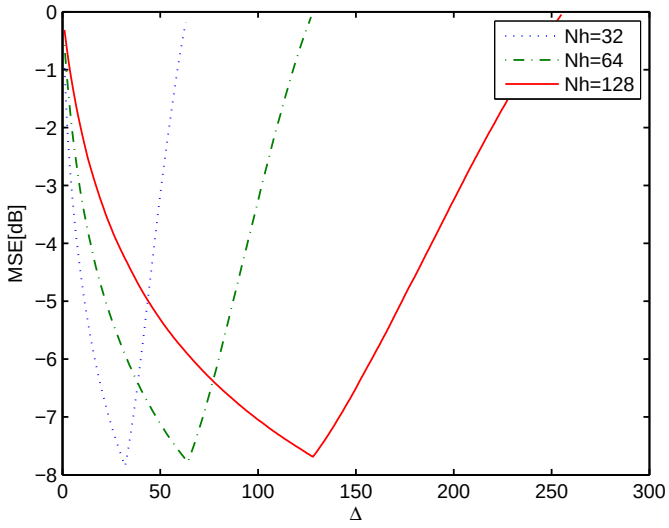

Figure 4.9: MSE of FIR-MMSE-DFE versus Δ at $f_{\mathsf{d}}T_{\mathsf{s}} = 0.003$, SNR= 10, and $N_f = N_h$.

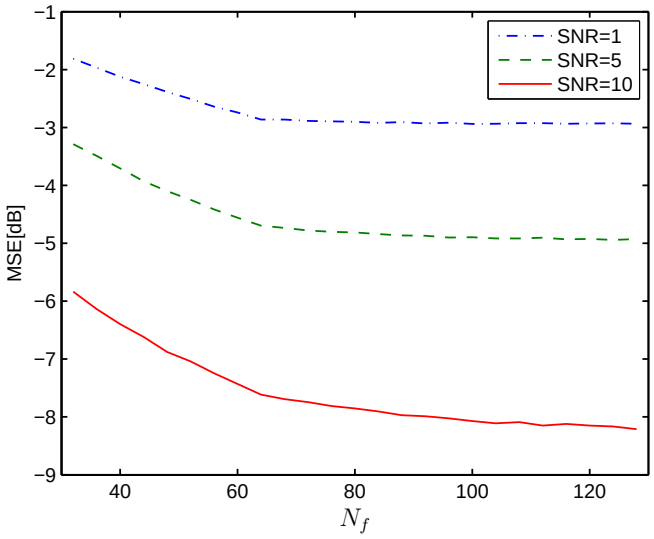

Figure 4.10: MSE of FIR-MMSE-DFE versus N_f when $f_{\mathsf{d}}T_{\mathsf{s}} = 0.003$, $N_h = 64$, and $\Delta = N_f - 1$.

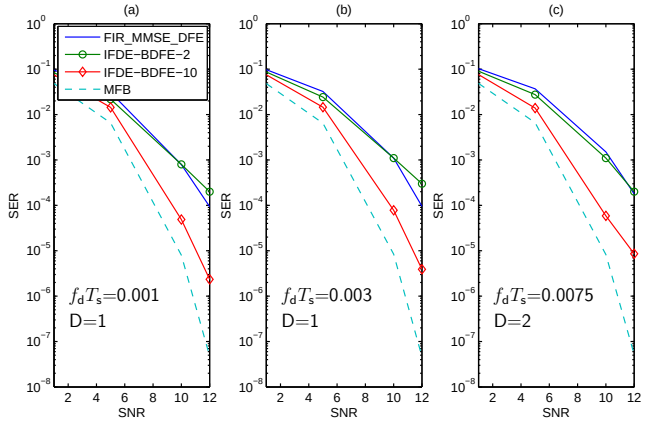

Figure 4.11: SER versus SNR for $N_h = 64$ and various $f_{\mathsf{d}}T_{\mathsf{s}}$ with perfect CSI.

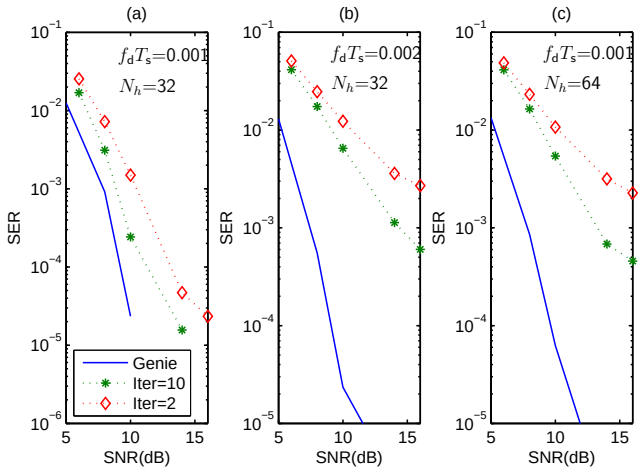

Figure 4.12: SER of IFDE/PACE versus SNR for various N_h and f_dT_s.

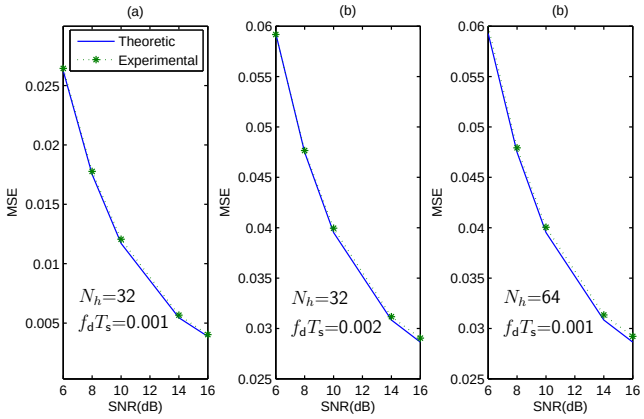

Figure 4.13: MSE of PACE versus SNR for various N_h and f_dT_s.

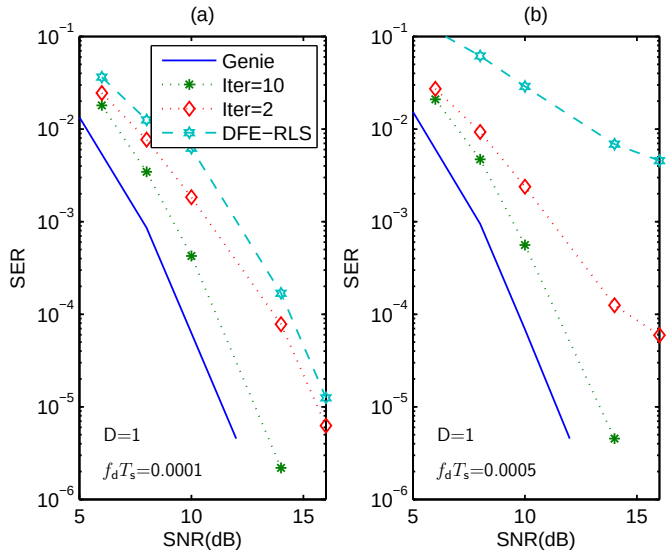

Figure 4.14: SER versus SNR for $N_h = 32$ and various $f_{\mathsf{d}}T_{\mathsf{s}}$.

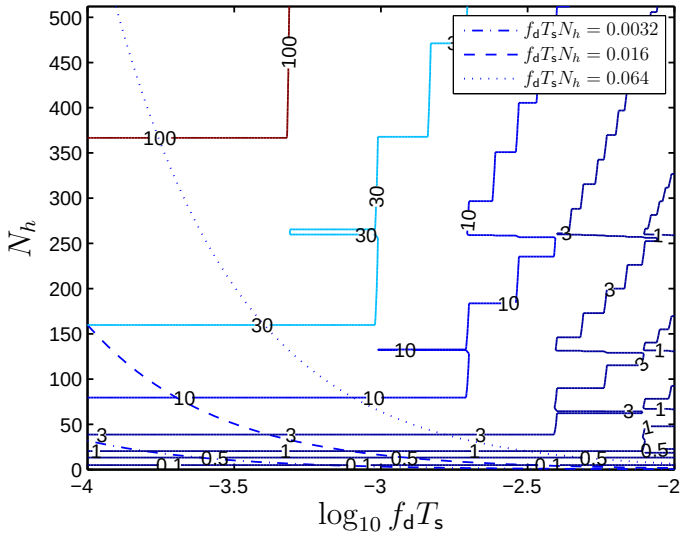

Figure 4.15: Computational complexity ratio of FIR-MMSE-DFE/RLS-CE to IFDE-noBDFE-2/PACE.

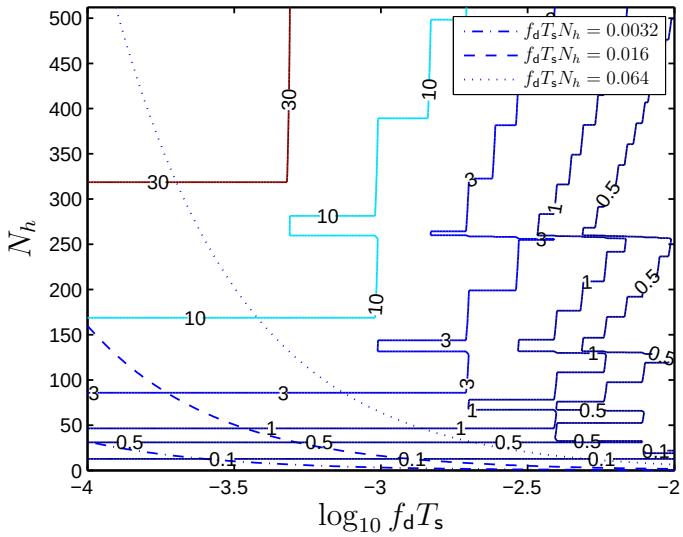

Figure 4.16: Computational complexity ratio of FIR-MMSE-DFE/RLS-CE to IFDE-noBDFE-10/PACE.

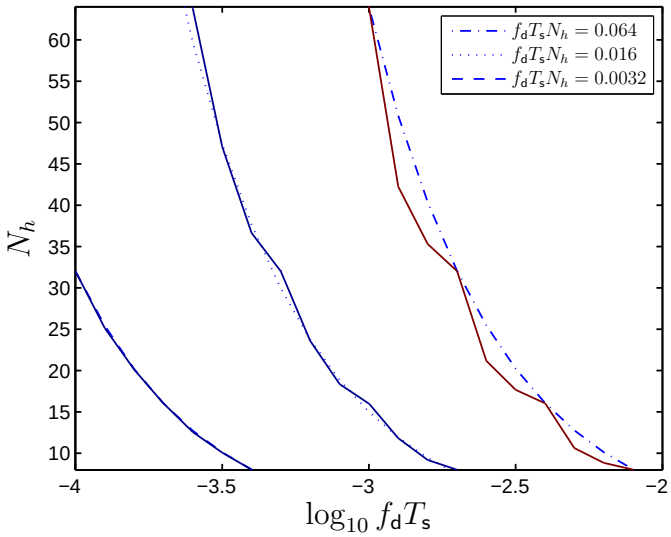

Figure 4.17: Contour of theoretical MSE of PACE.

$L_l^{(0)} = 0 \ \forall l$

for $n = 0, 1, 2, \ldots$

  for $l = 0 \ldots PN-1$,

    $\bar{s}_l^{(n)} = \tanh(L_l^{(n)}/2)$

    $v_l^{(n)} = 1 - (\bar{s}_l^{(n)})^2$

  end

  $\overline{\boldsymbol{t}}^{(n)} = \boldsymbol{F}\overline{\boldsymbol{s}}^{(n)}$

  for $k = 0 \ldots PN-1$,

    $\boldsymbol{g}_k^{(n)} = \left(\boldsymbol{\mathcal{H}}_k \boldsymbol{F}\, \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{F}^H \boldsymbol{\mathcal{H}}_k^H + \sigma^2 \boldsymbol{C}_k \boldsymbol{C}_k^H\right)^{-1}$

    $\qquad \times \boldsymbol{\mathcal{H}}_k \boldsymbol{F}\, \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{F}^H \boldsymbol{i}_k$

    $\hat{t}_k^{(n)} = \bar{t}_k^{(n)} + \boldsymbol{g}_k^{(n)H}(\boldsymbol{x}_k - \boldsymbol{\mathcal{H}}_k \overline{\boldsymbol{t}}^{(n)})$

  end

  $\hat{\boldsymbol{s}}^{(n)} = \boldsymbol{F}^H \hat{\boldsymbol{t}}^{(n)}$

  $\boldsymbol{Q}^{(n)} = \boldsymbol{F}^H \left(\sum_{k=0}^{PN-1} \boldsymbol{\mathcal{H}}_k^H \boldsymbol{g}_k^{(n)} \boldsymbol{i}_k^H \right) \boldsymbol{F}$

  $\boldsymbol{P}^{(n)} = \left(\sum_{k=0}^{PN-1} \boldsymbol{C}_k^H \boldsymbol{g}_k^{(n)} \boldsymbol{i}_k^H \right) \boldsymbol{F}$

  for $l = 0 \ldots PN-1$,

    $$L_l^{(n+1)} = L_l^{(n)} + \frac{4\left(\mathrm{Re}\{Q_{l,l}^{(n)}(\hat{s}_l^{(n)} - \bar{s}_l^{(n)})\} + |Q_{l,l}^{(n)}|^2 \bar{s}_l^{(n)}\right)}{\boldsymbol{q}_l^{(n)H}\, \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{q}_l^{(n)} - |Q_{l,l}^{(n)}|^2 v_l^{(n)} + \sigma^2 \|\boldsymbol{p}_l\|^2}$$

  end

end

Table 4.1: Summary of Iterative Symbol Detection.

	Step	Cost Per Step
1	$\boldsymbol{H}(i,\ell)$	$2\widetilde{D}(L_{\mathsf{pst}}+1)PNN_h$
2	compute $\boldsymbol{x}(i)$ from (4.56)	$(\widetilde{D}L_{\mathsf{pst}}+1)PN+\frac{3}{2}PN\log_2 PN$
3	$\boldsymbol{\Sigma}=\sqrt{PN}\boldsymbol{C}_0\boldsymbol{C}_0^H$	$2PN+\frac{1}{2}PN\log_2 PN$
	$L_l^{(0)}=0\ \forall l$	0
	for $n=0,1,\ldots,M-1$	
	for $l=0,1,\ldots,PN-1$	
	$\bar{s}_l^{(n)}=\tanh(L_l^{(n)}/2)$	1
	$v_l^{(n)}=1-(\bar{s}_l^{(n)})^2$	1
	end	
	$\overline{\boldsymbol{t}}^{(n)}=\boldsymbol{F}\overline{\boldsymbol{s}}^{(n)}$	$\frac{1}{2}PN\log_2 PN$
	$\boldsymbol{u}^{(n)}=\boldsymbol{F}\boldsymbol{v}^{(n)}$	$\frac{1}{2}PN\log_2 PN$
4	$\boldsymbol{R}_0^{(n)}=\boldsymbol{\mathcal{H}}_0\,\mathcal{C}(\boldsymbol{u}^{(n)})\boldsymbol{\mathcal{H}}_0^H+\boldsymbol{\Sigma}$	$3\widetilde{D}^3$
	for $k=0,1,\ldots,PN-1$	
5	compute $(\boldsymbol{R}_k^{(n)})^{-1}$	$\min\left\{2\widetilde{D}^2+\frac{1}{3}\widetilde{D}^3,\ 7\widetilde{D}^2\right\}$
6	$\boldsymbol{g}_k^{(n)}=(\boldsymbol{R}_k^{(n)})^{-1}\boldsymbol{\mathcal{H}}_k\,\mathcal{C}(\boldsymbol{u}^{(n)})\boldsymbol{i}_k$	$2\widetilde{D}^2$
	$\hat{t}_k^{(n)}=\bar{t}_k^{(n)}+\boldsymbol{g}_k^{(n)H}(\boldsymbol{x}_k-\boldsymbol{\mathcal{H}}_k\overline{\boldsymbol{t}}^{(n)})$	$\widetilde{D}^2$
	end	
	$\hat{\boldsymbol{s}}^{(n)}=\boldsymbol{F}^H\hat{\boldsymbol{t}}^{(n)}$	$PN\log_2 PN$
7	compute $\{y_l^{(n)}\}_{l=0}^{PN-1},\{\boldsymbol{\alpha}_d\}_{d=-2D}^{2D}$ from (4.57)-(4.59)	$(4\widetilde{D}^2+2\widetilde{D})PN\log_2 PN+5\widetilde{D}^2PN$
8	compute $\{z_l^{(n)}\}_{l=0}^{PN-1}$ from (4.60)-(4.62)	$\widetilde{D}^2PN+(\widetilde{D}^2+\widetilde{D})PN\log_2 PN$
	for $l=0,1,\ldots,PN-1$	
9	$Q_{l,l}^{(n)}=\frac{1}{\sqrt{PN}}\sum_{d=-2D}^{2D}[\boldsymbol{\alpha}_d^{(n)}]_0\,e^{j\frac{2\pi}{PN}ld}$	$2\widetilde{D}$
	$L_l^{(n+1)}=L_l^{(n)}+4\frac{\mathrm{Re}\{Q_{l,l}^{(n)}(\hat{s}_l^{(n)}-\bar{s}_l^{(n)})\}+\lvert Q_{l,l}^{(n)}\rvert^2\bar{s}_l^{(n)}}{y_l^{(n)}-\lvert Q_{l,l}^{(n)}\rvert^2v_l^{(n)}+\sigma^2z_l^{(n)}}$	6
	end	
	end	

Table 4.2: Fast Implementation of the Iterative Symbol Detector

IFDE-noBDFE: $3M\widetilde{D}^3/N + \left[2 + M(5\widetilde{D}^2 + 3\widetilde{D} + 2)\right] P \log_2 PN$ $+ \left[3 + M\left(\min\{\frac{1}{3}\widetilde{D}^3, 5\widetilde{D}^2\} + 11\widetilde{D}^2 + 2\widetilde{D} + 8\right)\right] P$ PACE: $2\widetilde{D}PN_h$
IFDE-BDFE: $3M\widetilde{D}^3/N + \left[2 + M(5\widetilde{D}^2 + 3\widetilde{D} + 2)\right] P \log_2 PN$ $+ \left[3 + L_{\mathsf{pst}}\widetilde{D} + M\left(\min\{\frac{1}{3}\widetilde{D}^3, 5\widetilde{D}^2\} + 11\widetilde{D}^2 + 2\widetilde{D} + 8\right)\right] P$ PACE: $2\widetilde{D}(L_{\mathsf{pst}} + 1)PN_h$
FIR-MMSE-DFE: $\frac{9}{2}N_f^2 - \frac{1}{2}N_f - 1$ RLS-CE: $N_h^2 + 3N_h$

Table 4.3: Relative Algorithm Complexity (Per Symbol).

Appendix

4.A Signal Energy Distribution

$$
\begin{aligned}
\mathcal{E}_{\mathsf{ss},n} &:= \sum_{d=0}^{PN-1} \mathrm{E}\left\{ \left| \sum_{k=0}^{PN-1} H_{d-k,k}(i,0) \cdot \frac{1}{\sqrt{PN}} e^{-j\frac{2\pi}{PN}nk} s_n(i) \right|^2 \right\} \\
&= \frac{1}{PN} \sum_{d=0}^{PN-1} \mathrm{E}\left\{ \left| \sum_{k=0}^{PN-1} \frac{1}{PN} \sum_{m} \sum_{l=0}^{N_h-1} h_{iN+m+N_o,l} b_m a_{m-l} e^{-j\frac{2\pi}{PN}m(d-k)} e^{-j\frac{2\pi}{PN}k(l+n)} \right|^2 \right\} \\
&= \frac{1}{(PN)^3} \sum_{d=0}^{PN-1} \sum_{k,k'=0}^{PN-1} \sum_{m,m'} \sum_{l=0}^{N_h-1} \rho_{m-m'} \sigma_l^2 b_m b_{m'}^* a_{m-l} a_{m'-l}^* \\
&\qquad \cdot e^{-j\frac{2\pi}{PN}d(m-m')} e^{j\frac{2\pi}{PN}k(m-l-n)} e^{-j\frac{2\pi}{PN}k'(m'-l-n)} \\
&= \frac{1}{(PN)^3} \sum_{d=0}^{PN-1} \sum_{m,m'} \sum_{l=0}^{N_h-1} \rho_{m-m'} \sigma_l^2 b_m b_{m'}^* a_{m-l} a_{m'-l}^* e^{-j\frac{2\pi}{PN}d(m-m')} \\
&\qquad \cdot \sum_{k,k'=0}^{PN-1} e^{j\frac{2\pi}{PN}k(m-l-n)} e^{-j\frac{2\pi}{PN}k'(m'-l-n)} \\
&= \frac{1}{(PN)^3} \sum_{d=0}^{PN-1} \sum_{m,m'} \sum_{l=0}^{N_h-1} \rho_{m-m'} \sigma_l^2 b_m b_{m'}^* a_{m-l} a_{m'-l}^* e^{-j\frac{2\pi}{PN}d(m-m')} \cdot PN\delta_{m-l-n} \\
&\qquad \cdot PN\delta_{m'-l-n} \\
&= \sum_{l=0}^{N_h-1} \sigma_l^2 |b_{l+n}|^2 a_n^2 \\
&= \sum_{l=0}^{N_h-1} \sigma_l^2 |b_{l+n}|^2
\end{aligned}
\tag{4.53}
$$

4.B Conditional Mean and Variance

From (4.38), (4.40), and the definition of $\mu_l^{(n)}(b)$,

$$\begin{aligned}\mu_l^{(n)}(b) &= \boldsymbol{i}_l^H \boldsymbol{F}^H \sum_k \boldsymbol{i}_k \,\mathrm{E}\{\hat{t}_k^{(n)}|s_l = b\} \\ &= \boldsymbol{i}_l^H \boldsymbol{F}^H \sum_k \boldsymbol{i}_k \Big(\bar{t}_k^{(n)} + \boldsymbol{g}_k^{(n)H}\big(\mathrm{E}\{\boldsymbol{x}_k|s_l = b\} - \boldsymbol{\mathcal{H}}_k \overline{\boldsymbol{t}}^{(n)}\big)\Big) \\ &= \bar{s}_l^{(n)} + \boldsymbol{i}_l^H \boldsymbol{Q}^{(n)H} \boldsymbol{i}_l (b - \bar{s}_l^{(n)})\end{aligned}$$

which leads to (4.42). In the last step above, we used the fact that $\mathrm{E}\{\boldsymbol{x}_k|s_l = b\} = \boldsymbol{\mathcal{H}}_k \boldsymbol{F}\big(\overline{\boldsymbol{s}}^{(n)} + \boldsymbol{i}_l(b - \bar{s}_l^{(n)})\big) = \boldsymbol{\mathcal{H}}_k \overline{\boldsymbol{t}}^{(n)} + \boldsymbol{\mathcal{H}}_k \boldsymbol{F} \boldsymbol{i}_l (b - \bar{s}_l^{(n)})$. Next we find an expression for $[\sigma_l^{(n)}(b)]^2$. Before doing so, however, it will be convenient to note from (4.38) and (4.40) that

$$\begin{aligned}\hat{s}_l^{(n)} &= \boldsymbol{i}_l^H \boldsymbol{F}^H \sum_k \boldsymbol{i}_k \Big(\bar{t}_k^{(n)} + \boldsymbol{g}_k^{(n)H}\big(\boldsymbol{x}_k - \boldsymbol{\mathcal{H}}_k \overline{\boldsymbol{t}}^{(n)}\big)\Big) \\ &= \boldsymbol{i}_l^H \boldsymbol{F}^H \sum_k \boldsymbol{i}_k \Big(\bar{t}_k^{(n)} + \boldsymbol{g}_k^{(n)H}\big(\boldsymbol{\mathcal{H}}_k \boldsymbol{F} \boldsymbol{s} + \boldsymbol{C}_k \boldsymbol{\nu} - \boldsymbol{\mathcal{H}}_k \boldsymbol{F} \overline{\boldsymbol{s}}^{(n)}\big)\Big) \\ &= \bar{s}_l^{(n)} + \boldsymbol{i}_l^H \boldsymbol{Q}^{(n)H}(\boldsymbol{s} - \overline{\boldsymbol{s}}^{(n)}) + \boldsymbol{i}_l^H \boldsymbol{P}^{(n)H} \boldsymbol{\nu} \\ &= \mu_l^{(n)}(b) + \boldsymbol{i}_l^H \boldsymbol{Q}^{(n)H}\big(\boldsymbol{s} - \overline{\boldsymbol{s}}^{(n)} + \boldsymbol{i}_l(\bar{s}_l^{(n)} - b)\big) + \boldsymbol{i}_l^H \boldsymbol{P}^{(n)H} \boldsymbol{\nu} \qquad (4.54)\end{aligned}$$

and that, since $\mathrm{E}\{\boldsymbol{s}|s_l = b\} = \overline{\boldsymbol{s}}^{(n)} - \boldsymbol{i}_l(\bar{s}_l^{(n)} - b)$,

$$\begin{aligned}&\mathrm{E}\Big\{\big(\boldsymbol{s} - \overline{\boldsymbol{s}}^{(n)} + \boldsymbol{i}_l(\bar{s}_l^{(n)} - b)\big)\big(\boldsymbol{s} - \overline{\boldsymbol{s}}^{(n)} + \boldsymbol{i}_l(\bar{s}_l^{(n)} - b)\big)^H |s_l = b\Big\} \\ &\quad = \mathrm{cov}(\boldsymbol{s}, \boldsymbol{s}|s_l = b) \\ &\quad = \mathcal{D}(\boldsymbol{v}^{(n)}) - \boldsymbol{i}_l \boldsymbol{i}_l^H v_l^{(n)}. \qquad (4.55)\end{aligned}$$

Using (4.54), (4.55), and the definition of $\sigma_l^{(n)}(b)$,

$$\begin{aligned}[\sigma_l^{(n)}(b)]^2 &= \mathrm{E}\Big\{\big(\hat{s}_l^{(n)} - \mu_l^{(n)}(b)\big)\big(\hat{s}_l^{(n)} - \mu_l^{(n)}(b)\big)^H |s_l = b\Big\} \\ &= \boldsymbol{i}_l^H \boldsymbol{Q}^{(n)H}\Big(\mathcal{D}(\boldsymbol{v}^{(n)}) - \boldsymbol{i}_l \boldsymbol{i}_l^H v_l^{(i)}\Big)\boldsymbol{Q}^{(n)} \boldsymbol{i}_l + \sigma^2 \boldsymbol{i}_l^H \boldsymbol{P}^{(n)H} \boldsymbol{P}^{(n)} \boldsymbol{i}_l\end{aligned}$$

which leads to (4.43).

4.C Fast-IFDE Details

The details of each step are enumerated below in correspondence with the left column of Table 4.2. For brevity, we use $\widetilde{D} := 2D + 1$ in the sequel. We make frequent use of the property $\boldsymbol{F}\,\mathcal{D}(\boldsymbol{a})\boldsymbol{F}^H = \mathcal{C}(\boldsymbol{F}\boldsymbol{a}/\sqrt{PN})$. Finally, we assume that PN-length FFTs require $\frac{1}{2}PN\log_2(PN)$ and $PN\log_2(PN)$ complex multiplies for real- and complex-valued inputs, respectively (as per the radix-2 Cooley-Tukey algorithm [103]).

Detail 1: At each block index i, we must compute the frequency domain coefficients $\boldsymbol{\mathcal{H}}(i,0)$, or $\{\boldsymbol{\mathcal{H}}(i,\ell)\}_{\ell=0}^{L_{\mathsf{pst}}}$ when BDFE is used, using the PACE. From (4.25), we can see that $\boldsymbol{R}_{\underline{\boldsymbol{g}}\underline{\boldsymbol{r}}_{\mathsf{p}}}\boldsymbol{R}^{-1}_{\underline{\boldsymbol{r}}_{\mathsf{p}}\underline{\boldsymbol{r}}_{\mathsf{p}}}$ only need to be computed once at the beginning, therefore the computations needed to estimate $\{\boldsymbol{\mathcal{H}}(i,\ell)\}$ from $\underline{\boldsymbol{r}}_{\mathsf{p}}^{(q)}$ is $2\widetilde{D}PNN_h$ complex multiplies for each ℓ.

Detail 2: In the BDFE case, the frequency domain observation is computed as

$$\boldsymbol{x}(i) \;=\; \boldsymbol{F}\boldsymbol{J}\,\mathcal{D}(\boldsymbol{b})\boldsymbol{r}(i) - \sum_{\ell=1}^{L_{\mathsf{pst}}} \boldsymbol{\mathcal{H}}(i,\ell)\hat{\hat{\boldsymbol{t}}}(i-\ell P), \tag{4.56}$$

where $\hat{\hat{\boldsymbol{t}}}(i) := \boldsymbol{F}\hat{\boldsymbol{s}}(i)$. The non-BDFE case is similar, but without the IBI cancellation. The first term in $\boldsymbol{x}(i)$ requires $N_b + PN\log_2 PN$ multiplications per block to compute, while the second requires $L_{\mathsf{pst}}\widetilde{D}PN$ since $\boldsymbol{\mathcal{H}}(i,\ell)$ contains only $\widetilde{D}PN$ non-zero elements. Since $\hat{\hat{\boldsymbol{t}}}(i)$ needs to be computed only when i is a multiple of P, it requires an average of $\frac{1}{P}PN\log_2 PN$ multiplications per block. Using $P = 2$ and the approximation $N_b \approx PN$, we get a total of $(\widetilde{D}L_{\mathsf{pst}} + 1)PN + 1.5PN\log_2 PN$ multiplications per block.

Detail 3: From (4.11) and the property $\boldsymbol{J}\,\mathcal{D}(\boldsymbol{b}) = \mathcal{D}(\boldsymbol{J}\boldsymbol{b})$, it follows that $\boldsymbol{C}\boldsymbol{C}^H = \boldsymbol{F}\,\mathcal{D}(\sigma^2\boldsymbol{J}\boldsymbol{b}\odot\boldsymbol{J}\boldsymbol{b}^*)\boldsymbol{F}^H = \mathcal{C}\left(\sqrt{PN}\sigma^2\boldsymbol{F}\left(\boldsymbol{J}\boldsymbol{b}\odot\boldsymbol{J}\boldsymbol{b}^*\right)\right)$. Thus, the PN coefficients that specify $\boldsymbol{C}\boldsymbol{C}^H$ can be computed in roughly $2PN + \frac{1}{2}PN\log_2 PN$ multiplies. Notice that $\boldsymbol{C}_k\boldsymbol{C}_k^H$ is a sub-block of $\boldsymbol{C}\boldsymbol{C}^H$, and that the Toeplitz nature of $\boldsymbol{C}\boldsymbol{C}^H$ implies that $\boldsymbol{C}_k\boldsymbol{C}_k^H$ is identical for every k.

Detail 4: This step initializes the recursive computation of $\boldsymbol{R}_k^{(n)} := \sqrt{PN}\left(\boldsymbol{\mathcal{H}}_k \boldsymbol{F}\, \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{F}^H \boldsymbol{\mathcal{H}}_k^H + \boldsymbol{C}_k \boldsymbol{C}_k^H\right)$, where we note $\sqrt{PN}\boldsymbol{F}\, \mathcal{D}(\boldsymbol{v}^{(n)})\boldsymbol{F}^H = \mathcal{C}(\boldsymbol{u}^{(n)})$. For computation of $\boldsymbol{\mathcal{H}}_0\, \mathcal{C}(\boldsymbol{u}^{(n)}) \boldsymbol{\mathcal{H}}_0^H$, we first compute $\boldsymbol{\mathcal{H}}_0\, \mathcal{C}(\boldsymbol{u}^{(n)})$, then post-multiply the result by $\boldsymbol{\mathcal{H}}_0^H$. But since $\boldsymbol{\mathcal{H}}_0^H$ contains only $4D+1 \approx 2\widetilde{D}$ non-zero rows, only $2\widetilde{D}$ non-zero columns of $\boldsymbol{\mathcal{H}}_0\, \mathcal{C}(\boldsymbol{u}^{(n)})$ need be computed. This requires $2\widetilde{D}^3$ multiplications, since $\boldsymbol{\mathcal{H}}_0$ contains $\widetilde{D}$ rows, each with only $\widetilde{D}$ non-zero elements. Using a similar reasoning, the post-multiplication also requires $\widetilde{D}^3$ multiplications.

Detail 5: $\boldsymbol{R}_k^{(n)}$ can be inverted directly or recursively since $\boldsymbol{R}_{k+1}^{(n)} = \begin{bmatrix} \boldsymbol{\Theta}_k & \tilde{\boldsymbol{\theta}}_k \\ \tilde{\boldsymbol{\theta}}_k^H & \tilde{\theta}_k \end{bmatrix}$ when $\boldsymbol{R}_k^{(n)} = \begin{bmatrix} \theta_k & \boldsymbol{\theta}_k^H \\ \boldsymbol{\theta}_k & \boldsymbol{\Theta}_k \end{bmatrix}$. In the direct method, we first compute $[\tilde{\boldsymbol{\theta}}_k^t \;\; \tilde{\theta}_k]^t$ to obtain $\boldsymbol{R}_{k+1}^{(n)}$ from $\boldsymbol{R}_k^{(n)}$. Cost-wise, this is similar to computing one column (i.e., $1/\widetilde{D}$ of the total elements) of $\boldsymbol{R}_0^{(n)}$, requiring $2\widetilde{D}^2$ multiplies. The direct inversion of Hermitian $\boldsymbol{R}_{k+1}^{(n)}$ then requires an additional $\frac{1}{3}\widetilde{D}^3$ multiplies (using LDL* factorization [98]). The procedure for recursive computation of $(\boldsymbol{R}_{k+1}^{(n)})^{-1}$ follows directly from the well-known block-matrix inversion formula [104] $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1}(\boldsymbol{I} + \boldsymbol{B}\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}) & -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{P}^{-1} \\ -\boldsymbol{P}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{P}^{-1} \end{bmatrix}$, where $\boldsymbol{P} := \boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}$, and is detailed in Table 4.4. In summary, the total cost of the direct and recursive inversions are approximately $2\widetilde{D}^2 + \frac{1}{3}\widetilde{D}^3$ and $7\widetilde{D}^2$ multiplications, respectively.

Detail 6: Since $\boldsymbol{\mathcal{H}}_k$ contains $\widetilde{D}$ rows, each with only $\widetilde{D}$ nonzero elements, the calculation of $\boldsymbol{\mathcal{H}}_k \tilde{\boldsymbol{V}}^{(n)} \boldsymbol{i}_k$ consumes only $\widetilde{D}^2$ multiplies. Multiplication by $(\boldsymbol{R}_k^{(n)})^{-1}$ consumes an additional $\widetilde{D}^2$.

Detail 7: LLR updating requires $\{y_k^{(n)}\}_{k=0}^{PN-1}$, where $y_k^{(n)} := \boldsymbol{q}_k^{(n)H}\, \mathcal{D}(\boldsymbol{v}^{(n)}) \boldsymbol{q}_k^{(n)}$. Note that the explicit calculation of $\boldsymbol{Q}^{(n)}$, as defined in (4.44), would involve $2PN$

FFTs of length PN, and thus a total complexity of $\mathcal{O}(P^2N^2\log_2 PN)$. In Appendix 4.C.1 we show that

$$\boldsymbol{y}^{(n)} = \frac{1}{\sqrt{PN}} \sum_{d,l=-2D}^{2D} \boldsymbol{F}\,\mathcal{D}\left(\boldsymbol{T}_{l-d}\overline{\boldsymbol{u}}^{(n)}\right)\boldsymbol{F}^H(\boldsymbol{\alpha}_d^{(n)} \odot \boldsymbol{\alpha}_l^{(n)*}) \tag{4.57}$$

where $[\boldsymbol{y}^{(n)}]_k = y_k^{(n)}$, $\overline{\boldsymbol{u}}^{(n)} := \boldsymbol{F}^H\boldsymbol{v}^{(n)}$, $\boldsymbol{T}_k := \mathcal{C}(\boldsymbol{i}_{\langle k\rangle_{PN}})$ is the right circular k-shift matrix, and where

$$\boldsymbol{\alpha}_d^{(n)} = \boldsymbol{F}\mathrm{diag}_d(\boldsymbol{G}^{(n)}) \tag{4.58}$$

$$\boldsymbol{G}^{(n)} = \sum_{k=0}^{PN-1} \boldsymbol{\mathcal{H}}_k^H \boldsymbol{g}_k^{(n)} \boldsymbol{i}_k \tag{4.59}$$

Note that $\overline{\boldsymbol{u}}^{(n)}$ is simply a rearrangement of $\boldsymbol{u}^{(n)}$. The k^{th} column of $\boldsymbol{G}^{(n)}$ equals $\boldsymbol{\mathcal{H}}_k^H\boldsymbol{g}_k^{(n)}$ and requires $\widetilde{D}^2$ multiplies to compute, and so $\boldsymbol{G}^{(n)}$ requires $PN\widetilde{D}^2$ multiplies to compute. Computation of $\{\boldsymbol{\alpha}_d\}_{d=-2D}^{2D}$ involves $4D+1 \approx 2\widetilde{D}$ FFTs for a total cost of $2\widetilde{D}PN\log_2 PN$ multiplies. For each (d,l) pair, the computation of (4.57) requires an additional $2PN + 2PN\log_2 PN$ multiplies. However, due to conjugate symmetry, only about half of the $\approx 4\widetilde{D}^2$ pairs need be evaluated. Hence, using (4.57) rather than direct computation of $\boldsymbol{Q}^{(n)}$, the calculation of $\{y_k^{(n)}\}_{k=0}^{PN-1}$ requires only about $4\widetilde{D}^2(PN + PN\log_2 PN) + 2\widetilde{D}PN\log_2 PN + PN\widetilde{D}^2$, or $5\widetilde{D}^2PN + (4\widetilde{D}^2 + 2\widetilde{D})PN\log_2 PN$, multiplies.

Detail 8: LLR updating also requires $\{z_k^{(n)}\}_{k=0}^{PN-1}$, where $z_k^{(n)} := \|\boldsymbol{p}_k^{(n)}\|^2$. In Appendix 4.C.2 we show that

$$\boldsymbol{z}^{(n)} = \frac{1}{\sqrt{PN}} \sum_{d,l=-D}^{D} \boldsymbol{F}\,\mathcal{D}\left(\boldsymbol{T}_{l-d}\boldsymbol{F}^H\overline{\boldsymbol{b}}\right)\boldsymbol{F}^H(\overline{\boldsymbol{\alpha}}_d^{(n)} \odot \overline{\boldsymbol{\alpha}}_l^{(n)*}), \tag{4.60}$$

where $[\boldsymbol{z}^{(n)}]_k = z_k^{(n)}$, $[\overline{\boldsymbol{b}}]_m := \sum_{n=0}^{N_b-1} |b_{\langle n\rangle_{PN}}|^2\delta_{\langle n\rangle_{PN}-m}$, and

$$\overline{\boldsymbol{\alpha}}_d^{(n)} = \boldsymbol{F}\mathrm{diag}_d(\overline{\boldsymbol{G}}^{(n)}) \tag{4.61}$$

$$\overline{\boldsymbol{G}}^{(n)} = \sum_{k=0}^{PN-1} \overline{\boldsymbol{g}}_k^{(n)} \boldsymbol{i}_k \tag{4.62}$$

Note that $\boldsymbol{F}^H\overline{\boldsymbol{b}}$ can be computed in advance, $\overline{\boldsymbol{G}}^{(n)}$ requires no computation, and $\{\overline{\boldsymbol{\alpha}}_d^{(n)}\}_{d=-D}^{D}$ involves $\widetilde{D}$ FFTs, for a total cost of $\widetilde{D}PN\log_2 PN$ multiplies. For each

(d, l) pair, (4.60) requires an additional $2PN + 2PN \log_2 PN$ multiplies, but only about half of the $\widetilde{D}^2$ pairs need be evaluated (due to conjugate symmetry). Hence, calculation of $\{z_k^{(n)}\}_{k=0}^{PN-1}$ requires about $\frac{1}{2}\widetilde{D}^2(2PN + 2PN \log_2 PN) + \widetilde{D}PN \log_2 PN$, or $\widetilde{D}^2 PN + (\widetilde{D}^2 + \widetilde{D})PN \log_2 PN$, multiplies.

Detail 9: LLR updating also requires $\{Q_{k,k}^{(n)}\}_{k=0}^{PN-1}$. From (4.44), (4.58), (4.59), and Lemma 1, it follows that

$$Q_{k,k}^{(n)} \;=\; \frac{1}{\sqrt{PN}} \sum_{d=-2D}^{2D} [\boldsymbol{\alpha}_d^{(n)}]_0 \, e^{j\frac{2\pi}{PN}kd}. \tag{4.63}$$

As reported in Table 4.2, direct evaluation of (4.63) requires $4D + 1 \approx 2\widetilde{D}$ multiplies for each k. Note that, if $2\widetilde{D} > \log_2 PN$, it would be more efficient to compute $\{Q_{k,k}^{(n)}\}_{k=0}^{PN-1}$ using a single PN-point FFT. However, since the cost of this step is relatively small, the difference is insignificant.

4.C.1 Derivation of (4.57)

Here we derive an expression for $\boldsymbol{y}^{(n)}$ enabling fast computation. First, however, we present a useful lemma. Without loss of generality, we omit superscripts in this appendix.

Lemma 1. *If $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{PN \times PN}$ has the banded structure of Fig. 4.2 with $2\rho D + 1$ non-zero diagonals, and if $\boldsymbol{B} = \boldsymbol{F}^H \boldsymbol{\mathcal{H}} \boldsymbol{F}$, then*

$$[\boldsymbol{B}]_{n,m} \;=\; \frac{1}{\sqrt{PN}} \sum_{d=-\rho D}^{\rho D} e^{j\frac{2\pi}{PN}nd} [\boldsymbol{F}\, diag_d(\boldsymbol{\mathcal{H}})]_{\langle m-n \rangle_{PN}}.$$

Proof. Denote $\boldsymbol{a}_d = \mathrm{diag}_d(\boldsymbol{\mathcal{H}})$, so that $[\boldsymbol{\mathcal{H}}]_{n,m} = a_{n-m,m}$ where $a_{k,l} := [\boldsymbol{a}_k]_l$. Then, since $\boldsymbol{a}_d = \boldsymbol{0}$ for $d \notin \{-\rho D, \ldots, \rho D\}$,

$$\begin{aligned}
[\boldsymbol{B}]_{n,m} &= \frac{1}{PN} \sum_{k=0}^{PN-1} \sum_{l=0}^{PN-1} e^{j\frac{2\pi}{PN}nk} a_{k-l,l} e^{-j\frac{2\pi}{PN}lm} \\
&= \frac{1}{PN} \sum_{d=-\rho D}^{\rho D} e^{j\frac{2\pi}{PN}nd} \sum_{l=0}^{PN-1} a_{d,l} e^{-j\frac{2\pi}{PN}l(m-n)} \\
&= \frac{1}{\sqrt{PN}} \sum_{d=-\rho D}^{\rho D} e^{j\frac{2\pi}{PN}nd} [\boldsymbol{F}\boldsymbol{a}_d]_{\langle m-n\rangle_{PN}},
\end{aligned}$$

where we used the substitution $d = k - l$. □

From (4.44), (4.58), (4.59), and Lemma 1

$$Q_{n,m} = \frac{1}{\sqrt{PN}} \sum_{d=-2D}^{2D} e^{j\frac{2\pi}{PN}nd} [\boldsymbol{\alpha}_d]_{\langle m-n\rangle_{PN}} \tag{4.64}$$

where $\boldsymbol{\alpha}_d := \boldsymbol{F}\mathrm{diag}_d(\boldsymbol{G})$. With $\alpha_{d,m} := [\boldsymbol{\alpha}_d]_m$, we find

$$\begin{aligned}
y_k &= \sum_{n=0}^{PN-1} |Q_{n,k}|^2 v_n \\
&= \frac{1}{PN} \sum_{n=0}^{PN-1} v_n \sum_{d,l=-2D}^{2D} e^{-j\frac{2\pi}{PN}n(l-d)} \alpha_{d,\langle k-n\rangle_{PN}} \alpha^*_{l,\langle k-n\rangle_{PN}} \\
&= \frac{1}{PN} \sum_{m=0}^{PN-1} v_{\langle k-m\rangle_{PN}} \sum_{d,l=-2D}^{2D} e^{-j\frac{2\pi}{PN}(l-d)(k-m)} \alpha_{d,m} \alpha^*_{l,m} \qquad (4.65)
\end{aligned}$$

where, for (4.65), $m = \langle k-n\rangle_{PN}$ so that $n = \langle k-m\rangle_{PN}$. Defining the matrix $\boldsymbol{D}_k := \mathcal{D}(\boldsymbol{F}\boldsymbol{i}_k)$ and the vector $\boldsymbol{\beta}(d,l)$ such that $[\boldsymbol{\beta}(d,l)]_m = \alpha_{d,m}\alpha^*_{l,m}$,

$$\boldsymbol{y} = \sum_{d,l=-2D}^{2D} \boldsymbol{D}_{l-d}\, \mathcal{C}(\boldsymbol{v}) \boldsymbol{D}^H_{l-d} \boldsymbol{\beta}(d,l) \tag{4.66}$$

Using the property $\boldsymbol{D}_k\boldsymbol{F} = \frac{1}{\sqrt{PN}}\boldsymbol{F}\boldsymbol{T}_k$,

$$\begin{aligned}
\boldsymbol{D}_{l-d}\, \mathcal{C}(\boldsymbol{v}) \boldsymbol{D}^H_{l-d} &= \sqrt{PN}\boldsymbol{D}_{l-d}\boldsymbol{F}\, \mathcal{D}(\boldsymbol{F}^H\boldsymbol{v})\boldsymbol{F}^H\boldsymbol{D}^H_{l-d} \\
&= \frac{1}{\sqrt{PN}}\boldsymbol{F}\boldsymbol{T}_{l-d}\, \mathcal{D}(\boldsymbol{F}^H\boldsymbol{v})\boldsymbol{T}^H_{l-d}\boldsymbol{F}^H \\
&= \frac{1}{\sqrt{PN}}\boldsymbol{F}\, \mathcal{D}\left(\boldsymbol{T}_{l-d}\boldsymbol{F}^H\boldsymbol{v}\right)\boldsymbol{F}^H. \qquad (4.67)
\end{aligned}$$

Substituting (4.67) into (4.66) yields (4.57).

4.C.2 Derivation of (4.60)

Here we derive an expression for $\boldsymbol{z}^{(n)}$ enabling fast computation. Without loss of generality, we omit superscripts in this appendix. From the definition of $\overline{\boldsymbol{g}}_k$ we notice $\boldsymbol{C}_k^H \boldsymbol{g}_k = \boldsymbol{C}^H \overline{\boldsymbol{g}}_k$, and thus, with (4.11), (4.45), and (4.62), we have $\boldsymbol{P} = \boldsymbol{C}^H \sum_k \overline{\boldsymbol{g}}_k \boldsymbol{i}_k^H \boldsymbol{F} = \boldsymbol{C}^H \overline{\boldsymbol{G}} \boldsymbol{F} = \mathcal{D}(\boldsymbol{b}^*) \boldsymbol{J}^H \boldsymbol{F}^H \overline{\boldsymbol{G}} \boldsymbol{F}$. Then we can write $z_k = \|\boldsymbol{p}_k\|^2 = \sum_{n=0}^{N_b-1} |b_n^* [\boldsymbol{F}^H \overline{\boldsymbol{G}} \boldsymbol{F}]_{\langle n \rangle_{PN}, k}|^2$. Since $\overline{\boldsymbol{G}}$ is banded with $2D$+1 non-zero diagonals, Lemma 1 implies $[\boldsymbol{F}^H \overline{\boldsymbol{G}} \boldsymbol{F}]_{\langle n \rangle_{PN}, k} = \frac{1}{\sqrt{PN}} \sum_{d=-D}^{D} e^{j \frac{2\pi}{PN} nd} \bar{\alpha}_{d, \langle k-n \rangle_{PN}}$ for $\bar{\alpha}_{d,m} := [\overline{\boldsymbol{\alpha}}_d]_m$. Thus

$$
\begin{aligned}
z_k &= \frac{1}{PN} \sum_{n=0}^{N_b-1} \left| b_n^* \sum_{d=-D}^{D} e^{j \frac{2\pi}{PN} nd} \bar{\alpha}_{d, \langle k-n \rangle_{PN}} \right|^2 \\
&= \frac{1}{PN} \sum_{n=0}^{N_b-1} |b_n|^2 \sum_{d,l=-D}^{D} e^{-j \frac{2\pi}{PN} n(l-d)} \bar{\alpha}_{d, \langle k-n \rangle_{PN}} \bar{\alpha}^*_{l, \langle k-n \rangle_{PN}} \\
&= \frac{1}{PN} \sum_{m=0}^{PN-1} \bar{b}_{\langle k-m \rangle_{PN}} \sum_{d,l=-D}^{D} e^{-j \frac{2\pi}{PN} (l-d)(k-m)} \bar{\alpha}_{d,m} \bar{\alpha}^*_{l,m}
\end{aligned}
$$

where $\bar{b}_m := \sum_{n=0}^{N_b-1} |b_{\langle n \rangle_{PN}}|^2 \delta_{\langle n \rangle_{PN} - m}$. Using $\boldsymbol{D}_k$ from Appendix 4.C.1, and defining $\overline{\boldsymbol{\beta}}(d,l)$ such that $[\overline{\boldsymbol{\beta}}(d,l)]_m = \bar{\alpha}_{d,m} \bar{\alpha}^*_{l,m}$, we find that

$$
\boldsymbol{z} = \sum_{d,l=-D}^{D} \boldsymbol{D}_{l-d} \mathcal{C}(\overline{\boldsymbol{b}}) \boldsymbol{D}_{l-d}^H \overline{\boldsymbol{\beta}}(d,l) \tag{4.68}
$$

where $[\overline{\boldsymbol{b}}]_m = \bar{b}_m$. Similar to Appendix 4.C.1, we substitute $\boldsymbol{D}_{l-d} \mathcal{C}(\overline{\boldsymbol{b}}) \boldsymbol{D}_{l-d}^H = \frac{1}{\sqrt{PN}} \boldsymbol{F} \mathcal{D}(\boldsymbol{T}_{l-d} \boldsymbol{F}^H \overline{\boldsymbol{b}}) \boldsymbol{F}^H$ into (4.68) to get (4.60).

Step	Cost
$\begin{bmatrix} \alpha_k & \boldsymbol{a}_k^H \\ \boldsymbol{a}_k & \boldsymbol{A}_k \end{bmatrix} = \boldsymbol{R}_k^{-1}$	0
$\boldsymbol{\Theta}_k^{-1} = \boldsymbol{A}_k - \alpha_k^{-1}\boldsymbol{a}_k\boldsymbol{a}_k^H$	$\widetilde{D}^2$
compute $\tilde{\boldsymbol{\theta}}_k$ and $\tilde{\theta}_k$	$3\widetilde{D}^2$
$\boldsymbol{b}_k = -\boldsymbol{\Theta}_k^{-1}\tilde{\boldsymbol{\theta}}_k$	$\widetilde{D}^2$
$\beta_k = \left(\tilde{\theta}_k - \tilde{\boldsymbol{\theta}}_k^H\boldsymbol{\Theta}_k^{-1}\tilde{\boldsymbol{\theta}}_k\right)^{-1}$	$\widetilde{D}^2$
$\boldsymbol{R}_{k+1}^{-1} = \begin{bmatrix} \boldsymbol{\Theta}_k^{-1} + \boldsymbol{b}_k\boldsymbol{b}_k^H\beta_k & \boldsymbol{b}_k\beta_k \\ \boldsymbol{b}_k^H\beta_k & \beta_k \end{bmatrix}$	$\widetilde{D}^2$

Table 4.4: Recursive Update of $(\boldsymbol{R}_k^{(n)})^{-1}$

CHAPTER 5

CONCLUSION

In the dissertation, we considered the problem of receiver design for single carrier transmissions over time-varying channels with long delay spread. The conventional solutions for this problem are various time-domain channel equalizers and estimators. These schemes often suffer from heavy computational burden due to deconvolution of the long channels. Inspired by the FDE idea behind OFDM and SCCP modulation schemes, we concentrated on designing receivers operating mainly in the frequency-domain, so as to achieve efficient implementation in favor of practical applications. In order to improve the performance of FDE, we adopted the idea of Turbo equalization (TE) and used soft information to iteratively equalize and estimate the channels. In the framework of FDE and TE, we designed and developed a group of frequency-domain joint channel estimation and equalization algorithms, and evaluated their performance in terms of BER and computational complexity through simulations and analysis.

5.1 Summary of Original Work

For single carrier transmissions over relatively moderately fast fading frequency-selective channels, we investigated iterative FDE (IFDE) with explicit frequency-domain channel estimation (FDCE). First, an improved IFDE algorithm was presented based on soft iterative interference cancellation. Second, soft-decision-directed

channel estimation algorithms were derived and analyzed both in time and frequency domain. As it turns out, the frequency-domain approach is more computationally efficient than the time-domain approach. Therefore a new adaptive FDCE (AFDCE) algorithm based on per-tone Kalman filtering was proposed to track and predict the frequency-domain channel coefficients. The AFDCE algorithm employed across-tone noise reduction, which could perfectly remove the redundant noise when channels could be modeled by an AR process, and employed Chu sequences as training sequences. In addition, the AFDCE algorithm exploited temporal correlation between successive blocks to adaptively update the AR model coefficients, bypassing the need for prior knowledge of channel statistics. Finally, a block overlapping idea was proposed for the joint operation of IFDE and AFDCE. Simulation results show that, compared to other existing IFDE and adaptive channel estimation schemes, the proposed schemes offer lower MSE in channel prediction, lower BER after decoding, and robustness to non-stationary channels.

We further extended the IFDE/AFDCE scheme to fit in the application of digital television (DTV) signal reception, where trellis coded vestigial sideband modulation is employed, as specified by the ATSC North American terrestrial DTV standard. The proposed FDTE/AFDCE scheme estimates and equalizes channels only on active subcarriers in the frequency-domain, and therefore achieves low-cost and high-performance reception of highly impaired DTV signals. Through numerical simulation, we demonstrated that our FDTE/AFDCE scheme outperformed the traditional joint DFE/decoding plus FDLMS-CE approach at a fraction of the implementation cost.

For single carrier transmissions over very fast fading large-delay-spread channels, the traditional FDE methods fail when the influence of virtual ICI is not negligible. We first applied Doppler channel shortening to concentrate the energy of virtual ICI coefficients into a banded structure in the Doppler and frequency domain, and then derived a pilot-aided channel estimator (PACE) to estimate those significant virtual

ICI coefficients based on MMSE criteria. Finally a soft iterative interference cancellation algorithm was proposed to efficiently detect transmitted symbols by leveraging the banded structure of ICI, while block decision feedback and block overlapping were employed to further combat ISI and the virtual ICI. Numerical results showed that the proposed scheme had advantages over the well-known FIR-MMSE-DFE/RLS-CE scheme in both performance and complexity.

5.2 Possible Future Work

In our joint CE and equalization design, we only considered the influence of symbol detection errors on CE in soft-decision-directed CE case, while we have not considered the influence of CE errors on channel equalization yet. Possible performance gain can be achieved by characterizing CE errors as stochastic processes and incorporating those information into channel equalization.

For the receiver design of DTV receiver, we investigated the FDTE/AFDCE algorithm based on the assumption that there is a pilot block available to initialize the process. In practice, the pilot symbols transmitted may be not long enough to initialize the FDTE/AFDCE, therefore a frequency-domain blind CE or blind channel equalization algorithm is expected to start up the process. Blind DFE algorithm is a promising candidate, which is widely adopted in current DTV receivers for reliable reception over slow fading channels [105–108]. How to efficiently combine the startup process with our FDTE/AFDCE algorithm would be a interesting problem to work on in order to complete the receiver design for DTV signal reception.

In very fast fading channels, we only considered PACE, where as decision-directed CE could also be explored to further reduce CE errors and enable reception of transmitted symbols over a wider Doppler and delay spread region.

In this dissertation, we only considered single transmitter and receiver antenna case, while it would be interesting to see how similar performance gain and implementation gain can be achieved in multiple-input multiple-output systems jointly with space-time coding.

BIBLIOGRAPHY

[1] D. Falconer, S. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, pp. 58–66, Apr. 2002.

[2] "ATSC digital television standard, rev. d-w/amendment 1." http://www.atsc.org/standards/a_53d.pdf.

[3] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Journal Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.

[4] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-time Communications.* Cambridge University Press, 2003.

[5] G. Stuber, J. Barry, S. Mclaughlin, Y. Li, M. ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *IEEE Proceedings*, vol. 92, pp. 271–294, Feb. 2004.

[6] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Trans. on Wireless Communications*, vol. 12, pp. 56–65, Apr. 2005.

[7] "Doppler effect." http://en.wikipedia.org/wiki/Doppler_effect.

[8] J. G. Proakis, *Digital Communications, 4th ed.* New York: McGraw-Hill, 2001.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication.* Cambridge University Press, Sep. 2004.

[10] M. Ghosh, "Blind decision feedback equalization for terrestrial television receivers,"

[11] A. Shah, S. Biracree, R. Casas, T. Endres, S. Hulyalkar, T. Schaffer, and C. Strolle, "Global convergence of a single-axis constant modulus algorithm," in *Proc. of the Workshop on Statistical Signal and Array Processing*, pp. 645–649, Aug. 2000.

[12] S. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.

[13] "Orthogonal frequency-division multiplexing." http://en.wikipedia.org/wiki/OFDM.

[14] Y. S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective rayleigh fading chanels," *IEEE Trans. Commun.*, vol. 49, pp. 1375–1387, Aug. 2001.

[15] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An Equalization technique for orthogonal frequency-division multiplexing systems in time-varying multipath channels," *IEEE Trans. Commun.*, vol. 47, pp. 27–32, Jan. 1999.

[16] W. S. Hou and B. S. Chen, "ICI cancellation for OFDM communication systems in time-varying multipath fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2100–2110, Sep. 2005.

[17] X. Cai and G. B. Giannakis, "Bounding performance and suppressing intercarrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, pp. 2047–2056, Dec. 2003.

[18] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 1002–1011, Apr. 2004.

[19] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE commun. Letter*, vol. 9, pp. 619–621, July 2005.

[20] A.Gorokhov and J.-P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, pp. 572–583, Apr. 2004.

[21] I. Barhumi, G. Leus, and M. Moonen, "Time-domain and frequency-domain per-tone equalization for OFDM in doubly selective channels," *Signal processing*, vol. 84, pp. 2055–2066, Nov. 2004.

[22] C. B. A. Glavieux and P. Thitimajshima, "Near shannon limit errorcorrecting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. on Communications*, May 1993.

[23] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo equalization," *European Trans. on Telecommunications*, vol. 6, pp. 507–511, Sept.-Oct. 1995.

[24] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. commun.*, vol. 47, pp. 1046–1061, July 1999.

[25] M. Tüchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using a priori information," *IEEE Trans. Signal Processing*, vol. 50, pp. 673–683, Mar. 2002.

[26] M. Tüchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Signal Processing*, vol. 50, pp. 673–683, Mar. 2002.

[27] R. Koetter, A. Singer, and M. Tüchler, "Turbo equalization," *IEEE Signal Processing Mag.*, vol. 21, pp. 67–80, Jan. 2004.

[28] M. Tüchler and J. Hagenauer, "Turbo equalization using frequency domain equalizers," in *Proc. Allerton Conf., Monticello, IL.*, Oct. 2000.

[29] M. Tüchler and J. Hagenauer, "Linear time and frequency domain turbo equalization," in *Proc. Vehicular Technology Conf., Rhodes, Greece.*, May 2001.

[30] B. Ng, C. Lam, and D. Falconer, "Turbo frequency domain equalization for single-carrier broadband wireless systems," *IEEE Trans. on Wireless Communications*, vol. 6, pp. 759–767, Feb. 2007.

[31] R. Visoz, A. O. Berthet, and S. Chtourou, "Frequency-domain block Turbo equalization for single-carrier transmission over MIMO broadband wireless channel," *IEEE Trans. on Communications*, vol. 54, pp. 2144–2149, Dec. 2006.

[32] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Information Theory*, vol. 20, pp. 284–287, Mar. 1974.

[33] F. Kschischang, B.J.Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.

[34] R. McEliece, D. MacKay, and J. Cheng, "Turbo decoding as an instance of pearl's "belief propagation" algorithm," *IEEE Jour. Selected Areas in Communs.*, vol. 47, pp. 498–519, Feb. 2001.

[35] R.Otnes and M. Tüchler, "Iterative channel estimation for Turbo equalization of time-varying frequency-seletive channels," *IEEE Trans. on Wireless Comm.*, vol. 3, pp. 1918–1923, Nov. 2004.

[36] S. Song, A. C. Singer, and K.-M. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Processing*, vol. 52, pp. 2885–2894, Oct. 2004.

[37] N. Nefedov, M. Pukkila, R. Visoz, and A. O. Berthet, "Iterative data detection and channel estimation for advanced TDMA systems," *IEEE Trans. Commun.*, vol. 51, pp. 141–144, Feb. 2003.

[38] P. Banelli, R. C. Cannizzaro, and L. Ruguni, "Data-aided Kalman-tracking for channel estimation in Doppler-affected OFDM system," in *Proc. International Acoustic, speech and signanl processing Conf., Hawai, USA.*, 2007.

[39] R. C. Cannizzaro, P. Banelli, and G. Leus, "Adaptive channel estimation for OFDM systems with Doppler spread," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, 2006.

[40] I. Barhumi, G. Leus, and M. Moonen, "MMSE estimation of basis expansion model for rapidly time-varying channels," in *EUSIPCO*, Sep. 2005.

[41] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. signal process*, vol. 53, pp. 3597–3607, Sep. 2005.

[42] K. Ann, D. Teo, and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Globecom*, Dec. 2005.

[43] S. Tomasin, A. Gorokhov, H. Yang, and J. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. on Wireless Comm.*, vol. 4, pp. 238–245, Jan. 2005.

[44] P. Schniter, "On doubly dispersive channel estimation for pilot-aided pulse-shaped multi-carrier modulation," in *Proc. Conference on Information Sciences and Systems*, Mar. 2006.

[45] A. P. Kannu and P. Schniter, "Reduced-complexity decision-directed pilot-aided tracking of doubly-selective channels," in *Proc. Conference on Information Sciences and Systems*, Mar. 2004.

[46] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of time-varying channels in wireless MIMO-OFDM systems," in *Proc. Signals, Systems and Computers Conf., Asilomar, CA. USA.*, Nov. 2003.

[47] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Trans. Wireless Communications*, vol. 4, no. 2, pp. 593–602, 2005.

[48] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid disppersive fading channels," *IEEE Trans. Communications*, vol. 46, no. 7, pp. 902–914, 1998.

[49] M. Tüchler, R. Otnes, and A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalization," in *Proc. International Communications Conf., New York, USA.*, May 2002.

[50] S. Jiun, J. Coon, R. J. Piechocki, A. Dowler, A. Nix, M. Beach, S. Armour, and J. McGeehan, "A channel estimation algorithm for MIMO-SCFDE," *IEEE Communications Letters*, vol. 8, no. 9, pp. 555–557, 2004.

[51] R. Dinis, R. Kalbasi, D. Falconer, and A. Banihashemi, "Channel estimation for MIMO systems employing single-carrier modulations with iterative frequency-domain equalization," in *Proc. Vehicular Technology Conference, Los Angeles, CA. USA.*, 2004.

[52] W. Liu, L. Yang, and L. Hanzo, "Wideband channel estimation and prediction in single-carrier wireless systems," in *Proc. Vehicular Technology Conf., Stockholm, Sweden.*, May 2005.

[53] F. A. Dietrich and W. Utschick, "Pilot-assisted channel estimation based on second-order statistics," *IEEE Trans. Signal Processing*, vol. 53, no. 3, pp. 1178–1193, 2005.

[54] O. Rousseaux and G. Leus, "An iterative method for improved training-based estimation of doubly selective channels," in *ICASSP*, May 2004.

[55] M. Niedżwiecki and T.Kłaput, "Fast recursive basis function estimators for identification of time-varying processes," *IEEE Trans. on Signal Proc.*, vol. 50, pp. 1925–1934, Aug. 2002.

[56] M. Morelli, L. Sanguinetti, and U. Mengali, "Channel estimation for adaptive frequency-domain equalization," *IEEE Trans. Wireless Communications*, vol. 4, no. 5, pp. 2508–2518, 2005.

[57] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Processing*, vol. 45, pp. 658–672, Mar. 1997.

[58] C. Johnson, P. Schniter, T. Endres, J. Behm, D. Brown, and R. Casas, "Blind equalization using the constant modulus criterion: A review," *Proceedings of the IEEE: Special Issue on Blind System Identification and Estimation*, vol. 86, pp. 1927–1950, Oct. 1998.

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

[60] H. Liu and P. Schniter, "Iterative frequency-domain channel estimation and equalization for single-carrier transmissions without cyclic-prefix." submitted to IEEE Trans. on Wireless Communicationn, 2007.

[61] H. Liu and P. Schniter, "Iterative frequency-domain channel estimation and equalization for single-carrier transmission without cyclic prefix," in *Proc. Conference on Information Sciences and Systems*, Mar. 2007.

[62] H. Liu, P. Schniter, H. Fu, and R. A. Casas, "A frequency domain turbo equalization for trellis coded vestigial sideband modulation," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, 2006.

[63] H. Liu, R. A. Casas, and H. Fu, "A frequency-domain turbo equalizer for dtv signals," Jan. 2006. Patent filed.

[64] P. Schniter and H. Liu, "Iterative frequency-domain equalization of single-carrier transmission over douly dispersive channel." submitted to IEEE Trans. on Signal processing, 2005.

[65] P. Schniter and H. Liu, "Iterative equalization for single carrier cyclic-prefix in doubly-dispersive channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Oct. 2003.

[66] P. Schniter and H. Liu, "Iterative frequency-domain equalization for single-carrier systems in doubly-dispersive channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2004.

[67] D. Kim and G. Stüber., "Residual ISI cancellation for OFDM with application to HDTV broadcasting," *IEEE Journal Select. Areas in Comm.*, vol. 16, pp. 1590–1599, Oct. 1998.

[68] C.-J. Park and G.-H. Im, "Efficient cyclic prefix reconstruction for coded OFDM systems," *IEEE Commun. Letter*, vol. 8, pp. 274–276, May 2004.

[69] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Information Theory*, vol. 42, pp. 429–445, Mar. 1996.

[70] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, NJ: Prentice-Hall, 4th ed., 2001.

[71] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Information Theory*, vol. 18, pp. 531–532, July 1972.

[72] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[73] T. Laud, M. Aitken, W. Bretl, and K. Kwak, "Performance of 5th generation 8-VSB receivers," *IEEE Trans. on Consumer Electronics*, vol. 50, pp. 1076–1080, Nov. 2004.

[74] Behara, "An integrated VSB/QAM/NTSC/BTSC receiver: recent advances in television design," in *Proc. IEEE Broadcasting Symposium*, Oct. 2003.

[75] M. Ghosh, "Blind decison feedback equalization for terrestrial television receivers," *Proceedings of the IEEE*, vol. 86, pp. 2070–2081, Oct. 1998.

[76] J. Henderson, W. Bretl, M. Deiss, A. Goldberg, B. Markwalter, and M. Muterspaugn, "ATSC DTV Receiver Implementation," *Proceedings of the IEEE*, vol. 94, pp. 119–147, Jan. 2006.

[77] Y. Wu, X. Wang, R. Citta, B. Ledoux, S. Lafleche, and B. Caron, "An ATSC DTV receiver with improved robustness to multipath and distributed transmission environments," *IEEE Trans. on Broadcasting*, vol. 50, pp. 32–41, Mar. 2004.

[78] G. Sgrignoli, "Meansuring peak/average power ratio of the zenith/AT&T DSC-HDTV signal with vector signal analyzer," *IEEE Trans. on Broadcast*, vol. 39, pp. 255–264, Jun. 1993.

[79] S. L. Ariyavisitakul and Y. Li, "Joint coding and decision feedback equalization for broadband wireless channels," *IEEE Jour. Selected Areas in Communs.*, vol. 16, pp. 1670–1678, Dec. 1998.

[80] N. Al-Dhahir and J. M. Cioffi, "Fast computation of channel-estimate based equalizers in packet data transmission," *IEEE Trans. on Signal Processing*, vol. 43, pp. 2462–2473, Nov. 1995.

[81] "ATSC recommended practice: Receiver performance guidelines." http://www.atsc.org/standards/practices/a_74_rfs.pdf.

[82] W. C. Jakes, *Microwave Mobile Communication*. Wiley, 1974.

[83] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Information Theory*, vol. 28, pp. 55–67, Jan. 1982.

[84] S. Öz, "Topics on channel estimation and equalization for sparse channels with applications to digital TV systems." Ph.D dissertation, 2003.

[85] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," vol. 50, pp. 2451–2464, Oct. 2002.

[86] X. Cai and G. Giannakis, "Bounding performance and suppressing inter-carrier interference in wireless mobile ofdm," *TCOM*, vol. 51, pp. 2047–2056, Dec 2003.

[87] A. Gorokhov and J.-P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: Equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, pp. 572–583, Apr. 2004.

[88] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Processing*, vol. 52, pp. 1002–1011, Apr. 2004.

[89] G. Leus, I. Barhumi, and M. Moonen, "Equalization techniques for fading channels," in *Handbook on Signal Processing for Communications* (M. Ibnkahla, ed.), CRC Press, 2004.

[90] P. Schniter and H. Liu, "Iterative equalization for single carrier cyclic-prefix in doubly-dispersive channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Oct. 2003.

[91] U. S. A. T. S. Committee, "Guide to the use of the ATSC digital television standard." ATSC Doc. A/54, Oct. 1995.

[92] D. B. Kilfoyle and A. B. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE J. Oceanic Eng.*, vol. 25, Jan. 2000.

[93] N. Al-Dhahir and J. M. Cioffi, "MMSE decision feedback equalizers: Finite-length results," *IEEE Trans. Inform. Theory*, vol. 41, pp. 961–976, July 1995.

[94] J. Proakis, *Digital Communications*. Boston: McGraw-Hill, 4 ed., 2001.

[95] M. Stojanovic, June 2007. Private communication.

[96] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," vol. 52, pp. 1541–1562, Nov. 1973.

[97] S. Das and P. Schniter, "A new pulse shaped frequency division multiplexing technique for doubly dispersive channels," in *Proc. Asilomar Conf. on Signals, Systems, and Computersm*, Nov. 2004.

[98] G. Golub and C. V. Loan, *Matrix Computations*. Baltimore,MD: Jonh Hopkins University Press, 3 ed., 1996.

[99] H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 2 ed., 1994.

[100] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using a priori information," vol. 50, pp. 673–683, Mar 2002.

[101] P. Schniter and S. Das, "Multi-carrier modulation over doubly dispersive channels: Pulse shaping and equalization," *IEEE Trans. Signal Processing*, 2006. Accepted.

[102] P. K. Shukla and L. F. Turner, "Channel-estimation-based adaptive DFE for fading multipath radio channels," *IEEE Proceedings*, vol. 38, pp. 525–543, Dec. 1991.

[103] N. Ahmed and K. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer, 1975.

[104] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, New York, 1985.

[105] J. Labat, O. Macchi, and C. Laot, "Adaptive decision feedback equalization: Can you skip the training period?," *IEEE Trans. on Communications*, vol. 46, pp. 921–930, July 1998.

[106] J. Labat and C. Laot, "Blind adaptive multiple-input decision feedback equalizer with a self-optimized configurationa," *IEEE Trans. on Communications*, vol. 49, pp. 646–654, Apr. 2001.

[107] S. N. Hulyalkar and M. Ghosh, "Blind equalizer method and apparatus for HDTV transmission using an NTSC rejection filter for mitigating co-channel interference." US patent issued, 1998.

[108] A. K. Rodney and D. A. Thushara, "VSV-MOE pre-equalizer for 8-VSB DTV." US patent issued, 2006.