## EXACT BLIND CHANAE BSTNOMOR

A THESIS<br>SUBMITEDTOTHEDEPARTMENT OF ELEOTHCAMADD ELECTRONACE ENCHEEMMG AND THE WSTHUTE OF ENGHEERNS AND SCENOES   ROR TNE DECREE Q<br>MA.STER OF SCIEM:

## $3 y$

A. Kema: Ozatow:
Juno 159

# EXACT BLIND CHANNEL ESTIMATOR 

A THESIS<br>SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING AND THE INSTITUTE OF ENGINEERING AND SCIENCES OF BILKENT UNIVERSITY in Partial fulfillment of the requirements FOR THE DEGREE OF MASTER OF SCIENCE

## By

A. Kemal Özdemir

June 1998

TK
5102.5
.093
1998
3042045

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

$\overline{\text { Assist. Prof. Dr. Orhan Arıkan(Supervisor) }}$

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.


Prof. Dr. Enol Sezer

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.


Approved for the Institute of Engineering and Sciences:


Director of Institute of Engineering and Sciences

# ABSTRACT <br> EXACT BLIND CHANNEL ESTIMATOR 

A. Kemal Özdemir<br>M.S. in Electrical and Electronics Engineering Supervisor: Assist. Prof. Dr. Orhan Arıkan<br>June 1998

Recently blind identification of single-input multiple-output (SIMO) FIR channels has received considerable attention. The obtained exact identification approaches place over-restrictive constraints on the channels. In this thesis least set of constraints on the channels are placed and the noise-free blind channel identification problem is solved in two stages: The identification of the uncommon zeros followed by the identification of the common zeros of the channels. The minimum number of samples required to identify the uncommon zeros is specified, and closed form solutions are obtained. Also a binary-tree algorithm is proposed for the computation of the uncommon zeros efficiently. Then the common zeros of the channels are identified by a novel pruning algorithm. Finally a simulation example is presented to illustrate these ideas.

Keywords: Blind channel identification, blind deconvolution, system identification, fractional sampling

## ÖZET

# GİRİşí BİLİNMEYEN İLETİ̧̧̇M KANALLARININ HATASIZ TANINMASI 

A. Kemal Özdemir<br>Elektrik ve Elektronik Mühendisliḡi Bölümü Yüksek Lisans Tez Yöneticisi: Yrd. Doç. Dr. Orhan Arıkan<br>Haziran 1998

Son zamanlarda tek girişli çok çıkı̧̧lı (SIMO) sonlu itmeli (FIR) süzgeçlerin gözü kapalı olarak tanınması sorunu büyük ilgi toplamaktadır. Yakın zamanda yapılan araştırmalar kanallar üzerinde çok fazla kısıt içermektedir. Bu tezde kanallar üzerinde en az sayıda kısıt kabul edilerek gürültüsüz gözü kapalı kanal kestirimi sorunu iki aşamalı olarak çözülmüştür: Çoklu kanalların ortak olmayan sıfırlarının kestirimi ve ortak olan sifirlarımın kestirimi. Ortak olmayan sıfırların kestirimi için gerekli en küçük örnek sayısı belirtilmiş ve kapalı biçimde çözümler elde edilmiştir. Bu sıfırların verimli bir şekilde hesaplanabilmesi için bir ikili ağaç algoritması önerilmiştir. Daha sonra kanalların ortak sıfırları yeni bir budama algoritması ile kestirilmiştir. Son olarak öğretici bir benzetim örneği ile bu düşünceler somutlaştırılmıştır.

Anahtar Kelimeler: Gözü kapalı kanal tanınması, gözü kapalı ters evrişim, sistem tanınması, kesirli örnekleme

## ACKNOWLEDGMENTS

I would like to use this opportunity to express my deep gratitude to my supervisor Assist. Prof. Dr. Orhan Arıkan for his guidance, suggestions and invaluable encouragement throughout the development of this thesis.

I would like to thank Prof. Dr. Erol Sezer and Assoc. Prof. Dr. Ömer Morgül for reading and commenting on the thesis.

I express my special thanks to my family for their constant support, patience and sincere love.

Finally, many thanks to all of my close friends.

## Contents

1 INTRODUCTION ..... 1
2 MODEL OF THE COMMUNICATION SYSTEM ..... 5
3 BLIND IDENTIFICATION OF THE UNCOMMON ZEROS ..... 9
3.1 Characterization of the Minimizers of the Cost Function ..... 12
3.2 Identification of $h_{1}[n], \ldots, h_{M}[n]$ from a Minimizer of the Cost Function ..... 16
3.3 Constrained Minimization of the Cost Function ..... 18
3.3.1 Energy Constraint 1 ..... 19
3.3.2 Energy Constraint 2 ..... 19
3.3.3 Constant Vector Constraint ..... 22
3.4 The Equivalence of the Constrained Solutions ..... 23
3.5 A Binary-Tree Algorithm ..... 24
4 BLIND IDENTIFICATION OF THE COMMON ZEROS ..... 28
5 SIMULATION ..... 34
5.1 The Estimation of the Delay ..... 35
5.2 The Binary-tree Algorithm ..... 36
5.3 The Pruning Algorithm ..... 37
6 CONCLUSIONS ..... 41
APPENDICES ..... 46
A Multi-channel Filter Model ..... 46
B Proof of Lemma 1 ..... 48
C Proof of Lemma 2 ..... 50
D Proof of Lemma 3 ..... 52
E Proof of Lemma 4 ..... 53

## List of Figures

2.1 Model of the communication system. ..... 5
2.2 Multi-channel filter model. ..... 6
3.1 Generation of the error signal $e_{i j}[n]$ associated with the $i^{\text {th }}$ and $j^{\text {th }}$ sub-channels. ..... 10
3.2 The generation of the error signal $e_{i j}[n]$ associated with the $i^{\text {th }}$ and $j^{\text {th }}$ sub-channels. ..... 25
3.3 Estimation of the input sequence $u_{i j}[n]$ ..... 26
3.4 The binary-tree representation of a 4-channel FIR filter. ..... 27
4.1 Identification of the common zeros. ..... 31
5.1 The pole-zero plots of the channels: (a) Channel 1, (b) Channel 2, (c) Channel 3, (d) Channel 4. ..... 39
5.2 The magnitudes of the outputs of the channels: (a) Channel 1,
(b) Channel 2, (c) Channel 3, (d) Channel 4. ..... 39
5.3 The singular values. ..... 40
5.4 The running energy of the error sequences associated with dif-ferent channel estimates (identification of the common zeros).40

To my parents ...

## Chapter 1

## INTRODUCTION

The transmission of high speed data in a communication system is subject to channel distortion which is known as the inter-symbol interference (ISI). This ISI, if left uncompensated, leads to high error rates at the receiver. To combat with the ISI receivers are usually designed with some provisions of channel equalization and / or identification. Traditionally in many applications a known training sequence of sufficiently long duration is transmitted through the channel. Since during the training period both the input and the output of the channel are known at the receiver, it is possible to adjust the parameters of the receiver by using a suitable adaptation algorithm. Once the channel is equalized or identified the receiver switches to the decision directed mode, i.e., the receiver uses its estimate (decision) of the input symbols based on the output of the equalizer. In these communication systems the transmission of the training sequence should be periodically repeated to avoid the propagation of the past decision errors into the future decisions. However this repeated transmission of a training sequence which does not convey any information is
a waste of the precious bandwidth which we might consider to use for some other purposes such as the error control coding.

There are practical situations [1] such as multi-point data networks, mobile communication systems and linear predictive deconvolution where the transmission of a training sequence is impossible, impractical or costly. Equalization or identification schemes based on the identification of the receiver parameters without the benefit of a training sequence are said to be blind, unsupervised or self recovering. In other words the objective of the blind algorithms is to identify the channel or the input to the channel using only the observed channel output and possibly some prior knowledge about the statistics of the input symbols. Beginning with the seminal paper of Sato [2] we may identify two broad family of algorithms in this subject: The symbol-spaced and the fractionally-spaced algorithms.

It is assumed that the input to the communication system is a discrete-time stationary sequence. If the received signal is sampled at the symbol rate ${ }^{1}$, it can be shown that the output of the sampler is also a discrete-time stationary sequence. It is well known that the second order statistics of the stationary received sequence provides information about the magnitude of the channel but not phase information. Therefore the symbol-spaced algorithms use the higher-order statistics (HOS) of the received sequence such as cumulants or polyspectra (the discrete-time fourier transform of cumulants) to retrieve the phase information. HOS-based algorithms can be analyzed in two categories depending on how these higher-order statistics are used in the algorithm.

[^0]The algorithms which implicitly exploit the higher-order statistics of the received signal [2-5] are termed as implicit HOS-based algorithms. They try to minimize some non-linear cost function of the unknown equalizer or channel coefficients, because the minimization of this cost function leads to smaller ISI. Since the algorithm is blind the cost function depends on the received sequence but not on the input sequence. The cost function is usually minimized using a simple stochastic gradient descent type of algorithm. Although these algorithms are easy to implement they require long output sequences to reduce the ISI to reasonably low levels. Also the nonlinear cost functions to be minimized are usually multi-modal, thus these algorithms might stuck at one of the local minima of the cost function [6].

The algorithms which explicitly use the higher-order statistics of the received signal are termed as explicit HOS-based algorithms. These type of algorithms are generally reliable and effective but they are computationally intensive and they require large amount of data. In fact it has been shown that the sample size required to estimate the higher-order statistics of a stochastic process increases almost exponentially with the order of the statistics for a fixed variance and bias [7].

It is known that if a digitally modulated stationary signal is sampled at a fraction of the symbol period, then the resulting sequence becomes cyclostationary. As it will be briefly shown in Chapter 2, such a communication system can be modeled as a single-input multiple-output discrete-time multichannel FIR filter as in Fig. 2.2. Recently Tong et. al. [8] has exploited this result to show that for a certain class of channels it is possible to identify this multi-channel filter upto a scalar multiplicative factor using only the second
order statistics of the received sequence. This type of algorithms which use only the second order statistics of the received signal are called as second order cyclostationary statics (SOCS) based algorithms. These algorithms converge faster than HOS-based algorithms. However they are applicable only under certain restrictions on the channel. If the channel does not comply with these restrictions, these algorithms fail to identify the channel even if the channel observations are free of noise.

In this thesis we present theoretical results on the exact estimation of the channel response based on the noise-free observation of the channel output sequence. Our purpose is to fully characterize what can be done with the least set of assumptions on the channel model. Also, discussions on how to extend the obtained results to the more realistic case of blind channel estimation based on noisy channel outputs are provided.

The organization of the thesis is as follows: First the multi-channel filter model is introduced. In Chapter 3, formulation of the BCI problem for the noise-free case is presented. Then, an important result of this thesis is given in the form of a theorem which states the minimum number of required channel observations to identify the uncommon zeros of the sub-channels $h_{1}[n], \ldots, h_{M}[n]$ shown in Fig. 2.2. After providing closed form estimators for these sub-channels under different constraints and proving their equivalence in the absence of noise disturbance an algorithm is proposed for the efficient computation of the sub-channels. In Chapter 4 the estimation of $h_{c}[n]$, the common part of the channels is investigated and the estimation of the input sequence $a[n]$ is discussed. In Chapter 5 an instructive simulation is given and then the thesis is concluded.

## Chapter 2

## MODEL OF THE COMMUNICATION SYSTEM

The block-diagram model of a serial transmission system is shown in Fig. 2.1. In this figure $\{a[n]\}_{n=0}^{\infty}$ is the input symbol sequence chosen from a finite al-


Figure 2.1: Model of the communication system.
phabet, $y(t)$ is the received signal and $T$ is the symbol duration. The impulse response of the linear channel, $h(t)$, is the cascade of the transmitter filter, transmission path and the receiver filter. We will assume that $h(t)$ is of finite
duration:

$$
\begin{equation*}
h(t)=0 \text { for } t \notin\left[D T,\left(D+L_{h}\right) T\right), \tag{2.1}
\end{equation*}
$$

where $D T$ represents the transmission delay.

Over-sampling of the channel output with a factor of $M^{\prime}$ provides channel diversity. Consequently as shown in Appendix A, the system in Fig. 2.1 can be equivalently represented as a single-input $M^{\prime}$-output discrete-time multichannel FIR filter. In the following chapters, we investigate the exact identification of either all or a subset of these channels. Without loss of generality, assuming that first $M \leq M^{\prime}$ of these channels are to be identified, the corresponding multi-channel model is shown in Fig. 2.2. In this figure the outputs of the multi-channel filter are the samples of the received signal $y(t)$ :

$$
\begin{equation*}
y_{i}[n]=y\left(n T+(i-1) \frac{T}{M^{\prime}}\right) \quad, \quad 1 \leq i \leq M \tag{2.2}
\end{equation*}
$$



Figure 2.2: Multi-channel filter model.
The FIR filter $h_{c}[n]$ in Fig. 2.2 corresponds to the common zeros of the sub-channels. If the sub-channels share $L_{1}$ common zeros, then the length of $h_{c}[n]$ is $L_{1}+1$. If the sub-channels do not share any common zeros then $h_{c}[n]=\delta[n]$.

Since all the common zeros are placed in $h_{c}[n]$, the filters $h_{1}[n], \ldots, h_{M}[n]$ do not share any common zeros.

The only assumption we make about the model in Fig. 2.2 is that there is no channel noise: $v_{i}[n]=0$ for $i=1, \ldots, M$. Specifically, we state below some of the assumptions and/or constraints that are avoided in our work:

1. The knowledge of the exact channel order as in [8-12].
2. The constraints on the length of the channels and the number of the channels as in [13].
3. The assumption that the sub-channels do not share any common zeros (i.e., $h_{c}[n]=\delta[n]$ ) as in almost every second-order statistics based algorithm [9,12-14]. It has been shown that these algorithms lose their robustness when this assumption is not true [15]. This is a severe limitation because there exists classes of multi-path communication systems [16,17] for which this assumption is not valid for any over-sampling factor $M^{\prime}$.
4. The assumption that the number of sub-channels is the same as the over-sampling factor.

In Fig. 2.2 we allow the existence of possibly a non-zero delay $D \geq 0$ as a consequence of (2.1). If at the receiver end the starting time of transmission is known, the delay $D$ can be easily obtained from the observation time of first non-zero channel output. Otherwise, it is impossible to identify the value of $D$. As will be detailed, even if $D$ is not known, the exact identification of the channels can be carried out.

In this thesis we propose a two stage procedure for the blind-channelidentification (BCI). In the first stage channels $h_{1}[n], \ldots, h_{M}[n]$ are identified. Once, these channels are identified several of the existing algorithms in the literature $[18]$ can be used to estimate the sequence $x_{c}[n]$ in Fig. 2.2. Then in the second stage, the blind identification of $h_{c}[n]$ is carried out by using the estimated output of this filter $x_{c}[n]$.

## Chapter 3

## BLIND IDENTIFICATION OF THE UNCOMMON ZEROS

In this chapter we set up the framework within which the channels $h_{1}[n], \ldots, h_{M}[n]$ are identified blindly by using outputs $y_{1}[n], \ldots, y_{M}[n]$. As shown in Fig. 2.2, the output signals $y_{1}[n], \ldots, y_{M}[n]$ are the responses of the channels $h_{1}[n], \ldots, h_{M}[n]$ to the same input $x_{c}[n]$. Thus it can be conjectured that the input-output relationship between pairs of channels might produce sufficient information $[12,19]$ to estimate the channels $h_{1}[n], \ldots, h_{M}[n]$ without any prior knowledge about the input $x_{c}[n]$. In this and the next section we will show that indeed this is the case.

For $1 \leq i \leq M$, let $g_{i}[n]$ be an estimate of $h_{i}[n]$. Here we assume that the assumed order $\hat{L}_{2}$ of the channel estimates is larger than or equal to $L_{2}$ which is the largest order of the channels $h_{i}[n]$. Consider two arbitrary and distinct channels, say $i$ and $j$, in Fig. 2.2. After filtering the outputs of these channels
with FIR filters $g_{j}[n]$ and $g_{i}[n]$ we generate the error signal $e_{i j}[n]$ as shown in Fig. 3.1.


Figure 3.1: Generation of the error signal $e_{i j}[n]$ associated with the $i^{\text {th }}$ and $j^{\text {th }}$ sub-channels.

We will base the optimality of a set of channel estimates at the sampling index $N$ of the received data $y_{i}[n], 1 \leq i \leq M$, to the following cost function:

$$
\begin{equation*}
J_{\hat{L}_{2}}(\boldsymbol{g} ; N)=\frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{i-1} J_{i j}\left(\boldsymbol{g}_{i}, \boldsymbol{g}_{j} ; N\right), \tag{3.1}
\end{equation*}
$$

where $J_{i j}\left(\boldsymbol{g}_{\boldsymbol{i}}, \boldsymbol{g}_{j} ; N\right)$, the cost function associated with channels $i$ and $j$, is defined as:

$$
\begin{align*}
J_{i j}\left(\boldsymbol{g}_{i}, \boldsymbol{g}_{j} ; N\right) & =\frac{1}{C_{w, N}} \sum_{k=0}^{N} w_{N-k}\left|e_{i j}[k]\right|^{2}  \tag{3.2}\\
& =\frac{1}{C_{w, N}} \sum_{k=0}^{N} w_{N-k}\left|\boldsymbol{g}_{i}{ }^{T} \boldsymbol{y}_{j}[k]-\boldsymbol{g}_{j}^{T} \boldsymbol{y}_{i}[k]\right|^{2} \tag{3.3}
\end{align*}
$$

where $\boldsymbol{g}_{i}$ and $\boldsymbol{y}_{i}[\boldsymbol{k}]$ are defined as:

$$
\begin{align*}
\boldsymbol{g}_{\boldsymbol{i}} & =\left[\begin{array}{llll}
g_{i}[0] & g_{i}[1] & g_{i}\left[\hat{L}_{2}\right]
\end{array}\right]^{T}  \tag{3.4}\\
\boldsymbol{y}_{i}[\boldsymbol{k}] & =\left[\begin{array}{llll}
y_{i}[k] & y_{i}[k-1] & \cdots & y_{i}\left[k-\hat{L}_{2}\right]
\end{array}\right]^{T} \tag{3.5}
\end{align*}
$$

$C_{w, N}$ in (3.3) is a normalization constant defined as $C_{w, N}=\sum_{k=0}^{N} w_{N-k}$ and $w_{k}$ is a weighting sequence that satisfies

$$
\begin{equation*}
0 \leq w_{k} \leq 1 \quad, \quad 0 \leq k \leq N \tag{3.6}
\end{equation*}
$$

We observe that if the channel estimates in Fig. 3.1 are replaced with true channels, then the cost function in (3.1) becomes identically zero. In other words the true channel coefficients constitute one of the possibly many minimizers of (3.1).

By using (3.4) a more compact representation for the cost $J_{\hat{L}_{2}}(\boldsymbol{g} ; N)$ is obtained as:

$$
\begin{equation*}
J_{\hat{L}_{2}}(\boldsymbol{g} ; N)=\boldsymbol{g}^{\boldsymbol{H}} \boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}[\boldsymbol{N}] \boldsymbol{g} \tag{3.7}
\end{equation*}
$$

where $\boldsymbol{g}$ is the concatenated channel vector estimates

$$
\boldsymbol{g}=\left[\begin{array}{llll}
g_{1}^{T} & g_{2}^{T} & \cdots & g_{M}^{T} \tag{3.8}
\end{array}\right]^{T}
$$

and the hermitian positive semidefinite matrix $\boldsymbol{R}_{y y}[N]$ is

$$
\begin{equation*}
R_{y y}[N]=P[N]-Q[N] \tag{3.9}
\end{equation*}
$$

where

$$
\begin{align*}
P[N]= & \frac{1}{M}\left[\begin{array}{ccc}
R_{a}[N] & 0 & 0 \\
0 & R_{a}[N] & 0 \\
\vdots & \vdots & \\
0 & 0 & R_{a}[N]
\end{array}\right]  \tag{3.10}\\
& R_{a}[N]=\sum_{i=1}^{M} R_{y_{i} y_{i}}[N] \tag{3.11}
\end{align*}
$$

and

$$
Q[N]=\frac{1}{M}\left[\begin{array}{ccc}
R_{y_{1} y_{1}}[N] & R_{y_{2} y_{1}}[N] & R_{y_{M} y_{1}}[N]  \tag{3.12}\\
R_{y_{1} y_{2}}[N] & R_{y_{2} y_{2}}[N] & R_{y_{M} y_{2}}[N] \\
\vdots & \vdots & \vdots \\
R_{y_{1} y_{M}}[N] & R_{y_{2} y_{M}}[N] & R_{y_{M} y_{M}}[N]
\end{array}\right]
$$

$\boldsymbol{R}_{y_{i} y_{j}}[N]$ can be interpreted as the weighted cross-correlation matrix of the multi-channel filter outputs $\boldsymbol{y}_{\boldsymbol{i}}$ and $\boldsymbol{y}_{\boldsymbol{j}}$ :

$$
\begin{equation*}
\boldsymbol{R}_{y_{i} \boldsymbol{y}_{j}}[\boldsymbol{N}]=\frac{1}{C_{w, N}} \sum_{k=0}^{N} w_{N-k} \boldsymbol{y}_{i}^{*}[\boldsymbol{k}] \boldsymbol{y}_{j}^{T}[\boldsymbol{k}] \tag{3.13}
\end{equation*}
$$

In the following section we will characterize the minimizers of (3.7):

$$
\begin{equation*}
\hat{\boldsymbol{g}}=\arg \min _{\boldsymbol{g}} \boldsymbol{g}^{\boldsymbol{H}} \boldsymbol{R}_{\boldsymbol{y y}}[\boldsymbol{N}] \boldsymbol{g} \tag{3.14}
\end{equation*}
$$

### 3.1 Characterization of the Minimizers of the Cost Function

In this section, we will present detailed properties of the optimal solutions to the channel identification problem. We begin with the following powerful theorem which states that the optimal solutions are the FIR filtered versions of the actual channel responses.

## Theorem 1.

$$
\begin{equation*}
J_{\hat{L}_{2}}(\boldsymbol{g} ; N)=0 \Leftrightarrow g_{i}[n]=h_{i}[n] * f[n] \tag{3.15}
\end{equation*}
$$

provided that

$$
\begin{align*}
& \text { 1) } \hat{L}_{2} \geq L_{2}  \tag{3.16}\\
& \text { 2) } N \geq D+L_{2}+\hat{L}_{2} \tag{3.17}
\end{align*}
$$

where $f[n]$ is an arbitrary $F I R$ filter of order at most $\hat{L}_{2}-L_{2}$.

Proof of if part. $g_{i}[n]=h_{i}[n] * f[n]$
$\Rightarrow J_{\hat{L}_{2}}(\boldsymbol{g} ; N)=\sum_{i=1}^{M} \sum_{j=1}^{i-1}\left\|h_{i} * y_{j} * f-h_{j} * y_{i} * f\right\|_{W, N}^{2}$

$$
\begin{aligned}
& =\sum_{i=1}^{M} \sum_{j=1}^{i-1}\left\|\left(h_{i} * y_{j}-h_{j} * y_{i}\right) * f\right\|_{W, N}^{2} \\
& =0
\end{aligned}
$$

where we defined the norm $\|\cdot\|_{W, N}$ as $\|u\|_{W, N}^{2}=1 / C_{w, N} \sum_{k=0}^{N} w_{N-k}|u[k]|^{2}$.

Proof of only if part. $J_{\hat{L}_{2}}(\boldsymbol{g} ; N)=0$ for $N \geq D+L_{2}+\hat{L}_{2} \triangleq D+L_{r}$

$$
\begin{align*}
& \Rightarrow \quad \sum_{i=1}^{M} \sum_{j=1}^{i-1}\left\|g_{i} * y_{j}-g_{j} * y_{i}\right\|_{W, N}^{2}=0 \\
& \Rightarrow \quad\left\|g_{i} * y_{j}-g_{j} * y_{i}\right\|_{W, N}^{2}=0 \quad, \quad \forall i, j \\
& \Rightarrow\left\|\left(g_{i} * h_{j}-g_{j} * h_{i}\right) * x_{c}\right\|_{W, N}^{2}=0 \quad, \quad \forall i, j . \tag{3.18}
\end{align*}
$$

To complete the proof of the theorem we will make use of several lemmas. In order not to break the continuity of the text, the proofs of the lemmas will be left to appendices.

Lemma 1. If $\left\|\left(g_{i} * h_{j}-g_{j} * h_{i}\right) * x_{c}\right\|_{W, N}^{2}=0$ for $N \geq D+L_{r}$ then $g_{i}[n] *$ $h_{j}[n]-g_{j}[n] * h_{i}[n]=0 \quad \forall n$.

This lemma can be used in (3.18) to conclude the equality of $g_{i}[n] * h_{j}[n]=$ $g_{j}[n] * h_{i}[n]$ for all $n$. Hence, in the $z$-transform domain:

$$
\begin{equation*}
G_{i}(z) H_{j}(z)=G_{j}(z) H_{i}(z) \tag{3.19}
\end{equation*}
$$

Since for $i \neq j, H_{i}(z)$ and $H_{j}(z)$ may have common zeros, let monic polynomial $H_{i j}(z)$ be the greatest common divisor of $H_{i}(z)$ and $H_{j}(z)$. Hence, $H_{i}(z)$ can be decomposed as:

$$
\begin{equation*}
H_{i}(z)=H_{i j}(z) Q_{i / j}(z) \tag{3.20}
\end{equation*}
$$

where the quotient polynomial $Q_{i / j}(z)$ is equal to $H_{i}(z) / H_{i j}(z)$. Then, (3.19) can be written as:

$$
\begin{equation*}
G_{i}(z) H_{i j}(z) Q_{j / i}(z)=G_{j}(z) H_{i j}(z) Q_{i / j}(z) \tag{3.2I}
\end{equation*}
$$

After cancellation of the $H_{i j}(z)$ we get;

$$
\begin{equation*}
G_{i}(z) Q_{j / i}(z)=G_{j}(z) Q_{i / j}(z) \tag{3.22}
\end{equation*}
$$

Since, $Q_{j / i}(z)$ and $Q_{i / j}(z)$ have no common zeros;

$$
\begin{equation*}
Q_{i / j}(z) \mid G_{i}(z) \quad, \quad \forall i, j \tag{3.23}
\end{equation*}
$$

At this stage we use the following lemma:
Lemma 2. If $Q_{i / j}(z) \mid G_{i}(z)$ for all $i, j$ then $H_{i}(z) \mid G_{i}(z)$ for all $i$.

Lemma 2 implies that $G_{i}(z)$ can be factored as:

$$
\begin{equation*}
G_{i}(z)=F_{i}(z) H_{i}(z) . \tag{3.24}
\end{equation*}
$$

We complete the proof of the Theorem 1 using the following lemma:
Lemma 3. $F_{i}(z)=F_{j}(z)$ for $1 \leq i, j \leq M$.

This completes the proof of $G_{i}(z)=F(z) H_{i}(z)$ for all $i$.

Theorem 1 states that the minimizers of (3.7) are FIR filtered versions of the actual $h_{i}[n]$ 's where the filter $f[n]$ has length at most $\hat{L}_{2}-L_{2}+1$. In the following theorem, the solution space is further characterized.

Theorem 2. The set of vectors $\boldsymbol{g}=\left[\begin{array}{llll}\boldsymbol{g}_{1}^{\boldsymbol{T}} & \boldsymbol{g}_{2}^{\boldsymbol{T}} & \cdots & \boldsymbol{g}_{M}^{\boldsymbol{T}}\end{array}\right]^{T}$ that satisfies (3.15) constitutes an $\hat{L}_{c}+1$ dimensional vector space, where $\hat{L}_{c}=\hat{L}_{2}-L_{2}$.

The proof of the theorem is based on the following lemma which will be used to construct a basis for the solution space. The proof of the lemma is given in Appendix E.

Lemma 4. Any vector $\boldsymbol{g}$ in the solution space can be expressed as

$$
\begin{equation*}
\boldsymbol{g}=\sum_{k=0}^{\hat{L}_{c}} f\left[\hat{L}_{c}-k\right] \boldsymbol{v}_{k} \tag{3.25}
\end{equation*}
$$

where $\boldsymbol{v}_{\boldsymbol{k}}$ is related to the true channel coefficients as:

$$
v_{k}=J^{k}\left[\begin{array}{lll|l|ll}
\boldsymbol{h}_{1}^{\boldsymbol{T}} & \mathbf{0}^{\boldsymbol{T}} & \mid & \cdots & \boldsymbol{h}_{M}^{T} & \mathbf{0}^{T} \tag{3.26}
\end{array}\right]^{T}
$$

In (3.26) $\boldsymbol{J}$ is a shifting matrix with 1 's on the lower sub-diagonal

$$
\boldsymbol{J}=\left[\begin{array}{ccccc}
0 & 0 & \cdots & 0 & 0  \tag{3.27}\\
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & & & & \vdots \\
0 & 0 & & 1 & 0
\end{array}\right]
$$

and $\mathbf{0}$ is an $\hat{L}_{c}$ dimensional zero vector.

Proof of Theorem 2. It is easy to see that the vector space assertion holds. Indeed if $\left(f^{(1)}, g^{(1)}\right)$ and $\left(f^{(2)}, g^{(2)}\right)$ pairs satisfy the relation (3.2.5) in Lemma 4, then their linear combination $\alpha_{1}\left(f^{(1)}, \boldsymbol{g}^{(1)}\right)+\alpha_{2}\left(f^{(2)}, \boldsymbol{g}^{(2)}\right)$ also satisfies it. The rest of the theorem is proven by using the Lemma 4 to construct a basis for the solution space.

Based on Lemma 4 we conclude that $\left\{\boldsymbol{v}_{0}, \ldots, \boldsymbol{v}_{\hat{\boldsymbol{L}}_{c}}\right\}$ constitutes a spanning set for the solution space. Thus the dimension of the solution space is no larger then $\hat{L}_{c}+1$. We can show that it is exactly $\hat{L}_{c}+1$ by proving that $\boldsymbol{v}_{\boldsymbol{k}}$ 's are linearly independent. Suppose that $\sum_{i=0}^{\dot{L}_{c}} \alpha_{i} \boldsymbol{v}_{i}=0$. Then

$$
\left[\begin{array}{lllll}
\boldsymbol{v}_{0} & \boldsymbol{v}_{1} & \boldsymbol{v}_{2} \cdots & \boldsymbol{v}_{\hat{\boldsymbol{L}}_{c}}
\end{array}\right]\left[\begin{array}{c}
\alpha_{0}  \tag{3.28}\\
\alpha_{1} \\
\alpha_{\hat{L}_{c}}
\end{array}\right]=\mathbf{0}
$$

Without loss of generality it can be assumed that $v_{0}[0]=h_{1}[0]$ is nonzero ${ }^{1}$. Hence, the above system of equations can be written as:

$$
\left[\begin{array}{ccccc}
v_{0}[0] & 0 & 0 & & 0  \tag{3.29}\\
\mathrm{x} & v_{0}[0] & 0 & & 0 \\
\mathrm{x} & \mathrm{x} & v_{0}[0] & & 0 \\
& & \vdots & & \\
\mathrm{x} & \mathrm{x} & \mathrm{x} & \ldots & v_{0}[0] \\
& & & & \vdots \\
\mathrm{x} & \mathrm{x} & \mathrm{x} & & \mathrm{x}
\end{array}\right]\left[\begin{array}{c}
\alpha_{0} \\
\alpha_{1} \\
\vdots \\
\alpha_{\tilde{L}_{c}}
\end{array}\right]=\mathbf{0}
$$

which accepts the unique solution of $\alpha_{0}=\alpha_{1}=\cdots=\alpha_{\hat{L}_{c}}=0$. Thus $\boldsymbol{v}_{i}$ 's are linearly independent.

An important implication of this theorem is stated as:
Corollary 1. The matrix $\boldsymbol{R}_{\boldsymbol{y y}}[\boldsymbol{N}]$ has an $\hat{L}_{c}+1$ dimensional null-space.

In the next section we will show how the exact identification of $h_{1}[n], \ldots, h_{M}[n]$ can be achieved from a given minimizer of (3.7).

### 3.2 Identification of $h_{1}[n], \ldots, h_{M}[n]$ from $\boldsymbol{a}$ Minimizer of the Cost Function

Suppose that using some optimization algorithm we have computed a particular set of channel estimates $\hat{g}_{1}[n], \ldots, \hat{g}_{M}[n]$ that satisfies (3.15) for some nonzero

[^1]$\hat{f}[n]$. In the following by using the fact that $h_{1}[n], \ldots, h_{M}[n]$ do not share any common zeros, we will show how to exactly obtain $h_{1}[n], \ldots, h_{M}[n]$ from $\hat{g}_{1}[n], \ldots, \hat{g}_{M}[n]$.

Let's use the notation $\mathcal{Z}\{\hat{f}\}$ to denote the zeros of an arbitrary FIR filter $\hat{f}$. Then the zeros of $\hat{g}_{i}[n]$ are given as

$$
\begin{equation*}
\mathcal{Z}\left\{\hat{g}_{i}\right\}=\mathcal{Z}\{\hat{f}\} \cup \mathcal{Z}\left\{h_{i}\right\} \tag{3.30}
\end{equation*}
$$

Since the common zeros of $\hat{g}_{i}[n]$ 's are the zeros of the greatest common divisor of $\hat{g}_{i}[n]$ 's, there exist many algorithms that can be used to obtain them [20]. Formally, these common zeros can be obtained as:

$$
\begin{align*}
\cap_{i=1}^{M} \mathcal{Z}\left\{\hat{g}_{i}\right\} & =\cap_{i=1}^{M}\left[\mathcal{Z}\{\hat{f}\} \cup \mathcal{Z}\left\{h_{i}\right\}\right]  \tag{3.31}\\
& =\mathcal{Z}\{\hat{f}\} \cup\left[\cap_{i=1}^{M} \mathcal{Z}\left\{h_{i}\right\}\right]  \tag{3.32}\\
& =\mathcal{Z}\{\hat{f}\} \tag{3.33}
\end{align*}
$$

where the fact that $h_{1}[n], \ldots, h_{M}[n]$ do not share any common zeros is used in the last step. Hence, we showed that the common zeros of $\hat{g}_{i}[n]$ 's are the zeros of $\hat{f}[n]$. Thus, we can compute $\hat{f}[n]$ upto a scaling factor. Finally $h_{i}[n]$ can be found in the $z$-domain by polynomial division:

$$
\begin{equation*}
H_{i}(z)=\hat{G}_{i}(z) / \hat{F}(z) \quad \text { for } i=1, \ldots, M \tag{3.34}
\end{equation*}
$$

Since, $\hat{f}[n]$ is not identically zero, the above polynomial division is well defined.

### 3.3 Constrained Minimization of the Cost Function

As detailed in the previous section, by starting with an order estimate $\hat{L}_{2}$ which is larger than the actual order $L_{2}$ of channels $h_{1}[n], \ldots, h_{M}[n]$, one can obtain a minimizer of the cost function from which the actual channels $h_{1}[n], \ldots, h_{M}[n]$ are computed by first finding the greatest common divisor and then performing a polynomial division (3.34). Alternatively, by starting with $\hat{L}_{2}$ which is larger than $L_{2}$, one can form $\boldsymbol{R}_{y y}[\boldsymbol{N}]$ by using (3.9). Then by using Corollary 1, $L_{2}$ can be obtained as:

$$
\begin{equation*}
L_{2}=\hat{L}_{2}-\eta+1 \tag{3.35}
\end{equation*}
$$

where $\eta$ is the dimension of the null-space of $\boldsymbol{R}_{y y}[N]$. Once the actual order $L_{2}$ is obtained, the minimization problem (3.14) is solved with $\hat{L}_{2}=L_{2}$. Since $\hat{L}_{c}=\hat{L}_{2}-L_{2}=0$, Theorem 1 states that any minimizer of $J_{L_{2}}(\boldsymbol{g} ; N)$ would be in the form

$$
\begin{equation*}
g_{i}[n]=f[0] h_{i}[n] \quad, \quad \text { for } i=1, \ldots, M \tag{3.36}
\end{equation*}
$$

where $f[0]$ is an arbitrary constant. To avoid the undesired trivial solution $f[0]=0$, we have to introduce some constraints into the minimization problem. The constraints should be selected so that they do not alter the global minima of the cost function, but they avoid the above undesired solution. We will investigate the solution to the minimization problem under the following constraints:
(i) $\|\boldsymbol{g}\|^{2}=\mu^{2}$
(ii) $\left\|\boldsymbol{g}_{i}\right\|^{2}=\mu_{i}^{2}$ for some $i$
(iii) $\boldsymbol{c}^{\boldsymbol{H}} \boldsymbol{g}=\alpha$ for some fixed vector $\boldsymbol{c}$.

In the first two, constraints are placed on the energies of the sub-channels. The last one requires the specification of an arbitrary constant $c$. In the following, closed form solutions of these constrained minimization problems are obtained. Then in Section 3.4 we prove that the minimizers of (3.7) under the constraints (i)-(iii) are identical to the true channel within a multiplicative scale factor.

### 3.3.1 Energy Constraint 1

Let the eigen decomposition of $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}$ be given as ${ }^{2}$ :

$$
\boldsymbol{R}_{y y}=\sum_{i=1}^{M\left(L_{2}+1\right)} \lambda_{i} \boldsymbol{q}_{i} \boldsymbol{q}_{\boldsymbol{i}}^{\boldsymbol{H}}, \quad \lambda_{1} \geq \lambda_{2} \geq \cdots \geq \lambda_{M\left(L_{2}+1\right)} \geq 0
$$

It is well known that the minimizer of (3.7) under the total energy constraint is the eigenvector of $\boldsymbol{R}_{\boldsymbol{y y}}$ corresponding to the smallest eigenvalue $\lambda_{M\left(L_{2}+1\right)}$ :

$$
\begin{equation*}
\hat{\boldsymbol{g}}=\mu \boldsymbol{q}_{M\left(L_{2}+1\right)} \tag{3.38}
\end{equation*}
$$

Since $\hat{L}_{c}=0$, the Corollary 1 states that $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}$ has a 1-dimensional null space. Consequently the smallest eigenvalue of $\boldsymbol{R}_{\boldsymbol{y}}$ is 0 and its multiplicity is 1 .

### 3.3.2 Energy Constraint 2

Without loss of generality we will assume that the energy constraint is imposed only on to the first channel. Then the constrained optimization problem can

[^2]be restated as
\[

$$
\begin{array}{ll}
\min _{\boldsymbol{g}} & \boldsymbol{g}^{\boldsymbol{H}} \boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}} \boldsymbol{g} \\
\text { s.t. } & \left\|\boldsymbol{g}_{\mathbf{1}}\right\|^{2}=\mu_{1}^{2} . \tag{3.40}
\end{array}
$$
\]

We partition the matrix $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}$ and the vector $\boldsymbol{g}$ as follows:

$$
\begin{align*}
R_{y y} & =\left[\begin{array}{ll}
R_{11} & R_{12} \\
R_{21} & R_{22}
\end{array}\right]  \tag{3.41}\\
g & =\left[\begin{array}{ll}
g_{1}^{T} & s^{T}
\end{array}\right]^{T} \tag{3.42}
\end{align*}
$$

where $\boldsymbol{R}_{11}=\left(\boldsymbol{R}_{a}-\boldsymbol{R}_{y_{1} y_{1}}\right) / M$. Then the cost function becomes:

$$
\begin{align*}
J_{L_{2}}\left(g_{1}, s\right) & =g^{H} R_{y y} g  \tag{3.43}\\
& =g_{1}^{H} R_{11} g_{1}+g_{1}^{H} R_{12} s+s^{H} R_{21} g_{1}+s^{H} R_{22} s \tag{3.44}
\end{align*}
$$

We form the lagrangian $\mathcal{L}$

$$
\begin{equation*}
\mathcal{L}\left(\boldsymbol{g}_{1}, \boldsymbol{s}, \lambda\right)=J_{L_{2}}\left(\boldsymbol{g}_{1}, \boldsymbol{s}\right)+\lambda\left(\mu_{1}^{2}-\left\|\boldsymbol{g}_{\mathbf{1}}\right\|^{2}\right) \tag{3.45}
\end{equation*}
$$

and set the gradients of $\mathcal{L}$ with respect to $\boldsymbol{g}_{\mathbf{1}}^{*}$ and $\boldsymbol{s}^{*}$ to zero:

$$
\left[\begin{array}{ll}
\boldsymbol{R}_{11} & \boldsymbol{R}_{12}  \tag{3.46}\\
\boldsymbol{R}_{21} & \boldsymbol{R}_{22}
\end{array}\right]\left[\begin{array}{c}
\hat{g}_{1} \\
\hat{s}
\end{array}\right]=\hat{\lambda}\left[\begin{array}{c}
\hat{g}_{1} \\
0
\end{array}\right]
$$

In (3.46) $\boldsymbol{R}_{\mathbf{2 2}}$ cannot be singular, unless $\hat{\boldsymbol{g}}=\left[\begin{array}{lll}\mathbf{0}^{\boldsymbol{T}} & \hat{\boldsymbol{g}}_{2}^{\boldsymbol{T}} & \cdots\end{array} \hat{\boldsymbol{g}}_{\boldsymbol{M}}^{T}\right]^{T} \neq \mathbf{0}$ is a minimizer of (3.7). However this requires $\boldsymbol{h}_{\mathbf{1}}=\mathbf{0}$ or $y_{1}[n]=0 \forall n$. In this case the minimization problem should be solved under the constraint $\left\|\boldsymbol{g}_{\mathbf{1}}\right\|=0$. However the optimal solution to this constraint minimization problem accepts the trivial solution $\boldsymbol{g}=\mathbf{0}$. To avoid this undesired situation, the energy constraint is imposed on a sub-channel which does not have an identically zero
output. Without loss of generality we assume that $\boldsymbol{h}_{\mathbf{1}} \neq \mathbf{0}$. Therefore (3.46) can be equivalently expressed as

$$
\begin{align*}
\left(\boldsymbol{R}_{11}-\boldsymbol{R}_{12} \boldsymbol{R}_{22}^{-1} \boldsymbol{R}_{21}\right) \hat{\boldsymbol{g}}_{1} & =\hat{\lambda} \hat{g}_{1}  \tag{3.47}\\
\hat{\boldsymbol{s}} & =-\boldsymbol{R}_{22}^{-1} \boldsymbol{R}_{21} \hat{\boldsymbol{g}}_{1} \tag{3.48}
\end{align*}
$$

Since the hermitian matrix $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}$ is nonnegative definite, its Schur complement, $\boldsymbol{R}_{\mathbf{1 1}}-\boldsymbol{R}_{\mathbf{1 2}} \boldsymbol{R}_{\mathbf{2 2}}^{\mathbf{- 1}} \boldsymbol{R}_{\mathbf{2 1}}$, is also a nonnegative definite matrix [21]. Also $\hat{\lambda}$ in (3.47) is an eigenvalue of the Schur complement and $\hat{\boldsymbol{g}}_{1}$ is the corresponding eigenvector. The ambiguity of which eigenvalue to choose is resolved once it is verified that the substitution of (3.47) and (3.48) into (3.44) yields the minimum of the cost function as $\hat{\lambda}$. Consequently we should choose $\hat{\lambda}$ as the smallest eigenvalue. Thus, the following procedure provides the optimal solution:

1. Find the smallest eigenvalue and the corresponding eigenvector of the Schur complement

$$
\begin{equation*}
\left(\boldsymbol{R}_{11}-\boldsymbol{R}_{12} \boldsymbol{R}_{22}^{-1} \boldsymbol{R}_{21}\right) \hat{\boldsymbol{g}}_{1}=\lambda_{\min } \hat{\boldsymbol{g}}_{1} \tag{3.49}
\end{equation*}
$$

2. Normalize $\hat{\boldsymbol{g}}_{1}$

$$
\begin{equation*}
\hat{\boldsymbol{g}}_{1}:=\mu_{1} \hat{\boldsymbol{g}}_{1} /\left\|\hat{\boldsymbol{g}}_{1}\right\| \tag{3.50}
\end{equation*}
$$

3. Compute $\hat{\boldsymbol{s}}$

$$
\begin{equation*}
\hat{s}=-R_{22}^{-1} R_{21} \hat{g}_{1} \tag{3.51}
\end{equation*}
$$

Since the dimension of the eigenvalue problem in (3.49) is $M$ times smaller than that of in Section 3.3.1, computationally it is easier to obtain the solution under the second constraint than the first constraint.

### 3.3.3 Constant Vector Constraint

We know that the unconstrained minimizer of (3.7) is in the null-space of $\boldsymbol{R}_{\boldsymbol{y y}}$. If the constant vector $\boldsymbol{c}$ used in this constraint is in the column-space of the hermitian matrix $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}, \boldsymbol{c} \in \mathcal{R}\left(\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}\right)$, then $\boldsymbol{c}^{\boldsymbol{H}} \boldsymbol{g}=0$. Therefore the constraint minimization problem should be solved with $\alpha=0$. However the minimizer of (3.7) under this constraint accepts the trivial solution $\boldsymbol{g}=\mathbf{0}$. To avoid this undesired situation we assume that $c \notin \mathcal{R}\left(\boldsymbol{R}_{\boldsymbol{y y}}\right)$.

We form the lagrangian $\mathcal{L}$ :

$$
\begin{equation*}
\mathcal{L}(\boldsymbol{g}, \lambda)=\boldsymbol{g}^{\boldsymbol{H}} \boldsymbol{R}_{y \boldsymbol{y}} \boldsymbol{g}+\lambda^{*}\left(\alpha-\boldsymbol{c}^{\boldsymbol{H}} \boldsymbol{g}\right)+\lambda\left(\alpha^{*}-\boldsymbol{g}^{\boldsymbol{H}} \boldsymbol{c}\right) \tag{3.52}
\end{equation*}
$$

Setting the gradients of $\mathcal{L}$ with respect to $\boldsymbol{g}^{*}$ and $\lambda^{*}$ to zero we obtain the following set of equations in $\hat{\boldsymbol{g}}$ and $\hat{\lambda}$ :

$$
\begin{align*}
\boldsymbol{R}_{y y} \hat{\boldsymbol{g}} & =\hat{\lambda} \boldsymbol{c}  \tag{3.53}\\
\boldsymbol{c}^{\boldsymbol{H}} \hat{\boldsymbol{g}} & =\alpha \tag{3.54}
\end{align*}
$$

Since $\boldsymbol{c} \notin \mathcal{R}\left(\boldsymbol{R}_{y y}\right),(3.53)$ is inconsistent unless $\hat{\lambda}=0$. If we set $\hat{\lambda}=0$, then (3.53) reads as

$$
\begin{equation*}
\hat{\boldsymbol{g}}=\beta \boldsymbol{n} \tag{3.55}
\end{equation*}
$$

where $\beta$ is an arbitrary complex constant and $\boldsymbol{n}$ is a vector in the $1-D$ nullspace of $\boldsymbol{R}_{y y}$ normalized to have unit norm. Using (3.54) to eliminate $\beta$ we obtain the optimal solution as

$$
\begin{equation*}
\hat{\boldsymbol{g}}=\frac{\alpha}{c^{H} \boldsymbol{n}} \boldsymbol{n} \tag{3.56}
\end{equation*}
$$

### 3.4 The Equivalence of the Constrained Solutions

When solving the constrained optimization problems we assumed that $\hat{L}_{2}=L_{2}$, because we can find $L_{2}$ using the dimension of the null-space of $\boldsymbol{R}_{\boldsymbol{y} \boldsymbol{y}}$. On the other hand we know from Theorem 1 that if $\hat{L}_{2}=L_{2}$, then $J_{L_{2}}(\boldsymbol{g} ; N)=0$ if and only if $\boldsymbol{g}$ is a non-zero multiple of the true channel vector:

$$
\begin{equation*}
\boldsymbol{g}=f[0] \boldsymbol{h} \quad, \quad f[0] \in \mathcal{C} \tag{3.57}
\end{equation*}
$$

Therefore to prove that the minimizer of (3.7) under the constraints in Section 3.3 is a non-zero multiple of the true channel, it is sufficient to show that the constrained solutions can be expressed as in (3.57). The following list provides the appropriate multiplicative constants $f[0]$ for each of the optimal solutions:

1. Under the energy constraint 1 , choose $f[0]=\mu /\|\boldsymbol{h}\|$ so that

$$
\begin{equation*}
\boldsymbol{g}=(\mu /\|\boldsymbol{h}\|) \boldsymbol{h} \tag{3.58}
\end{equation*}
$$

2. Under the energy constraint 2 , choose $f[0]=\mu_{1} /\left\|\boldsymbol{h}_{1}\right\|$ so that

$$
\begin{equation*}
\boldsymbol{g}=\left(\mu_{1} /\left\|\boldsymbol{h}_{1}\right\|\right) \boldsymbol{h} \tag{3.59}
\end{equation*}
$$

3. Under the constant vector constraint choose $f[0]=\alpha / \boldsymbol{c}^{\boldsymbol{H}} \boldsymbol{h}$ so that

$$
\begin{equation*}
\boldsymbol{g}=\left(\alpha / \boldsymbol{c}^{\boldsymbol{H}} \boldsymbol{h}\right) \boldsymbol{h} \tag{3.60}
\end{equation*}
$$

As shown in this section, any of the constraints in Section 3.3 can be used to obtain a non-zero multiple of the true channel when there is no noise in the channel output. However, in the presence of noise, the Theorem 1 does not hold
and the minimization of (3.7) under different constraints in general produces different channel estimates. Currently we are working on the comparison of the channel estimates obtained under the above mentioned constraints when there is noise in the channel outputs.

### 3.5 A Binary-Tree Algorithm

As shown in the previous section, true channel coefficients can be obtained as the minimizers of (3.7) under the constraints in Section 3.3. Unfortunately, direct computation of the obtained closed form solutions are computationally intensive. For $\hat{L}_{2}=L_{2}$, the direct implementation requires $O\left(M^{3} L_{2}^{3}+M^{2} L_{2}^{2} N\right)$ operations for estimation of all the sub-channels $h_{1}[n], \ldots, h_{M}[n]$ using $N$ consecutive channel outputs. In this section we propose to use a binary-tree algorithm to reduce the required computational load to $O\left(M L_{2}^{3}+M L_{2}^{2} N\right)$.

In order to motivate the idea behind the tree-based algorithm, we first investigate the application of the constrained minimization algorithms to only two arbitrary sub-channels. Suppose that we have the setup shown in Fig. 3.1 and we want to minimize the weighted energy of $e_{i j[n]}$ denoted as $J_{i j}\left(\boldsymbol{g}_{\boldsymbol{i}}, \boldsymbol{g}_{\boldsymbol{j}} ; N\right)$ in (3.3) using one of the constrained minimization algorithms of Section 3.3.

Using the notation in Section 3.1 we denote the common zeros between the $i^{\text {th }}$ and $j^{\text {th }}$ channels as $H_{i j}(z)$. Then Fig. 3.1 can be equivalently represented as in Fig. 3.2 where the FIR filters $Q_{i / j}(z) \triangleq H_{i}(z) / H_{i j}(z)$ and $Q_{j / i}(z) \triangleq H_{j}(z) / H_{i j}(z)$ in this latter figure are coprime. Therefore the estimation of the FIR filters in the dashed box in Fig. 3.2 is a special case of the


Figure 3.2: The generation of the error signal $e_{i j}[n]$ associated with the $i^{\text {th }}$ and $j^{\text {th }}$ sub-channels.
$M$-channel blind identification problem with $M=2$. As it is implied by Theorem 1 , we can identify $Q_{i / j}(z)$ and $Q_{j / i}(z)$ using exactly $\hat{L}_{2}+\max \left\{L_{i / j}, L_{j / i}\right\}+1$ output samples where $L_{i / j}$ is the order of the filter $Q_{i / j}(z)$ and $L_{j / i}$ is the order of the filter $Q_{j / i}(z)$. The identification of $Q_{i / j}(z)$ and $Q_{j / i}(z)$, can be carried out by using any of the algorithms in Section 3.3. Since these filters do not share any common zeros, their common input $u_{i j}[n]$ can be recovered exactly by using the corresponding 2 -channel deconvolution filter. This follows from the Bezout identity [22] which guarantees the existence of a pair of FIR deconvolution filters $E_{i / j}(z), E_{j / i}(z)$ of order $\max \left\{L_{i / j}, L_{j / i}\right\}$ such that

$$
\begin{equation*}
E_{i / j}(z) Q_{i / j}(z)+E_{j / i}(z) Q_{j / i}(z)=1 \tag{3.61}
\end{equation*}
$$

The deconvolution filters $E_{i / j}(z)$ and $E_{j / i}(z)$ can be computed by using the following partial fraction expansion:

$$
\begin{equation*}
\frac{1}{Q_{i / j}(z) Q_{j / i}(z)}=\frac{E_{i / j}(z)}{Q_{j / i}(z)}+\frac{E_{j / i}(z)}{Q_{i / j}(z)} . \tag{3.62}
\end{equation*}
$$

In conclusion, once the FIR filters in the dashed box in Fig. 3.2 and the FIR filters $E_{i / j}(z)$ and $E_{j / i}(z)$ in (3.61) are computed the configuration in Fig. 3.3 provides the desired input sequence $u_{i j}[n]$. This forms the basic idea


Figure 3.3: Estimation of the input sequence $u_{i j}[n]$.
of the binary-tree algorithm which will be illustrated for $M=4$ using the re-arranged multi-channel filter model in Fig. 3.4. In this figure the filters in the dashed boxes 1,2 and 3 are coprime (see the argument that leads to Fig. 3.2.). Therefore the above results derived for the 2-channel FIR filters can be used to identify the filters in the dashed boxes 1 and 2 and their inputs $u_{12}[n]$ and $u_{34}[n]$. Since the outputs of the coprime filters $H_{12}(z)$ and $H_{34}(z)$ shown in Fig. 3.4 can be exactly identified, the above results for the 2-channel FIR filters can be used once more to identify these filters together with their common input $x_{c}[n]$. This completes the identification of the sub-channels $h_{1}[n], \ldots, h_{4}[n]$.


Figure 3.4: The binary-tree representation of a 4-channel FIR filter.

## Chapter 4

## BLIND IDENTIFICATION OF THE COMMON ZEROS

In the previous chapter we have presented a method for blind identification of the sub-channels $h_{1}[n], \ldots, h_{M}[n]$ in Fig. 2.2 and their common input $x_{c}[n]$. Thus the overall problem is reduced to the blind identification of the filter $h_{c}[n]$ in Fig. 2.2. In this chapter we give a solution to this problem.

We write the input-output relation of the filter $h_{c}[n]$ using the vector/matrix notation ${ }^{1}$ :

$$
\underbrace{\left[\begin{array}{c}
x_{c}[D]  \tag{4.1}\\
x_{c}[D+1] \\
x_{c}\left[D+L_{1}\right]
\end{array}\right]}_{\boldsymbol{x}_{c, L_{1}}}=\underbrace{\left[\begin{array}{cccc}
a[0] & 0 & 0 \\
a[1] & a[0] & 0 \\
\vdots & \vdots & \\
a\left[L_{1}\right] & a\left[L_{1}-1\right] & \cdots & a[0]
\end{array}\right]}_{\boldsymbol{A}_{\boldsymbol{L}_{\mathbf{1}}}} \cdot \underbrace{\left[\begin{array}{c}
h_{c}[0] \\
h_{c}[1] \\
\vdots \\
h_{c}\left[L_{1}\right]
\end{array}\right]}_{\boldsymbol{h}_{\boldsymbol{c}}}
$$

[^3]In the above equation the vector $\boldsymbol{x}_{\mathbf{c}, \boldsymbol{L}_{1}}$ has been computed as explained in the previous chapter. We also have the a priori information that the information symbols $\left\{a[0], \ldots, a\left[L_{\mathbf{1}}\right]\right\}$ which completely determine $\boldsymbol{A}_{\boldsymbol{L}_{\mathbf{1}}}$ are chosen from a finite alphabet $\mathcal{A}$. If the size of this alphabet is $\mathcal{N}$ then there are only $\mathcal{N}^{L_{1}+1}$ distinct and possible $\boldsymbol{A}_{\boldsymbol{L}_{1}}$ matrices. Actually depending on the shape of the input symbol constellation this number can be somewhat reduced. For example, if the Binary Phase Shift Keying (BPSK) input symbol constellation is used then the alphabet is $\mathcal{A}=\{-1,+1\}$. In this case it is not possible to distinguish the $\left(\boldsymbol{A}_{\boldsymbol{L}_{1}}, \boldsymbol{h}_{\boldsymbol{c}}\right)$ pair from the $\left(-\boldsymbol{A}_{\boldsymbol{L}_{1}},-\boldsymbol{h}_{\boldsymbol{c}}\right)$ pair using only the blind techniques, because both of them satisfies (4.1). However such a constant phase ambiguity is acceptable in digital communications. In practice this phase ambiguity at the receiver can be removed by requiring that the first transmitted information symbol $a[0]$ is always 1 .

The algorithm that we propose in this chapter first computes all possible channel estimates that can lead to the output sequence $x_{c}[D], \ldots, x_{c}\left[D+L_{1}\right]$. For example, if $\boldsymbol{A}_{\boldsymbol{L}_{1}}^{(\boldsymbol{k})}$ is one of the possible $\boldsymbol{A}_{\boldsymbol{L}_{1}}$ matrices then the corresponding channel estimate $\hat{\boldsymbol{h}}_{c}^{(k)}$ is obtained by solving (4.1) using the forwardsubstitution:

$$
\begin{align*}
\hat{h}_{c}^{(k)}[0] & =x_{c}[D] / a^{(k)}[0] \\
\hat{h}_{c}^{(k)}[n] & =\frac{1}{a^{(k)}[0]}\left[x_{c}[n+D]-\sum_{l=1}^{n-1} a^{(k)}[n-l] \hat{h}_{c}^{(k)}[l]\right] \quad \text { for } 1 \leq n \leq L_{1} . \tag{4.2}
\end{align*}
$$

In the noise-free case the sequence $x_{c}[n]$ can be computed exactly. Therefore one of the computed channel estimates, $\hat{\boldsymbol{h}}_{c}^{(k)}$, should be the actual channel. The pruning algorithm given in the next page can be used to identify the actual channel among all those possible channel estimates. Basically at each
sample index $n$, the algorithms discards the channel estimates which cannot produce the most recent output sample $x_{c}[n]$. The algorithm terminates when there remains only one channel estimate. After the correct channel is identified using the pruning algorithm, the configuration in Fig. 4.1 provides a solution for the input sequence $a[n]$.

```
Algorithm 1 The pruning algorithm
Initialization:
Define the current sample index \(n\) and the set of remaining channel estimates \(\mathcal{S}^{n-1}\) at the previous sample index:
\[
\begin{aligned}
n & =D+L_{1}+1 \\
\mathcal{S}^{n-1} & =\left\{\boldsymbol{h}_{c}^{(0)}, \boldsymbol{h}_{c}^{(1)}, \ldots, \boldsymbol{h}_{c}^{\mathcal{N}^{L_{1}+1}}\right\}
\end{aligned}
\]
```


## Pruning loop:

while $\operatorname{size}\left(\mathcal{S}^{(n-1)}\right)>1$ do
Set $\mathcal{S}^{(n)}:=\mathcal{S}^{(n-1)}$.
Compute $x_{c}[n]$ using the results in Chapter 3.
for each channel estimate $\left\{\boldsymbol{h}_{\boldsymbol{c}}^{(k)}\right\} \in \mathcal{S}^{(n-1)}$ do
Compute $o^{(k)}[n]$ and $\hat{a}^{(k)}[n-D]$ as in Fig. 4.1.
if the residual error

$$
\begin{equation*}
\left|e^{(k)}[n]\right| \triangleq\left|o^{(k)}[n]-\hat{a}^{(k)}[n-D]\right| \tag{4.3}
\end{equation*}
$$

exceeds a threshold then
Set $\mathcal{S}^{n}:=\mathcal{S}^{n}-\left\{\boldsymbol{h}_{c}^{(\boldsymbol{k})}\right\}$.
end if
end for
Set $n:=n+1$.
end while

There exist some pathological cases such that the pruning algorithm does not terminate. In other words there may exist two or more input sequence-channel pairs such that they produce the same output $x_{c}[n]$. Actually this would always be the case if there were no constraints on the input sequence. But in the context of digital communications the samples of the input sequence (i.e., the information symbols) are chosen from a finite alphabet $\mathcal{A}$ with $\mathcal{N}$


Figure 4.1: Identification of the common zeros.
elements. Unfortunately even under such a stringent constraint there exists such pathological cases. For example consider the following input sequence and channel pairs:

$$
\begin{align*}
& \left(a_{1}[n], h_{c_{1}}[n]\right)=\left((-1)^{n} u[n], \delta[n]+\delta[n-1]\right)  \tag{4.4}\\
& \left(a_{2}[n], h_{c_{2}}[n]\right)=(u[n], \delta[n]-\delta[n-1]) . \tag{4.5}
\end{align*}
$$

A straightforward calculation shows that $x_{c_{1}}[n]=x_{c_{2}}[n]=\delta[n]$. Thus knowing only the output sequence $x_{c}[n]=\delta[n]$ it is impossible to tell which input sequence has been actually transmitted ${ }^{2}$. So this identifiability problem applies to all blind channel identification algorithms.

The above discussion proves the existence of input sequence-channel pairs which produce the same output. However in practice what is important is the probability of coming across to such pathological cases. Now we show that this probability is zero.

Theorem 3. Suppose that the output sequence $x_{c}[n]$ which is not identically zero is observed for $n \geq 0$. The vectors of filter coefficients

$$
\boldsymbol{h}_{\boldsymbol{c}}=\left[\begin{array}{llll}
h_{c}[0] & h_{c}[1] & \cdots & h_{c}\left[L_{1}\right] \tag{4.6}
\end{array}\right]^{T}
$$

[^4]which produce the output $x_{c}[n]$ for a suitable input sequence a[n] are isolated points in $L_{1}+1$ dimensional space.

Proof. The proof is by contradiction. Let $\epsilon>0$ be given. Then there exists $L_{1}+1$ dimensional vectors $\boldsymbol{h}_{\boldsymbol{c}_{\mathbf{1}}}, \boldsymbol{h}_{\boldsymbol{c}_{\mathbf{2}}}$ such that

$$
\begin{equation*}
0<\left\|\boldsymbol{h}_{c_{1}}-\boldsymbol{h}_{c_{2}}\right\|<\epsilon, \tag{4.7}
\end{equation*}
$$

where the FIR filters $h_{c_{1}}[n], h_{c_{2}}[n]$ associated with these vectors produce the same response $x_{c}[n]$ to the input sequences ${ }^{3} a_{1}[n], a_{2}[n]$ :

$$
\begin{equation*}
X_{c}(z)=z^{-D} A_{1}(z) H_{c_{1}}(z)=z^{-D} A_{2}(z) H_{c_{2}}(z) . \tag{4.8}
\end{equation*}
$$

Let $H_{c_{2}}(z)=H_{c_{1}}(z)+\delta H_{c}(z), A_{2}(z)=A_{1}(z)+\delta A(z)$. Substituting these into (4.8) yields:

$$
\begin{equation*}
H_{c_{1}}(z) \delta A(z)+\delta H_{c}(z) A_{1}(z)+\delta H_{c}(z) \delta A(z)=0 . \tag{4.9}
\end{equation*}
$$

In this equation, $H_{c_{1}}(z)$ is fixed, $\delta A(z)$ is the $z$-transform of a sequence whose samples can take values only the in the set $\{-2,-1,0,1,2\}$, and $\delta H_{c}(z)$ is an arbitrary FIR filter. Let $\epsilon \rightarrow 0$ in (4.7). Then (4.9) reads as

$$
\begin{equation*}
H_{c_{1}}(z) \delta A(z)=0 \tag{4.10}
\end{equation*}
$$

Since $\delta A(z) \neq 0$ (because $H_{c_{1}}(z) \neq H_{c_{2}}(z)$ in (4.8)), the above condition can be satisfied only if $H_{c_{1}}(z)=0$. Then, it follows from (4.8) that $H_{c_{2}}(z)=0$. $A$ contradiction.

As a consequence of this theorem we conclude that the measure of such pathological input sequence-channel pairs is zero among all admissible pairs.

[^5]Therefore in practice the pruning algorithm always identifies $h_{c}[n]$ in a finite number of iterations. As shown in the next chapter when the common zeros are identifiable from $x_{c}[n]$ the pruning algorithms converges to the solution very fast. On the other hand if the identification of the common zeros is desired under all circumstances, then the transmission of a known initialization sequence of length $L_{1}+1$ should be tolerated so that (4.1) can be solved for the common zeros $h_{c}[n]$. Note that unlike a training sequence which is transmitted periodically, the initialization is performed only once.

## Chapter 5

## SIMULATION

In this chapter an illustrative simulation is presented to highlight the order of the operations used in BCI problem. Consider the 4 -channel FIR system ( $M=M^{\prime}=4$ ) with the pole-zero plots shown in Fig. 5.1. Each pair of channels in this figure share exactly 3 common finite zeros, where the two of these zeros are common to all of the 4 channels (i.e, $H_{c}(z)=1+z^{-2}$ ). The delay associated with a particular channel is given as the number of the zeros of that channel at infinity ${ }^{1}$, which can found in Fig. 5.1 as the difference between the number of poles at the origin and the number of finite zeros. For example the delay associated with the first channel is $10-6=4$ samples. Similarly the delays associated with the other channels are 3,2 and 3 samples respectively.

The FIR system with these pole-zero plots can be decomposed into a binary-tree structure as shown in Fig. 3.4. It is known that such a decomposition is always ambiguous upto constant scale factors. For example consider

[^6]multiplying all the filters in the dashed boxes 1 and 2 by a non-zero $\gamma \in \mathcal{C}$ and multiplying all the filters in the dashed box 3 by $1 / \gamma$. To remove this ambiguity we will scale the filters in Fig. 3.4 such that the first non-zero samples of the filters except for those at the leaves (dashed boxes 1 and 2) are always 1.

The first box in Fig. 3.4 is the overall delay $D=2$ of the communication system which is defined as the minimum of the delays of each sub-channel. The filter $H_{c}(z)=1+z^{-2}$ in this figure represents the common finite zeros to the 4 channels, $H_{i j}(z), i \neq j$, denotes the zeros common to the $i^{\text {th }}$ and $j^{\text {th }}$ channels apart from those in $z^{-D} H_{c}(z)$ and finally $Q_{i / j}(z), i \neq j$, denotes the zeros of the $i^{\text {th }}$ channel which are not also in the $j^{\text {th }}$ channel. The importance of this representation is that the filters in a dashed box do not share any common zeros. For example the zeros common to $1^{\text {st }}$ and $2^{\text {nd }}$ channels are in $z^{-D} H_{c}(z) H_{12}(z)$ therefore $Q_{1 / 2}(z)$ and $Q_{2 / 1}(z)$ cannot have any zeros in common.

In this simulation, the independent identically distributed input sequence $a[n]$ is drawn from a BPSK symbol constellation, hence the causal input sequence consists of $+/-1$ 's for $n \geq 0$.

### 5.1 The Estimation of the Delay

The outputs $y_{1}[n], \ldots, y_{4}[n]$ of the channels are shown in Fig. 5.2. We use this figure to identify the delays associated with the channels. For example we observe that $y_{1}[n]=0$ for $n<4$ and $y_{1}[4] \neq 0$. Thus the delay associated with the first channel is 4 samples. With similar reasoning the delays associated with other channels are found as 3,2 and 3 samples respectively. Finally we compute the common delay of the communication system as $D=\min \{4,3,2,3\}=2$.

### 5.2 The Binary-tree Algorithm

The next step is the identification of the filters in the boxes labeled as 1,2 and 3 in Fig. 3.4 (the binary-tree of section 3.5). We will first estimate the leaves of the tree (boxes 2,3 ) and finally the root (box 1). The following comments refer to the computation of the FIR filters $Q_{1 / 2}(z)$ and $Q_{2 / 1}(z)$ in box 1 and their common input $u_{12}[n]$ :

1) The data in Fig. 5.2 is used. Therefore $N=20$.
2) The filter orders are overestimated as $\hat{L}_{2}=8$.
3) The matrix $\boldsymbol{R}_{\boldsymbol{y y}}[\boldsymbol{N}]$ in (3.9) is computed with $M=2$ and $\hat{L}_{2}=8$ using $y_{1}[n]$ and $y_{2}[n], 0 \leq n \leq N$. The singular values of $\boldsymbol{R}_{\boldsymbol{y y}}[\boldsymbol{N}]$ are plotted in Fig. 5.3. Clearly $\eta$, the number of zero singular values, is 5 .
4) The maximum of the orders of $Q_{1 / 2}(z)$ and $Q_{2 / 1}(z)$ is computed as $L_{2}=$ $\hat{L}_{2}-\eta+1=4$.
5) $\boldsymbol{R}_{y y}[\boldsymbol{N}]$ is recomputed using the true filter order $L_{2}=4$. Any of the constrained minimization algorithms of Section 3.3 can be used to compute the estimates $\hat{Q}_{1 / 2}(z)$ and $\hat{Q}_{2 / 1}(z)$ from $\boldsymbol{R}_{y y}[\boldsymbol{N}]$. In this example, Energy Constraint 1 is used to compute $\hat{Q}_{1 / 2}(z)$ and $\hat{Q}_{2 / 1}(z)$ as:

$$
\begin{align*}
& \hat{Q}_{1 / 2}(z)=(-0.37403-0.13025 j) Q_{1 / 2}(z)  \tag{5.1}\\
& \hat{Q}_{2 / 1}(z)=(-0.37403-0.13025 j) Q_{2 / 1}(z) \tag{5.2}
\end{align*}
$$

6) The FIR deconvolution filters $\hat{E}_{1 / 2}(z)$ and $\hat{E}_{2 / 1}(z)$ are computed using (3.62). The output of the deconvolution filters is computed as (see

Fig. 3.3)

$$
\begin{equation*}
\hat{u}_{12}[n]=(-2.3844+0.83038 j) u_{12}[n], \quad 0 \leq n \leq N . \tag{5.3}
\end{equation*}
$$

Adapting the same steps to the $2^{\text {nd }}$ box, we identified the filters in the $2^{\text {nd }}$ box and their common input $\hat{u}_{34}[n]$ for $0 \leq n \leq N$ as

$$
\begin{align*}
\hat{Q}_{3 / 4}(z) & =(-0.24927) Q_{3 / 4}(z)  \tag{5.4}\\
\hat{Q}_{4 / 3}(z) & =(-0.24927) Q_{4 / 3}(z)  \tag{5.5}\\
\hat{u}_{34}[n] & =(-4.0117) u_{34}[n] \tag{5.6}
\end{align*}
$$

At this point we know the output of the filters in the $3^{\text {rd }}$ box, $\hat{u}_{12}[n]$ and $\hat{u}_{34}[n]$, for $0 \leq n \leq N$ which is sufficient to estimate $H_{12}(z)$ and $H_{34}(z)$. The estimates of these filters and their common input $x_{c}[n], 0 \leq n \leq N$ are found as

$$
\begin{align*}
\hat{H}_{12}(z) & =(0.064595+0.37838 j) H_{12}(z)  \tag{5.7}\\
\hat{H}_{34}(z) & =(-0.10080+0.60151 j) H_{34}(z)  \tag{5.8}\\
\hat{x}_{c}[n] & =(1.0871+6.4872 j) x_{c}[n] \tag{5.9}
\end{align*}
$$

### 5.3 The Pruning Algorithm

The last problem is the identification of $H_{c}(z)$ in Fig. 3.4 from its estimated output $x_{c}[n]$. This is accomplished using the pruning algorithm of Chapter 4. In this example we overestimate the number of common zeros as $\hat{L}_{1}=4$. We note that for $0 \leq n \leq 4$ there are only 32 possible input sequences $a^{(k)}[n]$ half of which differ only in sign. Therefore there are effectively only 16 non
trivially distinct input sequences. Since $x_{c}[n]$, the output of the filter $H_{c}(z)$, has been already identified, we use (4.2) to compute the 16 channel estimates corresponding to each of the distinct input sequence $a^{(k)}[n]$. Then for each channel estimate we compute the residual error $e^{(k)}[n], n \geq D+\hat{L}_{1}+1=7$, defined in Chapter 4. The running energy of the residual errors

$$
\begin{equation*}
E_{k}[n]=\sum_{i=0}^{n}\left|e^{(k)}[i]\right|^{2}, \tag{5.10}
\end{equation*}
$$

for all the channel estimates are plotted in Fig. 5.4. In this figure we see that the energy of only one of the error sequences remains around $10^{-30}$ at all samples. Hence this channel is identified as the estimate of the common zeros, $\hat{H}_{c}(z):$

$$
\begin{equation*}
\hat{H}_{c}(z)=(1.0871+6.4872 j) H_{c}(z) \tag{5.11}
\end{equation*}
$$

The use of $\hat{H}_{c}(z)$ in Fig. 4.1 provides the estimate of the input sequence

$$
\begin{equation*}
\hat{a}[n]=a[n] \tag{5.12}
\end{equation*}
$$



Figure 5.1: The pole-zero plots of the channels: (a) Channel 1, (b) Channel 2, (c) Channel 3, (d) Channel 4.


Figure 5.2: The magnitudes of the outputs of the channels: (a) Channel 1, (b) Channel 2, (c) Channel 3, (d) Channel 4.


Figure 5.3: The singular values.


Figure 5.4: The running energy of the error sequences associated with different channel estimates (identification of the common zeros).

## Chapter 6

## CONCLUSIONS

The blind channel identification problem is investigated using the multichannel filter model of the communication system. The main contribution of this thesis is to give the exact solution to this problem under the noise-free case virtually with the least set of constraints on the channel. Furthermore this is achieved based on a very short observation of the channel output sequence. This is an important result because the current algorithms in the literature are in general either fast converging to the solution but applicable to only a certain restricted class of channels (SOCS-based algorithms) or generally applicable but computationally intensive and large amount of data demanding (HOS-based algorithms).

In this thesis the problem is divided into two sub-problems which are independently formulated and solved: The identification of the uncommon zeros (i.e., $h_{1}[n], \ldots, h_{M}[n]$ ) followed by the identification of the common zeros (i.e., $\left.h_{c}[n]\right)$ in the multi-channel filter model. The first sub-problem is formulated
in Chapter 3. In that chapter a cost function in terms of the sub-channels $h_{1}[n], \ldots, h_{M}[n]$ is derived and its minimizers are fully characterized with a powerful theorem. An important aspect of this theorem is that it also specifies the minimum number of samples required for the exact identification of the sub-channels $h_{1}[n], \ldots, h_{M}[n]$ upto a scalar multiplicative constant. Based on this theorem several closed-form solutions are provided to estimate these subchannels and the equivalence of these estimates under the noise-free case is proven. The investigation of the behavior of these closed-form solutions under the more realistic noisy channels is a future research topic. Finally a novel binary-tree algorithm is provided to estimate these channels in a computationally efficient way. The utilization of this structure in noisy channels is also a future research topic.

The solution to the second sub-problem, i.e., the identification of the filter $h_{c}[n]$ corresponding to the common zeros, is given in Chapter 4 based on a novel pruning algorithm. This algorithm first identifies all possible estimates of $h_{c}[n]$ based on first few samples of the output. This is achieved using the fact that the input symbols are not arbitrary but drawn from a known symbol constellation. The basic step of the algorithm is to eliminate the channel estimates which are inconsistent with the observation of the current output sample. If the BCI problem is solvable then the pruning algorithm always converges to the true solution in a finite number of steps. It is shown with simulation that this algorithm is practical for moderate orders of $h_{c}[n]$ and in practice it converges to the true solution very fast.

## Bibliography

[1] S. Haykin, Adaptive Filter Theory. New Jersey: Prentice-Hall, third ed., 1996.
[2] Y. Sato, "A method of self-recovering equalization for multi-level amplitude modulation," IEEE Trans. on Communications, vol. COM-23, pp. 679-682, June 1975.
[3] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications," IEEE Trans. Automatic Control, vol. COM-25, pp. 385-399, June 1980.
[4] D. N. Godard, "Self-recovering equalization and carrier tracking in twodimensional data communication systems," IEEE Trans. on Communications, vol. COM-28, pp. 1867-1875, 1980.
[5] G. Picci and G. Prati, "Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm," IEEE Trans. on Communications, vol. COM-35, pp. 877-887, 1987.
[6] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," IEEE Trans. Signal Process., vol. 43, pp. 2120-2129, Sept. 1995.
[7] D. R. Brillinger, Time Series: Data analysis and theory. New York: Holt, Rinehart, and Winston, 1974.
[8] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approah," IEEE Trans. Information Theory, vol. 40, pp. 340-349, Mar. 1994.
[9] G. B. Giannakis and S. D. Halford, "Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions," IEEE Trans. Signal Process., vol. 45, pp. 2277-2292, Sept. 1997.
[10] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," IEEE Trans. Signal Process., vol. 43, pp. 516-525, Feb. 1995.
[11] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filter banks and multi-channel linear prediction," in Proc. IEEE, Int. Conf. Acoust. Speech Signal Process., vol. 4, pp. 585-588, 1994.
[12] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," IEEE Trans. Signal Process., vol. 43, pp. 29822993, Dec. 1995.
[13] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital cornmunications," IEEE Trans. Signal Process., vol. 43, pp. 2714-2723, Nov. 1995.
[14] D. Gesbert, P. Duhamel, and S. Mayrargue, "On-line blind multichannel equalization based on mutually referenced filters," IEEE Trans. Signal Process., vol. 45, pp. 2307-2317, Sept. 1997.
[15] A. Touzni and I. Fijalkow, "Robustness of blind fractionally-spaced identification/equalization to loss of channel disparity," in Proc. IEEE Int. Conf. Acoust. Speech Signal Process., pp. 3937-3940, 1997.
[16] Z. Ding, "Characteristics of band-limited channels unidentifiable form second-order cyclostationary statistics," IEEE Trans. Signal Processing Letters, vol. 3, pp. 150-152, May 1996.
[17] J. Tugnait, "On blind identifiability of the multipath channels using fractional sampling and second-order cyclostationary statistics," IEEE Trans. Information Theory, vol. 41, pp. 308-311, Jan. 1995.
[18] B. Farhang-Boroujney, "Channel equalization via channel identification: Algorithms and simulation results for rapidly fading HF channels," IEEE Trans. on Communications, vol. 44, pp. 1409-1412, Nov. 1996.
[19] M. I. Gürelli and C. L. Nikias, "Evam: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," IEEE Trans. Signal Process., vol. 43, pp. 134-149, Jan. 1995.
[20] W. Qui and Y. Hua, "A GCD method for blind channel identification," Digital Signal Processing, pp. 199-205, 1997.
[21] L. L. Scharf, Statistical signal processing : detection, estimation, and time series analysis, ch. 2. Addison-Wesley, 1991.
[22] T. Kailath, Linear Systems. New Jersey: Prentice-Hall, 1980.

## APPENDIX A

## Multi-channel Filter Model

The fact that the channel model in Fig. 2.1 can be represented as in Fig. 2.2 is widely used in the current blind channel identification literature. Here we provide a proof of this result.

The filtered and noise corrupted received signal $y(t)$ is given as

$$
\begin{equation*}
y(t)=\sum_{m=0}^{\infty} a[m] h(t-m T)+v(t) \tag{A.1}
\end{equation*}
$$

where $h(t)$ is the baseband model of the channel defined in Chapter 2 and $v(t)$ is the additive channel noise. The output of the sampler in Fig. 2.1 is the samples of $y(t)$ taken at an $M^{\prime}$ times faster rate then the baud rate. It is given as

$$
\begin{equation*}
y\left(k \frac{T}{M^{\prime}}\right)=\sum_{m=0}^{\infty} a[m] h\left(k \frac{T}{M^{\prime}}-m T\right)+v\left(k \frac{T}{M^{\prime}}\right) . \tag{A.2}
\end{equation*}
$$

Substituting $k=n M^{\prime}+(i-1)$ into (A.3) it turns out that
$y\left(n T+(i-1) \frac{T}{M^{\prime}}\right)=\sum_{m=0}^{\infty} a[m] h\left((n-m) T+(i-1) \frac{T}{M^{\prime}}\right)+v\left(n T+(i-1) \frac{T}{M^{\prime}}\right)$

$$
\begin{equation*}
\text { for } i=1, \ldots, M^{\prime} \tag{A.3}
\end{equation*}
$$

If we use the following definitions

$$
\begin{align*}
& y_{i}[n] \triangleq y\left(n T+(i-1) \frac{T}{M^{\prime}}\right)  \tag{A.4}\\
& \tilde{h}_{i}[n] \triangleq h\left(n T+(i-1) \frac{T}{M^{\prime}}\right)  \tag{A.5}\\
& v_{i}[n] \triangleq v\left(n T+(i-1) \frac{T}{M^{\prime}}\right) \tag{A.6}
\end{align*}
$$

then (A.3) simplifies to

$$
\begin{equation*}
y_{i}[n]=a[n] * \tilde{h}_{i}[n]+v_{i}[n] \quad, \quad \text { for } i=1, \ldots, M^{\prime} \tag{A.7}
\end{equation*}
$$

The Fig. 2.2 directly follows from (A.7) where $\tilde{h}_{i}[n]$ is decomposed as

$$
\begin{equation*}
\tilde{h}_{i}[n]=\delta[n-D] * h_{c}[n] * h_{i}[n] \tag{A.8}
\end{equation*}
$$

## APPENDIX B

## Proof of Lemma 1

Let $r[n]$ be the filter defined as

$$
\begin{equation*}
r[n]=g_{i}[n] * h_{j}[n]-g_{j}[n] * h_{i}[n] . \tag{B.1}
\end{equation*}
$$

Since $r[n]$ is constructed from causal FIR filters, it is itself a causal FIR filter with its support constrained into the interval $0 \leq n \leq L_{r}$ :

$$
\begin{equation*}
r[n]=0 \quad, \quad \text { for } n<0 \text { for } n>L_{r} . \tag{B.2}
\end{equation*}
$$

We are given that $e[n] \triangleq r[n] * x_{c}[n]=0$ for $0 \leq n \leq D+L_{r} \leq N$. We also deduce from the presence of delay $D$ in Fig. 2.2 that $x_{c}[n]=0$ for $n<D$ and $x_{c}[D]=a[0] h_{c}[0] \neq 0$. These require that

$$
\begin{equation*}
e[D]=r[0] x_{c}[D]=0 \Rightarrow r[0]=0 . \tag{B.3}
\end{equation*}
$$

Using the relation

$$
\begin{equation*}
r[n-D]=-\frac{1}{x_{c}[D]} \sum_{k=D+1}^{n} r[n-k] x_{c}[k] \tag{B.4}
\end{equation*}
$$

for $n=D+1, \ldots, D+L_{r}$ with the fact that $r[0]=0$ it follows that $r[n]=0$ for $0 \leq n \leq L_{r}$. Since this interval is the support of $r[n]$ we conclude that $r[n]=0$ for all $n$.

## APPENDIX C

## Proof of Lemma 2

To start with the proof of the lemma, let $\mathcal{Z}_{i}$ be a set of touplets $\left\{\left(z_{i k}, m_{i k}\right)\right\}$ where $z_{i k}$ is the $k^{\text {th }}$ zero of the $H_{i}(z)$ with multiplicity $m_{i k} \geq 1$. Our conclusion will be based on the following claims which show that all the zeros of $H_{i}(z)$ should be a zero of $G_{i}(z)$.

Claim 1. For any $z_{i k}$, there exists at least one $j_{k}$ such that $H_{i j_{k}}(z)$ does not accept $z_{i k}$ as its root.

Proof. Assume that for a particular $z_{i k}$, we do not have any $H_{i j}(z)$ which does not accept $z_{i k}$ as its root. Hence, $z_{i k}$ is a common zero for all $H_{i j}(z), 1 \leq j \leq M$. Therefore, it is a common zero between $H_{i}(z)$ and $H_{j}(z)$ for $1 \leq j \leq M$. This contradicts with the definition of $H_{i}(z)$ 's, since they are defined as polynomials which share no common zeros (Note that all the common zeros of the subchannels are represented with $H_{c}(z)$ in Fig. 2.2.).

By using this result, we can state and prove the following claim:

Claim 2. Let $z_{i k}$ be the $k^{\text {th }}$ zero of $H_{i}(z)$ with multiplicity $m_{i k}$, then $z_{i k}$ is a zero of $G_{i}(z)$ with multiplicity larger than or equal to $m_{i k}$.

Proof. In Claim 1 we have proved that for any $z_{i k}$, there exists a $H_{i j_{k}}(z)$ such that $H_{i j_{k}}(z)$ does not accept $z_{i k}$ as its root. Hence, invoking (3.23) with $i$ and $j_{k}$ we get:

$$
\begin{equation*}
Q_{i / j_{k}}(z) \mid G_{i}(z) \tag{C.1}
\end{equation*}
$$

Since $H_{i}(z)=H_{i j_{k}}(z) Q_{i / j_{k}}(z)$, and $z_{i k}$ is not a root of $H_{i j_{k}}(z), z_{i k}$ is a root of $Q_{i / j_{k}}(z)$ with multiplicity $m_{i k}$. Since $Q_{i / j_{k}}(z)$ divides $G_{i}(z), z_{i k}$ should be a root of $G_{i}(z)$ with multiplicity greater than or equal to $m_{i k}$.

Proof of Lemma 2. Claim 2 implies that all the roots of $H_{i}(z)$ are also roots of $G_{i}(z)$. Thus :

$$
\begin{equation*}
H_{i}(z) \mid G_{i}(z) \tag{C.2}
\end{equation*}
$$

This completes the proof of Lemma 2.

## APPENDIX D

## Proof of Lemma 3

If $F_{i}(z) \neq F_{j}(z)$ for a particular $i$ and $j$ pair, by using (3.19) we get:

$$
\begin{equation*}
F_{i}(z) H_{i}(z) H_{j}(z)=F_{j}(z) H_{j}(z) H_{i}(z) \tag{D.1}
\end{equation*}
$$

which implies the equality of $F_{i}(z)$ and $F_{j}(z)$ at all $z$ except at a discrete set of $z$ values where $H_{i}(z) H_{j}(z)=0$. Since $F_{i}(z)$ and $F_{j}(z)$ are both analytic functions of $z$,

$$
\begin{equation*}
F_{i}(z)=F_{j}(z) \text { for all } z \tag{D.2}
\end{equation*}
$$

completing the proof of the lemma.

## APPENDIX E

## Proof of Lemma 4

To prove the Lemma 4 it is convenient to express the convolution $g_{i}[n]=$ $f[n] * h_{i}[n]$ using the matrix notation:

$$
\begin{equation*}
\boldsymbol{g}_{\boldsymbol{i}}=\boldsymbol{F} \boldsymbol{h}_{i}, \quad \text { for } 1 \leq i \leq M, \tag{E.1}
\end{equation*}
$$

where the filter vectors $\boldsymbol{g}_{\boldsymbol{i}}, \boldsymbol{h}_{\boldsymbol{i}}$ and the $\left(\hat{L}_{2}+1\right) \mathrm{X}\left(L_{2}+1\right)$ dimensional convolution matrix $\boldsymbol{F}$ are defined as

$$
\begin{align*}
\boldsymbol{g}_{\boldsymbol{i}} & =\left[\begin{array}{llll}
g_{i}[0] & g_{i}[1] & \cdots & g_{i}\left[\hat{L}_{2}\right]
\end{array}\right]^{T}  \tag{E.2}\\
\boldsymbol{h}_{\boldsymbol{i}} & =\left[\begin{array}{cccc}
h_{i}[0] & h_{i}[1] & \cdots & h_{i}\left[L_{2}\right]
\end{array}\right]^{T}  \tag{E.3}\\
\boldsymbol{F} & =\left[\begin{array}{cccc}
f[0] & & \\
f[1] & f[0] & \\
\vdots & f[1] & \\
f\left[\hat{L}_{c}\right] & \vdots & & f[0] \\
& f\left[\hat{L}_{c}\right] & f[1] \\
& & \vdots \\
& & f\left[\hat{L}_{c}\right]
\end{array}\right] \tag{E.4}
\end{align*}
$$

We can write (E.1) as

$$
\left[\begin{array}{c}
\boldsymbol{g}_{1}  \tag{E.5}\\
\boldsymbol{g}_{2} \\
\vdots \\
\boldsymbol{g}_{M}
\end{array}\right]=\left[\begin{array}{llllll}
\boldsymbol{F} & \boldsymbol{A} & & & & \\
& \boldsymbol{B} & \boldsymbol{F} & \boldsymbol{A} & & \\
& & & & & \\
& & & B & \boldsymbol{F} & \boldsymbol{A}
\end{array}\right] \cdot\left[\begin{array}{c}
\boldsymbol{h}_{1} \\
\mathbf{0} \\
\boldsymbol{h}_{2} \\
\mathbf{0} \\
\\
\mathbf{0} \\
\boldsymbol{h}_{M} \\
\mathbf{0}
\end{array}\right] \text { (E }
$$

where $\left(\hat{L}_{2}+1\right) \mathrm{X} \hat{L}_{c}$ dimensional $\boldsymbol{A}$ and $\boldsymbol{B}$ matrices can be arbitrarily chosen, because they are multiplied with $\mathbf{0}$ 's of $\boldsymbol{v}_{\mathbf{0}}$. If we choose $\boldsymbol{A}$ and $\boldsymbol{B}$ as the below toeplitz matrices

$$
\boldsymbol{A}=\left[\begin{array}{ccc}
0 & 0 & 0 \\
& & \\
0 & 0 & 0 \\
f[0] & 0 & 0 \\
f[1] & f[0] & \\
\vdots & & 0 \\
f\left[\hat{L}_{c}-1\right] & & f[1]
\end{array}\right][0]\left[\begin{array}{ccc}
f\left[\hat{L}_{c}\right] & f\left[\hat{L}_{c}-1\right] & f[1] \\
0 & f\left[\hat{L}_{c}\right] & \\
& \ddots & f\left[\hat{L}_{c}-1\right] \\
0 & 0 & f\left[\hat{L}_{c}\right] \\
0 & 0 & 0 \\
& & \\
0 & 0 & 0
\end{array}\right]
$$

then $\mathcal{F}$ becomes a toeplitz matrix with $f\left[\hat{L}_{c}-i+1\right]$ on the $i^{\text {th }}, 1 \leq i \leq \hat{L}_{c}$, lower sub-diagonal. Therefore $\mathcal{F}$ can be written as

$$
\begin{equation*}
\mathcal{F}=\sum_{k=0}^{\dot{L}_{c}} f\left[\hat{\mathcal{L}}_{c}-k\right] \boldsymbol{J}^{k}, \tag{E.7}
\end{equation*}
$$

where $\boldsymbol{J}$ is defined in (3.27). Substituting this into (E.6), (3.25) immediately follows.


[^0]:    ${ }^{1}$ The rate at which input symbols are transmitted through the channel. It is also known as the baud rate.

[^1]:    ${ }^{1}$ Since $h_{1}[n], \ldots, h_{M}[n]$ do not share any common zeros, the first coefficient of at least one of these sub-channels should be non-zero.

[^2]:    ${ }^{2}$ For notational convenience, the dependence on $N$ is not shown explicitly.

[^3]:    ${ }^{1}$ Using an over-estimate of $L_{1}$ does not change the results in this chapter. It only increases the computational load.

[^4]:    ${ }^{2}$ Notice that there was no identifiability problem in estimation of the uncommon zeros.

[^5]:    ${ }^{3}$ For simplicity we assume that the input sequences are drawn from a BPSK symbol constellation.

[^6]:    ${ }^{1}$ Except for this discussion, we will not discriminate between the finite zeros and the zeros at infinity (delays).

