Department of Electrical and Computer Engineering

Low Order Channel Estimation for CDMA Systems

Amar ABD EL-SALLAM

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University of Technology

December 2005

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

Signature: .....

Date: .....

To my father, my mother and to someone friends who are really so special to me.

Acknowledgments

First of all I would like to thank Curtin University of Technology and the Australian Telecommunication Cooperative Research Centre (ATcrc) for offering me the International Postgraduate Research (IPRS) and the ATcrc Top-Up scholarships and for supporting me with many facilities during my PhD.

I would like to express my gratitude to my supervisors, Prof. Abdelhak M. Zoubir and Dr. Yee H. Leung for their continuous scientific expertise and constructive advice which helped me to develop this research.

I would also like to thank all of the CSP group members especially Dr. Roy Howard for their great and continuous support. Thanks also go to all of the SPG members in Darmstadt for making my visit to Germnay memorable.

Additional thanks to the Department of Electrical & Computer Engineering especially administrative and technical staff.

Very special thanks to the office of Research and Development (R&D) and to the Devision of Engineering, Science and Computing including Student Services in building 314.

Finally thanks to many very good friends I have met in Australia who made my life happy.

Keywords

Code Division Multiple Access, Multipath, Rayleigh fading, Multiple Access Interference, RAKE receiver, FIR filter, Equal Gain Combiner, Maximum Ratio Combiner, Maximum Likelihood, Least Square, Maximum Length Binary Sequence, Minimum Description Length, Akaike Information Criterion, sphericity test, hypothesis test, F-Statistics, bootstrap, probability value, Bonferroni test, Backward Elimination, Singular Value Decomposition, Periodogram, cross-correlation, cross-spectra, Bit Error Rate, Adaptive, Blind estimation, Recursive Least Squares, subspace.

Abstract

New approaches and algorithms are developed for the identification and estimation of low order models that represent multipath channel effects in Code Division Multiple Access (CDMA) communication systems. Based on these parsimonious channel models, low complexity receivers such as RAKE receivers are considered to exploit these propagation effects and enhance the system performance.

We consider the scenario where multipath is frequency selective slowly fading and where the channel components including delays and attenuation coefficients are assumed to be constant over one or few signalling intervals. We model the channel as a long FIR-like filter (or a tapped delay line filter) with the number of taps related to the ratio between the channel delay-spread and the chip duration. Due to the high data rate of new CDMA systems, the channel length in terms of the chip duration will be very large. With classical channel estimation techniques this will result in poor estimates of many of the channel parameters where most of them are zero leading to a reduction in the system performance. Unlike classical techniques which estimate directly the channel response given the number of taps or given an estimate of the channel length, the proposed techniques in this work will firstly identify the significant multipath parameters using model selection techniques, then estimate these identified parameters.

Statistical tests are proposed to determine whether or not each individual parameter is significant. A low complexity RAKE receiver is then considered based on estimates of these identified parameters only. The level of significance with which we will make this assertion will be controlled based on statistical tests such as multiple hypothesis tests. Frequency and time domain based approaches and model selection techniques are proposed to achieve the above proposed objectives.

The frequency domain approach for parsimonious channel estimation results in an efficient implementation of RAKE receivers in DS-CDMA systems.
In this approach, we consider a training based strategy and estimate the channel delays and attenuation using the averaged periodogram and modified time delay estimation techniques. We then use model selection techniques such as the sphericity test and multiple hypotheses tests based on F-Statistics to identify the model order and select the significant channel paths. Simulations show that for a pre-defined level of significance, the proposed technique correctly identifies the significant channel parameters and the parsimonious RAKE receiver shows improved statistical as well as computational performance over classical methods.
The time domain approach is based on the Bootstrap which is appropriate for the case when the distribution of the test statistics required by the multiple hypothesis tests is unknown. In this approach we also use short training data and model the channel response as an FIR filter with unknown length. Model parameters are then estimated using low complexity algorithms in the time domain. Based on these estimates, bootstrap based multiple hypotheses tests are applied to identify the non-zero coefficients of the FIR filter. Simulation results demonstrate the power of this technique for RAKE receivers in unknown noise environments.
Finally we propose adaptive blind channel estimation algorithms for CDMA systems. Using only the spreading code of the user of interest and the received data sequence, four different adaptive blind estimation algorithms are proposed to estimate the impulse response of frequency selective and frequency non-selective fading channels. Also the idea is based on minimum variance receiver techniques. Tracking of a frequency selective varying fading channel is also considered. A blind based hierarchical MDL model selection method is also proposed to select non-zero parameters of the channel response. Simulation results show that the proposed algorithms perform better than previously proposed algorithms. They have lower complexity and have a faster convergence rate. The proposed algorithms can also be applied to the design of adaptive blind channel estimation based RAKE receivers.

Contents

List of Figures

List of Tables

Abbreviations, Acronyms and Symbols

TDMA	Time Division Multiple Access
FDMA	Frequency Division Multiple Access
CDMA	Code Division Multiple Access
DS-CDMA	Direct Sequence Code Division Multiple Access
TH-CDMA	Time Hopping Code Division Multiple Access
FH-CDMA	Frequency Hopping Code Division Multiple Access
BW	Band Width
0G	Zeroth Generation
1G	First Generation
2G	Second Generation
3G	Third Generation
4G	Forth Generation
ITU	International Telecommunication Union
IMT-2000	International Mobile Telecommunications-2000
3GPP	3rd Generation Partnership Project
APR	Autoradiopuhelin (car radio)
NMT	Nordisk Mobil Telefon
AMPS	Advanced Mobile Phone System
TAC	Total Access Communication System
GSM	Global System for Mobile Communications
iDEN	Integrated Digital Enhanced Network
IS-95	Interim Standard 95

PDC	Personal Digital Cellular
GPRS	General Packet Radio Service
SMS	Short Message Service
EDGE	Enhanced Data Rates for GSM Evolution
W-CDMA	Wideband Code Division Multiple Access
UMTS	Universal Mobile Telecommunications System
FOMA	Freedom of Mobile Multimedia Access
MC-CDMA	Multi-Carrier Code Division Multiple Access
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
SMS	Short Message Service
HSDPA	High-Speed Downlink Packet Access
w.r.t	with respect to
Km/h	Kilometer per hour
A/D	Analoge to Digital
ISI	Inter-Symbol Interference
MAI	Multiple Access Interference
MACSI	Multiple Access Cross Spectral Interference
MPI	Multipath Interference
EGC	Equal Gain Combiner
MRC	Maximum Ratio Combiner
AGC	Automatic Gain Control
AFC	Automatic Frequency Control
PSD	Power Spectral Density
AIC	Akaike Information Criterion
MDL	Minimum Description Length
FE	Forward Elimination
BE	Backward Elimination
#	Number of times
BER	Bit Error Rate
pdf	probability density function
cdf	cumulative distribution function

DOA	Direction of Arrival
FWE	Family Wise Error rate
iid	independent and identically distributed
LSE	Least Squares Estimate/Estimator
MHT	Multiple Hypothesis Test
MLE	Maximum Likelihood Estimate/Estimator
MMSE	Minimum Mean Squared Error
MSE	Mean Squared Error
SSE	Sum Squared error
SSR	Regression Sum Squared error
MUD	Multiuser Detection
NFR	Near Far Ratio
$\mathcal{P}$-value	Probability value
RV	Random Variable
SNR	Signal to Noise Ratio
SRB	Sequentially Rejective Bonferroni
$\approx$	Approximately
$(\cdot)^*$	Complex conjugation
$(\cdot)^{\mathsf{H}}$	Hermitian transpose
$(\cdot)^{\mathsf{T}}$	Transposition
$\lvert\cdot\rvert$	Absolute value
$\lvert\lvert\cdot\rvert\rvert$	Euclidean norm
$\lfloor\cdot\rfloor$	The integer part
$\hat{\theta}$	Estimate or estimator of a parameter θ
λ_i	Eigenvalues
N	Normal
μ	Mean
σ^2	Variance
χ^2_d	Chi-squared distribution with d degrees of freedom
$\mathsf{Cov}[X, Y]$	Covariance between X and Y
$\mathrm{diag}(\cdot)$	A diagonal matrix

$\mathsf{E}\{X\}$	Expectation of X
$\mathsf{Var}\{X\}$	Variance of X
$f_X(x)$	Probability density function of X
$F_X(x)$	Cumulative distribution function of X
$\mathcal{H}$	Null hypothesis
$\mathcal{K}$	Alternative hypothesis
$\mathrm{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$\mathbb{Z}$	The set of integers
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers
$\mathrm{Re}(\cdot)$	Real part of a complex variable
$\mathrm{Im}(\cdot)$	Imaginary part of a complex variable
$\mathrm{sgn}(\cdot)$	The sginum function
$\mathrm{U}(a, b)$	Uniform distribution over the interval $[a, b]$
MV	Minimum Variance
SVD	Singular Value Decomposition
EVD	Eigen Value Decomposition

Publications

The following publications have been produced during the period of the PhD.

Internationally Refereed Journal Articles

1. Amar Abd El-Sallam and Abdelhak M. Zoubir, "Selection of Significant Multipath Components for Low Complexity CDMA RAKE Receivers: A Frequency Domain Based Approach " *IEEE Trans. on Signal Processing (accepted)*

2. Amar Abd El-Sallam, Abdelhak M. Zoubir and Yee Hong Leung, "On the Selection of Significant Multipath Components for Low Complexity CDMA RAKE Receivers," *IEEE Trans. on Signal Processing (submitted)*

3. Amar Abd El-Sallam, Abdelhak M. Zoubir and Yee Hong Leung, "Fast Blind Adaptive Channel Estimation and Tracking Algorithms for CDMA Systems" *IEEE Trans. on Signal Processing (submitted)*

Internationally Refereed Conference Papers

1. Amar A. El-sallam, Abdelhak M. Zoubir and Samir Attallah, "A Bootstrap-Based RAKE Receiver for CDMA Systems," In *Proc. of IEEE Globecom'02*, vol. 2, pp. 1073-1077, Taiwan, November 2002.

2. Amar A. Abd el-sallam, Salim Kayhan and Abdelhak M. Zoubir, "Two Backward Elimination Based Approaches for Low Order Model Selection,"

In *Proc. of the 7th International Symposium on Signal Processing and its Applications (ISSPA'03)*, vol. 2, pp. 535-538, Paris, France, July 2003.

3. Amar A. Abd El-sallam, Salim Kayhan and Abdelhak M. Zoubir, "Bootstrap and Backward Elimination-Based Approaches for Model Selection," In *Proc. of the 3rd Intl. Symposium on Image and Signal Processing and Analysis (ISPA'03)*, vol. 1, pp. 152-157, Rome, Italy, September 2003.

4. Amar Abd El-sallam, Yee Hong Leung and Abdelhak M. Zoubir, "An F-Statistic Frequency Domain Based Approach To The Optimum Design of RAKE Receiver Structures For CDMA Systems," In *Proc. of the IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'03)*, vol. 1, pp. 367-370, Darmstadt, Germany, December 2003.

5. Amar A. Abd El-sallam and Abdelhak M. Zoubir, " A Frequency Domain Based Approach for Low Complexity RAKE Receivers in CDMA Systems," In *Proc.of the IEEE Workshop on Statistical Signal Processing (SSP'03)*, vol. 1, pp. 327-330, St. Louis, MO, USA, September 2003.

6. Amar Abd El-sallam and Abdelhak M. Zoubir, " Fast Adaptive Channel Estimation Algorithms For CDMA Systems," In *Proc.of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05)*, Philadelphia, PA, USA, March 2005

7. Amar Abd El-sallam, Yee Hong Leung and Abdelhak M. Zoubir, "A Frequency Domain Sphericity Test Based Approach For Parsimonious RAKE Receiver in CDMA systems," In *Proc. of the 8th International Symposium on Signal Processing and its Applications (ISSPA'05)*, Sydney, Australia, July 2005.

8. Amar Abd El-sallam, Yee Hong Leung and Abdelhak M. Zoubir, "To The Selection of Significant Multipath Components For Parsimonious DS-CDMA RAKE Receivers: A Covariance Bootstrap Based Approach," In *Proc. of the IEEE 11th Asia-Pacific Conf. in Commun., (APCC'05)*, Perth, Australia, October 2005.

Locally Refereed Conference Papers

1. Amar Abd El-Sallam, Abdelhak M. Zoubir and S. Attallah, "Low Order Channel Estimation for CDMA," In *Proc. of the 3rd Australian Communications Theory Workshop (AusCTW'02)*, pp. 105-108., Canberra, Australia, February. 2002.

2. Amar Abd El-Sallam, "Low Order Bootstrap Based RAKE Receivers for CDMA Systems," In *Proc. of the 3rd Inter University Postgraduate Electrical Engineering Symposium (IUPEES'02)*, Rockingham, Australia, October 2002 .

3. Amar Abd El-Sallam, "A Frequency Domain Based Approach For Reduced Complexity RAKE Receivers In CDMA Systems," In *Proc. of the Australian Telecommunications Cooperative Research Centre (ATcrc) Workshop*, pp. 42-45, Melbourne, Australia,October 2003 .

4. Amar Abd El-Sallam, "A Low Complexity Bootstrap-Based Rake Receiver For CDMA Systems," In *Proc. of the Postgraduate Electrical Engineering and Computing Symposium (PEECS'03)*, pp. 108-112, Perth, Australia, 2003.

5. Amar Abd El-Sallam and Yee Hong Leung "Rapidly Convergent Blind Adaptive Channel Estimation Algorithms for CDMA Systems," In *Proc. of the Australian Telecommunications Cooperative Research Centre (ATcrc) Workshop*, pp. 63-76, Perth, Australia, October 2004.

6. Amar Abd El-Sallam, "Blind Adaptive Channel Estimation & Tracking Algorithms for CDMA Systems ," In *Proc. of the Postgraduate Electrical Engineering and Computing Symposium (PEECS'04)*, pp. 136-143, Perth, Australia, 2004.

Chapter 1

Introduction

Direct Sequence Code Division Multiple Access (DS-CDMA) communication systems have attracted considerable attention for third-generation (3G) mobile systems. They have the ability to suppress a wide variety of interfering signals including Narrow-Band Interference (NBI), Multiple-Access Interference (MAI), and Multipath Interference (MPI) [1] [2] [3].

Multipath effects are one of the major factors that limit system performance. Depending on the channel characteristics, multipath can result in either frequency selective or in some cases can be approximated as a frequency non-selective (flat) fading [4].

Compensation of multipath is possible through the use of RAKE receivers that use several baseband correlators to coherently process the multipath components and exploit their diversity advantages. In the presence of fading, the capacity of the system can be improved through multipath diversity gained by utilizing an appropriately designed RAKE receiver [4] [5]. RAKE receiver correlators can be implemented as Tapped Delay Lines (TDL) or FIR filter-like which have complex structures if the number of taps necessary to accurately model the channel is high. For example, in 3G UMTS systems [6] [7], the delay spread can be up to 75 chips in a in a 3.84 Megachip per second (Mcps) chip-rate system, thus requiring a RAKE receiver covering a delay of 75 chip duration.

However, several distinct paths typically dominate the channel, hence only a few taps, or, in other words low order models, are necessary to capture most of the

signal energy. If all taps are used in the RAKE structure, even though most do not correspond to an actual signal path, channel noise would ensure any linear estimate of their tap weights would produce non-zero weights. This leads not only to undue complexity, but also reduces system performance. By discarding those taps with a non-significant contribution to the channel model, it is possible to simultaneously reduce system complexity and improve the performance [4] [5] [8]. There has been intensive research on the performance of RAKE receivers using different channel models, system models, and analytical techniques. In [9] Subspace Maximum-Likelihood and Euclidean Norm based methods have been proposed for channel estimation, however they assume the noise distribution to be Gaussian and are also computationally expensive. A discrete time multipath channel model was used in [10], where it is assumed that the temporal separation between each multipath component is greater than the chip duration so that the signals received at each tap are mutually independent. A Tap Delay Line (TDL) channel model was used in [4] [11], under the assumption that the bandwidth of the transmitted signal is limited to the inverse of the chip duration, and the number of required taps is equal to the maximum delay of the channel divided by the chip duration. This model is inappropriate for cases where some or most of the multipath attenuation coefficients are zeros. In [12], a RAKE receiver that uses more taps with reduced tap spacing was considered, while in [13] and [14], the optimum RAKE structure was found by searching in a window of potential delays for the finger positions combination that maximizes the Signal-to-Noise Ratio (SNR). In [15], model selection criteria such as Rissanen's Minimum Description Length (MDL) [16] [17] and subspace techniques were used to estimate and identify the channel, but the techniques do not provide a level of significance. In this work we propose new algorithms and approaches for the identification and estimation of the significant channel parameters. This will include low complexity time and frequency domain based methods. For a given level of significance, we construct statistical tests to identify whether the path parameters are significant or not. Finally, the identified parameters are used to construct a parsimonious RAKE receiver structure to enhance the system performance.

1.1 Aims and Objectives

The broad aim of this thesis is to develop new approaches and algorithms for the identification and estimation of low order models that represent multipath channel effects in CDMA communication systems. Based on these low order channel models, parsimonious receivers such as RAKE receivers are considered to exploit these propagation effects and enhance the system performance. More specifically, the aims and objectives of this thesis are,

1. To model the multipath channel propagations in CDMA systems
2. To propose statistical tests for low order model selection
3. To propose low complexity algorithms to estimate the model parameters
4. To design parsimonious receivers such as RAKE receivers based on the estimated parameters.
5. To enhance the CDMA system performance which includes low Bit Error Rate (BER).

1.2 Contributions

The original contributions made in this thesis include

1. Model selection algorithms based on the MDL, AIC, F-Statistics, sphericity tests, multiple hypothesis tests and the bootstrap.
2. A frequency domain sphericity test and F-Statistics based approach for the identification and estimation of significant multipath components.
3. Time domain bootstrap based approaches for the identification and estimation of significant multipath components where there is minimal knowledge about the test statistics.
4. Low computation complexity blind adaptive channel estimation and tracking algorithms.

5. A blind based hierarchical MDL model selection method is proposed to identify non-zero parameters of the channel response.

1.3 Scope and Overview

Chapter 2. Review of Multiple Access Techniques: This chapter reviews digital modulation techniques and the evolution of mobile communication systems.

Chapter 3. DS-CDMA Signals Over Multipath Fading Channels: This chapter describes the characteristics of multipath propagation effects on CDMA systems.

Chapter 4. DS-CDMA RAKE Receivers: This chapter discusses the modeling of multipath propagation effects and the use of RAKE receivers to exploit these effects.

Chapter 5. Model Selection Techniques: This chapter covers different model selection techniques that will be used to estimate low order channel models.

Chapter 6. Channel Estimation and Identification Using Frequency Domain Sphericity Test and F-Statistics Based Approaches: This chapter presents a frequency domain approach for channel estimation combined with a sphericity test or F-Statistics for the identification of the significant multipath parameters given the statistical test distribution.

Chapter 7. Channel Estimation and Identification Using Time Domain Bootstrap Based Approaches: This chapter presents proposed time domain bootstrap approaches for the identification and estimation of the significant multipath parameters. In this approach the bootstrap is used to estimate the statistical test distribution.

Chapter 8. Blind Adaptive Channel Estimation Approaches: In this chapter we propose four low complexity blind adaptive channel estimation algorithms based on minimum variance receivers. A blind based hierarchical MDL

model selection method is also proposed to identify non-zero parameters of the channel response.

Chapter 9. Conclusions and future directions.

Chapter 2

Review of Multiple Access Techniques

2.1 Introduction

Advancement in communication and information processing technologies creates more new applications and products. In particular, the demand for wireless communication services has increased rapidly and the trend is expected to continue. Therefore, stringent requirements on the capacity of communication systems are needed in terms of the number of users a system can serve simultaneously. In other and more appropriate words, as much information as possible should be transferred. Multiple Access refers to a technique that enable sharing a common communication channel between multiple users. The freedom in use when designing multiuser communication systems include space, time and frequency. The time and frequency domains are dual of each other via the Fourier transform so that the actual options to use are, space domain and time-frequency domain designs. In the space domain users can be separated by making their distance large enough or in other words by assigning different channel for different user. An example is to use different cables for each user to separate signals in wire-lines communication. More advanced techniques include Polarization Division Multiple Access (PDMA) and Space Division Multiple Access (SDMA) [18] [19]. In

PDMA two users can be separated by using electromagnetic waves with different polarization but in SDMA sectorized antennas are usually applied to separate users at the same frequency. In time domain, different time slot can be assigned to different user where in frequency domain different frequency band can be assigned to different user and then all user's signals are linearly combined to form one signal. Spread spectrum multiple access communication is a driving technology behind the rapidly advancing personal communications industry [4] [5], each user's transmitted data is modulated by a unique signature waveform then all user's signals are linearly combined to form one signal. Various signature waveform designs result in multiple access techniques, but to retrieve each user's data again , the signature waveforms must be known at the receiver. In the following sections we review the most famous multiple access systems followed by the evolution of communication systems.

2.2 Frequency Division Multiple Access (FDMA)

The oldest multiple access technique is Frequency Division Multiple Access (FDMA). In FDMA each user's waveform occupies its own frequency band and the receiver can separate each user by simple bandpass filtering, so that FDMA can be used in analog and digital modulation depending on the input data signal. The FDMA system is shown in Fig. 2.1.

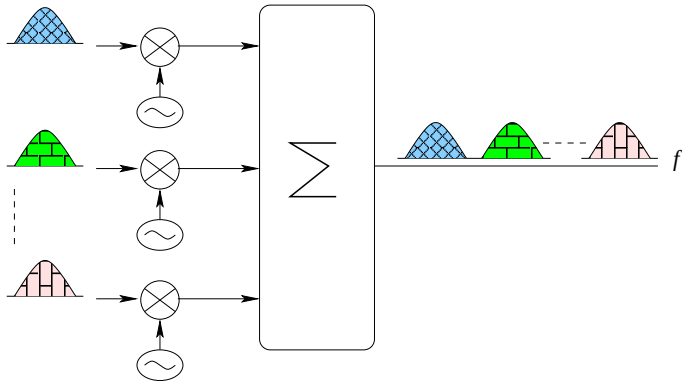


Figure 2.1: Frequency Division Multiple Access (FDMA) system

2.2.1 Advantages of FDMA

- Reducing the information bit rate and using efficient digital codes can increase the capacity.
- As FDMA systems use low bit rates (large symbol time) compared to average delay spread, they reduce the cost, and there is low Inter Symbol Interference (ISI).
- There is hardly any equalization required.
- Technological advances required for implementation are simple. A system can be configured so that improvements in terms of speech coder bit-rate reduction could be readily incorporated.
- Since the transmission is continuous, less number of bits are needed for synchronization and framing.

2.2.2 Disadvantages of FDMA

- It does not differ significantly from analog systems; capacity improvement depends on reducing signal to interference ratio, or signal to noise ratio (SNR).
- The maximum bit rate per channel is fixed and small.
- The guard bands between each two users result in wastage of capacity.
- Hardware involves narrow band filters, which cannot be realized in VLSI and thus increase cost.

2.3 Time Division Multiple Access (TDMA)

The introduction of digital modulation enabled the appearance of Time Division Multiple Access (TDMA). TDMA is digital transmission technology that allows a number of users to access a single Radio-Frequency (RF) channel without interference by allocating unique time slots to each user within each channel as shown

in Fig. 2.2.

TDMA is relatively simple to implement and is very flexible for providing variable bit rates. Increasing the bit rate can be implemented by assigning to a user more transmission intervals. However, the transmission of all users must be exactly synchronized to each other.

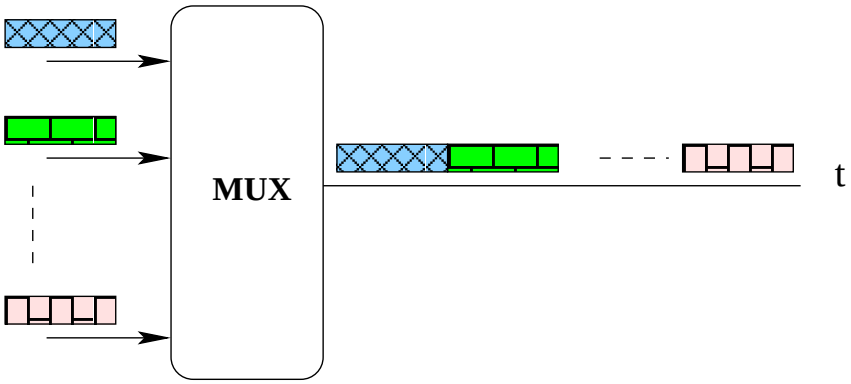


Figure 2.2: Time Division Multiple Access (TDMA) system

2.3.1 Advantages of TDMA

- Permits flexible bit rates by assigning more slots per frame to a certain user.
- Can support bursts or variable bit rate traffic.
- No guard bands required for wideband system.
- No narrowband filters required for wideband system.

2.3.2 Disadvantages of TDMA

- The high bit rates of wideband systems require complex equalization.
- Because of burst mode of operation, a large number of overhead bits for synchronization and framing are required.
- Guard time between each two users is required in each slot to accommodate time inaccuracies because of clock instability.

- Electronics operating at high bit rates increase power consumption.
- Complex signal processing is required for synchronize within a short slot time.

2.4 Code Division Multiple Access (CDMA)

The invention of spread spectrum techniques for communication systems with anti-jamming and low probability of undesired interception capabilities lead to the idea of Code Division Multiple Access (CDMA). CDMA systems can be implemented in numerous ways including Frequency Hopping (FH-CDMA), Time Hopping (TH-CDMA), Direct Sequence (DS-CDMA) spread spectrum techniques, and new systems such as Multi-Carrier (MC-CDMA) techniques.
TH-CDMA is a spread spectrum system in which the period and duty cycle of a pulsed RF carrier are varied in a pseudo random manner under the control of a coded sequence Fig. 2.3. Time hopping is often used effectively with frequency hopping to form a hybrid multiple access (TDMA) spread spectrum system.

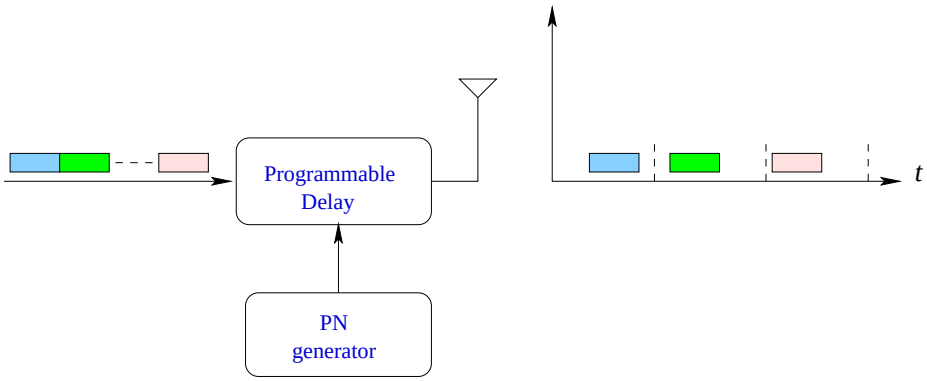


Figure 2.3: TH-CDMA system

In FH-CDMA user's signature waveforms are centered on different carrier frequencies at different time intervals, the hopping from a frequency to another is controlled according to pseudo random spreading sequences [5] [19] [20], as shown in Fig. 2.4.
In DS-CDMA systems each user's signal is modulated by a unique signature

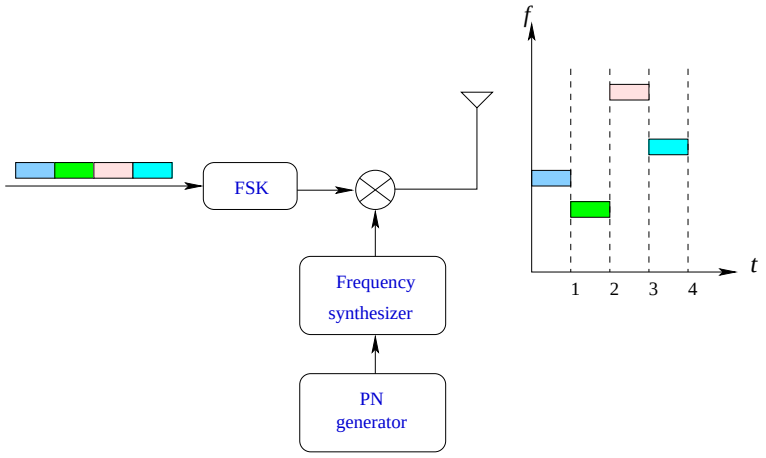


Figure 2.4: FH-CDMA system

known as spreading sequence. These signatures are continuous in the time domain and have a relatively flat spectrum, Fig. 2.5 shows an example of the system. Therefore, in DS-CDMA systems, users are separated neither in time nor in fre-

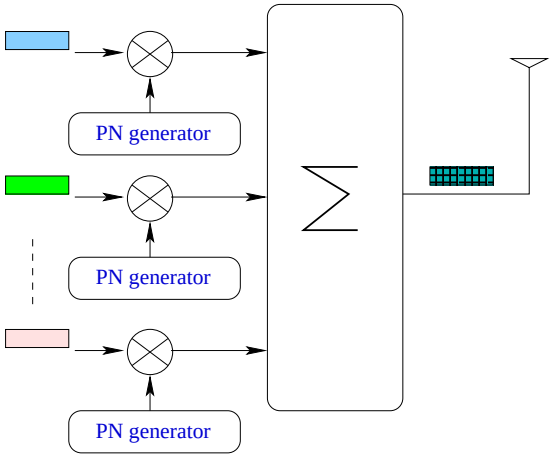


Figure 2.5: DS-CDMA system

quency, but all signature waveforms occupy the whole frequency band allocated for the transmission at all times. However the data of each user can be separated in the receivers using the same unique spreading code of that user. One of the famous receivers is known as RAKE receivers.

Fig. 2.6 describes the difference between TDMA, FDMA and CDMA systems. In multicarrier modulation MC-CDMA each user's data is first modulated by its spreading code and then transmitted using different carrier frequencies called

sub-carriers [21].

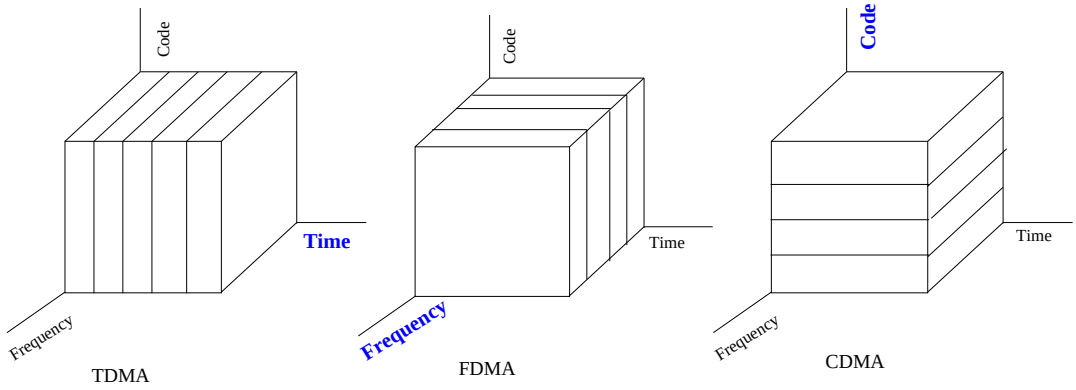


Figure 2.6: TDMA, FDMA and CDMA systems

2.4.1 Advantages of CDMA

- CDMA can support many users in the same channel, i.e., a high capacity.
- Lower mobile transmit power, i.e., longer battery life and better power control.
- Improved performance in multipath environments, RAKE receivers can be used to improve signal reception by exploiting diversity.
- Soft handoffs can be used. Mobiles can switch base stations without switching carriers. Two base stations receive the mobile signal and the mobile is receiving from two base stations.
- High peak data rates can be accommodated.
- Burst transmission - reduces interference.

2.4.2 Disadvantages of CDMA

- The code length has to be carefully selected. A large code length can induce delay or even cause interference.
- CDMA requires tight power control as it suffers for far-near effect.

- Time synchronization is necessary.
- Soft handoff increases use of radio resources and hence can reduce capacity.
- As the sum of the power received at and transmitted from a base station has to be constant, a tight power control is needed. This can result in more handoffs.

2.5 Evolution of Mobile Communication Systems

This section reviews the evolution of mobile communication systems. It will cover briefly the period between 1971 to 2005. For more details about these systems see [22] [23] [24] [25] [26] [27].

2.5.1 Zeroth generation

0G refers to pre-cellphone mobile technology, such as radio telephones that some had in cars before the advent of cellphones. One such technology is the Autoradiopuhelin (ARP) launched in 1971 in Finland as the country's first public commercial mobile phone network.

2.5.1.1 ARP

Autoradiopuhelin (ARP) or Car Radio Phone in English was the first commercially operated public mobile phone network in Finland around 1970. The system operated on 150 MHz frequency (147.9 - 154.875 MHz). Transmission power ranged from 1 watt to 5 watts. It used half-duplex transmission, meaning that receiving and transmitting voice could not happen at the same time. Being analog, it had no encryption and calls could be listened to with scanners. It started as a manually switched service, but was fully automatated in 1990 although by that time the number of subscribers had dwindled down to 980 users. ARP did not support handover, so calls would disconnect when moving to a new cell area. Cell size was approximately 30 km. ARP mobile terminals were extremely large

for the time and could only be fitted in cars' trunks, with a handset near the driver's seat. ARP was also expensive.

2.5.2 First generation

1G (or 1-G) is a short term for first-generation wireless technology, cellphones. These are the analog cellphone standards that were introduced in the 80's and continued until being replaced by 2G digital cellphones. One such standard is Nordic Mobile Telephone (NMT), used in Nordic countries, Eastern Europe and Russia. Another is Advanced Mobile Phone System (AMPS) used in the United States.

2.5.2.1 NMT

Nordisk Mobil Telefon (Nordic Mobile Telephony in English) is a mobile phone system that was created in 1981 as a response to the increasing congestion and heavy requirements of the ARP mobile phone network. It is based on analog technology (first generation or 1G) and two variants exist: NMT450 and NMT900. The numbers indicate the frequency bands uses. NMT900 was introduced in 1986 because it carries more channels than the previous NMT450 network.
Compared to the previous ARP network, NMT had automatic switching built into the standard from the beginning. Additionally, the NMT standard specified billing and roaming. A disadvantage of the original NMT specification is that traffic was not encrypted. So anyone willing to listen in would just have to buy a scanner and tune it to the correct frequency but later versions of the NMT specifications defined optional analog encryption which was based on two-band audio frequency inversion. NMT also supported a primitive data transfer mode called DMS or NMT-Text, which used the network's signalling channel for data transfer. Transfer speeds vary between 600 and 1200 bits per second, using Fast Frequency Shift Keying (FFSK) modulation. Another data transfer mode was called NMT Mobidigi with transfer speeds of 380 bits per second.

2.5.2.2 AMPS

Advanced Mobile Phone System or AMPS is the analog mobile phone system standard, introduced in the Americas during the early 1980s. It was a first-generation technology, using FDMA but due to security reasons it was later replaced by the newer Digital TDMA system which brought improved security as well as increased capacity.

For the systems frequency bands, each network is authorized to use 416 channels in the 800 MHz band. Each channel is composed of 2 frequencies. 416 of these are in the 824-849 MHz range for transmissions from mobile stations to the base stations, paired with 416 frequencies in the 869-894 MHz range for transmissions from base stations to the mobile stations. Each cell site uses a subset of these channels, and must use a different set than neighboring cells to avoid interference. This significantly reduces the number of channels available at each site in real-world systems. The Band Width (BW) of each AMPS frequency is 30kHz wide.

2.5.2.3 TAC

Total Access Communication System or TACS is the European version of AMPS. ETACS was an extended version of TACS with more channels. TACS and ETACS are now obsolete in Europe, having been replaced by the more scalable and all-digital GSM system.

2.5.3 Second generation

2G (or 2-G) is a short term for second-generation wireless telephone technology. It cannot normally transfer data, such as email or software, other than the digital voice call itself, and other basic ancillary data such as time and date. Nevertheless, Short Message Service (SMS) messaging is also available as a form of data transmission for some standards. 2G services are frequently referred as Personal Communications Service or PCS in the US. 2G technologies can be divided into TDMA-based and CDMA-based standards depending on the type of multiplexing

used. The main 2G standards are:

- GSM, TDMA-based
- iDEN, TDMA-based
- IS-136 (D-AMPS), TDMA-based
- IS-95 (cdmaOne), CDMA-based
- PDC (TDMA-based)

2.5.3.1 GSM

Global System for Mobile Communications (GSM) is the most popular standard for mobile phones in the world introduced around 1982 and used from 1990. GSM is a cellular network, which means that mobile phones connect to it by seaching for cells in the immediate vicinity. GSM networks can operate at various radio frequencies. Currently it uses dual, Tri and quad bands. GSM phones are used by over a billion people across more than 200 countries. The ubiquity of the GSM standard makes international roaming very easy with roaming agreements between mobile phone operators. GSM differs significantly from its predecessors in that both signalling and speech channels are digital, which means that it is seen as 2G mobile phone system. This fact has also meant that data communication was built into the system from very early on. GSM is an open standard which is developed by the 3rd Generation Partnership Project (3GPP).
One of the key features of GSM is the Subscriber Identity Module (SIM), commonly known as a SIM card. The card is a detachable smartcard containing the user's subscription information and phonebook. This allows the user to retain his information while switching handsets. Alternatively, the user can also change operators while retaining the handset simply by changing the SIM. Some operators will block this by allowing the phone to use only a single SIM card, or only a SIM card issued by them; this practice is known as SIM locking. There are four different cell sizes in a GSM network macro, micro, pico and umbrellacells. The coverage area of each cell is different in different environments. Macro cells

can be regarded as cells where the base station antenna is installed in a mast or a building above the average roof top level. However, micro cells are cells where the antenna height is under the average roof top level and they are typically used in urban areas. Picocells are small cells whose diameter is a few dozen metres and are mainly used indoors. On the other hand, umbrellacells are used to cover shadow regions of smaller cells and fill in gaps in coverage between those cells.

2.5.3.2 iDEN

Integrated Digital Enhanced Network (iDEN) is a mobile communications technology, developed by Motorola, which provides its users the benefits of a trunked radio and a cellular phone based on a TDMA digital systems. The iDEN has six communication channels that share a 25 kHz space, some competing technologies place only one channel in 12.5 kHz. The new generation of the iDEN systems support voice communication, paging, text messaging and a medium speed downloads, but there is no path for high speed wireless data.

2.5.3.3 D-AMPS

IS-54 and IS-136 are 2G mobile phone systems, known as Digital AMPS (D-AMPS). It is used throughout the United States and Canada. D-AMPS is known as a TDMA based system but currently the system is being replaced by new systems such as GSM/GPRS and cdma2000 technologies.

D-AMPS uses existing AMPS channels and allows for smooth transition between digital and analog systems in the same area. Capacity was increased over the preceding analog design by dividing each 30 kHz channel pair into three time slots and digitally compressing the voice data, yielding three times the call capacity in a single cell. A digital system also made calls more secure because analog scanners could not access digital signals. Calls were encrypted, although the algorithm used (CMEA) was later found to be weak.

IS-136 added a number of features to the original IS-54 specification, including text messaging, Circuit Switched Data (CSD), and an improved compression protocol. SMS and CSD were both available as part of the GSM protocol, and

IS-136 implemented them in a nearly identical fashion.

2.5.3.4 cdmaOne

Interim Standard 95 (IS-95), is the first CDMA-based digital cellular standard pioneered by Qualcomm. The brand name for IS-95 is cdmaOne. IS-95 is also known as TIA-EIA-95. It is now being supplanted by IS-2000. The system is widely used around the world specially in the US and Korea. The CDMA technology used in IS-95 is in technical competition with the TDMA technology used in GSM.

2.5.3.5 PDC

Personal Digital Cellular (PDC) is a 2G mobile phone based on TDMA standard developed and used exclusively in Japan. PDC uses 25 kHz carrier, 3 time slots, $\pi/4$-DQPSK modulation and low bit-rate 11.2 kbit/s and 5.6 kbit/s (half-rate) voice codecs. PDC is implemented in the 800 MHz (downlink 810-888 MHz, uplink 893-958 MHz), and 1.5 Ghz (downlink 1477-1501 MHz, uplink 1429-1453 MHz) bands. The services include voice (full and half-rate), supplementary services (call waiting, voice mail, three-way calling, call forwarding , and so on), data service (up to 9.6 kbit/s CSD), and packet-switched wireless data (up to 28.8 kbit/s PDC-P).

2.5.3.6 2.5G

2.5G is a stepping stone between 2G and 3G cellular wireless technologies. The term "second and a half generation" is used to describe 2G-systems that have implemented a packet switched domain in addition to the circuit switched domain. It does not necessarily provide faster services because bundling of timeslots is used for circuit switched data services (HSCSD) as well. While the terms 2G and 3G are officially defined, 2.5G is not. It was invented for marketing purposes only. 2.5G provides some of the benefits of 3G (e.g. it is packet-switched) and can use some of the existing 2G infrastructure in GSM and CDMA networks. The commonly known 2.5G technique is GPRS.

- (GPRS)

  General Packet Radio Service (GPRS) is a mobile data service which was available to users of GSM mobile phones around 1997. It provides moderate speed data transfer, by using unused TDMA channels in the GSM network. GPRS is packet-switched which means that multiple users share the same transmission channel, only transmitting when they have data to send. This means that the total available bandwidth can be immediately dedicated to those users who are actually sending at any given moment, providing higher utilization where users only send or receive data intermittently. Web browsing, receiving e-mails as they arrive and instant messaging are examples of uses that require intermittent data transfers, which benefit from sharing the available bandwidth. A consequence of this is that packet switched data has a poor bit rate in busy cells. The theoretical limit for packet switched data is approx. 170 kbit/s. A realistic bit rate is 30-70 kbit/s. GPRS upgrades GSM data services by providing:

  - Point-to-point (PTP) service: internetworking with the Internet (IP protocols) and X.25 networks.
  - Point-to-multipoint (PT2MP) service: point-to-multipoint multicast and point-to-multipoint group calls.
  - SMS
  - Anonymous service: anonymous access to predefined services.
  - Future enhancements: flexible to add new functions, such as more capacity, more users, new accesses, new protocols, new radio networks

2.5.3.7 2.75G

The term 2.75G has not been officially defined anywhere, but as of 2004 is beginning to be used quite often in media reports. 2.75G is the term which has been decided on for systems which do not meet the 3G requirements but are marketed as if they do (e.g. CDMA-2000 without multi-carrier) or which do, just, meet the requirements but are not strongly marketed as such. (e.g. EDGE systems).

- EDGE

  Enhanced Data Rates for GSM Evolution (EDGE) is a digital mobile phone technology which acts as a bolt-on enhancement to 2G and 2.5G (GPRS) networks. This technology works in TDMA and GSM networks. EDGE uses the same spectrum allocated for GSM850, GSM900, GSM1800 and GSM1900 operation.

  In addition to Gaussian minimum-shift keying (GMSK), EDGE uses 8PSK (8 Phase Shift Keying) for its upper five of the nine modulation and coding schemes. EDGE produces a 3bit word for every change in carrier phase. This effectively triples the gross data rate offered by GSM. EDGE, like GPRS, uses a rate adaptation algorithm that adapts the modulation and coding scheme (MCS) used for the quality of the radio channel, and thus the bit rate and robustness of data transmission. It introduces a new technology not found in GPRS, Incremental Redundancy, which, instead of retransmitting disturbed packets, sends more redundancy information to be combined in the receiver. This increases the probability of correct decoding.

2.5.4 Third generation

3G (or 3-G) is short for third-generation mobile phone technology. 3G systems have a high degree of commonality of design worldwide, compatibility of services. They also allow use of many application such as:

- Small pocket terminals with worldwide roaming capability
- Internet and other multimedia applications
- Wide range of services and terminals.
- The ability to transfer both voice data such as phone call and non-voice data such as downloading information, exchanging email, and instant messaging.

Originally, 3G was supposed to be a single, unified, worldwide standard, but in practice, the 3G world has been split into three camps.

2.5.4.1 W-CDMA

Wideband Code Division Multiple Access (W-CDMA), a wideband spread-spectrum 3G mobile communication technology that utilizes CDMA, is a 3G mobile communications standard allied with the GSM standard. W-CDMA uses 5 MHZ channels according to the International Telecommunication Union (ITU) and the International Mobile Telecommunications-2000 IMT-2000 standards. The system is more than a multiplexing standard, it is a complete set of specifications, a detailed protocol that defines how a mobile phone communicates with the tower, how signals are modulated, how data are structured.

- UMTS

  Universal Mobile Telecommunications System (UMTS) or sometime marked as 3GSM is one of the 3G mobile phone technologies. UMTS is the European 3G mobile communications standard. It uses W-CDMA as the underlying standard by the 3GPP.

  UMTS offers mobile operators significant capacity and broadband capabilities to support greater numbers of voice and data customers especially in urban centers plus higher data rates at lower incremental cost than 2G. Making use of radio spectrum in bands identified by the ITU for 3G IMT-2000 mobile services and subsequently licensed to operators, UMTS employs a 5 MHz channel carrier width to deliver significantly higher data rates and increased capacity compared with second generation networks. This 5 MHz channel carrier provides optimum use of radio resources, especially for operators who have been granted large, contiguous blocks of spectrum typically ranging from 2x10 MHz up to 2x20 MHz to reduce the cost of deploying 3G networks.

  UMTS has been specified as an integrated solution for mobile voice and data with wide area coverage. Universally standardized via the 3GPP and using globally harmonized spectrum in paired and unpaired bands, UMTS in its initial phase offers theoretical bit rates of up to 384 kbps in high mobility situations, rising as high as 2 Mbps in stationary/nomadic user

environments. In the early standardization of 3GPP, several different chip rates were considered. These included multiples of the basic chip rate 1x, 2x and 4x or approximately 4 Mcps, 8 Mcps and 16Mcps. The limited amount of spectrum available in the core UMTS bands forced a choice of the lowest chip rate which ultimately became 3.84 Mcps. However it was the assumption that higher chip rates would not be precluded from future.

The specific frequency bands originally defined by the UMTS standard are 1885-2025 MHz for uplink and 2110-2200 MHz for downlink. Symmetry between uplink and downlink data rates when using paired (FDD) spectrum also means that UMTS is ideally suited for applications such as real-time video telephony, in contrast with other technologies such as ADSL where there is a pronounced asymmetry between uplink and downlink throughput rates. Specified and implemented as an end-to-end mobile system, UMTS also features the additional benefits of automatic international roaming plus integral security and billing functions, allowing operators to migrate from 2G to 3G while retaining many of their existing back-office systems. Offering increased capacity and speed at lower incremental cost compared with second generation mobile systems.

- FOMA

  FOMA or Freedom of Mobile Multimedia Access, is the brand name for the 3G services being offered by Japanese mobile phone operator NTT DoCoMo. FOMA was the world's first WCDMA 3G service when launched in 2001. However, FOMA's variant of the technology is currently incompatible with standard UMTS, and hence does not provide global roaming. Originally FOMA handsets were big, had poor battery life and the network had poor coverage. Unsurprisingly it was not widely used. With the introductions of attractive handsets and better coverage, the use of FOMA based technologies has increased and as of January 2005, FOMA has close to 10 million subscribers and is the fastest growing cellphone network in Japan.

2.5.4.2 CDMA2000

CDMA2000 is a 3G mobile telecommunications standard, one of the approved radio interfaces for the ITU's IMT-2000 standard, and a successor to 2G CDMA (IS-95, branded cdmaOne). The underlying signaling standard is known as IS-2000. CDMA2000 is an incompatible competitor of the other major 3G standard W-CDMA. CDMA2000 is a registered trademark of the Telecommunications Industry Association (TIA-USA) in the United States, not a generic term like CDMA. TIA has branded their 2G CDMA standard (IS-95) as cdmaOne. There are many different types of CDMA2000. In order of increasing complexity:

- CDMA2000 1x (3G1x or 1xRTT), 1.25 MHz radio channels. It supports up to 144kbit/s
- CDMA2000 1xEV, a High Data Rate capability (HDR). It supports up to 3.1 Mbit/s for downloading uplink and 1.8 Mbit/s for uplink.
- CDMA2000 3x (MC-CDMA) a 3.75 MHz radio channels. Very good system, but has not been deployed yet.

2.5.4.3 TD-SCDMA

Time Division Synchronous Code Division Multiple Access (TD-SCDMA) is a 3G mobile telecommunications standard, being pursued in the People's Republic of China by the Chinese Academy of Telecommunications Technology (CATT), Datung and Siemens, in an attempt to develop home-grown technology and not be dependent on Western technology. It is based on spread spectrum CDMA technology. The launch of an operational system is projected by 2005.

2.5.4.4 3.5G

High-Speed Downlink Packet Access or (HSDPA) is a mobile telephony protocol also called 3.5G. High Speed Downlink Packet Access (HSDPA) is a packet-based data service in W-CDMA downlink with data transmission up to 8-10 Mbit/s (and 20 Mbit/s for MIMO systems) over a 5MHz bandwidth in WCDMA

downlink. HSDPA implementations includes Adaptive Modulation and Coding (AMC), multiple Input multiple Output (MIMO), Hybrid Automatic Request (HARQ), fast cell search, and advanced receiver design. In 3GPP standards, Release 4 specifications provide efficient IP support enabling provision of services through an all-IP core network and Release 5 specifications focus on HSDPA to provide data rates up to approximately 10 Mbit/s to support packet-based multimedia services. MIMO systems are the work item in Release 6 specifications, which will support even higher data transmission rates up to 20 Mbit/s. HSDPA is evolved from and backward compatible with Release 99 WCDMA systems.

2.5.5 Fourth Generation

4G (or 4-G) is a short term for fourth-generation, the successor of 3G and is a wireless access technology. It describes two different but overlapping ideas.

- High-speed mobile wireless access with a very high data transmission speed, of the same order of magnitude as a local area network connection (10 Mbits/s and up). It has been used to describe wireless LAN technologies like Wi-Fi, as well as other potential successors of the current 3G mobile telephone standards.

- Pervasive networks. An amorphous and presently entirely hypothetical concept where the user can be simultaneously connected to several wireless access technologies and can seamlessly move between them. These access technologies can be Wi-Fi, UMTS, EDGE or any other future access technology. Included in this concept is also smart-radio technology to efficiently manage spectrum use and transmission power as well as the use of mesh routing protocols to create a pervasive network.

Chapter 3

DS-CDMA Signals Over Multipath Fading Channels

3.1 Introduction

Fading and multipath occur in many radio communication systems. In mobile communication systems, the mobile or the base station unit is often surrounded by various objects, such as buildings, trees, etc. These objects produce more than one path over which the signal can travel between the transmitter and the receiver.

In a multipath situation, the signals arriving along different paths will have different attenuation, delays and Direction Of Arrivals (DOA). They might add at the receiving antenna either constructively or destructively depending on their phases, resulting in phenomena known as fading [4] [8]. If the path length and/or the geometry change due to variations in the transmission medium or due to relative motion between the transmitter and receiver antennas, the signal level will exhibit some changes subjected to random fluctuations.

To become familiar with the effects of the multipath propagation, let us consider the following situations.

If we transmit say a short duration pulse to represent a digital sequence, due to reflections, the received signal may appear as a number of pulses with dif-

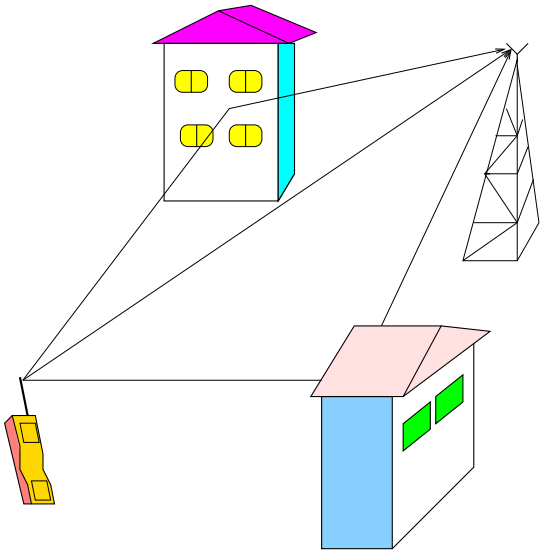


Figure 3.1: A simple multipath channel

ferent attenuation and delays. Also, if we repeat and transmit another pulse, attenuation and delays will be different from the previous case and the received number of pulses does quite often changes, see Fig. 3.2. Hence the first two characteristic of the channel are the time spread introduced in the signal and the channel medium structure is also time varying. Moreover, the time variation of the channel is unpredictable to the channel user. In some situations and due to the nature of the reflecting objects, the duration of the received pulses may also change and this will create other problems such as scattering, dispersion and frequency distortion [4] [8].

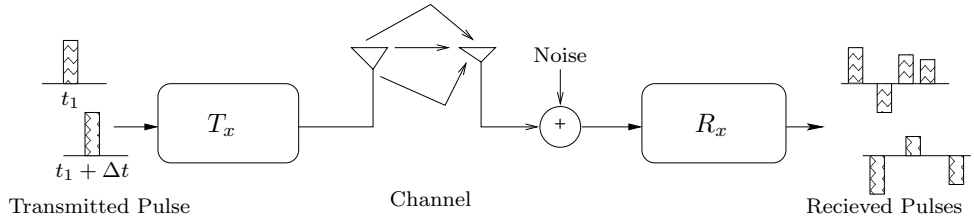


Figure 3.2: Multipath effects

3.2 Characteristics of Multipath Propagation

In this section we study the characteristics of the multipath propagation from signal processing point of view. Hence, let us consider the situation where a signal $x(t)$ is transmitted over a multipath channel. For continuous multipath propagation, the equivalent base-band response of the channel define $g(\tau, t)$ can be defined as a function of two variables, τ and t. The variable τ is considered to represent the effect of the channel delay and t is to represent the time variation nature. The received signal $r(t)$ as shown in Fig. 3.3, can then be defined as,

$$r(t) = \int_{-\infty}^{\infty} g(\tau, t)x(t-\tau)\, d\tau + v(t) \tag{3.1}$$

where $v(t)$ represents the channel noise.

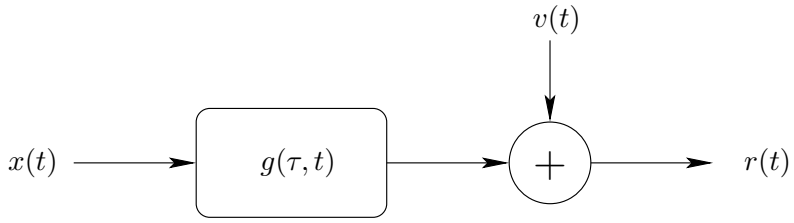


Figure 3.3: General system model for multipath fading channel

Due to the random nature of $g(\tau, t)$, in practice the channel effects can be characterized by four important factors: the multipath delay spread, T_m, the coherence time, Δt_c, the coherence bandwidth, Δf_c, and finally the Doppler spread, β_d. Given information about these factors, one can assume the situation over which the transmitted signal $x(t)$ can be affected then a specific or an approximated channel model can be considered [4] [5] [18].

To become familiar with the meaning of each of the above factors, we will briefly summarize some important functions of the channel $g(\tau, t)$.

The channel response, $g(\tau, t)$ is characterized in [4] as a wide-sense-stationary complex-valued random process with the autocorrelation function,

$$\varphi(\tau_1; \tau_2, \Delta t) = \frac{1}{2}\mathsf{E}\{g^*(\tau_1, t)g(\tau_2, t + \Delta t)\}\,. \tag{3.2}$$

In most radio transmission media, the attenuation and phase shift of the channel associated with the delay τ_1 is uncorrelated with the one presented at delay τ_2, then it follows from the above equation that,

$$\frac{1}{2}\mathsf{E}\{g^*(\tau_1,t)g(\tau_2,t+\Delta t)\} = \varphi(\tau_1-\tau_2,\Delta t).\delta(\tau_1-\tau_2) \tag{3.3}$$

If we replace $\Delta t = 0$, the resulting autocorrelation function, $\varphi(\tau') = \varphi(\tau',0)$ is known as the *multipath intensity profile* of the channel. The range of values of τ over which $\varphi(\tau') \neq 0$ is called the *multipath delay spread* of the channel and denoted by T_m (see Fig. 3.4).

Define the Fourier transform $G(f,t)$ of the channel response $g(\tau,t)$ with respected to (w.r.t) the delay τ,

$$G(f,t) = \int_{-\infty}^{\infty} g(\tau,t)\ e^{-j2\pi f\tau} d\tau \tag{3.4}$$

Since the Fourier transform will not change the wide-sense-stationary properties, then we can also define another function know as the *spaced-time spaced-frequency correlation function* of the channel,

$$\phi(\Delta f,\Delta t) = \frac{1}{2}\mathsf{E}\{G^*(f,t)G(f+\Delta f,t+\Delta t)\}\,. \tag{3.5}$$

After some manipulations one can show that $\phi(\Delta f,\Delta t)$ is the Fourier transform of the *multipath intensity profile* function, $\varphi(\tau',\Delta t)$ w.r.t τ' (see Fig. 3.4).

Using some practical measurements [4], it has been found that the relation between the coherence bandwidth Δf_c and the multipath delay spread T_m can be approximated by,

$$\Delta f_c \approx \frac{1}{T_m}. \tag{3.6}$$

In order to see the effect of the time variation of the channel, define the Fourier transform of $\phi(\Delta f,\Delta t)$ w.r.t the variable Δt to be the function $S(\Delta f,\lambda)$,

$$S(\Delta f,\lambda) = \int_{-\infty}^{\infty} \phi(\Delta f,\Delta t)\ e^{-j2\pi\lambda\Delta t} d\Delta t \tag{3.7}$$

This function gives us information about the relation between the channel time variation and the channel Doppler effects. With $\Delta f = 0$ the function $S(\lambda) =$

$S(0, \lambda)$ is known as the *Doppler power spectrum* of the channel (see Fig. 3.4). The range of values of λ over which $S(\lambda) \neq 0$ is called the Doppler spread of the channel, β_d. The reciprocal of β_d is a measure of another factor known as the coherence time of the channel,

$$\Delta t_c \approx \frac{1}{\beta_d}. \tag{3.8}$$

A summary of the previous channel factors is as follows,

- T_m : The multipath delay spread of the channel
- $\Delta f_c \approx \frac{1}{T_m}$: The coherence bandwidth of the channel
- β_d : The Doppler spread of the channel
- $\Delta t_c \approx \frac{1}{\beta_d}$: The coherence time of the channel

According to the above channel factors one can assume the situation over which the transmitted signal can be modelled then a proper channel model can be considered. In the following sections we summarize some of these situations.

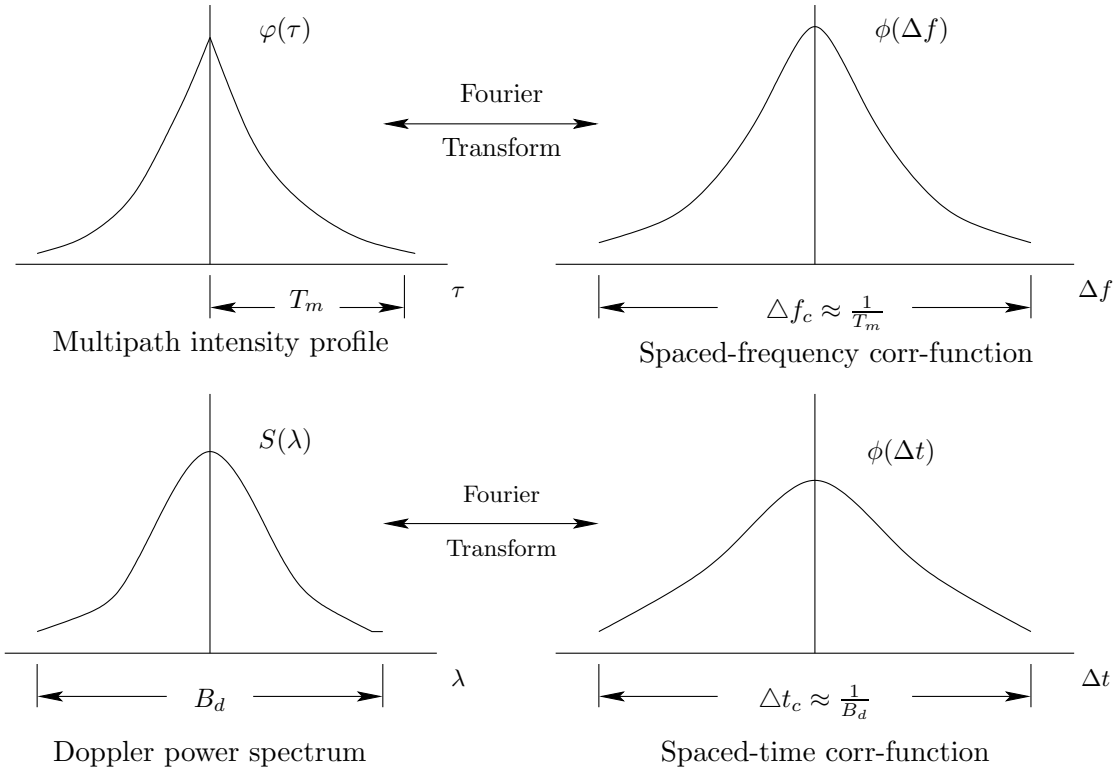


Figure 3.4: Characteristic functions of multipath fading channel

3.3 Frequency Variations of Multipath Fading Channels

3.3.1 Frequency non-selective fading channels

When the bandwidth of the transmitted signal, W, is smaller than the coherence bandwidth of the channel Δf_c i.e.

$$W < \Delta f_c \tag{3.9}$$

the channel in this case is defined as a frequency non-selective fading channel. Furthermore, the condition that the channel fades slowly implies that the multiplicative process or the channel effects can be regarded constant during at least one or few signaling intervals. In other words, the channel characteristics are the same over all the spectrum. In this case the signal will suffer simple amplitude changes, but there is no delay distortions i.e. no Inter-Symbol Interference (ISI).

3.3.2 Frequency selective fading channels

When the bandwidth of the transmitted signal, W is larger than the coherence bandwidth of the channel, Δf_c i.e.

$$W > \Delta f_c, \tag{3.10}$$

the channel is defined as frequency selective fading channel. This case happens due to motion of the mobile unit or changing of the path length due to motion of the object that the reflected signals come from. The channel characteristics in this case are different over all the spectrum and in some cases can be represented by time-varying random variables. Also, the signal will suffer amplitude distortion and there will be delay distortions i.e. ISI.

3.4 Time Variations of Multipath Fading Channels

3.4.1 Flat fading channels

When the pulse duration of the transmitted signal, T, is larger than the coherence time of the channel, Δt_c or in other words the Doppler spread is small compared to the system rate, i.e.,

$$T > \Delta t_c, \tag{3.11}$$

the channel is defined as a flat or slowly fading channel. In other words, all spectral components are attenuated equally.

3.4.2 Fast fading channels

When the pulse duration of the transmitted signal, T is smaller than the coherence time of the channel, Δt_c i.e.,

$$T < \Delta t_c, \tag{3.12}$$

the channel is defined as a fast fading channel or sometimes termed as time-selective fading channel. In this case, the time variation of the channel has to be taken into account.

3.5 Modeling Multipath Fading Channels

There are several models for multipath fading channels, each model depends on whether the signal is transmitted over a frequency selective, non-selective, fast or slowly fading channel.

Let us consider the general case system in Eqn. 3.1 and shown in Fig. 3.3 where the signal at the receiver $r(t)$ can be represented as,

$$r(t) = \int_{-\infty}^{\infty} g(\tau, t)x(t-\tau)\, d\tau + v(t) \tag{3.13}$$

Although radio transmissions are often continuous-time passband signals centered around the carrier frequencies, it is often conceptually convenient to work with

an equivalent discrete-time baseband model, i.e.

$$r(n) = \sum_{l} g(l,n)x(n-l) + v(n). \tag{3.14}$$

Accurate and tractable channel modeling is critical for reliable and efficient signal processing algorithms for transceiver design. The wireless channel is particularly challenging in this regard, since the orientation and material properties of the obstacles between transmitter and receiver are unknown in advance, or may be time varying. It is common to use statistical models for the fading channels in system design and performance evaluations. Our work is based on the well known frequency selection slowly Rayleigh fading channel model, where the fading coefficient $g(\tau,t)$ is modelled as zero mean, complex Gaussian random variable with variance σ_g^2 , which is in general a function of τ and t. The phase of $g(\tau,t)$ is therefore uniformly distributed in $[0,2\pi)$ and the amplitude $|g(\tau,t)|$ is a Rayleigh distributed Random Variable (RV) with density,

$$p_{|g|}(x) = \frac{2x}{\sigma_g^2}\, e^{\frac{-x^2}{\sigma_g^2}}. \tag{3.15}$$

3.5.1 The channel response as an FIR filter

In some situations where the channel time variation is not presenting or slow, one can then regard the channel response $g(\tau,t)$ as $g(\tau)$ over a certain interval $t \in [0,T]$. This assumption can be considered for situations such as: flat fading, frequency non-selective and frequency selective slowly fading channel. In these cases, one can model the channel response as a TDL or FIR filter then based on this model a proper receiver can be considered to overcome the channel effects. Over the interval $[0,T]$, the received signal in (3.14) can then be simplified by,

$$r(n) = \sum_{l} g(l)x(n-l) + v(n). \tag{3.16}$$

3.6 Simulation of Multipath Fading Channels

The channel characteristics can be represented by a complex time-varying random variable, $g(t) = g(\tau_o, t)$. In some cases the channel is called time non-selective

channel. There are two famous models for this case, a complex Gaussian model and sum of sinusoids (SOS) Jakes model.

3.6.1 Gaussian model

Gaussian model is a simple one where two independent normally distributed random variables can be generated then passed through a low pass filter whose power spectral density (PSD) is given by,

$$|H(f)|^2 = \begin{cases} \frac{E_p}{4\pi f_m} \frac{1}{\sqrt{1-\left(\frac{f-f_c}{f_m}\right)}} & |f - f_c| \leq f_m \\ 0 & \text{otherwise} \end{cases} \tag{3.17}$$

where,

E_p is the channel power, f_c is the signal carrier frequency and f_m is the channel maximum Doppler spread. The model block diagram is shown in Fig. 3.5.

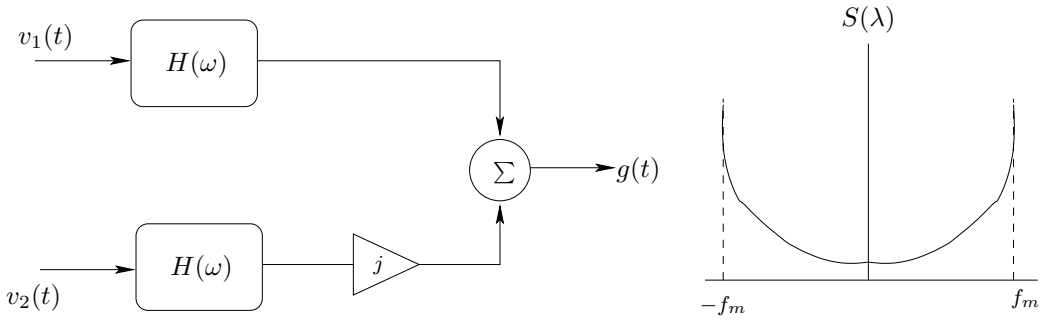


Figure 3.5: Gaussian model for multipath fading channel

It has been shown in [28, 29] that this model can be approximated by an AR(2) process,

$$g(t) = g_I(t) + jg_Q(t) \tag{3.18}$$

where,

$$\begin{aligned} &g_I(t) = a_1 g_I(t-1) + a_2 g_I(t-2) + v_1(t) \\ &\text{and} \\ &g_Q(t) = a_1 g_Q(t-1) + a_2 g_Q(t-2) + v_2(t) \end{aligned} \tag{3.19}$$

and $a_1 = 2r_d \cos(2\pi f_m T)$, $a_2 = -r_d^2$ which reflects the damping, $v_1(t)$ and $v_2(t)$ are two independent zero mean Gaussian noise.

3.6.2 Jakes model

The Jakes' model [30] is a sum of sinusoids (SOS) based fading channel simulator, which implements the channel as a superposition of a finite number of sinusoids. It has been widely used to approximate isotropic scattering environments. In general, a frequency selective and non-selective fading process of the based model is given by,

$$g(t) = \sqrt{\frac{2}{N}} \sum_{n=1}^{N} \cos\left(\omega_m t \cos(\alpha_n) + \theta_n\right) \tag{3.20}$$

where, $\omega_m = \frac{v}{c}\omega_c$ represents the maximum Doppler spread, v is the mobile speed, ω_c is the carrier frequency, and finally α_n and θ_n are two independent random variable uniformly distributed between $[0, 2\pi)$ representing the angle of incoming ray and the initial phase associated with the nth propagation path.

To reduce the number of distinct Doppler frequency shifts, Jakes uses the following parameters,

$$\begin{aligned} \alpha_n &= \tfrac{2\pi n}{N},\ n = 1, 2, \ldots, N \\ N_o &= \tfrac{N-2}{4} \end{aligned} \tag{3.21}$$

Thus the base-band equivalent fading channel can be constructed through only $(N_o + 1)$ low-frequency oscillators. With Jakes' simulator, the output signal in terms of quadrature components can be expressed as

$$g(t) = g_I(t) + g_Q(t) \tag{3.22}$$

where

$$\begin{aligned} g_I(t) &= \tfrac{2}{\sqrt{2}} \cos(\omega_m t + \theta_{N_o+1}) \cos(\alpha) + 2 \sum_{n=1}^{N_o} \cos(\omega_n t + \theta_n) \cos(\beta_n) \\ &\text{and} \\ g_Q(t) &= \tfrac{2}{\sqrt{2}} \cos(\omega_m t + \theta_{N_o+1}) \sin(\alpha) + 2 \sum_{n=1}^{N_o} \cos(\omega_n t + \theta_n) \sin(\beta_n) \end{aligned} \tag{3.23}$$

where, $f_n = f_m cos(2\pi n/N),\ n = 1, 2, \ldots, N_o$ represents the Doppler shift of the nth component, $\alpha = 0$ or $\pi/4$ and $\beta_n = \pi n/(N_o + 1),\ n = 1, 2, \ldots, N_o$. The complete model block diagram is shown in Fig. 3.6. In Fig 3.7 a typical simulated fading envelope is obtained by choosing $N_o = 16$, $f_c = 1$ GHZ for frequency selective Raleigh fading where the speed $v = 50$ Km/h and frequency

non-selective slowly fading channels where $v = 1$ Km/h. Another simulation for frequency selective slowly fading channel is also shown in Fig 3.8 for $v = 10$ Km/h.

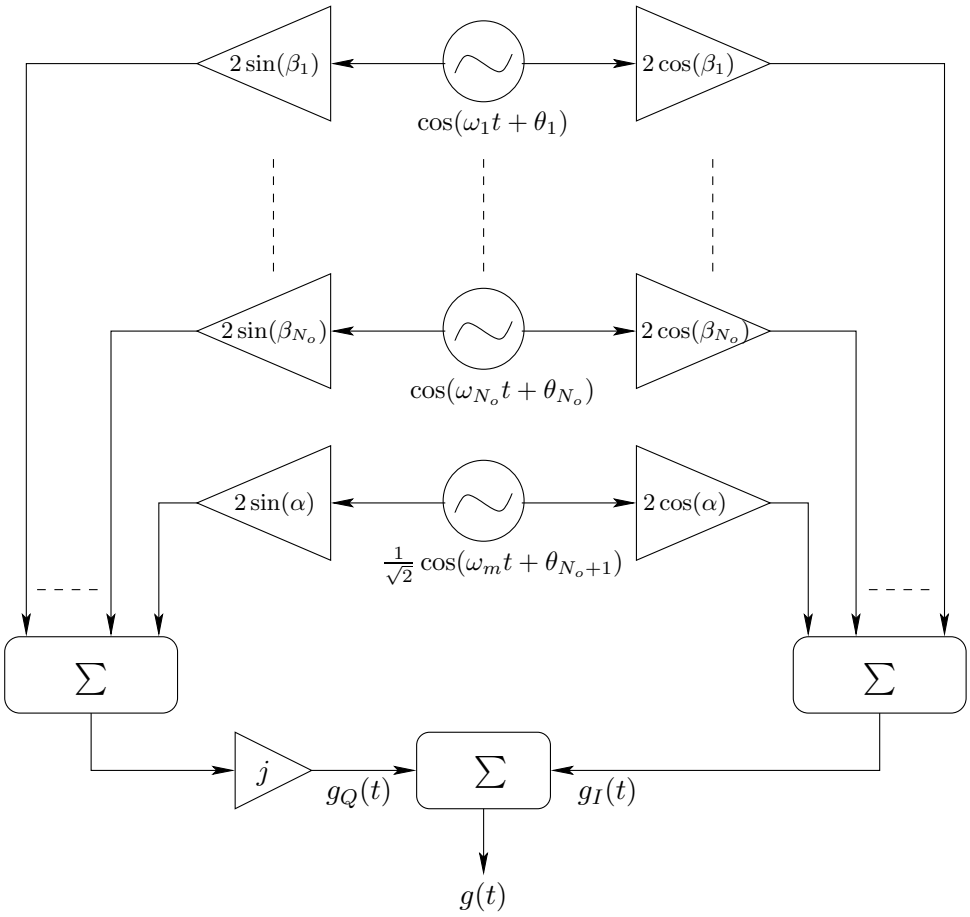


Figure 3.6: Jakes model for multipath fading channel

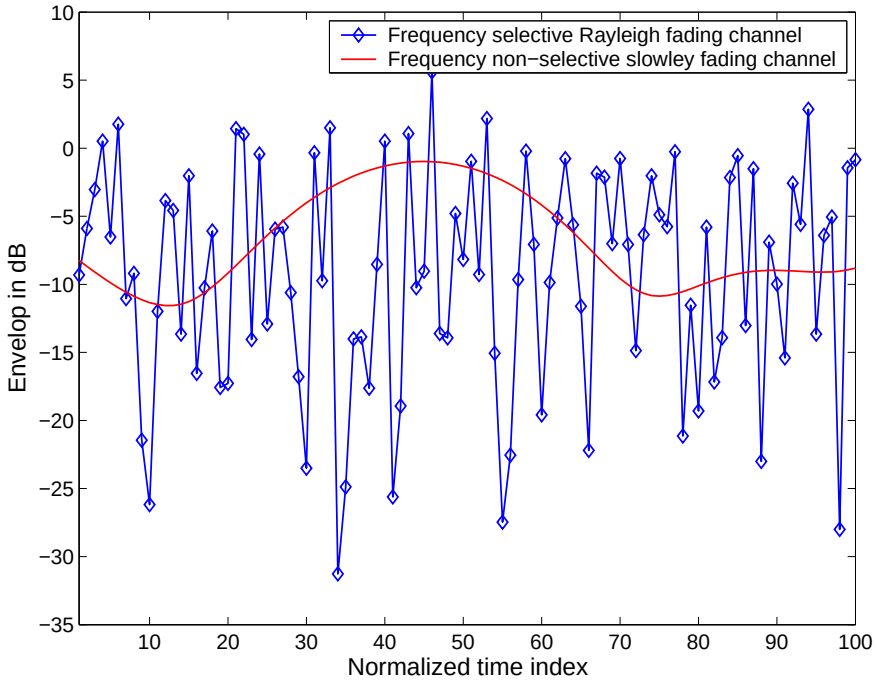


Figure 3.7: Faded envelope generated by using sum-of-sinusoids Jakes model

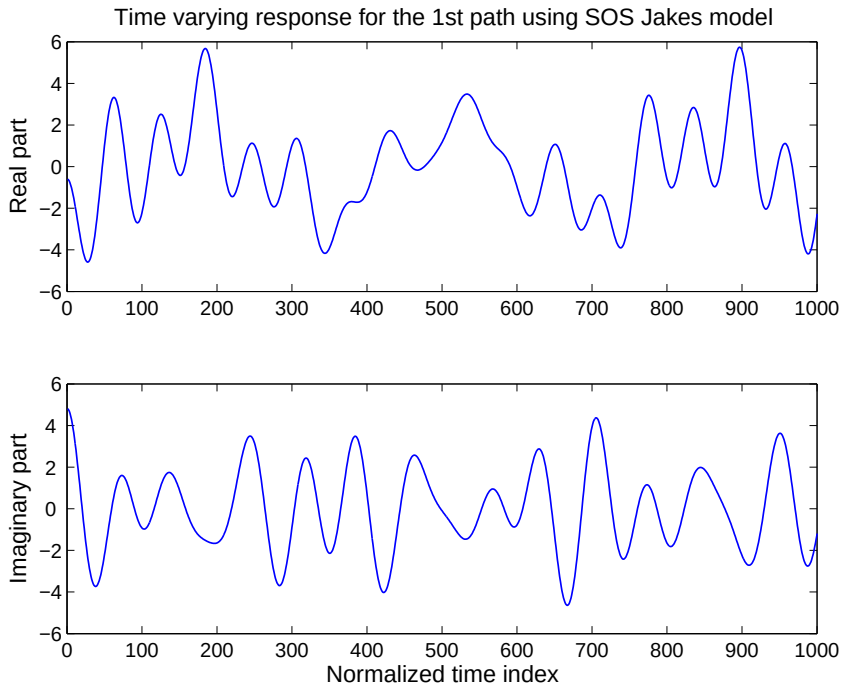


Figure 3.8: A frequency selective slowly fading channel using modified SOS Jakes model

Chapter 4

DS-CDMA RAKE Receivers

4.1 Introduction

Advancement of new technologies in wireless communication has increased rapidly since the proposed third and fourth generations of mobile communication systems. With these new technologies, mobile systems will operate in more diverse environmental conditions and they should be able to cope with the transmission medium effects such as multipath fading and channel noise. In this chapter we discuss some techniques which have been used in CDMA systems to overcome the multipath propagation effects and mitigate the channel noise. These techniques are based on the use of a well known CDMA receivers known as RAKE receivers. The chapter also presents a data model for the DS-CDMA system, this model will form the main bases for the work in this research.

4.2 DS-CDMA System Model

Consider the baseband model of a K user in DS-CDMA system shown in Fig. 4.1 operating with a coherent binary data modulation format. The kth user's continuous time baseband transmitted signal is modelled as:

$$x_k(t) = \sum_{i=1}^{N} A_k b_k(i)\ s_k(t - iT_s)\ ,\ t \in \mathbb{R}, \tag{4.1}$$

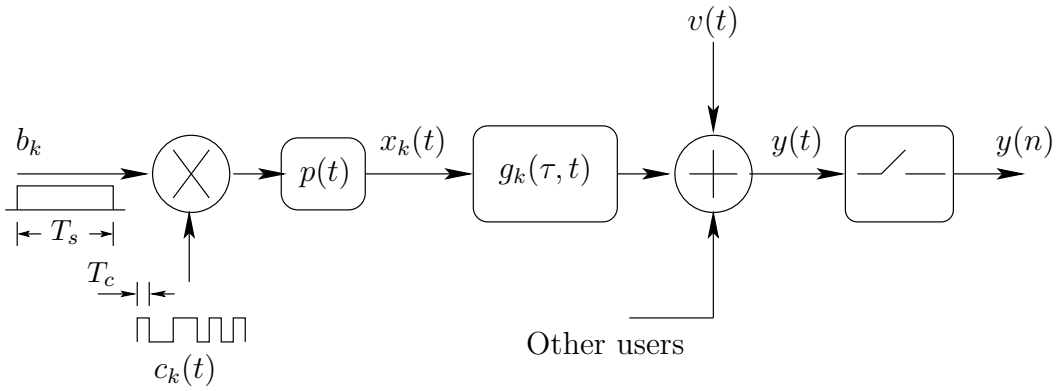


Figure 4.1: DS-CDMA data model

where A_k represents the kth user's amplitude, $b_k(i) \in \{+1, -1\}$ is the kth user data block with length N symbols, and T_s is the symbol period. The normalized waveform signature $s_k(t)$ is

$$s_k(t) = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} c_k(m) p(t - (m-1)T_c) \quad , \; t \in \mathbb{R}, \tag{4.2}$$

where M is the processing gain, $c_k(m) \in \{+1, -1\}$, $m = 1, 2, \ldots, M$ is the signature for user k, $p(t)$ is the chip waveform with chip duration $T_c = \frac{T_s}{M}$ and is assumed to be a raised cosine with a certain roll-off factor.

For a channel with a maximum duration of τ_{kd_k}, i.e.,

$$g_k(\tau_{kl}, t) = 0 \text{ for } \tau_{kl} > \tau_{kd_k} \tag{4.3}$$

where d_k represents the maximum delay index so that the continuous-time asynchronous baseband signal at the receiver is given by

$$y(t) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(\tau_{kl}, t) \; x_k(t - \tau_{kl}) + v(t), \; t \in \mathbb{R}, \tag{4.4}$$

where $v(t)$ represents the channel additive noise.

We consider a frequency selective slowly fading propagation (see chapter 3, section 3.3) then one can regard the multipath delays $\tau_{kl}(t) \approx \tau_{kl}$ and attenuation $g_k(\tau_{kl}, t) \approx g_k(\tau_{kl})$ during at least one or few signaling intervals [4], that is,

$$y(t) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(\tau_{kl}) \; x_k(t - \tau_{kl}) + v(t), \; 0 \le t \le T \quad , \quad T = T_s + \tau_{kd_k}, \tag{4.5}$$

After sampling with a rate (r/T_c), $r \geq 1$, then $g_k(\tau_{kl})$ can be simplified as $= g_k(l)$ and we obtain at the receiver

$$y(n) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(l)\ x_k\left(n - \frac{r\tau_{kl}}{T_c}\right) + v(n),\ n = 0, .., r\mathcal{M} - 1, \quad \mathcal{M} = M + \left\lfloor \frac{\tau_{kd_k}}{T_c} \right\rfloor \tag{4.6}$$

In the above equation, we consider the problem in its general form. In other words, we consider different delay and different attenuation for each path in each user so that the data model can cover the up-link systems. In case of down-link systems, delays and attenuation will be same for all users.

4.2.1 Research Primary Objectives

The primary objective of this work is to compensate for the effects of the multipath propagation $g_k(\tau_{kl})$ for each user k; In other words, it is to retrieve the transmitted symbols $b_k(l)$ from the baseband signal in Eqn. (4.6) using a parsimonious receiver which is based on an estimates of $g_k(\tau_{kl})$. This receiver will be only given the observations $y(n)$ and in some cases the training sequences $x_k(n)$, for $k = 1, \ldots, K,\ n = 0, \ldots, r\mathcal{M} - 1$, is also available. For the the design of such receiver we assume that $x_k(n)$, and the complex noise process $v(n)$ are stationary and independent with $\mathsf{E}\{v(n)\} = 0$ and $\mathsf{Var}\{v(n)\} < \infty$.

4.3 RAKE Receivers

The word "RAKE" is not an acronym and, in fact, is not always capitalized as it is in this writing. RAKE derives its name from its inventors Price and Green in 1958. For direct sequence spread spectrum modulation, the RAKE receiver is an indirect way of performing matched filtering to reduce the multipath effects. The received signal $y(t)$ as shown in Eqn. (4.6) is the summation of the delayed time arrivals of the transmitted signal. The RAKE receiver captures separately the multipath components of each user in $y(t)$ by exploiting the correlation properties of the spreading codes, $c_k(t)$. Specifically, if the arrivals are separated in time by more than one chip duration $\tau_{kl} \geq T_c$, the different paths can be resolved. A

path designed to extract one specific delay is referred to as a finger or correlator. Each finger despreads the signal, and the outputs of the respective fingers are weighted by the parameter $\hat{g}_k^*(l)$ which is the complex conjugate of the estimated coefficient of the channel and combined to produce a final signal estimate. A RAKE receiver usually employs a number of fingers to cover the delay spread of the channel. In IS-95, 1999 version, the base station combines the outputs of its RAKE receiver fingers incoherently. i.e. the outputs are added in power. The mobile receiver combines its RAKE receiver finger outputs coherently, i.e., the outputs are added in voltage. Currently, mobile receivers and base station receivers have between four to eight fingers, depending on the equipment manufacturer [6] [32] [33]. There are two primary methods used to combine the RAKE receiver finger outputs. One method weights each output equally and is, therefore, called Equal-Gain Combining (EGC). The second method uses the received data to estimate weights which maximize the SNR of the combined output. This technique is known as Maximal-Ratio Combining (MRC). In practice, but under some conditions [34] [35], it is not unusual for both combining techniques to have a close performance. In Fig. 4.2 a MRC-based RAKE receiver with L fingers is shown. Each finger is in charge of compensating one of the multipath components. Here $\hat{g}_k^*(l)$ is the complex conjugate of an estimate of the lth path parameter of user k, $g_k(l)$.

In practical RAKE receivers, the following functionality for symbol estimation and synchronization are needed:

1. Channel response acquisition and scanning to allocate RAKE fingers
2. Channel delay tracking to fine-adjust and track multipath components
3. Complex channel coefficient tracking to obtain coherent reception
4. Automatic Gain Control (AGC) to keep the receiver output within the dynamic range of the A/D converter
5. Automatic Frequency Control (AFC) to compensate for the drift of the local oscillator and possibly to compensate for high Doppler shifts

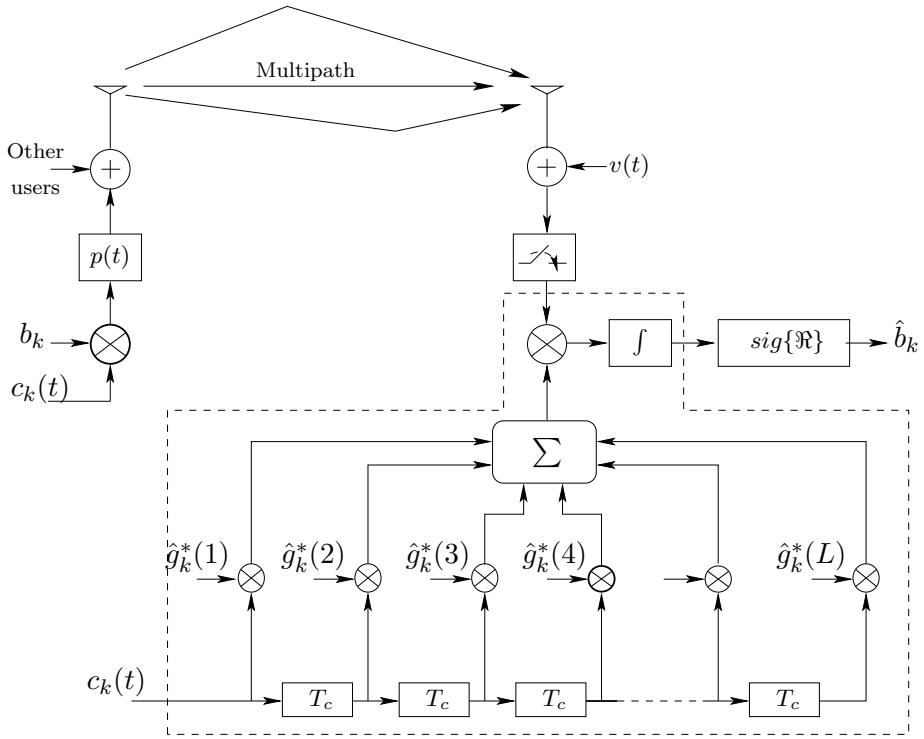


Figure 4.2: A single user MRC based RAKE receiver with L taps

4.4 Research Main Objectives

RAKE receivers as we mentioned above are used to overcome the effect of multipath propagation effects. However, it is not only important to design such receiver in its classical ways but also is to optimize its structure. Accordingly, the main objectives of this thesis is propose algorithms and techniques that can be used for the design of parsimonious structure RAKE receivers.

4.4.1 Parsimonious structure RAKE receivers

As the chip duration, T_c, becomes very short in new high rate systems such as 3G UMTS systems [6] [7], the channel delay becomes very long when expressed in chips. Since the number of significant paths and the order d_o are unknowns, in theory the number of fingers or taps of decorrelator receivers is limited to $L = \lfloor \frac{\tau_{max}}{T_c} \rfloor > d_o \gg d_k$, where τ_{max} represents the maximum delay-spread of the channel [4] and d_k is the number of the significant parameters which is 5 in Fig. 4.3. Consequently, this direct implementation of a RAKE receiver will

require many fingers to be implemented and many coefficients to be estimated.

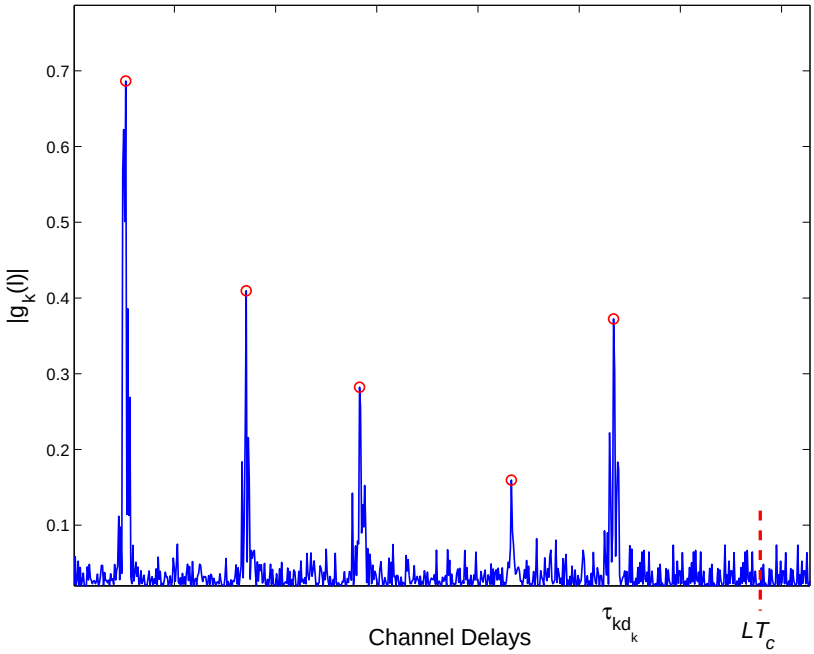


Figure 4.3: A sharp multipath propagation in a high SNR channel

In reality most of these coefficients are either zero or insignificant and this will lead to:

- Poor estimates of the channel response $\boldsymbol{g}_k = [g_k(1)\ g_k(2)\ ....\ g_k(L)]^T$
- Decorrelation of undesired signals through spurious fingers
- Increase in the RAKE receiver complexity and reduction in performance
- Loss of synchronization

The solution to this problem is to identify and estimate at once the non-zero or significant parameters $g_k(l)$ and delays τ_{kl}, for $k = 1, \ldots, K,\ l = 1, \ldots, L$, which also require estimation of the channel true length, $d_o = \lfloor \frac{\tau_{kd_k}}{T_c} \rfloor$ and the number of these significant parameters, d_k.

Once these goals are achieved,the RAKE receiver will have a minimal structure with only number of fingers much less than L and less than d_k itself.

The identification of the significant parameters can be achieved using model selection techniques based on statistical tests. For this goal, model selection techniques will be covered in the following chapter.

Chapter 5

Model Selection Techniques

5.1 Introduction

Model selection is a fundamental problem in many data analysis tasks. It has applications such as system identification, wireless communication systems, networks, radar and sonar [36] [37] [38]. Over the past two decades, many model selection techniques have been developed such as Akaike Information Criterion (AIC) or Rissanen's principle of Minimum Description Length (MDL) [16] [17] [39]. The widespread use of these criteria is mainly due to their intrinsic simplicity. Due to their better accuracy in estimating the correct model, other methods such as Forward or Backward Elimination (FE)(BE), sphericity test and F-statistic have also become widely used [40] [41] [42]. Recently, it has been proposed that bootstrap techniques can also be applied to model selection [38] [43] [44]. The main advantage of the bootstrap over classical statistical methods is that, it can be applied with minimal assumptions to scenarios where minimal information about the underlying distributions involved is available [45]. Although there exist many model selection techniques, the development of new techniques that outperform the popular ones is still ongoing.

In this chapter we present different model selection methods. These methods will be used to identify the significant multipath components for parsimonious DS-CDMA RAKE receivers. The chapter is organized as follows: In Section 5.2, we

introduce the model of interest and assumptions. In Section 5.3, we discuss the principles of the MDL, AIC and F-statistic. In Section 5.4, we present solutions to the full model selection problem. In Section 5.5 we summarize the proposed methods before we conclude in Section 5.6.

5.2 Data Model and Assumptions

Consider the following discrete signal model,

$$y(n) = \sum_{l=1}^{d} g(l)\, x(n,l) + v(n),\ \ n = 1, ..., N \tag{5.1}$$

where $y(n)$ represents the received signal, $\mathbf{g} = [g(1)\ g(2) \cdots g(d)]^T$ represents the model parameters, $x(n)$ is the model input signal and $v(n)$ is the additive noise. In matrix form,

$$\mathbf{y} = \mathbf{X}\mathbf{g} + \mathbf{v}. \tag{5.2}$$

where, without loss of generality,

$$\mathbf{X} = \begin{bmatrix} x(1,1) & x(1,2) & \cdots & x(1,d) \\ x(2,1) & x(2,2) & \cdots & x(2,d) \\ \vdots & \vdots & \vdots & \vdots \\ x(N,1) & x(N,2) & \cdots & x(N,d) \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}.$$

5.2.1 Objectives

The objective is to use statistical tests to identify the non-zero parameters in the model $\mathbf{g} = [g(1)\ g(2) \dots g(d)]^T$ and the order d itself, given observations $y(n)$ and in some cases the sequence $x(n)$ for $n = 1, \dots, N$ or in general $\mathbf{y}$ and $\mathbf{X}$. We will assume that $x(n)$ and $v(n)$ are stationary and independent. Minimal assumption will be considered for the noise $v(n)$ but in some cases we will make no assumption on its distribution except for $\mathsf{E}\{v(n)\} = 0$ and $\mathsf{Var}\{v(n)\} < \infty$.

5.3 Model Selection Methods

5.3.1 The AIC and MDL criteria

The AIC and MDL criteria are two well known methods for model selection. They are mainly applied by evaluating two terms: a data term which requires the maximization of the likelihood function and a penalty term which is a function of the complexity of the model [16] [17].

For a model with parameters $\mathbf{g}$ of length L assigns the likelihood $f(y/\mathbf{g})$ to a given set of observed data y, the AIC and the MDL criteria can be defined by,

$$\begin{aligned} \mathrm{AIC}(k) &= -\log(f(y/\hat{\mathbf{g}})) + p_{\mathrm{AIC}}(N,k) \\ \mathrm{MDL}(k) &= -\log(f(y/\hat{\mathbf{g}})) + p_{\mathrm{MDL}}(N,k) \end{aligned} \tag{5.3}$$

where $\hat{\mathbf{g}}$ is an estimate of the model parameters that maximize the likelihood, k is the number of parameters in the model, N is the sample size and $p_{\mathrm{AIC}}(N,k)$ and $p_{\mathrm{MDL}}(N,k)$ represent the penalty terms for the AIC and the MDL criteria respectively.

For the model given in Sec. 5.2, the AIC and the MDL criteria can be directly described by the following equations:

$$\begin{aligned} \mathrm{AIC}(k) &= \log(\sigma_{v_r}^2) + \tfrac{2(k-1)}{N} \\ \mathrm{MDL}(k) &= \log(\sigma_{v_r}^2) + \tfrac{k}{N}\log(N) \end{aligned} \tag{5.4}$$

where, k and L are the suggested model order and the maximum model order respectively, $\log(\sigma_{v_r}^2)$ represents maximization of the likelihood term, $p_{\mathrm{AIC}}(N,k) = \frac{2(k-1)}{N}$, $p_{\mathrm{MDL}}(N,k) = \frac{k}{N}\log(N)$ and residual errors $v_r(n)$ is defined by,

$$v_r(n) = y(n) - \sum_{l=1}^{k} \hat{g}(l)\; x(n,l),\;\; n = 1, ..., N \tag{5.5}$$

where, without loss of generality, $\hat{\mathbf{g}} = [\mathbf{X}^H\mathbf{X}]^{-1}\mathbf{X}^H\,\mathbf{y}$, represents the LSE of $\mathbf{g}$. Estimates of $\mathbf{g}$ can also be found using any other methods such as correlation and cross-correlation based methods. In that case $\mathbf{X}$ and $\mathbf{y}$ will be replaced by the correlation and cross-correlation matrices $\mathbf{R}_{xx}$ and $\mathbf{R}_{yx}$ respectively. Finally an estimate of the model order $\hat{d}$ using, say, the MDL method can then be determined

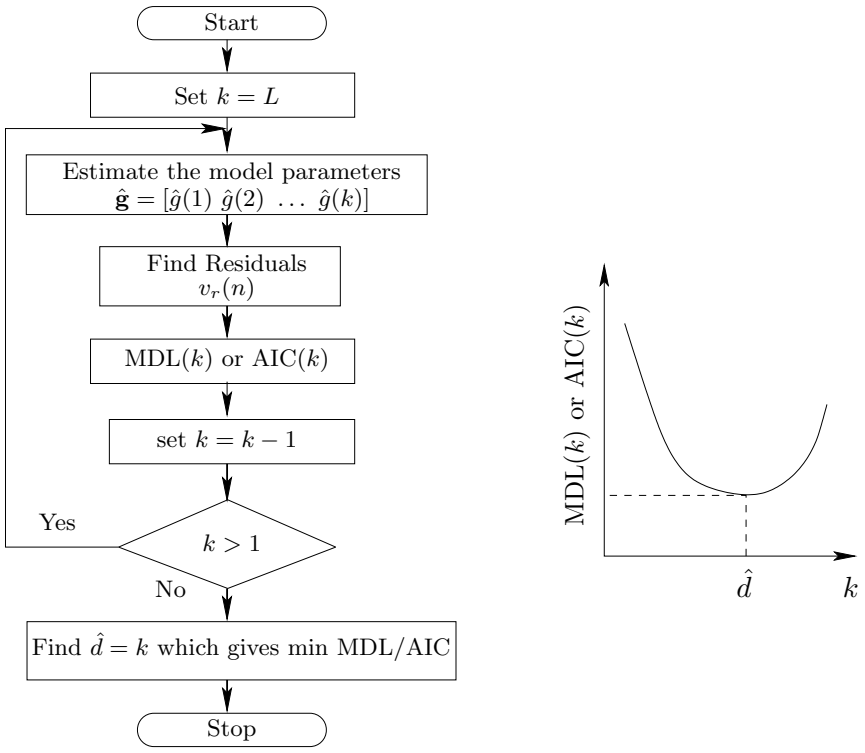


Figure 5.1: MDL/AIC algorithm for model order estimation

by,

$$\hat{d} = \underset{k\in\{1,2,\ldots,L\}}{\operatorname{argmin}} \ \mathrm{MDL}(k) \tag{5.6}$$

The order estimation algorithm shown in Fig. 5.1 summarizes the two methods.

5.3.2 AIC/MDL sphericity test based method

The sphericity test is one of the well known tests which can be used to check the multiplicity of a certain subject. In some cases it is known as the variance/covariance homogeneity test [46][47]. To become familiar with the test, let us consider the model in Sec. 5.2 in its matrix form, i.e.,

$$\mathbf{y} = \mathbf{X}\mathbf{g} + \mathbf{v}. \tag{5.7}$$

To apply the test, first we define the correlation matrix,

$$\mathbf{R} = \mathsf{E}\{\mathbf{y}\mathbf{y}^H\} = \mathbf{A}\mathbf{g}\mathbf{g}^H\mathbf{A}^H + \sigma_v^2\,\mathbf{I}. \tag{5.8}$$

Since the true number of parameters, d_k, is unknown we assume that $\mathbf{R}$ is a rank-L matrix.

The next step is to find the eigenvalues of $\mathbf{R}$ and sort them such that, $\lambda_1 > \lambda_2 > \ldots > \lambda_L$. These eigenvalues are biased and mutually correlated. Their finite sample joint distribution is known in the Gaussian case and is represented as a series of zonal polynomials [47][48]. A mathematically tractable form for their asymptotic joint distribution does exist in the Gaussian case [49], although it may be unreliable for small sample sizes. In addition, this joint distribution is sensitive to departures from Gaussianity [50]. Once the eigenvalues are found then the sphericity test (ratio) of a k-parameter model, $\mathrm{T}_{sph}(k)$ is defined by,

$$\mathrm{T}_{sph}(k) = \left[\frac{\prod_{l=k+1}^{L} \lambda_l^{1/(L-k)}}{\frac{1}{L-k} \sum_{l=k+1}^{L} \lambda_l} \right] \tag{5.9}$$

where the numerator and the denominator parts represent the geometric and the arithmetic means respectively. When $\mathrm{T}_{sph}(k) = 1$, it means that the geometric mean is the same as the arithmetic mean or in other words $\lambda_{k+1} = \lambda_{k+2} = \ldots = \lambda_L$. This suggests the matrix $\mathbf{R}$ has a rank $(k+1)$ and $\hat{d}_t = k$. Since the test requires an estimate of the eigenvalues λ's, which are biased, then in some cases $\mathrm{T}_{sph}(k)$ will be approximately close to one and it can occur at different values of k. Whereupon, the true number of parameters $\hat{d}_t$ will then be taken to be the smallest k. Due to the previous problem, one can fix it by using the MDL or the AIC criterion.

The AIC and MDL criteria use $\mathrm{T}_{sph}(k)$ as the likelihood function, in other words,

$$\begin{aligned} \mathrm{AIC}(k) &= -2 \log \left[\frac{\prod_{l=k+1}^{L} \lambda_l^{1/(L-k)}}{\frac{1}{L-k} \sum_{l=k+1}^{L} \lambda_l} \right]^{(L-k)N} + 2k(2L-k) \\ &\text{and} \\ \mathrm{MDL}(k) &= -\log \left[\frac{\prod_{l=k+1}^{L} \lambda_l^{1/(L-k)}}{\frac{1}{L-k} \sum_{l=k+1}^{L} \lambda_l} \right]^{(L-k)N} + \tfrac{1}{2} k(2L-k) \log(N) \end{aligned} \tag{5.10}$$

An estimate of the principal number of parameters in the model $\hat{d}_t$ is k over which the AIC or the MDL function is minimized, i.e.,

$$\begin{aligned} \hat{d}_t &= \underset{k\in\{1,2,\ldots,L\}}{\operatorname{argmin}} \ \mathrm{AIC}(k) \\ \hat{d}_t &= \underset{k\in\{1,2,\ldots,L\}}{\operatorname{argmin}} \ \mathrm{MDL}(k) \end{aligned} \tag{5.11}$$

5.4 Full Identification of Significant Model Parameters

In the previous section we showed how to estimate the order of a certain channel model. In CDMA systems, when dealing with multipath propagation, estimation of the channel order is not sufficient for low order channel models. It is also important to check the significance of each individual parameter in the model. For example, in new high rate systems such as 3G UMTS DS-CDMA systems [6] [7], the short chip duration, T_c, results in a very long model order, $L = \left\lfloor \frac{\tau_{max}}{T_c} \right\rfloor$, where τ_{max} is the maximum delay-spread of the channel [4]. The problem is that most of the model parameters are either zero or non-significant. If one considers only the channel order then he/she will end up with a high complexity RAKE receiver. Moreover, this receiver will also be based on poor estimates of the channel parameters.

In this section we proposed different model selection methods that can be used not only to estimate the model order but also to check the significance of each individual parameter in that model. Once this is done then low order estimates of the channel will be achieved and a parsimonious RAKE receiver based on these estimates can be considered.

5.4.1 Hierarchical MDL/AIC based method

In standard model selection and regression problems, it is important to decide how many parameters or predictors to enter. In hierarchical-based methods, it is not only how many parameters to enter but also the order in which they enter is important. Usually, the order of entry is based on logical or theoretical

considerations. In hierarchical BE multiple regression analysis, the number of parameters to be selected and the order of entry are both decided by statistical measures such as entry or removal criterion [41] [42].

In this hierarchical procedure we begin with the full model (parent model) that includes all parameters from $\hat{g}(1)$ to $\hat{g}(L)$. Define $\mathbf{p}$ as the parent hierarchical code at this step, $\mathbf{p} = [1\ 1 \ldots 1]_{(L\times 1)}$. That means we have to use the full LSE matrix $\mathbf{X}$ in Sec. 5.3 to find the LSE $\hat{\mathbf{g}}_{\mathbf{p}}$ of the parent model. From the parent hierarchical code $\mathbf{p}$, one can generate the child hierarchical codes $\mathbf{c}_1,\ \mathbf{c}_2, \ldots \mathbf{c}_L$ and estimate their LSEs, $\hat{\mathbf{g}}_{\mathbf{c}_1},\ \hat{\mathbf{g}}_{\mathbf{c}_2}, \ldots \hat{\mathbf{g}}_{\mathbf{c}_L}$.

A zero element in any code means a removal of its corresponding parameter and a removal of its corresponding column in the matrix $\mathbf{X}$. The next step is then to compare the MDL or AIC value of each child model and select the child which has the smallest value. If this child's MDL value is less than or equal to its parent one then we replace the parent code with the child code and continue; otherwise we keep it and stop.

In Table 5.1 and 5.2 two iterations for $L = 4$ describe how to generate the hierarchical codes and their corresponding models to be tested. The complete model selection procedure is shown in Fig. 5.2, and also summarized in Table 5.3.

Table 5.1: Iteration # 1

Item	Code				Parameters				MDL
$\mathbf{p}$	1	1	1	1	$g(1)$	$g(2)$	$g(3)$	$g(4)$	$v_{\mathbf{p}}$
$\mathbf{c_1}$	0	1	1	1	0	$g(2)$	$g(3)$	$g(4)$	$v_{\mathbf{c_1}}$
$\mathbf{c_2}$	1	0	1	1	$g(1)$	0	$g(3)$	$g(4)$	$v_{\mathbf{c_2}}$
$\mathbf{c_3}$	1	1	0	1	$g(1)$	$g(2)$	0	$g(4)$	$v_{\mathbf{c_3}}$
$\mathbf{c_4}$	1	1	1	0	$g(1)$	$g(2)$	$g(3)$	0	$v_{\mathbf{c_4}}$

$$v_{\mathbf{c_2}} \leq v_{\mathbf{p}} \Rightarrow \text{Yes, then } \mathbf{p} = \mathbf{c_2}$$

Table 5.2: Iteration # 2

Item	Code				Parameters				MDL
$\mathbf{p}$	1	0	1	1	$g(1)$	0	$g(3)$	$g(4)$	$v_{\mathbf{p}}$
$\mathbf{c_1}$	0	0	1	1	0	0	$g(3)$	$g(4)$	$v_{\mathbf{c_1}}$
$\mathbf{c_2}$	1	0	0	1	$g(1)$	0	0	$g(4)$	$v_{\mathbf{c_2}}$
$\mathbf{c_3}$	1	0	1	0	$g(1)$	0	$g(3)$	0	$v_{\mathbf{c_3}}$

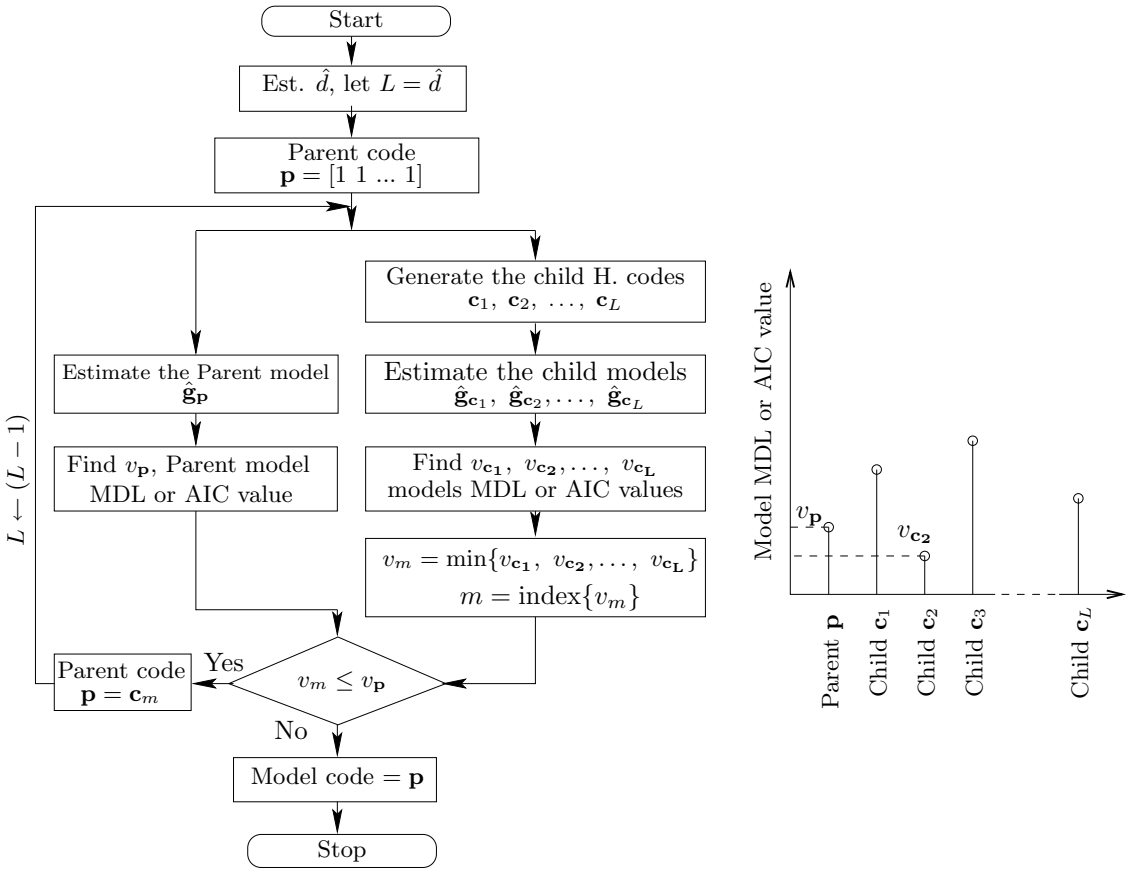


Figure 5.2: Hierarchical BE based method for model selection

Table 5.3: The hierarchical MDL/AIC based procedure

Step 1: Use the MDL or the AIC to estimate the model order, $\hat{d}$
Step 2: Set $L = \hat{d}$ and start the parent with the full code $\mathbf{p} = [1\ 1\ \ldots 1]_{L\times 1}$.
Step 3: Estimate the parent model $\hat{\mathbf{g}}_{\mathbf{p}} = [\hat{g}(1)\ \hat{g}(2)\cdots\hat{g}(L)]^T$, then calculate its MDL/AIC value $v_{\mathbf{p}}$ given by Sec. 5.3.
Step 4: Generate the child hierarchical codes $\mathbf{c}_1,\ \mathbf{c}_2, \ldots \mathbf{c}_L$, remove from the LSE matrix $\mathbf{X}$, columns corresponding to zero locations of each code, then estimate each code's model $\hat{\mathbf{g}}_{\mathbf{c}_l},\ l = 1, 2, .., L$.
Step 5: For each child model $\hat{\mathbf{g}}_{\mathbf{c}_l},\ l = 1, 2, .., L$, calculate the MDL/AIC value, $\mathbf{v} = [v_{\mathbf{c_1}},\ v_{\mathbf{c_2}}, \ldots v_{\mathbf{c_L}}]$.
Step 6: Find the minimum MDL/AIC value among the child models $v_m = \min\{\mathbf{v}\}$, where m is the index, then compare v_m with the parent one $v_{\mathbf{p}}$.
Step 7: If $v_k \leq v_{\mathbf{p}}$, replace the Parent code $\mathbf{p}$ by $\mathbf{c}_k$ then return to **Step 3**. otherwise **Stop** and $\mathbf{p}$ will be then the procedure decision about the model.

5.4.2 F-statistic

The F-statistic or F-test is one of the most widely known methods for model selection. The mathematical definition of the F-test is in general,

$$F_o = \frac{u_m/m}{v_n/n} \tag{5.12}$$

where u_m and v_n are two independent χ^2 distributed random variables with degrees of freedom (DF) m and n respectively. The F-statistic here will be formulated in a way which is different from some other methods given by [40] [41]. To understand the test let us consider this example where we have N samples of the linear model $y(n) = g(0) + g(1)x(n,1) + g(2)x(n,2) + \ldots + g(L)x(n,L)$. The question is, under a specific level of significance γ, does the term $g(l)x(n,l)$ add a significant contribution to the model, i.e., is $g(l) = 0$? The answer is to apply

the F-test for $g(l) = 0$ between the full and reduced models,

$$\begin{aligned} y_f(n) &= \sum_{i=0}^{L} g(i)x(n,i) \\ &\text{and} \\ y_r(n) &= \sum_{i=0, i\neq l}^{L} g(i)x(n,i) \end{aligned} \tag{5.13}$$

In other words, we need to test simultaneously for $l = 1, \ldots, L$ the hypothesis,

$$\begin{aligned} \mathcal{H}_l &: g(l) = 0 \\ &\text{against} \\ \mathcal{K}_l &: g(l) \neq 0 \end{aligned} \tag{5.14}$$

To do that, we define the following,

- SS stands for sum of squares and DF for degree of freedom.
- $SST = \sum_{n=1}^{N} |y_f(n) - \bar{y}_f|^2$, SS total with $DF_f = N - 1$, $\bar{y}_f$ is the mean value of $y_f(n)$.
- $SSE_f = \sum_{n=1^N} |y_f - \hat{y}_f|^2$, SS of the residual error for the full model:
  - $y_f(n) = \sum_{i=0}^{L} g(i)x(n,i)$ with $DF_r = N - 1 - L$
- $SSE_r = \sum |y_r(n) - \hat{y}_r|^2$, SS of the residual errors for the reduced model:
  - $y_r(n) = \sum_{i=0, i\neq l}^{L} g(i)x(n,i)$ with $DF_m = N - 1 - (L-1)$,
- $SSR_f = SST - SSE_f$ Regression SS with $DF_f = L$, (full model).
- $SSR_r = SST - SSE_r$ Regression SS with $DF_r = L - 1$, (reduced model).
- $MSE_f = \frac{SSE_f}{N-1-L}$ Mean squared error with $DF = N - 1 - L$.
- $F_l = \frac{(SSR_f - SSR_r)/(DF_f - DF_r)}{MSE_f} \sim F(DF_f - DF_r, DF_m)$.
- Define $\mathcal{P}$-value, $\mathcal{P}_l = Pr\{f > F_l\}$, where $f \sim F(DF_f - DF_r, DF_m)$.
- Test hypothesis $[\mathcal{H}_l : g(l) = 0]$ against $[\mathcal{K}_l : g(l) \neq 0]$. In other words as shown in Fig. 5.3,

$$\text{if } \mathcal{P}_l > \gamma \;\Rightarrow\; g(l) = 0$$
$$\text{if } \mathcal{P}_l \leq \gamma \;\Rightarrow\; g(l) \neq 0$$

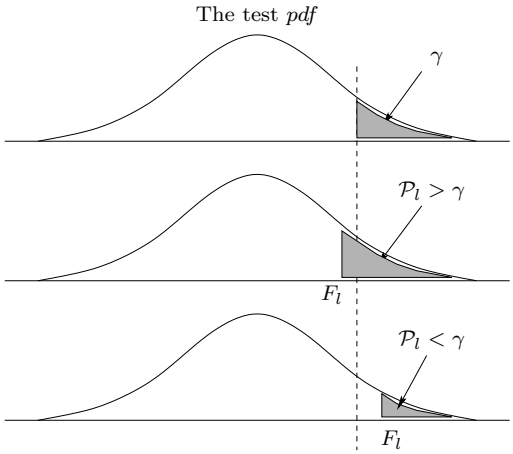


Figure 5.3: Comparing the F-statistic $\mathcal{P}$-value with γ

5.4.2.1 BE method based on F-statistic

This procedure begins with the full model that includes all considered parameters from $g(1)$ to $g(L)$. It then attempts to remove one parameter at a time by determining whether the least significant variable currently in the model can be removed or not because its $\mathcal{P}$-value is greater or less than the user-specified or default level of significance γ. Once a variable has been removed from the model it cannot re-enter at subsequent steps. The procedure is shown in Fig. 5.4 and is summarized in Table 5.4 as follows,

5.4.3 The bootstrap based method

For the given channel model in Section 5.2, the model length L can be very large, but usually only a few channel parameters or paths are non-zero. The objective is to identify the non-zero or significant channel parameters for an arbitrary large $L < N$, given observations $y(n)$ and $x(n)$. In other words, under a global level of significance γ, we need to test simultaneously for $l = 1, \ldots, L$ the hypothesis,

Table 5.4: The F-statistic based procedure

Step 1: Assume the model length L and then estimate the full model parameters $\hat{\mathbf{g}} = [\hat{g}(1)\ \hat{g}(2) \cdots \hat{g}(L)]^T$.
Step 2: Use the F-statistic described before, find the $\mathcal{P}$-values, $\mathcal{P}_1,\ \mathcal{P}_2, \ldots \mathcal{P}_L$ for each parameter.
Step 3: Remove from the model the parameter that is insignificant, i.e $\mathcal{P}_l > \gamma$ and has the highest $\mathcal{P}$-value (one parameter at a time).
Step 4: Re-estimate the new model and again the parameter that is insignificant and has the highest $\mathcal{P}$-value is deleted.
Step 5: Continue until all remaining parameters are significant.

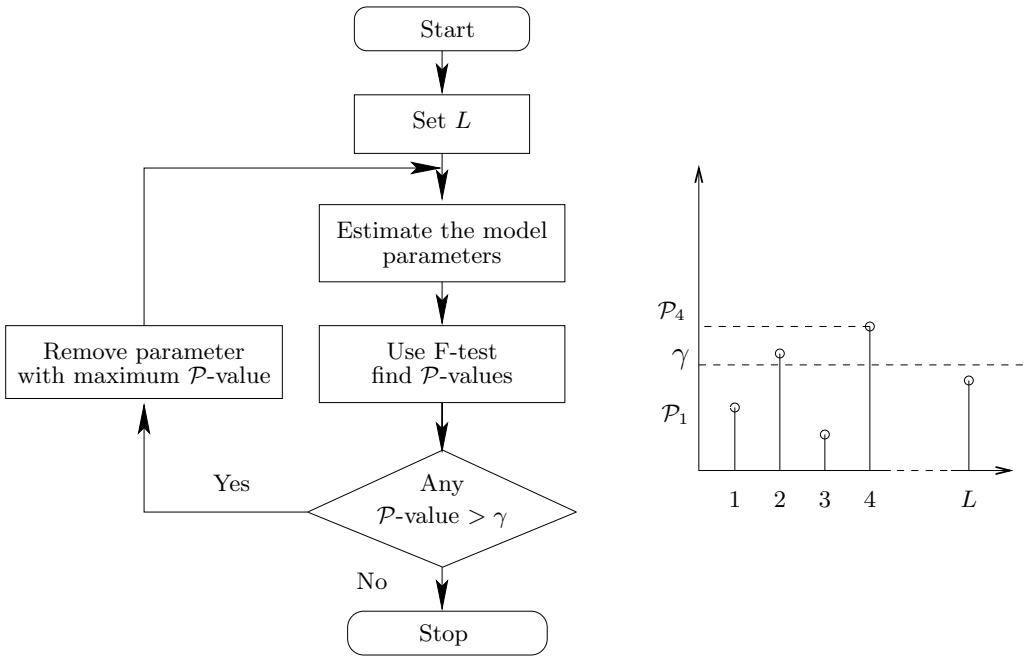


Figure 5.4: BE F-statistic based method

$$\begin{gathered} \mathcal{H}_l : g(l) = 0 \\ \text{against} \\ \mathcal{K}_l : g(l) \neq 0 \end{gathered} \tag{5.15}$$

When there is minimal assumption about the additive noise, it is important to use a good statistical technique to estimate the test distribution. The bootstrap is a good technique that can be considered for these problems [38] [45]. The

bootstrap is a recently developed statistical method that has many attractive properties such as, it is a powerful technique for assessing the accuracy of a parameter estimator in situations where conventional techniques are not valid.
Unlike the previous section which was based on a predefined F-statistic, this section we use the bootstrap multiple hypotheses test to identify the significant model parameters. The test can be summarized as follows.
First, we find the LSE $\hat{\mathbf{g}}$ of the parameters and form the test statistics,

$$\hat{T}_l = \frac{|\hat{g}(l)|}{\sigma_{\hat{g}(l)}} \ , \ l = 1, ..., L. \tag{5.16}$$

Then we calculate the residuals $v_r(n)$ as given by Eqn. 5.5, and draw resampled data $v_r^\star(n)$, using either the classical iid bootstrap or the surrogate data bootstrap as shown in Figure 5.5.
Unlike with the classical bootstrap where we draw at random with replacement samples or data blocks from $v_r(n)$, with surrogate data methods, we take the Fourier transform of the residuals $v_r(n)$ and modulate them with a random phase, distributed uniformly on $[-\pi, \pi)$, then inverse Fourier transform to generate the resampled $v_r^\star(n)$ [51].

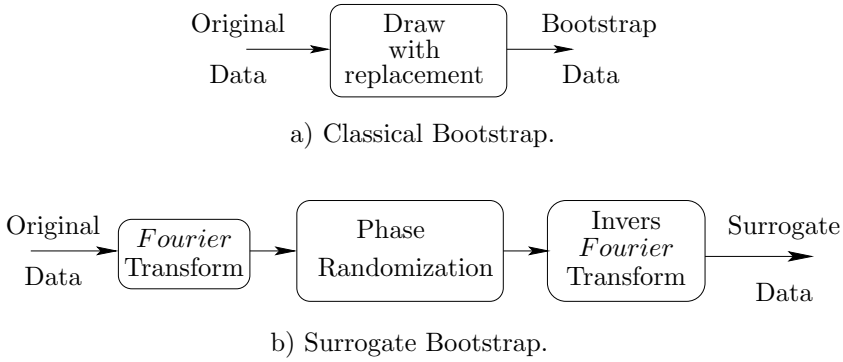


Figure 5.5: Bootstrap randomization techniques

The last step is to calculate the bootstrap data $y^\star(n)$ and re-compute the bootstrap channel estimates $\hat{\mathbf{g}}^\star = [\hat{g}(1)^\star \ \hat{g}(2)^\star \dots \hat{g}(L)^\star]^T$. Repeating the procedure B times leads to an approximate distribution of the test statistics, given by the bootstrap statistics.

$$\hat{T}_l^\star = \frac{|\hat{g}(l)^\star - \hat{g}(l)|}{\sigma_{\hat{g}(l)^\star}}, \ l = 1, ..., L. \tag{5.17}$$

With $\hat{T}_l^\star$ and $\hat{T}_l$ the $\mathcal{P}$-values are calculated (see Fig. 5.6),

$$\mathcal{P}_l = \frac{1}{B}\#\{\hat{T}_l^\star \geq \hat{T}_l\} \tag{5.18}$$

Then we check,

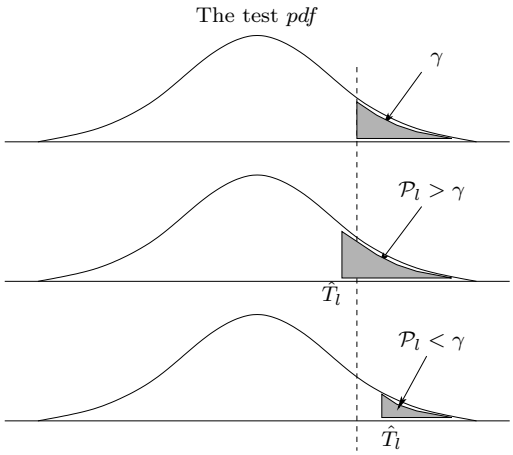


Figure 5.6: Comparing bootstrap statistical test $\mathcal{P}$-value with γ

$$\begin{aligned} \text{if } [\mathcal{P}_l > \gamma] &\Rightarrow g(l) = 0 \\ \text{if } [\mathcal{P}_l \leq \gamma] &\Rightarrow g(l) \neq 0. \end{aligned}$$

The full procedure is summarized in Table 5.5.

5.4.4 Classical and Sequentially Rejective Bonferroni tests

In the previous hypothesis test, we test the null hypothesis $\mathcal{H}_l : g(l) = 0$ against the alternative $\mathcal{K}_l : g(l) \neq 0$ w.r.t several parameters. When performing L multiple independent tests $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_L$ each at the γ level, the Family Wise Error, FWE, which is the probability of making at least one Type I error (rejecting the null hypothesis inappropriately) is $1-(1-\gamma)^L$ which is $\gg \gamma$. However, in order to maintain an FWE at a chosen level γ, if each hypothesis is tested separately using tests with level $\frac{\gamma}{L}$, then it follows immediately from the Boole inequality that the probability of rejecting any true hypothesis is less than or equal to γ. This constitutes a multiple test procedure with the global level of significance γ for free combinations. The classical test is depicted in Fig. 5.8. In this test

the hypotheses $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_L$, are assumed to be independent. For dependent hypotheses the test becomes overly conservative and is less powerful [52][53][54]. A more powerful test is the Sequentially Rejective Bonferroni (SRB)test, which can be used when the hypothesis tests are dependent [40]. Since the SRB test requires only the $\mathcal{P}$-values, then it can be used with the F-test or the bootstrap. First we sort the $\mathcal{P}$-values such that, $\mathcal{P}_{(1)} < \mathcal{P}_{(2)} < \ldots < \mathcal{P}_{(L)}$ and then to identify the significant model parameters we follow the algorithm shown in Fig. 5.7.

Table 5.5: The bootstrap procedure

Step 1: Calculate,

The model LSE, $\hat{\mathbf{g}} = [\hat{g}(1)\ \hat{g}(2)\ ...\ \hat{g}(L)]^T$

The parameter variances, $\sigma^2_{\hat{g}}$ =diag$\{[\mathbf{X}^T\mathbf{X}]^{-1}\hat{\sigma}^2_y\}$

where, $\hat{\sigma}^2_y \approx \frac{1}{N-1}\sum_{n=1}^{N}\left[y(n) - \frac{1}{N}\sum_{n=1}^{N} y(n)\right]^2$

Step 2: Define the test statistic, $\hat{T}_l = \frac{|\hat{g}(l)|}{\sigma_{\hat{g}(l)}},\ l = 1, \ldots, L.$

Step 3: Calculate the residuals $v_r(n)$ and remove its mean value,

$v_r(n) = y(n) - \sum_{l=1}^{L} \hat{g}(l)\ x(n,l),\ n = 1, \ldots, N$

Step 4: From $v_r(n)$ draw the bootstrap resampled data $v^\star_r(n)$, then calculate the bootstrap data,

$y^\star(n) = \sum_{l=1}^{L} \hat{g}(l)\ x(n,l) + v^\star_r(n),\ n = 1, ..., N$

Then use $y^\star(n)$ instead of $y(n)$ as in Step 1: to find the bootstrap LSE $\hat{\mathbf{g}}^\star$ and their $\sigma^2_{\hat{g}^\star}$

Step 5: Define the bootstrap test statistic, $\hat{T}^\star_l = \frac{|\hat{g}(l)^\star - \hat{g}(l)|}{\sigma_{\hat{g}(l)^\star}},\ l = 1, ..., L$

Step 6: Repeat steps (4 and 5) B times to obtain $\hat{\mathbf{T}}^\star$ where,

$\hat{\mathbf{T}}^\star = [\hat{\mathbf{T}}^\star_1\ \hat{\mathbf{T}}^\star_2 \ldots, \hat{\mathbf{T}}^\star_L],$

$\hat{\mathbf{T}}^\star_l = [\hat{T}^\star_l(1)\ \hat{T}^\star_l(2)\ \ldots\ \hat{T}^\star_l(B)]^T$

Step 7: For a global level of significance, γ equal say 5%, estimate the $\mathcal{P}$-values, $\mathcal{P}_l = \frac{1}{B}\#\{\hat{\mathbf{T}}^\star_l \geq \hat{T}_l\}$ and compare,

$$\text{if } \mathcal{P}_l > \gamma \Rightarrow \text{retain } \mathcal{H}_l : g(l) = 0$$

$$\text{if } \mathcal{P}_l \leq \gamma \Rightarrow \text{reject } \mathcal{H}_l : g(l) = 0$$

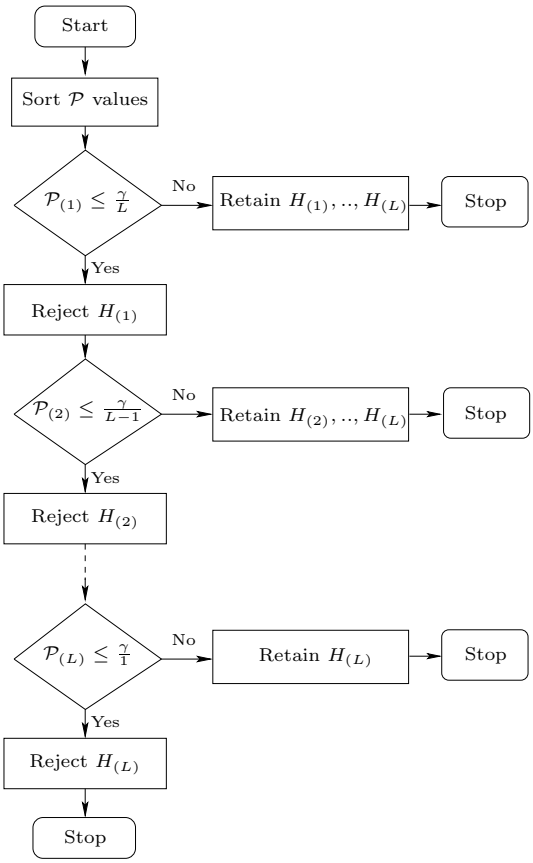


Figure 5.7: Sequentially Rejective Bonferroni tests (SRB)

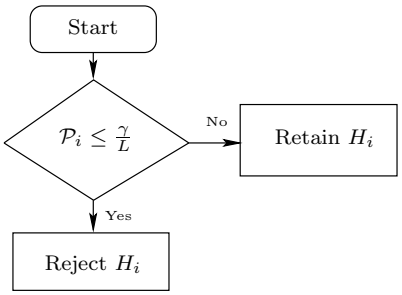


Figure 5.8: Classical Bonferroni test

5.5 Summary

The MDL and the AIC based methods are known to be simple techniques for model selection. In some situations, they can have better performance over other methods. However, some of the disadvantages of the MDL and the AIC are that, (1) both are based on the Gaussianity assumption (2) they over/under estimate the model order, also (3) they cannot control the level of significance. If one wants to control the level of significance then it is better to use either the *F*-statistic or the bootstrap based model selection techniques. The *F*-statistic and the bootstrap can use the Sequential Rejective Bonferroni test which is a powerful technique to control the global level of significance γ. In summary, one can choose any of the above methods to identify the non-zero parameters of a certain model. This selection of any method can be based on issues such as the the model selection

accuracy, the hardware and software implementations.

5.6 Conclusion

In this chapter we addressed the problem of model selection. Four different approaches for model selection are presented. In these approaches low computation complexity algorithms are developed based on the MDL, the AIC, F-statistic and the bootstrap. The MDL and AIC based methods can normally used when the model additive noise is Gaussian, but they do not control the level of significance. Under some assumptions, the F-statistic based methods can identify the significant model parameters and consider the level of significance. The performances of the F-statistic based method are better when the additive noise is Gaussian. Under minimal assumptions about the noise distribution and for a pre-defined level of significance, the bootstrap technique is a powerful technique to identify the significant model parameters.

Finally, all of the above model selection approaches are found to be capable of identifying the principle (the MDL and AIC cases) or the significant (F-test and bootstrap cases) model parameters with high probabilities. In the coming chapters, we use the above model selection techniques to identify the significant multipath parameters for parsimonious CDMA RAKE receivers.

Chapter 6

Channel Estimation and Identification Using Frequency Domain Sphericity Test and F-Statistics Based Approaches

6.1 Introduction

This chapter presents a frequency domain based approach for channel estimation resulting in an implementation of parsimonious DS-CDMA RAKE receivers. We consider frequency selective slowly fading channels (see Chapter 3. We estimate the channel delays and attenuation coefficients using the averaged periodogram and modified time delay estimation techniques. We then use model selection techniques to identify the model order and select the significant channel parameters. These techniques include the use of the MDL sphericity test based method and the multiple hypothesis tests F-Statistics based method discussed in Chapter 5. Simulations show that for a pre-defined level of significance, the proposed techniques correctly identify the significant channel parameters and the parsimonious RAKE receiver shows improved performance over the ones derived using the classical methods. The chapter is organized as follows. In Section 6.2, we in-

troduce the model and assumptions. In Section 6.3, we describe RAKE receivers. In Section 6.4, we present the estimation procedure. In Section 6.5 and 6.6, we consider the classification of significant channel parameters. In Section 6.7, we present simulation results before we conclude in Section 6.8.

6.2 Data Model

Consider K users in a baseband DS-CDMA up-link system operating with a coherent binary data modulation format, where the kth user's continuous time baseband transmitted tarining signal is modelled as:

$$x_k(t) = \sum_{i=1}^{N} A_k b_k(i)\ s_k(t - iT_s)\ ,\ t \in \mathbb{R}, \tag{6.1}$$

where A_k represents the kth user's amplitude, $b_k(i) \in \{+1, -1\}$ is the data block with length N symbols and T_s is the symbol period. The normalized waveform signature $s_k(t)$ is

$$s_k(t) = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} c_k(m) p(t - (m-1)T_c)\ \ ,\ t \in \mathbb{R}, \tag{6.2}$$

where M is the processing gain, $c_k(m) \in \{+1, -1\},\ m = 1, 2, \ldots, M$ is the signature for user k. $p(t)$ is the chip waveform with chip duration $T_c = \frac{T_s}{M}$ and is assumed, without loss of generality, to be a raised cosine with roll off factor 0.35.

We consider frequency selective slowly fading channels (see Section 3.3) with a maximum duration of τ_{kd_k} where d_k represents the maximum delay index so that the continuous-time asynchronous baseband signal at the receiver is given by

$$y(t) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(\tau_{kl}, t)\ x_k(t - \tau_{kl}) + v(t),\ t \in \mathbb{R}. \tag{6.3}$$

Where, $v(t)$ is a complex-valued additive Gaussian noise of $calN\,(0, \sigma_v^2)$ is independent with the channel response. Herein, one can regard the multipath parameters τ_{kl} nearly constant and $g_k(\tau_{kl}, t) \approx g_k(\tau_{kl})$ during at least one signaling interval [4], that is,

$$y(t) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(\tau_{kl})\ x_k(t - \tau_{kl}) + v(t),\ 0 \le t \le T\quad ,\quad T = T_s + \tau_{kd_k}, \tag{6.4}$$

After sampling with rate (r/T_c), $r \geq 1$, then $g_k(\tau_{kl})$ can be written as $g_k(l)$. We obtain at the receiver

$$y(n) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(l)\, x_k\left(n - \frac{r\tau_{kl}}{T_c}\right) + v(n),\ n = 0, .., r\mathcal{M} - 1, \quad \mathcal{M} = M + \lfloor \frac{\tau_{kd_k}}{T_c} \rfloor \tag{6.5}$$

6.2.1 Objectives

The objective is to retrieve the transmitted symbols $b_k(i)$ from the baseband signal in Eqn. (6.5) using a parsimonious RAKE receiver (see Fig. 6.1), given only observations $y(n)$ and training sequences $x_k(n)$ for $k = 1, \ldots, K,\ n = 0, \ldots, r\mathcal{M}-1$. We will assume that $x_k(n)$ and the additive noise $v(n)$ are stationary and independent under some regularity conditions for large $r\mathcal{M}$.

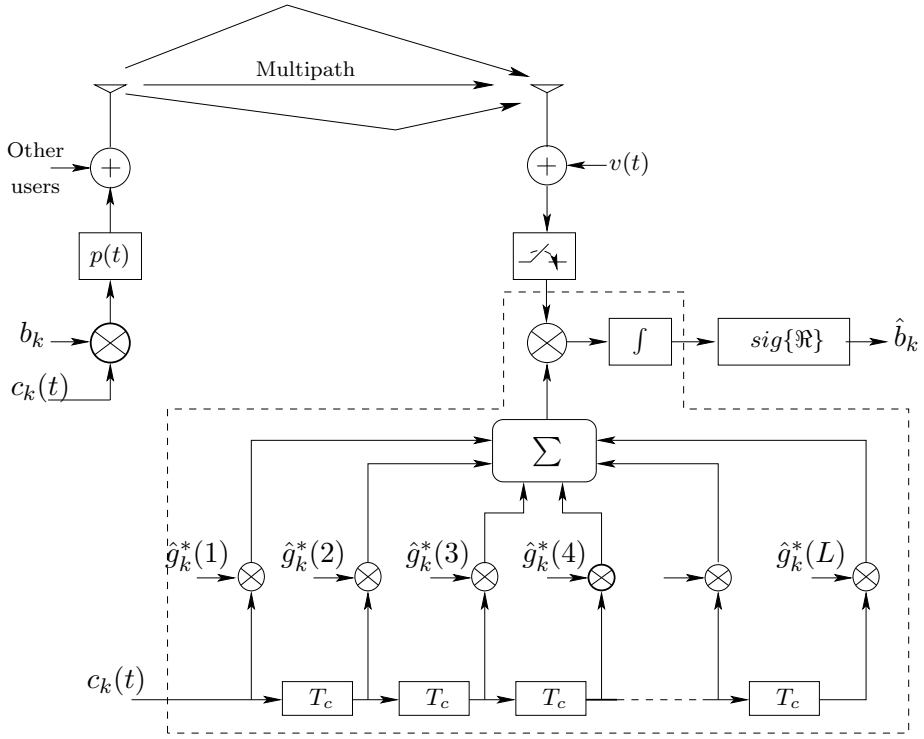


Figure 6.1: A single-user MRC based RAKE receiver with L taps

6.3 Parsimonious RAKE Receivers

As the chip duration, T_c, is shortened, in new high rate systems such as 3G UMTS systems (see Section 2.5.4), the maximum delay spread will be very long when expressed in chips. Since the number of significant channel paths, d_k is unknown, then in theory the number of fingers, or taps, of decorrelator receivers such as the RAKE receiver (see Chapter 4) will be limited to $L = \lfloor \frac{\tau_{max}}{T_c} \rfloor \gg d_k$, where τ_{max} is the maximum delay-spread of the channel [4]. In some situations L is greater than the channel length itself which is $d_o = \left\lfloor \frac{\tau_{kd_k}}{T_c} \right\rfloor$. Consequently, this direct implementation of a RAKE receiver will require many fingers to be implemented and many coefficients to be estimated. In reality most of these coefficients are either zero or insignificant as shown in Fig. 6.2, and this will lead to:

- Poor estimates of the channel response $\boldsymbol{g}_k = [g_k(1)\ g_k(2)\ ....\ g_k(L)]^T$
- Decorrelation of undesired signals through spurious fingers
- Increase in RAKE receiver complexity and reduction in performance
- Loss of synchronization

6.3.1 Solution

The solution to the above problem is to identify and estimate the non-zero or significant parameters $g_k(l)$ and delays τ_{kl}, for $k = 1, \ldots, K,\ l = 1, \ldots, L$. This will also require estimation of the channel length, define d_o and the number of these significant parameters d_k which is 5 true paths in the example shown in Fig. 6.2.

Once these goals are achieved then, the RAKE receiver will have a minimal structure with number of fingers much less than L and less than d_o. The identification of the principle or significant parameters can be achieved using model selection techniques.

In this work we model the problem in the frequency domain, then for the identification of the significant parameters we use two different model selection techniques:

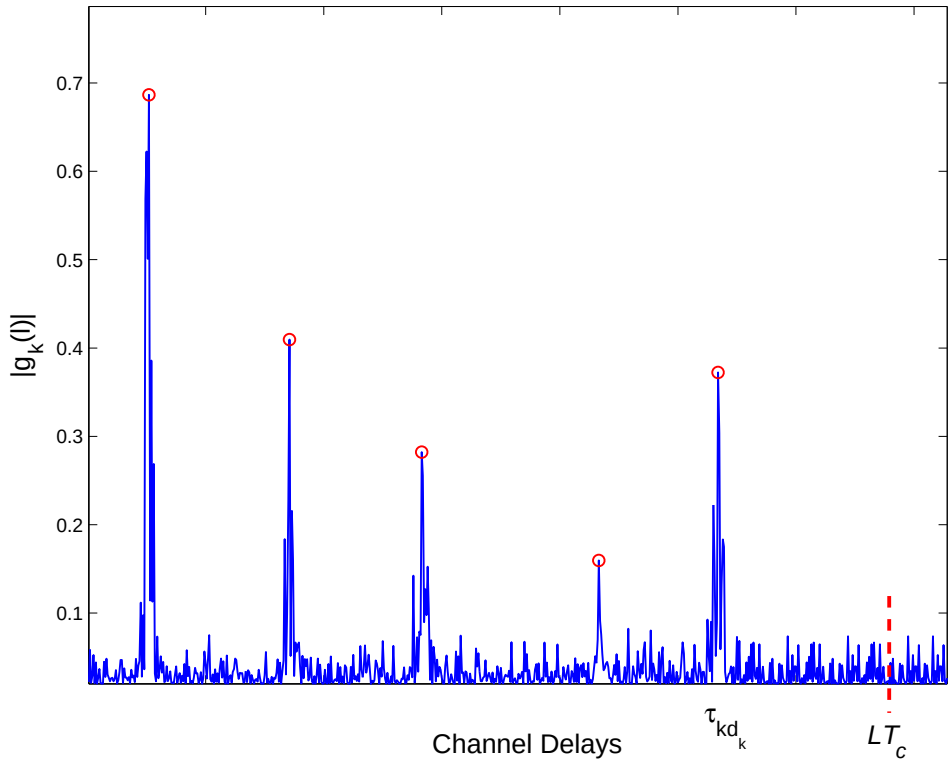


Figure 6.2: A scattered multipath propagation in a high SNR channel

- An MDL sphericity test based technique.
- A multiple hypothesis F-Statistics based technique.

Each technique requires an initial estimate of the channel parameters and this will be covered in the following section.

6.4 Channel Estimation: A Frequency Domain Based Approach

Consider user k to be the user of interest. After some manipulations, one can show from Eqns. (6.1) and (6.4) that the cross spectral density between the training

sequence $x_k(t)$ and the observed signal $y(t)$ is given by

$$C_{yx_k}(\omega) = C_{x_k x_k}(\omega) \sum_{l=1}^{d_k} g_k(l) e^{-j\omega\tau_{kl}} + C_R(\omega),$$

where

$$C_R(\omega) = \sum_{k'=1, k' \neq k}^{K} \left[C_{x_{k'} x_k}(\omega) \sum_{l=1}^{d_{k'}} g_{k'}(l) e^{-j\omega\tau_{k'l}} \right] + C_{v x_k}(\omega) \tag{6.6}$$

Define $\hat{C}_{yx_k}(\omega)$ and $\hat{C}_{x_k x_k}(\omega)$ as estimates for $C_{yx_k}(\omega)$ and $C_{x_k x_k}(\omega)$, respectively. Using the averaged periodogram [55] with length $\mathcal{M}$ to obtain $\hat{C}_{yx_k}(\omega)$ and $\hat{C}_{x_k x_k}(\omega)$, we establish the regression at discrete frequencies $\omega_m = \frac{2\pi m}{\mathcal{M}}$.

$$\begin{aligned} G_k(m) &= \hat{C}_{yx_k}(\omega_m) / \hat{C}_{x_k x_k}(\omega_m) \\ &= \sum_{l=1}^{d_k} g_k(l) e^{-j\frac{2\pi m}{\mathcal{M}}\tau_{kl}} + \varepsilon_k(m), \quad m = 0, \ldots, \mathcal{M}-1 \end{aligned} \tag{6.7}$$

where

$$\varepsilon_k(m) = \hat{C}_R(\omega_m) / \hat{C}_{x_k x_k}(\omega_m) \tag{6.8}$$

are estimation errors due to noise and the cross spectra term. Herein $\varepsilon_k(m),\ m = 0, \ldots, \mathcal{M}-1$ is assumed to be independent and identically distributed, which is valid asymptotically for large $\mathcal{M}$ under some regularity conditions (chapter 2) in [56].

6.4.1 The estimation procedure

From Eqn. (6.7), since d_k is unknown we assume the physical length $L > d_k$ paths, we consider the frequency domain model

$$G_k(m) = \sum_{l=1}^{L} g_k(l) e^{-j\Omega_{kl} m} + \varepsilon_k(m), \quad m = 0, \ldots, \mathcal{M}-1 \tag{6.9}$$

where, $\Omega_{kl} = \frac{2\pi\tau_{kl}}{\mathcal{M}},\ l = 1, \cdots, L$ are the unknown "frequencies" to be estimated, from which $\tau_{kl},\ l = 1, \cdots, L$ are found.

In (6.9) one can see that finding Ω_{kl} is a frequency estimation problem. Solution to this problem can be performed using methods such as MUSIC, ROOTMUSIC or ESPRIT [58] [59] [60]. However, one can also use other different methods to estimate these "frequencies". The estimation procedure based on the above mentioned methods is summarized in Table 6.1.

Table 6.1: Delays detection and estimation procedure

Step 1:

For a normalized sampling rate of $r/T_c = 1$, N training symbols, and given the training sequence $x_k(n)$ and observations $y(n)$, $n = 0, \ldots, N\mathcal{M} - 1$, compute $\hat{C}_{yx_k}(\omega_m)$, $\hat{C}_{x_kx_k}(\omega_m)$ using the averaged periodogram with length $\mathcal{M}$

Step 2

Find $G_k(m) = \hat{C}_{yx_k}(\omega_m)/\hat{C}_{x_kx_k}(\omega_m)$, $\omega_m = \frac{2\pi m}{\mathcal{M}}$, $m = 0, \ldots, \mathcal{M} - 1$.

Step 3:

Use MUSIC, ROOTMUSIC or ESPRIT techniques to estimate location of peaks of $G_k(m)$, Ω_{kl}, then estimate the delays $\tau_{kl} = \frac{\mathcal{M}\Omega_{kl}}{2\pi}$

Step 4:

Use LSE to estimate the multipath parameters for user k,

$\hat{\mathbf{g}}_k = [\hat{g}_k(1)\ \hat{g}_k(2) \ldots \hat{g}_k(L)]^T$,

$\hat{\mathbf{g}}_k = \left[\mathbf{A}^H\mathbf{A}\right]^{-1}\mathbf{A}^H\mathbf{G}_k$

where,

$(\mathbf{A})_{ml} = e^{-j\Omega_{kl}m}, m = 0, \ldots, \mathcal{M} - 1, l = 1, \ldots, L$

$\mathbf{G}_k = [G_k(0)\ G_k(1)\ \cdots\ G_k(\mathcal{M} - 1)]^T_{(\mathcal{M}\times 1)}$.

6.4.2 Remarks on the estimation procedure

By using directly the previous algorithm, we will end up with L different "frequencies" for each user, i.e., KL for all users. This will lead to a very long KL channel parameters, where most of them are either zero or non-significant. Thus, model selection techniques are required to decide which of these parameters are to be consider in the model, in other words, which parameter is significant. In the coming section we proposed two different approaches for this task, the first is an MDL sphericity test based approach and the second is a multiple hypothesis test F-Statistics based approach.

6.5 An MDL Sphericity Test Based Approach

The MDL criteria has a widespread use in model selection techniques mainly due to their intrinsic simplicity (see Chapter 5). The MDL criteria are mainly applied by evaluating two terms: a data term which requires the maximization of the log-likelihood and a penalty term which is a function of the complexity of the model [16] [17].

For the proposed model in Eqn. (6.9) with parameters $\mathbf{g}_k$ assigns the likelihood $f(G_k/\mathbf{g}_k)$ to the set of observed data G_k. The full form of the MDL measure for such a model family is given as,

$$\mathrm{MDL}(d) = -\log(f(G_k/\hat{\mathbf{g}}_k)) + p(d, \mathcal{M}) \tag{6.10}$$

where $\hat{\mathbf{g}}_k$ is an estimate of the parameters that maximize the likelihood, d is the number of parameters in the model, $\mathcal{M}$ is the sample size and $p(d, \mathcal{M})$ represents the MDL penalty term. In this approach we use the sphericity test to represent the likelihood function $f(G_k/\mathbf{g}_k)$.

6.5.1 The sphericity test

The sphericity test (ratio) is a simple and powerful test for model selection [46], the main use of this test is to check for variance or covariance multiplicity. To be familiar with the test let us consider the model in Eqn. (6.9) expressed in matrix notation,

$$\mathbf{G}_k = \mathbf{A}\mathbf{g}_k + \boldsymbol{\varepsilon}_k \tag{6.11}$$

where, $\mathbf{G}_k = [G_k(0)\ G_k(1) \dots G_k(\mathcal{M}-1)]^T$,
$\mathbf{A}$ is an $\mathcal{M} \times L$ matrix with $\mathbf{A}_{ml} = e^{-j\Omega_{kl}m}$,
$\mathbf{g}_k = [g_k(1)\ g_k(2) \dots g_k(L)]^T$,
$\boldsymbol{\varepsilon}_k = [\varepsilon_k(0)\ \varepsilon_k(1) \dots \varepsilon_k(\mathcal{M}-1)]^T$.
Define the covariance matrix $\mathbf{R} = \mathsf{E}\{\mathbf{G}_k\mathbf{G}_k^H\}$,

$$\mathbf{R} = \mathbf{A}\mathbf{g}_k\mathbf{g}_k^H\mathbf{A}^H + \sigma_{\varepsilon_k}^2\mathbf{I} \tag{6.12}$$

Since the true number of parameters, d, is unknown then we assume that $\mathbf{R}$ is a rank-L matrix.

The next step is then to use the SVD to find the eigenvalues of $\mathbf{R}$ and sort them such that, $\lambda_{(1)} > \lambda_{(2)} > \ldots > \lambda_{(L)}$.

These eigenvalues are biased and mutually correlated. Their finite sample joint distribution is known in the Gaussian case and is represented as a series of zonal polynomials [47] [48]. A mathematically tractable form for their asymptotic joint distribution does exist in the Gaussian case [49], although it may be unreliable for the small sample sizes considered here. In addition, this joint distribution is sensitive to departures from Gaussianity [50].

Once the eigenvalues are found then the sphericity test, $\mathrm{T}_{sph}(d)$ is defined by,

$$\mathrm{T}_{sph}(d) = \left[\frac{\prod_{l=d+1}^{L} \lambda_{(l)}^{1/(L-d)}}{\frac{1}{L-d} \sum_{l=d+1}^{L} \lambda_{(l)}} \right] \tag{6.13}$$

where the numerator and the denumerator parts represent the geometric mean and the the arithmetic mean respectively. When $\mathrm{T}_{sph}(d) \approx 1$, it means that the geometric mean is the same as the arithmetic mean or in other words $\lambda_{(d+1)} = \lambda_{(d+2)} = \ldots = \lambda_{(L)} = \sigma^2_{\varepsilon_k}$, i.e., the matrix $\mathbf{R}$ has a rank $(d+1)$ and the model has only d principle parameters instead of L (see Chapter 5 for more details).

6.5.2 MDL based on the sphericity test

The MDL criteria use $\mathrm{T}_{sph}(d)$ as the likelihood function, then it can defined by,

$$\mathrm{MDL}(d) = -\log \left[\mathrm{T}_{sph}(d)\right]^{(L-d)N} + p(d, \mathcal{M})$$

i.e.,

$$\mathrm{MDL}(d) = -\log \left[\frac{\prod_{l=d+1}^{L} \lambda_{(l)}^{1/(L-d)}}{\frac{1}{L-d} \sum_{l=d+1}^{L} \lambda_{(l)}} \right]^{(L-d)\mathcal{M}} + \tfrac{1}{2} d(2L-d) \log(\mathcal{M}) \tag{6.14}$$

and finally, an estimate of the number of principle parameters is given by,

$$\hat{d} = \underset{d \in \{1,2,\ldots,L\}}{\operatorname{argmin}} \ \mathrm{MDL}(d) \tag{6.15}$$

Once $\hat{d}$ is estimated, which obviously will be much less than L, then the following procedure can be followed to estimate the channel principle parameters and based on these estimates, obtain parsimonious receivers. The estimation and identification procedure is summarized in Table 6.2.

Table 6.2: The MDL spherecity based estimation and identification procedure

Step 1: Select the user of interest say the k'th user and compute $\hat{C}_{yx_k}(\omega_m)$, $\hat{C}_{x_kx_k}(\omega_m)$ by the averaged periodogram technique over the symbol length $\mathcal{M}$ given the training sequence $x_k(n)$ and observations $y(n)$, $n = 0, \cdots, N\mathcal{M}-1$.

Step 2: Find $G_k(m) = \hat{C}_{yx_k}(\omega_m)/\hat{C}_{x_kx_k}(\omega_m)$, $\omega_m = \frac{2\pi m}{\mathcal{M}}$, $m = 0, .., \mathcal{M}-1$.

Step 3: Estimate the frequency domain covariance matrix $\mathbf{R}$, and sort its eigenvalues such that $\lambda_{(1)} > \lambda_{(2)} > \ldots > \lambda_{(L)}$

Step 4: For $d = 1, 2, \ldots, L$ use the MDL sphericity based method in Eqn (6.14) to estimate the true number of delays (parameters), $\hat{d}$.

Step 5: Supply the estimated number of delays $\hat{d}$ and the covariance matrix $\mathbf{R}$ to say the ROOTMUSIC algorithm and estimate the true channel delays.

Step 6: For each identified delay, estimate non-zero multipath parameters,
$\hat{\mathbf{g}}_k = \left[\mathbf{A}^H\mathbf{A}\right]^{-1}\mathbf{A}^H\mathbf{G}_k$
where,
$(\mathbf{A})_{ml} = e^{-j\Omega_{kl}m}$, $m = 0, \ldots, M-1$ and $\Omega_{kl} = \frac{2\pi\tau_{kl}}{\mathcal{M}}$ are identified delays only,
$\mathbf{G}_k = [G_k(0)\ G_k(1)\ \cdots\ G_k(\mathcal{M}-1)]^T$.

6.6 Hypothesis Tests F-Statistics Based Approach

Assume the model in Eqn. (6.5) where for each user k there are L channel parameters, $\mathbf{g}_k = [g_k(1)\ g_k(2) \ldots g_k(L)]^T$. We need to identify whether a path parameter, say, $g(l) = g_k(l)$, is significant in the model or not, i.e., we wish to test at a specific level of significance γ the hypothesis

$$\begin{aligned} &\mathcal{H}_l : g(l) = 0 \\ &\qquad \text{versus} \\ &\mathcal{K}_l : g(l) \neq 0 \end{aligned} \tag{6.16}$$

where $\mathcal{H}_l$ represents the null hypothesis that there is no significant path parameter at the corresponding delay τ_l and $\mathcal{K}_l$ is the alternative hypothesis that there is a significant path at that delay.

To implement this test, first we define a test statistic, T, then we find the distri-

bution of T under the null hypothesis $\mathcal{H}_l$.

For each parameter, we calculate the probability values ($\mathcal{P}$-values),

$$\mathcal{P}_l = Pr(T \geq T_{g(l)}|\mathcal{H}_l),\ l = 1, 2, \ldots, L \tag{6.17}$$

where $T_{g(l)}$ is the threshold for the parameter $g(l)$ under the null hypothesis $\mathcal{H}_l$, see Fig. 6.3. The $\mathcal{P}$-values are then compared to the level of significant γ such that if $\mathcal{P}_l > \gamma$, then we retain $\mathcal{H}_l$, and if $\mathcal{P}_l \leq \gamma$ we reject $\mathcal{H}_l$

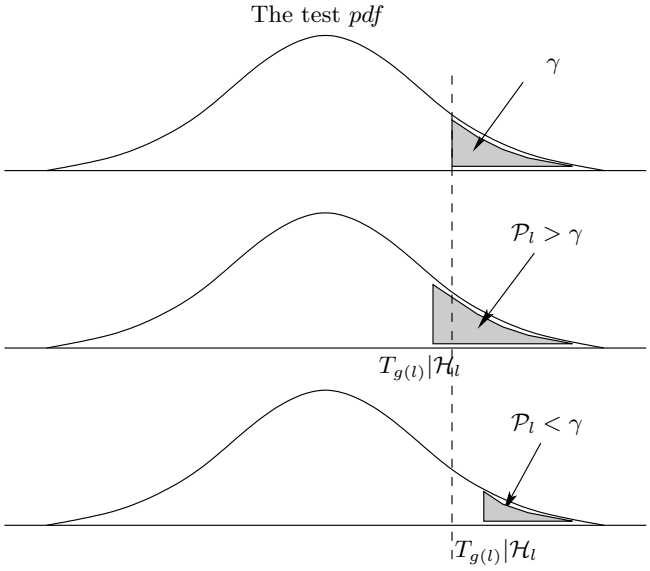


Figure 6.3: Testing the probability value of $g(l)$

6.6.1 Classical and Sequentially Rejective Bonferroni tests

In the above hypothesis test just described, one can see that the test is to be performed with respect to several parameters. When performing L multiple independent tests $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_L$ each at the γ level of significance, the Family Wise Error, FWE, which is the probability of making at least one Type I error (rejecting the null hypothesis inappropriately) is $1 - (1 - \gamma)^L$ is $\gg \gamma$. However, in order to maintain an FWE at a chosen level γ, if each hypothesis is tested separately using tests with level $\frac{\gamma}{L}$, then it follows immediately from the Boole inequality that the probability of rejecting any true hypothesis is less than or equal to γ. This constitutes a multiple test procedure with the global level of

significance γ for free combinations. The classical test is depicted in Fig. 6.5. In this test the hypotheses $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_L$ are assumed to be independent. For dependent hypotheses the test becomes overly conservative and is less powerful [52] [54] [53].

A more powerful test is the Sequentially Rejective Bonferroni (SRB) test, which can be used when the hypothesis tests are dependent. In this work we use a SRB test based on F-Statistics [40] [51]. First we sort the $\mathcal{P}$-values such that, $\mathcal{P}_{(1)} < \mathcal{P}_{(2)} < \ldots < \mathcal{P}_{(L)}$. Then we follow the algorithm shown in Fig. 6.4 to identify the significant model parameters.

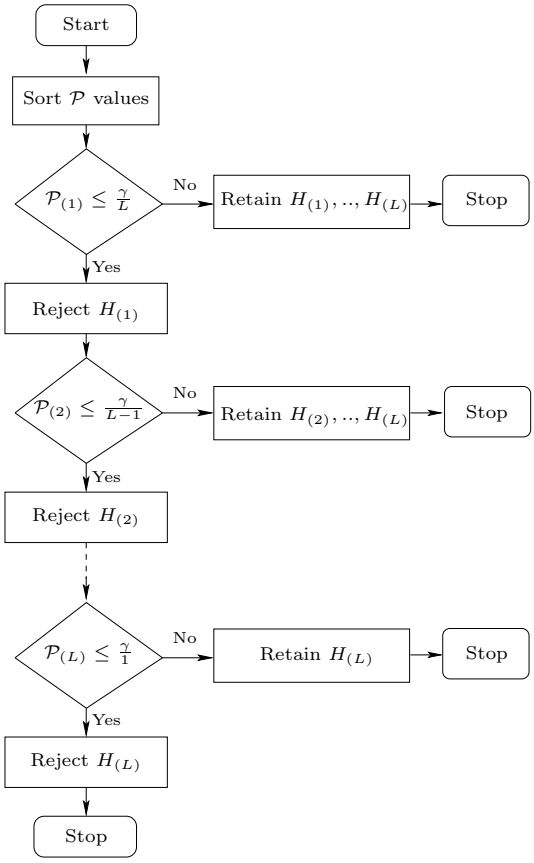


Figure 6.4: Sequentially Rejective Bonferroni tests (SRB)

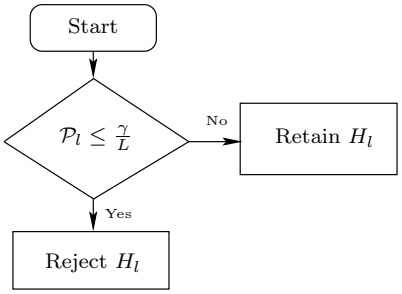


Figure 6.5: Classical Bonferroni test

6.6.2 Backward Elimination

Using the $\mathcal{P}$-values, the Backward Elimination (BE) method can also be used to identify the significant model parameters. The procedure begins with the full model that includes all considered parameters from $g(1)$ to $g(L)$. It then attempts to remove one parameter at a time by determining whether the least significant parameter currently in the model can be removed or not because its $\mathcal{P}$-value is greater or less than the level of significance γ. Once a variable has been removed from the model it cannot re-enter at a subsequent step. The procedure is shown in Fig. 6.6 and can be summarized in Table 6.3 as follows,

Table 6.3: The Backward Elimination based method

Step 1: Set the model length L and then estimate the full model parameters $\hat{\mathbf{g}}_k = [\hat{g}_k(1)\ \hat{g}_k(2) \cdots \hat{g}_k(L)]^T$.
Step 2: Use a pre-defined statistical test, find the parameters $\mathcal{P}$-values, $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_L$.
Step 3: Remove from the model the parameter that is insignificant, i.e., $\mathcal{P}_l > \gamma$ and has the highest $\mathcal{P}$-value (one parameter at a time).
Step 4: Re-estimate the new model and again the parameter that is insignificant and has the highest $\mathcal{P}$-value is deleted.
Step 5: Continue until all remaining parameters are significant.

To construct the statistical test mentioned in this section and find the $\mathcal{P}$-values that Section 6.6.1 and Section 6.6.2 require, we need to consider the test statistics distribution T under the null hypothesis $\mathcal{H}_l$, i.e., the *pdf* of $T|\mathcal{H}_l$. This task will be addressed in the following section.

6.6.3 Calculation of the $\mathcal{P}$-values

Suppose channel parameters have been estimated. The next step is to find the $\mathcal{P}$-values so that one can apply the BE or the SRBT as described in Section 6.6. In this section we follow an F-Statistic to find the $\mathcal{P}$-values see (Chapter 5 and [40] [41] [51]).

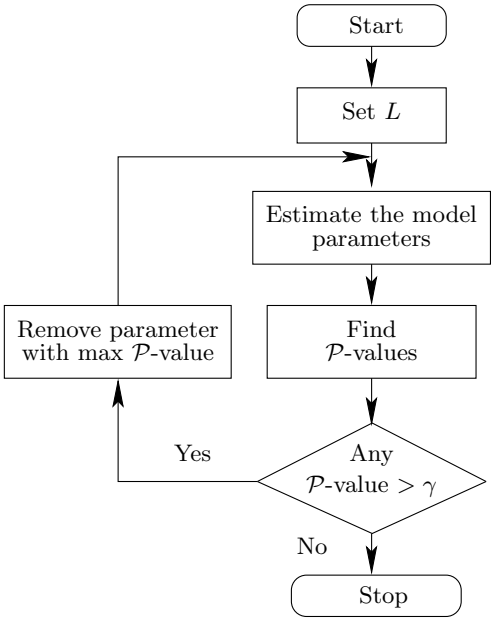


Figure 6.6: Backward Elimination based method for model selection

Assume we have $\mathcal{M}$ samples of the linear model given in Eqn. (6.9). When comparing two models one of which is a reduced from the other, we consider the following test statistic for T in Section 6.6,

$$T = \frac{(SSR_f - SSR_r)/(DF_f - DF_r)}{SSE_f/DF_m}, \tag{6.18}$$

where

- SSR_f and DF_f are the Regression Sum Squared Error (SSR) and the number of degree of freedom (DF) of the full model respectively
- SSR_r and DF_r are SSR and DF of the reduced model respectively
- SSE_f is the full model Sum Squared Error (SSE) with degree of freedom DF_m

Assume that the distribution of the model residual error is complex Gaussian for a large data size under some regularity conditions [57]. The test statistic, T then becomes a ratio between two χ^2 distributions with DF equal to $(DF_f - DF_r)$ and DF_m respectively, which corresponds to an F-distribution with DF $(DF_f - DF_r)$ and DF_m, (see [41] [51] and Chapter 5 for more details), i.e.,

$$T \sim F(DF_f - DF_r, DF_m) \tag{6.19}$$

For calculation of the $\mathcal{P}$-values, we reduce the full model by one parameter at a time say $g_k(l)$, then we find the parameter's $\mathcal{P}$-value,

$$\mathcal{P}_l = Pr\left(T > F_{g_k(l)}|\mathcal{H}_l\right) \tag{6.20}$$

where $F_{g_k(l)}$ is calculated using Eqn. (6.18) for a model which does not include the parameter $g_k(l)$.

Once one has found the full model $\mathcal{P}$-values, $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_L$, then either the BE or the SRBT in Section 6.6 based on the F-Statistics above can be used to find the best fit model. Obviously this model includes only the significant parameters.

6.7 Simulation Results

The following examples demonstrate the performance of the delay estimation algorithm in Section 6.4.1, the SRB test in section 6.6.1, and the performance of the proposed minimal complexity RAKE receiver. In each example, the kth user's training data $x_k(t)$ is generated using Eqn. (1). Each symbol is modulated by a Maximum Length Binary Sequence (MLBS) spreading code $s_k(t)$. In all examples we show the first user results where d_k represents the true number of paths, d_o represents the channel length,L represents the physical number of paths and M_g is the maximum lag which is used in MUSIC, ROOTMUSIC and ESPRIT covariance matrices. Finally, the channel attenuation $g_k(t)$ corresponds to a frequency selective slowly fading channel (see Section 3.3) and $v(t)$ is a zero mean complex Gaussian noise process. The SNR is the value of $\frac{E_s}{\sigma_n^2}$ in dB, where E_s represents the Energy per symbol and σ_v^2 is the noise power. The channel response is generated such that each user's channel response has a unit norm, i.e., the SNR all multipaths/symbol.

6.7.1 The estimation approach performance

6.7.1.1 Example 1

This example investigates the proposed algorithm in Table 6.1 for delay detection and estimation. It also considers the effect of increasing the number of users in

the estimation results, the Multiple Access Cross Spectral Interference (MACSI) contribution. The input settings and the channel parameters for this case are given in Table 6.4. Simulation results for delay estimation are shown in Fig. 6.7 and Table 6.5.

6.7.1.2 Discussion of Example 1

In Example 1, we applied the estimation algorithm for $K = 8$ users. Results by MUSIC, ROOTMUSIC and ESPRIT given in Fig. 6.7 and Table 6.5 show that the contribution of the noise and the MACSI term, $\varepsilon_k(m)$ clearly has a significant effect. They produced undesired false peaks with semi-significant amplitudes for delays $\tau > 10T_c$. Results of this example show that, with these conditions, the algorithm successfully estimated the channel delays. Since the results in Table 6.5 are quite similar when rounded for the RAKE receiver, then any of the frequency estimation methods can be used. In this work, we use the MUSIC algorithm.

Table 6.4: Input settings and multipath parameters for Example 1

Input settings	
K	8
N	64
M	32
d_o	10
d_k	5
L	16
M_g	16
SNR(dB)	4

Delay τ_{1l}	Multipath parameters $g_1(l)$
T_c	$-0.0027 - j0.2281$
$2T_c$	0
$3T_c$	$0.2584 + j0.4406$
$4T_c$	$-0.0078 - j0.4460$
$5T_c$	0
$6T_c$	0
$7T_c$	0
$8T_c$	$0.3534 - j0.0768$
$9T_c$	$0.0900 - j0.2823$

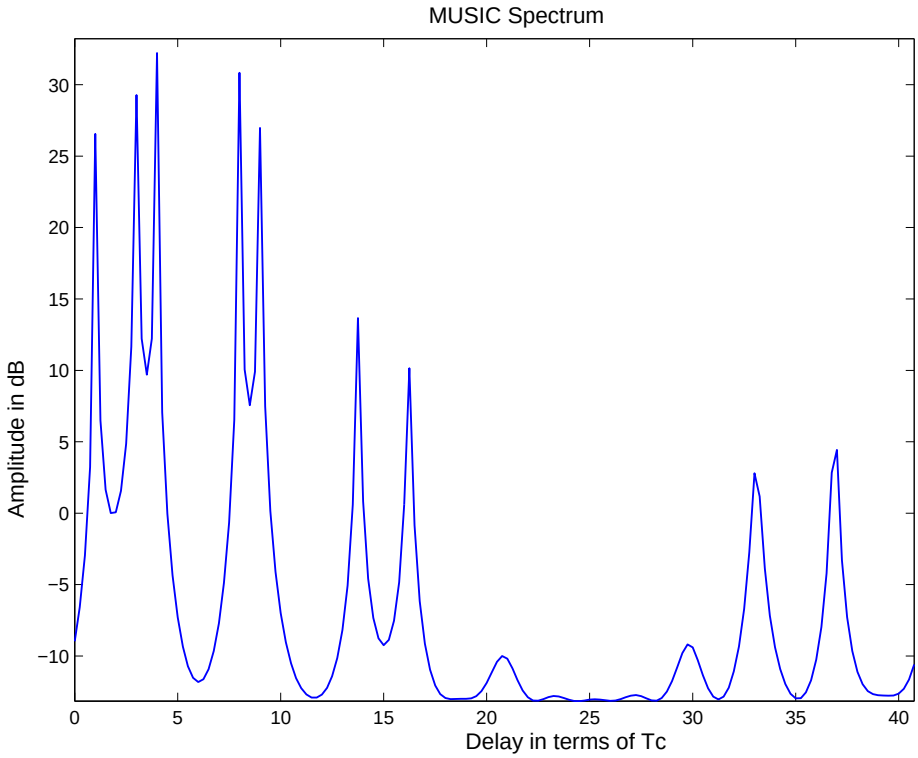


Figure 6.7: Example 1: Delay estimation using MUSIC

Table 6.5: Example 1: Detection results

True Delays $\tau_k(l)/T_c$	1	3	4	8	9			
MUSIC	1.00	3.00	4.00	8.00	9.00	13.75	16.25	...
RMUSIC	1.0193	2.9705	3.9887	7.9949	9.0183	13.0457	16.2345	...
ESPRIT	1.0115	2.9592	3.9565	7.9833	9.0619	13.5788	16.5174	...

6.7.1.3 Example 2

This example consider the case where some delays are closer enough such that the algorithm cannot correctly identify them. The input data and the channel parameters are given in Table 6.6, simulation results are also shown Fig. 6.8 and Table 6.7.

6.7.1.4 Discussion of Example 2

In situations where the channel delays are close enough or the MACSI has highly significant effects, the algorithm may fail to detect one or more of the channel

delays. This situation is presented in this example when the algorithm failed to detect the channel delays at $3T_c$ and $6T_c$. To overcome this problem one can either increases the signal power (SNR), the training symbols and/or the sampling rate.

Table 6.6: Input settings and multipath parameters for Example 2

Input settings	
K	8
N	64
M	32
d_o	10
d_k	7
L	16
M_g	16
SNR(dB)	4

Delay τ_{1l}	Multipath parameters $g_1(l)$
$2T_c$	$-0.5611 + j0.1865$
$3T_c$	$-0.2246 + j0.2523$
$4T_c$	$-0.2246 + j0.2523$
$5T_c$	$0.1649 + j0.1218$
$6T_c$	$-0.1327 - j0.4014$
$7T_c$	$-0.2246 + j0.2523$
$8T_c$	0
$9T_c$	$-0.2717 - j0.0298$

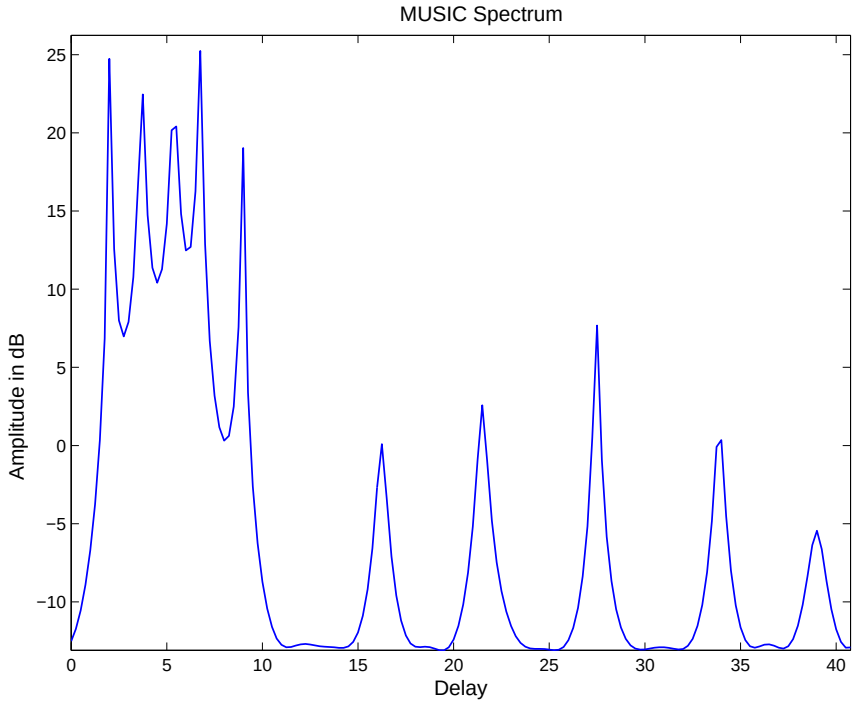


Figure 6.8: Example 2: Delay estimation using MUSIC

Table 6.7: Example 2: Detection results

True Delays τ_{1l}/T_c		2.0	3.0	4.0	5.0	6.0	7.0	9.0		
MUSIC		2.0000		3.7500	5.5000		6.7500	9.0000	16.2500	...
RMUSIC		2.0379		3.6943	5.3761		6.7405	8.9671	10.0804	...
ESPRIT	1.3019	2.1470		3.8838	5.5505		6.8140	8.9793	17.0152	...

6.7.2 Identification performance

6.7.2.1 Example 3

This example demonstrates the use of the MDL sphericity test described in Section 6.5 and the SRB F-Statistics based test in Section 6.6.1 and Section 6.6.3 in identifying the significant channel parameters of a frequency selective slowly fading channel introduced in a $K = 1, 2, 3, 4$ users system with maximum channel length of $d_o = 10$ with $d_k = 3$ significant paths, $L = 16$ required physical taps, $M = 32$ chips and $N = 32,\ 64,\ 96$ and 128 symbols.The channel components are given in Table 6.8. Let $p_k(l)$ be the probability of identifying each individual path parameter $g_k(l)$ correctly, then one can show that the probability of identifying the full channel parameters of user k correctly p_f can be given by,

$$p_f = \prod_{l=1}^{L} p_k(l) \tag{6.21}$$

Results are shown in Fig. 6.9, Fig. 6.10 and Fig. 6.11, Fig. 6.12 for $\gamma = 0.05$.

6.7.2.2 Discussion of Example 3

From the results, one can see that both of the identification tests work well at a low SNR and short N compared with many techniques that are based on an $N \geq 512$ symbols. For example when $N = 32$ it is better to run the system with an SNR > 4 dB and for $N \geq 64$ symbols the system can run over an SNR > 2 dB. In practice, many wireless communication systems operate with SNRs greater than 8 dB. If the system is running at SNRs less than 4 dB, a longer training sequence is then required. Using the averaged periodogram, long training sequences will reduce the effect of MACSI and the detection rate will be then

Table 6.8: Input settings and multipath parameters for Example 3

Delay	Multipath parameters $g_k(l)$ for			
τ_{1l}	User 1	User 2	User 3	User 4
T_c	-0.0207 - j0.4087	0	0	-0.2290 + j0.1217
$2T_c$	0	0.0347 + j0.4313	0.1423 -j0.0278	0.1732 - j0.2147
$3T_c$	0	0	0	-0.4983 - j0.1737
$4T_c$	0	0.0347 + j0.4313	-0.4201 + j0.4713	0
$5T_c$	0.5350 + j0.0278	0	0.6119 + j0.4301	0
$6T_c$	0.6730 + j0.3042	-0.5138 - j0.1791	0	-0.2825 + j0.4796
$7T_c$	0	-0.0842 - j0.3971	0	-0.3364 + j0.3774
$8T_c$	0	0	0	-0.0598 - j0.0957
$9T_c$	0	-0.0842 - j0.3971	0.1423 - j0.0278	0
$10T_c$	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$16T_c$	0	0	0	0

enhanced. It is also clear that the *F*-Statistics SRB test based method has a better performance over the MDL sphericity test based method. For a certain level of significance, the *F*-Statistics based method can check separately the significance of each individual parameter where the MDL method does not have a level of significance also the full algorithm also gave the over all number of paths instead of identifying each individual path. However both methods can be used for the identification of the channel non-zero parameters with high probabilities as shown in Figs. 6.10 and Figs. 6.12.

In Fig. 6.9 it is also clear that the MDL based method can correctly identify the true number of paths and in Fig. 6.11 one can see that at low SNR's the *F*-Statistics SRB test can easily identify the channel zero or non-significant parameters with high probabilities.

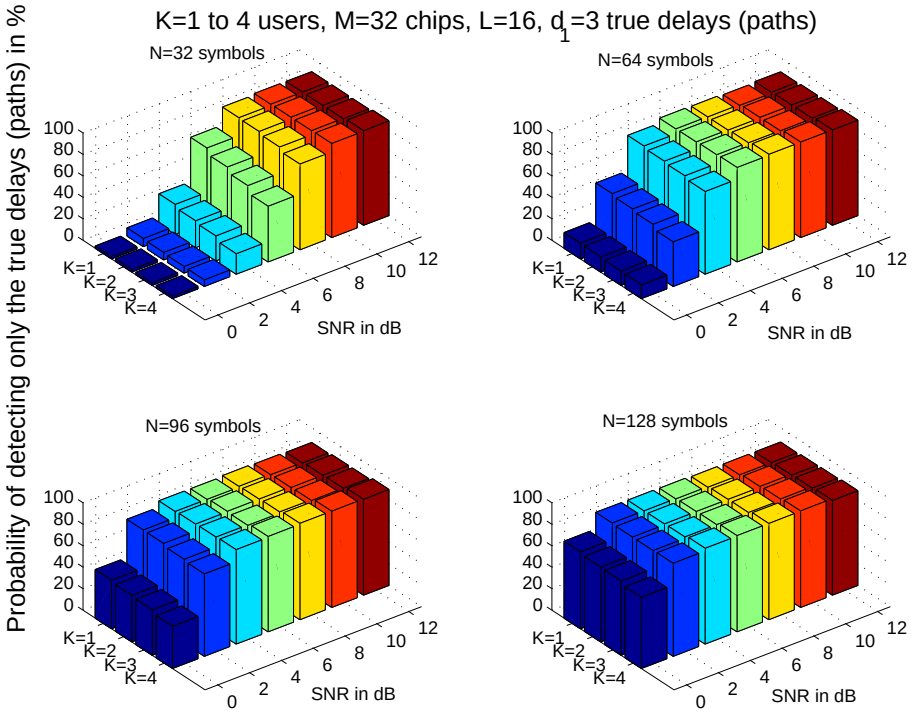


Figure 6.9: The probability of correctly detecting user's one true delays (paths), MDL method

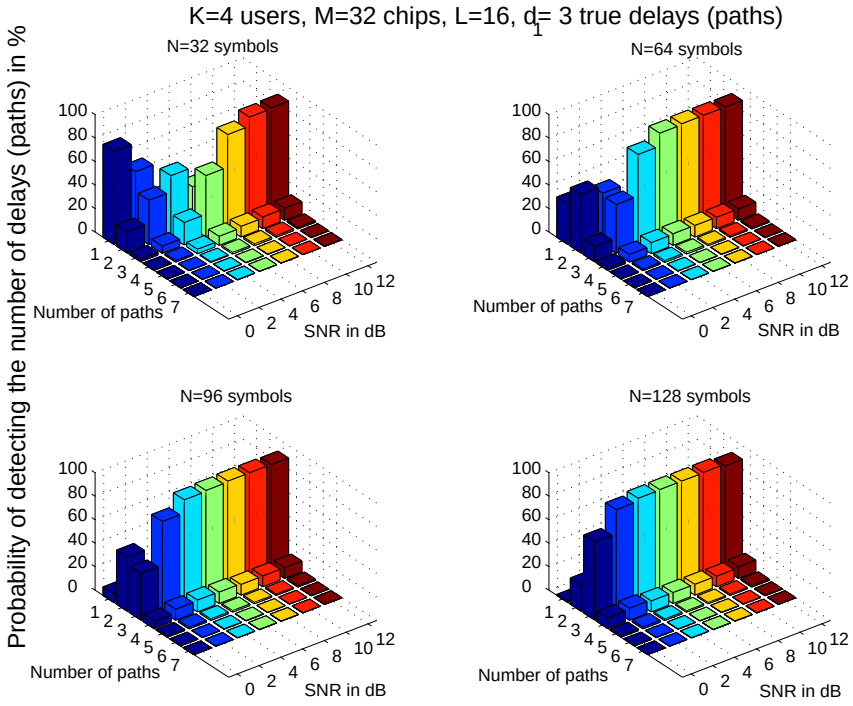


Figure 6.10: The probability of detecting all users delays (paths), MDL method

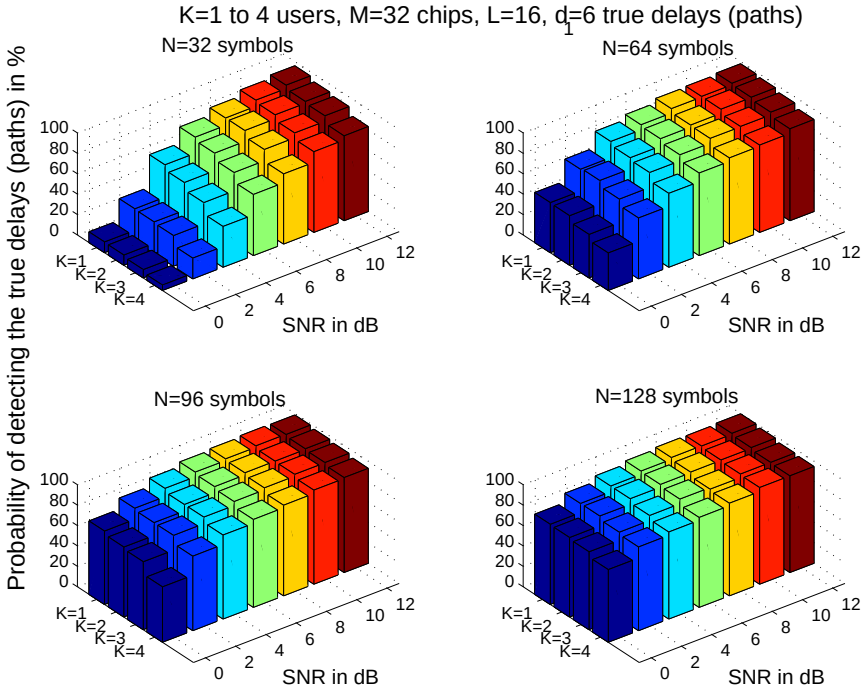


Figure 6.11: The probability of correctly detecting user's one non-zero parameters (paths), SRB test

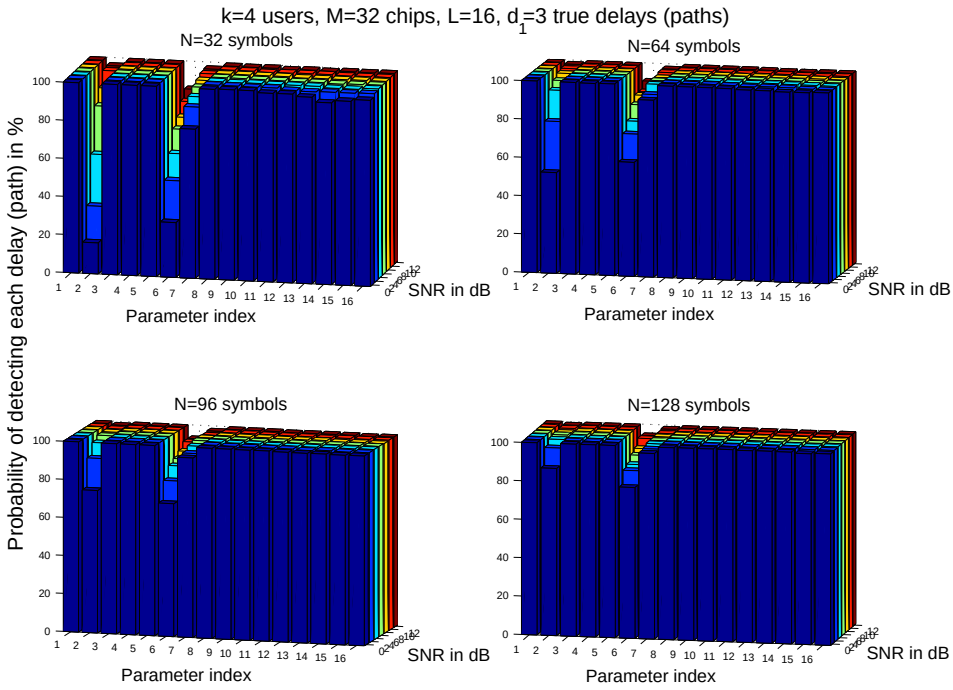


Figure 6.12: The probability of correctly detecting user's one all parameters (paths), SRB test

6.7.3 System performance

To assess the performance of the proposed RAKE receiver which is based on the identified parameters, the following examples compare the Bit Error Rate (BER) when:

(1) Perfect knowledge of the channel, $\mathbf{g}_P$ is used

(2) Channel estimated with the same known length d_o, $\hat{\mathbf{g}}_{d_o}$

(3) Channel estimated with length L, $\hat{\mathbf{g}}_L$

(4) Channel estimated using the new frequency based approach, $\hat{\mathbf{g}}_n$

The channel is assumed to be a frequency selective slowly fading channel with parameters chosen randomly from the Rayleigh distribution. A summary of the comparison procedure is shown in Table 6.9.

6.7.3.1 Example 4:

In this simulation we apply the procedure in Table 6.9 and the MDL sphericity test based approach to calculate the performance of two different systems over a frequency selective slowly fading channel. The first system has $K = 2$ users, $M = 32$ chips, length $d_o = 8$ paths, $L = 16$ taps and $N = 32$ symbols, where the second has $K = 4$ users, $M = 32$ chips, length $d_o = 8$ paths, $L = 16$ taps and $N = 32$ symbols. In both cases the channel has $d_k = 3$ significant paths. Results are shown in Fig. 6.13 and Fig. 6.14 respectively for the first user.

6.7.3.2 Discussion of Example 4:

In Fig. 6.13 and Fig. 6.14. Simulation BERs show that the performance of the proposed frequency domain MDL sphericity test based RAKE receiver is much improved over other classical methods such as direct implementation and or when the channel order is known. For $K = 2$ or 4 users, the new approach has the smallest BER except for, when the channel is known exactly. Using a few training symbols, $N = 32$, and increasing the number of users to 4, as in Fig. 6.14, the MACSI reduces the quality of the estimated parameters which reduces the performance of all methods.

Table 6.9: The performance comparison procedure

- For $k = 1, .., K$, define the channel parameters:

  1- $\mathbf{g}_P$ for perfect channel.

  2- $\hat{\mathbf{g}}_{d_o}$ estimated with the same channel length d_o.

  3- $\hat{\mathbf{g}}_L$ estimated with direct implementation of length L.

  4- $\hat{\mathbf{g}}_n$ identified and estimated using the proposed frequency domain approach.

- Let $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_K]^T_{(K\times 1)}$ be the transmitted symbols

- Compute

  - $\mathbf{S} = [\mathbf{s}_1 \ \mathbf{s}_2 \ \dots \ \mathbf{s}_K]_{(M\times K)}$, the normalized signature matrix.

  - $\mathbf{R} = \mathbf{S}^T\mathbf{S}$, the signature covariance matrix

- Consider the BER for the channel case, say, $\hat{\mathbf{g}}_L$,

  - $\hat{\mathbf{g}}_L = [\hat{\mathbf{g}}_1 \ \hat{\mathbf{g}}_2 \ \dots \ \hat{\mathbf{g}}_K]_{(L\times K)}$,

  - $\hat{\mathbf{g}}_k = [\hat{g}_k(1) \ \hat{g}_k(2) \ \dots \ \hat{g}_k(L)]^T$

  - Find $\mathbf{D} = \mathbf{R}^{-1}\mathbf{W}^H$

  where,

  $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_K]_{(\mathcal{M}\times K)}$,

  $$\mathbf{w}_k = \begin{bmatrix} s_k(1) & & \mathbf{0} \\ \vdots & \ddots & s_k(1) \\ s_k(M) & & \vdots \\ \mathbf{0} & \ddots & s_k(M) \end{bmatrix}_{(\mathcal{M}\times L)} \times \begin{bmatrix} \hat{g}_k(1) \\ \hat{g}_k(2) \\ \vdots \\ \hat{g}_k(L) \end{bmatrix}_{(L\times 1)}$$

- Estimate the transmitted symbols,

  - $\hat{\mathbf{b}} = \text{sgn}\{\Re(\mathbf{D}\ \mathbf{y})\}$

  - $\mathbf{y} = [y(0) \ y(2) \ \dots y(\mathcal{M}-1)]^T_{(\mathcal{M}\times 1)}$

- Choose another estimated channel and compare the BER for each case.

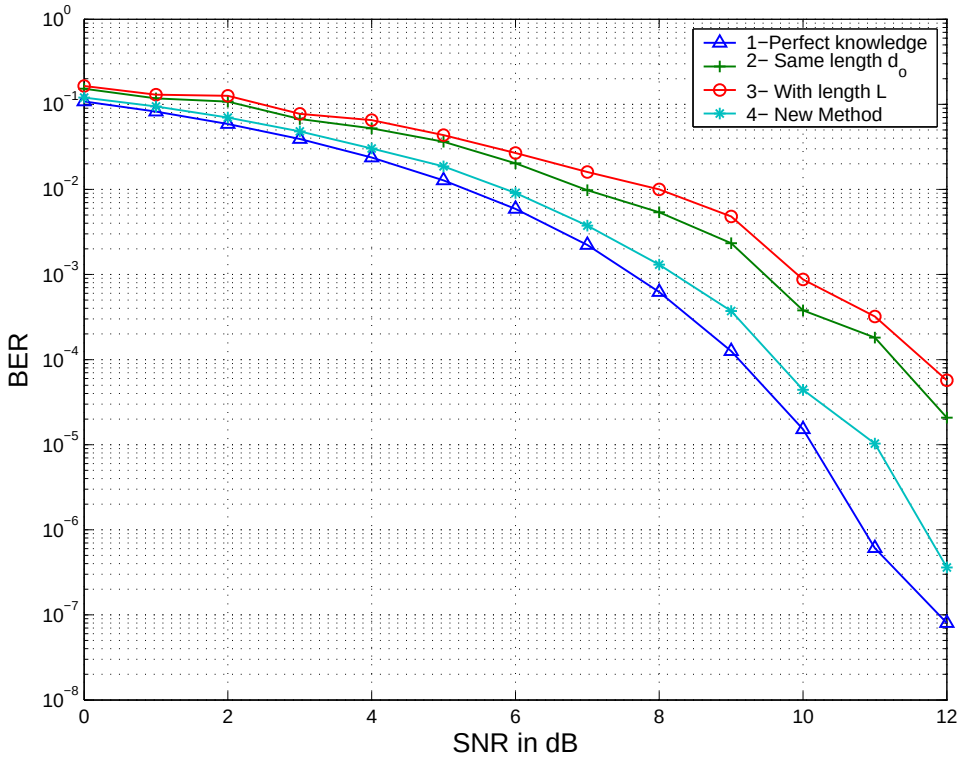


Figure 6.13: Example 4: BER in the case of 2 users, MDL sphericity test

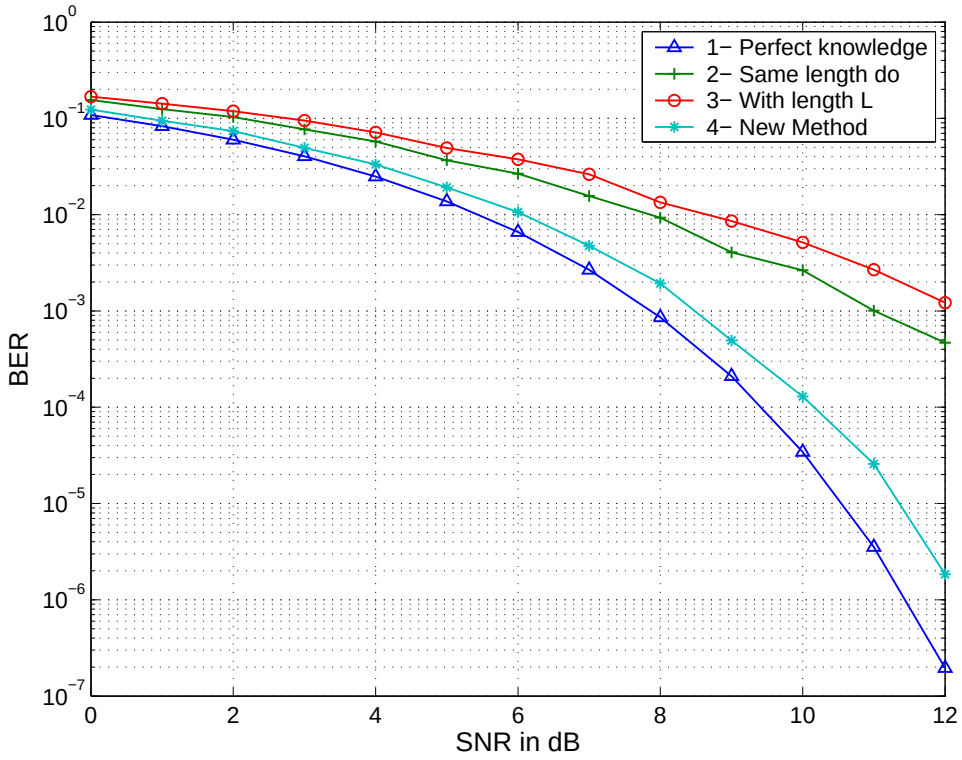


Figure 6.14: Example 4: BER in the case of 4 users, MDL sphericity test .

6.7.3.3 Example 5

In this simulation we apply the previous procedure and the F-Statistics SRB test based method to calculate the performance of two different systems over a frequency selective slowly fading channel. In this simulation the channel response is more difficult than the previous example since it has more users and more significant channel paths. The first system has $K = 8$ users, $M = 32$ chips, length $d_o = 10$ paths, $L = 16$ taps and $N = 64$ symbols, where the second has $K = 16$ users, $M = 32$ chips, length $d_o = 10$ paths, $L = 16$ taps and $N = 64$ symbols. In both cases the channel has $d_k = 6$ significant paths. Results are shown in Fig. 6.15 and Fig. 6.16 respectively for the first user.

6.7.3.4 Discussion of Example 5:

In Fig. 6.15 and Fig. 6.16, simulation BERs show that the performance of the proposed frequency domain based RAKE receiver is much improved over other classical methods such as direct implementation and or when the channel order is known. For $K = 8$ or 16 users, the new approach has the smallest BER except for when the channel is known exactly. Using a few training symbols, $N = 128$, and increasing the number of users to 16, as in Fig. 6.16, the MACSI reduced the quality of the estimated parameters which reduced the performance of all methods.

6.7.4 Structure limitation

When implementing RAKE receivers, there is a limitation to the number of fingers that can be used. A practical RAKE receiver is usually a combination of a software and hardware circuits that covers a certain delay range. According to the estimated channel delays, the software part controls certain digital switches on or off such that only these delays that are significant are considered. However in rare situations the number of identified significant paths d_k or the channel length d_o itself, can be greater than the number of the implemented RAKE fingers. In this situation we have to modify the algorithm such that it will only choose the

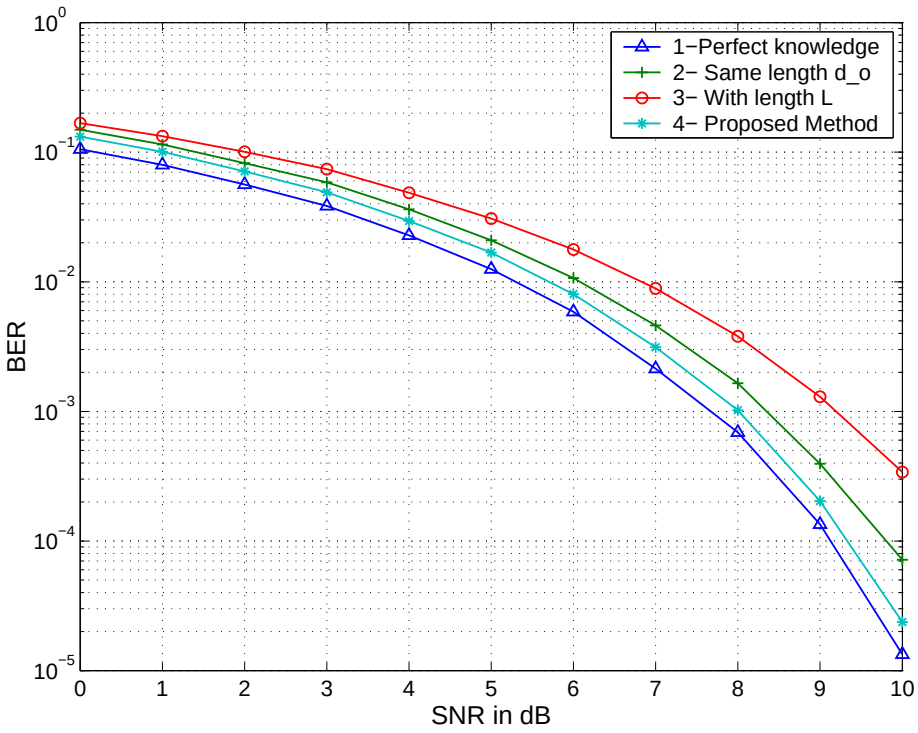


Figure 6.15: Example 5: BER in the case of 8 users, F-Statistics SRB test

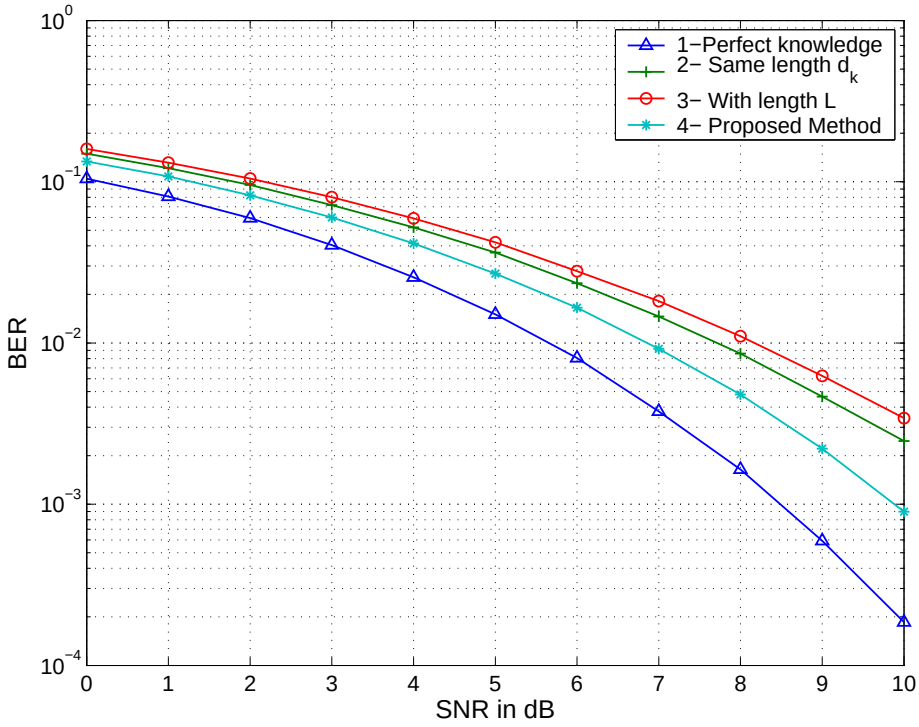


Figure 6.16: Example 5: BER in the case of 16 users, F-Statistics SRB test.

most significant paths. This task can be accomplished using each path's $\mathcal{P}$-value. For clarification, let d_f be the number of the implemented RAKE fingers, then one can stop the SRB test given before when the decision involves only the d_f most significant paths. To become familiar with this situation let us consider the following example.

6.7.4.1 Example 6

This example considers the above structure limitation for $K = 4$ and $K = 2$ users, $N = 128$ symbols, $M = 32$ chips, $L = 12$, $d_f = 4$ fingers $d_o = 10$ and $d_k = 6$ significant paths. In Fig. 6.17 the channel response is chosen so that one of the six paths is very less significant, but in Fig. 6.18 all of the channel paths are significant.

6.7.4.2 Discussion of Example 6

Due to some practical limitations, one may not be able to consider all identified paths. Fig. 6.17 and Fig. 6.18 show the BERs of two RAKE receivers that consider only $d_f = 4$ fingers out of the required 6 fingers of the full model. In Fig. 6.17 the performance of the practical 4 fingers RAKE receiver is close to the new proposed method than the classical method or the known length one. This situation can happen when one or more paths are weak or not significant. In Fig. 6.18 where all the channel paths are significant, the performance of the practical RAKE receiver is still better than the direct method with L fingers, but not as good as the case where d_k fingers are considered or the fully identified channel.

6.8 Conclusion

This chapter addressed the problem of minimizing the RAKE receiver structure complexity and enhancing its performance over classical methods. This is done through the identification and estimation of the significant channel parameters. A frequency domain based delay estimation approach is first proposed to estimate the channel delays and attenuation. Then two model selection approaches

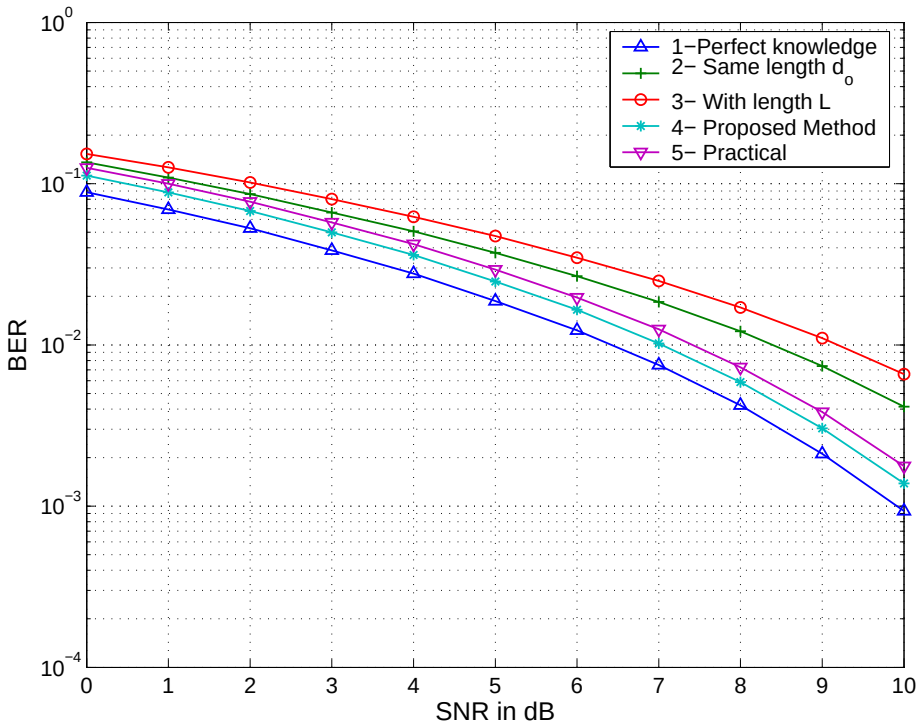


Figure 6.17: Example 6: BER for $K = 4$, $d_f = 4$, $d_k = 6$ with a weak path

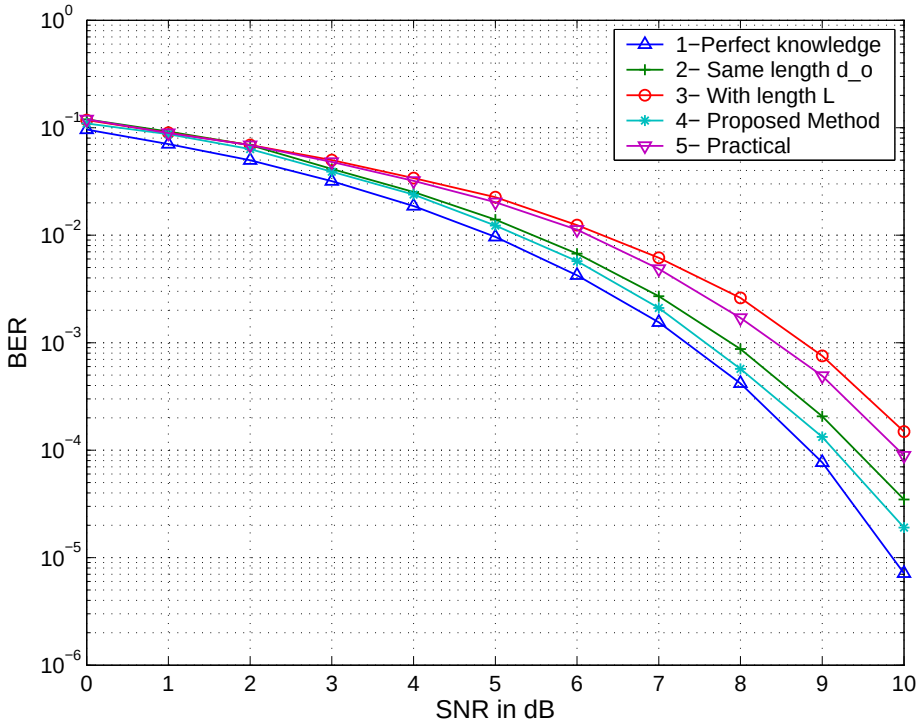


Figure 6.18: Example 6: BER for $K = 2$, $d_f = 4$, $d_k = 6$ with no weak path

including the MDL sphericity test based approach and the multiple hypothesis F-Statistics based tests were used to achieve the identification goal. Both approaches are low in computation complexity since they only involves the FFT to estimate the spectra. The F-Statistics SRB based approach is found to be more robust than the MDL based method, firstly because it can control the level of significance and secondly because it uses the SVD only once for the calculation of the channel delays using MUSIC algorithm. The identification and BER simulation results support the claim that with short training data, the structure complexity of the RAKE receiver is reduced and the system performance is also enhanced. The usage of the significant multipath components only is strongly recommended based on our findings.

Chapter 7

Channel Estimation and Identification Using Time Domain Bootstrap Based Approaches

7.1 Introduction

In this chapter we consider two bootstrap based approaches for the design of parsimonious DS-CDMA RAKE receivers. We consider the case where the multipath propagation channels are frequency selective slowly fading. A training based strategy is proposed to identify low order estimates of the channel parameters. This task is done in the situation where there are minimal information about the noise distribution, in other words, no specific distribution of the statistical test as in Chapter 6 is considered. In this approach, we first model the channel as an FIR filter with length related to the maximum delay spread of the channel. Then with low complexity, we estimate the model parameters directly in the time domain. Finally, based on the estimates, bootstrap based multiple hypothesis tests are then applied to identify the non-zero coefficients of the FIR filter, i.e., low order estimates of the channel response. The main advantage of using the bootstrap over other classical statistical methods is that, it can be applied with minimal assumptions, to scenarios where no information is available about

the underlying distributions [45]. Simulation results demonstrate the power of using this technique in identifying significant channel parameters for parsimonious RAKE receivers in unknown noise environments.

The chapter is organized as follows. In Section 7.2, we introduce the model and assumptions. In Section 7.3, we discuss the time domain based approach. In Section 7.4 we discuss the bootstrap based multiple hypothesis tests. In Section 7.5 we give simulation results and discussion before we conclude in Section 7.6.

7.2 Data Model

The data model is the same model in Chapter 6. Thus, the baseband discrete time received signal of a K users' DS-CDMA up-link system operating with a binary data modulation format is given by,

$$y(n) = \sum_{k=1}^{K} \sum_{l=1}^{d_k} g_k(l)\; x_k\left(n - \frac{r\tau_{kl}}{T_c}\right) + v(n),\; n = 0, .., r\mathcal{M} - 1, \quad \mathcal{M} = M + \left\lfloor \frac{\tau_{kd_k}}{T_c} \right\rfloor \tag{7.1}$$

where, $x_k(n)$ represents the training signals for $k = 1, 2, \ldots, K$ and $g_k(l),\; k = 1, 2, \ldots, K, l = 1, 2, \ldots, L$ represent the channel attenuation coefficients.

7.2.1 Objectives

By using only observations $y(n)$ and training sequences $x_k(n)$ for $k = 1, \ldots, K,\; n = 0, \ldots, r\mathcal{M} - 1$, the objective is to retrieve the transmitted symbols $b_k(i)$ from the baseband received signal in Eqn. (7.1) using parsimonious CDMA RAKE receivers. Since CDMA RAKE receivers are based on estimates of the channel response, then this approach can be translated as to identify and estimate only the significant parameters of the channel. We will assume that $x_k(n)$ and the complex noise process $v(n)$ are stationary and independent with $\mathsf{E}\{v(n)\} = 0$ and $\mathsf{Var}\{v(n)\} < \infty$. Unlike the proposed methods in Chapter 6, where the noise distribution and/or the statistical test distribution is known, the proposed bootstrap method in this chapter will estimate the statistical test distribution distributions under minimal assumption about the additive noise $v(n)$.

7.3 A Structured FIR Filter Model

Referring to Eqn. (7.1) and for a sampling ratio of $r = 1$, if we assume that the channel delays, τ_{kl} are integer multiple of the chip duration T_c, i.e., $\tau_{kl} = lT_c$ then one can model the channel as an FIR filter,

$$y(n) = \sum_{k=1}^{K} \sum_{l=1}^{L} g_k(l) x_k(n - l + 1) + v(n),\ n = 1, \ldots, M \tag{7.2}$$

Since the true number of paths, d_k is unknown, we assume the FIR model has L taps where, $L = \lfloor \frac{\tau_{max}}{T_c} \rfloor \gg d_k$ and τ_{max} is the maximum delay-spread of the channel [4]. Due the higher data rates in latest generation of mobile systems such as UMTS, the length L will be very large. This length will lead to a very long channel order with many parameters where only a few of them are non-zero or significant. With short training sequences this will also lead to poor estimates of the channel response and low system performance. Thus, under a certain level of significance we need to identify and estimate these non-zero parameters $g_k(l)$ for an arbitrary large $L < M$, given observations $y(n)$ and the training sequences $x_k(n),\ k = 1, ..., K\ ,\ n = 1, \ldots, M$.

In other words for a global level of significance, γ, we need to test simultaneously for $l = 1, \ldots, L,\ k = 1, \ldots, K$,

$$\begin{gathered} \mathcal{H}_k(l) : g_k(l) = 0 \\ \text{against} \\ \mathcal{K}_k(l) : g_k(l) \neq 0 \end{gathered} \tag{7.3}$$

where $\mathcal{H}_k(l)$ represents the null hypothesis that there is no significant path parameter $g_k(l)$ at the corresponding delay τ_{kl}, and $\mathcal{K}_k(l)$ is the alternative hypothesis that there is a significant path at that delay.

To implement this test, first for each parameter $g_k(l)$, we define a test statistic, T_{kl}. Then we find the distribution of T_{kl} under the null hypothesis $\mathcal{H}_k(l)$, i.e., the *pdf* of $T_{kl}|\mathcal{H}_k(l)$.

Since the noise distribution is not known, we use the bootstrap technique to find the above test distribution and construct the multiple hypothesis tests. These tests require an estimate of the channel parameters $g_k(l)$. This is covered in the following sections.

7.3.1 Estimation of channel parameters

In this section we present two different methods to estimate the channel response, $g_k(l),\ k = 1, \ldots, K,\ l = 1, \ldots, L$. Since we previously assumed that the channel delays, τ_{kl} are multiple integers of the chip duration, i.e., $\tau_{kl} = lT_c$ then one can use the following LSE based methods to estimate the channel response.

7.3.1.1 First method: Estimation based on signals

The first method is a straight forward LSE based on the training and the observed signals which is given by,

$$\hat{\mathbf{g}} = \left[\mathbf{X}^H \mathbf{X}\right]^{-1} \mathbf{X}^H\, \mathbf{y} \tag{7.4}$$

where,

$\hat{\mathbf{g}} = \left[\hat{\mathbf{g}}_1^T\ \hat{\mathbf{g}}_2^T \ldots\ \hat{\mathbf{g}}_K^T\right]^T$, $\hat{\mathbf{g}}_k = \left[\hat{g}_k(1)\hat{g}_k^(1) \ldots \hat{g}_k^(L)\ \right]^T$,

$\mathbf{y} = [y(1)\ y(2)\ \ldots\ y(\mathcal{M})]^T$,

$\mathbf{X} = [\mathbf{X}_1\ \mathbf{X}_2\ \ldots\ \mathbf{X}_K]$,

$$\mathbf{X}_k = \begin{bmatrix} x_k(1) & & \mathbf{0} \\ \vdots & \ddots & x_k(1) \\ x_k(M) & & \vdots \\ \mathbf{0} & \ddots & x_k(M) \end{bmatrix}_{(\mathcal{M}\times L)}$$

7.3.1.2 Second method: Estimation based on cross-correlations

After some manipulations, one can show from Eqn. (7.2) that the cross-correlation $R_{yx_k}(m)$ between, say, the training sequence of user k, $x_k(n)$, and the observed signal $y(n)$ at a lag m, is given by,

$$\begin{aligned} &R_{yx_k}(m) = \sum_{l=1}^{L} g_k(l) R_{x_k x_k}(m - l + 1) + \epsilon_k(m),\ m = -M_g, \ldots, M_g \\ &\text{where,} \\ &\epsilon_k(m) = \sum_{k'=1, k'\neq k}^{K} \left[\sum_{l=1}^{L} g_k(l) R_{x'_k x_k}(m - l)\right] + R_{v x_k}(m) \end{aligned} \tag{7.5}$$

$\epsilon_k(m)$ are estimation errors due to the noise and the MAI terms. We assume they are identically and independently distributed under some regularity condition [57].

In matrix form Eqn. (7.5) becomes,

$$\mathbf{R}_{yx_k} = \mathbf{R}_{x_kx_k}\ \mathbf{g}_k + \boldsymbol{\epsilon}_k \tag{7.6}$$

where $\mathbf{R}_{x_kx_k}$ is the correlation matrix of dimension $(2M_g+1)\times(L+1)$, $\mathbf{R}_{yx_k} = [R_{yx_k}(-M_g : M_g)]^T$ is the cross-correlation vector, M_g is the maximum lag and $\mathbf{g}_k = [g_k(1)\dots g_k(L)]^T$.

Define estimates $\hat{R}_{yx_k}(m)$ and $\hat{R}_{x_kx_k}(m)$ for $R_{yx_k}(m)$ and $R_{x_kx_k}(m)$ respectively. Using the average techniques in [55] with length equal to $\mathcal{M}$ samples for $\hat{R}_{yx_k}(m)$ and $\hat{R}_{x_kx_k}(m)$, the channel response for user k is given by the LSE,

$$\hat{\mathbf{g}}_k = \left[\mathbf{R}^H_{x_kx_k}\mathbf{R}_{x_kx_k}\right]^{-1}\mathbf{R}^H_{x_kx_k}\ \mathbf{R}_{yx_k} \tag{7.7}$$

Using either of the estimators in Eqn. (7.4) or Eqn. (7.7) one can estimate the channel response. Once these estimates are available the next step is then to construct the multiple hypothesis tests in Eqn. (7.3) to identify the significant model parameters. In the following section we use the bootstrap technique for that goal.

7.4 Bootstrap Based Multiple Hypothesis Tests

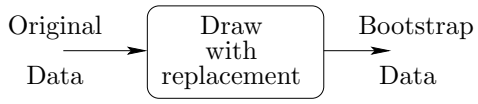
To form the hypothesis test given in Eqn. (7.3), we define the test statistics,

$$\hat{T}_{kl} = \frac{|\hat{g}_k(l)|}{\sigma_{\hat{g}_k(l)}}\ ,\ k = 1,...,K,\ l = 1,...,L. \tag{7.8}$$

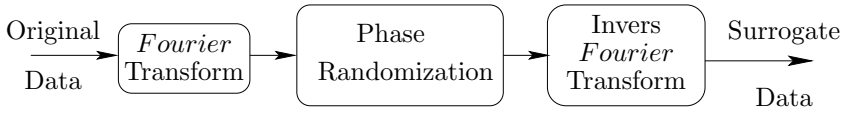
where $\sigma_{\hat{g}_k(l)}$ is the standard deviation of the parameter $\hat{g}_k(l)$. Since the distribution of $\hat{T}_{kl}$ is unknown then the bootstrap is an appropriate technique for this task. The technique is simple and attractive. It starts first with the calculation of the residual errors,

$$v_r(n) = y(n) - \sum_{k=1}^{K}\sum_{l=1}^{L}\hat{g}_k(l)x_k(n-l+1), n = 1,\dots,\mathcal{M} \tag{7.9}$$

We then draw the resampled data $v_r^\star(n)$, $n = 1, \ldots, \mathcal{M}$, using either the classical iid bootstrap or the surrogate data bootstrap as shown in Fig 7.1. For more details and see Chapter 5and [61] [62].



a) Classical Bootstrap.



b) Surrogate Bootstrap.

Figure 7.1: Bootstrap randomization techniques.

With each resampled data we calculate the bootstrap signal, or the bootstrap cross-correlations,

$$\begin{aligned} & y^*(n) = \sum_{k=1}^{K} \sum_{l=1}^{L} \hat{g}_k(l) x_k(n-l+1) + v_r^\star(n), \ n = 1, \ldots, \mathcal{M} \\ & \text{or,} \\ & R_{y^\star x_k}(m) = \sum_{l=1}^{L} \hat{g}_k(l) R_{x_k x_k}(m-l) + \epsilon_k^\star(m), \ m = -M_g, \ldots, M_g \end{aligned} \tag{7.10}$$

The bootstrap LSE $\hat{g}_k^\star(l)$ is then calculated,

$$\begin{aligned} & \hat{\mathbf{g}}^\star = \left[\mathbf{X}^H \mathbf{X}\right]^{-1} \mathbf{X}^H \ \mathbf{y}^\star \\ & \text{where,} \\ & \hat{\mathbf{g}}^\star = [\hat{\mathbf{g}}_1^\star; \ \hat{\mathbf{g}}_2^\star; \ldots; \hat{\mathbf{g}}_K^\star] \end{aligned} \tag{7.11}$$

or by cross-correlation,

$$\hat{\mathbf{g}}_k^\star = \left[\mathbf{R}_{x_k x_k}^H \mathbf{R}_{x_k x_k}\right]^{-1} \mathbf{R}_{x_k x_k}^H \ \mathbf{R}_{y^\star x_k} \tag{7.12}$$

Repeating the procedure say B times leads to an approximate distribution of the test statistics, given by the bootstrap,

$$\hat{T}_{kl}^\star = \frac{|\hat{g}_k^\star(l) - \hat{g}_k(l)|}{\sigma_{\hat{g}_k^\star(l)}}, \ l = 1, ..., L \ , \ k = 1, ..., K. \tag{7.13}$$

where the parameter variances $\boldsymbol{\sigma}_{\hat{g}_k^\star} = \left[\sigma_{\hat{g}_k^\star(1)} \ \sigma_{\hat{g}_k^\star(2)} \ldots \sigma_{\hat{g}_k^\star(L)}\right]^T$ can be found, based on how the channel parameters are estimated in Eqn. (7.11) and Eqn. (7.12).

Based on the first estimator in Eqn. (7.11), the parameter variances can be estimated by,

$$\boldsymbol{\sigma}^2_{\hat{g}^\star} = \mathrm{diag}\Big(\left[\mathbf{X}^H\mathbf{X}\right]^{-1} \hat{\sigma}^2_{y^\star} \Big) . \tag{7.14}$$

For the second estimator in Eqn. (7.12), since the statistical properties of the correlation and the cross-correlation functions $R_{x_k x_k}(m)$ and $R_{y x_k}(m)$ and the cross-correlation errors $\epsilon_k(m)$ are not the same as the signals $x(n)$ and $y(n)$ and the noise $v(n)$, the above variance expression will not hold. In this case, we can also use the bootstrap to estimate these variances. In contrast, this task can be done by resampling again from each bootstrap data $v^\star_r(n)$ another new bootstrap data $v^{\star\star}_r(n)$. Repeating this procedure say B_σ time will lead to the bootstrap estimates $\hat{g}^{\star\star}_k(l)_q,\ \hat{g}^{\star\star}_k(l)_{(q+1)}, \ldots,\ \hat{g}^{\star\star}_k(l)^{\#B_\sigma}$, where q represents the resampling index number and

$$\begin{aligned} \hat{\mathbf{g}}^{\star\star}_k &= \left[\mathbf{R}^H_{x_k x_k}\mathbf{R}_{x_k x_k}\right]^{-1} \mathbf{R}^H_{x_k x_k}\ \mathbf{R}_{y^{\star\star} x_k}, \\ \hat{\mathbf{g}}^{\star\star}_k &= [\hat{g}^{\star\star}_k(1)\ \hat{g}^{\star\star}_k(1)\ \ldots\ \hat{g}^{\star\star}_k(L)\ ]^T . \end{aligned} \tag{7.15}$$

Finally the variances can be estimated by,

$$\sigma^2_{\hat{g}^\star}(l) = \frac{1}{B_\sigma} \sum_{q=1}^{B_\sigma} \Big(\hat{g}^{\star\star}_k(l)_q - \overline{\hat{g}^{\star\star}_k(l)} \Big) \Big(\hat{g}^{\star\star}_k(l)_q - \overline{\hat{g}^{\star\star}_k(l)} \Big)^\dagger \tag{7.16}$$

where † represents the complex conjugate and the mean is defined by,

$$\overline{\hat{g}^{\star\star}_k(l)} = \frac{1}{B_\sigma} \sum_{q=1}^{B_\sigma} \hat{g}^{\star\star}_k(l)_q \tag{7.17}$$

With $\hat{T}^\star_{kl}$ and $\hat{T}_{kl}$ the probability values, $\mathcal{P}$-values clarified by Fig. 7.2 are calculated,

$$\mathcal{P}_{kl} = \frac{1}{B} \#\{ \hat{\mathbf{T}}^\star_{kl} \geq \hat{T}_{kl} \} \tag{7.18}$$

where, $\hat{\mathbf{T}}^\star_{kl} = [\hat{T}^\star_{kl}(1)\ \hat{T}^\star_{kl}(2)\ \ldots\ \hat{T}^\star_{kl}(B)]^T,\ k = 1, ..., K,\ l = 1, ..., L$ and finally to control the global level of significance, the Bonferroni multiple hypothesis test is used [52]. The full procedures are summarized in Table 7.1, Table 7.2 and shown in Fig. 7.3, Fig. 7.4 and Fig. 7.5.

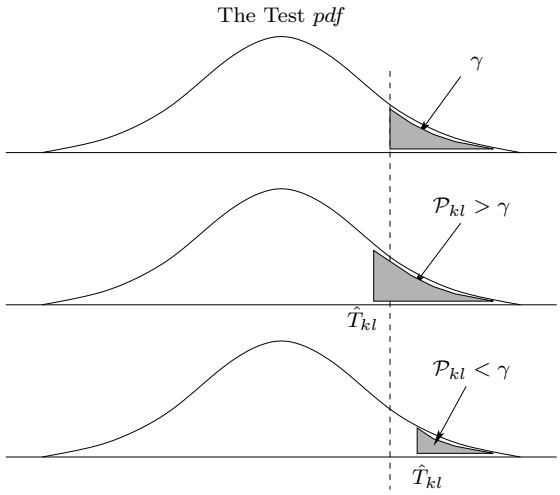


Figure 7.2: Testing the probability value for the parameter $g_k(l)$

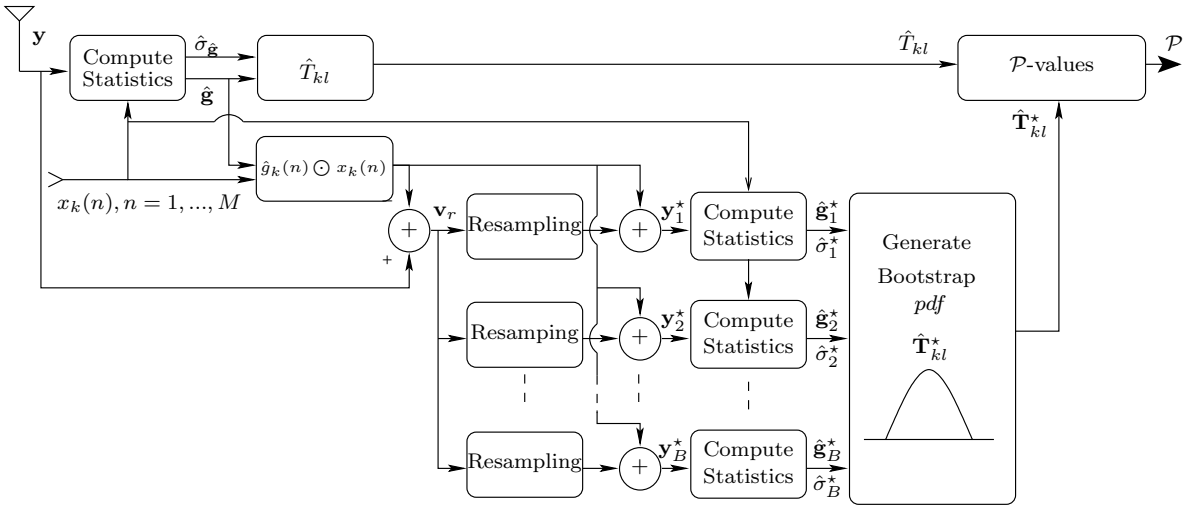


Figure 7.3: The bootstrap algorithm

Table 7.1: The bootstrap test procedure for the first method

Step 1: Given $x_k(n)$ and $y(n),\ n = 1, \cdots, \mathcal{M}$ calculate the channel LSE and its variance,

$$\hat{\mathbf{g}} = \left[\mathbf{X}^H\mathbf{X}\right]^{-1}\mathbf{X}^H\ \mathbf{y}$$

$$\boldsymbol{\sigma}^2_{\hat{\mathbf{g}}} = \mathrm{diag}\Big(\left[\mathbf{X}^H\mathbf{X}\right]^{-1}\hat{\sigma}^2_y\Big)$$

where,

$$\hat{\sigma}^2_y = \tfrac{1}{\mathcal{M}-1}\sum_{n=1}^{\mathcal{M}}\left[y(n) - \tfrac{1}{\mathcal{M}}\sum_{n=1}^{\mathcal{M}} y(n)\right]^2$$

$\mathbf{X}$ and $\mathbf{y}$ are the same as given before in sec. 7.3.1.1.

Step 2: Define the test statistic,

$$\hat{T}_{kl} = \tfrac{|\hat{g}_k(l)|}{\sigma_{\hat{g}_k(l)}},\ k = 1, \ldots, K,\ l = 1, \ldots, L.$$

Step 3: Calculate the residual errors,

$$v_r(n) = y(n) - \sum_{k=1}^{K}\sum_{l=1}^{L}\hat{g}_k(l)\ x_k(n-l+1)\ ,\ n = 1, \cdots, \mathcal{M}$$

and remove the mean value of $v_r(n)$.

Step 4: From $v_r(n)$ draw the resampled data $v^\star_r(n)$ [62], and calculate the bootstrap data,

$$y^\star(n) = \sum_{k=1}^{K}\sum_{l=1}^{L}\hat{g}_k(l)\ x_k(n-l+1) + v^\star_r(n),\ n = 1, \cdots, \mathcal{M}$$

Then find the bootstrap LSE and its variance

$$\hat{\mathbf{g}}^\star = \left[\mathbf{X}^H\mathbf{X}\right]^{-1}\mathbf{X}^H\ \mathbf{y}^\star$$

$$\boldsymbol{\sigma}^2_{\hat{\mathbf{g}}^\star} = \mathrm{diag}\Big(\left[\mathbf{X}^H\mathbf{X}\right]^{-1}\sigma^2_{y^\star}\Big)$$

Step 5: Define the bootstrap test statistics,

$$\hat{T}^\star_{kl} = \tfrac{|\hat{g}^\star_k(l) - \hat{g}_k(l)|}{\sigma_{\hat{g}^\star_k(l)}},\ k = 1, ..., K,\ l = 1, ..., L$$

Step 6: Repeat steps (4 to 5) B times to obtain $\hat{T}^\star_{kl}$ such that,

$$\hat{\mathbf{T}}^\star_{kl} = [\hat{T}^\star_{kl}(1)\ \hat{T}^\star_{kl}(2)\ \ldots\ \hat{T}^\star_{kl}(B)]^T,\ k = 1, ..., K,\ l = 1, ..., L$$

Step 7: Calculate the $\mathcal{P}$-values, $\mathcal{P}_{kl} = \frac{1}{B}\#\{\hat{\mathbf{T}}^\star_{kl} \geq \hat{T}_{kl}\}$ and for a global level of significance γ, say 5%, construct the Bonferroni tests as shown in Fig. 7.4 and Fig. 7.5 (see Chapter 5 for more details).

Step 8: Using the multiple hypothesis test with Bonferroni test results in Step 7:, identify the model order and the non-zero parameters and then re-estimate the significant channel parameters with the new low order estimate.

Table 7.2: The bootstrap test procedure for the second method

Step 1: Given $x_k(n)$ and $y(n),\ n = 1, \cdots, \mathcal{M}$ calculate the channel LSE,

$\hat{\mathbf{g}}_k = \left[\mathbf{R}^H_{x_k x_k}\mathbf{R}_{x_k x_k}\right]^{-1}\mathbf{R}^H_{x_k x_k}\ \mathbf{R}_{y x_k}$

Step 2: Calculate the residual errors,

$v_r(n) = y(n) - \sum_{k=1}^{K}\sum_{l=1}^{L}\hat{g}_k(l)\ x_k(n-l+1)\ ,\ n = 1, \cdots, \mathcal{M}$

Step 3: From $v_r(n)$ draw the resampled data $v^\star_r(n)$ and calculate the bootstrap data,

$y^\star(n) = \sum_{k=1}^{K}\sum_{l=1}^{L}\hat{g}_k(l)\ x_k(n-l+1) + v^\star_r(n),\ n = 1, \cdots, \mathcal{M}$

then find the bootstrap LSE,

$\hat{\mathbf{g}}^\star_k = \left[\mathbf{R}^T_{x_k x_k}\mathbf{R}_{x_k x_k}\right]^{-1}\mathbf{R}^H_{x_k x_k}\ \mathbf{R}_{y^\star x_k}$

Step 4: For $q = 1, 2, \ldots B_\sigma$ repeat Step 3, then find the parameter variances using the bootstrap procedure in Eqn. (7.16) and Eqn. (7.17).

$\sigma^2_{\hat{g}}(l) = \frac{1}{B_\sigma}\sum_{q=1}^{B_\sigma}\left(\hat{g}^\star_k(l)_q - \overline{\hat{g}^\star_k(l)}\right)\left(\hat{g}^\star_k(l)_q - \overline{\hat{g}^\star_k(l)}\right)^\dagger,$

$\overline{\hat{g}^\star_k(l)} = \frac{1}{B_\sigma}\sum_{q=1}^{B_\sigma}\hat{g}^\star_k(l)_q$

Step 5: Define the test statistic,

$\hat{T}_{kl} = \frac{|\hat{g}_k(l)|}{\sigma_{\hat{g}_k(l)}},\ k = 1, \ldots, K,\ l = 1, \ldots, L.$

Step 6: Resample as Step 3: and find the bootstrap LSE,

$\hat{\mathbf{g}}^\star_k = \left[\mathbf{R}^H_{x_k x_k}\mathbf{R}_{x_k x_k}\right]^{-1}\mathbf{R}^H_{x_k x_k}\ \mathbf{R}_{y^\star x_k}$

Step 6-1: For each resample in Step 6, resample a new B_σ times from the bootstrap residuals $v^\star_r(n)$ a new bootstrap data $v^{\star\star}_r(n)$ then find the bootstrap parameter variances using the bootstrap procedure in Eqn. (7.16) and 7.17.

Step 7: Define the bootstrap test statistics,

$\hat{T}^\star_{kl} = \frac{|\hat{g}^\star_k(l) - \hat{g}_k(l)|}{\sigma_{\hat{g}^\star_k(l)}},\ k = 1, ..., K,\ l = 1, ..., L$

Step 8: Repeat steps (6 to 7) B times to obtain $\hat{T}^\star_{kl}$ such that,

$\hat{\mathbf{T}}^\star_{kl} = [\hat{T}^\star_{kl}(1)\ \hat{T}^\star_{kl}(2)\ \ldots\ \hat{T}^\star_{kl}(B)]^T,\ k = 1, ..., K,\ l = 1, ..., L$

Step 9: Calculate the $\mathcal{P}$-values, $\mathcal{P}_{kl} = \frac{1}{B}\#\{\hat{\mathbf{T}}^\star_{kl} \geq \hat{T}_{kl}\}$ and for a global level of significance γ, say 5%, construct the Bonferroni tests as shown in Fig. 7.4 and Fig. 7.5.

Step 10: Using results of Step 9, identify and estimate the significant channel parameters.

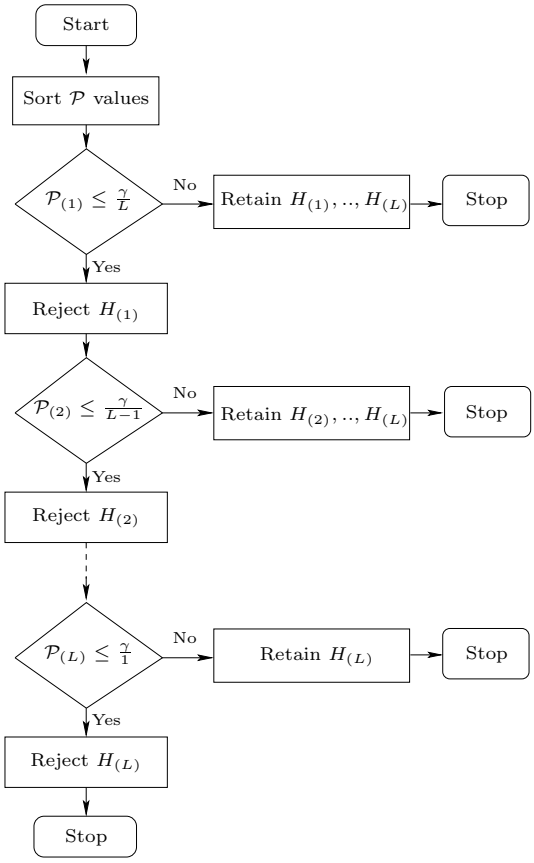


Figure 7.4: Sequentially Rejective Bonferroni tests (SRB)

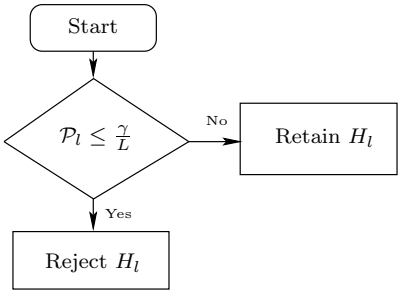


Figure 7.5: Classical Bonferroni test

7.5 Simulation Results

7.5.1 Estimation and identification performances

To check the accuracy of the statistical bootstrap techniques, the following simulations are conducted in a DS-CDMA system with $K = 4$ users. The training signals $x_k(t)$ are assumed to be sequences of random symbols $b_k(j) \in \{+1, -1\}$ modulated by a maximum length binary sequence (MLBS) as spreading codes waveform, $s_k(t) \in \{+1, -1\}$. These signals are transmitted over a frequency-selective slowly fading channel of length $d_o = 6$ and with parameters shown in Table 7.3, and additive circular Gaussian noise.

Table 7.3: Multipath parameters for Example 1

g_{l1}	g_{l2}	g_{l3}	g_{l4}
-1.163-j0.917	1.133-j0.956	0.637+0.385	0.303+1.816
-0.071-j0.701	0	0.374-j0.235	-0.656-j1.640
-0.052+j0.277	0	0	0.015+j0.360
0	-0.319+j0.204	1.775-j0.769	-0.925-j0.091
-0.227+j0.490	-0.071-j0.703	0	0.205+j0.630
1.774-j0.769	-0.925-j0.092	0.478-j0.025	-0.656-j1.640

7.5.1.1 Example 1

To check how powerful is the bootstrap technique in estimating the distribution of a certain estimated parameter, the bootstrap randomization methods mentioned in Section 7.4 are used to estimate the mean $\overline{\hat{g}_k(l)}$ and the variance $\mathsf{Var}\{\hat{g}_k(l)\} = \sigma^2_{\hat{g}_k(l)}$ of two parameters, $g_1(1) = -1.163 - j0.917$ and $g_3(2) = 0$. The mean and variance based on the bootstrap estimates can be defined by,

$$\begin{aligned} &\overline{\hat{g}_k(l)} = \mathsf{E}\{\hat{g}^\star_k(l)\} \\ &\text{and} \\ &\sigma^2_{\hat{g}_k(l)} = \mathsf{E}\left\{ \left(\hat{g}^\star_k(l) - \overline{\hat{g}_k(l)}\right) \left(\hat{g}^\star_k(l) - \overline{\hat{g}_k(l)}\right)^\dagger \right\} \end{aligned} \tag{7.19}$$

Repeating the bootstrap randomization B times, this can be approximated by,

$$\begin{aligned} &\overline{\hat{g}_k(l)} \approx \tfrac{1}{B}\sum_{q=1}^{B} \hat{g}_k^{\star}(l)_q \\ &\text{and} \\ &\sigma^2_{\hat{g}_k(l)} \approx \tfrac{1}{B}\sum_{q=1}^{B} \left(\hat{g}_k^{\star}(l)_q - \overline{\hat{g}_k(l)}\right)\left(\hat{g}_k^{\star}(l)_q - \overline{\hat{g}_k(l)}\right)^{\dagger} \end{aligned} \tag{7.20}$$

where, $\hat{g}_k^{\star}(l)_q$ is the bootstrap estimates at a randomization index number q. Fig. 7.6 and Fig. 7.7 present histograms of the estimated $\hat{g}_1(1)$, Fig. 7.8 and Fig. 7.9 present histograms of the estimated $\hat{g}_3(2)$ using the bootstrap when $M = 32$, $N = 64$ symbols, $B = 100$ at SNR=$4dB$ for the following two different noise cases:

- Uncorrelated zero mean circular complex Gaussian noise with a unity variance.
- Correlated zero mean circular complex Gaussian noise generated by passing a zero mean circular complex Gaussian noise with a unit variance through an AR process of order 8 with parameters, $a_1 = -3.4846,\ a_2 = 7.3032,\ a_3 = -10.4278,\ a_4 = 10.9317,\ a_5 = -8.4956,\ a_6 = 4.7928,\ a_7 = -1.8176,\ a_8 = 0.3749$.

7.5.1.2 Discussion of Example 1

In the results shown in Fig. 7.6, 7.7, 7.8 and 7.9, one can see that the bootstrap techniques can give sufficient statistics such as the mean and the variance of the estimated parameters. These results are obvious when the parameter is significant as shown in Fig. 7.6 for $\hat{g}_1(1)$ or when the parameter is not significant or zero as shown in Fig 7.8 for $\hat{g}_3(2)$. The results in Fig. 7.7 and Fig. 7.9 also show that the technique is not affected by the situation when the additive noise is correlated. In conclusion, one can see that the bootstrap is a good technique for estimating the mean and the variance which are the two important factors for the required statistical test in Eqn (7.13).

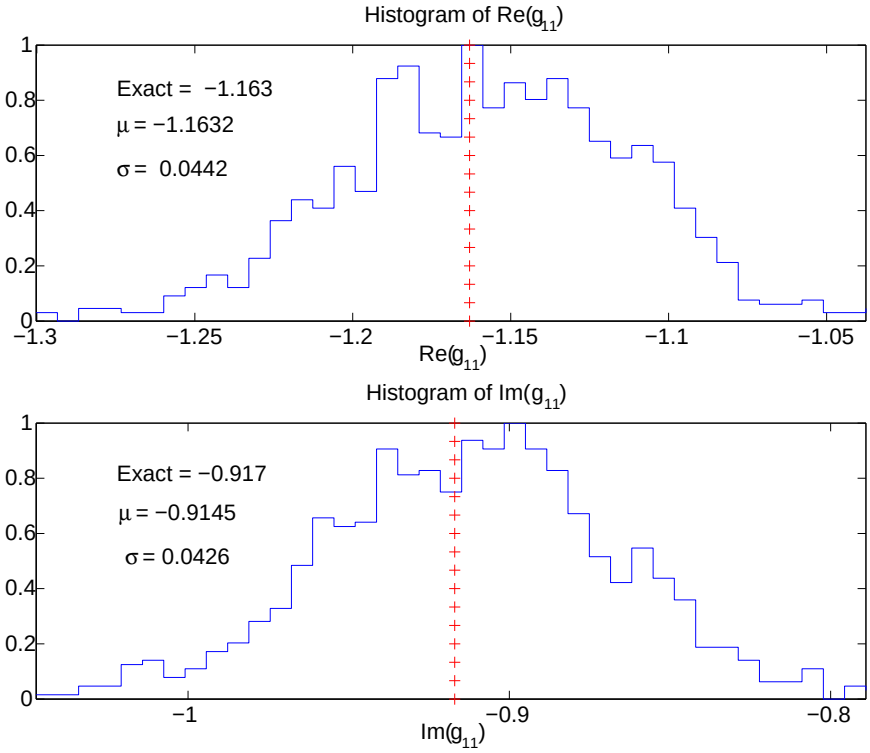


Figure 7.6: Histogram of the $\mathrm{Re}(\hat{g}_1(1))$ and $\mathrm{Im}(\hat{g}_1(1))$, uncorrelated noise

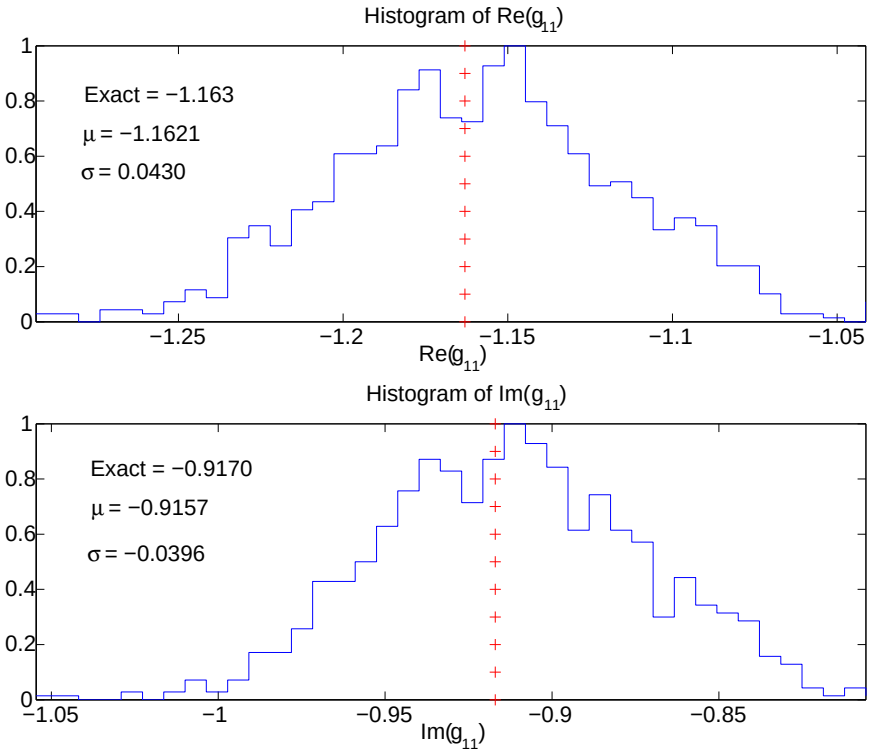


Figure 7.7: Histogram of the $\mathrm{Re}(\hat{g}_1(1))$ and $\mathrm{Im}(\hat{g}_1(1))$, correlated noise

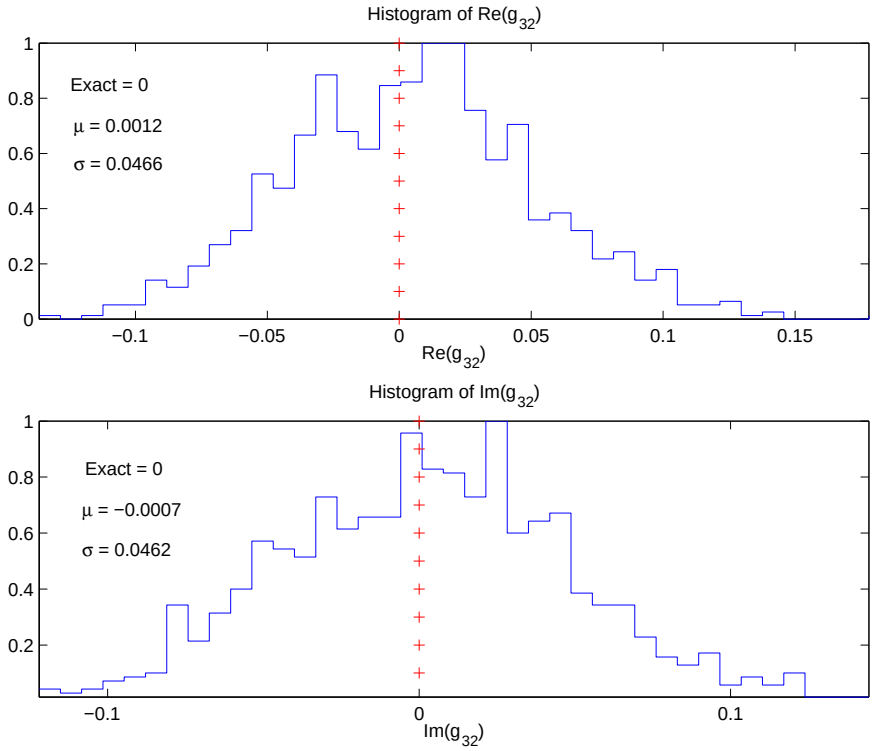


Figure 7.8: Histogram of the $\mathrm{Re}(\hat{g}_3(2))$ and $\mathrm{Im}(\hat{g}_3(2))$, uncorrelated noise

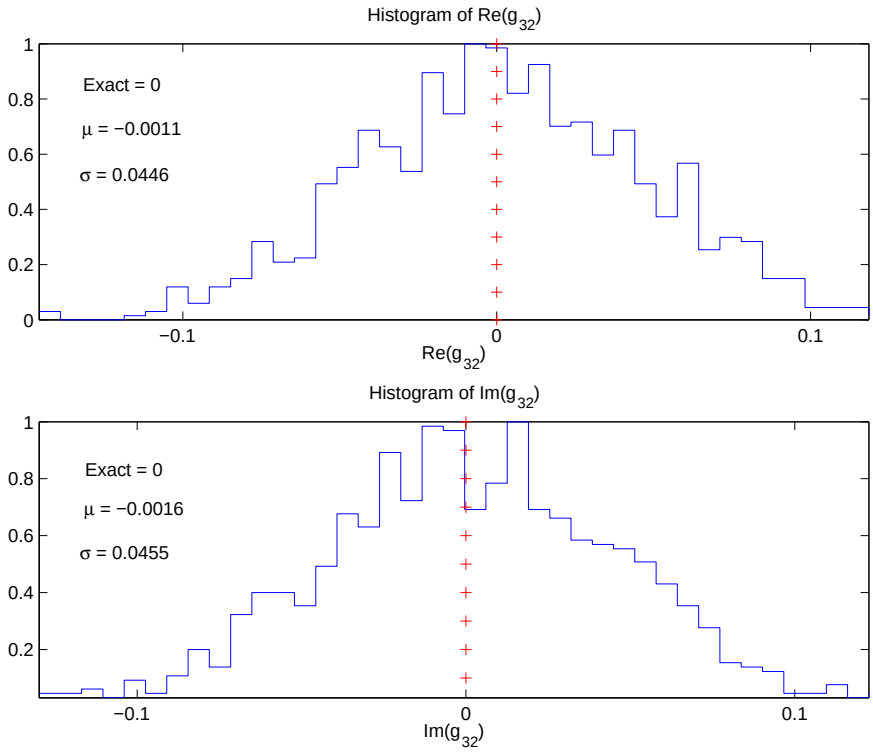


Figure 7.9: Histogram of the $\mathrm{Re}(\hat{g}_3(2))$ and $\mathrm{Im}(\hat{g}_3(2))$, correlated noise

7.5.1.3 Example 2

Using the input data of Example #1, Table 7.4 compares the percentages of correct identification of the exact order and the channel non-zero parameters using the classical iid bootstrap and surrogate bootstrap methods. Techniques used for comparison are:

- Classical Bonferroni test based on Classical iid Bootstrap (CBCB)
- Classical Bonferroni test based on Surrogate Bootstrap (CBSB)
- Sequentially Rejective Bonferroni test based on Classical iid Bootstrap (SRBCB)
- Sequentially Rejective Bonferroni test based on Surrogate Bootstrap (SRBSB).

The classical Bonferroni test can be used when the statistical tests are independent. Otherwise the SRB test is used. The two procedures are shown in Fig. 7.3 and Fig. 7.4.

7.5.1.4 Discussion of Example 2

The results in Table 7.4 show that the performance of the classical and the sequential Rejective Bonferroni tests is almost the same. This is valid for both randomization methods, the classical iid and the surrogate data bootstrap. This situation can happen when the multiple hypothesis tests are independent. Results also show that the iid bootstrap resampling method performs better than the surrogate data bootstrap method. At larger SNR, clearly, the percentage of correct identification is increased.

7.5.1.5 Example 3

This example demonstrates the use of the bootstrap SRB test in identifying the significant channel parameters of a frequency selective slowly fading channel introduced in a $K = 1, 2, 3, 4$ users system with maximum channel length of $d_o = 10$, $d_k = 3$ significant paths, $L = 16$ required physical taps, $M = 32$

Table 7.4: Percentage of correct identification in (%)

Case	SNR in dB	CBCB	CBSB	SRBCB	SRBSB
Uncorrelated	0	82	84	81	81
	2	95	94	95	86
	4	98	91	96	87
	6	99	93	95	85
	8	99	94	95	89
	10	98	91	96	87
Correlated	0	82	75	78	72
	2	92	83	86	77
	4	89	86	84	76
	6	93	83	85	74
	8	92	84	85	75
	10	94	85	84	77

chips and $N = 32,\ 64,\ 96$ and 128 symbols. The channel components are given in Table 7.5. Let $p_k(l)$ be the probability of identifying each individual path parameter $g_k(l)$ correctly, then one can show that the probability of identifying the full channel parameters of user k correctly p_f can be given by,

$$p_f = \prod_{l=1}^{L} p_k(l) \tag{7.21}$$

The results are shown in Fig. 7.10 and Fig. (7.11) for $\gamma = 0.05$.

7.5.1.6 Discussion of Example 3

From the results, one can see that the bootstrap SRB test is a powerful statistical test in identifying the significant and the nonsignificant channel parameters with high probability. The identification test works well at low SNRs and short N symbols compared with many other techniques that are based on larger number of symbols and/or with pre-defined statistical tests. In Fig. 7.10 and 7.11 one can

see that, for a better performance, say, when $N = 32$ symbols, it is better to run the system with an SNR > 4 dB, and for $N \geq 64$ symbols the system can run over an SNR > 2 dB . In practice, many wireless communication systems operate with SNRs greater than 8 dB but if the system is running at SNRs less than 4 dB, a longer training sequence is then required or a higher level of significance can be assumed.

Table 7.5: Multipath parameters for Example 3.

Delay	Multipath parameters $g_k(l)$ for			
τ_{1l}	User 1	User 2	User 3	User 4
0	0	0	0	0
T_c	-0.0207 - j0.4087	0	0	-0.2290 + j0.1217
$2T_c$	0	0.0347 + j0.4313	0.1423 -j0.0278	0.1732 - j0.2147
$3T_c$	0	0	0	-0.4983 - j0.1737
$4T_c$	0	0.0347 + j0.4313	-0.4201 + j0.4713	0
$5T_c$	0.5350 + j0.0278	0	0.6119 + j0.4301	0
$6T_c$	0.6730 + j0.3042	-0.5138 - j0.1791	0	-0.2825 + j0.4796
$7T_c$	0	-0.0842 - j0.3971	0	-0.3364 + j0.3774
$8T_c$	0	0	0	-0.0598 - j0.0957
$9T_c$	0	-0.0842 - j0.3971	0.1423 - j0.0278	0
$10T_c$	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$16T_c$	0	0	0	0

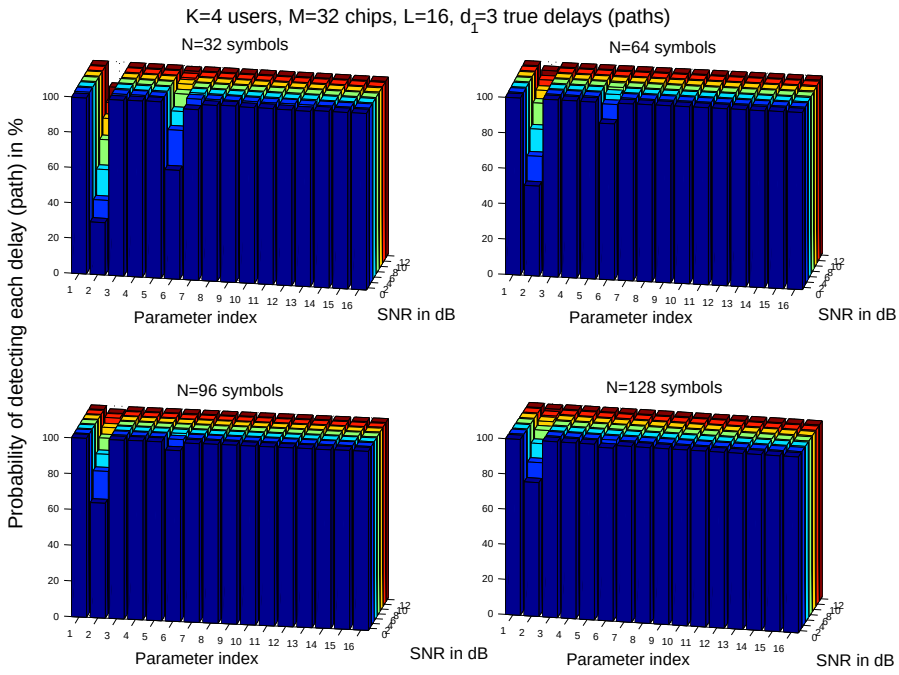


Figure 7.10: The probability of correctly detecting user's one non-zero parameters (paths), SRB test

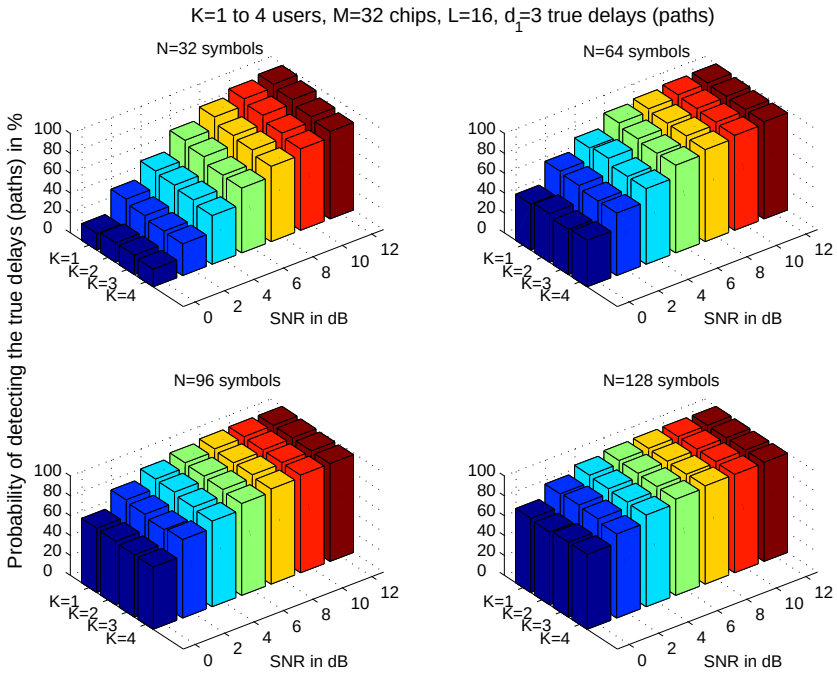


Figure 7.11: The probability of correctly detecting user's one all parameters (paths), SRB test

7.5.2 System performance

To assess the performance of the proposed RAKE receiver which is based on the identified parameters, the following examples compares the Bit Error Rate (BER) for systems with:

- Perfect knowledge of the channel
- Channel estimated with known length d_o
- Channel estimated with length L
- Channel identified and estimated using the proposed bootstrap approach.

The comparison algorithm is summarized in Table 7.6. In the following simulations the channel is assumed to be a frequency selective slowly fading channel with parameters chosen randomly from the Rayleigh distribution.

7.5.2.1 Example 4:

In this simulation we apply the full bootstrap procedure to calculate the performance of two different systems over a frequency selective slowly fading channel. The first system has $K = 4$ users, $M = 32$ chips, length $d_o = 8$ paths, $L = 16$ taps and $N = 64$ symbols, while the second has $K = 8$ users, $M = 32$ chips, length $d_o = 8$ paths, $L = 16$ taps and $N = 64$ symbols. In both cases the channel has $d_k = 4$ significant paths. The results are shown in Fig. 7.12 and Fig. 7.13 respectively for the first user.

7.5.2.2 Discussion of Example 4

In Fig. 7.12 and Fig. 7.13, simulation BERs show that the performance of the proposed frequency domain based RAKE receiver is much improved over other classical methods such as direct implementation based RAKE and/or when the channel order is known. For $K = 4$ or 8 users, the new approach has the smallest BER except for when the channel is known exactly. Using a few training symbols, $N = 64$, and increasing the number of users to 8, as in Fig. 7.13, the MAI reduced

Table 7.6: System performance

For $k = 1, \ldots, K$, define the channel parameters:

1- $\mathbf{g}_P$ for perfect channel.

2- $\hat{\mathbf{g}}_{d_o}$ estimated with the true channel length d_o.

3- $\hat{\mathbf{g}}_L$ estimated with direct implementation of length L.

4- $\hat{\mathbf{g}}_B$ estimated using the proposed bootstrap approach.

- Let $\mathbf{b} = [b_1\ b_2\ \ldots\ b_K]^T_{(K\times 1)}$ be the transmitted symbols
- Calculate
  - $\mathbf{S} = [\mathbf{s}_1\ \mathbf{s}_2\ \ldots\ \mathbf{s}_K]_{(M\times K)}$, the normalized signature matrix.
  - $\mathbf{R} = \mathbf{S}^T\mathbf{S}$, the signature covariance matrix
- Consider the BER for one of the channel cases say $\hat{\mathbf{g}}_L$,
  - $\hat{\mathbf{g}}_L = [\hat{\mathbf{g}}_1\ \hat{\mathbf{g}}_2\ \ldots\ \hat{\mathbf{g}}_K]_{(L\times K)}$
  - $\hat{\mathbf{g}}_k = [\hat{g}_k(1)\ \hat{g}_k(2)\ \ldots\ \hat{g}_k(L)]^T$
  - Find $\mathbf{D} = \mathbf{R}^{-1}\mathbf{G}^H$

  where

  $\mathbf{G} = [\mathbf{G}_1\ \mathbf{G}_2\ \ldots\ \mathbf{G}_K]_{(\mathcal{M}\times K)}$,

  $$\mathbf{G}_k = \begin{bmatrix} s_k(1) & & \mathbf{0} \\ \vdots & \ddots & s_k(1) \\ s_k(M) & & \vdots \\ \mathbf{0} & \ddots & s_k(M) \end{bmatrix}_{(\mathcal{M}\times L)} \times \begin{bmatrix} \hat{g}_k(1) \\ \hat{g}_k(2) \\ \vdots \\ \hat{g}_k(L) \end{bmatrix}_{(L\times 1)}$$

  Estimate the transmitted symbols,

  - $\hat{\mathbf{b}} = \mathrm{sgn}\{\Re(\mathbf{D}\ \mathbf{y})\}$

  where,

  - $\mathbf{y} = [y(1)\ y(2)\ \ldots y(\mathcal{M}-1)]^T_{(\mathcal{M}\times 1)}$
- Choose another estimated channel and compare the BER for each case.

the quality of the estimated parameters which reduced the performance of all methods.

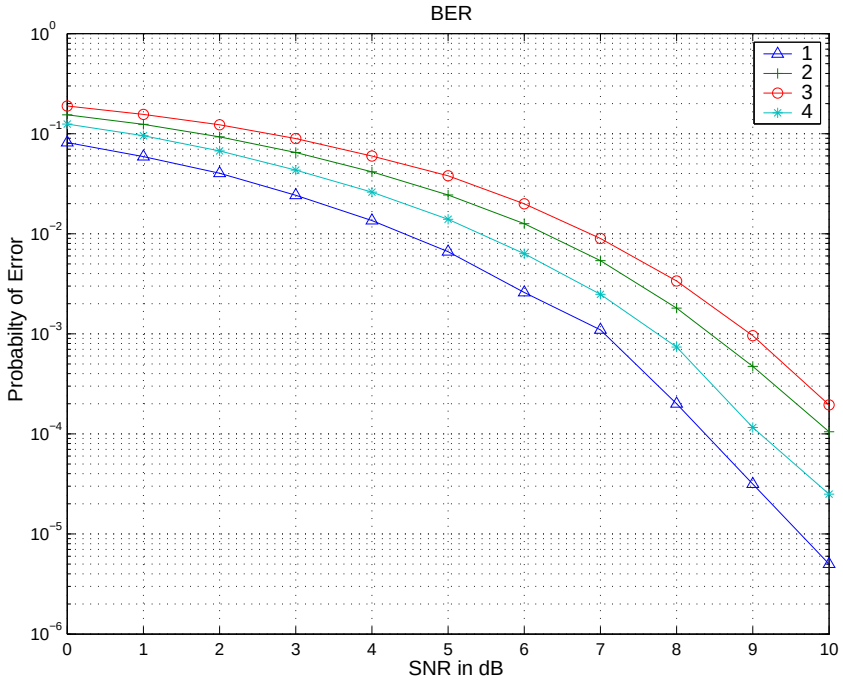


Figure 7.12: BER in the case of 4 users

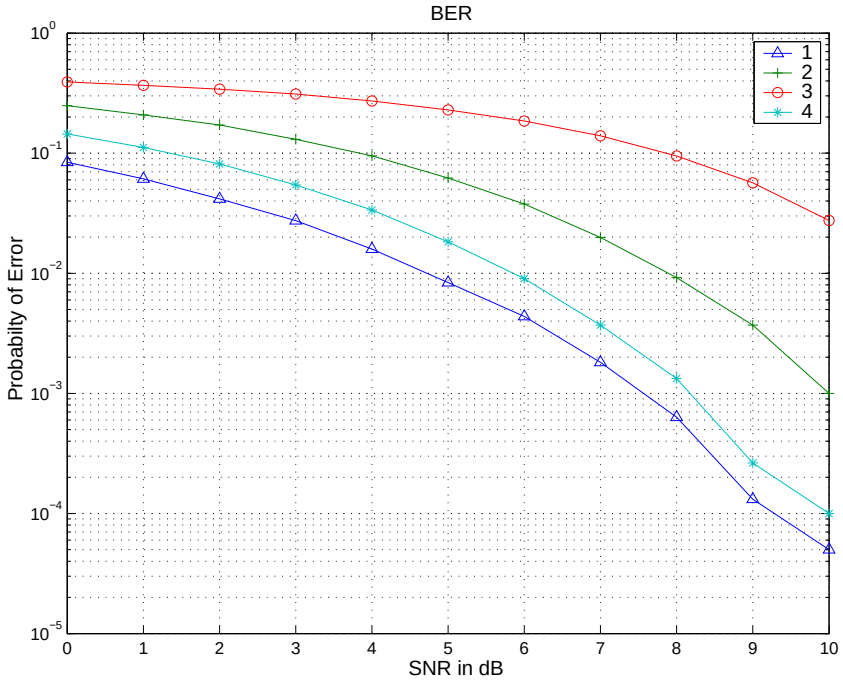


Figure 7.13: BER in the case of 8 users

7.6 Conclusion

A bootstrap time-domain based approach is proposed for the identification and estimation of the significant multipath parameters in multiuser DS-CDMA systems. The approach is used to enhance the RAKE receiver performance over other classical methods. The proposed approach used the bootstrap resampling techniques to estimate the unknown distribution of the test statistic. The bootstrap is found to be more flexible than classical statistical methods as it is able to compute a statistic and estimate its sampling distribution without any assumption on the model or knowledge of the noise distribution. Two different bootstrap resampling techniques are used: the classic bootstrap and the surrogate data bootstrap. The approach is also simple since it assumes the channel as an FIR filter with taps located at multiple integers of the chip duration. Results show that with the proposed approach the performance of the parsimonious RAKE receiver can increased substantially with a small processing gain. Increasing the processing gain or at higher SNR, the accuracy of the estimates obviously increases.

Chapter 8

Blind Adaptive Channel Estimation Approaches

.

8.1 Introduction

Channel propagation effects dominate as one of the major factors that limit system performance. These effects can result in many different type of fading such frequency selective and frequency non-selective fading [4]. Compensation of channel fading due to multipath propagation as discussed in the previous chapters, is possible through the use of decorrelator receivers such as the RAKE receiver. These receivers require estimation of the channel response.

There has been and still intensive research on channel estimation. Some are training based methods, while others are blind or semi-blind methods [63]-[84]. However, more attention is being paid to blind and semi-bind channel estimation techniques since they reduce the complexity of the system, increase its capacity and minimize the amount of data checking between the transmitters and the receivers. The presence of multipath delays, unfortunately, destroys the assumed orthogonality between the users' spreading codes. As a result, the accuracy of training based estimators is severely limited by the cross interference between data and pilot symbols.

Our goals in this chapter are first to derive low computational complexity blind adaptive channel estimation algorithms. We then use these algorithms to implement low complexity receivers such as RAKE receivers for multiuser CDMA systems. This will be achieved by employing constrained optimization techniques based on minimum variance (MV) receivers [63] [64]. We recursively minimize the output variance of the received signal subject to some constraints which are also jointly updated. Four different algorithms are proposed in this chapter from different viewpoints, then the results are compared with some previously proposed algorithms [63] [64] [65] [66] [67]. Finally the new algorithms are used to track frequency selective varying fading channels. The proposed algorithms are applied to the design of adaptive blind estimation based RAKE receivers.

The chapter is organized as follows. In Section 8.2, we introduce the model and assumptions. In Section 8.3, we discuss the MV receivers. In Section 8.4 we review some of the current existing blind channel estimation algorithms. In Section 8.5 we present the new algorithms. In Section 8.7 we present simulation results and a discussion and finally we we conclude in Section 8.8.

8.2 Data Model

Consider the up-link received signal of a K-user DS-CDMA system shown in Fig. 8.1.

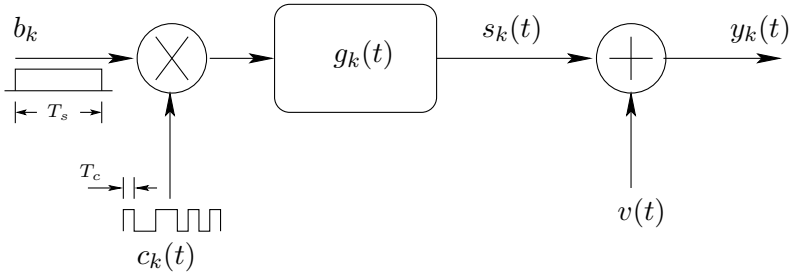


Figure 8.1: Data model

Herein, and after sampling,

$$y(n) = \sum_{k=1}^{K} y_k(n) + v(n) \tag{8.1}$$

where

$$y_k(n) = \sum_{m=-\infty}^{\infty} b_k(m) s_k(n - mT_s) \tag{8.2}$$

is the k'th user received complex signal, $b_k(m) \in \{1, -1\}$ is the transmitted symbol, T_s is the symbol duration, and $s_k(n)$ is given by

$$s_k(n) = \sum_{l=-\infty}^{\infty} g_k(l) c_k(n - lT_c) \tag{8.3}$$

where $g_k(n)$ is the channel response, $c_k(n)$ is the k'th user unit energy spreading code of length N and $T_c = \frac{T_s}{N}$ is the chip duration.

Let user k be the user of interest. Then,

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{s}_k^{(n)} b_k(n) + \underbrace{\mathbf{s}_k^{(n-1)} b_k(n-1) + \mathbf{s}_k^{(n+1)} b_k(n+1)}_{\text{ISI}} \\ &+ \underbrace{\sum_{k'=1, k' \neq k}^{K} \left\{ \mathbf{s}_{k'}^{(n)} b_{k'}(n) + \mathbf{s}_{k'}^{(n-1)} b_{k'}(n-1) + \mathbf{s}_{k'}^{(n+1)} b_{k'}(n+1) \right\}}_{\text{MAI}} + \mathbf{v}(n). \end{aligned} \tag{8.4}$$

i.e.,

$$\mathbf{y}(n) = \mathbf{s}_k^{(n)} b_k(n) + \text{ISI} + \text{MAI} + \mathbf{v}(n), \tag{8.5}$$

where for q different paths,

$\mathbf{y}(n) = [y(1)\ y(2)\ \cdots\ y(N+q-1)]^T$

$\mathbf{s}_k^{(n)} = \mathbf{C}_k \mathbf{g}_k$

$$\mathbf{C}_k = \begin{bmatrix} c_k(1) & & \mathbf{0} \\ \vdots & \ddots & c_k(1) \\ c_k(N) & & \vdots \\ \mathbf{0} & \ddots & c_k(N) \end{bmatrix}_{(N+q-1)\times q} \qquad \mathbf{g}_k = \begin{bmatrix} g_k(1) \\ g_k(2) \\ \vdots \\ g_k(q) \end{bmatrix}_{(q\times 1)}$$

$\mathbf{s}_k^{(n-1)} = [s_k(N)\ \cdots\ s_k(N+q-1)\ 0\ \cdots\ 0]^T$

$\mathbf{s}_k^{(n+1)} = [0\ \cdots\ 0\ s_k(1)\ \cdots\ s_k(q)]^T$

$\mathbf{v}(n)$ is the complex noise process with $\mathsf{E}\{v(n)\} = 0$ and $\mathsf{Var}\{v(n)\} < \infty$ is assumed independent with $s_k(n)$.

8.3 Minimum Variance Receivers

The idea of estimating the transmitted symbols $b_k(n)$ is to find a complex vector $\mathbf{f}$ of length $(N+q-1)$ such that

$$\hat{b}_k(n) = \mathbf{f}^H \mathbf{y}(n) \tag{8.6}$$

For MV receivers [63] [64], it has been shown that the vector $\mathbf{f}$ can be found by minimizing the variance of the zero mean output symbols $\hat{b}_k(n)$, or in other words minimizing the cost function ζ given by,

$$\begin{aligned} &\zeta = \mathsf{E}\Big\{ ||\hat{b}_k(n)||^2 \Big\} = \mathbf{f}^H \mathbf{R}_y \mathbf{f} \\ &\text{where} \\ &\mathbf{R}_y = \mathsf{E}\big\{ \mathbf{y}(n)\mathbf{y}^H(n) \big\} \end{aligned} \tag{8.7}$$

The minimization of ζ is subject to the constraint that the response of the user of interest, k, has to be constant, i.e.,

$$\mathbf{f}^H \mathbf{C}_k \mathbf{g}_k = 1 \tag{8.8}$$

Since the solution of Eqn. (8.8) includes a scaling factor and a phase ambiguity in $\mathbf{g}$, then assuming that $\mathbf{g}^H\mathbf{g} = 1$, Eqn. (8.8) becomes

$$\mathbf{C}_k^H \mathbf{f} = \mathbf{g}_k \tag{8.9}$$

8.3.1 Optimum solution

Using Lagrange multipliers, it has been shown in [63] that for a given unknown channel response $\mathbf{g} = \mathbf{g}_k$ with toeplitz spreading matrix $\mathbf{C} = \mathbf{C}_k$, the optimum solution to the vector $\mathbf{f}$ is

$$\mathbf{f}_{opt} = \mathbf{R}_y^{-1}\mathbf{C}\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g}. \tag{8.10}$$

This optimum solution leads to the minimum output variance,

$$\zeta_{opt} = \mathbf{f}_{opt}^H \mathbf{R}_y \mathbf{f}_{opt} = \mathbf{g}^H\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g} \tag{8.11}$$

8.4 Existing Algorithms

8.4.1 Adaptive LMS algorithms

In [64] [65] [68], different adaptive algorithms are proposed for the estimation of $\mathbf{f}$ and $\mathbf{g}$. The first three LMS algorithms are based on the following similar cost functions,

$$\zeta_1 = \mathbf{f}^H \mathbf{R}_y \mathbf{f} + \boldsymbol{\lambda}^H \left(\mathbf{C}^H \mathbf{f} - \mathbf{g}\right) + \boldsymbol{\lambda} \left(\mathbf{f}^H \mathbf{C} - \mathbf{g}^H\right) + \rho \left(\mathbf{g}^H \mathbf{g} - 1\right), \tag{8.12}$$

$$\zeta_2 = \mathbf{f}^H \mathbf{R}_y \mathbf{f} + \boldsymbol{\lambda}^H \left(\mathbf{C}^H \mathbf{f} - \mathbf{g}\right) + \boldsymbol{\lambda} \left(\mathbf{f}^H \mathbf{C} - \mathbf{g}^H\right), \tag{8.13}$$

$$\begin{aligned} \zeta_3 = {} & \mathbf{g}^H \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{C}^H \mathbf{R}_y \mathbf{C} \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{g} + \mathbf{u}^H \mathbf{C}_n^H \mathbf{R}_y \mathbf{C}_n \mathbf{u} \\ & - \mathbf{g}^H \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{C}^H \mathbf{R}_y \mathbf{C}_n \mathbf{u} - \mathbf{u}^H \mathbf{C}_n^H \mathbf{R}_y \mathbf{C} \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{g}. \end{aligned} \tag{8.14}$$

The main idea of estimating $\mathbf{f}$ and $\mathbf{g}$ using the above cost functions was to minimize any of them w.r.t $\mathbf{f}$ and maximize it w.r.t $\mathbf{g}$.

Firstly one initializes $\mathbf{f}$ and $\mathbf{g}$ with certain values, then when a new symbol arrives at the instant $(n+1)$, their values can be updated according to the following LMS algorithm,

$$\mathbf{f}_{n+1} = \mathbf{f}_n - \mu_{\mathbf{f}} \nabla_{\mathbf{f}^*} \zeta \tag{8.15}$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_{\mathbf{g}} \nabla_{\mathbf{g}^*} \zeta \tag{8.16}$$

and for the third cost function ζ_3

$$\mathbf{f}_{n+1} = \mathbf{C} \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{g}_{n+1} - \mathbf{C}_n \mathbf{u}_{n+1}$$

where

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \mu_{\mathbf{u}} \nabla_{\mathbf{u}^*} \zeta, \tag{8.17}$$

$\nabla_{\mathbf{x}^*} \zeta$ is the partial gradient vector of ζ w.r.t the vector $\mathbf{x}^*$, or $\nabla_{\mathbf{x}^*} \zeta = \frac{\partial \zeta}{\partial \mathbf{x}^*}$.

8.4.2 Adaptive RLS algorithm

In [64] [69] an RLS algorithm was also used. In this algorithm the channel parameters $\mathbf{g}$ is found as the eigenvector which corresponds to the minimum eigenvalue of the quadratic function $\mathbf{C}^H \mathbf{R}_y^{-1} \mathbf{C}$.

The main problem with this method was the calculation of $\mathbf{R}_y^{-1}$ where, a Kalman

RLS recursive algorithm was used. After initializing $\mathbf{R}_y^{-1}(n-1)$, the algorithm can be summarized as follows,

$$\mathbf{k}(n) = \frac{\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)}{v + \mathbf{y}^H(n)\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)} \tag{8.18}$$

$$\hat{\mathbf{R}}_y^{-1}(n) = \frac{1}{v}\hat{\mathbf{R}}_y^{-1}(n-1)\left[\mathbf{I} - \mathbf{k}(n)\mathbf{y}^H(n)\right] \tag{8.19}$$

$$\mathbf{g}_{opt} = \underset{||\mathbf{g}||=1}{\operatorname{argmin}}\ \mathbf{g}^H\mathbf{C}^H\mathbf{R}_y^{-1}(n)\mathbf{C}\mathbf{g} \tag{8.20}$$

It has been suggested that if q, the length of $\mathbf{g}$ is small enough then the SVD method can be applied to estimate $\mathbf{g}_{opt}$.

The optimum response $\mathbf{g}_{opt}$ is considered as the eigenvector which corresponds to the minimum eigenvalue of the quadratic matrix $\left[\mathbf{C}^H\mathbf{R}_y^{-1}(n)\mathbf{C}\right]_{(q\times q)}$

8.4.3 A subspace based algorithm

In [65] [66] [67] [68], the optimum channel parameters $\mathbf{g}_{opt}$ are determined by

$$\mathbf{g}_{opt} = \underset{||\mathbf{g}||=1}{\operatorname{argmin}}\ \mathbf{g}^H\mathbf{C}^H\mathbf{R}_y^{-m}\mathbf{C}\mathbf{g} \tag{8.21}$$

The difference between this algorithm and the previous RLS algorithm is that, $\mathbf{R}_y^{-m}$ is found using the subspaced Eigen Value Decomposition (EVD) technique. First the covariance matrix $\mathbf{R}_y$ is decomposed by EVD as

$$\mathbf{R}_y = [\mathbf{U}_s\ \mathbf{U}_n]\begin{bmatrix} \mathbf{\Lambda}_s + \sigma_v^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_v^2\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \tag{8.22}$$

where $\mathbf{\Lambda}_s = \operatorname{diag}\{\lambda_1^2, \ldots, \lambda_q^2\}$, $\mathbf{U}_s$ and $\mathbf{U}_n$ represent the signal and noise subspaces respectively. $\mathbf{R}_y^{-m}$ is determined using the noise subspace as

$$\sigma_v^{2m}\mathbf{R}_y^{-m} = \mathbf{U}_n\mathbf{U}_n^H + \mathbf{U}_s\ \operatorname{diag}\left\{\left(\frac{\sigma_v^2}{\lambda_i + \sigma_v^2}\right)^m\right\}\mathbf{U}_s^H \tag{8.23}$$

Since $\left(\frac{\sigma_v^2}{\lambda_i+\sigma_v^2}\right)^m < 1$

$$\lim_{m\to\infty} \sigma_v^{2m}\mathbf{R}_y^{-m} = \mathbf{U}_n\mathbf{U}_n^H \tag{8.24}$$

Once $\mathbf{U}_n$ is found, $\mathbf{g}_{opt}$ can be estimated as the eigenvector which corresponds to the minimum eigenvalue of the quadratic matrix $\mathbf{C}^H\mathbf{U}_n\mathbf{U}_n^H\mathbf{C}$.

8.4.4 Disadvantages of the existing methods

In the previous algorithms there are some disadvantages that can be summarized as follows,

- In [64], the three LMS cost functions gave the same results. So one can consider only the low costs one.
- Cost functions in the previous approaches depend on the true error terms, for example the term $\boldsymbol{\lambda}^H \left(\mathbf{C}^H\mathbf{f} - \mathbf{g}\right) + \boldsymbol{\lambda} \left(\mathbf{f}^H\mathbf{C} - \mathbf{g}^H\right) + \rho \left(\mathbf{g}^H\mathbf{g} - 1\right)$ in ζ_1 and similar terms in ζ_2 and ζ_3. In adaptive LMS algorithms [70], it is better to consider the square of the error rather than the error itself specially when the additive noise is Gaussian. Using the error squares will also speed up the convergence rate
- From costs and estimation quality point of views
  - The vector $\mathbf{f}$ or $\mathbf{u}$ has the same dimension $(N + q - 1)$ as the received vector $\mathbf{y}(n)$ in [64] and this will increase the computation costs
  - For each received symbol, the adaptation of $\mathbf{f}$ and $\mathbf{g}$ will require around $(N + q - 1)$ and $(q \times 1)$ processing cycles respectively
  - In each adaptation cycle we need as well a $(q \times 1)$ λ's to be calculated
  - In case of long spreading code systems such as UMTS [6], calculations of $\mathbf{f}$will slow down the convergence rate
  - The LMS algorithms require estimation of two dependent vectors $\mathbf{f}$ and $\mathbf{g}$
  - In new higher bit rate CDMA systems such as 3G, the channel physical length q is large and this makes the SVD techniques costly
  - Calculation of $\mathbf{R}_y^{-1}$, $\mathbf{R}_y^{-m}$ or using SVD techniques in [64]-[68] obviously require heavy computation.

8.5 New Proposed Algorithms

In CDMA systems, since the spreading code of each user, $\mathbf{c}_k(n)$, is known to both, the transmitter and the receiver, then in order to reduce the system complexity and enhance the accuracy of the estimated channel parameters, we consider the adaptive based RAKE receiver shown in Fig. 8.2, where the vector $\mathbf{f}$ in Eqn. (8.6) is replaced by,

$$f(n) = c(n) * h(n) \qquad \text{or} \qquad \mathbf{f} = \mathbf{Ch} \tag{8.25}$$

where $\mathbf{C}$ is the same Toeplitz spreading matrix as before and the vector $\mathbf{h}$ contains the RAKE finger taps for user k.

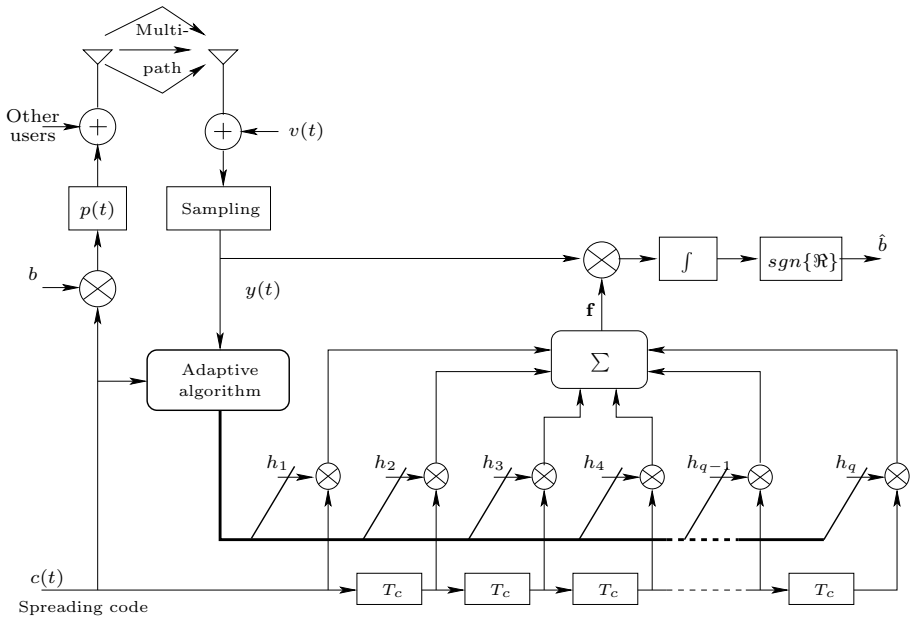


Figure 8.2: A blind adaptive based RAKE receiver

Since $\mathbf{C}$ is known, then one can see that instead of estimating an $(N + q - 1)$ parameters, $\mathbf{f}$ as in [64], we need to estimate only a q parameters vector, $\mathbf{h}$. Clearly this will reduce the system complexity and give better estimates.
With this definition, the constraint in Eqn. (8.8) $\mathbf{C}^H\mathbf{f} = \mathbf{g}^H$ becomes,

$$\mathbf{h}^H\mathbf{C}^H\mathbf{C} = \mathbf{g}^H \ , \ \mathbf{g}^H\mathbf{g} = 1 \tag{8.26}$$

The interesting point about Eqn. (8.26) is that the term $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}_{q\times q}$. For example for $q = 4$ in a maximum length binary sequence (MLBS) spreading code

based system with $N = 32$ chips,

$$\mathbf{C}^H\mathbf{C} = \begin{bmatrix} 1.0000 & -0.0312 & -0.0000 & 0.0312 \\ -0.0312 & 1.0000 & -0.0312 & -0.0000 \\ -0.0000 & -0.0312 & 1.0000 & -0.0312 \\ 0.0312 & -0.0000 & -0.0312 & 1.0000 \end{bmatrix} \tag{8.27}$$

Obviously the off-diagonal elements are very small compared to the diagonal elements.

Under this assumption, Eqn. (8.26) then becomes $\mathbf{h}^H \approx \mathbf{g}^H$. With $\mathbf{g}^H\mathbf{g} = 1$ we can enforce it with the original cost function in Eqn. (8.11) such that $\mathbf{h}^H\mathbf{h} = 1$ as $\mathbf{g}^H\mathbf{g} = 1$.

Therefore, let us consider the optimum solution, $\mathbf{h}_{opt}$ which is the Lagrange multipliers given in Section 8.3 by,

$$\mathbf{h}_{opt} = \underset{||\mathbf{h}||=1}{\operatorname{argmin}}\ \mathbf{h}^H\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\mathbf{h}. \tag{8.28}$$

Using the SVD, $\mathbf{R}_y^{-1}$ can be expressed by,

$$\mathbf{R}_y^{-1} = \mathbf{U}\ \mathbf{S}\ \mathbf{V} \tag{8.29}$$

where, $\mathbf{U}$ and $\mathbf{V}$ are the two SVD unitary matrices such that $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^H\mathbf{V} = \mathbf{I}$ and the eigenvalues matrix $\mathbf{S} = \operatorname{diag}(\lambda_l)\,,\ l = 1, 2, \ldots, q$.

Eqn. (8.28) then becomes,

$$\mathbf{h}_{opt} = \underset{||\mathbf{h}||=1}{\operatorname{argmin}}\ \mathbf{h}^H\mathbf{U}'\ \mathbf{S}\ \mathbf{V}'\mathbf{h}. \tag{8.30}$$

where $\mathbf{U}' = \mathbf{C}^H\mathbf{U}$ and $\mathbf{V}' = \mathbf{C}^H\mathbf{V}$.

Under the assumption that $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}$ one can show that $\mathbf{U}'$ and $\mathbf{V}'$ are also approximately unitary matrices, in other words,

$$\begin{aligned} \mathbf{U}'\mathbf{U}'^H &= \mathbf{C}^H\mathbf{U}\mathbf{U}^H\mathbf{C} \\ &= \mathbf{C}^H\mathbf{C} \\ &\approx \mathbf{I} \end{aligned} \tag{8.31}$$

and,

$$\begin{aligned} \mathbf{V}'\mathbf{V}'^H &= \mathbf{C}^H\mathbf{V}\mathbf{V}^H\mathbf{C} \\ &= \mathbf{C}^H\mathbf{C} \\ &\approx \mathbf{I} \end{aligned} \tag{8.32}$$

Using matrix proparties, the optimum channel response $\mathbf{h}_{opt}$ can then also be found by,

$$\mathbf{h}_{opt} = \underset{||\mathbf{h}||=1}{\text{argmax}}\ \mathbf{h}^H \mathbf{U}'\ \mathbf{S}\ \mathbf{V}'\mathbf{h} \tag{8.33}$$

i.e.,

$$\mathbf{h}_{opt} = \underset{||\mathbf{h}||=1}{\text{argmax}}\ \mathbf{h}^H \left(\mathbf{C}^H \mathbf{R}_y \mathbf{C}\right) \mathbf{h} \tag{8.34}$$

8.5.1 Advantages of the new methods

In [64], channel estimation requires minimization of the cost function w.r.t $\mathbf{f}$ and maximization w.r.t $\mathbf{g}$. In this approach, obviously the optimum solution in Eqn. (8.34) is less complex, does not require estimation of the two vectors, and requires only one vector, $\mathbf{h}_{opt}$. Also in the equation, one can see that finding $\mathbf{h}_{opt}$ does not require the estimation of $\mathbf{R}_y^{-1}$ or $\mathbf{R}_y^{-m}$ compared to the solution proposed by [64]-[68].

Estimation of $\mathbf{h}_{opt}$ in this work will be determined by,

1. A fast adaptive LMS algorithm
2. A low cost SVD algorithm
3. A maximum (dominant) eigenvalue power based algorithm

The three methods can be summarized as follows,

8.5.2 Proposed adaptive LMS algorithm

8.5.2.1 First adaptive LMS algorithm

In this algorithm we use a cost function similar to the one in [65]. Our cost function will be low in complexity, also it will not assume that $\mathbf{C}\mathbf{C}^H \approx \mathbf{I}$. Since we assumed that $\mathbf{f} = \mathbf{C}\mathbf{h}$, then based on the optimization problem in Eqns. (8.10) and (8.11), one can estimate the channel response $\mathbf{g}$ by minimizing the following cost function ζ w.r.t $\mathbf{h}$ and maximizing it w.r.t $\mathbf{g}$ such that the norm of $\mathbf{g}$ remain constant. The cost function ζ is given by,

$$\zeta = \left[\mathbf{f}^H \mathbf{R}_y \mathbf{f} + \lambda \left(\mathbf{C}^H \mathbf{f} - \mathbf{g}\right) \left(\mathbf{f}^H \mathbf{C} - \mathbf{g}^H\right)\right]_{\mathbf{f}=\mathbf{Ch}}$$

i.e.,

$$\zeta = \mathbf{h}^H \left(\mathbf{C}^H \mathbf{R}_y \mathbf{C}\right) \mathbf{h} + \lambda \left(\mathbf{C}^H \mathbf{C}\mathbf{h} - \mathbf{g}\right) \left(\mathbf{h}^H \mathbf{C}^H \mathbf{C} - \mathbf{g}^H\right) \tag{8.35}$$

With the previous cost function one can gain the following,

- Instead of adaptively estimating $(N + q - 1)$ parameters, $\mathbf{f}$ as in [64], we only need to estimate $2q$ parameters, $\mathbf{g}$ and $\mathbf{h}$.
- The process requires a single λ instead of a $(q \times 1)$ vector
- The error will be minimized in a square sense instead and this will enhance the convergence rate.

By using the cost function in Eqn. (8.35) and the LMS techniques, both $\mathbf{h}$ and $\mathbf{g}$ can be estimated by minimizing ζ in Eqn. (8.35) w.r.t $\mathbf{h}$ and maximizing it w.r.t $\mathbf{g}$, i.e.,

$$\mathbf{h}_{n+1} = \mathbf{h}_n - \mu_{\mathbf{h}} \nabla_{\mathbf{h}^*} \zeta \tag{8.36}$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_{\mathbf{g}} \nabla_{\mathbf{g}^*} \zeta \tag{8.37}$$

Since $\mathbf{P}_{\mathbf{g}_n} = \left[\mathbf{I} - \frac{\mathbf{g}_n \mathbf{g}_n^H}{\mathbf{g}_n^H \mathbf{g}_n}\right]$ is orthogonal to $\mathbf{g}_n$, one can also use

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_{\mathbf{g}} \mathbf{P}_{\mathbf{g}_n} \nabla_{\mathbf{g}^*} \zeta \tag{8.38}$$

where,

$$\nabla_{\mathbf{h}^*} \zeta = 2 \left(\mathbf{C}^H \mathbf{R}_y \mathbf{C}\right) \mathbf{h}_n + 2\lambda_n \left(\mathbf{C}^H \mathbf{C}\right) \left(\mathbf{C}^H \mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right) \tag{8.39}$$

$$\nabla_{\mathbf{g}^*} \zeta = -2\lambda_n \left(\mathbf{C}^H \mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right) \tag{8.40}$$

By using Eqn. (8.9) and enforcing the constraint of $\mathbf{h}$ as $\left(\mathbf{C}^H \mathbf{C}\right) \mathbf{h}_{n+1} = \mathbf{g}_n$, one can solve for λ_n.

With a simple manipulation one can show that λ_n is the smallest or the least square solution of $\mathbf{A}\lambda_n = \mathbf{B}$, where

$$\begin{aligned} \mathbf{A} &= -2\mu_h \left(\mathbf{C}^H \mathbf{C}\right) \left(\mathbf{C}^H \mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right) \\ \mathbf{B} &= \left(\mathbf{C}^H \mathbf{C}\right)^{-1} \mathbf{g}_n - \mathbf{h}_n + 2\mu_h \left(\mathbf{C}^H \mathbf{R}_y \mathbf{C}\right) \mathbf{h}_n \end{aligned} \tag{8.41}$$

The algorithm shown in Table 8.1 summarizes the full procedure.

Table 8.1: First proposed LMS algorithm

Step 1: At $n = 0$ initialize $\mathbf{h}_n$ and $\mathbf{g}_n$ and choose the step sizes μ_h and μ_g, $0 < \mu_g, \mu_h < 1$

Step 2: Compute $\left(\mathbf{C}^H\mathbf{C}\right)$ and $\left(\mathbf{C}^H\mathbf{C}\right)^{-1}$

Step 3: For $n = 1, 2, \ldots$

1. Find λ_n, using $\mathbf{A}_n\lambda_n = \mathbf{B}_n$,
   $\mathbf{A}_n = -2\mu_h \left(\mathbf{C}^H\mathbf{C}\right)\left(\mathbf{C}^H\mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right)$,
   $\mathbf{B}_n = \left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_n - \mathbf{h}_n + 2\mu_h \left(\mathbf{C}^H\mathbf{y}(n)\mathbf{y}^H(n)\mathbf{C}\right)\mathbf{h}_n$
2. Find the gradients,
   $\nabla_{\mathbf{h}^*}\zeta = 2\left(\mathbf{C}^H\mathbf{y}(n)\mathbf{y}^H(n)\mathbf{C}\right)\mathbf{h}_n + 2\lambda_n\left(\mathbf{C}^H\mathbf{C}\right)\left(\mathbf{C}^H\mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right)$
   $\nabla_{\mathbf{g}^*}\zeta = -2\lambda_n\left(\mathbf{C}^H\mathbf{C}\mathbf{h}_n - \mathbf{g}_n\right)$
3. Update $\mathbf{h}_n$ and $\mathbf{g}_n$,
   $\mathbf{h}_{n+1} = \mathbf{h}_n - \mu_{\mathbf{h}}\nabla_{\mathbf{h}^*}\zeta$
   $\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_{\mathbf{g}}\left[\mathbf{I} - \frac{\mathbf{g}_n\,\mathbf{g}_n^H}{\mathbf{g}_n^H\mathbf{g}_n}\right]\nabla_{\mathbf{g}^*}\zeta$
4. Normalize $\mathbf{g}_{n+1}$ such that $\mathbf{g}_{n+1} = \frac{\mathbf{g}_{n+1}}{||\mathbf{g}_{n+1}||}$ becomes a unity norm vector
5. Continue until convergence

8.5.2.2 Second adaptive LMS algorithm

In this algorithm the estimation of the optimum response $\mathbf{h}_{opt}$ will be based on Eqn. (8.34) and the assumption that $\mathbf{C}\mathbf{C}^H \approx \mathbf{I}$. For each incoming data vector $\mathbf{y}(n)$, $\mathbf{h}_{opt}$ can be found by maximizing $\zeta = \mathbf{h}^H\left(\mathbf{C}^H\mathbf{R}_y\mathbf{C}\right)\mathbf{h}$ w.r.t $\mathbf{h}$ according to the gradient search algorithm,

$$\mathbf{h}_{n+1} = \mathbf{h}_n + \mu_{\mathbf{h}}\nabla_{\mathbf{h}_n^*}\zeta \tag{8.42}$$

where,

$$\nabla_{\mathbf{h}_n^*}\zeta = 2\left(\mathbf{C}^H\mathbf{R}_y(n)\mathbf{C}\right)\mathbf{h}_n \tag{8.43}$$

This is followed by normalizing the vector $\mathbf{h}_{n+1}$ by its norm so that the constraint $||\mathbf{h}_{n+1}|| = 1$ is valid. The full algorithm is shown in Table 8.2.

Table 8.2: Second proposed LMS algorithm

Step 1: Initialize $\mathbf{h}_n$ and choose a step size $0 < \mu_h < 1$

Step 2: Let $\mathbf{R}_y(n-1) = 0$ and choose the correlation matrix updating factor $0.9 \leq \beta < 1$

Step 3: For $n = 1, 2, \ldots$

1. $\mathbf{R}_y(n) = \beta\, \mathbf{R}_y(n-1) + \mathbf{y}(n)\mathbf{y}^H(n)$
2. $\nabla_{\zeta/\mathbf{h}_n^*} = 2\left(\mathbf{C}^H\mathbf{R}_y\mathbf{C}\right)\mathbf{h}_n$
3. Update: $\mathbf{h}_{n+1} = \mathbf{h}_n + \mu_{\mathbf{h}}\nabla_{\zeta/\mathbf{h}_n^*}$
4. Normalize $\mathbf{h}_{n+1}$ so that $\mathbf{h}_{n+1} = \frac{\mathbf{h}_{n+1}}{||\mathbf{h}_{n+1}||}$ becomes a unity norm vector
5. Continue until convergence

8.5.3 Proposed SVD algorithm

It is well known that the SVD based techniques require heavy computation. In this method we also used SVD to estimate the channel response, but with less calculations than the method given in [64] [65], as the new proposed method here does not involve the estimation of $\mathbf{R}_y^{-1}$ using a Kalman filter or $\mathbf{R}_y^{-m}$ as in [66] [67] [68] using two successive SVD operations, the first is to find the noise subspace of $\mathbf{R}_y$, $\mathbf{U}_n$, and the second is to find the eigenvalues and eigenvectors of $\mathbf{C}^H\mathbf{U}_n\mathbf{U}_n^H\mathbf{C}$.

In this method, using the same assumption as before that $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}_{q\times q}$, the optimum response $\mathbf{h}_{opt}$ is found as the eigenvector which corresponds to the maximum eigenvalue of $\mathbf{C}^H\mathbf{R}_y\mathbf{C}$. The SVD based algorithm in Table 8.3 can be applied for the estimation of $\mathbf{h}_{opt}$.

8.5.4 Proposed power algorithm

The eigenvectors or characteristic vectors of an $(L \times L)$ matrix $\mathbf{A}$ are the set of L-vectors $\mathbf{x} = \mathbf{u}_i$, $i = 1, \cdots, L$, $\mathbf{u}^H\mathbf{u} = 1$ which are the non-trivial solutions of

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \qquad i.e., \qquad \mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i \tag{8.44}$$

Table 8.3: Proposed SVD based algorithm

Step 1: Let $\mathbf{R}_y(n-1) = 0$ and choose the correlation matrix updating factor $0.9 \leq \beta < 1$

Step 2: For $n = 1, 2, \ldots$

1. $\mathbf{R}_y(n) = \beta\, \mathbf{R}_y(n-1) + \mathbf{y}(n)\mathbf{y}^H(n)$
2. Use SVD to find $\mathbf{h}_{n+1} = \mathsf{eigenvector}\{\mathbf{C}^{\mathsf{H}}\mathbf{R}_{\mathsf{y}}\mathbf{C}\}$ which corresponds to the maximum eigenvalue
3. Normalized $\mathbf{h}_{n+1}$ so that $\mathbf{h}_{n+1} = \frac{\mathbf{h}_{n+1}}{||\mathbf{h}_{n+1}||}$ becomes a unity norm vector
4. Continue until convergence

Any vector $\mathbf{x}$ can be expressed in terms of these orthonormal sets as,

$$\mathbf{x} = a_1\mathbf{u}_1 + a_1\mathbf{u}_1 + \cdots + a_L\mathbf{u}_L \tag{8.45}$$

where the orthonormal sets $\mathbf{u}_1, \mathbf{u}_1, \ldots, a_L\mathbf{u}_L$ are also called eigenvectors. If we assume there is a unique (only one) largest eigenvector say $\lambda_1 > \lambda_2 > \ldots > \lambda_L$ then, we can find λ_1 and its corresponding eigenvector $\mathbf{u}_1$ of the matrix $\mathbf{A}$ by the power method [71]. The method can be described by the following iterative equations.

$$\begin{aligned} \lambda_1 &= \lim_{n\to\infty} \frac{||\mathbf{x}_{n+1}||}{||\mathbf{x}_n||} \\ \mathbf{u}_1 &= \lim_{n\to\infty} \frac{\mathbf{x}_{n+1}}{||\mathbf{x}_{n+1}||} \end{aligned} \tag{8.46}$$

where,

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n \tag{8.47}$$

The previous iterative method is only valid for an initial choice of $\mathbf{x}_n$ which is not orthogonal to the matrix $\mathbf{A}$, otherwise the method will fail. The iterative steps shown in Table 8.4. clarify Eqn. (8.46).

Since $\mathbf{h}_{opt}$ is the eigenvector which corresponds to the maximum eigenvalue of $\mathbf{C}^H\mathbf{R}_y\mathbf{C}$, then based on the previous power method the algorithm shown in Table 8.5 can be used to estimate the channel response.

Table 8.4: Fining the dominant eigenvalue and eigenvector using power method

- Choose an initial guess of $\mathbf{x}_0$, which mathematically represents,
  $\mathbf{x}_0 = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_L\mathbf{u}_L$
- Find $\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = a_1\mathbf{A}\mathbf{u}_1 + a_2\mathbf{A}\mathbf{u}_2 + \cdots + a_L\mathbf{A}\mathbf{u}_L$
  i.e., $\mathbf{x}_1 = a_1\lambda_1\mathbf{u}_1 + a_2\lambda_2\mathbf{u}_2 + \cdots + a_L\lambda_L\mathbf{u}_L$
- Find $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1$
  i.e., $\mathbf{x}_2 = a_1\lambda_1^2\mathbf{u}_1 + a_2\lambda_2^2\mathbf{u}_2 + \cdots + a_L^2\lambda_L\mathbf{u}_L$
- $\vdots$
- $\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n$
  i.e., $\mathbf{x}_{n+1} = a_1\lambda_1^n\mathbf{u}_1 + a_2\lambda_2^n\mathbf{u}_2 + \cdots + a_L\lambda_L^n\mathbf{u}_L$
- Since λ_1 is the largest eigenvalue, eventually $\mathbf{x}_{n+1} \approx a_1\lambda_1^n\mathbf{u}_1$ and $\mathbf{x}_n \approx a_1\lambda_1^{n-1}\mathbf{u}_1$
- For large n, $\lambda_1 = \frac{||\mathbf{x}_{n+1}||}{||\mathbf{x}_n||}$ and $\mathbf{u}_1 = \frac{\mathbf{x}_{n+1}}{||\mathbf{x}_{n+1}||}$

Table 8.5: The proposed power method based algorithm

Step 1: Initialize $\mathbf{h}_n$

Step 2: Let $\mathbf{R}_y(n-1) = 0$ and choose the correlation matrix updating factor $0.9 \leq \beta < 1$

Step 3: For $n = 1, 2, \ldots$

1. $\mathbf{R}_y(n) = \beta\ \mathbf{R}_y(n-1) + \mathbf{y}(n)\mathbf{y}^H(n)$
2. $\mathbf{h}_{n+1} = \mathbf{C}^H\mathbf{R}_y(n)\mathbf{C}\mathbf{h}_n$
3. Normalize $\mathbf{h}_{n+1} = \frac{\mathbf{h}_{n+1}}{||\mathbf{h}_{n+1}||}$
4. Continue until convergence

8.5.5 Remarks on the case when $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$

In the previous algorithms we assumed that $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}$, this can be valid only for some spreading codes. In an $N = 31$ chip Gold code system with $q = 4$

multipaths, one may find,

$$\mathbf{C}^H\mathbf{C} = \begin{bmatrix} 1.0000 & 0.2581 & -0.0323 & -0.3226 \\ 0.2581 & 1.0000 & 0.2581 & -0.0323 \\ -0.0323 & 0.2581 & 1.0000 & 0.2581 \\ -0.3226 & -0.0323 & 0.2581 & 1.0000 \end{bmatrix} \tag{8.48}$$

It is clear that $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$. Then we cannot neglect the effect of off-diagonal elements. By returning back to the constraint in Eqn. (8.26) where, $\mathbf{h}^H\mathbf{C}^H\mathbf{C}\mathbf{g} = 1$, one will find the resultant estimated parameters of the previous algorithms will no longer be $\mathbf{h}$. These parameters let us call them $\mathbf{h}_b$ will then be biased by the term $\mathbf{C}^H\mathbf{C}$.

To validate Eqn. (8.26) and the minimization constraint that $\mathbf{h}_b^H\mathbf{h}_b = 1$, the unbiased parameters $\mathbf{h}_{ub}$ which represent an estimate of $\mathbf{g}$ are then determined by,

$$\mathbf{h}_{ub} = \left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{h}_b \tag{8.49}$$

Clearly we can still use the previous algorithms. We calculate the term $\mathbf{C}^H\mathbf{C}$ once and fix the biased parameters $\mathbf{h}_b$ at each iteration. The algorithm shown in Table 8.6 summarizes the method,

Table 8.6: The proposed algorithm for the case when $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$

Step 1: Initialize $\mathbf{h}_{b_n}$, the biased estimator

Step 2: Let $\mathbf{R}_y(n-1) = 0$ and choose the correlation matrix updating factor $0.9 \leq \beta < 1$

Step 3: For $n = 1, 2, \ldots$

1. $\mathbf{R}_y(n) = \beta\,\mathbf{R}_y(n-1) + \mathbf{y}(n)\mathbf{y}^H(n)$
2. Determine $\mathbf{h}_{b_{n+1}}$ using any of the previous algorithms in Tables 8.2, 8.3, 8.5
3. Normalize $\mathbf{h}_{b_{n+1}} = \frac{\mathbf{h}_{b_{n+1}}}{||\mathbf{h}_{b_{n+1}}||}$
4. Find the unbiased parameters $\mathbf{h}_{ub_{n+1}} = \left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{h}_{b_{n+1}}$
5. Continue until convergence

8.6 Identification of Significant Channel Parameters

In the current high data rate systems, the shortened chip duration made the channel length required for RAKE receivers or equalization very long (see Chapter 4 for more details). If one uses this length for channel parameters estimation,then he/she will end up with estimating many parameters where most of them are zero. In this section we use the above blind channel estimation techniques combined with a Hierarchical MDL based model selection method in Chapter 5 method to identify only non-zero parameters of the channel response.

8.6.1 The identification based criteria

For the k'th user model, say $\mathbf{g} = \mathbf{g}_k$ in Section. 8.2 with L complex parameters, the MDL criteria is given by,

$$\mathrm{MDL}(L) = \log(\sigma_{e_r}^2) + 2\frac{L}{\mathcal{N}}\log(\mathcal{N}) \tag{8.50}$$

where $\mathcal{N} = N + L - 1$ and $\sigma_{e_r}^2$ represent the observed data length and the residual errors variance respectively.

Since the optimum response $\mathbf{g}_{opt}$ is the vector which minimize the variance

$$\sigma^2 = \mathbf{g}^H \mathbf{C}^H \mathbf{R}_y^{-1}(n)\mathbf{C}\mathbf{g} \tag{8.51}$$

i.e.,

$$\mathbf{g}_{opt} = \underset{||\mathbf{g}||=1}{\operatorname{argmin}}\ \mathbf{g}^H \mathbf{C}^H \mathbf{R}_y^{-1}(n)\mathbf{C}\mathbf{g} \tag{8.52}$$

Then, if we assume that the residual errors, say, $\mathbf{e}_r$ which correspond to the optimum response $\mathbf{g}_{opt}$ are Gaussian, then one can replace the variance $\sigma_{e_r}^2$ in Eqn. (8.50) by

$$\sigma_{e_r}^2 = \mathbf{g}_{opt}^H \mathbf{C}^H \mathbf{R}_y^{-1}(n)\mathbf{C}\mathbf{g}_{opt} \tag{8.53}$$

At this stage the MDL criteria in Eqn. (8.50) can be used to choose between different models with different number of parameters, in other words to select the best fit model. However, using this directly, we will end up with $\left(2^L\right)$ different

models to compare.

In this work we use the hierarchical MDL model selection procedure we proposed in Chapter 5 to choose the best fit model. The order of entry in our method is based on some logical and theoretical considerations and that will lead us to compare between L differnet models only (see Chapter 5). The decision of the best fit model in our method can occur within one to $\left(\frac{L(L+1)}{2} - 1\right)$ cycle(s) only.

8.6.2 The identification procedure

In this hierarchical procedure we begin with the full model (parent model) that includes all parameters from $\hat{g}_k(1)$ to $\hat{g}_k(L)$ where L is the channel length. Then we, define $\mathbf{p}$ as the parent hierarchical code at this step, $\mathbf{p} = [1\ 1 \dots 1]_{(L\times 1)}$ and this means we have to estimate the full model (parent model) $\mathbf{g}_{opt}^{\mathbf{p}}$.

From the parent hierarchical code $\mathbf{p}$, one can generate the child hierarchical codes $\mathbf{c}_1,\ \mathbf{c}_2, \dots \mathbf{c}_L$ and estimate their model parameters, $\mathbf{g}_{opt}^{\mathbf{c}_1},\ \mathbf{g}_{opt}^{\mathbf{c}_2}, \dots \mathbf{g}_{opt}^{\mathbf{c}_L}$. For example, if one uses the SVD to estimate the above model parameters then the optimum solution of, say, the model corresponds to the code $\mathbf{c}_l$ can be found as the non-trivial eigenvector which corresponds to the minimum eigenvalue (greater than zero) of the cost function

$$\zeta_l = (\mathbf{C}\ \mathrm{diag}(\mathbf{c}_l))^H\, \mathbf{R}_y^{-1}(n)\, (\mathbf{C}\ \mathrm{diag}(\mathbf{c}_l)) \tag{8.54}$$

The final step is then to compare the MDL value based on Eqns. (8.50) and (8.53) of each child model and select the child which has the smallest value. If this child's MDL value is less than or equal to its parent one then we replace the parent code with the child code and continue, otherwise we keep it and stop.

The complete model selection procedure is shown in Fig. 8.3, and also summarized in Table 8.7.

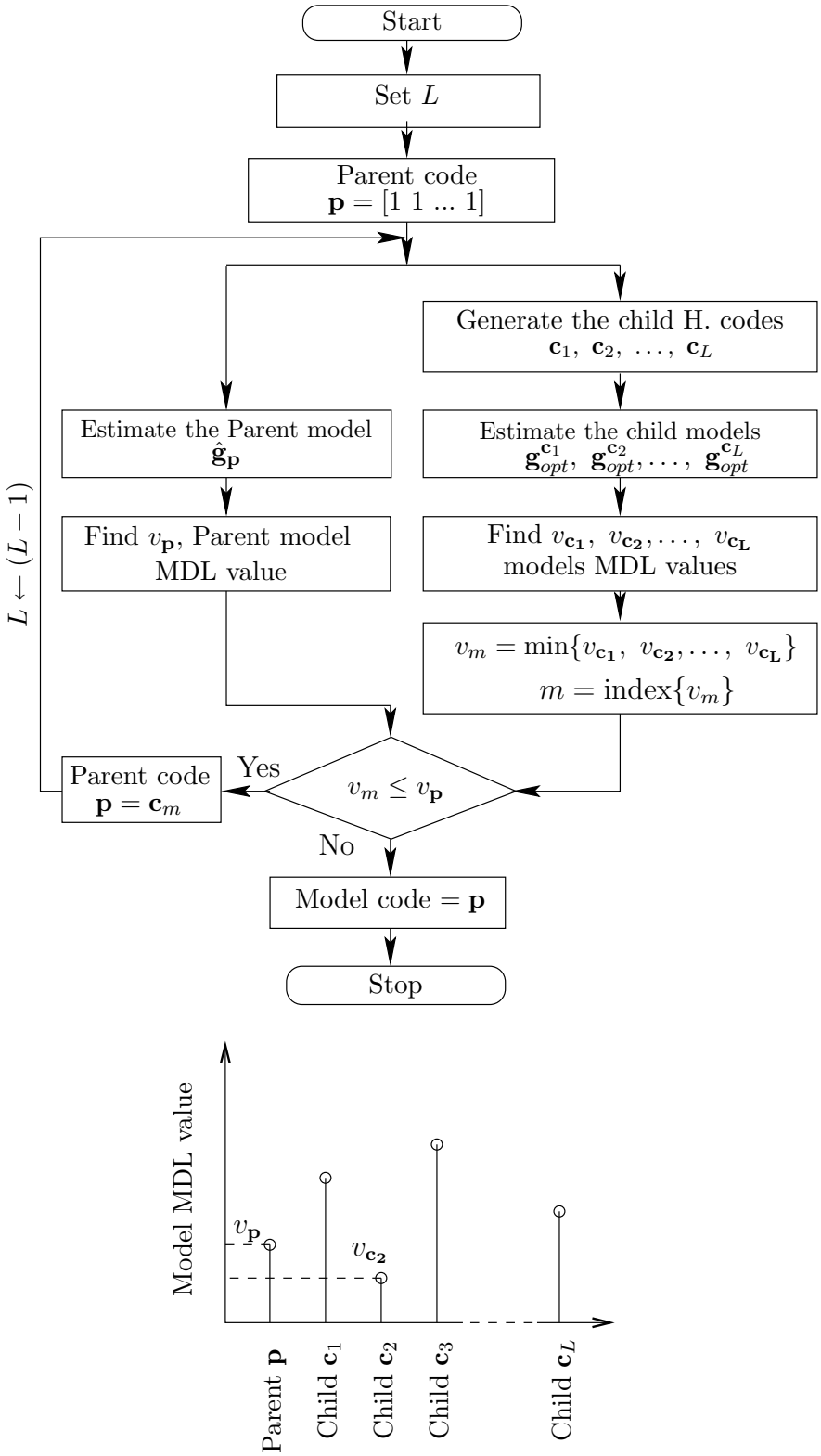


Figure 8.3: Hierarchical BE based method for model selection

Table 8.7: The hierarchical MDL based model selection procedure

Step 1:

Set the channel length L and/or use the MDL to estimate the model order, $L = \hat{d}$

Step 2:

Start the parent with the full code $\mathbf{p} = [1\ 1\ \ldots 1]_{L\times 1}$.

Step 3:

Estimate the parent model $\mathbf{g}_{opt}^{\mathbf{p}}$ using SVD based on Eqn. (8.54) and calculate its MDL value $v_{\mathbf{p}}$ based on Eqns. (8.50) and (8.53).

Step 4:

Generate the child hierarchical codes $\mathbf{c}_1,\ \mathbf{c}_2, \ldots \mathbf{c}_L$ then use SVD based on Eqn. (8.54) to estimate each code's optimum model, $\mathbf{g}_{opt}^{\mathbf{c}_l},\ l = 1, 2, .., L$.

Step 5:

For each child estimated model $\mathbf{g}_{opt}^{\mathbf{c}_l},\ l = 1, 2, .., L$, calculate the MDL values, $\mathbf{v} = [v_{\mathbf{c_1}},\ v_{\mathbf{c_2}}, \ldots v_{\mathbf{c_L}}]$.

Step 6:

Find the minimum MDL value among the child's values, i.e., find $v_m = \min\{\mathbf{v}\}$, where m is the index, then compare v_m with the parent one $v_{\mathbf{p}}$.

Step 7:

If $v_k \leq v_{\mathbf{p}}$, replace the parent code $\mathbf{p}$ by $\mathbf{c}_k$ then return to Step 3. otherwise **Stop** and $\mathbf{p}$ will be then the procedure decision about the model. One means a non-zero parameter and zero means a zero-parameter.

8.7 Simulation Results

In the following examples, we show the following,

- First we compare the performance of one of the new proposed algorithms with the algorithms in [64]-[68] that have the lowest MSE.
  - the proposed algorithm of Section 8.5.2.1 is compared with the RLS one of Section 8.4.1 proposed by [64] [65]
  - the proposed algorithm of Section 8.5.2.2 is compared with the SVD one of Section 8.4.1 proposed by [64]-[67]
- Simulation results for the new proposed methods,
  - the LMS algorithm in Section 8.5.2.2
  - the SVD algorithm in Section 8.5.3
  - the Power algorithm in Section 8.5.4
- MSE calculation
  - due to the scaling factor and phase ambiguity, the estimated vectors $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{h}$ in each case are normalized by their first parameters
  - for comparisons, the optimum solution $\mathbf{g}_o$ and $\mathbf{h}_o$ are estimated using the SVD technique calculated using an estimate of the correlation matrix $\mathbf{R}_y$ using the full signal length.
  - the MSE is then calculated as the MSE between the optimum solution and the estimated vector.
- Simulation for channel tracking

  In this simulation we track slow and frequency selective Rayleigh fading channels as follows,
  - the normalized second element say $\frac{h(2)}{h(1)}$ of each estimated channel is compared with the normalized second element of the optimum and the true channel

– The proposed algorithm in Section 8.5.2.1 and 8.5.2.2 are used to track frequency selective varying fading channels

- In all examples MLBS spreading codes are used except when mentioned, MC represents the number of Monte Carlo runs

8.7.1 Convergence rate performance

8.7.1.1 Example 1

In this example we compare the MSE between a channel which is estimated using the first proposed LMS algorithm given in Section 8.5.2.1 and the lowest MSE LMS algorithm presented by [64] in Section 8.4.1. The step size of each algorithm is chosen so that it gave the lowest MSE. The example input settings is given in Table 8.8 and the results are shown in Fig. 8.4

Table 8.8: Input settings for Example 1

Method	K	N	q	μ_f	μ_g	μ_h	SNR	MC
Ref. [64]	16	32	4	0.0009	0.02	-	6 dB	50
Section 8.5.2.1	16	32	4	-	0.01	0.0001	6 dB	50

8.7.1.2 Discussion of Example 1

Fig 8.4 compares the MSE between the blind adaptive LMS algorithm in [64] and the proposed LMS algorithm in Section 8.5.2.1. Simulation results show that, the proposed approach has a lower MSE and a faster convergence rate than the one in [64]. The proposed approach also required adaptation of $(2q+1)$ parameters only, ($\mathbf{h}$, $\mathbf{g}$ and λ) where the other approach required adaptation of $(N+3q-1)$ parameters, ($\mathbf{f}$, $\mathbf{g}$ and $\boldsymbol{\lambda}$). Obviously the estimation and the adaptation of a shorter length vector such as $\mathbf{h}$ in the proposed approach will be more accurate and faster than the longer length vector $\mathbf{f}$ as in [64].

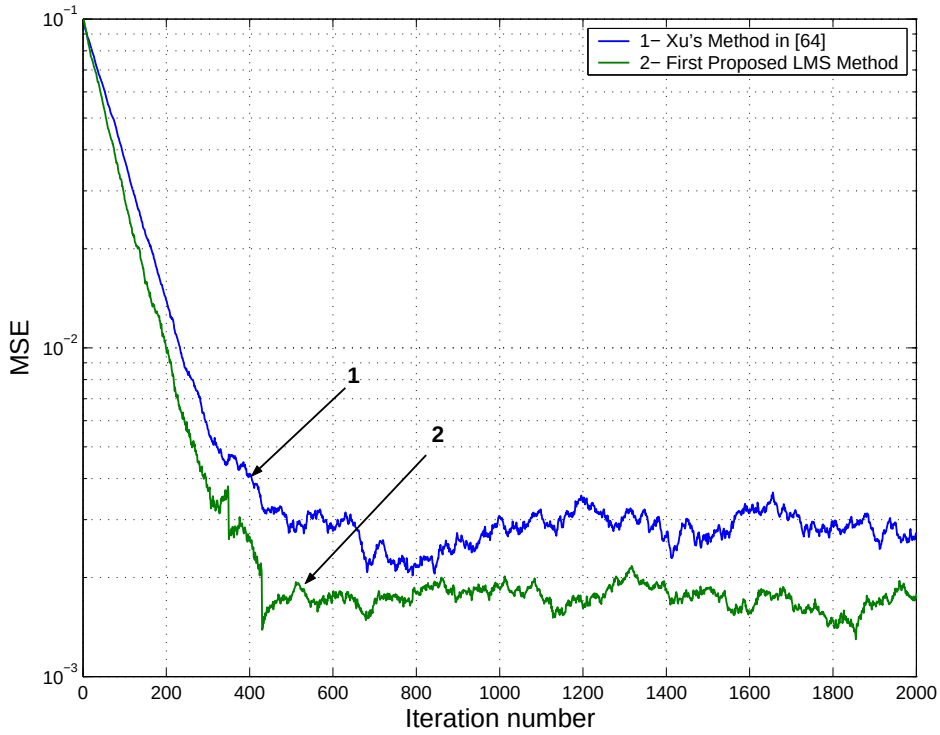


Figure 8.4: MSE in case of 16 users, first LMS algorithm

8.7.1.3 Example 2

In this example we compare the MSE between a channel which is estimated using the proposed LMS, SVD and power algorithms given in Sections 8.5.2.1, Sections 8.5.2.2, Sections 8.5.3 and Sections 8.5.4 and the adaptive RLS algorithm in Section 8.4.1 chosen to be the one which has the best performance in [64] [65] [66]. The input settings are given in Table 8.9 and the results are shown in Fig. 8.5.

Table 8.9: Input settings for Example 2

Method	K	N	q	μ_f	μ_g	μ_h	SNR	MC
Ref. [64]	16	32	4	0.0009	0.02	-	6 dB	50
Section 8.5.2.2	16	32	4	-	-	0.01	6 dB	50

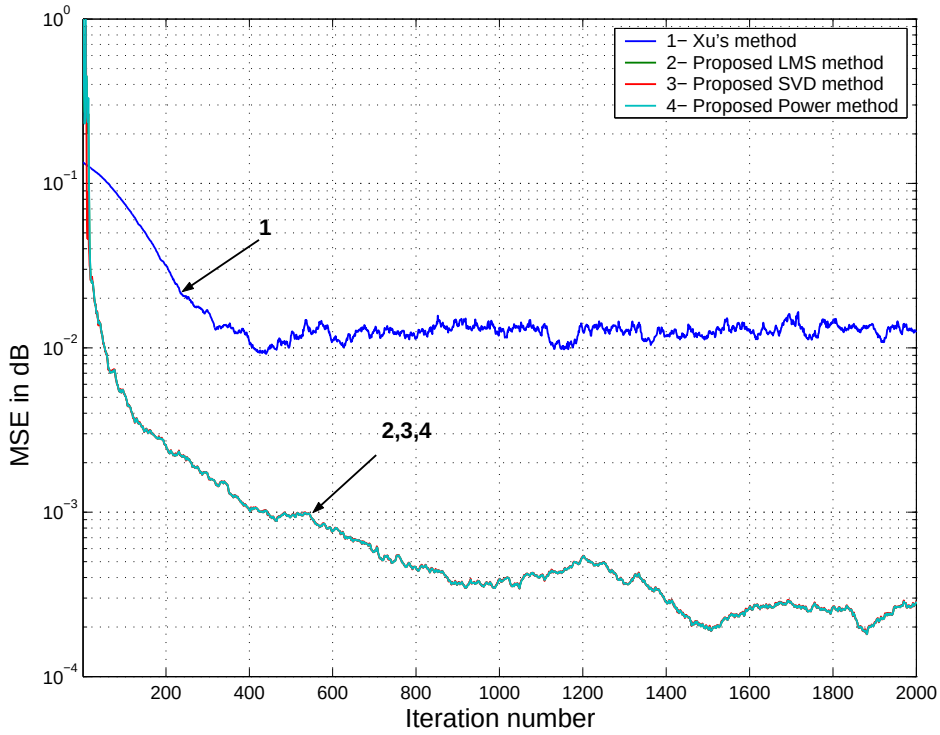


Figure 8.5: MSE in case of 16 users, second LMS algorithm

8.7.1.4 Discussion of Example 2

Fig 8.5 compares the MSE between the blind adaptive RLS algorithm in [64] and the three proposed algorithms in Section 8.5.2.2, Section 8.5.3 and Section 8.5.4. Simulation results show that, the proposed algorithms performed much better than the RLS algorithm in [64] [66]. The proposed algorithms complexity are also much cheaper than the RLS algorithms in [64] [66] which use Kalman filters for the estimation of $\mathbf{R}_y^{-1}$ or $\mathbf{R}_y^{-m}$.

8.7.1.5 Example 3

In this example we show the performance of the SVD algorithm given in Section 8.5.3 using the assumption that $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}$. The algorithm is applied to a system with $K = 16$ users, $N = 32$ chips, $q = 8$ paths. Results in Fig. 8.6 compares the MSE between:

- the estimated and the optimum channels, $\mathbf{h}$ and $\mathbf{g}_o$

- the estimated and the true channels, $\mathbf{h}$ and $\mathbf{g}$

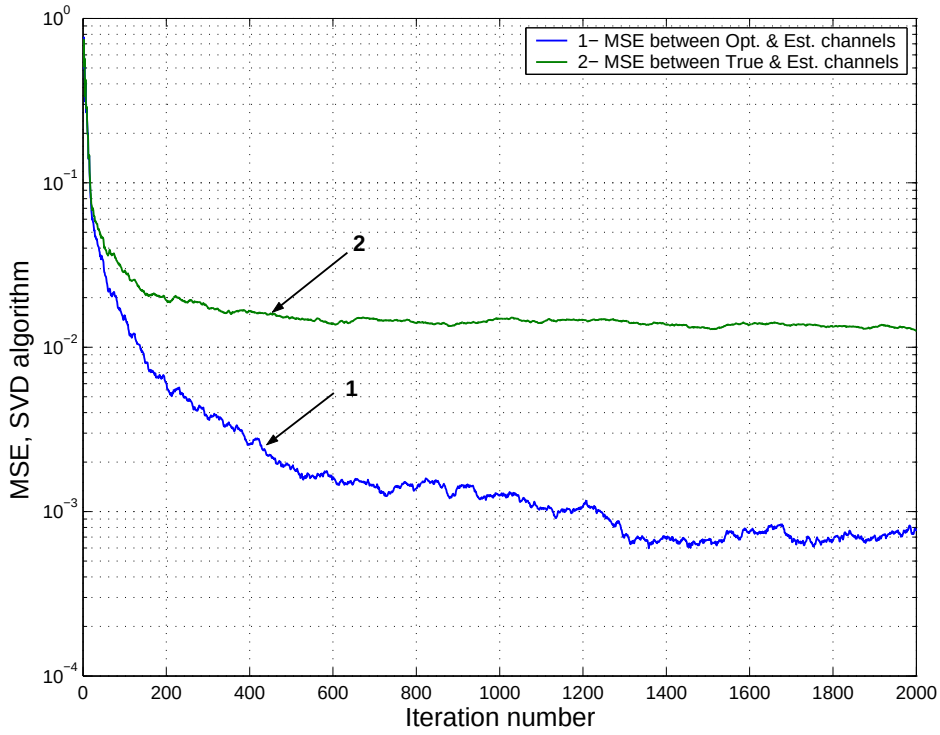


Figure 8.6: MSE between optimum, true and estimated channels, 16 users

8.7.1.6 Discussion of Example 3

Using the assumption that the term $\mathbf{C}^H\mathbf{C} \approx \mathbf{I}$, simulation results in Fig 8.6 show that the proposed adaptive SVD algorithm converges very fast to the optimum solution, also the MSE between the estimated channel, $\mathbf{h}$ and the true channel, $\mathbf{g}$ is small compared with MSE of the methods in [64] [67] [68]. The algorithm is low in complexity since it does not require the estimation of $\mathbf{R}_y^{-1}$ or $\mathbf{R}_y^{-m}$ as in [64] [67] [68] or the noise subspace as in [65]-[68].

8.7.1.7 Example 4

In this example we use the proposed power algorithms given in Section 8.5.4 to the case where $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$. The algorithm is applied to the case when $K = 16$ users, $N = 32$ chips, $q = 8$ paths. The MSE is compared as shown in Fig. 8.7 between:

- the biased optimum and the biased estimated channels, $\mathbf{g}_o$ and $\mathbf{h}_b$
- the unbiased optimum and the unbiased estimated channels, $\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_o$ and $\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{h}_b$
- the true and the biased estimated channels, $\mathbf{g}$ and $\mathbf{h}_b$
- the true and the unbiased estimated channels $\mathbf{g}$ and $\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{h}_b$

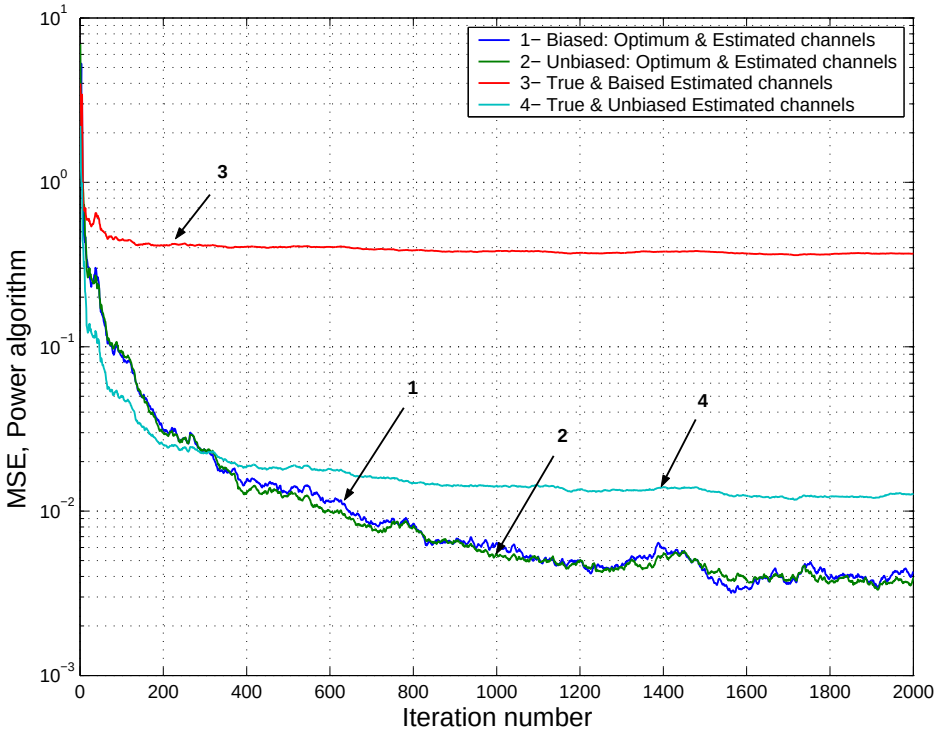


Figure 8.7: MSE between, optimum, true and estimated channels , 16 users, $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$

8.7.1.8 Discussion of Example 4

In this example we used the proposed power algorithms in Section 8.5.4 and the algorithm in Section 8.5.5 for the case when $\mathbf{C}^H\mathbf{C} \neq \mathbf{I}$. Simulation results in Fig 8.7 show that the MSE between the true channel response $\mathbf{g}$ and the biased estimated response $\mathbf{h}_b$ is high. After fixing $\mathbf{h}_b$ using, $\mathbf{h}_{ub} = \left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{h}_b$ in Section 8.5.5, the MSE between $\mathbf{g}$ and $\mathbf{h}_{ub}$ become very small, i.e., the algorithm succeeded in estimating the channel response.

8.7.2 Channel tracking performance

8.7.2.1 Example 5

In this example we use the proposed LMS algorithm in Section 8.5.2.2 to track the transient response of a time varying fading channel where the channel parameters are generated from a Rayleigh distribution and the channels vary each 200 symbols. The algorithm is applied to the case with $K = 4$ users, $N = 32$ chips, $q = 8$ paths, $\mu_h = 0.01$ at SNR = 6 dB and MC=20 runs. The results are shown in Fig. 8.8.

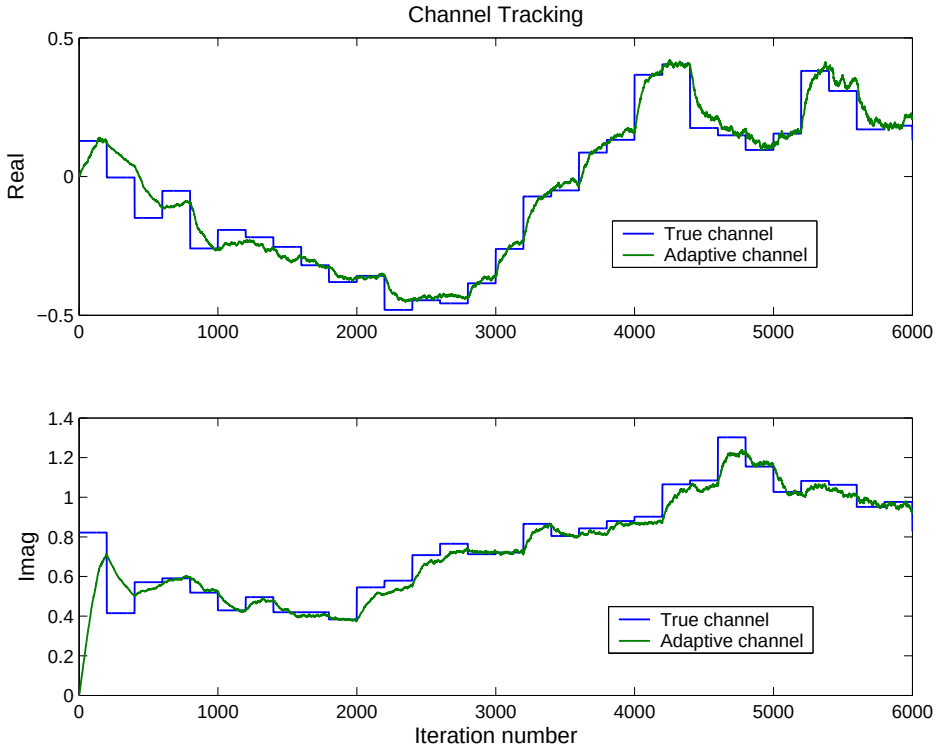


Figure 8.8: Tracking of a time varying fading channel

8.7.2.2 Discussion of Example 5

In the previous Examples 1-4, it is obvious that the convergence rate of the proposed blind adaptive algorithms is fast. Within 100 or 200 symbols any of the proposed algorithm can reach the optimum solution. In this example we used the proposed LMS algorithm to track a varying fading channel. Simulation results

support the claim that, within a few number of symbols the proposed algorithms can track the change in a varying fading channel with good quality estimates for the real and imaginary parts. Also it is clear that the algorithm can track a faster channel with 100 symbols. In the case of high SNRs a smaller number of symbols is then required.

8.7.2.3 Example 6

In this example we use the proposed LMS algorithm in Section 8.5.2.2 to track a time varying channel where the channel parameters are generated using the SOS Jakes model (see chapter 3). The mobile speed $v = 20$ Km/h, the carrier frequency, $f_c = 1$ GHz and $N_o = 16$ Oscillators. The algorithm is applied to the case where $K = 1$ users, $N = 32$ chips, $q = 4$ paths, $\mu_h = 0.01$ at SNR = 6 dB and MC=20 runs. Results are shown in Fig. 8.9.

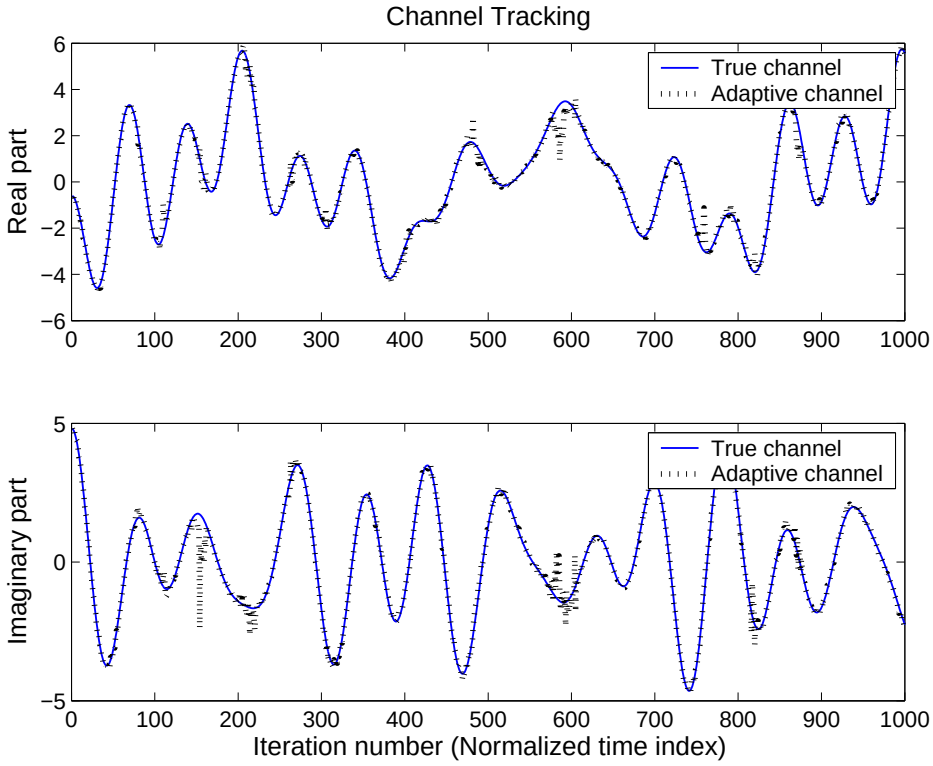


Figure 8.9: Tracking of a time varying fading SOS Jakes channel model

8.7.2.4 Discussion of Example 6

In this example we used the proposed LMS algorithm to track a strong time varying fading channel generated using the SOS Jakes model, which is used to represent a realstic channel response. Simulation results show clearly that the proposed algorithm can easily track the variations of the channel response.

8.7.3 Identification performance

8.7.3.1 Example 7

This example demonstrates the use of the Hierachial MDL based model selection method described in Section 5.4.1 and Chapter 5 in identifying the significant channel parameters of a frequency selective slowly fading channel introduced in a $K = 1, 2, 3, 4$ users system with maximum channel length of $d_o = 10$,$L = 16$ required physical taps, with $d_1 = 3$ significant paths, $N = 32$ chips and length $M = 64,\ 96,\ 128$ and 160 symbols. The channel components are given in Table 8.10. The results are shown in Fig. 8.10 and Fig. 8.11. Let $p_k(l)$ be the probability of identifying each individual path parameter $g_k(l)$ correctly, then one can show that the probability of identifying the full channel parameters of user k correctly p_f can be given by,

$$p_f = \prod_{l=1}^{L} p_k(l) \tag{8.55}$$

Table 8.10: Input settings and multipath parameters for Example 7

Delay	Multipath parameters $g_k(l)$ for			
τ_{1l}	User 1	User 2	User 3	User 4
0	0	0	0	0
T_c	-0.0207 - j0.4087	0	0	-0.2290 + j0.1217
$2T_c$	0	0.0347 + j0.4313	0.1423 -j0.0278	0.1732 - j0.2147
$3T_c$	0	0	0	-0.4983 - j0.1737
$4T_c$	0	0.0347 + j0.4313	-0.4201 + j0.4713	0
$5T_c$	0.5350 + j0.0278	0	0.6119 + j0.4301	0
$6T_c$	0.6730 + j0.3042	-0.5138 - j0.1791	0	-0.2825 + j0.4796
$7T_c$	0	-0.0842 - j0.3971	0	-0.3364 + j0.3774
$8T_c$	0	0	0	-0.0598 - j0.0957
$9T_c$	0	-0.0842 - j0.3971	0.1423 - j0.0278	0
$10T_c$	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$16T_c$	0	0	0	0

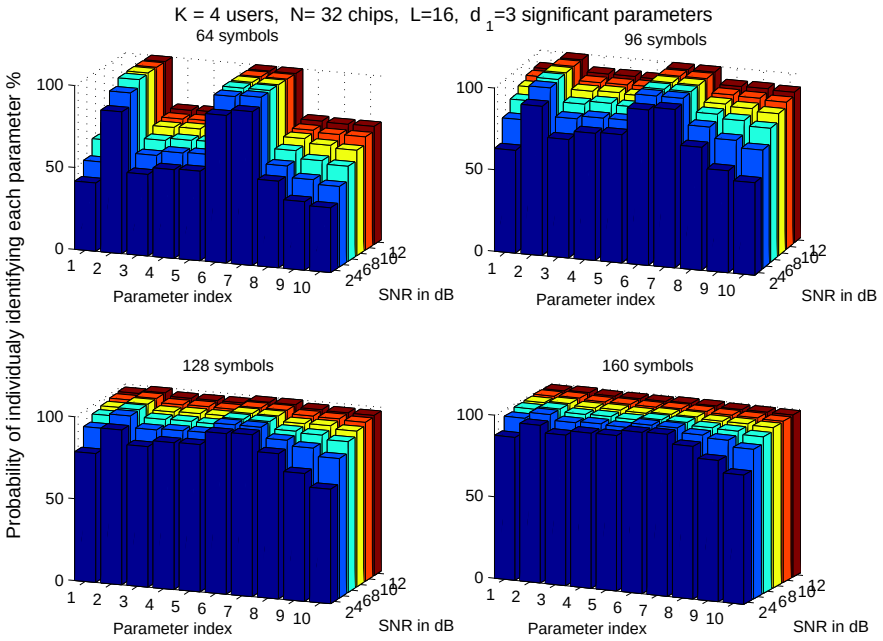


Figure 8.10: The probability of correctly identifying each channel parameter of user one, $p_k(l)$

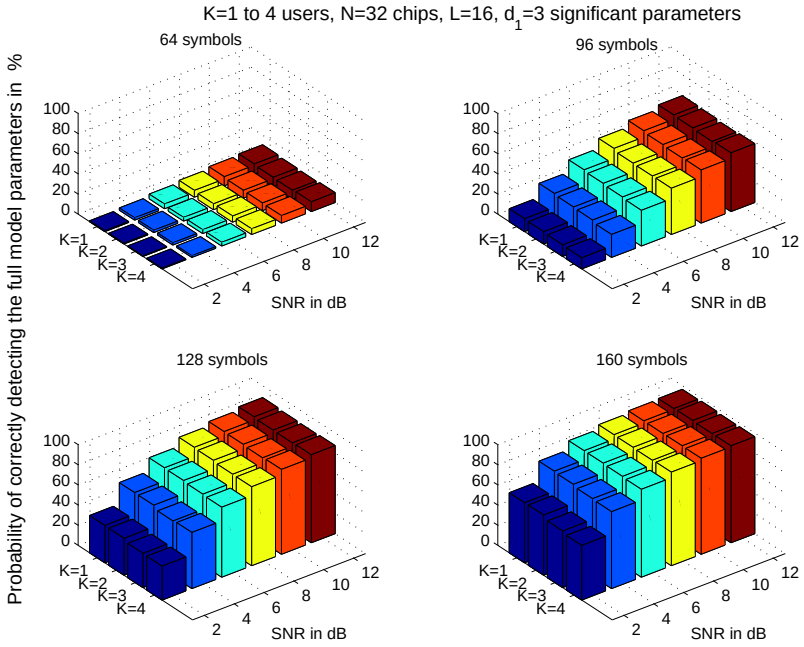


Figure 8.11: The probability of identifying the full model parameters of user one, p_f

8.7.3.2 Discussion of Example 7

From the results, one can see that the proposed model selection method works well when number of symbols M is greater than 96 symbols. For example it is better if one runs the system with $N \geq 128$ symbols for SNRs ≥ 4 dB . However, in practice, many wireless communication systems operate with SNRs greater than 8 dB. If the system is running at SNRs less than 4 dB, a longer sequence length such as $M \geq 160$ symbols is then required.

8.8 Conclusion

The problem of blind channel estimation and tracking in CDMA systems is considered. Using only the spreading code of the user of interest and the received data sequence, four low complexity algorithms are proposed to estimate and track the impulse response of the multipath channel. While existing blind methods suffer from high computational complexity, due to large SVDs and matrices inversion, the proposed methods overcome both problems. The computation complexity of the new algorithms are less in comparison with the existing approaches. Simulation results show that, the proposed algorithms' convergence rates are fast and their estimation accuracy are much better. The proposed approaches are also able to track the variation of frequency selective varying fading channels. The algorithms can be implemented as blind adaptive based RAKE receivers in CDMA systems.

Chapter 9

Conclusions and Future Directions

9.1 Conclusion

In this thesis we developed new approaches and algorithms for the identification and estimation of low order models that represent multipath propagation effects in CDMA communication systems. Based on these parsimonious channel models, low complexity receivers such as the RAKE receivers are considered. RAKE receivers were able to exploit the multipath channel propagation effects to enhance the system performance. We considered three different cases of channel propagation:

1. Flat fading channels where the channel attenuation and delays are constant over the full signaling intervals

2. Frequency selective slowly fading channels where the channel attenuation and delays are constant over a few signaling intervals

3. Frequency selective varying fading channels where the channel attenuation and delays can vary over each signaling interval

We modeled the above channel effects in the frequency and time domains as follows:

1. In the frequency domain: as complex attenuated exponentials

2. In the time domain: as an FIR-like filter (or a tapped delay line filter)

The number of taps of each model was related to the ratio between the channel delay-spread and the chip duration. Unlike classical techniques which estimate directly the channel response given the number of taps or given an estimate of the channel true length, the proposed techniques in this work firstly used model selection techniques to identify only the significant parameters of the channel model. Specifically, under a certain level of significance, statistical tests are proposed to determine whether or not each individual parameter is significant. The level of significance with which we made this assertion was controlled based on statistical tests such as the multiple hypothesis tests. Three training and a low complexity blind channel estimation based algorithms are then used to estimate only the identified or classified significant parameters. Finally, parsimonious RAKE receivers are then considered based on these low order models. The proposed receivers are found to have a better BERs over the classical receivers. The usage of the significant multipath components only is strongly recommended based on our findings. A full conclusion about these approaches are summarized as follows:

- Chapter 5: We addressed the problem of model selection. Four different approaches for model selection are presented. In these approaches low computation complexity algorithms are developed based on model selection methods such as MDL, AIC, F-statistic and the bootstrap techniques. The MDL and AIC based methods are normally used when the model additive noise is Gaussian, but they do not control the level of significance. They are nevertheless still widely used in many different areas in signal processing.
  Under some assumptions, the F-statistic based methods can identify the significant model parameters and consider the level of significance. The performances of the F-statistic based method is found to be better when the additive noise is Gaussian.
  Under minimal assumptions about the noise distribution and for a predefined level of significance, the bootstrap technique is a powerful technique to identify the significant model parameters.

Finally, all of the above model selection approaches are found to be capable of identifying the principle (the MDL and AIC cases) or the significant (F-test and bootstrap cases) model parameters with high probabilities.

- Chapter 6: A frequency domain based approach is proposed to estimate the channel delays and attenuation coefficients. Then two model selection approaches including the MDL sphericity test based approach and the multiple hypothesis F-statistic based tests were used to achieve the identification goal. Both approaches are low in computation complexity since they only involve the FFT to estimate the spectra. The F-statistic SRB based approach is found to be more robust than the MDL one, firstly because it can control the approach level of significance and secondly because it uses the EVD only once for the calculation of the channel delays using MUSIC algorithm. The identification and BER simulation results support the claim that with a short training data the structure complexity of the proposed RAKE receiver is reduced and the system performance is also enhanced.

- Chapter 7: Two bootstrap time domain based approaches are proposed for the identification and estimation of the significant multipath parameters in multiuser DS-CDMA systems. The approach is used to enhance the RAKE receiver performance over other classical methods. Unlike the proposed approach in Chapter 6 where the statistical test distribution is known, this approach used the bootstrap resampling techniques to estimate the unknown distribution of the test statistic. The bootstrap is found to be more flexible than classical statistical methods as it is able to compute a statistic and estimate its sampling distribution without any assumption on the model or knowledge of the noise distribution. Two different bootstrap resampling techniques are used: the classic bootstrap and the surrogate data bootstrap. The approach model is also simple since it assumes the channel as an FIR filter with taps located at multiple integers of the chip duration. Results show that with the proposed approach the performance of the parsimonious RAKE receiver can increased substantially with a small

processing gain. Increasing processing gain or at higher SNRs, the accuracy of the estimates obviously increases.

- Chapter 8: The problem of blind channel estimation and tracking in CDMA systems is considered. Using only the spreading code of the user of interest and the received data sequence, four low computation complexity algorithms are proposed to estimate and track the impulse response of the multipath channel. While existing blind methods suffer from high computation complexity, due to large SVDs and matrix inversions, the proposed methods overcome both problems. The computation complexity of the new algorithms are less in comparison with the existing approaches. Also a low computation complexity blind hierarchical MDL model selection based method is proposed to select non-zero parameters of the channel response. Simulation results show that, the proposed algorithms convergence rates are fast and their estimation accuracy are much better. The proposed approaches are also used to track the variation a frequency selective varying fading channels. The algorithms can be used to implement blind adaptive based RAKE receivers in CDMA systems.

9.2 Future Work

The following points summarize some future research directions:

- DS-CDMA communication systems have attracted considerable attention for 3G mobile systems. They have the ability to suppress a wide variety of interfering signals including narrow-band interference (NBI), MAI, and MPI. However, the need for higher capacity systems has resulted in the duration of the transmitted symbols being shorter compared to the channel time dispression. Accordingly, the ISI is increased and the CDMA system performance is reduced. Around 1993, MC-CDMA systems are proposed to overcome this ISI problem and add more orthogonality between users. This is done by splitting the fast data rate (stream) into N low-rate sub-streams,

where N is the number of sub-carrier. Thus, the data symbol rate per carrier is reduced by a factor N in other words the ISI is reduced. As the ISI become less, this can be interprated as the channel fading become flat. Thus, one can use the frequency domain approach proposed in Chapter 6 for channel estimation in MC-CDMA system. In case of flat fading channel, the frequency approach can estimate the channel response with low computation complexity in the frequency domain and less estimation errors. OFDM is a form of MC-CDMA modulation, it can be generated as MC DS-CDMA modulation where all users simultaneously share the available BW and the separation of the users' signal is carried out in the code domain similar to DS-CDMA. The channel response of OFDM systems is been regarded as flat fading, i.e., the frequency approach can also be used for channel estimation.

- The channel Doppler spread is one of the big issues which affects the RAKE receiver performance if not considered. The proposed RAKE in this thesis was based on a certain number of fingers (taps) where each of them has an estimate of the corresponding path attenuation presented at a certain delay, say, τ. The performance of such RAKE will be reduced if the presense of the channel Doppler spread is significant. One of the solutions to such problem is knows as time-frequency RAKE. In time-frequency RAKE each tap will not only consider a single attenuation coefficient at τ but also considers what is know as the side band attenuation. In other words, if the channel Doppler frequency is f_D then the tap L coefficient at τ can be model as $\sum_{m=-M}^{M} g_{lm}(\tau)e^{j\frac{2\pi m}{M}f_D}$. The interesting thing about this model is that, $g_{lm}(\tau)$ can be consider flat over the signaling interval. Thus, the proposed algorithms in this can also be used to estimated and identify the significant parameters for parsimonious time-frequency RAKEs.

- In this thesis we considered only a single antenna based RAKE receiver. As a new task, one can consider multiple antennas based RAKE receiver which is known as 2D-RAKE or Space-Time RAKE. The concept of a 2D-

RAKE is that a beam be formed toward each multipath of the desired user using an antenna array and the resultant combined using a 1D-RAKE receiver (time-only Rake). Thus, instead of using only the RAKE receiver time diversity, space diversity can also be used. Based on that the channel estimation and identification algorithms proposed in Chapters 6, 7 and 8 will have a better performance in estimating and identifying the significant channel parameters in case of multiple antenna.

Bibliography

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*, Revised Ed. New York: McGraw- Hill, 1994.

[2] R. Price and P. E. Green, "A Communication Technique for Multipath Channels," In *Proc. of the IRE, vol. 46, pp. 555-570 1958*

[3] Hui Liu and Kemin Li, "A Decorrelating RAKE Receiver for CDMA Communications Over Frequency-Selective Fading Channels," *IEEE Trans. Commun.*, vol. 47, pp. 1036-1045, July 2001.

[4] J. G. Proakis, *Digital Communications*. McGraw-Hill series In Electrical Engineering, 2001.

[5] G. L. Stuber, *Principles of Mobile Communication*. 2nd Edition, Kluwer Academic Publisher, London, 2001.

[6] A. Springer and R. Weigel, *UMTS: The Universal Mobile Telecommunications Systems*. Springer-Verlag Berlin Heidelberg New York, 2002.

[7] E. Dahlman, B. Gudmundson, M. Nilson and J. Skold, "UMTS/IMT-2000 based on wideband CDMA," *IEEE Commun. Magzine*, vol 36, pp. 70-80, June 1998.

[8] C. Sengupta, J. R. Cavallaro, B. Aazhang, "On Multipath Channel Estimation for CDMA Systems Using Multiple Sensors," *IEEE Trans. Commun.*, vol. 49, pp. 543-553, March 2001.

[9] S. E. Bensley and B. Aazhang, "Subspaced-Based Channel Estimation for Code Division Multiple Access Communication," *IEEE Trans. Commun.*, vol. 44, pp. 1009-1020, August 1996.

[10] U. Madhow and M. L. Honig, "MMSE Interference Suppression for Direct-Sequence Spread Spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178-3188, December 1994

[11] A. A. M. Saleh and R. A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Trans. Select. Areas Commun.*, vol. SAC-5, pp.128-137, February 1987

[12] K. J. Kim, S. Y. Kwon, E. K. Hong and K. C. Whang, "Effect of Tap Spacing on the Performance of Direct-Sequence Spread-Spectrum RAKE Receiver," *IEEE Trans. Commun.*, vol. 48, No 6, pp. 1029-1036, June 2000.

[13] G. Kutz, A. Chass, "Low Complexity Implementation of a Downlink CDMA Generalized RAKE Receiver," In *Proc. of IEEE 56th Vehicular Technology Conf.*, vol. 3, pp.1357-1361, September 2002.

[14] G. Kutz, A. Chass, "On the Performance of a Practical Downlink CDMA Generalized RAKE Receiver," In *Proc. of IEEE 56th Vehicular Technology Conf.*, vol. 3, pp. 1352-1356, September 2002.

[15] J. Joutsensalo, "Algorithms for Delay Estimation and Tracking in CDMA," In *Proc. IEEE Intl. Conf. Commun. (ICC'97)*, vol. 1, pp. 366-370, Montreal, June 1997.

[16] A. Akaike, "Information Theory and an Extensions of the Maximum Likelihood Principle," In *Proc. 2nd IEEE Intl. Symposium Info. Theory*, 1973.

[17] A. Barron, J. Rissanen and B. Yu, "The Minimum Description Length Principle in Coding and Modeling," *(Special Commemorative Issue: Information Theory, 1948-1998) IEEE. Trans. Inform. Theory* pp. 2743-2760, October 1998.

[18] J. Michel, P. Balaban and K. Sam Shanmugan, *Simulation of The Communication Systems*. 2nd Edition, Kluwer Academic/Plenum Publishers, New York, 2000.

[19] A. Viterbi, *Principle of Spread Spectrum Communication*. Addison-Wesley Publishing, 1995.

[20] S. Glisic and B. Vucetic, *Spread Spectrum CDMA Systems for Wireless Communication*. Artech House Publisher, Boston, London, 1997.

[21] S. Verdu, *Multiuser Detection*. Cambrdige University Press, 1998.

[22] http://www.3gpp.org

[23] http://www.itu.int

[24] http://www.wikipedia.org

[25] C. Y. Lee William, *The mobile communications handbook*, 2nd Ed., Wiley Series in Telecommunications, 1995.

[26] M. Paetsch, *The evolution of mobile communications in the U.S. and Europe: regulation, technology, and markets*, Artech House, Boston, 1993.

[27] V. B. Lawrence, M. R. Karim, M. Sarraf, *W-CDMA and cdma2000 for 3G Mobile Networks* McGraw-Hill series in Tecommunications.

[28] M. J. Omidi, S. Gazor, P. G. Gulak, S. Pasupathy, "Differential Kalman filtering for tracking Rayleigh fading channels," In *Proc. of IEEE signal processing systems (SIP'98)* , vol. 1, pp. 376-385, October, 1998.

[29] L. Lindbom, "Simplified Kalman estimation of fading mobile radio channels: high performance at LMS computational load," In *Proc. of IEEE Acoustics, Speech, and Signal Processing* , vol. 3, pp. 352-355, April 1993.

[30] J. Wang, Y. Qiang, B. Zhang and D. Li, "Comparison of statistical properties between an improved Jakes' model and the classical one," In *Proc. of IEEE*

Personal, Indoor and Mobile Radio Commun. (PIMRC'03), vol. 1, pp. 380-384, Septemeber 2003.

[31] A. A. El-sallam, A. M. Zoubir and Samir Attallah, "A Bootstrap-Based RAKE Receiver for CDMA Systems," In *Proc. of IEEE Globecom'02*, vol. 2, pp. 1073-1077, Taiwan, Novemeber 2002.

[32] H. Meyer, M. Moeneclaey and S. A. Fechtel, *Digital Communication Receivers Synchronization, Channel Estimation and Signal Processing.* John Wiley & Sons, 1998.

[33] H. Holma, A. Toskala, *WCDMA for UMTS Radio Access For Third Generation Mobile Communications.* John Wiley & Sons 2001.

[34] F. J. Altman and W. Sichak, "A Simplified Diversity Communication System for Beyond-the-horizon Links," *IRE Trans. Commun. Systems*, vol. CS-4, pp. 50-55, March 1956.

[35] D. G. Brennan, "Linear Diversity Combining Techniques," In *Proc. of the IEEE*, vol. 91, no. 2, January 1959, February 2003.

[36] A. W. Rihaczek, *Principles of High-Resolution Radar.* Peninsula Publishing, 1985.

[37] L. Ljung, *System Identification. Theory for the User.* Prentice-Hall, 1987.

[38] A. M. Zoubir and B. Boashash, "The Bootstrap and Its Application in Signal Processing," *IEEE Signal Processing Magazine*, vol. 15, pp. 56-76, 1998.

[39] M Wax and T. Kailath, "Detection of Signals by Information Theoretic Criteria," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 33, pp. 387-392, 1985.

[40] H. Linhart and W. Zucchini, *Model Selection.* Wiley, 1986.

[41] M. H. DeGroot and M. J. Schervish, *Probability and Statistics.* Addison Wesely, 3rd Edition 2002.

[42] T. Wonnhacott, *Wonnhacott Regression: Asecond course in statistics.* John Wiley & Sons, September 1981.

[43] P. M. Djuri, "Using the Bootstrap to Select Models," In *Proc. of the IEEE ICASSP 97*, vol. V, pp. 3729-3732, Munich, Germany, 1997.

[44] J. Shao, "Bootstrap Model Selection," *Journal of American Statist. Assoc.*, vol. 91, pp. 655-665, 1996.

[45] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap.* Chapman and Hall, 1993.

[46] D. Williams and D. Johnson, "Using the sphericity test for source detection with narrow-band passive arrays," *IEEE Trans. Acoust., Speech, Signal Processing, vol. 38, pp. 2008-2014, November 1990.*

[47] A. James, "The distribution of the latent roots of the covariance matrix," *Ann. Math. Statist., vol. 31, pp. 151-158, 1960*

[48] R. F. Brcich, A. M. Zoubir and P. Pelin, "Detection of Sources Using Bootstrap Techniques," *IEEE Trans. on signal processing, vol. 50, no. 2, pp. 206-215 February 2002*

[49] T. Anderson, "Asymptotic theory for principal component analysis," *Ann. Statist.,vol. 34, pp. 122-148, 1963.*

[50] C. Waternaux "Asymptotic distribution of the sample roots for a nonnormal population" *Biometrika,vol. 63, no. 3, pp. 639-645, 1976.*

[51] A. A. Abd El-sallam, S. Kayhan and A. M. Zoubir, "Bootstrap and Backward Elimination-Based Approaches for Model Selection," In *Proc. of 3rd Intl. Symposium on Image and Signal Processing and Analysis (ISPA'03), Rome*, vol. 1, pp. 152-157, Rome, September 2003.

[52] Y. Hochberg and A. C. Tamhane, *Multiple Comparison Procedures.* John Wiley, 1987.

[53] E. L. Lehmann, *Testing Statistical Hypothesis*. John Wiley & Sons, 1986.

[54] A. M. Zoubir and R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge University Press, 2004.

[55] S. M. Kay, *Modern Spectral Estimation*. Prentice-Hall Signal Processing Series 1988.

[56] D. R. Brillinger and P. R. Krishnaiah *Time Series in the Frequency Domain*. Elsevier Science Publishing, 1983

[57] D. R. Brillinger, *Time Series: Data Analysis and Theory*. Holden-Day 1981.

[58] P. Stocia and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao Bound," *IEEE Trans. Acoustics Speech and Signal Processing*, vol. 37, pp. 720-741, May 1989.

[59] R. Roy, A. Paulraj and T. Kailath, "ESPRIT-A Subspace Rotation Approach to Estimation of Parameters of Cisoids in Noise," *IEEE Trans. Acoustics Speech and Signal Processing*, vol. ASSP-34, pp. 1340-1342, February 1986.

[60] T. Ostman, S. Parkvall and B. Ottersten, "An improved MUSIC Algorithm For Estimation of Time Delays in Asynchronous DS-CDMA Systems," *IEEE Trans. Commun.*, vol. 47, pp 1628-1631, Novemeber 1999.

[61] A. M. Zoubir and B. Boashash, "The Bootstrap and Its Application in Signal Processing", *IEEE Signal Processing Magazine*, vol. 1, pp. 56-76, 1998.

[62] A. Abd El-Sallam, A. M. Zoubir and S. Attallah, "Low order Channel Estimation For CDMA", In *Proc. of the 3rd Australian Commun. Theory Workshop (AusCTW)*, vol 1., pp. 105-108, Canberra, Australia, Feb. 4-5, 2002.

[63] M. Tastsains and Z. Xu, "Performance Analysis of Minimum Variance CDMA Receivers", *IEEE Trans. Signal Processing*, vol. 46, no. 11, pp. 3014-3022, Novemeber 1998.

[64] Z. Xu. and M. Tastsains, "Blind Adaptive Algorithms for Minimum Variance CDMA Receivers", *IEEE Trans. Commun.*, vol. 49, no. 1, , pp. 180-194, January 2001.

[65] Z. Xu, "Blind Channel Estimation Via Subspace Approximation", *Proc. of the 37 Asilomar conf. on Signals, Systems & Computers*, vol. 2, pp. 1653-1657, Novemeber 2003.

[66] X. G. Doukopoulos and G. V. Moustakides, "Blind Channel Estimation for Downlink CDMA Systems", In *Proc. of the IEEE Intl. Conf. on Commun. (ICC'03)*, vol. 4, pp 2416-2420, May. 2003.

[67] X. G. Doukopoulos and G. V. Moustakides, "Power Techniques for Blind Adaptive Channel Estimation in CDMA Systems", In *Proc. of the IEEE GLOBECOM'03*, vol. 4, pp 2330 - 2334, December 2003.

[68] Z. Xu, P. Liu and X. Wang, "Blind Multiuser Detection: From MOE to Subspace Methods", *IEEE Trans. Signal processing*, vol. 52, no. 2, pp. 510-524, February 2004.

[69] W. A. Hamouda and P. J. McLane, "A Fast Adaptive Algorithm for MMSE Receivers in DS-CDMA Systems", *IEEE Signal processing Letter*, vol. 11, no. 2, pp. 86-89, February 2004.

[70] S. Haykin, *Adaptive Filter Theory* Prentice Hall, 4th Edition, September, 2001.

[71] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge University Press, New York, 2001.

[72] G. Yue, X. Zhou, X. Wang, "Performance comparisons of channel estimation techniques in multipath fading CDMA", *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 716-724, May. 2004.

[73] K. Zheng, G. Zheng and W. Wang, "A novel uplink channel estimation in OFDM-CDMA systems", *IEEE Trans. Consumer Electronics*, vol. 50, no. 1, pp. 125-129, Feb 2004.

[74] A. D'Amico, U. Mengali and M. Morelli, "Channel estimation for the uplink of a DS-CDMA system ", *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1132-1137, Novemeber 2003.

[75] A. A. Abd El-sallam and A. M. Zoubir, "A Frequency Domain Based Approach For Low complexity RAKE Receivers In CDMA Systems", In *Proc. of the IEEE Workshop on Statis. Signal Processing. (SSP'03)*, vol. , pp. 313-316, St. Louis, USA, September 2003.

[76] Y. Hwan You, W.i Jeon, J. Paik, D. Hong and H. Song, "Training sequence design and channel estimation of OFDM-CDMA broadband wireless access networks with diversity techniques", *IEEE Trans. Broadcasting*, vol. 49, no. 4, pp. 354-361, December 2003.

[77] A. J. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes", *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1561-1569, October 1999.

[78] E. Aktas and U. Mitra, "Semiblind Channel Estimation for CDMA Systems With Parallel Data and Pilot Signals", *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1102-1112, July 2004.

[79] S. Lasaulce, P. Loubaton and E. Moulines, "A semi-blind channel estimation technique based on second-order blind method for CDMA systems", *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1894-1904, July 2003.

[80] E. de Carvalho and D. Slock, "Cramer Rao bounds for semi-blind, blind and training sequence based channel estimation " In *Proc. of the First IEEE Workshop on Signal Processing Advances in Wireless Commun.*, pp. 129-132, 1997.

[81] Z. Xu and M. K. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems", *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 988-1001, April 2000.

[82] C. Sengupta, J. R. Cavallaro and B. Aazhang, "On multipath channel estimation for CDMA systems using multiple sensors", *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 543-553, March 2001.

[83] M. Torlak and X. Guanghan, "Blind multiuser channel estimation in asynchronous CDMA systems", *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 137-147, January 1997.

[84] S. Roy and Y. Hongbo, "Blind channel estimation in multi-rate CDMA systems", *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 995-1004, June 2002.

Vision without action is dreaming.
Action without vision is passing the time.
Action with vision can change the world.