2002

Optimal channel equalization for filterbank transceivers in presence of white noise

Ehab Farouk Badran
Louisiana State University and Agricultural and Mechanical College

OPTIMAL CHANNEL EQUALIZATION FOR FILTERBANK TRANSCEIVERS IN PRESENCE OF WHITE NOISE

A Dissertation

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

in

The Department of Electrical and Computer Engineering

by
Ehab Farouk Badran
B.S.E.E., Assiut University, Egypt, 1995
M.S.E.E., Assiut University, Egypt, 1998
M.S.E.E., Louisiana State University, 2001
May 2002

To whom I love,
specially my professor Dr. Guoxiang Gu,
my mother, my father,
and my brothers Sherif and Tamer

Acknowledgments

First of all, I thank Almighty God "Allah", the most gracious and the most merciful, for everything he has given to me.

Deep thanks and appreciation are not enough to express my gratitude towards my major professor Guoxiang Gu for his kind supervision, discussions, suggestions, and guidance throughout the stages of this research, that helped to propel my research work and to prepare this dissertation. My gratitude and thanks also go to Professor Ahmed El-Amawy for helping me to get a teaching assistantship in my first semester at LSU, and for serving on my advisory committee.

Appreciation is also extended to my other advisory committee members: Prof. W. George Cochran, Prof. Peter Kelle, and Prof. Kemin Zhou for their willingness to serve on this committee and for their review of this dissertation. The financial support from the Department of Electrical and Computer Engineering at LSU, and the Air Force Office for Scientific Research is also acknowledged.

No words can express my sincere gratitude to my family for their support, endless love they endowed, encouragement and understanding that motivated me to finish this work, even they are at a far distance.

Table of Contents

List of Figures

List of Symbols

A^*:	Complex conjugate transpose of A
A^T:	Transpose of A
A^{-1}:	Inverse of A
I_s:	Identity matrix of size $s \times s$
$\rho(A)$:	Spectral radius of A
ℓ_+:	The space of causal discrete-time signals
$\mathbf{R}^n$:	A real space of n dimensions
$\Lambda^*(\bar{z}^{-1})$:	Para-Hermitian of a transfer function matrix $\Lambda(z)$, with $\bar{\cdot}$ conjugation

List of Acronyms

ARE:	Algebraic Riccati equation
CDMA:	Code Division Multiple Access
DFT:	Discrete-time Fourier Transform
DMT:	Discrete Multi-Tone
DRE:	Difference Riccati equation
DS-CDMA:	Discrete Sequence Code Division Multiple Access
DWMT:	Discrete Wavelet Multi-Tone
FDMA:	Frequency Division Multiple Access
FIR:	Finite Impulse Response
IIR:	Infinite Impulse Response
ISI:	Inter-Symbol-Interference
LS:	Least Square
MIMO:	Multi-Input, Multi-Output
MSE:	Mean-Squared-Error
OFDM:	Orthogonal Frequency Division Multiplexing
PR:	Perfect Reconstruction
SISO:	Single-Input, Single-Output
SVD:	Singular Value Decomposition
TDMA:	Time Division Multiple Access
WSS:	Wide-Sense Stationary
ZF:	Zero Forcing

Abstract

Filterbank transceivers are widely employed in data communication networks to cope with inter-symbol-interference (ISI) through the use of redundancies. This dissertation studies the design of the optimal channel equalizer for both time-invariant and time-varying channels, and wide-sense stationary (WSS) and possible non-stationary white noise processes. Channel equalization is investigated via the filterbank transceivers approach. All perfect reconstruction (PR) or zero-forcing (ZF) receiver filterbanks are parameterized in an affine form, which eliminate completely the ISI. The optimal channel equalizer is designed through minimization of the mean-squared-error (MSE) between the detected signals and the transmitted signals. Our main results show that the optimal channel equalizer has the form of state estimators, and is a modified Kalman filter. The results in this dissertation are applicable to discrete wavelet multitone (DWMT) systems, multirate transmultiplexers, orthogonal frequency division multiplexing (OFDM), and direct-sequence/spread-spectrum (DS/SS) based code division multiple access (CDMA) networks. Design algorithms for the optimal channel equalizers are developed for different channel models, and white noise processes, and simulation examples are worked out to illustrate the proposed design algorithms.

Chapter 1

Introduction

1.1 Literature Review

Equalization is one of the techniques that can be used to improve the received signal quality in telecommunications, especially in digital communications. Equalization compensates for the *inter-symbol-interference* (ISI), which is a common problem in telecommunication systems and wireless communication systems, such as television broadcasting, digital data communications, and cellular mobile communication systems. In telecommunication systems, ISI occurs when the modulation bandwidth exceeds the coherence bandwidth of the radio channel where modulation pulses are spread in time. For wireless communications, ISI is caused by multipath in bandlimited time dispersive channels, and it distorts the transmitted signal, causing bit errors at the receiver. ISI has been recognized as the major obstacle to high-speed data transmission and multipath fading over radio channels.

An equalizer within a receiver compensates for the average range of the expected channel amplitude and delay characteristics. In a broad sense, the term equalization can be used to describe any signal processing operation that minimizes the ISI.

Channel equalization has received great attention for the past two decades, and it has received renewed interest in the research community due to the need for wireless communications, where the channels are inherently distorted due to multipath and dispersion, and due to the increased demand for high speed data transmission.

There have been considerable efforts devoted to solving the ISI problem (see [75] and references therein). One of the techniques is post-equalization, such as the least-mean-squared (LMS) equalizer, and the decision feedback equalization (DFE), which were studied in [6, 50, 51].

The idea of DFE is to cancel the ISI that may be caused by the current detected symbol on future symbols by estimating and subtracting its ISI out prior to the detection of subsequent symbols. DFE consists of two filters, feedforward filter, and feedback filter, and decision device. The feedback filter coefficients are adjusted to cancel the ISI on the current symbol, resulted from the previously detected symbols. This adjustment is controlled by decisions of the output of the detector. The coefficients adjustment may be achieved by a linear equalizer with LMS algorithms. Also, the feedforward filter coefficients may be adjusted to help canceling ISI on the current symbol. For conventional DFE, the coefficients adjustment of the feedforward filter, and the feedback filter is carried out jointly, while separately for predictive DFE. The convergence of these iterative algorithms is dependent on the channel characteristics. For spectral null or frequency selective fading channels, these algorithms converge very slowly and, therefore, become computationally expensive.

There exist many research results [56, 60, 65, 68] for post-equalization techniques, with emphasis on blind equalization, where the channel is assumed to be unknown. The existing work on blind equalization techniques, can be approximately divided into three groups:

- high-order statistics techniques;
- second-order statistics techniques with oversampling;
- antenna array multi-receiver techniques;

Multicarrier modulation is also a technique to cope with the ISI, which aims to increase the transmission system length to eventually cancel the ISI [14, 52, 75, 77]. Another group of techniques is the use of precoding. Examples include Tomlinson-Harashima (TH) precoding [29, 62], trellis precoding [20, 22], matched spectral null precoding in partial response channels [33], and other precoding schemes [9, 34, 71, 73, 74]. The idea of TH precoding is to equalize the signal before the transmission. The transmitter using the TH precoding needs to know the channel characteristics, and thus the precoding is not reliable when the channel has spectral null or frequency selective fading characteristics. In fact, the pre-equalizer oscillates in a dramatic way when the channel is close to zero. The trellis precoding scheme proposed in [20] whitens the noise at the equalizer output. This scheme combines precoding and trellis shaping. Similar to the TH precoding, the transmitter using the trellis precoding also needs to know the channel characteristics. The trellis shaping depends on the channel and may not be suitable for spectral-null channels either. An approach used for magnetic recording systems, is the matched spectral null

precoding scheme [33]. The matched spectral null precoding, when dealing with partial response channels, chooses certain error control codes to match the spectral nulls of partial response channels in order to lose less signal information through the channel.

A mutual influence and interaction have been recently recognized between the fields of multirate signal processing, and digital communications. Some of the popular communication applications can be described in terms of filterbanks configurations (transmultiplexers) of subband transforms. Such applications include the following.

- Frequency division multiple access (FDMA).

  In FDMA individual channels are assigned to individual users. Each user is allocated a unique frequency band or channel. These channels are assigned on demand to users who request service.

- Time division multiple access (TDMA) systems.

  In TDMA the radio spectrum is divided into time slots, and in each slot only one user is allowed to either transmit or receive. Each user occupies a cyclically repeating time slot, so a channel may be thought of as particular time slot that reoccurs every frame, where a group of time slots comprises a frame.

- Code division multiple access (CDMA) systems.

  In CDMA, the narrowband message signal is multiplied by a very large bandwidth signal called the spreading signal. The spreading signal is a pseudo-noise code sequence based on a chip rate which is orders of magnitudes greater than the data rate of the message. All users in a CDMA system use the same carrier frequency,

and may transmit simultaneously. Each user has its own pseudo-random codeword, which is approximately orthogonal to all other codewords. The receiver performs a time correlation operation to detect only the specific desired codeword.

In particular, FDMA which is also called orthogonal frequency division multiplexing (OFDM) or discrete multitone (DMT) modulation-based systems have been more widely used than the others.

The orthogonality of multicarriers was recognized early on as the proper way to pack more subchannels into the same channel spectrum [49, 53, 72], (also see [77] and references therein). Originally, a bank of analog Nyquist filters was used for the multicarrier modulation techniques, which provide a set of continuous-time orthogonal functions. However, it was impossible for a set of strictly orthogonal analog filters to be realized. Therefore, the formulation was reworked into the discrete-time model. This discrete-time model makes explicit use of the orthogonal filterbanks or transmultiplexers.

Transmultiplexers (filterbanks) were studied in the early 1970's in [7] for telephony applications. Their seminal work was one of the first dealing with multirate signal processing, which has matured lately in the signal processing field. Among the recent research on channel equalization, a multirate filterbank was proposed as a precoder before transmission in [75], which introduces the use of redundancy. An FIR (which is a non-maximally decimated multirate filterbank) decoder was used at the receiver to eliminate the ISI. Another framework for multirate filterbanks was proposed in [25, 57]. This framework is of particular interest, which encompasses existing modulations and equalization schemes, as well as general channel identifiability conditions, leading to possible optimal design

of transmitter/receiver due to the redundancy for an FIR channel, and block transmission introduced by filterbanks. Although the results in [57] cover only a limited class of channels having the drawback of high complexity, the filterbank approach shows the promise for its future role in optimization for channel equalization, and minimization of the mean-squared-error (MSE).

Some of the applications in digital communications have recently generated significant research activities (see [77] and references therein). In particular two of them in multicarrier modulation (OFDM or DMT) techniques are discussed briefly as follows. First, the coded orthogonal frequency division multiplexing (COFDM) has been forwarded for terrestrial digital audio broadcasting (T-DAB) in Europe. Consequently, an European T-DAB standard has been defined, and actual digital broadcasting systems are being built. Second, a DFT-based DMT modulation scheme has become the standard for asymmetric digital subscriber line (ADSL) communications, which provides an efficient solution to the last mile problem (eg., providing high-speed connectivity to subscribers over the unshielded twisted pair (UTP) copper cables) [17, 19, 31, 32]. In addition to these two successful applications of multicarrier modulation, several emerging application areas, including spread-spectrum orthogonal transmultiplexers for CDMA [3] and low probability of intercept (LPI) communications [21, 54] have received great attention as well.

Filterbank transceivers (transmultiplexers) have been studied for the past several years in the signal processing community, due to their capability for block transmission, and use of redundancies for precoding, enabling complete elimination of ISI, which is one of the major obstacles for data communications. Indeed for high-speed transmission over chan-

nels such as digital subscriber loops, DWMT is a widely used technique [2, 15, 31, 32, 36], of which realization relies on the design of redundant filterbank transceivers that effectively divide the channel into subbands. Band separation is important in terms of allocation of transmission power and bits, because the noise can be highly frequency selective, or nonflat. For FDMA (frequency division multiple access) in wireless communications, multirate transmultiplexers are essential for dispersive channels with additive noise, due to again their frequency separability and selectivity [11, 40, 55]. If the up-samplers and down-samplers have higher ratio than the number of frequency bands, we have a model of redundant filterbank transceivers. While DS/SS-based CDMA network is less obvious in its connection with redundant filterbank transceivers, [57, 66] have succeeded in this regard. Thus in our opinion, redundant filterbank transceivers provide a common mathematical model in data communications.

Because of the introduced redundancies, filterbank transceivers are capable of achieving channel equalization, which eliminate completely ISI under some mild conditions. However optimal detection of the transmitted data symbols remains a major issue for dispersive channels with additive noise, giving the rise of the optimal design problem for the transceiver receivers, which is the main goal of this dissertation.

1.2 Dissertation Contributions

This dissertation is a continuation of the existing research in channel equalization, emphasizing on the design of filterbank transceivers (transmultiplexers) for the purpose of eliminating the ISI completely, and minimizing the MSE of the received signals. Many

researchers have contributed to this important problem area, focusing on the design of FIR transceiver receivers to achieve MMSE [40, 66, 70]. It should be clear that MMSE receivers in the existing literature are actually suboptimal by the fact that optimization is carried out over all FIR filters of a fixed order. Their optimality over all possible linear, or nonlinear filters of possibly infinite order can not be claimed thus far. Although Kalman filters have been employed in [11, 12, 39] for the design of optimal transceiver receivers, the perfect reconstruction (PR), or the zero-forcing condition is neglected for the noise-free case, raising the issue of ISI. Moreover a true MMSE detection without the zero-forcing condition remains unsolved in general.

The contributions of this dissertation are as follows.

- Development of the new design algorithm for transceiver receivers achieving optimal channel equalization for time-invariant channels and wide-sense stationary (WSS) white noise processes.

  Contrasting to the time-domain method used as in [25, 57], we will employ transfer function and state-space methods to tackle the optimal design of channel equalizers, which enable us to generalize the results in [57] to more general channels, to remove the constraints on the filterbanks, and to reduce the complexity. Our contributions include the necessary and sufficient condition for PR on the transmitter and receiver filterbanks in the noise-free environment. Although joint transceiver optimization is not investigated, a similar condition to the one in [57] can be used for the design of the transmitter filters to ensure good SNR (signal-to-noise ratio). More importantly, if the PR condition holds, all causal and stable receiver filterbanks which achieve PR

are parameterized. Under the condition that the noise is white and WSS, we show further that the optimal receiver filterbank or optimal channel equalizer has the form of state estimators, and is a modified Kalman filter, which minimizes the MSE among all possible linear time-invariant filters. The rich theory of Kalman filtering enables the optimal design for transmitter/receiver filterbanks by employing efficient and reliable numerical algorithms.

- Development of the new design algorithm for transceiver receivers achieving optimal channel equalization for time-varying channels and non-stationary white noise processes.

  We consider the case where the estimated channel is available, and is time-varying, due to the mobile and ubiquitous nature of cellular phones, and applications of the filterbank transceivers to wireless data communications. We will study design of the optimal receivers that achieve not only PR or zero-forcing, but also MMSE for possible non-stationary white noise processes, among all possible linear and time-varying filters of arbitrary orders. This is in contrast to the existing results in the present literature. We will show that for redundant filterbank transceivers, the zero-forcing receivers are equivalent to causal and stable left inverses of some "tall" matrix of linear dynamic systems, consisting of the channel and transmitters. Such left inverses are not unique. Any zero-forcing receiver filterbank accomplishes the goal of channel equalization, and eliminates completely ISI. We will first parameterize all zero-forcing channel equalizers in an affine form, and then seek one of them to minimize the MSE caused by possible non-stationary white noises corrupted at

the receivers, thereby converting the constrained optimization into unconstrained optimization for receivers design. It will be shown that the design of optimal channel equalizers is equivalent to the design of optimal state estimators for some augmented system, subject to the same noise process. Hence the celebrated Kalman filtering can be used successfully to design the optimal detectors among all channel equalizers, which are truly optimal over all linear and time-varying receiver filters of arbitrary order.

This dissertation is comprehensive in using the system theory and Kalman filtering theory to approach channel equalization via filterbank transceivers. The results are submitted to the journal of *IEEE transaction on signal processing* of which two are under review [5, 26]. It is hoped that the results in this dissertation will improve the technology for channel equalization in filterbank transceivers.

1.3 Background Materials

This section introduces some standard techniques used in this dissertation to treat periodic/multirate systems, and fundamental concepts in linear systems.

1.3.1 Sampling Rate Conversion

When dealing with signals having different sampling rate, sampling rate conversion is needed. To achieve sampling rate conversion for discrete-time data streams in multirate signal processing, two basic devices are used. One is called down-sampler (decimator, or compressor), and the other is called up-sampler (interpolator, or expander).

• Down-sampler and Up-sampler

A down-sampler by M and an up-sampler by P are shown in Figure 1.1 where M and P are integers. Passing an input data stream $s(n)$ through a down-sampler by M produces an output signal $X_d(n)$ where

$$X_d(n) = s(Mn).$$

In the $\mathcal{Z}$-transform domain,

$$X_d(z) = \frac{1}{M}\sum_{k=0}^{M-1} S(z^{1/M}W_M^k), \quad W_M = e^{-j2\pi/M}, \quad j = \sqrt{-1}. \tag{1.1}$$

The down-sampler retains only those samples of $s(n)$ with $n = kM$ and $k = 0, 1, 2, \cdots$. The resulting signal $x_d(n)$ has a sampling rate $1/M$ of that of $s(n)$. A properly designed lowpass filter is usually used before the down-sampling process to eliminate the aliasing. Otherwise it may not be possible to recover $s(n)$ from $x_d(n)$ because of the loss of information due to the aliasing.

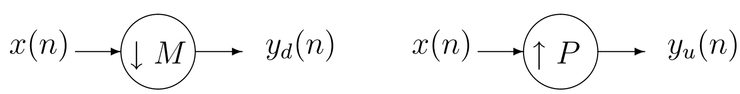


Figure 1.1: A down-sampler by M and an up-sampler by P

Passing an input data stream $s(n)$ through an up-sampler by P produces an output signal $x_u(n)$ where

$$x_u(n) = \begin{cases} s(n/P), & n = kP, k = 0, 1, 2, \cdots \\ 0, & n \neq kP. \end{cases}$$

which has a sampling rate P times that of the original signal $s(n)$. In $\mathcal{Z}$-transform domain,

$$X_u(z) = S(z^P). \tag{1.2}$$

The input signal $s(n)$ can be completely recovered from $x_u(n)$. In practice, a properly designed lowpass filter is used after the up-sampling process to remove the unwanted images in the spectrum of the up-sampled signal, which was caused by the up-sampling process. The down-sampler and the up-sampler are linear but (periodically) time-varying discrete-time systems.

- **Cascade Equivalencies**

Two useful and simple cascade equivalence relations, which are known as noble identities are illustrated in Figure 1.2. These relations are often used to convert periodically time-varying systems into time-invariant systems, thus making the mathematical manipulations of the multirate networks simpler.

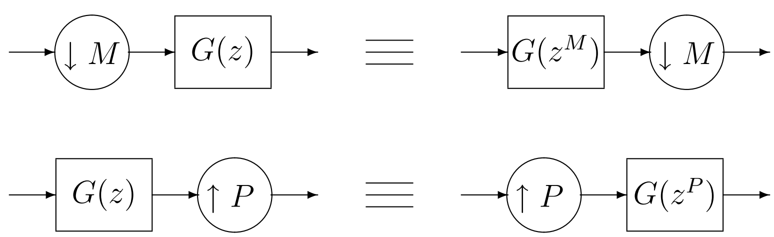


Figure 1.2: The noble identities for multirate systems

1.3.2 Blocking Signals

The standard technique for treating periodic/multi-rate systems is called blocking in signal processing. Let ℓ_+ be the space of causal discrete-time signals defined on the time index set {0,1,2,....}. A signal s in ℓ_+ can be written as

$$s = \{s(0), s(1), s(2), \cdots\} = \{s(n)\}_{n=0}^{\infty}.$$

For an integer $M > 0$, define the M-fold blocking operator, L_M, via $\underline{s} = L_M s$ (where underlining denotes the blocked signal). Blocking a signal $\{s(n)\}_{n=0}^{\infty}$ by a factor M converts the serial data stream into M parallel substreams $\{s_m(n) := s(nM + m)\}_{n=0}^{\infty}$ for $m = 0, 1, \cdots, M-1$. Thus $s_m(n)$ is the mth symbol in the nth block of symbols, yielding a blocked signal $\{\underline{s}(n)\}_{n=0}^{\infty}$:

$$\begin{aligned} s = \{s(0), s(1), s(2), \cdots\} \mapsto \underline{s} &= \{\underline{s}(0), \underline{s}(1), \underline{s}(2), \cdots\} \\ &= \left\{ \begin{bmatrix} s(0) \\ s(1) \\ \vdots \\ s(M-1) \end{bmatrix}, \begin{bmatrix} s(M) \\ s(M+1) \\ \vdots \\ s(2M-1) \end{bmatrix}, \cdots \right\} \end{aligned} \tag{1.3}$$

The blocking operator L_M maps ℓ_+ to ℓ_+^M, the external direct sum of M copies of ℓ_+. If the underlying sampling period for s is T_s, then the underlying sampling period for $\underline{s}$ is MT_s. The blocking operation results in no loss of information.

The blocking operation can be achieved using unit advance elements and down-samplers by the blocking factor M as in Figure 1.3.

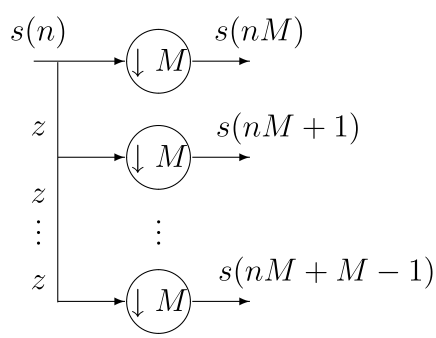


Figure 1.3: Blocking a signal

In reality the unit advance element can not be implemented but blocking operation can be achieved also by using unit delay elements which is implementable.

1.3.3 Blocked Signal and Polyphase Decomposition

Applying the blocking operator L_M to a signal $\{s(n)\}_{n=0}^{\infty}$ yields a blocked signal

$$\underline{s} = \{\underline{s}(n)\}_{n=0}^{\infty}, \quad \underline{s}(n) = \begin{bmatrix} s_0(n) \\ s_1(n) \\ \vdots \\ s_{M-1}(n) \end{bmatrix}, \tag{1.4}$$

where $s_m(n) = s(nM + m)$ for $m = 0, 1, \cdots, M-1$ as earlier.

In the frequency domain, bring in the M-fold polyphase decomposition for the $\mathcal{Z}$-transform of $s(n)$:

$$S(z) = \mathcal{Z}\{s\} = \sum_{n=0}^{\infty} s(n)z^{-n} = \sum_{m=0}^{M-1} z^{-j} S_m(z^M), \quad S_m(z) = \sum_{n=0}^{\infty} s(nM+m)z^{-n}. \tag{1.5}$$

The representation is unique for the given integer M. It follows that the mth polyphase component $S_m(z)$ is the $\mathcal{Z}$-transform of $\{s_m(n)\}_{n=0}^{\infty} = \{s(nM+m)\}_{n=0}^{\infty}$, and there holds

$$\underline{S}(z) = \sum_{n=0}^{\infty} \underline{s}(n)z^{-n} = \begin{bmatrix} S_0(z) & S_1(z) & \cdots & S_{M-1}(z) \end{bmatrix}^T, \tag{1.6}$$

where the superscript ^{T} denotes transpose.

1.3.4 Blocking Systems

Consider the system in Figure 1.4(a) where an up-sampler is followed by a filter $H(z) = \sum_{j=0}^{P-1} z^{-j} H_j(z^P)$ which is represented by its polyphase components. The system is equivalent to Figure 1.4(b) where the up-sampler by factor P goes to each polyphase component branch. Applying noble identity to each branch of Figure 1.4(b) leads to an equivalent system as in Figure 1.4(c).

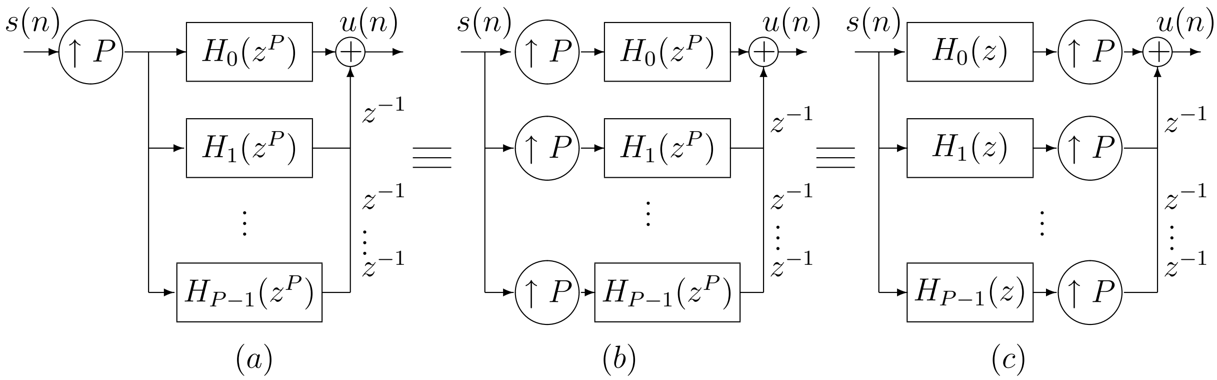


Figure 1.4: Blocking a system

The input-output relationship can be written as

$$U(z) = \sum_{j=0}^{P-1} z^{-j} S(z^P) H_j(z^P).$$

That is, the polyphase components of $U(z)$ are given by

$$U_j(z) = S(z)H_j(z), \qquad j = 0, 1, \cdots, P-1.$$

The blocked input-output relationship is given by

$$\underline{U}(z) = \begin{bmatrix} U_0(z) \\ U_1(z) \\ \vdots \\ U_{P-1}(z) \end{bmatrix} = \begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{P-1}(z) \end{bmatrix} S(z). \tag{1.7}$$

Let the time-invariant system $H(z)$ be represented by its state-space model as

$$\begin{aligned} H(z) &= D_h + C_h(zI - A_h)^{-1}B_h \\ &= D_h + z^{-1}C_h(I - A_h z^{-1})^{-1}B_h = D_h + z^{-1}C_h \left[\sum_{i=0}^{\infty}(A_h z^{-1})^i\right] B_h \\ &= D_h + z^{-1}C_h \left[I + A_h z^{-1} + A_h^2 z^{-2} + \cdots\right] B_h. \end{aligned} \tag{1.8}$$

The state-space representation of the polyphase components of $H(z)$ can be found using the fact that the coefficients for the jth polyphase component are the coefficients of $z^{-(nP+j)}$ for $n = \{0, 1, \cdots\}$. Therefore, the zeroth polyphase component of $H(z)$ is given by

$$\begin{aligned} H_0(z^P) &= D_h + C_h A_h^{P-1} B_h z^{-P} + C_h A_h^{2P-1} B_h z^{-2P} + \cdots = D_h + C_h \sum_{i=1}^{\infty} A_h^{iP-1} B_h z^{-iP} \\ &= D_h + C_h A_h^{P-1} \sum_{i=1}^{\infty} (A_h^P)^{i-1} B_h z^{-iP} = D_h + C_h A_h^{P-1} z^P \sum_{i=1}^{\infty} (A_h^P z^{-P})^{i-1} B_h \\ &= D_h + C_h A_h^{P-1} z^P (I - A_h^P z^{-P})^{-1} B_h = D_h + C_h A_h^{P-1} (z^P I - A_h^P)^{-1} B_h. \end{aligned}$$

Similarly, the jth polyphase component of $H(z)$ where $j = 1, 2, \cdots$ has the form:

$$\begin{aligned} H_j(z^P) &= C_h A_h^{j-1} B_h + C_h A_h^{P+j-1} B_h z^{-P} + C_h A_h^{2P+j-1} B_h z^{-2P} + \cdots \\ &= C_h A_h^{j-1} B_h + C_h \sum_{i=1}^{\infty} A_h^{iP+j-1} B_h z^{-iP} \\ &= C_h A_h^{j-1} B_h + C_h A_h^{P+j-1} z^{-P} [\sum_{i=1}^{\infty} (A_h^P z^{-P})^{i-1}] B_h \\ &= C_h A_h^{j-1} B_h + C_h A_h^{P+j-1} (z^P I - A_h^P)^{-1} B_h. \end{aligned}$$

Therefore, the vector of the polyphase components of $H(z)$ can be written as

$$\begin{bmatrix} H_0 \\ H_1 \\ H_2 \\ \vdots \\ H_{P-1} \end{bmatrix} = \begin{bmatrix} D_h \\ C_h B_h \\ C_h A_h B_h \\ \vdots \\ C_h A_h^{P-2} B_h \end{bmatrix} + \begin{bmatrix} C_h \\ C_h A_h \\ C_h A_h^2 \\ \vdots \\ C_h A_h^{P-1} \end{bmatrix} A_h^{P-1} (zI - A_h^P)^{-1} B_h = D + C(zI - A)^{-1} B. \tag{1.9}$$

1.3.5 State-Space Models for Linear Systems

Let a finite dimensional discrete-time linear time-invariant system G be described by the following state-space model:

$$x(n+1) = Ax(n) + Bu(n), \quad x(0) = x_0, \tag{1.10}$$

$$y(n) = Cx(n) + Du(n), \quad n \geq 0, \tag{1.11}$$

where $x(n) \in \mathbf{R}^N$ is called system *state variable*, $x(0)$ is called *initial condition*, $u(n) \in \mathbf{R}^m$ is called the system *input*, and $y(n) \in \mathbf{R}^p$ is the system *output*. The matrices A, B, C, and D are appropriately dimensioned constant matrices. The corresponding transfer matrix from $u(n)$ to $y(n)$ is defined as

$$Y(z) = G(z)U(z)$$

where $U(z)$ and $Y(z)$ are the $\mathcal{Z}$-transform of $u(n)$ and $y(n)$ respectively with zero initial conditions. Hence,

$$G(z) = C(zI_N - A)^{-1}B + D = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

The four-tuple matrix (A, B, C, D) is called a realization of $G(z)$.

For time-varying case, let a finite dimensional discrete linear time-varying $g(n,k)$ be described by the following state-space model:

$$\begin{aligned} x(n+1) &= A(n)x(n) + B(n)u(n), \quad x(0) = x_0, \\ y(n) &= C(n)x(n) + D(n)u(n), \quad n \geq 0. \end{aligned} \tag{1.12}$$

The matrices $A(n)$,$B(n)$, $C(n)$, and $D(n)$ are appropriately dimensioned time-dependent constant matrices, which vary with time n. The corresponding convolution relation from

$u(n)$ to $y(n)$ is defined as

$$y(n) = g(n,k) \star u(n) = \sum_{k=0}^{n} g(n,k)u(k)$$

where $g(n,k)$ is a causal system

$$g(n,k) = \left[\begin{array}{c|c} A(n) & B(n) \\ \hline C(n) & D(n) \end{array}\right].$$

- **Important Concepts in Linear System Theory**

Some important concepts and notions in linear system theory are reviewed briefly as follows (see references [18, 76]).

Definition 1.1 *The system described by state-space equation (1.10), or the pair* (A, B) *is said to be controllable if, for some* $K > 0$*, initial state* $x(0) = x_0$*, and final state* x_1*, there exists an input* $u(\cdot)$ *such that the solution of (1.10) satisfies* $x(K) = x_1$*. Otherwise, the system or the pair* (A, B) *is said to be uncontrollable.*

Controllability is a system property. It can be shown that (A, B) is controllable, if and only if the eigenvalues $\{\lambda_i(A + BF)\}$ can be arbitrarily assigned through appropriate design of the state feedback gain F.

Definition 1.2 *The dynamical system described by equations (1.10) and (1.11) or by the pair* (C, A) *is said to be observable if, for some* $K > 0$*, the initial state* $x(0) = x_0$ *can be determined from the time history of the input* $u(n)$ *and the output* $y(n)$ *in the interval of* $[0, K]$*. Otherwise, the system or the pair* (C, A) *is said to be unobservable.*

Again the observability is a system property. It can be shown that (C, A) is observable, if and only if the eigenvalues $\{\lambda_i(A+LC)\}$ can be arbitrarily assigned through appropriate design of the state estimation gain L.

Controllability and observability of a state-space system are dependent on its realization, and in particular related to minimality of the realization.

Definition 1.3 *A state-space realization (A, B, C, D) of $G(z)$ is said to be minimal realization of $G(z)$ if A has the smallest possible dimension.*

Theorem 1.1 *A state-space realization (A, B, C, D) of $G(z)$ is minimal, if and only if (A, B) is controllable, and (C, A) is observable.*

Stability is also an important notion for system theory which is given in the next definition.

Definition 1.4 *An unforced dynamical system $x(n+1) = Ax(n)$ is said to be stable if for every nonzero initial condition $x(0) \in \mathbf{R}^N$, the state variable $x(n) \to 0$ asymptotically. That is, stability is equivalent to the condition that all eigenvalues of A lie strictly inside of unit circle. A system matrix A is called stable, if the corresponding unforced dynamic system is stable.*

Stabilizability and detectability are two parallel notions to controllability and observability, respectively.

Definition 1.5 *The dynamical system (1.10), or the pair (A, B), is said to be stabilizable if there exists a state feedback $u(n) = Fx(n)$ such that the overall system is stable, i.e.,*

$A + BF$ is stable. The system, or the pair (C, A), is said to be detectable if there exists an output injection $v(n) = Ly(n)$ such that the overall system is stable, i.e., $A + LC$ is stable.

Stability of the dynamic system (1.10) can be determined by using the Lyapunov method.

Theorem 1.2 *(Lyapunov) Suppose that (A, B) is stabilizable. Then a matrix A is a stability matrix, if and only if the Lyapunov equation*

$$P - APA^* = BB^*$$

admits a unique solution P which is positive semi-definite. Similarly, suppose that (C, A) is detectable. Then a matrix A is a stability matrix, if and only if the Lyapunov equation

$$Q - A^*QA = C^*C$$

admits a unique solution Q which is positive semi-definite.

For linear systems, a stronger notion can be used for stability.

Definition 1.6 *The time-varying system described by the state-space model (1.12) is exponentially stable, or simply stable, if there exists an $N_0 > 0$ such that*

$$\rho \left(\prod_{n=n_0}^{N+n_0} A(n) \right) \leq \alpha \beta^N, \quad \alpha > 0, \quad 0 < \beta < 1,$$

for all $N \geq N_0 > n_0 \geq 0$, where $\rho(\cdot)$ denotes the spectral radius, and α, β are independent of N.

If the state-space model is time-invariant, or $A(n) \equiv A$ is a constant matrix, then exponential stability reduces to $\rho(A) < 1$. That is, for linear, time-invariant, and finite-dimensional systems, stability and exponential stability are equivalent to each other. However such an equivalence does not hold in general for nonlinear and time-varying systems.

Chapter 2

Signal Models and Analysis

In this chapter, different signal and system models will be presented, and the input-output analysis will be carried out for both time-invariant and time-varying systems, as well as for the wide-sense stationary (WSS) and non-stationary white noises.

Channel equalization has been studied extensively in digital communications to combat inter-symbol-interference (ISI) [50, 75]. It is shown in [41] that in the absence of noise, a whitening filter is in fact a perfect equalizer. In fact the same holds for the multiuser communication systems such as digital subscribe lines (DSL), and wireless code division multiple access (CDMA) data networks, which are multi-input and multi-out (MIMO) systems. In this chapter we will analyze several different system models, including multi-channel, multirate, multirate filterbank transceiver models. The multirate filterbank transceivers are widely used in discrete wavelet multitone (DWMT) and discrete multitone (DMT) systems, and its mathematical form has a close connection with the DS/SS (direct-sequence/spread-sequence) CDMA data networks, which are widely used in the wireless communication systems. However we will begin with the noise model, and then discuss different system models.

2.1 Noise Model

A random process becomes a random variable when the time is fixed at some particular value. The random variable possesses statistical properties, such as mean, variance and moments. More generally, N-random variables possess statistical properties related to their N-dimensional joint density function. Strictly speaking, a random process is said to be *stationary*, if all its statistical properties do not change with time. Other processes are called non-stationary. Stationary random processes require very stringent conditions, which often do not hold in practice. A class of random processes satisfying less stringent conditions than stationary is the following (see references [47, 48]).

Definition 2.1 *A random process $V(n)$ is called wide-sense stationary (WSS) noise process, if the following two conditions are true:*

$$E\left[V(n)\right] = \overline{V} = \text{constant}$$

$$E\left[V(n)V(n+n_0)\right] = R_V(n_0)$$

where $E[\ \cdot\ ]$ denotes expectation, or mean.

The sequence $\{R_V(n_0)\}$ is called autocorrelation of the random process $V(n)$. By analogy with "white" light that contains all visible light frequencies in its spectrum, we define the WSS white noise as follows.

Definition 2.2 *A sample noise signal $v(n)$ of a WSS random process $V(n)$ with zero mean for all n is called white noise, if the power spectral density of $V(n)$ is constant at*

all frequencies. That is, the power spectral density of a WSS white noise process $V(n)$, which is the Fourier transform of its autocorrelation function, is given by

$$S_{VV}(w) = \frac{\mathcal{N}_0}{2} \ \forall \ \omega, \tag{2.1}$$

where $\mathcal{N}_0$ is a real positive constant.

The autocorrelation function of $V(n)$ can be found by inverse Fourier transform of (2.1) as

$$R_{VV}(n_0) = \frac{\mathcal{N}_0}{2}\delta(n_0). \tag{2.2}$$

The white noise may not necessarily be WSS.

Definition 2.3 *A sample noise signal $v(n)$ of a non-stationary random process $V(n)$ is called white noise, if it has zero mean for all n, and if*

$$E\left[V(n)V(n+n_0)\right] = R_{VV}(n)\delta(n_0) \tag{2.3}$$

where $\delta(n_0) = 0$ for $n_0 \neq 0$, and $\delta(0) = 1$.

By definition white noise processes are independent random processes. A more general white noise process is bandlimited white noise process, defined for WSS random processes.

Definition 2.4 *A WSS random process $V(n)$ having nonzero and constant power spectral density over a finite frequency band and zero elsewhere is called band-limited white noise.*

Random processes can also be vector valued at each time n. Consider $\{\underline{V}(n)\}$, where $\underline{V}(n)$ is d-dimensional random vector for each n.

Definition 2.5 *A random vector process $\underline{V}(n)$ is called wide-sense stationary (WSS) noise process, if the following two conditions are true:*

$$E\left[\underline{V}(n)\right] = \overline{\underline{V}} = \text{constant},$$

$$E\left[\underline{V}(n)\underline{V}^*(n+n_0)\right] = R_{\underline{V}}(n_0).$$

White noise random vector processes are defined as follows.

Definition 2.6 *A sample noise signal $\underline{v}(n)$ of a WSS random process $\underline{V}(n)$ with zero mean for all n is called white noise, if*

$$E[\underline{V}(n+n_0)\underline{V}^*(n)] = R_{\underline{V}}\delta(n_0). \tag{2.4}$$

A sample noise signal $\underline{v}(n)$ of a non-stationary random process $\underline{V}(n)$ with zero mean for all n is called white noise, if

$$E[\underline{V}(n)\underline{V}^*(n+n_0)] = R_{\underline{V}}(n)\delta(n_0). \tag{2.5}$$

Hence for the WSS white noise, its spectral density is a constant matrix $R_{\underline{V}}$ for all frequencies. But for non-stationary white noise, its frequency domain interpretation is lost. Thought out this dissertation, white noise will be assumed for the signal model used, due to its popularity, and close approximation to the noise processes encountered in various communication channels.

2.2 Time-Invariant Models for Channel Equalization

Spacial diversity, and time diversity are often employed in wireless communication to improve the quality of receiving signals. Spacial diversity corresponds to Figure 2.1 in

which more than one channels are used to transmit a single signal, and thus is termed as the multichannel model. The ith channel is represented by its transfer function $H_i(z)$. The receiver filters are represented by their transfer functions $G_i(z)$ for $i = 0, 1, \cdots, P-1$, respectively.

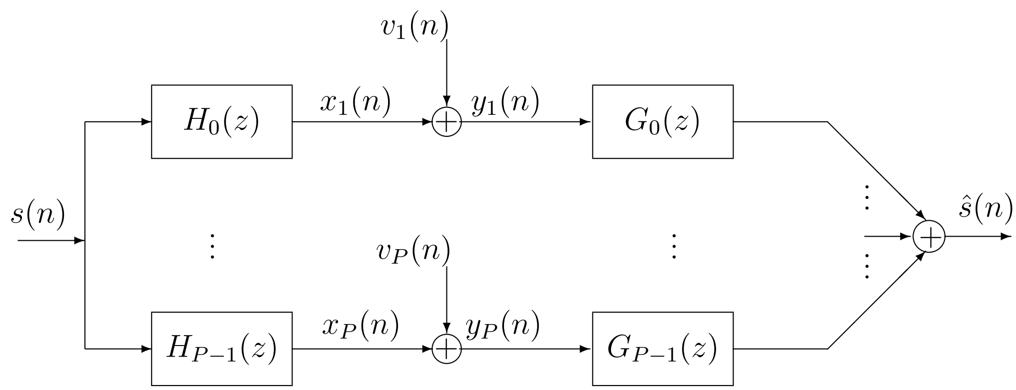


Figure 2.1: Time-Invariant Multichannel models

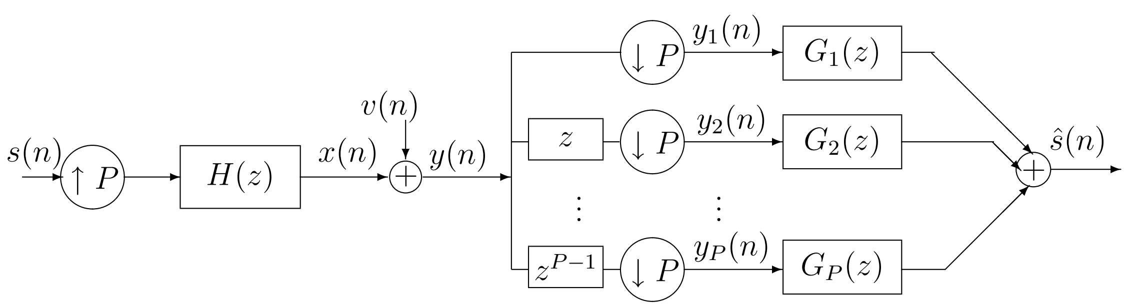


Figure 2.2: Time-Invariant Multirate model

On the other hand, the time diversity illustrated in Figure 2.2 employs an up-sampling by P to increase the sampling rate, and z represents the unit advance operator. It can be easily shown that the spacial diversity, and time diversity are equivalent to each other in terms of channel estimation, and channel equalization. Indeed for Figure 2.2, let $H(z)$ be the channel transfer function, and $S(z)$ be the $\mathcal{Z}$-transform of the input data. Since

the output of the the up-sampler with size P is $S(z^P)$, $Y(z)$ is given by:

$$Y(z) = H(z)S(z^P) + V(z). \tag{2.6}$$

Let the polyphase decomposition of $H(z)$ and $V(z)$ be defined by

$$H(z) = \sum_{p=0}^{P-1} z^{-p} H^{(p)}(z^P), \quad V(z) = \sum_{p=0}^{P-1} z^{-p} V^{(p)}(z^P). \tag{2.7}$$

Substituting the above expression into (2.6) yields

$$Y(z) = \sum_{p=0}^{P-1} z^{-p} \left[H^{(p)}(z^P)S(z^P) + V^{(p)}(z^P) \right]. \tag{2.8}$$

So the pth polyphase component of $Y(z)$ for $0 \le p \le P-1$ is given by

$$Y^{(p)}(z) = H^{(p)}(z)S(z) + V^{(p)}(z).$$

Thus Figure 2.1 and Figure 2.2 are equivalent with $H_p(z) = H^{(p)}(z)$ for $0 \le p \le P-1$.

For recovery of the transmitted signal $S(z)$, the received signal $\hat{S}(z)$ is the sum of filtered $\{Y^{(p)}(z)\}_{p=0}^{P-1}$:

$$\hat{S}(z) = \sum_{p=0}^{P-1} \left[G_p(z)H^{(p)}(z)S(z) + G_p(z)V^{(p)}(n) \right], \tag{2.9}$$

where $G_p(z)$ is the transfer function of the pth receiver filter. Let the system transfer functions, and the noise be blocked with size P via

$$\underline{G}(z) = \begin{bmatrix} G_1(z) & \cdots & G_P(z) \end{bmatrix}, \quad \underline{H}(z) = \begin{bmatrix} H^{(1)}(z) \\ \vdots \\ H^{(P)}(z) \end{bmatrix}, \quad \underline{V}(z) = \begin{bmatrix} V^{(1)}(z) \\ \vdots \\ V^{(P)}(z) \end{bmatrix}. \tag{2.10}$$

Then the perfect reconstruction (PR) condition (which is a slight modification of the zero-forcing condition in (2.38)) requires that

$$\sum_{p=1}^{P} G_p(z)H^{(p)}(z) = \mathcal{K}z^{-d}, \tag{2.11}$$

for some integer $d \geq 0$ and constant $\mathcal{K} \neq 0$, where $H^{(p)}(z)$ and $G_p(z)$ are $\mathcal{Z}$-transforms of the pth polyphase component of the channel, and pth receiver filter, respectively.

In contrast to the multichannel and multirate models in Figure 2.1, and Figure 2.2, [25, 57] proposed a novel filterbank model as shown in Figure 2.3. It is clearly seen that Figure 2.2 is a special case of Figure 2.3. Indeed, if $M = 1$ and $F_0(z) = 1$, then the filterbank in Figure 2.3 is identical to the multirate model in Figure 2.2. It consists of the transmitter filterbank $\{F_m(z)\}_{m=0}^{M-1}$ having down-samplers and up-samplers, the communication channel represented by the transfer function $H(z)$ which is assumed to be causal and stable, and the receiver filterbank $\{G_m(z)\}_{p=0}^{P-1}$ having down-samplers and up-samplers. The observation noise at the output of the channel is assumed to be wide-sense stationary (WSS), and white. It was demonstrated in [57] that the filterbank transceiver provides a useful framework that unifies modulation, precoding, and equalization. Examples of its applications include orthogonal frequency division multiplexing (OFDM), DMT, time-division multiple access (TDMA), CDMA, and discrete-wavelet multiple access (DWMA).

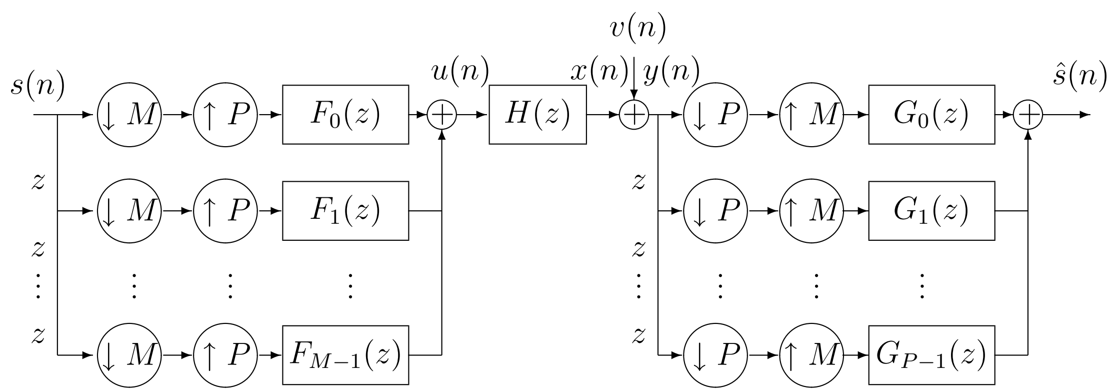


Figure 2.3: Multirate baseband equivalent transmitter/channel/receiver model

2.2.1 Analysis of the Multirate Filterbank Model

It is seen from Figure 2.3 that the input serial data stream $\{s(n)\}_{n=0}^{\infty}$ is converted into M parallel substreams $\{s_m(n) := s(nM+m)\}_{n=0}^{\infty}$ for $m = 0, 1, \cdots, M-1$. Thus $s_m(n)$ is the mth symbol in the nth block of symbols, which is the signal between the down sampler by M and the upsampler by P in the mth branch of the transmitter filterbank. The upsamplers by P insert $(P-1)$ zeros after each symbol. Assuming $P > M$, the redundancy introduced per transmitted block is measured by the ratio $(P - M)/P$, whereas at the receiver, the rate is reduced by the same amount to restore the original input data rate.

By applying blocking operator L_M to the input signal $s(n)$, then

$$\underline{s}(n) = \begin{bmatrix} s(nM) & s(nM+1) & \cdots & s(nM+M-1) \end{bmatrix}^T$$

is the nth block of data with size M. It follows that the mth polyphase component $S_m(z)$ is the $\mathcal{Z}$-transform of $\{s_m(n)\}_{n=0}^{\infty} = \{s(nM+m)\}_{n=0}^{\infty}$, and there holds

$$\underline{S}(z) = \sum_{n=0}^{\infty} \underline{s}(n) z^{-n} = \begin{bmatrix} S_0(z) & S_1(z) & \cdots & S_{M-1}(z) \end{bmatrix}^T. \tag{2.12}$$

Because in $\mathcal{Z}$-transform, the signal $S_m(z)$ after the up-sampler by P becomes $S_m(z^P)$ for $0 \le m < M$, the transmitted signal at the input of the channel, represented by $H(z)$, is given by

$$U(z) = \sum_{m=0}^{M-1} S_m(z^P) F_m(z). \tag{2.13}$$

Let the polyphase decomposition of $F_m(z)$ be defined by

$$F_m(z) = \sum_{p=0}^{P-1} z^{-p} F_{mp}(z^P), \quad m = 0, 1, \cdots, M-1. \tag{2.14}$$

Substituting the above expression into (2.13), and interchanging the order of summations yield

$$U(z) = \sum_{p=0}^{P-1} z^{-j} \sum_{m=0}^{M-1} F_{mp}(z^P) S_m(z^P) = \sum_{p=0}^{P-1} z^{-p} U_p(z^P). \tag{2.15}$$

So the pth polyphase component of $U(z)$ is given by

$$U_p(z) = \sum_{m=0}^{M-1} F_{mp}(z) S_m(z) = \begin{bmatrix} F_{0p}(z) & F_{1p}(z) & \cdots & F_{M-1,p}(z) \end{bmatrix} \underline{S}(z).$$

If $\{u(n)\}_{n=0}^{\infty}$ is blocked with size P at the input of the channel, then

$$\underline{U}(z) := \begin{bmatrix} U_0(z) & U_1(z) & \cdots & U_{P-1}(z) \end{bmatrix}^T = \underline{F}(z)\underline{S}(z), \tag{2.16}$$

where $\underline{F}(z)$ is the blocked transmitter filters of the form

$$\underline{F}(z) = \begin{bmatrix} F_{00}(z) & F_{10}(z) & \cdots & F_{M-1,0}(z) \\ F_{01}(z) & F_{11}(z) & \cdots & F_{M-1,1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ F_{0,P-1}(z) & F_{1,P-1}(z) & \cdots & F_{M-1,P-1}(z) \end{bmatrix} \tag{2.17}$$

which has $\{F_{mp}(z)\}_{p=0}^{P-1}$ as the polyphase components of the mth transmitter filter.

Because the sampling rate through the channel is P times of the sampling rate for the blocked signal $\{\underline{s}(n)\}$, we block the input and output signals of the channel and the corrupting noise with size P as follows:

$$\underline{U}(z) = \begin{bmatrix} U_0(z) \\ U_1(z) \\ \vdots \\ U_{P-1}(z) \end{bmatrix}, \quad \underline{V}(z) = \begin{bmatrix} V_0(z) \\ V_1(z) \\ \vdots \\ V_{P-1}(z) \end{bmatrix}, \quad \underline{Y}(z) = \begin{bmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{P-1}(z) \end{bmatrix}. \tag{2.18}$$

After passing the signal $\{u(n)\}_{n=0}^{\infty}$ through the channel represented by its transfer function $H(z)$, the output of the channel is given by $X(z) = H(z)U(z)$. Upon substituting the

polyphase decompositions of $H(z)$ and $U(z)$, $X(z)$ can be written as

$$\begin{aligned}
X(z) &= H(z)U(z) = \left[\sum_{i=0}^{P-1} H_i(z^P)z^{-i}\right]\left[\sum_{k=0}^{P-1} U_k(z^P)z^{-k}\right] \\
&= \sum_{k=0}^{P-1}\sum_{i=0}^{P-1} H_i(z^P)U_k(z^P)z^{-(i+k)} = \sum_{k=0}^{P-1}\left[\sum_{n=k}^{k+P-1} H_{n-k}(z^P)z^{-n}\right]U_k(z^P) \\
&= \sum_{k=0}^{P-1}\left[\sum_{n=k}^{P-1} H_{n-k}(z^P)z^{-n} + \sum_{n=P}^{k+P-1} H_{n-k}(z^P)z^{-n}\right]U_k(z^P) \\
&= \sum_{k=0}^{P-1}\left[\sum_{n=k}^{P-1} H_{n-k}(z^P)z^{-n} + \sum_{n=0}^{k-1} z^{-n}\left(H_{n-k+P}(z^P)z^{-P}\right)\right]U_k(z^P) \\
&= \sum_{k=0}^{P-1}\sum_{n=0}^{P-1} z^{-n}\tilde{H}_{n,k}(z^P)U_k(z^P) = \sum_{n=0}^{P-1} z^{-n}\left\{\sum_{k=0}^{P-1}\tilde{H}_{n,k}(z^P)U_k(z^P)\right\}
\end{aligned}$$

where by inspection,

$$\tilde{H}_{n,k}(z) = \begin{cases} H_{n-k}(z), & n \geq k, \\ z^{-1}H_{n-k+P}(z), & n < k. \end{cases}$$

That is, the pth polyphase component of $X(z)$ is given by

$$\begin{aligned}
X_p(z) &= \sum_{k=0}^{M-1}\tilde{H}_{n,k}(z^M)U_k(z^M) && (2.19) \\
&= \begin{bmatrix} H_p(z) & H_{p-1}(z) & \cdots & H_0(z) & z^{-1}H_{P-1}(z) & \cdots & z^{-1}H_{P-p+1}(z) \end{bmatrix}\underline{U}(z)
\end{aligned}$$

where $\underline{U}(z)$ as in (2.18)

Block the channel output $X(z)$ as

$$\underline{X}(z) = \begin{bmatrix} X_0(z) & X_1(z) & \cdots & X_{P-1}(z) \end{bmatrix}^T$$

with the same size P. Then $\underline{X}(z) = \underline{H}(z)\underline{U}(z) = \underline{H}(z)\underline{F}(z)\underline{S}(z)$. The blocked channel transfer function $\underline{H}(z)$ follows from (2.19), and is given by

$$\underline{H}(z) = \begin{bmatrix} H_0(z) & z^{-1}H_{P-1}(z) & \cdots & z^{-1}H_1(z) \\ H_1(z) & H_0(z) & \cdots & z^{-1}H_2(z) \\ \vdots & \vdots & \cdots & \vdots \\ H_{P-1}(z) & H_{P-2}(z) & \cdots & H_0(z) \end{bmatrix} \qquad (2.20)$$

with $\{H_k(z)\}_{k=0}^{P-1}$ the polyphase components of $H(z)$. Taking the corrupting noise into consideration yields the (blocked) received signal

$$\underline{Y}(z) = \underline{X}(z) + \underline{V}(z) = \underline{H}(z)\underline{F}(z)\underline{S}(z) + \underline{V}(z). \tag{2.21}$$

The use of $\underline{V}(z)$ should be interpreted as $S_{\underline{v}}(\omega) = \underline{V}(e^{j\omega})[\underline{V}(e^{j\omega})]^*$, the spectral density function of the observation noise. Similarly the pth element of $\underline{Y}(z)$ is the $\mathcal{Z}$-transform of the signal in the pth branch of the receiver filterbank between the down-sampler and up-sampler. This can be verified by noting the polyphase decomposition

$$Y(z) = \sum_{k=0}^{P-1} z^{-k} Y_k(z^P),$$

and that the kth polyphase component $Y_k(z)$ is also the kth element of $\underline{Y}(z)$, which, after passing the up-sampler by M, becomes $Y_k(z^M)$. Therefore the reconstructed signal at the output of the receiver filterbank is given by

$$\begin{aligned}
\hat{S}(z) &= \sum_{k=0}^{P-1} G_k(z) Y_k(z^M) \\
&= \sum_{k=0}^{P-1} \left(\sum_{i=0}^{M-1} z^{-i} G_{ki}(z^M) \right) Y_k(z^M) \\
&= \sum_{i=0}^{M-1} z^{-i} \left(\sum_{k=0}^{P-1} G_{ki}(z^M) Y_k(z^M) \right) = \sum_{i=0}^{M-1} z^{-i} \hat{S}_i(z^M),
\end{aligned}$$

where $\{G_{ki}(z)\}_{i=0}^{M-1}$ is the polyphase components of the kth receiver filter $G_k(z)$ and $\{\hat{S}_i(z)\}_{i=0}^{M-1}$ is the polyphase components of $\hat{S}(z)$. It follows that by blocking the reconstructed signal $\{\hat{s}(n)\}_{n=0}^{\infty}$ with size M or $\hat{\underline{s}} = L_M\{\hat{s}\}$, we have

$$\hat{\underline{S}}(z) = \underline{G}(z)\underline{Y}(z) = \underline{G}(z)\underline{H}(z)\underline{F}(z)\underline{S}(z) + \underline{G}(z)\underline{V}(z), \tag{2.22}$$

where the blocked receiver filterbank $\underline{G}(z)$ is given by

$$\underline{G}(z) = \begin{bmatrix} G_{00}(z) & G_{10}(z) & \cdots & G_{P-1,0}(z) \\ G_{01}(z) & G_{11}(z) & \cdots & G_{P-1,1}(z) \\ \vdots & \vdots & \cdots & \vdots \\ G_{0,M-1}(z) & G_{1,M-1}(z) & \cdots & G_{P-1,M-1}(z) \end{bmatrix} \tag{2.23}$$

2.2.2 Transmultiplexers and Its Analysis

A typical filterbank transceiver in data communications is the multirate transmultiplexer as shown in Figure 2.4, which is widely used in FDMA. It is also called OFDM or DMT. DMT or multicarrier modulation utilizes a set of frequency selective orthogonal functions in digital communications. Since functions in the orthogonal set are designed to be frequency localized, DMT is of a frequency division multiplexing type (FDM) [10, 49]. The terms DMT and OFDM are used interchangeably. Note that this technique multiplexes the incoming bit stream of a single or multiple users into the frequency-selective subcarriers or subchannels. DMT modulation has been widely used in applications such as ADSL, high bit-rate digital subscriber line (HDSL), and very high bit-rate digital subscriber line (VDSL) communications for the single-user case. The digital subscriber line is the local UTP telephone line. In fact, ADSL communication techniques provide the means for high-speed digital transmission, up to 7 Mbps, over plain old telephone service (POTS) for limited distances. A size 512 DFT-based DMT modulation scheme has been standardized for the ADSL communications [17, 19].

DWMT system as proposed in [55] has the same structure as in Figure 2.4, which is also common in DSL (digital subscriber lines). The differences between the DMT

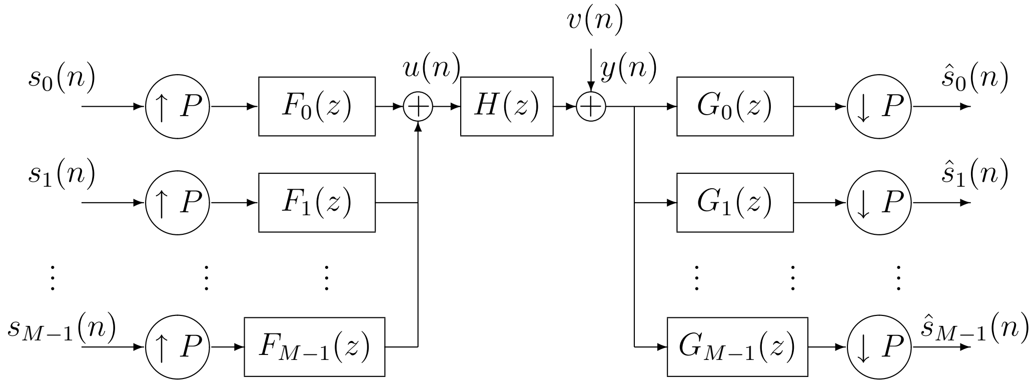


Figure 2.4: Time Invariant Multirate Transmultiplexer model for DMT/DWMT systems

systems is that for DWMT systems, different symbol blocks overlap in time, and thus the DWMT system is also called overlapped DMT. On the other hand, for DMT systems, different symbol blocks do not overlap. Also the DWMT system uses fast wavelet transform (FWT), and inverse fast wavelet transform (IFWT), instead of using discrete Fourier transform (DFT) and inverse DFT (IDFT) to perform multicarrier demodulation and multicarrier modulation, respectively.

Consider the transmultiplexer shown in Figure 2.4. On the transmitter side, the M input data sequences $\{s_i(n)\}_{i=0}^{M-1}$ are up-sampled by P, filtered or frequency shaped, and then transmitted through the distorted channel. On the receiver end, the received signal is contaminated by additive noise $v(n)$, which is filtered, and then down-sampled, in hope to recover the transmitted data sequences. It is assumed that $P > M$, which introduces the redundancies, enabling PR in absence of noise. The advantages of using such multirate transmultiplexers lie in the fact that the channel can be divided into M subbands, and M transmitter filters can be designed to confine the modulated signals within their respective

subbands, so that the transmission power and bits can be judiciously allocated according to the SNR (signal-to-noise ratio) in each band to improve the bit error rate.

In [66], a new model for the DS/SS CDMA network was obtained, which is shown in Figure 2.5. It is easy to see that the CDMA model shown in Figure 2.5 is a special case of the DMT/DWMT model in Figure 2.4 with the channel transfer function $H(z) = 1$ and $F_m(z) = C_m(z)H_m(z)$. However the DS/SS CDMA network has no frequency domain interpretation such as subbands division. In fact all signals of different users use the same frequency channel without dividing into the M subbands. It will be seen later that our optimal channel equalization results apply to the CDMA systems as well.

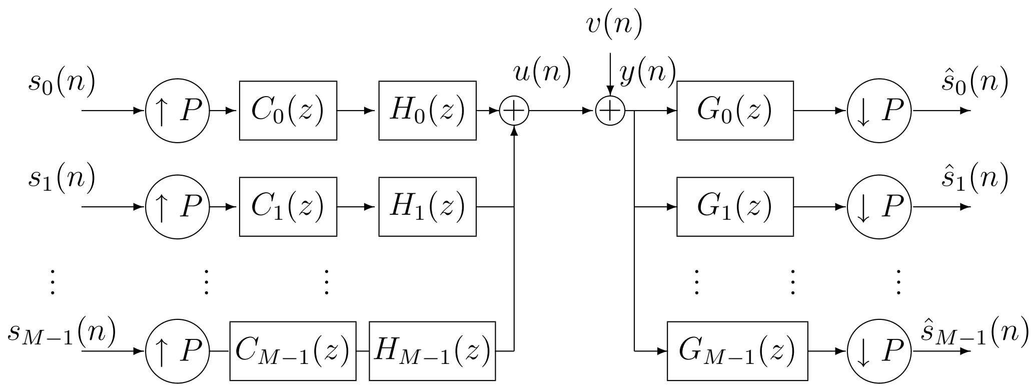


Figure 2.5: Time Invariant Multirate Transmultiplexer model for DS-CDMA Network

For the multirate transmultiplexer shown in Figure 2.4, the channel distortion is represented by the transfer function $H(z)$ which is assumed to be causal and stable. By using the polyphase decomposition of $F_m(z)$ as in (2.14), the polyphase decomposition of $G_m(z)$ as follows:

$$G_m(z) = \sum_{p=0}^{P-1} z^{-p} G_{mp}(z^P) = \sum_{i=1}^{P} z^{-i} \left(z^P G_{m,P-i}(z^P) \right), \quad m = 0, 1, \cdots, M-1, \tag{2.24}$$

and by blocking both the transmitters and receivers filters (see Subsection 1.3.4), the transmultiplexer as in Figure 2.4 has an equivalent filterbank transceiver as in Figure 2.6, where $\underline{F}(z)$ is the same as in (2.17), and $\underline{G}(z)$ is given by

$$\underline{G}(z) = \begin{bmatrix} G_{0,0}(z) & z^{-1}G_{0,P-1}(z) & \cdots & z^{-1}G_{01}(z) \\ G_{1,0}(z) & z^{-1}G_{1,P-1}(z) & \cdots & z^{-1}G_{11}(z) \\ \vdots & \vdots & \cdots & \vdots \\ G_{M-1,0}(z) & z^{-1}G_{M-1,P-1}(z) & \cdots & z^{-1}G_{M-1,1}(z) \end{bmatrix}, \tag{2.25}$$

with $\{G_{ki}(z)\}_{i=0}^{M-1}$ as the polyphase components (of size P) of the kth receiver filter.

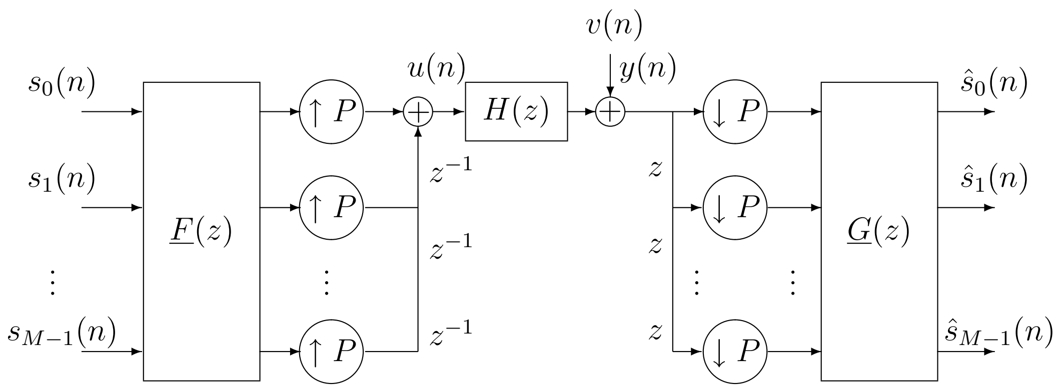


Figure 2.6: Time-Invariant Blocked system

Although the expression of $\underline{G}(z)$ contains pure delay in each column except the first one, which imposes a constraint on the design of the receiver filterbank, this problem can be easily resolved by assuming that the receiver filters are strictly causal. Otherwise we may assume that either the channel is strictly causal, or the transmitter filters are all strictly causal. If none of the above assumptions holds, then we may either make a suitable modification for the blocking in (2.18), or simply add a delay element before or

after the channel in Figure 2.6, then the blocked receiver filters $\underline{G}(z)$ becomes

$$\underline{G}(z) = z^{-1} \begin{bmatrix} G_{0,P-1}(z) & G_{0,P-2}(z) & \cdots & G_{0,0}(z) \\ G_{1,P-1}(z) & G_{1,P-2}(z) & \cdots & G_{1,0}(z) \\ \vdots & \vdots & \cdots & \vdots \\ G_{M-1,P-1}(z) & G_{M-1,P-2}(z) & \cdots & G_{M-1,0}(z) \end{bmatrix}, \tag{2.26}$$

which again removes the constraint in (2.25). Clearly $\underline{U}(z)$, $\underline{V}(z)$, and $\underline{Y}(z)$ as in (2.18) have the same sampling rate as the symbol rate for $\{S_i(z)\}_{i=0}^{M-1}$. There holds

$$\underline{Y}(z) = \underline{H}(z)\underline{U}(z) + \underline{V}(z) \tag{2.27}$$

where the blocked channel $\underline{H}(z)$ as in (2.20), and the blocked signal $\underline{U}(z)$ is given by

$$\underline{U}(z) = \underline{F}(z)\mathbf{S}(z), \qquad \mathbf{S}(z) = \begin{bmatrix} S_0(z) & S_1(z) & \cdots & S_{M-1}(z) \end{bmatrix}^T, \tag{2.28}$$

and $\underline{F}(z)$ as in (2.17). Combining (2.27) and (2.28) gives the rise of the following input/output relation:

$$\underline{Y}(z) = \underline{\Phi}(z)\mathbf{S}(z) + \underline{V}(z), \quad \underline{\Phi}(z) = \underline{H}(z)\underline{F}(z). \tag{2.29}$$

On the receivers end, the blocked output signal $\hat{\mathbf{S}}(z)$ is given by

$$\hat{\mathbf{S}}(z) = \underline{G}(z)\underline{\Phi}(z)\mathbf{S}(z) + \underline{G}(z)\underline{V}(z), \tag{2.30}$$

and $\hat{\mathbf{S}}(z) = \begin{bmatrix} \hat{S}_0(z) & \hat{S}_1(z) & \cdots & \hat{S}_{M-1}(z) \end{bmatrix}^T$.

In summary, both multirate filterbank model in Figure 2.3, and the transmultiplexer model in Figure 2.4 admit a mathematically equivalent input-output relation of the following form

$$\underline{\hat{S}}(z) = \underline{G}(z)\underline{\Phi}(z)\underline{S}(z) + \underline{E}(z), \qquad \underline{E}(z) = \underline{G}(z)\underline{V}(z), \tag{2.31}$$

where the system in consideration is time-invariant, $\underline{S}(z)$, and $\hat{\underline{S}}(z)$ are the $\mathcal{Z}$-transforms of the blocked input symbol and output symbol vectors, respectively, $\underline{G}(z)$ is the transfer function matrix with size $M \times P$ of the blocked receiving filters to be designed, $\underline{\Phi}(z)$ is the transfer function matrix with size $P \times M$ of the composite system consisting of the transmitter filters, and the channel. It is assumed that $P > M$, and $\underline{\Phi}(z)$ is causal, stable, and known, which can be estimated based on observed data. See [63] and references therein for channel estimation. Strictly speaking, $\mathcal{Z}$-transform of the blocked noise $\underline{V}(z)$ is not meaningful mathematically, because the statistical information of the noise is lost. As mentioned earlier, it should be interpreted from the point view of the spectral density function of the noise via

$$S_{\underline{v}}(f) = \underline{V}(e^{j\omega})[\underline{V}(e^{j\omega})]^*.$$

2.2.3 The Problem of Time-invariant Channel Equalization

Consider the input-output relation in (2.31) which is common for both multirate filterbanks and transmultiplexers. Denote Hermitian of a complex matrix Γ by Γ^*, and para-Hermitian of a transfer function matrix $\Lambda(z)$ by $\Lambda^*(\bar{z}^{-1})$ with $\bar{\cdot}$ conjugation. Then the optimal channel equalizer, according to [57], is the one that minimizes the mean-squared-error (MSE):

$$J = \text{trace}\left\{ \frac{1}{j2\pi} \oint_{|z|=1} \underline{G}(z) R_{\underline{v}} \underline{G}^*(\bar{z}^{-1}) \, \frac{dz}{z} \right\}, \quad j = \sqrt{-1}, \tag{2.32}$$

where $R_{\underline{v}}$ is the covariance matrix of the blocked noise $\{\underline{v}(n)\}$, subject to the PR condition: $\hat{S}(z) = \mathcal{K} z^{-q} S(z)$ in the noise-free case for some integer $q \geq 0$, and constant $\mathcal{K} \neq 0$. By

(2.22), the PR condition is equivalent to

$$\underline{G}(z)\underline{H}(z)\underline{F}(z) = \mathcal{K}z^{-q_0}\begin{bmatrix} 0 & I_r \\ z^{-1}I_{M-r} & 0 \end{bmatrix}, \quad q = Mq_0 + r + M - 1, \tag{2.33}$$

in light of [69] where I_s is the identity matrix of size s, and $0 \leq r < M$. By the WSS and white assumption, the noise covariance matrix $R_{\underline{v}} = E\left[\underline{v}(n)\underline{v}^*(n)\right]$ is Hermitian, and positive definite, independent of n. Due to the need for real time implementation, both the transmitter and the receiver filterbanks are required to be causal and stable. Thus the design of optimal channel equalizer is a constrained optimization problem: Design causal and stable transmitter and receiver filters $\{F_k(z)\}_{k=0}^{M-1}$, $\{G_k(z)\}_{k=0}^{P-1}$ that satisfy the PR condition in (2.33), and minimize the performance index J as in (2.32).

2.3 Time-varying Models for Channel Equalization

While the transfer function representation provides a powerful tool for time-invariant systems, it is not applicable to time-varying systems. We will develop parallel analysis for the time-varying models in this section, using impulse responses and convolution operators.

The time-varying multichannel model is illustrated in Figure 2.7. The ith time-varying channel is represented by its impulse responses $h_i(n,k)$ at time n, with unit impulse input applied at time k, where n and k are integer valued. The receiver filters are represented by their impulse responses $g_i(n,k)$ for $i = 0, 1, \cdots, P-1$. On the other hand, the time-varying multirate model (time diversity) illustrated in Figure 2.8 employs an up-sampling by P to increases the sampling rate, and q represents the advance operator.

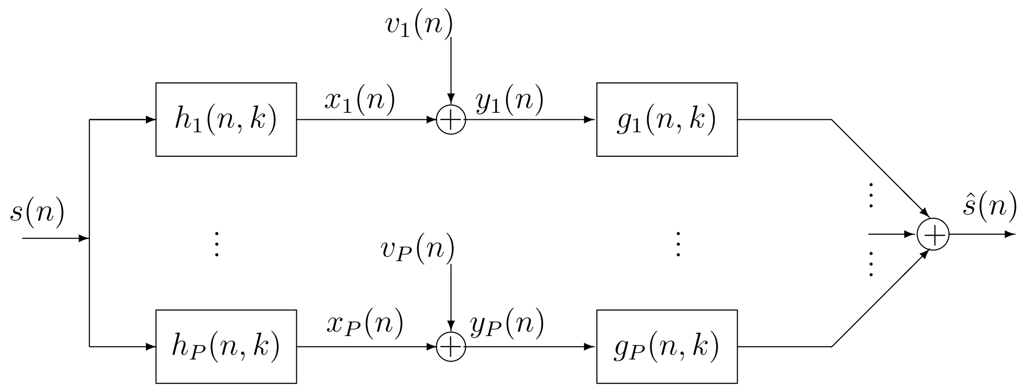


Figure 2.7: Timevarying Multichannel model

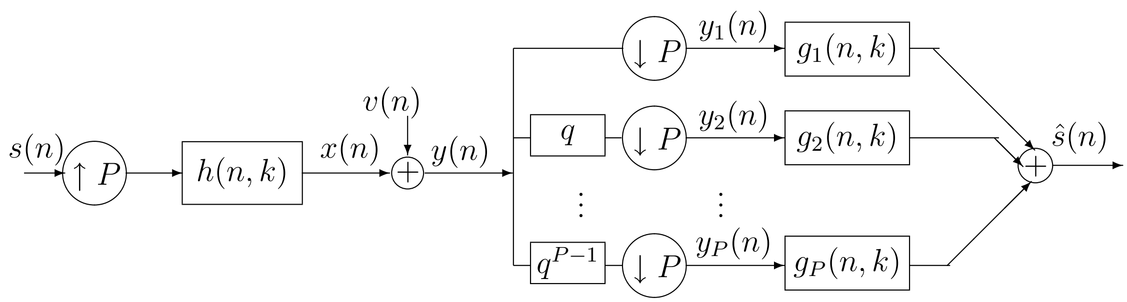


Figure 2.8: Timevarying Multirate model

It was shown that the spacial diversity, and time diversity are equivalent to each other for the time-invariant models. For the sake of completion, the equivalence for the time-varying case will be shown in the following. For the time-varying model in Figure 2.8, let $h(n,k)$ be the channel impulse response at time n with impulse input applied at time k. Denote the pth polyphase component of the channel impulse response at time n by

$$h^{(p)}(n,k) = h(n, kP + p - 1), \quad 0 \leq p \leq P - 1. \tag{2.34}$$

Let $y^{(p)}(k) = y(kP + p - 1)$, and $v^{(p)}(k) = v(kP + p - 1)$ for $0 \le p \le P - 1$. Then

$$y^{(p)}(n) = y_p(n) = h^{(p)}(n,k) * s(n) + v^{(p)}(n) = \sum_{k=-\infty}^{n} h^{(p)}(n,k)s(k) + v^{(p)}(n).$$

Thus Figure 2.7 and Figure 2.8 are equivalent as well for the time-varying case with $h_p(n,k) = h^{(p)}(n,k)$ for $0 \le p \le P - 1$.

For recovery of the transmitted signal $\{s(n)\}$, the received signal $\{\hat{s}(n)\}$ is the sum of filtered $\{y^{(p)}(n)\}_{p=0}^{P-1}$:

$$\begin{aligned} \hat{s}(n) &= \sum_{m=1}^{P} g_p(n,k) * \Big(h^{(p)}(n,k) * s(n)\Big) + e(n) \\ &= \sum_{m=1}^{P} \Big(g_p(n,k) * h^{(p)}(n,k)\Big) * s(n) + e(n) \\ &= \sum_{i=-\infty}^{n} \left[\sum_{m=0}^{P-1} \left(\sum_{k=i}^{n} g_p(n,k) h^{(p)}(k,i)\right) s(i)\right] + e(n), \end{aligned} \tag{2.35}$$

where $\{g_p(n,k)\}$ is the impulse response of the mth (channel/polyphase) receiver filter at time n with impulse input applied at time k, and

$$e(n) := \sum_{m=1}^{P} g_p(n,k) * v^{(p)}(n) = \sum_{m=1}^{P} \sum_{i=-\infty}^{n} g_p(n,k) v^{(p)}(k). \tag{2.36}$$

Block the system impulse responses, and the noise with size P via

$$\underline{g}(n,k) = \begin{bmatrix} g_1(n,k) \\ \vdots \\ g_P(n,k) \end{bmatrix}^T, \quad \underline{h}(n,k) = \begin{bmatrix} h^{(1)}(n,k) \\ \vdots \\ h^{(P)}(n,k) \end{bmatrix}, \quad \underline{v}(n) = \begin{bmatrix} v^{(1)}(n) \\ \vdots \\ v^{(P)}(n) \end{bmatrix}. \tag{2.37}$$

Then the zero forcing condition requires that

$$\underline{g}(n,k) * \underline{h}(n,k) = \sum_{p=0}^{P-1} \left(\sum_{k=0}^{n} g_p(n,k) h^{(p)}(k,0)\right) = \delta(n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0. \end{cases} \tag{2.38}$$

There still holds that the filterbank model proposed in [25, 57] is a general framework for signal models. It is clearly seen that Figure 2.8 is a special case of Figure 2.9. Indeed,

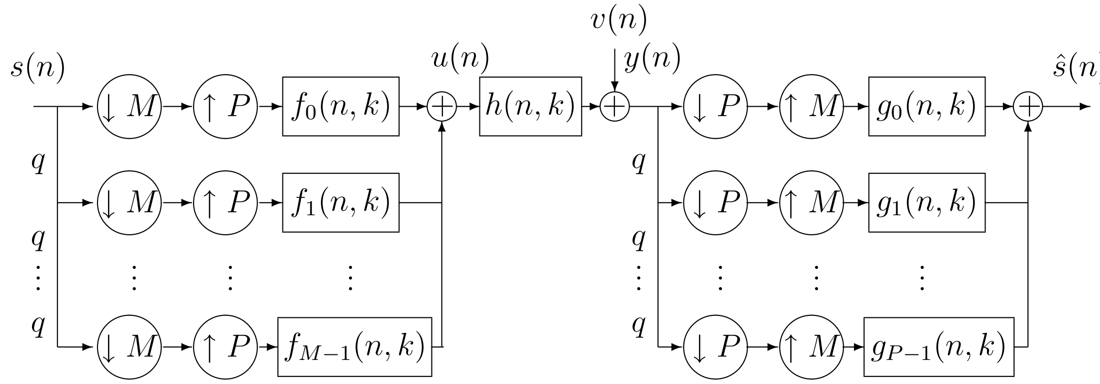


Figure 2.9: A Timevarying Versatile Multirate Filterbank Transceiver

if $M = 1$ and $f_1(n,k) = \delta(n-k)$, then the filterbank in Figure 2.9 is identical to the multirate model in Figure 2.8.

2.3.1 Analysis of the Time-varying Filterbank Model

Similar to the time-invariant case the input serial data stream $\{s(n)\}_{n=0}^{\infty}$ in Figure 2.9 is converted into M parallel substreams $\{s_m(n) := s(nM+m)\}_{n=0}^{\infty}$ for $m = 0, 1, \cdots, M-1$. Thus the signal after the downsampler by M in the mth transmitter filter branch becomes $s(nM+m)$, and after passing it through the upsampler by P it becomes

$$\sum_{k=0}^{\lfloor n/P \rfloor} s(kM+m)\delta(n-kP), \qquad 0 \le m < M,$$

with $\lfloor \cdot \rfloor$ denoting the integer part.

Let $\hat{u}_m(n)$ denote the output of the mth transmitter filter. Denote $\star$ as the convolution operation. There holds

$$\hat{u}_m(n) = s(kM+m) \star f_m(n,kP) = \sum_{k=0}^{\lfloor n/P \rfloor} s(kM+m) f_m(n,kP). \tag{2.39}$$

It follows that the input $u(n)$ to the time-varying channel, represented by its impulse response $h(n,k)$, is the summation of the outputs of the transmitter filters $\hat{u}_m(n)_{m=0}^{M-1}$,

given by

$$u(n) = \sum_{m=0}^{M-1} \sum_{k=0}^{\lfloor n/P \rfloor} s(kM+m) f_m(n, kP), \tag{2.40}$$

where the impulse responses of the transmitter filters are given by $f_m(n,k)$ for $0 \leq m < M$. The input of the channel $\{u(n)\}_{n=0}^{\infty}$ is blocked with size P, and the blocked $\underline{u}(n)$ is given by

$$\underline{u}(n) = \begin{bmatrix} u(nP) & u(nP+1) & \cdots & u(nP+P-1) \end{bmatrix}^T. \tag{2.41}$$

The pth element of the blocked signal $\underline{u}(n)$ for $0 \leq p < P-1$ can be written as

$$\begin{aligned} u(nP+p) &= \sum_{m=0}^{M-1} \sum_{k=0}^{n} s(kM+m) f_m(nP+p, kP) \qquad (2.42) \\ &= \sum_{k=0}^{n} \begin{bmatrix} f_0(nP+p,kP) & f_1(nP+p,kP) & \cdots & f_{M-1}(nP+p,kP) \end{bmatrix} \underline{s}(k) \end{aligned}$$

where $\underline{s}(k) = \begin{bmatrix} s(kM) & s(kM+1) & \cdots & s(kM+M-1) \end{bmatrix}^T$ is the blocked $\{s(k)\}_{k=-\infty}^{\infty}$ with size M. It follows that

$$\underline{u}(n) = \underline{f}(n,k) \star \underline{s}(n) = \sum_{k=0}^{n} \underline{f}(n,k) \underline{s}(k) \tag{2.43}$$

where the blocked transmitter impulse response matrix is given as follows

$$\underline{f}(n,k) = \begin{bmatrix} f_0(nP,kP) & f_1(nP,kP) & \cdots & f_{M-1}(nP,kP) \\ f_0(nP+1,kP) & f_1(nP+1,kP) & \cdots & f_{M-1}(nP+1,kP) \\ f_0(nP+2,kP) & f_1(nP+2,kP) & \cdots & f_{M-1}(nP+2,kP) \\ \vdots & \vdots & \vdots & \vdots \\ f_0(n'P-1,kP) & f_1(n'P-1,kP) & \cdots & f_{M-1}(n'P-1,kP) \end{bmatrix} \tag{2.44}$$

with $n' = n+1$, and $\{f_m(n,k)\}_{p=0}^{P-1}$ the impulse response of the mth transmitter filter at time n to the unit impulse input applied at time k. Thus $f_m(nP+p, kP)$ is the pth polyphase component (of size P) of the mth transmitter filter.

The channel output $x(n)$ is the convolution of the channel input signal $\{u(n)\}_{n=0}^{\infty}$ and the channel impulse response $h(n,k)$ given by

$$x(n) = h(n,k) \star u(k) = \sum_{k=0}^{n} h(n,k)u(k).$$

Applying the blocking operator L_P to the channel output signal yields

$$\underline{x}(n) = \begin{bmatrix} x(nP) & x(nP+1) & \cdots & x(nP+P-1) \end{bmatrix}^T.$$

The pth element of the blocked signal $\underline{x}(n)$ for $0 \leq p < P$ can be written as

$$\begin{aligned} x(nP+p) &= \sum_{k=0}^{\infty} h(nP+p,k)u(k) \qquad (2.45) \\ &= \sum_{k=0}^{n} \sum_{j=0}^{P-1} h(nP+p,kP+j)u(kP+j) \\ &= \sum_{k=0}^{n} \begin{bmatrix} h(nP+p,kP) & \cdots & h(nP+p,kP+P-1) \end{bmatrix} \underline{u}(k). \end{aligned}$$

The output of the channel is also blocked with the same size P, or $\underline{x} = L_P\{x\}$, which results in $\underline{x}(n) = \underline{h}(n,k) \star \underline{u}(n) = \sum_{k=0}^{n} \underline{h}(n,k)\underline{u}(k)$ where the blocked channel impulse response matrix $\underline{h}(n,k)$ following from (2.45) is given by

$$\underline{h}(n,k) = \begin{bmatrix} h(nP,kP) & h(nP,kP+1) & \cdots & h(nP,kP+P-1) \\ h(nP+1,kP) & h(nP+1,kP+1) & \cdots & h(nP+1,kP+P-1) \\ h(nP+2,kP) & h(nP+2,kP+1) & \cdots & h(nP+2,kP+P-1) \\ \vdots & \vdots & \vdots & \vdots \\ h(n'P-1,kP) & h(n'P-1,kP+1) & \cdots & h(n'P-1,k'P-1) \end{bmatrix} \qquad (2.46)$$

with $h(n,k)$ the impulse response of the channel at time n to the unit impulse input applied at time k. Thus $h(nP+p,kP)$ is the pth polyphase component (of size P) of the channel. Taking the corrupting noise into consideration yields the (blocked) received

signal

$$\begin{aligned}\underline{y}(n) &= \underline{h}(n,k) \star \underline{u}(n) + \underline{v}(n) = \underline{h}(n,l) \star \underline{f}(l,k) \star \underline{s}(k) + \underline{v}(n) \\ &= \sum_{l=0}^{n} \sum_{k=0}^{l} \underline{h}(n,l)\underline{f}(l,k)\underline{s}(k) + \underline{v}(n),\end{aligned} \tag{2.47}$$

where $\underline{y}(n)$ and $\underline{v}(n)$ are given by

$$\underline{y}(n) = \begin{bmatrix} y(nP) \\ y(nP+1) \\ \vdots \\ y(nP+P-1) \end{bmatrix}, \quad \underline{v}(n) = \begin{bmatrix} v(nP) \\ v(nP+1) \\ \vdots \\ v(nP+P-1) \end{bmatrix}, \tag{2.48}$$

with $y(nP+p)$ the pth element of $\underline{y}(n)$. Similar analysis as in the transmitter filterbank can be applied to the receiver filterbank, which gives the rise of the following equation for the blocked reconstructed signal at the output of the receiver filterbank:

$$\begin{aligned}\hat{\underline{s}}(n) &= \underline{g}(n,k) \star \left[\underline{h}(n,k) \star \underline{f}(n,k) \star \underline{s}(n) + \underline{v}(n)\right] \\ &= \sum_{j=0}^{n} \sum_{l=0}^{j} \sum_{k=0}^{l} \underline{g}(n,j)\underline{h}(j,l)\underline{f}(l,k)\underline{s}(k) + \sum_{k=0}^{n} \underline{g}(n,k)\underline{v}(k)\end{aligned} \tag{2.49}$$

where $\underline{G}_j(n)$ is the blocked receiver filterbank given by

$$\underline{g}(n,k) = \begin{bmatrix} g_0(nM,kM) & g_1(nM,kM) & \cdots & f_{M-1}(nM,kM) \\ g_0(nM+1,kM) & g_1(nM+1,kM) & \cdots & g_{M-1}(nM+1,kM) \\ g_0(nM+2,kM) & g_1(nM+2,kM) & \cdots & g_{M-1}(nM+2,kM) \\ \vdots & \vdots & \vdots & \vdots \\ g_0(n'M-1,kM) & g_1(n'M-1,kM) & \cdots & g_{M-1}(n'M-1,kM) \end{bmatrix} \tag{2.50}$$

with $\{g_p(n,k)\}_{p=0}^{P-1}$ the impulse response of the pth receiver filter at time n to the unit impulse input applied at time k. Thus $g_p(nM+m,kM)$ is the mth polyphase component (of size M) of the pth receiver filter.

2.3.2 Time-varying Transmultiplexers and Its Analysis

For the time-varying case, there still holds that the DS/SS CDMA model shown in Figure 2.10 is a special case of the transmultiplexer or the DMT/DWMT model with the channel $h(n,k) = \delta(n-k)$, and $f_m(n,k) = c_m(n,k) \star h_m(n,k)$ as in Figure 2.11.

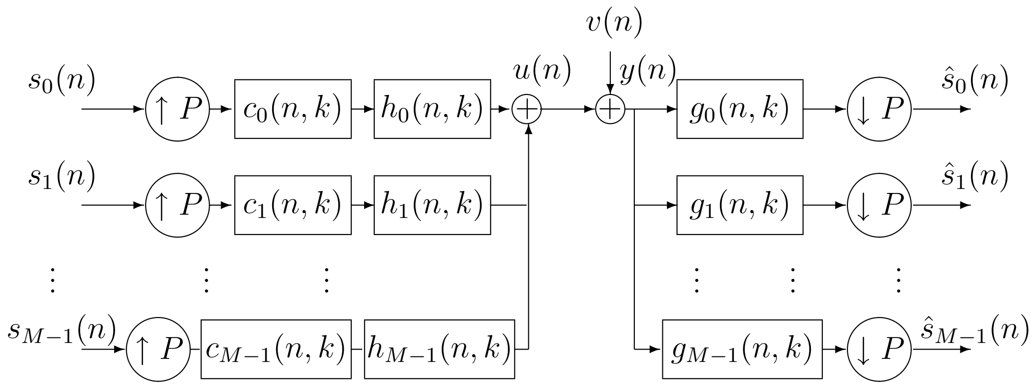


Figure 2.10: Timevarying Multirate Transmultiplexer model for DS-CDMA Network

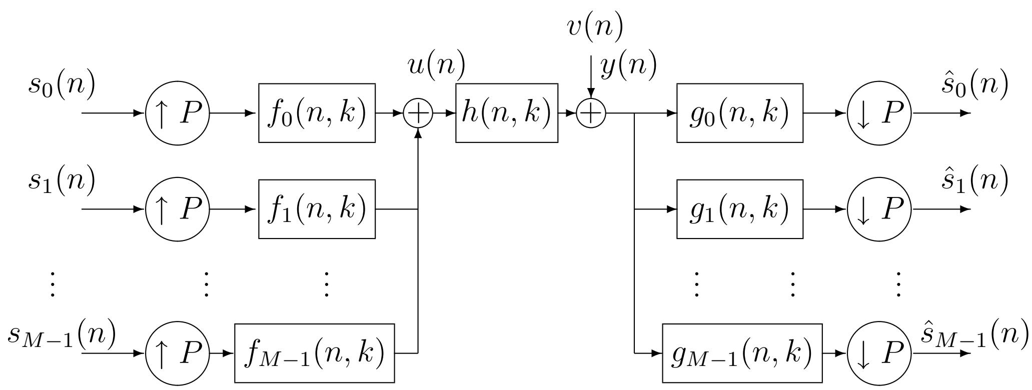


Figure 2.11: Timevarying Multirate Transmultiplexer model for DMT/DWMT systems

Consider the multirate transmultiplexer shown in Figure 2.11. The channel distortion is represented by the time-varying dynamic system $h(n,k)$, which is the impulse response of the channel at time n to the unit impulse input applied at time k, where n and k are integer valued. Causality of $h(n,k)$ implies that $h(n,k) = 0 \;\forall\; n < k$. If the channel is time-

invariant, then $h(n,k) = h(n-k)$. The reason to consider time-varying channels is the need for its applications to wireless communications, where cellular users move constantly, and the channels are dispersive. Because the sampling rate through the channel is P times of the symbol rate, we block the input and output signals of the channel with size P as in (2.41) and (2.48). Clearly $\underline{u}(n)$, $\underline{v}(n)$, and $\underline{y}(n)$ have the same sampling rate as the symbol rate for $\{s_i(n)\}_{i=0}^{M-1}$. If the channel is causal, then there holds

$$\underline{y}(n) = \underline{h}(n,k) \star \underline{u}(n) + \underline{v}(n) = \sum_{k=0}^{n} \underline{h}(n,k)\underline{u}(k) + \underline{v}(k), \tag{2.51}$$

where we assume that the system was at rest at time $n = 0$, and $\underline{h}(n,k)$ is the impulse response of the blocked channel at time n to the unit impulse input applied at time k, given as in (2.46). The expression for $\underline{h}(n,k)$ can be easily verified through routine calculations. It can also be verified that the blocked signal $\underline{u}(n)$ satisfies the convolution relation

$$\underline{u}(n) = \underline{f}(n,k) \star \mathbf{s}(n) = \sum_{k=0}^{n} \underline{f}(n,k)\mathbf{s}(k) \tag{2.52}$$

where $\mathbf{s}(n) = \begin{bmatrix} s_0(n) & s_1(n) & \cdots & s_{M-1}(n) \end{bmatrix}^T$, and $\underline{f}(n,k)$ as in (2.44). Combining (2.51) with (2.52) gives the rise of the following input/output relation:

$$\underline{y}(n) = \underline{\phi}(n,k) \star \mathbf{s}(n) + \underline{v}(n), \quad \underline{\phi}(n,k) = \underline{h}(n,k) \star \underline{f}(n,k) = \sum_{i=k}^{n} \underline{h}(n,k)\underline{f}(i,k), \tag{2.53}$$

by the convolution property, the causality of the input, and the fact that the system was at rest at time $n = 0$. On the receivers end, there holds dynamic relation

$$\hat{\mathbf{s}}(n) = \underline{g}(n,k) \star \underline{y}(n) = \sum_{k=0}^{n} \underline{g}(n,k)\underline{y}(k) \tag{2.54}$$

where $\hat{\mathbf{s}}(n) = \begin{bmatrix} \hat{s}_0(n) & \hat{s}_1(n) & \cdots & \hat{s}_{M-1}(n) \end{bmatrix}^T$, and

$$\underline{g}(n,k) = \begin{bmatrix} g_0(nP,kP) & g_0(n'P-1,kP)\star q^{-1} & \cdots & g_0(nP+1,kP)\star q^{-1} \\ g_1(nP,kP) & g_1(n'P-1,kP)\star q^{-1} & \cdots & g_1(nP+1,kP)\star q^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ g_{M-1}(nP,kP) & g_{M-1}(n'P-1,kP)\star q^{-1} & \cdots & g_{M-1}(nP+1,kP)\star q^{-1} \end{bmatrix}, \tag{2.55}$$

with q^{-1} the delay operator; That is, for any time domain signal $x(n)$, $q^{-1}x(n) = x(n-1)$ for each n, and for any dynamic system with impulse response $h(n,k)$, $q^{-1} \star h(n,k) = h(n-1,k)$ and $h(n,k) \star q^{-1} = h(n,k-1)$ for each n and k.

In light of (2.28), (2.53), and (2.54), we arrive at

$$\hat{\mathbf{s}}(n) = \underline{g}(n,k) \star [\phi(n,k) \star \mathbf{s}(n) + \underline{v}(n)], \quad \underline{\phi}(n,k) = \underline{h}(n,k) \star \underline{f}(n,k), \tag{2.56}$$

yielding an equivalent filterbank transceiver as in Figure 2.12. Thus, the time-varying filterbank transceiver model admits again the same input-output relation as in (2.58).

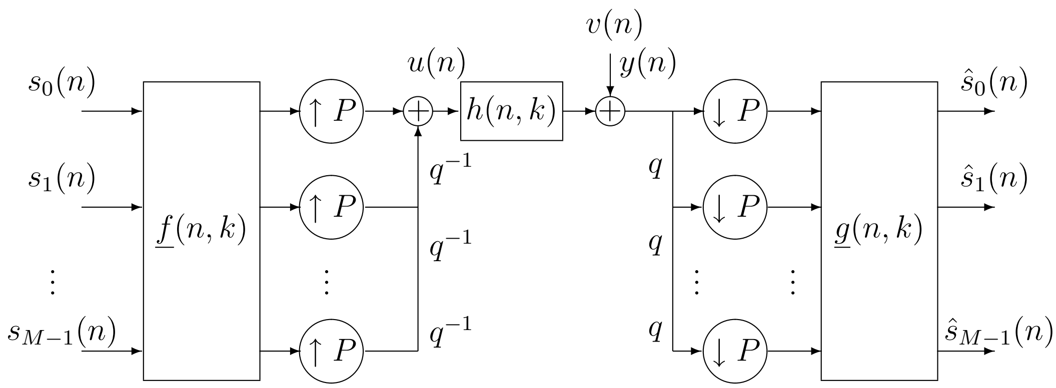


Figure 2.12: Timevarying Blocked system

Although the expression of $\underline{g}(n,k)$ contains pure delay in each column except the first one, which imposes a constraint on the design of the receiver filterbank, this problem can

be easily resolved as in the time-invariant case. We may assume either strictly causal receiver filters, or strictly causal channel, or strictly causal transmitter filters, or simply adding a delay element before or after the channel in Figure 2.12. Then the blocked receiver filters $\underline{g}(n,k)$ becomes

$$\underline{g}(n,k) = \begin{bmatrix} g_0(n'P-1,kP) & g_0(n'P-2,kP) & \cdots & g_0(nP,kP) \\ g_1(n'P-1,kP) & g_1(n'P-2,kP) & \cdots & g_1(nP,kP) \\ \vdots & \vdots & \vdots & \vdots \\ g_{M-1}(n'P-1,kP) & g_{M-1}(n'P-2,kP) & \cdots & g_{M-1}(nP,kP) \end{bmatrix} \star q^{-1}. \quad (2.57)$$

In summary, both time-varying multirate filterbank model in Figure 2.9 and the time-varying transmultiplexer model in Figure 2.11 admit a mathematical equivalent input-output relation of the following form

$$\underline{\hat{s}}(n) = \underline{g}(n,k) \star \underline{\phi}(n,k) \star \underline{s}(n) + \underline{e}(n), \qquad \underline{e}(n) = \underline{g}(n,k) \star \underline{v}(n), \quad (2.58)$$

where $\underline{g}(n,k)$ is the impulse response of the time-varying receiver filter matrix with size $M \times P$ to be designed, and $\underline{\phi}(n,k)$ is the impulse response of the time-varying system matrix with size $P \times M$ that represent the transmitter filters and the channel. It is again assumed that $\underline{\phi}(n,k)$ is known or can be estimated, which is causal and exponentially stable. It should be clear that (2.58) is parallel to (2.31), which is for the time-invariant case.

2.3.3 The Problem of Time-varying Channel Equalization

For optimal channel equalization, our objective is to recover the input data stream perfectly in absence of the noise, and to minimize the MSE, if the noise is present. That is, for the given transmitter filterbank $\underline{f}(n,k)$, and the distorted channel $\underline{h}(n,k)$, which are

possibly time-varying, we seek a causal and stable receiver filterbank $\underline{g}(n,k)$ such that it minimizes

$$J(\underline{g})(n) := E[\underline{e}^*(n)\underline{e}(n)] = \text{trace}\left\{E[\underline{e}(n)\underline{e}^*(n)]\right\}, \quad \underline{e}(n) = \underline{g}(n,k) \star \underline{v}(n), \tag{2.59}$$

with $\underline{v}$ possibly non-stationary white noise, subject to the zero-forcing or PR condition

$$\underline{g}(n,k) \star \underline{\phi}(n,k) = \delta(n-k)I_M = \begin{cases} I_M, & t = k, \\ 0, & t \neq k, \end{cases} \tag{2.60}$$

where the superscript * denotes conjugate transpose. The zero-forcing condition is equivalent to that the receiving filterbank $\underline{g}(n,k)$ is a causal and stable left inverse of $\underline{\phi}(n,k)$, and any $\underline{g}(n,k)$ achieving (2.60) is a channel equalizer, which eliminates the ISI completely.

Chapter 3

Optimal Equalization for Time-invariant Channels with WSS White Noise

3.1 Introduction

The importance of channel equalization has been discussed in the previous chapters. This chapter will investigate optimal channel equalization for time-invariant channels, in presence of WSS white noises. Different from the conventional channel equalization, multiuser data networks allow perfect reconstruction (PR) for the multiple input data signals due to the use of redundancies, provided that some mild conditions on the channel and transmitter filters are satisfied. We will employ the transfer function and state-space methods to tackle the optimal design for channel equalization. The focus will be on the filterbank transceivers introduced, and studied in [25, 57], which provide an important framework for data communication network, although our results also apply to other data networks, which are discussed in Chapter 2. Specifically our work generalizes the results in [57] to more general channels, removes the constraints on the filterbanks, and reduces the complexity. Our contributions include the necessary and sufficient condition

for PR on the transmitter and receiver filterbanks in absence of the noise. It should be mentioned that while joint transceiver optimization is not investigated in this chapter, a similar condition to [57] can be used for the design of the transmitter filters to ensure good signal-to-noise ratio (SNR). More importantly, if the PR condition holds, all causal and stable receiver filterbanks which achieve PR are parameterized. Under the condition that the noise is WSS and white, we show further that the optimal receiver filterbank or optimal channel equalizer has the form of state estimators, and is a modified Kalman filter, which minimizes the MSE among all possible linear time-invariant filters. The rich theory of Kalman filtering enables the optimal design for transmitter/receiver filterbanks by employing efficient and reliable numerical algorithms. Our proposed design algorithm will be illustrated by a simulation example, which has a substantially smaller overall MSE than in [57].

As discussed in the previous chapter, design of optimal channel equalizer is a constrained optimization problem: Design causal and stable transmitter and receiver filters $\{F_k(z)\}_{k=0}^{M-1}$, $\{G_k(z)\}_{k=0}^{P-1}$ that satisfy the PR condition in (2.33), and minimize the performance index J as in (2.32). In reference to the filterbank transceivers as in Figure 2.3, this problem is partially solved in [57] by assuming the following:

(i) The channel, represented by $H(z)$, is an Lth order FIR filter with impulse response $\{h(n)\}_{n=0}^{L}$ satisfying $h(0) \neq 0$, and $h(L) \neq 0$.

(ii) (P, M, L) are chosen such that $P > M$ and $P > L$.

(iii) Each transmitter filter $F_k(z)$ is causal and FIR with order no greater than P for

$0 \leq k < M$, and each receiver filter $G_p(z)$ for $0 \leq p < P$ is also causal and FIR with order QM for some integer Q satisfying $P \geq M + \lceil L/Q \rceil$, with $\lceil \cdot \rceil$ denoting the ceiling integer.

Under the above three assumptions, a design procedure is developed in [57] to design both the transmitter and the receiver filterbanks which achieve PR in the noise-free case, and which minimize the performance index function J as in (2.32). Moreover various other joint optimization criterion for the transceiver filterbank as in Figure 2.3 are investigated in [57] with complete results. It is important to notice that the joint transceiver optimization is necessary, due to the absence of the blocked transmitter filterbank $\underline{F}(z)$ in the performance index (2.32). Indeed, ideally one would prefer large gains for the blocked transmitter filterbank $\underline{F}(z)$, which implies large SNR at the output of the channel. Thus if the PR condition (2.33) holds for some transmitter filterbank $\underline{F}(z)$ and receiver filterbank $\underline{G}(z)$, then we can always replace $\underline{F}(z)$ by $\rho\underline{F}(z)$, and $\underline{G}(z)$ by $\rho^{-1}\underline{G}(z)$, so that the PR condition (2.33) holds for any $\rho \neq 0$. By taking $\rho \to \infty$, the performance index (2.32) can be made arbitrarily small. Such a drawback is overcome by the joint transceiver optimization in [57], leading to the following design constraints:

$$\text{(a)}\;\; G(z) = c\underline{F}^*(\bar{z}^{-1})\underline{H}^*(\bar{z}^{-1})R_{\underline{v}}^{-1}, \quad \text{(b)}\;\; \underline{F}^*(\bar{z}^{-1})\underline{H}^*(\bar{z}^{-1})R_{\underline{v}}^{-1}\underline{H}(z)\underline{F}(z) = \mathcal{K}I_M, \quad (3.1)$$

where $c \neq 0$ can be arbitrary, and $\mathcal{K} > 0$. (a) implies that the composite transfer function matrix $\underline{F}(z)\underline{H}(z)$ has to be FIR, and (b) implies that $\sqrt{\mathcal{K}}^{-1}R_v\underline{F}(z)\underline{H}(z)$ is power complementary [69], where $R_{\underline{v}}^{-1} = \tilde{R}_{\underline{v}}^*\tilde{R}_{\underline{v}}$ for some $\tilde{R}_{\underline{v}}$. Such design constraints can only be ensured by assuming (i) – (iii), which give rise of high complexity due to $P > L$.

Because transmitter filters need take wave shaping, filtering, precoding, and modulation into consideration, $\underline{F}(z)$ may not be completely free for channel equalization. While (b) can always be made true[1], (a) will be removed in this dissertation. Thus it will be assumed in this dissertation that the transmitter filterbank is either given/estimated, or satisfies (b). Our goal is to design an optimal receiver filterbank which minimizes J as in (2.32), subject to the PR condition as in (2.33), without assuming (i) – (iii).

3.2 Causal and Stable Left Inverses and the PR Condition

In this section, we will assume that the channel transfer function $H(z)$, and the transmitter filters $\{F_k(z)\}_{k=0}^{M-1}$ are given to us, and are known. We will first investigate the PR condition (2.33), or the existence of the receiver filters $\{G_k(z)\}_{k=0}^{P-1}$ such that the PR condition (2.33) holds for some integer $q > 0$, and constant $\mathcal{K} \neq 0$.

Theorem 3.1 *Suppose that the channel transfer function $H(z)$, and transmitter filters $\{F_k(z)\}_{k=0}^{M-1}$ are causal and stable. Then there exist causal and stable receiver filters $\{G_k(z)\}_{k=0}^{P-1}$ such that the PR condition (2.33) holds, if and only if there exist some integers q_0 and r such that for $0 \leq r < M$,*

$$\operatorname{rank}\left\{ z^{q_0}\underline{H}(z)\underline{F}(z)\begin{bmatrix} 0 & zI_{M-r} \\ I_r & 0 \end{bmatrix}\right\} = M \quad \forall |z| \geq 1 \text{ (including } |z| = \infty). \tag{3.2}$$

If (3.2) holds, then for the filterbank transceiver model in Figure 2.3, $\hat{S}(z) = z^{-q}S(z)$ in the noise-free case with $q = Mq_0 + r + M - 1$.

[1] If (b) is not satisfied, then we may set $\underline{F}_{\text{new}}(z) = \underline{F}\Omega^{-1}(z)$, where $\Omega(z)$ is the spectral factor satisfying $\Omega^*(\bar{z}^{-1})\Omega(z) = \underline{F}^*(\bar{z}^{-1})\underline{H}^*(\bar{z}^{-1})R_{\underline{v}}^{-1}\underline{F}(z)\underline{H}(z)$. In this case $\underline{F}_{\text{new}}(z)$ satisfies (b), in place of $\underline{F}(z)$.

Proof: The PR condition (2.33) is equivalent to

$$\underline{G}(z)\left(\mathcal{K}^{-1}z^{q_0}\underline{H}(z)\underline{F}(z)\left[\begin{array}{cc} 0 & zI_{M-r} \\ I_r & 0 \end{array}\right]\right) = I_M \quad \forall\ |z| \geq 1$$

for some integers q_0 and r with $0 \leq r < M$. Thus PR implies that $\underline{G}(z)$ is a causal and stable left inverse of the transfer function matrix in (3.2), which implies the rank condition in (3.2). Conversely if the rank condition in (3.2) is true, then a causal and stable $\underline{G}(z)$ exists such that the PR (2.33) holds [44] (which will be made clear subsequently). ∎

The PR condition established in Theorem 3.1 is not as transparent as what we would like it to be, because of the use of the blocked transfer function matrices of $\underline{H}(z)$ and $\underline{F}(z)$, rather than the channel transfer function $H(z)$ and the transmitter filters $\{F_k(z)\}_{k=0}^{M-1}$. However we do have the following characterization of the PR condition for the case $P > M = 1$.

Corollary 3.1 *Suppose that $P > M = 1$ with $F_0(z) = F_{00}(z^P)$, and the hypothesis as in Theorem 3.1 holds. Then the PR condition as in (2.33) is true, if and only if $F_{00}(z)$ is strictly minimum phase, i.e., $F_{00}(z) \neq 0, \forall |z| \geq 1$, and the channel transfer function $H(z)$ has no P finite zeros equally distributed on the circle of radius greater than or equal to 1.*

Proof: The hypotheses imply that

$$\underline{H}(z)\underline{F}(z) = \left[\begin{array}{cccc} H_0(z) & H_1(z) & \cdots & H_{P-1}(z) \end{array}\right]^T F_{00}(z)$$

which is a column of transfer functions. Thus there exists an integer $q_0 \geq 0$ such that

$$\Phi(z) = z^{q_0}\underline{H}(z)\underline{F}(z) = \phi_0 + \phi_1 z^{-1} + \phi_2 z^{-2} + \cdots$$

with ϕ_0, a nonzero column vector of size P. Then z_0 is a zero of $\Phi(z)$, if and only if it is a zero of $F_{00}(z)$, or it is a common zero of $H_k(z)$ for $0 \le k < M$. If z_0 is not a zero of $F_{00}(z)$, then $H_k(z) = (z - z_0)\tilde{H}_k(z)$ for some $\tilde{H}_k(z)$ having no pole at $z = z_0$ and for each k. The fact that $H_k(z)$ is the kth polyphase component of $H(z)$ implies that

$$H(z) = (z^P - z_0) \sum_{k=0}^{P-1} z^{-k} \tilde{H}_k(z^P).$$

Because the roots of $z^P - z_0 = 0$ are equally distributed on the circle of of radius $\rho = |z_0|^{1/P}$, $\underline{H}(z)\underline{F}(z)$ has rank 1 for all $|z| \ge 1$, if and only if $F_{00}(z)$ or $F_0(z)$ is strictly minimum phase, and the channel transfer function $H(z)$ has no P finite zeros equally distributed on the circle of radius greater than or equal to 1. The zeros at infinity can be taken care of by choosing integer $q_0 \ge 0$ appropriately in $\Phi(z)$. Thus the corollary is true, in light of Theorem 3.1. ∎

Corollary 3.1 shows that at least for the case $M = 1$, PR condition can be made true by designing a strictly minimum phase filter $F_0(z) = F_{00}(z^P)$, and by taking adequately large P. A simple case is to take $F_0(z) \equiv 1$, and generically for $P \ge 3$, the PR condition (3.2) holds. In fact very often the PR condition in Corollary 3.1 is true even for the case $P = 2$. It is noted that the strict minimum phase condition is not required in [57] due to assumptions (i) – (iii). That is, the strict minimum phase condition can be removed by increasing the complexity or number of transmitter and receiver filters. Conversely the high complexity issue in [57] can be resolved only if the strict minimum phase condition is restored.

Theorem 3.1 shows that the PR condition is equivalent to the existence of a causal and stable left inverse $\underline{G}(z)$. Assume $\mathcal{K} = 1$. Denote

$$\Phi(z) = z^{q_0}\underline{H}(z)\underline{F}(z) \begin{bmatrix} 0 & zI_{M-r} \\ I_r & 0 \end{bmatrix}. \tag{3.3}$$

We will use a similar method as in [44] to derive a specific left inverse of $\Phi(z)$. For this purpose, we assume that the PR condition as in (3.2) holds, and there exists a minimum state-space realization for $\Phi(z)$ of the form

$$\Phi(z) = D + C(zI_N - A)^{-1}B = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \tag{3.4}$$

with A of size $N \times N$, B of size $N \times M$, C of size $P \times N$, and D of size $P \times N$. The minimum realization of $\Phi(z)$ implies that its McMillan degree is exactly N. The PR condition in Theorem 3.1 is now equivalent to that D has full column rank, and

$$\text{rank} \begin{bmatrix} A - zI_N & B \\ C & D \end{bmatrix} = N + M \quad \forall |z| \geq 1. \tag{3.5}$$

Since $P > M$, D having full column rank implies the existence of $D_\perp$ such that $D_a = \begin{bmatrix} D & D_\perp \end{bmatrix}$ is square and nonsingular. In particular, $D_\perp$ can be chosen from the minimum rank Cholesky factorization

$$D_\perp D_\perp^* = D_{0\perp} D_{0\perp}^* := I_P - D(D^*D)^{-1}D^*. \tag{3.6}$$

It follows that left inverses of D and $D_\perp$ exist, denoted by D^+ and $D_\perp^+$, respectively, and satisfy

$$\begin{bmatrix} D^+ \\ D_\perp^+ \end{bmatrix} \begin{bmatrix} D & D_\perp \end{bmatrix} = \begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}. \tag{3.7}$$

A specific left inverse of D is $D^+ = D_0^+ := (D^*D)^{-1}D^*$. It follows that

$$\Phi(z) = \left[I_P + C(zI_N - A)^{-1}(BD^+ - KD_\perp^+)\right] D$$

for any constant matrix K of size $N \times (P-M)$. Hence with $B_K := BD^+ - KD_\perp^+$,

$$\tilde{\Phi}(z) = D^+ \left[I_P + C(zI_N - A)^{-1}B_K\right]^{-1} = D^+ \left[I_P - C(zI_N - A_K)^{-1}B_K\right] \tag{3.8}$$

is a left inverse of $\Phi(z)$ with $A_K := A - BD^+C + KD_\perp^+C$. It is noted that K has the form of state estimation gain. One may speculate that the optimal left inverse, the one having the smallest mean-squared value, is a Kalman filter which turns out to be incorrect. It will be shown in the next section that the optimal state estimator, or the optimal left inverse is a modified Kalman filter. But it should be clear now that $\Phi(z)$ is strictly minimum phase, if and only if A_K is asymptotically stable, i.e., all eigenvalues of A_K are strictly inside the unit circle. If A_K is asymptotically stable, we will say that K is stabilizing. By the property of $D_\perp$,

$$\Phi_a(z) = \begin{bmatrix} \Phi(z) & \Phi_\perp(z) \end{bmatrix} := \left[I_P + C(zI_N - A)^{-1}B_K\right] \begin{bmatrix} D & D_\perp \end{bmatrix}$$

is square, and strictly minimum phase, if $\Phi(z)$ is, and K is stabilizing. In this case,

$$\begin{aligned} \Phi_a^{-1}(z) &= \begin{bmatrix} D & D_\perp \end{bmatrix}^{-1} \left[I_P + C(zI_N - A)^{-1}B_K\right]^{-1} \\ &= \begin{bmatrix} D^+ \\ D_\perp^+ \end{bmatrix} \left[I_P - C(zI_N - A_K)^{-1}B_K\right] =: \begin{bmatrix} \tilde{\Phi}(z) \\ \tilde{\Phi}_\perp(z) \end{bmatrix} \end{aligned} \tag{3.9}$$

is causal and stable. Similar to the constant case, there holds

$$\begin{bmatrix} \tilde{\Phi}(z) \\ \tilde{\Phi}_\perp(z) \end{bmatrix} \begin{bmatrix} \Phi(z) & \Phi_\perp(z) \end{bmatrix} = \begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}. \tag{3.10}$$

In the next section, we will derive an optimal state estimation gain K_{opt}, and show that with $\underline{G}(z) = \tilde{\Phi}(z)$ and $K = K_{\text{opt}}$, the channel equalization performance index J as in (2.32) is minimized among all linear and time-invariant filters.

3.3 Optimal Channel Equalizer

Although the left inverse $\tilde{\Phi}(z)$ in (3.8) involves state estimation gain K, it does not have the form of Kalman filter, and K is not the same as the Kalman gain. The main reason is that the left inverse D^+ is not unique. Indeed, all left inverses of D are given by

$$D^+ = D_0^+ + \Theta D_{0\perp}^+, \quad D_0^+ := (D^*D)^{-1}D^*, \tag{3.11}$$

where Θ of size $M \times (P - M)$ is a free matrix. Thus the left inverse $\tilde{\Phi}(z)$ as in (3.8) involves two free parameters in optimization, which is beyond the Kalman filter. Furthermore not all causal and stable left inverses of $\Phi(z)$ are in the form of $\tilde{\Phi}(z)$ as in (3.8), which poses the difficulty in design of optimal channel equalizers. Our strategy is to first parameterize all causal and left inverses so that the constrained optimization problem as in Subsection 2.2 can be converted into unconstrained optimization, and then derive the optimal receiver filterbank.

It is noted that with $D_{0\perp}$ defined as in (3.6), $D_{0\perp}^+ = D_{0\perp}^*$. For any left inverse D^+ as in (3.11), we can choose

$$D_\perp = -D\Theta + D_{0\perp}, \quad D_\perp^+ = D_{0\perp}^+ = D_{0\perp}^*, \tag{3.12}$$

so that the identity (3.7) holds true. The next lemma parameterizes all causal and stable left inverses of $\Phi(z)$.

Lemma 3.1 *Suppose that $\Phi(z)$ as in (3.3) satisfies the strict minimum phase condition (3.5), and D has full column rank M. Let $\tilde{\Phi}(z)$ be as in (3.8), and $\tilde{\Phi}_\perp(z)$ as in (3.9) for some stabilizing gain K with D^+ as in (3.11), and $D_\perp$ and $D_\perp^+$ as in (3.12). Then $\Phi^+(z)$ is a causal and stable left inverse of $\Phi(z)$, if and only if*

$$\Phi^+(z) = \tilde{\Phi}(z) + Q(z)\tilde{\Phi}_\perp(z) \tag{3.13}$$

for some causal and stable $Q(z)$.

Proof: It is noted that $\Phi^+(z)$ as in (3.13) is causal and stable, by causality and stability of $\tilde{\Phi}(z)$, $\tilde{\Phi}_\perp(z)$, and $Q(z)$. Employing the identity (3.10) yields $\Phi^+(z)\Phi(z) = I_M$. Thus $\Phi^+(z)$ is indeed a causal and stable left inverse of $\Phi(z)$. Conversely consider any causal and stable left inverse $\Phi^+(z)$ of $\Phi(z)$, and set $Q(z) = \Phi^+(z)\Phi_\perp(z)$. Then $Q(z)$ is causal and stable, and

$$\left[\tilde{\Phi}(z) + Q(z)\tilde{\Phi}_\perp(z)\right]\Phi_\perp(z) = Q(z) = \Phi^+(z)\Phi_\perp(z).$$

The above implies that $\Phi^+(z) = \left[\tilde{\Phi}(z) + Q(z)\tilde{\Phi}_\perp(z)\right] + \Phi(z)\tilde{\Phi}(z)$ for some causal and stable $\Phi(z)$. But $\Phi^+(z)\Phi(z) = I_M$, which concludes that $\Phi(z) = 0$, and thus $\Phi^+(z)$ must be of the form in (3.13) for some causal and stable $Q(z)$. ∎

The constrained optimization formulated as in the previous section is now reduced to unconstrained optimization which is equivalent to minimization of the mean square error J as in (2.32) with $\underline{G}(z) = \tilde{\Phi}(z)+Q(z)\tilde{\Phi}_\perp(z)$ over all possible causal and stable $Q(z)$. While such an optimization procedure of searching for $Q(z)$ can be carried out theoretically, numerical algorithms for design of optimal $\underline{G}(z)$ can be very time consuming, and ill-conditioned. The main result in this chapter shows that the optimal channel equalizer is

actually a modified Kalman filter as demonstrated next. Hence numerically reliable and efficient algorithms can be used for design of optimal channel equalizers [4].

Theorem 3.2 *Suppose that the PR condition (3.2) as in Theorem 3.1 holds for some positive integers r and q_0, with $0 \le r < M$. Let $\Phi(z)$ be as in (3.3) having a realization in (3.4) which satisfies (3.5). Then the following algebraic Riccati equation (ARE)*

$$X - A_0 X A_0^* + (A_0 X C_0^* + BS_0)\tilde{R}^{-1}(A_0 X C_0^* + BS_0)^* - BR_0B^* = 0 \qquad (3.14)$$

has a unique positive (semi)definite solution X, where $\tilde{R} := D_{0\perp}^* R_{\underline{v}} D_{0\perp} + C_0 X C_0^*$, and

$$A_0 = A - BD_0^+ C, \quad C_0 = D_{0\perp}^* C, \quad S_0 = -D_0^+ R_{\underline{v}} D_{0\perp}, \quad R_0 = D_0^+ R_{\underline{v}} (D_0^+)^*.$$

Let X be the unique positive (semi)definite solution to the ARE (3.14). Then $\underline{G}(z) = \tilde{\Phi}(z)$ as in (3.8) with

$$K = K_{\text{opt}} := -AXC^* D_{0\perp}\tilde{R}^{-1}, \quad \Theta = \Theta_{\text{opt}} := (S_0 - D_0^+ CXC_0^*)\tilde{R}^{-1}, \qquad (3.15)$$

is the blocked receiver filterbank achieving optimal channel equalization, where $D^+ = D_0^+ + \Theta_{\text{opt}} D_\perp^+$ and $D_\perp^+ = D_{0\perp}^*$ with $D_{0\perp}$ and D_0^+ as in (3.6) and (3.11), respectively.

Proof: It is noted that the ARE (3.14) is associated with the Kalman filter for the state estimation problem of the discrete-time state-space model (see Section 5.4 of [4]):

$$x(t+1) = A_0 x(t) + B\mu(t), \qquad y(t) = C_0 x(t) + \nu(t), \qquad (3.16)$$

where $\mu(t)$ and $\nu(t)$ are two stationary and white noise processes with covariance

$$\begin{aligned} E\left\{ \begin{bmatrix} \mu(t) \\ \nu(t) \end{bmatrix} \begin{bmatrix} \mu^*(t) & \nu^*(t) \end{bmatrix} \right\} &= \begin{bmatrix} R_0 & S_0 \\ S_0^* & D_{0\perp}^* R_{\underline{v}} D_{0\perp} \end{bmatrix} \\ &= \begin{bmatrix} D_0^+ \\ -D_{0\perp}^* \end{bmatrix} R_{\underline{v}} \begin{bmatrix} (D_0^+)^* & -D_{0\perp} \end{bmatrix} \end{aligned}$$

which is positive semidefinite, by $R_0 = D_0^+ R_{\underline{v}} (D_0^+)^*$ and $S_0 = -D_0^+ R_{\underline{v}} D_{0\perp}$. Since A is asymptotically stable, and the strict minimum phase assumption implies the existence of the state estimation gain L such that $A_0 + LC_0$ is asymptotically stable, the ARE has a unique positive (semi)definite solution X [4]. Now with $K = K_{\rm opt}$ and $\Theta = \Theta_{\rm opt}$ as in (3.15),

$$
\begin{aligned}
A_K &= A - BD^+C + K_{\rm opt} D_\perp^+ C = A - BD_0^+ C + (K_{\rm opt} - B\Theta_{\rm opt}) D_{0\perp}^* C = A_0 + LC_0, \\
B_K &= BD^+ - K_{\rm opt} D_\perp^+ = BD_0^+ - (K_{\rm opt} - B\Theta_{\rm opt}) D_{0\perp}^* = BD_0^+ - LD_{0\perp}^*,
\end{aligned}
$$

where $L = K_{\rm opt} - B\Theta_{\rm opt}$, and (3.11) and (3.12) are used. It follows that A_K is asymptotically stable, and the following Lyapunov equation

$$
Y - A_K Y A_K^* = B_K R_{\underline{v}} B_K^*, \tag{3.17}
$$

has a unique solution. Substituting the expressions of A_K and B_K into (3.17) yields

$$
Y - A_0 Y A_0^* - L\tilde{R}L^* - L(A_0 Y C_0^* + BS_0)^* - (A_0 Y C_0^* + BS_0)L^* - BR_0 B^* = 0. \tag{3.18}
$$

It is noted that by $L = K_{\rm opt} - B\Theta_{\rm opt}$ with $K_{\rm opt}$ and $\Theta_{\rm opt}$ as in (3.15), we have

$$
L = -\left(AXC^* D_{0\perp} + B(S_0 - D_0^+ CXC_0^*)\right)\tilde{R}^{-1} = -(A_0 X C_0^* + BS_0)\tilde{R}^{-1}. \tag{3.19}
$$

Substituting the above L into (3.18) concludes that the Lyapunov equation (3.17) is identical to the ARE (3.14), by setting $Y = X$. To show the optimality of $K = K_{\rm opt}$ and $\Theta_{\rm opt}$ in (3.15), we substitute $\underline{G}(z) = \Phi^+(z)$ as in (3.13) into the performance index (2.32), yielding

$$
J = \text{trace}\left\{ \frac{1}{j2\pi} \oint_{|z|=1} \left[\tilde{\Phi}(z) R_{\underline{v}} \tilde{\Phi}^*(\bar{z}^{-1}) + Q(z)\tilde{\Phi}_\perp(z) R_{\underline{v}} \tilde{\Phi}_\perp^*(\bar{z}^{-1}) Q^*(\bar{z}^{-1}) \right] \frac{dz}{z} \right\} + 2J_1, \tag{3.20}
$$

where J_1 is the contour integral of $z^{-1}Q(z)\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1})$ on the unit circle, or

$$J_1 = \text{trace}\left\{\frac{1}{j2\pi}\oint_{|z|=1} Q(z)\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1})\ \frac{dz}{z}\right\}. \tag{3.21}$$

Assume that A_K has no zero eigenvalues[2]. Then with $\tilde{\Phi}(z)$ as in (3.8),

$$\begin{aligned}
\tilde{\Phi}^*(\bar{z}^{-1}) &= \left[(I_M + B_K^*A_K^{*-1}C^*) + B_K^*A_K^{*-1}(zI_N - A_K^{*-1})^{-1}A_K^{*-1}C^*\right](D^+)^* \\
&= \left[\begin{array}{c|c} A_K^{*-1} & A_K^{*-1}C^* \\ \hline B_K^*A_K^{*-1} & I_M + B_K^*A_K^{*-1}C^* \end{array}\right](D^+)^*.
\end{aligned}$$

Hence we have a composite state-space realization:

$$\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1}) = \overline{D} + \overline{C}(zI_{2N} - \overline{A})^{-1}\overline{B} = \left[\begin{array}{c|c} \overline{A} & \overline{B} \\ \hline \overline{C} & \overline{D} \end{array}\right],$$

where, by $D_\perp = D_{0\perp}$,

$$\overline{A} = \begin{bmatrix} A_K & B_K R_{\underline{v}} B_K^* A_K^{*-1} \\ 0 & A_K^{*-1} \end{bmatrix}, \quad \overline{B} = \begin{bmatrix} B_K R_{\underline{v}} + B_K R_{\underline{v}} B_K^* A_K^{*-1} C^* \\ A_K^{*-1}C^* \end{bmatrix}(D^+)^*, \tag{3.22}$$

$$\overline{C} = D_{0\perp}^*\begin{bmatrix} -C & R_{\underline{v}}B_K^*A_K^{*-1} \end{bmatrix}, \quad \overline{D} = D_{0\perp}^*R_{\underline{v}}(I_P + B_K^*A^{*-1}C^*)(D^+)^*. \tag{3.23}$$

Since similarity transform does not alter transfer function matrices, we have

$$\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1}) = \left[\begin{array}{c|c} \overline{A} & \overline{B} \\ \hline \overline{C} & \overline{D} \end{array}\right] = \left[\begin{array}{c|c} S^{-1}\overline{A}S & S^{-1}\overline{B} \\ \hline \overline{C}S & \overline{D} \end{array}\right], \quad S = \begin{bmatrix} I_N & X \\ 0 & I_N \end{bmatrix},$$

where X is the unique positive definite solution to the ARE in (3.14). Straightforward calculation yields, by (3.17) and $X = Y$,

$$\begin{aligned}
S^{-1}\overline{A}S &= \begin{bmatrix} A_K & (A_K X A_K^* + B_K B_K^* - X)A_K^{*-1} \\ 0 & A_K^{*-1} \end{bmatrix} = \begin{bmatrix} A_K & 0 \\ 0 & A_K^{*-1} \end{bmatrix}, \\
\overline{C}S &= D_{0\perp}^*\begin{bmatrix} -C & R_{\underline{v}}B_K^*A_K^{*-1} - CX \end{bmatrix}.
\end{aligned} \tag{3.24}$$

---

[2]If A_K has zero eigenvalues, then A_K can be replaced by $A_K^\epsilon = \epsilon I_N + A_K$ with ϵ a complex number sufficiently close to zero. Limit $\epsilon \to 0$ can be taken at the end of the proof. See [16] (page 214).

It is claimed that $D_{0\perp}^*(R_{\underline{v}}B_K^* - CXA_K^*)A_K^{*-1} = 0$, which can be verified by noting that

$$\begin{aligned} D_{0\perp}^*(R_{\underline{v}}B_K^* &- CXA_K^*) = D_{0\perp}^*\left(R_{\underline{v}}(BD_0 - LD_{0\perp}^*)^* - CX(A_0 + LD_{0\perp}^*C)^*\right) \\ &= D_{0\perp}^*R_{\underline{v}}(BD_0)^* - D_{0\perp}^*R_{\underline{v}}D_{0\perp}L^* - D_{0\perp}^*CXA_0^* - D_{0\perp}^*CXC^*D_{0\perp}L^* \\ &= D_{0\perp}^*R_{\underline{v}}D_0^*B^* - D_{0\perp}^*CXA_0^* - D_{0\perp}^*(R_{\underline{v}} + CXC^*)D_{0\perp}L^* \\ &= -(A_0XC_0^* + BS_0)^* - \tilde{R}L^* = 0 \end{aligned}$$

in light of (3.19), and $X = Y$. This implies that $\overline{C}S = D_{0\perp}^*\begin{bmatrix} -C & 0 \end{bmatrix}$. Moreover the above equality implies that $D_{0\perp}^+R_{\underline{v}}B_K^* = D_{0\perp}^*CXA_K^*$, and thus $\overline{D}$ as in (1.11) is given by

$$\begin{aligned} \overline{D} &= D_{0\perp}^*R_{\underline{v}}(D^+)^* + D_{0\perp}^*R_{\underline{v}}B_K^*A_K^{*-1}C^*(D^+)^* \\ &= D_{0\perp}^*R_{\underline{v}}(D^+)^* + D_{0\perp}^*CXC^*(D^+)^* \\ &= \left(D_{0\perp}^*R_{\underline{v}} + D_{0\perp}^*CXC^*\right)\left(D_0^+ + \Theta D_{0\perp}^*\right)^* \\ &= \left(D_0^+R_{\underline{v}}D_{0\perp} + D_0^+CXC^*D_{0\perp}\right)^* + \left(D_{0\perp}^*R_{\underline{v}}D_{0\perp} + D_{0\perp}^*CXC^*D_{0\perp}^*\right)\Theta^* \\ &= -\left(S_0 - D_0^+CXC_0^*\right)^* + \tilde{R}\Theta^* = 0 \end{aligned}$$

by the expression of Θ as in (3.15). We can now conclude that

$$\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1}) = -D_{0\perp}^*C(zI_N - A_K)^{-1}B_0$$

for some B_0 (which can be obtained from $S^{-1}\overline{B}$, but is not important here). Thus the integrand $Q(z)\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1})$ of J_1 as in (3.21) is strictly causal, and stable. That is,

$$Q(z)\tilde{\Phi}_\perp(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1}) = \sum_{k=1}^{\infty} w_k z^{-k}, \quad \sum_{k=1}^{\infty} \text{trace}\{w_k w_k^*\} < \infty.$$

It follows from the Cauchy integral theorem that $J_1 = 0$, and thus by (3.20),

$$J \geq \text{trace}\left\{\frac{1}{j2\pi}\oint_{|z|=1} \tilde{\Phi}(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1})\,\frac{dz}{z}\right\} =: J_{\text{opt}}, \tag{3.25}$$

for any causal and stable $Q(z)$, concluding the optimality of $\underline{G}(z) = \tilde{\Phi}(z)$, associated with $K = K_{\text{opt}}$ and Θ_{opt} as in (3.15). The proof is now complete. ∎

The next result gives the expression of the optimal value of the mean square error in (2.32).

Corollary 3.2 *Let X be the unique positive definite solution for* (3.17), *and K_{opt} and Θ_{opt} be given as in (3.15). Then the least mean square error J_{opt} as defined in (3.25) is given by*

$$J_{\text{opt}} := \text{trace}\left\{\left(D_0^+ + \Theta_{\text{opt}} D_{0\perp}^+\right)\left(R_{\underline{v}} + CXC^*\right)\left(D_0^+ + \Theta_{\text{opt}} D_{0\perp}^+\right)^*\right\}. \tag{3.26}$$

Proof: We use a similar derivation as in [27]. By the fact that $Y = X$, (3.17) is equivalent to

$$(zI_N - A_K)X(z^{-1}I_N - A_K^*) + A_K X(z^{-1}I_N - A_K^*) + (zI_N - A_K)XA_K^* = B_K R_{\underline{v}} B_K^*.$$

Multiplying $C(zI_N - A_K)^{-1}$ from left, and $(z^{-1}I_N - A_K^*)^{-1}C^*$ from right yield

$$\begin{aligned} & CXC^* + C(zI_N - A_K)^{-1}A_K XC^* + CXA_K^*(z^{-1}I_N - A_K^*)^{-1}C^* \\ & \qquad = \; C(zI_N - A_K)^{-1}B_K R_{\underline{v}} B_K^*(z^{-1}I_N - A_K^*)^{-1}C^*. \end{aligned}$$

It follows from the above expression that

$$\begin{aligned} & \tilde{\Phi}(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1}) = \\ = \;& D^+\left[I_P - C(zI_N - A_K)^{-1}B_K\right]R_{\underline{v}}\left[I_P - B_K^*(z^{-1}I_N - A_K^*)^{-1}C^*\right](D^+)^* \\ = \;& D^+\left(R_{\underline{v}} + CXC^*\right)(D^+)^* + D^+C(zI_N - A_K)^{-1}(A_K XC^* - B_K R_{\underline{v}})(D^+)^* \\ & \qquad + D^+(CXA_K^* - R_{\underline{v}}B_K^*)(z^{-1}I_N - A_K^*)^{-1}(D^+C)^* \end{aligned}$$

Taking contour integral, and noting that A_K is asymptotically stable lead to

$$\frac{1}{j2\pi}\oint_{|z|=1}\tilde{\Phi}(z)R_{\underline{v}}\tilde{\Phi}^*(\bar{z}^{-1})\,\frac{dz}{z} = D^+(R_{\underline{v}} + CXC^*)(D^+)^*$$

from which the expression of (3.26) follows by using D^+ as in (3.11), and Θ_{opt} as in (3.15).

■

We would like to comment that if the realization of $\Phi(z)$ is minimal, then the solution X to ARE (3.14) is positive definite. But if the realization of $\Phi(z)$ is not minimal, then the solution X to ARE (3.14) can be positive semidefinite. But minimal realization of $\Phi(z)$ is not required. Rather only asymptotic stability of A and the strict minimum phase condition (3.5) need be satisfied for the realization of $\Phi(z)$. We would also like to comment that the proof of Theorem 3.2 shows that the optimal state estimation gain associated with ARE (3.14) is $L = -(A_0XC_0^* + BS_0)\tilde{R}^{-1}$ as in (3.19), not K_{opt} as in (3.15). Hence the optimal blocked receiver filterbank $\underline{G}(z)$ is not exactly the optimal state estimator, or the Kalman filter, due to the free matrix Θ involved in D^+. Hence we call the optimal receiver filterbank $\underline{G}(z)$ (associated with K_{opt} and Θ_{opt} as in (3.15)) a modified Kalman filter. It is important to note that Theorem 3.2 has implications to time-varying channels as well, which will be investigated in the next chapter.

We summarize the proposed design algorithm for the optimal receiver filterbank, assuming that the channel and the transmitter filterbank are known:

- **Design Algorithm for Optimal Channel Equalizers:**
  - Step 1: Find minimal realizations for the blocked transfer function matrices: $\underline{H}(z)$ and $\underline{F}(z)$.

- Step 2: Search for integers $q_0 \geq 0$ and r with $0 \leq r < M$ such that the strict minimum phase condition in Theorem 3.1 is true.

- Step 3: Find a realization for $\Phi(z)$ in (3.3) satisfying (3.5). If $q_0 = 0$, then (3.33) can be used.

- Step 4: Solve ARE (3.14), and set $K = K_{\text{opt}}$ and Θ_{opt} as in (3.15), and set $\underline{G} = \tilde{\Phi}(z)$ as in (3.8).

- Step 5: Set $G_k(z) = \sum_{i=0}^{M-1} z^{-i} G_{ki}(z^M)$ for $0 \leq k < P$, where $G_{ki}(z)$ is the $(i\text{-}k)$th element of $\underline{G}(z)$, obtained in Step 4.

It is noted that if $\underline{H}(z)$ and $\underline{F}(z)$ have minimal realization, then the strict minimum phase condition (3.5) holds. Thus, if Step 3 fails, then the PR condition does not hold. New values of P and M, or/and new transmitter filters need be designed. The procedures in the previous section can be used to obtain a realization for $\Phi(z)$.

3.4 State-space Realizations for FIR/IIR Channels

State-space has become an increasingly important computational tool, and is instrumental in Kalman filtering theory. Our results in this chapter are also built upon the state-space method. In this section we illustrate how to find a simple state-space realization of $\Phi(z)$ as in (3.3), given FIR or IIR channel and transmitter filters.

If the channel $H(z)$ is an FIR filter of the form:

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \cdots + h_L z^{-L}, \quad h_L \neq 0. \tag{3.27}$$

Then a simple realization (a variation of the Brunovsky form) can be used for $H(z) = \left[\begin{array}{c|c} A_h & B_h \\ \hline C_h & D_h \end{array}\right]$:

$$A_h = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \qquad B_h = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix},$$

$$C_h = \begin{bmatrix} h_1 & h_2 & \cdots & h_n \end{bmatrix}, \qquad D_h = h_0. \tag{3.28}$$

The condition $h_L \neq 0$ ensures that the realization is minimal. If the channel has an IIR form (we assume that the numerator and denominator polynomials are coprime, which will also ensure the minimal realization):

$$H(z) = h_0 + \frac{b_1 z^{L-1} + b_2 z^{L-2} + b_2 z^{L-2} + \cdots + b_L}{z^L + a_1 z^{L-1} + \cdots + a_L}, \tag{3.29}$$

then we have the following simple realization (a variation of the Brunovsky form):

$$A_h = \begin{bmatrix} -a_1 & -a_2 & \cdots & \cdots & -a_L \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \qquad B_h = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \tag{3.30}$$

$$C_h = \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix}, \qquad D_h = h_0.$$

With the state-space realization for $H(z)$ available, and $\underline{F}(z) = \begin{bmatrix} I_M & 0 \end{bmatrix}^T$, the

blocked transfer function $\underline{H}(z)\underline{F}(z)$ for the case $P \geq L > M = 1$ is now given by

$$\underline{H}(z)\underline{F}(z) = \begin{bmatrix} H_0(z) \\ H_1(z) \\ H_2(z) \\ \vdots \\ H_{P-1}(z) \end{bmatrix} = \begin{bmatrix} D_h \\ C_h B_h \\ C_h A_h B_h \\ \vdots \\ C_h A_h^{P-2} B_h \end{bmatrix} + \begin{bmatrix} C_h \\ C_h A_h \\ C_h A_h^2 \\ \vdots \\ C_h A_h^{P-1} \end{bmatrix} A_h^{P-1}(zI_L - A_h^P)^{-1} B_h$$
$$= \underline{D}_h + \underline{C}_h (zI_L - \underline{A}_h)^{-1} \underline{B}_h. \tag{3.31}$$

It can be verified by straightforward calculation that for the FIR channel as in (3.27),

$$\underline{A}_h = \begin{bmatrix} 0_{P\times(L-P)} & 0_{P\times P} \\ I_{L-P} & 0_{(L-P)\times P} \end{bmatrix}, \quad \underline{B}_h = \begin{bmatrix} 0_{(P-1)\times 1} \\ e_{L-P} \end{bmatrix}, \quad e_{L-P} = \begin{bmatrix} 1 \\ 0_{(L-P-2)\times 1} \end{bmatrix},$$

$$\underline{C}_h = \begin{bmatrix} h_1 & h_2 & \cdots & h_{L-P+1} & h_{L-P+2} & \cdots & h_L \\ h_2 & h_3 & \cdots & h_{L-P+2} & h_{L-P+3} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & h_L & \cdots & \vdots \\ h_P & h_{P+1} & \cdots & h_L & 0 & \cdots & 0 \end{bmatrix}, \quad \underline{D}_h = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{P-1} \end{bmatrix},$$

where $0_{\alpha\times\beta}$ is the zero matrix of size $\alpha \times \beta$. For the IIR channel as in (3.29), we have

$$\underline{A}_h = \begin{bmatrix} \Phi \\ E_0 \end{bmatrix}, \quad \Phi = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,L} \\ \vdots & \vdots & \cdots & \vdots \\ \phi_{P-1,1} & \phi_{P-1,2} & \cdots & \phi_{P-1,L} \end{bmatrix}, \quad E_0 = \begin{bmatrix} I_{L-P} & 0_{(L-P)\times P} \end{bmatrix},$$

$$\underline{B}_h = \begin{bmatrix} \phi_{2,1} \\ \vdots \\ \phi_{P-1,1} \\ e_{L-P} \end{bmatrix}, \quad \underline{C}_h = \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,L} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & \gamma_{2,L} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{P,1} & \gamma_{P,2} & \cdots & \gamma_{P,L} \end{bmatrix}, \quad \underline{D}_h = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_P \end{bmatrix},$$

where by taking $a_0 = 1$, $\gamma_{1,i} = b_i$ for $1 \leq i \leq L$, and the convention $a_k = 0$ for $k > L$,

$\{\phi_{i,l}\}$, $\{\psi_i\}$, and $\{\gamma_{i,l}\}$ are given recursively by

$$
\begin{aligned}
\phi_{P-i,l} &= -\sum_{k=1}^{i} a_k \phi_{P-i+k,l} - a_{i+l}, \quad i = 1, 2, \cdots, P-1,\ l = 1, 2, \cdots, L, \\
\gamma_{i,l} &= \sum_{k=1}^{i-1} b_{i-k} \phi_{P-k+l,l} + b_{i+l-1}, \quad i = 1, \cdots, P-1,\ l = 1, 2, \cdots, L, \\
\psi_1 &= h_0, \quad \psi_i = \sum_{k=1}^{i-1} b_{i-k} \phi_{P-k+2,1}, \quad i = 2, 3, \cdots, P.
\end{aligned}
$$

For the case $P > M > 1$, we have

$$
\underline{H}(z) = \begin{bmatrix} H_0(z) & z^{-1}H_{P-1}(z) & \cdots & z^{-1}H_1(z) \\ H_1(z) & H_0(z) & \cdots & z^{-1}H_2(z) \\ \vdots & & & \\ H_{P-1}(z) & H_{P-2}(z) & \cdots & H_0(z) \end{bmatrix} =: \underline{D}_h + \underline{C}_h(zI_{LP} - \underline{A}_h)^{-1}\underline{B}_h, \quad (3.32)
$$

where $\underline{A}_h$ =diag (A_h), $\underline{B}_h$ =diag (B_h) with P repeated blocks, and

$$
\underline{C}_h = \begin{bmatrix} C_hA_h^{P-1} & C_hA_h^{P-2} & \cdots & C_h \\ C_hA_h^{P} & C_hA_h^{P-1} & \cdots & C_hA_h \\ \vdots & \vdots & \ddots & \vdots \\ C_hA_h^{2(P-1)} & C_hA_h^{2P-1} & \cdots & C_hA_h^{P-1} \end{bmatrix},
$$

$$
\underline{D}_h = \begin{bmatrix} D_h & 0 & \cdots & 0 \\ C_hB_h & D_h & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_hA_h^{P-2}B_h & C_hA_h^{P-1}B_h & \cdots & D_h \end{bmatrix}.
$$

Assume that the transmitter filters are given, and integers $q_0 \geq 0$ and $0 \leq r < M$ exist such that

$$
\Phi(z) = z^{q_0}\underline{H}(z)\underline{F}(z)\begin{bmatrix} 0 & zI_{M-r} \\ I_r & 0 \end{bmatrix} = D + C(zI_N - A)^{-1}B = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad D \neq 0.
$$

Suppose that a realization of $\underline{F}(z) = D_f + C_f(zI - A_f)^{-1}B_f$ can be obtained with the similar procedure as earlier. If $\underline{D}D_f \neq 0$, then we have

$$\Phi(z) = \left[\underline{D} + \underline{C}(zI_L - \underline{A})^{-1}\underline{B}\right]\left[D_f + C_f(zI - A_f)^{-1}B_f\right] = \left[\begin{array}{cc|c} \underline{A} & \underline{B}C_f & \underline{B}D_f \\ 0 & A_f & B_f \\ \hline \underline{C} & \underline{D}C_f & \underline{D}D_f \end{array}\right]. \quad (3.33)$$

If $\underline{D}D_f = 0$, then a smallest integer $q_0 > 0$ exists such that the column vector

$$D = \begin{bmatrix} \underline{C} & \underline{D}C_f \end{bmatrix} \begin{bmatrix} \underline{A} & \underline{B}C_f \\ 0 & A_f \end{bmatrix}^{q_0 - 1} \begin{bmatrix} \underline{B}D_f \\ B_f \end{bmatrix} \neq 0. \quad (3.34)$$

In this case, the remaining three matrices (A, B, C) of $\Phi(z)$ are given by

$$A = \begin{bmatrix} \underline{A} & \underline{B}C_f \\ 0 & A_f \end{bmatrix}, \quad B = \begin{bmatrix} \underline{B}D_f \\ B_f \end{bmatrix}, \quad C = \begin{bmatrix} \underline{C} & \underline{D}C_f \end{bmatrix} \begin{bmatrix} \underline{A} & \underline{B}C_f \\ 0 & A_f \end{bmatrix}^{q_0}. \quad (3.35)$$

3.5 BER and An Illustrative Example

A commonly used performance measurement is *bit error rate* (BER) probability, rather than the MSE. To find an expression for the BER at the receiver, let $\mathcal{J}$ be define as

$$\mathcal{J} = \frac{1}{j2\pi} \oint_{|z|=1} \underline{G}(z) R_{\underline{v}} \underline{G}^*(\bar{z}^{-1}) \frac{dz}{z}.$$

Thus there holds

$$J(\underline{G}(z) = \Phi^+_{\text{opt}}(z)) = \text{trace}\{\mathcal{J}\}\big|_{\underline{G}=\Phi^+_{\text{opt}}}. \quad (3.36)$$

The expression of $\mathcal{J}$ for the optimal channel equalizer is as in (3.26) given by

$$\mathcal{J} = D^+_{\text{opt}}\left(\underline{R}_{vv} + CXC^*\right)\left(D^+_{\text{opt}}\right)^*, \quad (3.37)$$

where $D^+_{\text{opt}} = \left(D^+_0 + \Theta^{\text{opt}} D^+_{0\perp}\right)$.

With $\underline{\mathbf{S}}(z)$ as in (2.22), the MSE value J_i, the ith diagonal element of $\mathcal{J}$, is given by

$$J_i := \frac{1}{j2\pi} \oint_{|z|=1} \underline{G}_i(z) R_{\underline{v}} \underline{G}_i^*(\bar{z}^{-1}) \ \frac{dz}{z} \tag{3.38}$$

where $\underline{G}_i(z)$ is the ith row of the blocked receivers filters $\underline{G}(z)$. Thus, the MSE for the ith polyphase component/user received data $\mathbf{S}_i(z)$ is J_i, and the average error probability for $\mathbf{S}(z)$ is defined as

$$P_e := \frac{1}{M} \sum_{i=0}^{M-1} P_e^{(i)} \tag{3.39}$$

where $P_e^{(i)}$ denotes the error probability of the ith symbol stream. For BPSK constellations, the error probability of the ith symbol in the case of additive Gaussian noise (AGN) is given by

$$P_e^{(i)} = \frac{1}{2} \text{erfc}\left(\frac{1}{\sqrt{J_i}}\right). \tag{3.40}$$

In the following, we consider an example studied in [57].

- **An Illustrative Example :** The channel is given by

$$H(z) = \sum_{i=0}^{L} h_i z^{-i} = 1 - 0.3z^{-1} + 0.5z^{-2} - 0.4z^{-3} + 0.1z^{-4} - 0.02z^{-5} + 0.3z^{-6} - 0.1z^{-7}$$

with $L = 7$. We choose $P = 4 > M = 3$, which are roughly one-tenth of P and M values in [57], respectively. We set initially $\underline{F}(z) = \begin{bmatrix} I_M & 0 \end{bmatrix}^T$, which yields:

$$\underline{H}(z)\underline{F}(z) = \sum_{i=0}^{2} \Gamma_i z^{-i}, \quad \Gamma_i = \begin{bmatrix} h(iP) & \cdots & h(iP-P+2) \\ h(iP+1) & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ h(iP+P-1) & \cdots & h(iP+1) \end{bmatrix}.$$

A simple realization is

$$\underline{H}(z)\underline{F}(z) = \left[\begin{array}{cc|c} 0_{M\times M} & 0_{M\times M} & I_M \\ I_M & 0_{M\times M} & 0_{M\times M} \\ \hline H_1 & H_2 & H_0 \end{array}\right].$$

Computing spectral factorization:

$$\underline{F}^*(\bar{z}^{-1})\underline{H}^*(\bar{z}^{-1})\underline{H}(z)\underline{F}(z) = \Omega^*(\bar{z}^{-1})\Omega(z),$$

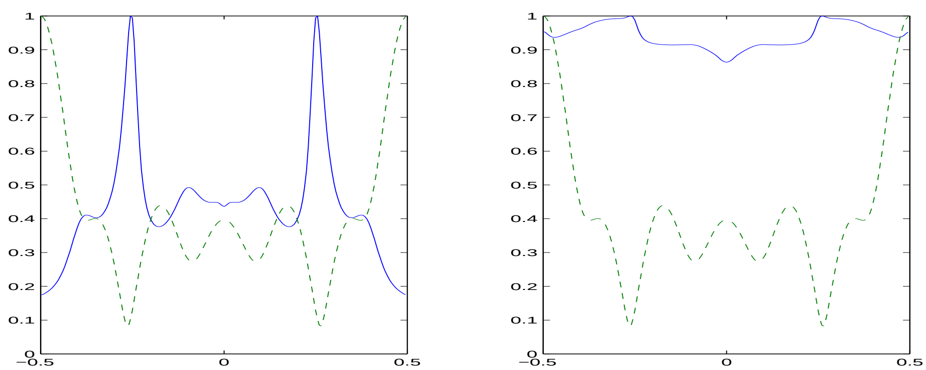
and setting $\underline{F}_{\text{new}}(z) = \begin{bmatrix} I_M & 0 \end{bmatrix}^T \Omega^{-1}(z)$ show that (b) of (3.1) is satisfied with $\underline{F}_{\text{new}}(z)$ in place of $\underline{F}(z)$. The transmitter filters $\{F_k(z)\}_{p=0}^2$ can be easily computed based on their polyphase components in $\underline{F}_{\text{new}}(z)$. To design optimal channel equalizers, we set $\Phi(z) = \underline{H}(z)\underline{F}_{\text{new}}(z)$, corresponding to $q_0 = 0$, $r = M$, and $\mathcal{K} = 1$ as in (3.3). We then use Theorem 3.2 to compute the optimal left inverse $\underline{G}(z)$. Because the stabilizing solution to ARE (3.14) exists such that $A_{K_{\text{opt}}} = A_0 - K_{\text{opt}}C_0$ is asymptotically stable, the strict minimum phase condition in (3.5) holds. The optimal receiver filters $\{G_m(z)\}_{m=0}^3$ are obtained subsequently. The MSE J in (2.32) corresponding to the optimal receivers is $J_{\text{opt}} = 3.188$ using (3.26) for the case with noise variance $\sigma_v = 1$ and $\mathcal{K} = 1$. We would like to emphasize that while our number of channels is roughly one tenth of those used in [57], the corresponding mean square error is also approximately one tenth ($\approx 1/10$) of that in [57] for the same channel.

The average frequency responses of the transmitter and receiver filters are plotted using

$$|\mathcal{F}(f)| := \frac{1}{M}\sum_{m=0}^{M-1} |F_m(e^{j2f\pi})|, \quad |\mathcal{G}(f)| := \frac{1}{M}\sum_{p=0}^{P-1} |G_p(e^{j2f\pi})|.$$

See (a) of Figure 3.1 for $|\mathcal{F}(f)|$ (solid line), and (b) of Figure 3.1 for $|\mathcal{G}(f)|$ (solid line),

versus the channel frequency response (dashed line), all normalized with respect to their respective maximum. It is noted that although we have much fewer filters (approximately one tenth) compared to [57], the transmitter filters have similar average frequency responses. Indeed the transmitter filterbank tends to pre-equalize the channel, to transmit more power at frequencies where the channel attenuation is higher, or vice-versa. But the average frequency responses of the receiver filterbank are much more flat than that in [57].



(a) $|\mathcal{F}(f)$ versus $|H(e^{j2f\pi})|$ (b) $|\mathcal{G}(f)|$ versus $|H(e^{j2f\pi})|$

Figure 3.1: Average frequency responses of transmitter and receiver filterbanks versus channel

To compare the BER, we compute the matrix $\mathcal{J}$ as in (3.36). The diagonal elements of $\mathcal{J}$ are then used to compute the BER probability for the BPSK case as in (3.39) and (3.40). we have used Theorem 3.2 to compute the optimal left inverse $\underline{G}(z)$ with $M = 3$ and $P = 8$, with BER shown in Figure 3.2 as a function of E_b/N_0 for both our design and the design in [57], where E_b/N_0 varies from 0.8264 to 100. It is interesting to observe that the two performance curves coincide with each other, with almost no difference in the BER performance. However we would like to emphasize that we have a much smaller

overall MSE J which corresponding to the optimal receivers, where $J_{\text{opt}} = 3$ using (3.26) for the case with noise variance $\sigma_v = 1$ and $\mathcal{K} = 1$. The corresponding mean square error is less than one tenth ($< 1/10$) of that in [57] for the same channel, because we use a much smaller number of transmitter and receiver filters than those used in [57]. ∎

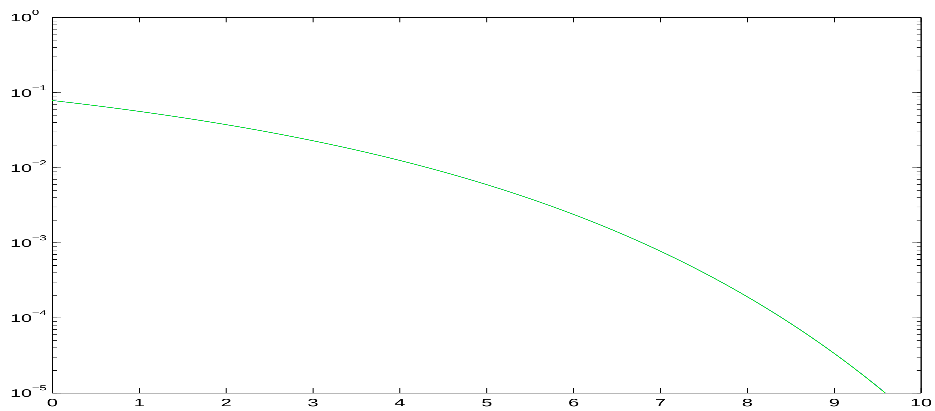


Figure 3.2: Bit error rate BER versus E_b/N_0 dB

Chapter 4

Optimal Equalization for Time-varying Channels with Non-stationary Noises

4.1 Introduction

While the previous chapter has focused on the single user filterbank transceivers in Figure 2.3, this chapter considers multiuser data networks, including DSL, FDMA, and CDMA communication systems. We will consider the case where the estimated channel is available, which is time-varying, due to the mobile and ubiquitous nature of cellular phones, and applications of the filterbank transceivers to wireless data communications. We will study design of optimal receivers that achieve not only PR or zero-forcing, but also MMSE for possible non-stationary white noise processes, among all possible linear and time-varying filters of arbitrary orders. This is contrast to the existing results in the present literature. We will show that for redundant filterbank transceivers, the zero-forcing receivers are equivalent to causal and stable left inverses of some "tall" matrix of linear dynamic systems, consisting of the channel and transmitters. Such left inverses are not unique. Any zero-forcing receiver filterbank accomplishes the goal of channel equaliza-

tion, and eliminates completely ISI. We will first parameterize all zero-forcing channel equalizers in an affine form, and then seek one of them to minimize the MSE caused by possible non-stationary white noises corrupted at the receivers, thereby converting the constrained optimization into unconstrained optimization for receivers design. It will be shown that the design of optimal channel equalizers is equivalent to the design of optimal state estimators for some augmented system subject to the same noise processes. Hence the celebrated Kalman filtering can again be used successfully to design the optimal detectors among all channel equalizers, which are truly optimal over all linear and time-varying receiver filters of arbitrary order. This chapter is also a continuation of our previous work in Chapter 3 on optimal channel equalization for filterbank models developed in [25, 57], in which channels are time-invariant, and the noise process is white and WSS. However the transfer function approach used in the previous chapter does not apply to time-varying channels or/and non-stationary white noise processes. Hence the solution method and the proofs of our main results in this chapter are new, which enable us to obtain more general results on optimal channel equalization under more practical constraints.

This chapter is organized as follows. We begin with the introduction section. Then in Section 4.2, all receiver filterbanks achieving PR or zero-forcing will be parameterized for time-varying channels in an affine form, which generalizes the corresponding result in the previous chapter. In Section 4.3, optimal channel equalization will be studied for time-varying channels and non-stationary white noise processes. It will be shown that the optimal channel equalizer is a modified Kalman filter, similar to the previous chapter, but the algebraic Riccati equation (ARE) in the previous chapter is replaced by the difference

Riccati equation (DRE), assuming that the covariance of the non-stationary white noise process is known for each time instant. Section 4.4 contains analysis of the computational complexity, and BER, associated with optimal channel equalization. An example is used to illustrate the proposed optimal channel equalizer. The notations in this chapter are standard, consistent with the previous chapters. New notations will be made clear as we proceed.

4.2 Parameterization of All Zero-Forcing Receiving Filters

In this section we focus on the multirate transmultiplexer in Figure 2.11, which is a multiuser data communication system. Thus the input and output relation (2.9) holds. We investigate the issue of channel equalization: Given $\underline{\phi}(n,k) = \underline{h}(n,k) \star \underline{f}(n,k)$, under what condition there exists a stable and causal dynamic system, represented by its impulse response $\underline{g}(n,k)$, such that (2.60) holds? If the existence condition is true, can we parameterize all channel equalizers? For this reason we assume through out this section that the noise process is absent, i.e., $v(n) \equiv 0$. The optimal channel equalization in the MMSE sense will be investigated in the next section.

We will assume that the composite system with impulse response $\underline{\phi}(n,k)$ is described by the following state-space model:

$$\begin{aligned} x(n+1) &= A(n)x(n) + B(n)\mathbf{s}(n), \quad x(0) = 0, \\ \underline{y}(n) &= C(n)x(n) + D(n)\mathbf{s}(n), \end{aligned} \tag{4.1}$$

where $\mathbf{s}(n)$ is the vector of M input data stream at time n, $\underline{y}(n)$ is the blocked output

as in (2.48) in absence of the noise, and $x(n)$ is the N-dimensional state vector at time n. Thus $A(n)$ has size $N \times N$, $B(n)$ has size $N \times M$, $C(n)$ has size $P \times N$, and $D(n)$ has size $P \times M$ for all integer valued time n. Clearly the impulse response $\underline{\phi}(n,k)$ for the state-space model in (4.1) is a linear map from $\mathbf{s}(n)$ to $\underline{y}(n)$. We are interested in characterizing all causal and stable left inverses of the linear map $\underline{\phi}(n,k)$, which is equivalent to the zero-forcing condition (2.60).

Throughout this chapter, exponential stability as in Definition 1.6 is used in place of stability. For time-invariant systems, exponential stability reduces to $\rho(A) < 1$. Because stability of the state-space system (4.1) depends only on $A(n)$, we will say that $A(n)$ is exponentially stable, if the condition in Definition 1.6 is true.

Remark 4.1 We will assume exponential stability, and moreover

$$\text{rank}\,[\ D(n)\ ] = M \quad \forall\ n. \tag{4.2}$$

If the condition (4.2) is violated for some n, then there always exists a factorization of the form

$$\underline{\phi}(n,k) = \underline{\phi}_0(n,k)\text{diag}(q^{-\ell_0}, q^{-\ell_1}, \cdots, q^{-\ell_{M-1}}),$$

for some positive integers $\ell_0, \ell_1, \cdots, \ell_{M-1}$, such that the state-space realization of $\underline{\phi}_0(n,k)$ has its $D(n)$-term full column rank. Since the diagonal matrix $\text{diag}(q^{-\ell_1}, \cdots, q^{-\ell_M})$ adds only ℓ_k delays in transmission from $s_k(n)$ to $\hat{s}_k(n)$ where $0 \le k < M$ under the zero-forcing condition, it has no adverse effect on MMSE design. Hence the assumption in (4.2) has no loss of generality. ∎

Since $P > M$, $D(n)$ having full column rank implies the existence of $D_\perp(n)$ such that

$D_a(n) = \left[\begin{array}{cc} D(n) & D_\perp(n) \end{array}\right]$ is square and nonsingular. In particular, $D_\perp(n)$ can be chosen from the minimum rank Cholesky factorization

$$D_\perp(n)D_\perp^*(n) = D_{0\perp}(n)D_{0\perp}^*(n) := I_P - D(n)(D^*(n)D(n))^{-1}D^*(n). \tag{4.3}$$

It follows that left inverses of $D(n)$ and $D_\perp(n)$ exist, denoted by $D^+(n)$ and $D_\perp^+(n)$, respectively, and satisfy

$$\left[\begin{array}{c} D^+(n) \\ D_\perp^+(n) \end{array}\right]\left[\begin{array}{cc} D(n) & D_\perp(n) \end{array}\right] = \left[\begin{array}{cc} I_M & 0 \\ 0 & I_{P-M} \end{array}\right]. \tag{4.4}$$

Lemma 4.1 *Let the state-space realization for $\underline{\phi}(n,k)$ be given as in (4.1) where $D(n)$ satisfies the condition (4.2). Then a causal and stable $\underline{g}(n,k)$ exists and satisfies the zero-forcing condition (2.60), if and only if for any $D^+(n)$ and $D_\perp^+(n)$ satisfying (4.4), there exists a state estimation gain $L(n)$ such that*

$$A_L(n) = A(n) - B(n)D^+(n)C(n) + L(n)D_\perp^+(n)C(n) \tag{4.5}$$

is exponentially stable. Such a state estimator gain $L(n)$ will be called stabilizing.

Proof: If $A_L(n)$ is exponentially stable for some $L(n)$, then it is claimed that $\underline{g}(n,k) = \underline{\phi}^+(n,k)$, described by the state-space model:

$$\begin{aligned} \tilde{x}(n+1) &= A_L(n)\tilde{x}(n) + B_L(n)\underline{y}(n), \quad B_L(n) = B(n)D^+(n) - L(n)D_\perp^+(n), \\ \hat{\mathbf{s}}(n) &= -D^+(n)C(n)\tilde{x}(n) + D^+(n)\underline{y}(n), \quad \tilde{x}(0) = 0, \end{aligned} \tag{4.6}$$

is a causal and stable left inverse of $\underline{\phi}(n,k)$, and thus the zero-forcing condition (2.60) holds. Indeed with the state-space models as in (4.1) and (4.6), the composite system

from $\mathbf{s}(n)$ to $\hat{\mathbf{s}}(n)$ is given by

$$\begin{aligned}\begin{bmatrix} x(n+1) \\ \tilde{x}(n+1) \end{bmatrix} &= \begin{bmatrix} A(n) & 0 \\ B_L(n)C(n) & A_L(n) \end{bmatrix} \begin{bmatrix} x(n) \\ \tilde{x}(n) \end{bmatrix} + \begin{bmatrix} B(n) \\ B(n) \end{bmatrix} \mathbf{s}(n) \\ \hat{\mathbf{s}}(n) &= \begin{bmatrix} D^+(n)C(n) & -D^+(n)C(n) \end{bmatrix} \begin{bmatrix} x(n) \\ \tilde{x}(n) \end{bmatrix} + \mathbf{s}(n)\end{aligned}$$

by $B_L(n)D(n) = B(n)$. Applying similarity transformation

$$\begin{bmatrix} x_s(n) \\ \tilde{x}_s(n) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -I_n & I_n \end{bmatrix} \begin{bmatrix} x(n) \\ \tilde{x}(n) \end{bmatrix}$$

yields

$$\begin{aligned}\begin{bmatrix} x_s(n+1) \\ \tilde{x}_s(n+1) \end{bmatrix} &= \begin{bmatrix} A(n) & 0 \\ 0 & A_L(n) \end{bmatrix} \begin{bmatrix} x_s(n) \\ \tilde{x}_s(n) \end{bmatrix} + \begin{bmatrix} B(n) \\ 0 \end{bmatrix} \mathbf{s}(n) \\ \hat{\mathbf{s}}(n) &= \begin{bmatrix} 0 & -D^+(n)C(n) \end{bmatrix} \begin{bmatrix} x_s(n) \\ \tilde{x}_s(n) \end{bmatrix} + \mathbf{s}(n)\end{aligned}$$

Hence by the initial condition of $\tilde{x}(0) = x(0) = 0$, we have that $\tilde{x}_s(0) = 0$ and $\tilde{x}_s(n) = 0 \ \forall\ n$. It follows that $\hat{\mathbf{s}}(n) = \mathbf{s}(n)$. Conversely, assume that a stable and causal $\underline{g}(n,k)$ exits such that the zero-forcing condition (2.60) is true. Since the state-space model (4.1) for $\underline{\phi}(n,k)$ has an equivalent form

$$x(n+1) = A(n)x(n) + B_L(n)\mathbf{d}(n), \quad \underline{y}(n) = C(n)x(n) + \mathbf{d}(n), \quad x(0) = 0, \tag{4.7}$$

for any $L(n)$, and any $D^+(n)$ and $D_\perp^+(n)$ satisfying (4.4), where $\mathbf{d}(n) = D(n)\mathbf{s}(n)$, and $B_L(n)$ is as in (4.6),

$$\underline{y}(n) = \underline{\phi}(n,k) \star \mathbf{s}(n) = \underline{\phi}_1(n,k) \star \mathbf{d}(n)$$

for some $\underline{\phi}_1(n,k)$. The existence of stable and causal $\underline{g}(n,k)$ such that (2.60) holds then implies that

$$\hat{\mathbf{s}}(n) = \underline{g}(n,k) \star \Big[\underline{\phi}(n,k) \star \mathbf{s}(n)\Big] = \Big[\underline{g}(n,k) \star \underline{\phi}_1(n,k)\Big] \star \mathbf{d}(n) = \mathbf{s}(n).$$

Because $\mathbf{d}(n) = D(n)\mathbf{s}(n)$, we have that $\underline{g}(n,k) \star \underline{\phi}_1(n,k) = D^{+}(n)$, a left inverse of $D(n)$, which in turn implies that $\underline{\phi}_1(n,k)$ as described in (4.7) has a stable and causal left inverse. Noticing that $\underline{\phi}_1(n,k)$, described by the state-space model as in (4.7), has an equal number of inputs and outputs, and the direct transmission from $\mathbf{d}(n)$ to $\underline{y}(n)$ is identity, we conclude that it has a unique inverse, given by

$$x(n+1) = A_L(n)x(n) + B_L(n)\underline{y}(n), \quad \mathbf{d}(n) = -C(n)x(n) + \underline{y}(n).$$

Causality and stability of $\underline{g}(n,k)$ then concludes that

$$A_L(n) = A(n) - B_L(n)C(n) = A(n) - B(n)D^{+}(n)C(n) + L(n)D_{\perp}^{+}(n)C(n)$$

is exponentially stable. ∎

For the case of time-invariant models, the equivalent zero-forcing condition in Lemma 4.1 reduces to the strictly minimum phase condition [26]:

$$\text{rank}\left\{\begin{bmatrix} zI_n - A & B \\ C & D \end{bmatrix}\right\} = n + M \;\; \forall \; |z| \geq 1. \tag{4.8}$$

With the existence condition established in Lemma 4.1 for zero-forcing receiving filters, we consider parameterization of all zero-forcing receiving filters. Let $\mathbf{v}(n)$ be of size $(P - M)\ \forall\ n$. We define a time-varying system $\underline{\phi}_{\perp}(n,k)$ by the state-space model as

$$\begin{aligned} x_{\perp}(n+1) &= A(n)x_{\perp}(n) + B_L(n)D_{\perp}(n)\mathbf{v}(n), \quad x_{\perp}(0) = 0, \\ \underline{w}(n) &= C(n)x_{\perp}(n) + D_{\perp}(n)\mathbf{v}(n) \end{aligned} \tag{4.9}$$

where $B_L(n)$ is the same as in (4.6). It is noted that $B(n) = B_L(n)D(n)$ by the expression of $B_L(n)$ in (4.6), and the relation in (4.4). Thus the state-space model for $\underline{\phi}_{\perp}(n,k)$ is

the same as for $\underline{\phi}(n,k)$ except that $D_\perp(n)$ replaces $D(n)$. Hence by Lemma 4.1, a specific left inverse $\underline{\phi}_\perp^+(n,k)$ to $\underline{\phi}_\perp(n,k)$ is described by the state-space model:

$$\begin{aligned} \tilde{x}_\perp(n+1) &= A_L(n)\tilde{x}_\perp(n) + B_L(n)\underline{v}(n), \quad \tilde{x}(0)=0, \\ \mathbf{v}(n) &= -D_\perp^+(n)C(n)\tilde{x}_\perp(n) + D_\perp^+(n)\underline{w}(n). \end{aligned} \tag{4.10}$$

Let $x_a(n) = x(n) + x_\perp(n)$ and $\underline{y}_a(n) = \underline{y}(n) + \underline{w}(n)$ for the state space models in (4.1) and (4.9). Then it can be verified that the augmented system $\underline{\phi}_a(n,k) = \begin{bmatrix} \underline{\phi}(n,k) & \underline{\phi}_\perp(n,k) \end{bmatrix}$ is square, and has the state-space model

$$\begin{aligned} x_a(n+1) &= A(n)x_a(n) + B_L(n)\begin{bmatrix} D(n) & D_\perp(n) \end{bmatrix}\begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}, \quad x_a(0)=0, \\ y_a(n) &= C(n)x_a(n) + \begin{bmatrix} D(n) & D_\perp(n) \end{bmatrix}\begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}. \end{aligned} \tag{4.11}$$

Lemma 4.2 *Consider the state-space model as in (4.1), which satisfies the condition (4.2), and that there exists a state estimation gain $L(n)$ such that $A_L(n)$ as in (4.5) is exponentially stable. Then for $\underline{\phi}_a(n,k)$ as in (4.11), there exists a unique $\underline{\phi}_a^{-1}(n,k) = \begin{bmatrix} \underline{\phi}^+(n,k) \\ \underline{\phi}_\perp^+(n,k) \end{bmatrix}$, which is a causal and stable inverse for the square augmented system $\underline{\phi}_a(n,k)$, and described by the state-space model:*

$$\begin{aligned} \tilde{x}_a(n+1) &= A_L(n)\tilde{x}_a(n) + B_L(n)y_a(n), \quad \tilde{x}_a(0)=0, \\ \begin{bmatrix} \hat{\mathbf{s}}(n) \\ \hat{\mathbf{v}}(n) \end{bmatrix} &= \begin{bmatrix} -D^+(n)C(n) \\ -D_\perp^+(n)C(n) \end{bmatrix}\tilde{x}_a(n) + \begin{bmatrix} D^+(n) \\ D_\perp^+(n) \end{bmatrix} y_a(n). \end{aligned} \tag{4.12}$$

That is, for each $D^+(n)$, $D_\perp^+(n)$ satisfying (4.4), and $L(n)$ stabilizing, there exists a unique causal and stable inverse $\underline{\phi}_a^{-1}(n,k)$ for $\underline{\phi}_a(n,k)$ such that

$$\begin{bmatrix} \underline{\phi}^+(n,k) \\ \underline{\phi}_\perp^+(n,k) \end{bmatrix} \star \begin{bmatrix} \underline{\phi}(n,k) & \underline{\phi}_\perp(n,k) \end{bmatrix} = \begin{bmatrix} \delta(n-k)I_M & 0 \\ 0 & \delta(n-k)I_{P-M} \end{bmatrix}. \tag{4.13}$$

Proof: Recall the relation in (4.4). With the state-space models as in (4.11) and (4.12), the composite system from $\begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}$ to $\begin{bmatrix} \hat{\mathbf{s}}(n) \\ \hat{\mathbf{v}}(n) \end{bmatrix}$ which is the output of $\begin{bmatrix} \underline{\phi}^{+}(n,k) \\ \underline{\phi}_{\perp}^{+}(n,k) \end{bmatrix}$, is given by

$$\begin{bmatrix} x_a(n+1) \\ \tilde{x}_a(n+1) \end{bmatrix} = \begin{bmatrix} A(n) & 0 \\ B_L(n)C(n) & A_L(n) \end{bmatrix} \begin{bmatrix} x_a(n) \\ \tilde{x}_a(n) \end{bmatrix} + \begin{bmatrix} B_L(n)D(n) & B_L(n)D_{\perp}(n) \\ B_L(n)D(n) & B_L(n)D_{\perp}(n) \end{bmatrix} \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix},$$

$$\begin{bmatrix} \hat{\mathbf{s}}(n) \\ \hat{\mathbf{v}}(n) \end{bmatrix} = \begin{bmatrix} D^{+}(n)C(n) & -D^{+}(n)C(n) \\ D_{\perp}^{+}(n)C(n) & -D_{\perp}^{+}(n)C(n) \end{bmatrix} \begin{bmatrix} x_a(n) \\ \tilde{x}_a(n) \end{bmatrix} + \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}.$$

Applying the similarity transformation $\begin{bmatrix} x_{as}(n) \\ \tilde{x}_{as}(n) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -I_n & I_n \end{bmatrix} \begin{bmatrix} x_a(n) \\ \tilde{x}_a(n) \end{bmatrix}$ yields

$$\begin{bmatrix} x_{as}(n+1) \\ \tilde{x}_{as}(n+1) \end{bmatrix} = \begin{bmatrix} A(n) & 0 \\ 0 & A_L(n) \end{bmatrix} \begin{bmatrix} x_{as}(n) \\ \tilde{x}_{as}(n) \end{bmatrix} + \begin{bmatrix} B_L(n)D(n) & B_L(n)D_{\perp}(n) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix},$$

$$\begin{bmatrix} \hat{\mathbf{s}}(n) \\ \hat{\mathbf{v}}(n) \end{bmatrix} = \begin{bmatrix} 0 & -D^{+}(n)C(n) \\ 0 & -D_{\perp}^{+}(n)C(n) \end{bmatrix} \begin{bmatrix} x_{as}(n) \\ \tilde{x}_{as}(n) \end{bmatrix} + \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}.$$

Hence by the initial condition of $\tilde{x}_a(0) = x_a(0) = 0$, we have that $\tilde{x}_{as}(0) = 0$ and $\tilde{x}_{as}(n) = 0 \ \forall\, n$. It follows that $\begin{bmatrix} \hat{\mathbf{s}}(n) \\ \hat{\mathbf{v}}(n) \end{bmatrix} = \begin{bmatrix} \mathbf{s}(n) \\ \mathbf{v}(n) \end{bmatrix}$. ∎

It is interesting to note that (4.13) is the dynamic version to (4.4), and $\underline{\phi}^{+}(n,k)$ as in (4.6) involves state estimation gain $L(n)$ which is stabilizing. However the left inverse system $\underline{\phi}^{+}(n,k)$ does not have the form of state estimator. Thus Kalman filtering can not be applied directly. Moreover, the left inverse $D^{+}(n)$ is not unique. Indeed, all left

inverses of $D(n)$ are given by

$$D^+(n) = D_0^+(n) + \Theta(n)D_{0\perp}^+(n), \quad D_0^+(n) := (D^*(n)D(n))^{-1}D^*(n), \tag{4.14}$$

where $\Theta(n)$ of size $M \times (P - M)$ is a free matrix, and $D_{0\perp}$ is as in (4.3). Thus the left inverse $\underline{\phi}^+(n, k)$ as in (4.6) involves two free parameters, which is beyond the Kalman filter. Furthermore not all left inverses of $\underline{\phi}(n, k)$ are in the form of $\underline{\phi}^+(n, k)$ as described in (4.6), which poses the difficulty in design of optimal channel equalizers in the next section. So we will first parameterize all causal and stable left inverses and then search for the optimal one in the next section.

It is noted that with $D_{0\perp}(n)$ defined as in (4.3), $D_{0\perp}^+(n) = D_{0\perp}^*(n)$. For any left inverse $D^+(n)$ as in (4.14), we can choose

$$D_\perp(n) = -D(n)\Theta(n) + D_{0\perp}(n), \quad D_\perp^+(n) = D_{0\perp}^+(n) = D_{0\perp}^*(n), \tag{4.15}$$

so that the identity (4.4) still holds true. The next lemma gives parameterization of all causal and stable left inverses of $\underline{\phi}(n, k)$ in an affine form, which satisfy the zero-forcing condition (2.60).

Theorem 4.1 *Under the same hypotheses as in Lemma 4.2, the set of all causal and stable left inverses of* $\underline{\phi}(n, k)$ *is given by*

$$\mathbf{\Phi}^+ := \left\{\underline{\phi}^+(n, k) + \gamma(n, k) \star \underline{\phi}_\perp^+(n, k), \;\; \gamma(n, k) \text{ is causal and stable}\right\}, \tag{4.16}$$

where $\underline{\phi}^+(n, k)$*, and* $\underline{\phi}_\perp^+(n, k)$ *are described by the state-space models (4.6), and (4.10), respectively.*

Proof: It is noted that any $\underline{\phi}^{\text{inv}}(n,k) \in \mathbf{\Phi}^+$ is causal and stable, by causality and stability of $\underline{\tilde{\phi}}(n,k)$, $\underline{\tilde{\phi}}_\perp(n,k)$, and $\gamma(n,k)$. Employing the identity (4.13) yields $\underline{\phi}^{\text{inv}}(n,k) \star \underline{\phi}(n,k) = \delta(n-k)I_M$. Thus any $\underline{\phi}^{\text{inv}}(n,k) \in \mathbf{\Phi}^+$ is indeed a causal and stable left inverse of $\underline{\phi}(n,k)$. Conversely consider any causal and stable left inverse $\underline{\tilde{\phi}}^+(n,k)$. Then with $\underline{\phi}^{\text{inv}}(n,k) \in \mathbf{\Phi}^+$,

$$\begin{aligned}
\underline{\phi}^{\text{inv}}(n,k) \star \begin{bmatrix} \underline{\phi}(n,k) & \underline{\phi}_\perp(n,k) \end{bmatrix} &= \left[\underline{\phi}^+(n,k) + \gamma(n,k) \star \underline{\phi}_\perp^+(n,k)\right] \star \left[\underline{\phi}(n,k) \quad \underline{\phi}_\perp(n,k)\right] \\
&= \begin{bmatrix} \delta(n-k)I_M & \gamma(n,k) \end{bmatrix}, \\
\underline{\tilde{\phi}}^+(n,k) \star \begin{bmatrix} \underline{\phi}(n,k) & \underline{\phi}_\perp(n,k) \end{bmatrix} &= \begin{bmatrix} \delta(n-k)I_M & \underline{\phi}^+(n,k) \star \underline{\phi}_\perp(n,k) \end{bmatrix}.
\end{aligned}$$

Thus with $\gamma(n,k) = \underline{\phi}^+(n,k) \star \underline{\phi}_\perp(n,k)$ which is causal and stable, $\underline{\tilde{\phi}}^+(n,k)$ is indeed in $\mathbf{\Phi}^+$, by the fact that $\underline{\phi}_a(n,k) = \begin{bmatrix} \underline{\phi}(n,k) & \underline{\phi}_\perp(n,k) \end{bmatrix}$ is square, and has a unique causal and stable inverse. ∎

Theorem 4.1 is the time-varying version for the parameterization of all stable left inverses in [26] associated with the time-invariant systems. It is instrumental to obtaining optimal channel equalization filterbanks transceiver models, which will be developed in the next section.

4.3 MMSE Channel Equalization

In this section, we study design of the receiving filterbank $\underline{g}(n,k)$ such that it not only achieves the zero-forcing condition (2.60), but also minimizes the MSE $J(\underline{g})$, as defined in (2.59). If the distorted channel, and the transmitter filters are all time-invariant, and the contaminating noise is WSS and white, this problem was solved in the previous

chapter, where it is shown that the MMSE receiving filterbank satisfying the zero-forcing condition is a modified Kalman filter, which is optimal among all linear and time-invariant filterbanks. However the transfer function approach used in [26] is not applicable to the time-varying channels/transmitter-filters, or/and non-stationary white noises. Hence a different approach is adopted, keeping in mind that the modified Kalman filter obtained in [26] is the key ingredient.

With the presence of the contaminating noise at the output of the channel, the composite system $\underline{\phi}(n,k)$ as in (4.1) is modified into

$$\begin{aligned} x(n+1) &= A(n)x(n) + B(n)\mathbf{s}(n), \quad x(0) = 0, \\ \underline{y}(n) &= C(n)x(n) + D(n)\mathbf{s}(n) + \underline{v}(n). \end{aligned} \tag{4.17}$$

Our objective is to seek $\underline{g}(n,k)$, among all the zero-forcing receiver filterbanks parameterized in Theorem 4.1, which minimizes the MSE $J(\underline{g})$. We assume that the noise process has zero mean, with known covariance for all n:

$$E[v(n)] = 0, \quad E\left[\underline{v}(n)\underline{v}^*(k)\right] = R_{\underline{v}}(n)\delta(n-k) = \underline{\psi}(n)\underline{\psi}^*(n)\delta(n-k), \tag{4.18}$$

where $\psi(n)$ has a minimum column rank, chosen from the minimum rank Cholesky factor. Under the zero-forcing condition, $\underline{g}(n,k) \in \mathbf{\Phi}^+$, by Theorem 4.1. Consequently the symbol detection error at the output of the receiving filterbank is

$$\underline{e}(n) = \hat{\mathbf{s}}(n) - \mathbf{s}(n) = \underline{g}(n,k) \star \underline{v}(n), \quad \underline{g}(n,k) \in \mathbf{\Phi}^+,$$

in light of (2.9), and Theorem 4.1. Thus the MSE, or the variance of the symbol detection error to be minimized is

$$J(\underline{g} \in \mathbf{\Phi}^+)(n) = \operatorname{trace}\left\{E\left[\left(\underline{g}(n,k) \star \underline{v}(n)\right)\left(\underline{g}(n,k) \star \underline{v}(n)\right)^*\right]\right\} \tag{4.19}$$

$$= \quad \text{trace}\left\{\sum_{k=-\infty}^{n} \underline{g}(n,k)R_{\underline{v}}(k)\underline{g}^*(n,k)\right\}.$$

Theorem 4.2 *Suppose that the hypotheses in Lemma 4.2 hold. Let* $\underline{g}(n,k) = \underline{\phi}^+_{\text{opt}}(n,k) \in \mathbf{\Phi}^+$ *be the optimal receiver filterbank to be designed with* $\mathbf{\Phi}^+$ *parameterized in Theorem 4.1. That is,*

$$\begin{aligned} J(\underline{g} = \underline{\phi}^+_{\text{opt}}) &= \text{trace}\left\{\sum_{k=-\infty}^{n}\left[\underline{\phi}^+_{\text{opt}}(n,k)R_{\underline{v}}(k)\left(\underline{\phi}^+_{\text{opt}}(n,k)\right)^*\right]\right\} \\ &= \inf_{\underline{g}\in\mathbf{\Phi}^+} \text{trace}\left\{\sum_{k=-\infty}^{n}\underline{g}(n,k)R_{\underline{v}}(k)\underline{g}^*(n,k)\right\}. \end{aligned}$$

Then the optimal receiver filterbank $\underline{g}(n,k) = \underline{\phi}^+_{\text{opt}}(n,k) \in \mathbf{\Phi}^+$ *which minimizes the MSE is equivalent to the optimal state estimator for the process*

$$\left[\begin{array}{c} x_K(n+1) \\ \hline y_K(n) \end{array}\right] = \left[\begin{array}{c|c} \underline{A}(n) & \underline{B}(n) \\ \hline \underline{C}(n) & \underline{D}(n) \end{array}\right]\left[\begin{array}{c} x_K(n) \\ \hline \underline{v}(n) \end{array}\right] \tag{4.20}$$

where $\underline{v}(n)$ *is the same as in (4.17) and*

$$\begin{aligned} \underline{A}(n) &= \begin{bmatrix} A_0(n) & 0 \\ -D_0^+(n)C(n) & 0 \end{bmatrix}, \quad \underline{B}(n) = \begin{bmatrix} B(n)D_0^+(n) \\ D_0^+(n) \end{bmatrix}, \\ \underline{C}(n) &= \begin{bmatrix} D_{0\perp}^*C(n) & 0 \end{bmatrix}, \quad \underline{D}(n) = -D_{0\perp}^*(n), \end{aligned} \tag{4.21}$$

and $A_0(n) = A(n) - B(n)D_0^+(n)C(n)$*. The state estimator is schematically illustrated in Figure 4.1.*

Proof: Let $\underline{g}(n,k) \in \mathbf{\Phi}^+$. Then by (4.16),

$$\begin{aligned} \underline{g}(n,k) = \underline{\phi}^+(n,k) + \gamma(n,k)\star\underline{\phi}^+_{\perp}(n,k) &= \begin{bmatrix} \delta(n-k)I_M & \gamma(n,k) \end{bmatrix}\star\phi_a^{-1}(n,k), \\ \underline{\phi}_a^{-1}(n,k) &= \begin{bmatrix} \underline{\phi}^+(n,k) \\ \underline{\phi}^+_{\perp}(n,k) \end{bmatrix}, \end{aligned}$$

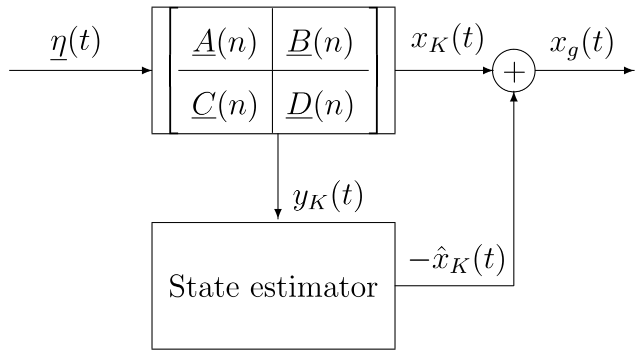


Figure 4.1: State estimator

which has $\underline{y}(n)$ as the input, and $\hat{\mathbf{s}}(n)$ as the output (cf. Figure 2.12). In light of the state-space model for $\underline{\phi}_a^{-1}(n,k)$ as in (4.12), we have that the above $\underline{g}(n,k)$ can be described by the following state-space model (recall that $A_L(n)$ is exponentially stable, by the hypotheses):

$$\begin{aligned} x_g(n+1) &= A_L(n)x_g(n) + B_L(n)\underline{v}(n), \qquad (4.22)\\ \hat{\mathbf{s}}(n) &= -\left[(D_0^+(n) + \chi(n,k) \star D_{0\perp}^+(n))C(n)\right] x_g(n) \\ &\quad + \left[D_0^+(n) + \chi(n,k) \star D_{0\perp}^+(n)\right] \underline{v}(n), \end{aligned}$$

with $\chi(n,k) = \Theta(n) + \gamma(n,k)$, where the expression of $D^+(n)$ as in (4.14) is used. That is, we may extend the left inverse of $D(n)$ to include the dynamics $\gamma(n,k)$ which results in a more general left inverse $D^+(n) = D_0^+(n) + \chi(n,k) \star D_{0\perp}^+(n)$. It follows that the state-space model of $\underline{g}(n,k) \in \mathbf{\Phi}^+$ as in (4.22) has the same form as the state-space model for $\underline{\phi}^+(n,k)$ as in (4.6). Since $J(\underline{g})$ as in (4.19) does not change with $\underline{g}(n,k)$ replaced by $\underline{\tilde{g}}(n,k) = q^{-1} \star \underline{g}(n,k)$, where q^{-1} is the unit delay operator, we have an equivalent minimization problem for $J(\underline{\tilde{g}})$ where $\underline{\tilde{g}}(n,k) = q^{-1} \star \underline{g}(n,k)$ is described by the state-space

model

$$\begin{aligned}\tilde{x}_g(n+1) &:= \begin{bmatrix} x_g(n+1) \\ \hat{\mathbf{s}}(n) \end{bmatrix} \\ &= \begin{bmatrix} A_L(n) & 0 \\ -D^+(n)C(n) & 0 \end{bmatrix} \tilde{x}_g(n) + \begin{bmatrix} B_L(n) \\ D^+(n) \end{bmatrix} \underline{v}(n), \qquad (4.23) \\ \tilde{\mathbf{s}}(n) &= \hat{\mathbf{s}}(n-1) = \underline{C}_1 \tilde{x}_g(n), \qquad \underline{C}_1 = \begin{bmatrix} 0 & I_M \end{bmatrix},\end{aligned}$$

by the state-space model of $\underline{g}(n,k)$ as in (4.22), where $D^+(n) = D_0^+(n) + \chi(n,k) \star D_{0\perp}^+(n)$, and

$$\begin{aligned}A_L(n) &= A(n) - B(n)D^+(n)C(n) + L(n)D_\perp^+(n)C(n) \\ &= A_0(n) + [L(n) - B(n) * \chi(n,k)] * C_0(n), \\ B_L(n) &= B(n)D^+(n) - L(n)D_\perp^+(n) = B(n)D_0^+(n) - [L(n) - B(n) * \chi(n,k)] * D_{0\perp}^*(n).\end{aligned}$$

Therefore with $\underline{A}(n), \underline{B}(n), \underline{C}(n), \underline{D}(n)$ as defined in (4.21), and

$$\underline{L}(n,k) = \begin{bmatrix} L_\chi(n,k) \\ -\chi(n,k) \end{bmatrix}, \quad L_\chi(n,k) = L(n) - B(n) * \chi(n,k), \qquad (4.24)$$

(4.23) is equivalent to

$$\begin{aligned}\tilde{x}_g(n+1) &= [\underline{A}(n) + \underline{L}(n,k) \star \underline{C}(n)]\, \tilde{x}_g(n) + [\underline{B}(n) - \underline{L}(n,k) \star D_{0\perp}^*(n)]\, \underline{v}(n), \qquad (4.25) \\ \tilde{\mathbf{s}}(n) &= C_1 \tilde{x}_g(n).\end{aligned}$$

Now consider the time-varying system as in (4.20). In light of the Kalman filtering theory [4], the optimal state estimate for $x_K(n)$ based on output measurements $y_K(\cdot)$ up to time $(n-1)$ is the conditional mean $\hat{x}_K(n|n-1)$, satisfying

$$E\left[\left\{x_K(n) - \hat{x}_{(n|n-1)}\right\}\left\{x_K(n) - \hat{x}_{(n|n-1)}\right\}^*\right] \geq$$

$$E\left[\left\{x_K(n) - \hat{x}_{K_{(n|n-1)}}\right\}\left\{x_K(n) - \hat{x}_{K_{(n|n-1)}}\right\}^*\right],$$

for any other estimate $\hat{x}(n|n-1)$. Moreover the optimal state estimator has the form

$$\hat{x}_K(n+1|n) = \underline{A}(n)\hat{x}_K(n|n-1) + \underline{K}(n,k) \star [y_K(n) - \underline{C}\hat{x}_K(n|n-1)], \tag{4.26}$$

with $\underline{K}(n,k) = \underline{K}_{\text{opt}}(n)$ the optimal state estimation gain which is non-dynamic, but time-varying. The above yields the error system

$$x_e(n+1) = [\underline{A}(n) + \underline{K}(n,k) \star \underline{C}(n)]\, x_e(n) + [\underline{B}(n) - \underline{K}(n,k) \star D^*_{0\perp}(n)]\, \underline{v}(n) \tag{4.27}$$

with $x_e(n) = x_K(n) - \hat{x}_K(n|n-1)$ which is identical to (4.25), if $x_e(n) = \tilde{x}_g(n)$, and $\underline{K}(n,k) = \underline{L}(n,k)$. Thus the augmented system $\underline{\tilde{g}}(n,k) = q^{-1} \star \underline{g}(n,k)$ has the form of state estimator. Hence minimization of $J(\underline{g})$ over all possible $\underline{g}(n,k) \in \mathbf{\Phi}^+$ is equivalent to optimal state estimator design over all possible state estimator gain $\underline{K}(n,k) = \underline{L}(n,k)$, which can be dynamical. ∎

In light of the celebrated Kalman filtering theory, the optimality of the state estimator is achieved by the static time-varying gain $\underline{K}_{\text{opt}}(n)$, and therefore our MMSE design for optimal channel equalization needs consider only the state estimation gain $\underline{L}(n)$, without searching over the dynamical gain $\underline{L}(n,k)$. However direct use of the Kalman filtering on the augmented system increases the order of the system by M, which can be large. The following demonstrate that we can obtain an nth order receiver filterbank to achieve optimal channel equalization.

Theorem 4.3 *Suppose that the set of all zero-forcing receiving filters $\mathbf{\Phi}^+$ is nonempty. Let the time-varying system $\underline{\phi}(n,k)$ be as in (4.1). Then the blocked time-varying receiver*

filterbank $\underline{g}(n,k) \in \mathbf{\Phi}^+$ achieving optimal channel equalization is described by the state-space model

$$\begin{aligned}\tilde{x}(n+1) &= \left[A_0(n) + L_\chi^{\rm opt}(n)C_0(n)\right]\tilde{x}(n) + \left[B(n)D_0^+(n) - L_\chi^{\rm opt}(n)D_{0\perp}^*(n)\right]\underline{y}(n),\\ \hat{\mathbf{s}}(n) &= -D^+(n)C(n)\tilde{x}(n) + D^+(n)\underline{y}(n), \end{aligned}\tag{4.28}$$

where $D^+(n) = D_0^+(n) + \Theta(n)D_\perp^+(n)$ and $D_\perp^+(n) = D_{0\perp}^*(n)$ with $D_{0\perp}(n)$ and $D_0^+(n)$ as in (4.3) and (4.14), respectively, and

$$A_0(n) = A(n) - B(n)D_0^+(n)C(n), \quad C_0(n) = D_{0\perp}^*(n)C(n). \tag{4.29}$$

The optimal state estimator gain and Θ are given by

$$\begin{aligned}L_\chi(n) &= L_\chi^{\rm opt}(n) := -\left[A_0(n)X(n|n-1)C_0^*(n) + B(n)S_0(n)\right]\tilde{R}^{-1}(n), \qquad (4.30)\\ \Theta(n) &= \Theta^{\rm opt}(n) := \left[S_0(n) - D_0^+(n)C(n)X(n|n-1)C_0^*(n)\right]\tilde{R}^{-1}(n),\end{aligned}$$

where $X(n|n-1)$ is the covariance of $\tilde{x}(n)$, calculated from the following recursive difference Riccati equation (DRE)

$$\begin{aligned}X(n+1|n) &= A_0(n)X(n|n-1)A_0^*(n) + B(n)R_0(n)B^*(n) \qquad (4.31)\\ &- \left[A_0(n)X(n|n-1)C_0^*(n) + B(n)S_0(n)\right]\tilde{R}^{-1}(n)\left[A_0(n)X(n|n-1)C_0^*(n) + B(n)S_0(n)\right]^*\end{aligned}$$

with $X(0|-1) = X_0$, the covariance for the initial value of the state vector $\tilde{x}(0)$, and

$$\tilde{R}(n) = D_{0\perp}^*(n)R_{\underline{v}}(n)D_{0\perp}(n) + C_0(n)X(n)C_0^*(n), \tag{4.32}$$

$$S_0(n) = -D_0^+(n)R_{\underline{v}}(n)D_{0\perp}(n), \quad R_0(n) = D_0^+(n)R_{\underline{v}}(n)(D_0^+(n))^*. \tag{4.33}$$

Proof: By the proof of Theorem 4.2, and [4], $\hat{x}_K(n|n-1) = E\left[x_K(n)|Y_{n-1}\right]$ is the optimal state estimate for the process in (4.20), based on measurements $Y_{n-1} = \{y_K(k)\}_{k=0}^{n-1}$,

which satisfies the following inequality

$$E\left[\left\{x_K(n) - \hat{x}_{(n|n-1)}\right\}\left\{x_K(n) - \hat{x}_{(n|n-1)}\right\}^*\right] \geq$$

$$E\left[\left\{x_K(n) - \hat{x}_{K_{(n|n-1)}}\right\}\left\{x_K(n) - \hat{x}_{K_{(n|n-1)}}\right\}^*\right],$$

for any other state estimate $\hat{x}(n|n-1)$. That is $\hat{\Sigma}(n|n-1) \geq \Sigma(n|n-1)$ with $\hat{\Sigma}(n|n-1)$ the covariance for $\hat{x}(n|n-1)$, and $\Sigma(n|n-1)$ for $\hat{x}_K(n|n-1)$, as proven in [4]. By the equivalence established in Theorem 4.2, and the state-space model (4.25), we have

$$E\left[\{x_K(n) - \hat{x}_K(n|n-1)\}\{x_K(n) - \hat{x}_K(n|n-1)\}^*\right] = E\left[\tilde{x}_g(n)\tilde{x}_g^*(n)\right],$$

with $\tilde{x}_g(n)$ as in (4.25). Denote $\Sigma_{(n|n-1)} = \Sigma(n|n-1)$. Then the error covariance for $\tilde{x}_g(n)$ is given by

$$\begin{aligned} \Sigma_{(n|n-1)} &= E\left[\tilde{x}_g(n)\tilde{x}_g^*(n)\right] = E\left\{\begin{bmatrix} x_g(n) \\ \hat{\mathbf{s}}(n-1) \end{bmatrix}\begin{bmatrix} x_g^*(n) & \hat{\mathbf{s}}^*(n-1) \end{bmatrix}\right\} \qquad (4.34) \\ &= E\begin{bmatrix} x_g(n)x_g^*(n) & x_g(n)\hat{\mathbf{s}}^*(n-1) \\ \hat{\mathbf{s}}(n-1)x_g^*(n) & \hat{\mathbf{s}}(n-1)\hat{\mathbf{s}}^*(n-1) \end{bmatrix} = \begin{bmatrix} \Sigma_{(n|n-1)}^{(1,1)} & \Sigma_{(n|n-1)}^{(12)} \\ \Sigma_{(n|n-1)}^{(21)} & \Sigma_{(n|n-1)}^{(22)} \end{bmatrix} \end{aligned}$$

Applying the Kalman filtering results [4], we have that $\Sigma_{(n|n-1)}$ satisfies the following DRE,

$$\begin{aligned} \Sigma_{n+1|n} &= \underline{A}(n)\Sigma_{(n|n-1)}\underline{A}^*(n) - \left[\underline{A}(n)\Sigma_{(n|n-1)}\underline{C}^*(n) + \underline{B}(n)R_{\underline{v}}(n)D_{0\perp}(n)\right]\underline{\tilde{R}}^{-1}(n) * \\ &* \left[\underline{A}(n)\Sigma_{(n|n-1)}\underline{C}^*(n) + \underline{B}(n)R_{\underline{v}}(n)D_{0\perp}(n)\right]^* + \underline{B}(n)R_{\underline{v}}(n)\underline{B}^*(n), \end{aligned}$$

where $\underline{\tilde{R}}(n)$ as in (4.32), and $R_{\underline{v}}(n)$ as in (4.18). By the expressions in (4.21), the (1,1) position of the above DRE is the same as

$$\begin{aligned} \Sigma_{(n|n-1)}^{(11)} &= A_0(n)\Sigma_{(n|n-1)}^{(11)}A_0^*(n) + B(n)R_0(n)B^*(n) \\ &- \left[A_0(n)\Sigma_{(n|n-1)}^{(11)}C_0^*(n) + B(n)S_0(n)\right]\tilde{R}^{-1}(n)\left[A_0(n)\Sigma_{(n|n-1)}^{(11)}C_0^*(n)B(n)S_0(n)\right]^* \end{aligned}$$

which is identical to (4.31), with $X(n|n-1) = \Sigma_{11}(n|n-1) = \Sigma^{(11)}_{(n|n-1)}$. Because

$$\begin{aligned} \underline{A}(n)\Sigma(n|n-1)\underline{C}^*(n) &= \begin{bmatrix} A_0(n) \\ -D_0^+C(n) \end{bmatrix} X(n|n-1)C^*(n)D_{0\perp}, \\ \underline{C}(n)\Sigma(n|n-1)\underline{C}^*(n) &= D_{0\perp}^*C(n)X(n|n-1)C^*(n)D_{0\perp}, \end{aligned}$$

the optimal state estimation gain formula as in [4] yields the expressions in (4.30). ∎

In light of the various properties of the Kalman filter, the following can be easily deduced.

Corollary 4.1 *The optimal channel equalizer given in Theorem 4.3 is stable in the sense that $A_0(n) + L_\chi^{\rm opt}(n)C_0(n)$ is exponentially stable. In particular, if $\underline{\phi}(n,k)$ converges to a time-invariant system, and the state space model matrices $A(n)$, $B(n)$, $C(n)$, and $D(n)$ converge to A, B, C, and D respectively for which (4.8) holds and D has full column rank, then the DRE as in (4.31) converges to the following ARE*

$$X - A_0XA_0^* - BR_0B^* + [A_0XC_0^* + BS_0]\,\tilde{R}^{-1}\,[A_0XC_0^* + BS_0]^* = 0, \tag{4.35}$$

with X the stabilizing solution, and the optimal state estimator gain and Θ converge to

$$L_\chi = L_\chi^{\rm opt} := -[A_0XC_0^* + BS_0]\,\tilde{R}^{-1}, \qquad \Theta = \Theta^{\rm opt} := \left[S_0 - D_0^+CXC_0^*\right]\tilde{R}^{-1},$$

where $A_L = A_0 + L_\chi C_0$ is exponentially stable.

We summarize this section with the following design algorithm:

- ## Design Algorithm for Optimal Channel Equalizers:

  - Step 1: Find state-space realizations for the blocked time-varying systems: $\underline{h}(n,k)$, and $\underline{f}(n,k)$ $\forall$ n. Then find a state space realization for $\underline{\phi}(n,k)$ in (2.47), which satisfies (4.2), and which are exponentially stabilizable.

- Step 2: Set the state-space model for the optimal channel equalizer as in (4.28), with the *a priori* initial condition $\tilde{x}(0) = \tilde{x}_0$, which has the covariance $X_0 \geq 0$.

- Step 3: For $n = 0, 1, \cdots$, do the following:

  – Compute $\hat{s}(n)$ according to (4.28).

  – Compute DRE (4.31). For $n = 0$, use $X(0|-1) = X_0$. Set $L_\chi(n) = L_\chi^{\text{opt}}(n)$ and $\Theta(n) = \Theta^{\text{opt}}(n)$ according to (4.30).

  – Compute $\tilde{x}(n+1)$ according to (4.28). For $n = 0$, use $\tilde{x}(0) = \tilde{x}_0$.

End.

It is noted that the initial condition $\tilde{x}(0) = \tilde{x}_0$, with covariance matrix X_0, is assumed *a priori*. Roughly speaking, X_0 measures the confidence on the *a priori* estimate $\tilde{x}_0$. If no knowledge on $\tilde{x}_0$ is available, then $\tilde{x}_0$ can be taken as a zero vector, and $X_0 = \rho I_n$ can be taken with ρ sufficiently large. Because of the optimality of the Kalman filter, $\tilde{x}(n)$ converges rather quickly to its steady-state value, and thus only the first a few estimates may have large errors, which is shown for the simulation example in the next section.

4.4 Computational Complexity, BER, and An Illustrative Example

The DRE as in (4.31) has several different forms, and thus a complete analysis for the associated computational complexity can be very lengthy. We will assume that the dimension of the state-space model for the composite system of the transmitters and the channel is $n > 0$, and $P > M$. In this case, the DRE as in (4.31) can be decomposed into

the following three equations:

$$\begin{aligned} X(n+1|n) &= A_0(n)Y(n|n-1)A_0^*(n) + B(n)\left(R_0(n) - S_0(n)\tilde{R}^{-1}(n)S_0^*(n)\right)B^*(n) \\ &\quad - A_0(n)K(n)\left(B(n)S_0(n)\right)^* - \left(A_0(n)K(n)\left(B(n)S_0(n)\right)^*\right) \quad (4.36) \\ K(n) &= X(n|n-1)C_0^*(n)\tilde{R}^{-1}(n) \quad (4.37) \\ Y(n|n-1) &= \left[I_n - K(n)C_0(n)\right]X(n|n-1) \quad (4.38) \end{aligned}$$

It can be verified that the recursive computation of the DRE (4.31) at each time sample n entails

$$(P-M)^2(P+n) + (P-M)(P^2+n^2) + nP(P-M) + \mathcal{O}\left((P-M)^3\right)$$

multiplications for the computation of $\tilde{R}^{-1}(n)$, $2Mn^2 + MP(4P+n-M)$ multiplications for (4.36), $n(P-M)^2 + n^2(P-M)$ multiplications for (4.37), and $n^2(P-M)$ multipilcations for (4.38). It is noted that if the noise covariance matrix $R_{\underline{v}}(n) = \sigma(n)I_p$, i.e., the noise samples are independent of each other, then (4.36) is reduced to

$$X(n+1|n) = A_0(n)Y(n|n-1)A_0^*(n) + B(n)R_0(n)B^*(n) \quad (4.39)$$

for which the number of multiplications is reduced to $2Mn^2 + MP(1+M)$ and the multiplications needed for the computation of $\tilde{R}^{-1}(n)$ are reduced to

$$(P-M)^2(P+n) + (P-M)(P+n^2) + nP(P-M) + \mathcal{O}\left((P-M)^3\right).$$

- **Bit Error Rate Probability**

A commonly used performance measurement is BER, rather than the MSE. To find an expression for the BER at the receiver, we denote $\mathcal{J}(n) = E\left[\underline{e}(n)\underline{e}^*(n)\right]$. Thus there

holds

$$J(\underline{g} = \underline{\phi}^{+}_{\text{opt}})(n) = \text{trace}\left\{\mathcal{J}(n)\right\}\big|_{\underline{g}=\underline{\phi}^{+}_{\text{opt}}}. \tag{4.40}$$

In light of [4], $\mathcal{J}(n)$ for the optimal channel equalizer can be found as

$$\mathcal{J}(n) = D^{+}_{\text{opt}}(n)\left(R_{\underline{v}}(n) + C(n)X(n|n-1)C^{*}(n)\right)\left(D^{+}_{\text{opt}}(n)\right)^{*}, \tag{4.41}$$

where $D^{+}_{\text{opt}}(n) = \left(D^{+}_{0}(n) + \Theta^{\text{opt}}(n)D^{+}_{0\perp}(n)\right)$.

With $\underline{\mathbf{s}}(n)$ as in (2.22), the MSE value $J_i(n)$, the ith diagonal element of $\mathcal{J}(n)$, is given by

$$J_i(n) := E[e_i(n)e_i^{*}(n)] \quad e_i(n) = \underline{g}_i(n,k) \star \underline{v}(n), \tag{4.42}$$

where $\underline{g}_i(n,k)$ is the ith row of the blocked receivers filters $\underline{g}(n,k)$. Thus, the MSE for the ith received data stream $\mathbf{s}_i(n)$ is $J_i(n)$, and the average error probability for the M symbol streams $\mathbf{s}(n)$ is defined as

$$P_e := \frac{1}{M}\sum_{i=0}^{M-1} P_e^{(i)} \tag{4.43}$$

where $P_e^{(i)}$ denotes the error probability of the ith symbol stream. For BPSK constellations, the error probability of the ith symbol in the case of additive Gaussian noise (AGN) is given by

$$P_e^{(i)} = \frac{1}{2}\text{erfc}\left(\frac{1}{\sqrt{J_i(n)}}\right). \tag{4.44}$$

- **An Illustrative Example**

We consider the time-varying channel given by its impulse response

$$\mathbf{h}(n) = \begin{bmatrix} h_0(n) & h_1(n) & \cdots & h_L(n) \end{bmatrix} = \mathbf{h}_{ss} + \delta\mathbf{h}(n),$$

where each of the element of $\delta\mathbf{h}(n)$ is given by α^{-n} with $\alpha = 1.2$, and

$$\mathbf{h}_{ss} = \begin{bmatrix} 1 & -0.3 & 0.5 & -0.4 & 0.1 & -0.02 & 0.3 & -0.1 \end{bmatrix}.$$

The variance for the non-stationary white noise $v(n)$ is given by $\sigma^2(n) = (\sigma_{ss} + \alpha^{-n})^2$ with $\sigma_{ss} = 0.1$. Thus the channel has order $L = 7$, and it converges to a time-invariant channel, and the noise $v(n)$ converges to a WSS white noise process. We choose $P = 8 > M = 3$, and set initially $\underline{f}(n,k) = \delta(n-k)\begin{bmatrix} I_M & 0 \end{bmatrix}^T$, which yields a simple state-space model of $\underline{\phi}(n,k) = \underline{h}(n,k) \star \underline{f}(n,k)$:

$$\underline{\phi}(n,k) = \left[\begin{array}{c|c} A(n) & B(n) \\ \hline C(n) & D(n) \end{array}\right] = \left[\begin{array}{c|c} 0_{M\times M} & I_M \\ \hline H_1(n) & H_0(n) \end{array}\right],$$

with $[A(n), B(n), C(n), D(n)]$ a state-space realization for $\underline{\phi}(n,k)$, and

$$\begin{aligned} H_0(n) &= \begin{bmatrix} h_0(n) & 0 & \cdots & 0 \\ h_1(n) & h_0(n) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h_{P-1}(n) & h_{P-2}(n) & \cdots & h_0(n) \end{bmatrix} \begin{bmatrix} I_M \\ 0 \end{bmatrix}, \\ H_1(n) &= \begin{bmatrix} 0 & h_P(n) & \cdots & h_1(n) \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & h_P(n) \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} I_M \\ 0 \end{bmatrix}, \end{aligned}$$

where $h_i(n) = 0$ is taken if $i > L$. In order to compare with the existing results in the literature, we will consider a different $\underline{\phi}(n,k)$, which is normalized in some sense. For this purpose we treat $\underline{h}(n,k)$, and $\underline{f}(n,k)$ as parameterized time-invariant systems, and compute spectral factorization:

$$\left(\underline{h}(n,k) \star \underline{f}(n,k)\right)^A \star \left(\underline{h}(n,k) \star \underline{f}(n,k)\right) = \Omega^A(n,k) \star \Omega(n,k),$$

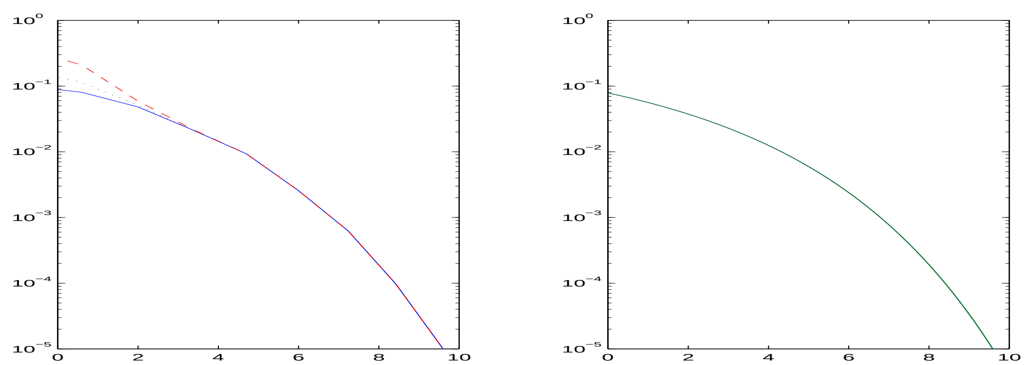
with superscript ^{A} adjoint operation, where $\Omega(n,k)$ having size $M \times M$ is causal and stable, and has a causal and stable inverse. Thus setting $\underline{f}_{\text{new}}(n,k) = \delta(n-k) \begin{bmatrix} I_M & 0 \end{bmatrix}^T \star \Omega^{-1}(n,k)$ shows that $\underline{\phi}_{\text{new}}(n,k) = \underline{h}(n,k) \star \underline{f}_{\text{new}}(n,k)$ is normalized in the sense that

$$\underline{\phi}^A_{\text{new}}(n,k) \star \underline{\phi}_{\text{new}}(n,k) = I_M.$$

Such normalization is necessary for achieving the joint transceiver optimization [57], and also allows comparisons of different designs on the equal basis. The transmitter filters $\{f_m(n,k)\}_{m=0}^2$ can be easily obtained from $\underline{f}_{\text{new}}(n,k)$. Theorem 4.3 can then be used to compute $\underline{g}(n,k)$ from the optimal left inverse of $\underline{\phi}_{\text{new}}(n,k)$.

Following the steps of the design algorithm in the previous section, we obtain a minimal realization $(A(n), B(n), C(n), D(n))$ for $\underline{\phi}_{\text{new}}(n,k)$. Because the system is exponentially stable for each frozen time, exponential stabilizability of $(A(n), B(n))$, and the full rank condition on $D(n)$ are satisfied for this particular example. The optimal receiver filters $\{g_p(n,k)\}_{p=0}^7$ can then be implemented according to Theorem 4.3. To compare with the existing results, we compute the matrix $\mathcal{J}(n)$ as in (4.19) for each time n. The diagonal elements of $\mathcal{J}(n)$ are then used to compute the BER probability at the same time n for the BPSK case as in (4.43) and (4.44). Figure 4.2(a) shows the BER as a function of $E_b/N_0(n)$, where $E_b/N_0(0) = 0.8264$ and $E_b/N_0(\infty) = 100$ with the same initial condition $\tilde{x}_0$, but three different covariance matrices X_0. It can be seen that it converges quickly to its steady-state value, which shows the robustness of the BER performance as time increases. Clearly the BER decreases as the $E_b/N_0(n)$ increases with time n, due to the decaying noise variance to σ_{ss}. Similarly as n increases, the time-varying channel $\mathbf{h}(n)$ converges to the time-invariant channel $\mathbf{h}_{ss}$, and $\sigma(n)$ converges to σ_{ss}, which is the same

as the example studied in [57]. In the steady-state case, the channel will be the same as the example in Chapter 3. BER of the time-invariant case is repeated in Figure 4.2(b) to compare between the time-varying and time-invariant designs. The two figures 4.2(a) and 4.2(b) seem to be similar, but the curves in Figure 4.2(a) have more curvature than the curves in Figure 4.2(b) and they become closer to each other as the time-varying channel converges to the time-invariant one. The deviations at the beginning part are caused by differences in initial condition $\tilde{x}_0$, and the covariance matrix X_0.



(a) BER versus $E_b/N_0(n)$ dB (b) BER versus E_b/N_0 dB

Figure 4.2: Bit error rate sketches for (a) time-varying channels, and (b) time-invariant channels.

Chapter 5

Conclusion

5.1 Introduction

Minimizing the error probability is an ultimate goal in digital communication. One of the techniques that can be used to improve the received signal quality in telecommunications is *equalization*. ISI is a common problem in telecommunication systems and wireless communication systems, such as television broadcasting, digital data communications, and cellular mobile communication systems can be eliminated by *equalization*. ISI has been recognized as the major obstacle to high-speed data transmission and multipath fading over radio channels. In a broad sense, the term equalization can be used to describe any signal processing operation that minimizes the ISI.

Considerable efforts have been devoted to solving the ISI problem (see [75] and references therein). Some of the techniques are post-equalization, multicarrier modulation, and some techniques that use the precoding. Post-equalization techniques include the least-mean-squared (LMS) equalizer, and the decision feedback equalization (DFE), which were studied in [6, 50, 51, 56, 60, 65, 68]. Multicarrier modulation is a technique to cope with the ISI, which aims to increase the transmission system length to eventu-

ally cancel the ISI [14, 52, 77]. Another group of techniques is the use of precoding such as Tomlinson-Harashima (TH) precoding [29, 62], trellis precoding [20, 22], matched spectral null precoding in partial response channels [33], and other precoding schemes [9, 34, 71, 73, 74]. Some of the popular communication applications can be described in terms of filterbanks configurations (transmultiplexers) of subband transforms. Such applications include FDMA,TDMA, and CDMA systems. In particular, FDMA which is also called orthogonal frequency division multiplexing (OFDM) or discrete multitone (DMT) modulation based systems has been more widely used than the others.

Among the recent research on channel equalization, a multirate filterbank was proposed as a precoder before transmission in [75], which introduces the use of redundancy. Another framework for multirate filterbanks was proposed in [25, 57]. This framework is of particular interest, which encompasses existing modulations and equalization schemes.

Some of the applications in digital communications have recently generated significant research activities, such as COFDM, based on which an European T-DAB standard has been defined. DFT-based DMT modulation scheme has become the standard for asymmetric digital subscriber line (ADSL) communications [17, 19, 31, 32]. DWMT [2, 15, 31, 32, 36] is also a common technique in DSL (digital subscriber lines). In addition to these successful applications of multicarrier modulation, several emerging application areas have received great attention as well.

Filterbank transceivers are capable of achieving channel equalization, because of the introduced redundancies, which eliminate completely ISI under some mild conditions. However optimal detection of the transmitted data symbols remains a major issue for

dispersive channels with additive noise, giving the rise of the optimal design problem for the transceiver receivers, which was the main goal of this dissertation.

5.2 Main Results

In this dissertation, we emphasized the design of filterbank transceivers (transmultiplexers) for the purpose of eliminating the ISI completely, and minimizing the MSE of the received signals. The results are applicable to wireless communication systems, due to the same MIMO form of the mathematical models for the equivalent communication channels. The main results of this dissertation are summarized as follows.

In Chapter 3, the new design algorithm is developed for transceiver receivers which achieves optimal channel equalization for linear time-invariant channels and wide-sense stationary (WSS) white noise processes. Transfer function and state-space methods were employed to tackle the optimal design of channel equalizers. A necessary and sufficient condition on PR is established in Theorem 3.1 for the noise-free case. Although joint transceiver optimization is not investigated, a similar condition to [57] can be used for the design of the transmitter filters to ensure good SNR (signal-to-noise ratio). Under the PR condition for the given channel and the transmitter filterbank, all causal and stable receiver filterbanks which achieve PR are parameterized. Under the condition that the noise is white and WSS, the optimal receiver filterbank is obtained in Theorem 3.2 which has the form of state estimators, and is a modified Kalman filter, which minimizes the MSE among all possible linear time-invariant filters. The rich theory of Kalman filtering enables the optimal design for transmitter/receiver filterbanks by employing efficient and reliable

numerical algorithms. Our proposed design algorithm was illustrated by a simulation example, which has a substantially smaller overall MSE than in [57]. The results obtained in Chapter 3 are useful for all time-invariant signal models discussed as in Chapter 2.

We would like to mention that we have assumed that the channel and the transmitter filters are known, and fixed. If the channel is unknown, or not fixed, then blind channel estimation and identification need be employed before applying equalization [8, 13, 25, 30, 46, 59, 61].

In Chapter 4, the new design algorithm is developed for transceiver receivers which achieve optimal channel equalization for time-varying channels and non-stationary white noise processes. We consider the case where the estimated channel is available, and which is time-varying. A necessary and sufficient condition on zero-forcing is established for the noise-free case. Under the zero-forcing condition for the given time-varying channel and the transmitter filterbank, all causal and stable time-varying receiver filterbanks which achieve zero-forcing are parameterized. Under the condition that the noise is non-stationary and white, the optimal time-varying receiver filterbank is obtained in Theorem 4.3, which has the form of state estimators, and is a modified Kalman filter, and which minimizes the MSE among all possible linear time-varying filters of arbitrary order. The rich theory of Kalman filtering enables the optimal design for transmitter/receiver filterbanks by employing efficient and reliable numerical algorithms. Our proposed design algorithm was illustrated by a simulation example. The results obtained in Chapter 4 are useful for all time-varying signal models discussed in Chapter 2.

Often adaptive schemes are employed to estimate the channel, which yields time-varying channel models. If the channel is unknown, then blind channel estimation and identification algorithms as in [8, 13, 25, 30, 59, 61] need be employed before applying the procedure of the optimal channel equalization, as proposed in Chapter 4.

5.3 Open Problems

Although optimal channel equalization for multiuser data networks has been completely solved for linear models under the white noise assumption, the issue of optimal channel equalization against ISI remains, if the corrupting noise is non-white, or the channel model is unknown, which needs be estimated adaptively and blindly. Thus examples of open problems that need further investigation include blind adaptive channel estimation, the negative impact of the incorrect statistical information on blind channel estimation, optimal blind adaptive channel equalization under the non-white noise, channel equalization for non-minimum phase channel model, and optimal detection without the zero-forcing condition. These problems remain unsolved in general, which are itemized as follows.

- Blind adaptive channel estimation, the associated modeling errors, and their negative impacts on optimal channel equalization.

  Blind adaptive channel estimation is a technique in which, the channel characteristics are unknown. A blind adaptive algorithm is employed to estimate the channel which is time-varying, due to the mobile and ubiquitous nature of cellular phones, and the applications to wireless data communications. The modeling error, which measures the error of the estimated channel with respect to the original channel,

may have the form $\|\Phi(z) - \hat{\Phi}(z)\|_\alpha$ with $\Phi(z)$ the true channel model, and $\hat{\Phi}(z)$ the approximate model, needs to be explicitly formulated and minimized in channel estimation. Moreover the norm used in the modeling error is an issue: Which is better for quantifying the modeling error? Convergence analysis is another issue with adaptive estimation algorithm. Although the existing literature contains a large number of research papers on blind channel estimation, the modeling errors, and their negative impacts on optimal channel equalization are overlooked.

- Optimal channel equalization under non-while noise, without the minimum phase condition.

  The identified channel model may not be accurate, due to the existence of the noise, and the use of blind and adaptive channel estimation algorithms. Thus the modeling error introduces non-white noise, in addition to the white noise in received signals. Hence optimal channel equalization becomes more complex, and the results obtained in this dissertation do not apply. It is necessary to develop new optimal channel equalization techniques. Furthermore due to again the existence of noise, the channel model obtained through blind and adaptive algorithms may not be minimum phase at all time. Thus PR condition can not be satisfied. How to minimize the ISI without the minimum phase condition is also an open and meaningful problem.

- Overall channel equalization performance when blind and adaptive channel estimation, and optimal channel equalization are combined to achieve blind channel equalization.

As demonstrated in this dissertation, optimal channel equalization admits a recursive algorithm, which can be implemented in real time. If it is used in conjunction with the blind and adaptive channel estimation algorithm, then blind channel equalization can be claimed. However the existence of the modeling error implies that the overall channel equalization performance may not be the best achievable. How to analyze, and improve the overall performance for blind channel equalization is a very important open problem, and will motivate new criteria for channel estimation, and optimal channel equalization.

5.4 A Concluding Remark

Broadband wireless communications have significantly increased transmission speed, which in turn demanded new technologies for suppression of ISI. This dissertation has contributed to optimal channel equalization for multiuser data networks, and successfully used Kalman filtering in channel equalization, which minimizes the MSE in presence of the white noise. However the research carried out in this dissertation does not stop here. Many new open research problems emerge, as discussed in the previous section. We feel that the approach taken in this dissertation is important. The use of state-space method is not common in communications. Due to the internet, the data networks are multiuser based, and the channels admit multi-input and multi-output (MIMO) for which state-space theory is a natural mathematical tool to approach. Thus state-space is believed to be the right method for many research issues in data communications, in addition to optimal channel equalization.

Bibliography

[1] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Inform. Theory*, vol. 43, pp. 499-511, March 1997.

[2] A.N. Akansu, P. Duhamel, X. Lin and M. Courville, "Orthogonal transmultiplexers in communication: A review," *IEEE Trans. Signal Processing*, vol. 46, pp. 979-995, Apr. 1998.

[3] A. Akansu, M. Tazebay, and R. Haddad, "A new look at digital orthogonal transmultiplexers for CDMA communications," *IEEE Trans. Signal Processing,* pp. 263-267, Jan. 1997.

[4] B.D. Anderson and J.B. Moore, *Optimal Filtering*, Prentice Hall, New Jersey, 1979.

[5] E.F. Badran, G. Gu, "Optimal channel equalization for timevarying channels with nonstationary noises" www.ee.lsu.edu/ggu/new_pubs.html. Submitted to *IEEE Trans. Signal Processing*, Dec. 2001.

[6] C. Belfiore and J. Park, "Decision feedback equalization," *Proc. IEEE*, vol. 67, pp. 1143-1156, Aug. 1979.

[7] M. Bellanger and J. Daguet, "TDM-FDM transmultiplexer: Digital polyphase and FFT", *IEEE Trans. Commun.*, vol. COMM-22, pp.1199-1205, Sept. 1974.

[8] V. Buchoux, C.E. Moulines and A. Grokhov, "On the performance of semi-blind subspace-based channel estimation," *IEEE Trans. Signal Processing*, vol. 48 pp. 1750-1759, June 2000.

[9] A. Calderbank, C. Heegard, and T. Lee, "Binary convolutional codes with application to magnetic recording," *IEEE Trans. Inform. Theory,* vol. IT-32, pp. 797-815, Nov. 1986.

[10] R. Chang, "High-speed multichannel data transmission with bandlimited orthogonal signals," *Bell Syst. Tech. I.*, vol. 45, pp. 1775-1796,Dec. 1966.

[11] B.S. Chen, C.L. Tsai, and Y.F. Chen, "Mixed H_2/H_∞ filtering design in multirate transmultiplexer systems: LMI approach," *IEEE Trans. Signal Processing*, vol. 49 pp. 2693-2701, November 2001.

[12] B.S. Chen, C.W. Lin, and Y.L. Chen, "Optimal signal reconstruction in noisy filter systems: Multirate Kalman synthesis filtering approach," *IEEE Trans. Signal Processing*, vol. 43 pp. 2496-2505, November 1995.

[13] A. Chevreuil, E. Serpedin, P. Loubaton, and G.B. Giannakis, "Blind channel identification and equalization using non-redundant periodic modulation precoders: performance analysis," *IEEE Trans. Signal Processing*, vol. 48, No. 6, pp. 1570-1586, June 2000.

[14] J. Chow, J. Cioffi, and J. Bingham, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 895-908, Aug. 1991.

[15] P.S. Chow, J.C. Tu, and J.M. Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and VHDSL services," *IEEE J. Select. Areas Commun.*, vol. 9, no. 6, pp. 909-919, Aug. 1991.

[16] C.K. Chui and G. Chen, *Discrete $\mathcal{H}_\infty$ Optimization*, Springer, 1997.

[17] I. Cioffi, "A multicarrier primer," Amati Commun. Corp.; Stanford Univ ., Tutorial.

[18] C-T. Chen, *Linear System Theory and Design*, 3rd Ed., Oxford University Press, New York, 1999

[19] Draft Amer. Nat. Std. Telecomm., ANSI; "Network and Customer Installation Interfaces. Asymmetric Digital Subscriber Line (ADSL) Metallic Interface", TIE1.4 (95-007R2).

[20] M. Eyuboglu and G. Fomey, Jr., "Trellis precoding: Combined coding, precoding and shaping for intersymbol interference channels," *IEEE Trans. Inform. Theory*, vol. 38, pp. 301-314, Mar. 1992.

[21] T. Farrell, "The detection and extraction of features of low probability of intercept signals using quadrature mirror filter bank trees," Ph.D. dissertation, Univ. Kansas, Lawrence, 1996.

[22] G. Fomey, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281-300, Mar. 1992.

[23] G.D. Forney, Jr. and M.V. Eyuboğlu, "Combined equalization and coding using precoding," *IEEE Trans. Commun. Mag.*, pp. 255-264, May 1992.

[24] G.H. Golub, and C.F. Van Loan *Matrix Computations*, Johns Hopkins university press, 1996.

[25] G.B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Letters*, vol. 4, pp. 184-187, June 1997.

[26] G. Gu and E.F. Badran, "Optimal design for channel equalization via filterbank approach," www.ee.lsu.edu/ggu/new_pubs.html. Submitted to *IEEE Trans. Signal Processing*, July 2001.

[27] G. Gu and J. Huang, "Design of QMF banks and nonlinear optimization", *Proceedings of 28th Southeastern Symposium on System theory*(Baton Rouge, Louisiana), pp.88-92, 1996.

[28] S.D. Halford and G.B. Giannakis, "Blind equalization of TDMA wireless channels exploiting guard time induced cyclostationarity," *Proc. 1st IEEE Signal Process. Workshop Wireless Commun.*, pp. 117-120, Paris, France, Apr. 16-18, 1997.

[29] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. CQMM-30, pp. 774-780, Aug. 1972.

[30] M. Honing, U. Madhow, and S. Verdu, "Blind multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944-960, July 1995.

[31] I. Kalet, "The multitone channel," *IEEE Trans. Commun.,* vol. 37, no. 2, pp. 119-224, Feb. 1989.

[32] I. Kalet, "Multitone modulation," in *Subband and wavelet Transforms: Design and Applications,* A.N. Akansu and M.J.T. Smith, Eds. Boston, MA: Kluwer, 1995.

[33] R. Karabed and P. Siegel "Matched spectral-null codes for partial- response channels," *IEEE Trans. Inform. Theory*, vol. 37, pp. 818-855, May 1991.

[34] S. Kasturia, J. Aslanis, and J. Cioffi, "Vector-coding for partial-response channels," *IEEE Trans. Inform. Theory*, vol. 36, pp. 741-762, July 1990.

[35] E. Kogbetliantz, "Solution of linear equations by diagonalization of the coeffcient matrices" *Quart. Appl. Math.*, vol 13, pp. 123-32, 1955.

[36] J.W. Lechleider, "High bit rate digital subscriber lines: A review of HDSL progress," *IEEE J. Select. Areas Commun.,* vol. 9, pp. 769-784, Aug. 1991.

[37] X. Li, H.H. Fan, "Blind channel identification: Subspace tracking method without rank estimation," *IEEE Trans. Signal Processing*, vol. 49, pp. 2372-2382, October 2001.

[38] X. Li, H.H. Fan, "QR factorization based blind channel identification and equalization with second-order statistics," *IEEE Trans. Signal Processing*, vol. 48, pp. 60-69, Jan. 2000.

[39] T.J. Lim, L.K. Rasmussen, and H. Sugimoto, "An asynchronous Multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.,* vol. 16, no. 9, pp. 1711-1722, December 1998.

[40] Y.P. Lin and S.M. Phoong, "ISI-free FIR filterbank transceivers for frequency-selective channels," *IEEE Trans. Signal Processing*, vol. 49, pp. 2648-2658, 2001.

[41] R. Liu and D. Dong, "A fundamental theorem for multi-channel blind equalization," *IEEE Trans. Circuits Syst. I*, vol. 44, pp.472-473, may 1997.

[42] F.T. Luk, " A Triangular Processor Array for Computing the Singular Value Decomposition", *Lin. Alg. Appl.*, vol 77, pp.259-273, 1986.

[43] H. Luo and R.W. Liu, "Blind equalization for MIMO FIR channels based only on second order statistics by use of pre-filters," *Proc. of IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Annapolis, MD, USA, pp. 106-109, 1999.

[44] K.D. Minto, *Design of Reliable Control Systems: Theory and Computations*, Ph.D Dissertation, Dept. Elec. Eng., Univ. Waterloo, Waterloo, Canada, 1985.

[45] M. Moonen, P.V. Dooren, and J. Vandewalle, "An SVD updating algorithm for subspace tracking," *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 1015-1038 , 1992.

[46] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516-525, Feb. 1995.

[47] A. Papoulis, *Probability, random variables, and stochastic processes*, 4th Ed., McGraw Hill, 2001

[48] P.Z. Peebles, *Probability, random variables and random signal principles*, 4th Ed., McGraw Hill, 2001

[49] A. Peled and A. Ruiz, "Frequency domain transmission using reduced computational 9omplexity algorithms," *in Proc. IEEE ICASSP*, 1980, pp. 964-967.

[50] J. G. Proakis, *Digital communications*, McGraw-Hill, 2001

[51] J. Proakis, "Adaptive equalization for TDMA digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 333-341, May 1991.

[52] M. Sablatilsh, "Transmission of all-digital advanced television: State of the art and future directions," *IEEE Trans. Broadcasting*, vol. 40, pp. 102-121, June 1994.

[53] B. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Trans. Commun.*, vol. COMM-15, pp. 805-811, Dec.1967.

[54] S. Sandberg, "A family of wavelet-related sequences as a basis for an LPI/D communications system prototype," *Proc. IEEE MILCOM*, 1993, pp. 537-542.

[55] S.D. Sandberg and M.A. Tzannes, "Overlapped discrete multitone modulation for high speed copper wire communications," *IEEE J. Select. Areas Commun.,* vol. 13, pp. 1571-1585, Dec. 1995.

[56] Y. Sato, "A method of self-recoving equalization for multi-level amplitude modulation," *IEEE Trans. Commun.*, vol. COMM-23, pp. 679-682, June 1975.

[57] A. Scaglione, G.B. Giannakis and S. Barbarossa "Redundant filterbank precoders and equalizers, part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1983-2006, July 1999.

[58] A. Scaglione, G.B. Giannakis and S. Barbarossa "Redundant filterbank precoders and equalizers, part II: Blind Channel Estimation, Synchronization, and Direct Equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 2007-2022, July 1999.

[59] E. Serpedin and G.B. Giannakis, "Blind channel identification and equalization with modulation induced cyclostationarity," *IEEE Trans. Signal Processing*, vol. 46, pp. 1930-1944, July 1998.

[60] D. Slock, "Blind fractionally-spaced equalization based on Cyclostationarity and second-order statistics," *in Proc. ATHOS Work.fhop Syst. Identification Higher Order Statist.*, Sophia Antipolis, France, Sept. 1993.

[61] C. Tepedelenlioglu and G.B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Processing*, Vol. 48, No. 7, pp. 2029-2043, July 2000.

[62] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electron. Lett.*, vol. 7, pp. 138-139, Mar. 1971.

[63] L. Tong and S. Perreau, "Multichannl blind identification: from subspace to maximum likelihood methods," *Proceedings of IEEE*, vol. 86, 1951-1968, Oct. 1998.

[64] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol.45, pp. 137-147, Jan 1997.

[65] M. Tsatsanis and G. Giannakis, "Fractional spacing or channel coding for blind equalization?," *in Proc. Int. Conf. Commun. Contr. Telecom./Signal Processing Multimedia Era*, Rithymna/Crete, Greece, June 26-30, 1995, pp: 451-462, invited paper.

[66] M.K. Tsatsanis and G.B. Giannakis, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, 3044-3055, 1996.

[67] M.K. Tsatsanis, G.B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *IEEE Trans. Signal Processing*, vol. 53, pp. 211-229, 1996.

[68] J. Tugnait, “Blind equalization and estimation of FIR communications channels using fractional sampling,” *IEEE Trans. Commun.*, vol. 44, pp. 324-336, Mar. 1996.

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

[70] X. Wang and H.V. Poor, “Blind equalization and multiuser detection in dispersive CDMA channels,” *IEEE Trans. Commun.*, vol. 46, pp. 91-103, Jan 1998.

[71] L. Wei, “Precoding technique for partial-response channels with applications to HDTV transmission,” *IEEE J. Select. Areas Commun.*, vol. 11, pp. 127-135, Jan. 1993.

[72] S. Weinstein and P. Ebert, “Data transmission by frequency- division multiplexing using the discrete fourier transform,” *IEEE Trans. Commun.*, vol. COMM-19, pp. 628-634, Oct. 1971.

[73] J. Wolf and G. Ungerboeck, “Trellis coding for partial-response channels,” IEEE Trans. Commun., vol. COMM”34, pp. 765-773, Aug. 1986.

[74] G. Womell, “Emerging applications of multirate signal processing and wavelets in digital communications,” *in Proc. IEEE*, vol. 84, pp. 586-603, Apr. 1996.

[75] X. Xia, “New precoding for intersymbol interference cancellation using nonmaximally decimated multirate filterbanks with ideal FIR equalizers,” *IEEE Trans. Signal Processing*, vol. 45, No. 10 pp.2431-2441, Oct. 1997.

[76] K. Zhou, J.C. Doyle, *Essentials of Robust Control*, Prentice Hall, New Jersey, 1998.

[77] W. Zou and Y. Wu, “COFDM: An overview,” *IEEE Trans. Broadcasting*, vol. 40, pp. 1-8, Mar. 1995.

Vita

Ehab Farouk Badran was born on July 10, 1973, in Assiut, Egypt. After graduation from Naser Secondary School (high school) in 1990, with the 3rd ranking on Assiut governorate, he joined the Faculty of Engineering, Assiut University, where he obtained the degree of Bachelor of Science in Electrical Engineering with honors in May 1995. As the top of his class, he was assigned to work as an instructor in the Electrical Engineering Department, Assiut University. He obtained the degree of Master of Science in Electrical Engineering in March 1998 with a thesis titled "Speaker recognition using artificial neural networks". He was promoted in May 1998 to work as an assistant lecturer.

In January 7, 2000, Ehab got a studying vocation from Assiut University to pursue a doctoral degree in Electrical Engineering at Louisiana State University. During his studies for the doctoral degree he obtained another master of science in electrical engineering from Louisiana State University in May 2001. Under the guidance and the support of Professor Guoxiang Gu, he worked in the research area of channel equalization for wireless communications via mutlirate filterbank transceivers. He worked as teaching assistant and research assistant during his graduate study at LSU from Spring 2000 to Spring 2002. He is presently a candidate for the degree of Doctor of Philosophy in Electrical Engineering at Louisiana State University.