ADAPTIVE FILTER THEORY

FOURTH EDITION

SIMON HAYKIN

PRENTICE HALL INFORMATION AND SYSTEM SCIENCES SERIES

Thomas Kailath Series Editor

ADAPTIVE FILTER THEORY

Fourth Edition

Simon Haykin
Communications Research Laboratory
McMaster University
Hamilton, Ontario, Canada

PRENTICE HALL
Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

Haykin, Simon S.,
Adaptive filter theory / Simon Haykin.—4th ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-090126-1
1. Adaptive filters. I. Title.
TK7872.F5 H368 2001
621.3815'324—dc21

2001021604

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Publisher: *Tom Robbins*
Associate Editor: *Alice Dworkin*
Editorial Assistant: *Jody McDonnell*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Managing Editor: *David A. George*
Production Services / Compositor: *Preparé, Inc.*
Director of Creative Services: *Paul Belfanti*
Creative Director: *Carole Anson*
Art Director: *Jayne Conte*
Art Editor: *Adam Velthaus*
Cover Designer: *Joseph Sengotta*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lynda Castillo*
Marketing Manager: *Holly Stark*
Marketing Assistant: *Karen Moon*

Prentice Hall

The author and publisher of this book have used their best efforts in preparing this book. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, expressed or implied, with regard to these programs or the documentation contained in this book. The author and publisher shall not be liable in any event for incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of these programs.

MATLAB is a registered trademark of The MathWorks, Inc.,
3 Apple Hill Drive, Natick, MA 01760-2098.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

0-13-090126-1

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educacíon de Mexico, S.A. de C.V.
Pearson Education—Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.

This book is dedicated to

- The many researchers whose significant contributions to adaptive filtering have made it possible to write the book.
- The many reviewers, readers, and students who have helped me to improve the book through insightful and critical comments.

Preface

The subject of adaptive filters constitutes an important part of statistical signal processing. Whenever there is a requirement to process signals that result from operation in an environment of unknown statistics or one that is inherently nonstationary, the use of an adaptive filter offers a highly attractive solution to the problem as it provides a significant improvement in performance over the use of a fixed filter designed by conventional methods. Furthermore, the use of adaptive filters provides new signal-processing capabilities that would not be possible otherwise. We thus find that adaptive filters have been successfully applied in such diverse fields as communications, control, radar, sonar, seismology, and biomedical engineering, among others.

Aims of the Book

The primary aim of this book is to develop the mathematical theory of various realizations of *linear adaptive filters*. Adaptation is accomplished by adjusting the free parameters (coefficients) of a filter in accordance with the input data, which, in reality, makes the adaptive filter nonlinear. When we speak of an adaptive filter being "linear," we mean the following: The input–output map of the filter obeys the principle of superposition whenever, at any particular instant of time, the filter's parameters are all fixed.

There is no unique solution to the linear adaptive filtering problem. Rather, we have a "kit of tools" represented by a variety of recursive algorithms, each of which offers desirable features of its own. This book provides such a kit. It also provides an introduction to neural networks, which are basic to nonlinear adaptive filtering.

In terms of background, it is assumed that the reader has taken introductory undergraduate courses on probability theory and digital signal processing; undergraduate courses on communication and control systems would also be an advantage.

Organization of the Book

The book begins with an introductory chapter, where the operations and different forms of adaptive filters are discussed in general terms. The chapter ends with historical notes, which are included to provide a source of motivation for the interested reader to plough through the rich history of the subject. The concepts and algorithms introduced in this chapter are explained in detail in subsequent parts of the book.

The main chapters of the book, 17 in number, are organized as follows:

- *Stochastic processes and models.* This material, presented in Chapter 1, emphasizes partial characterization (i.e., second-order statistical description) of stationary stochastic processes. As such, it is basic to much of what is presented in the rest of the book.
- *Wiener filter theory and its application to linear prediction.* The Wiener filter, presented in Chapter 2, defines the optimum linear filter for a stationary environment, and therefore provides a framework for the study of linear adaptive filters. Linear prediction theory, encompassing both of its forward and backward forms and variants thereof, is discussed in Chapter 3; the chapter finishes with the application of linear prediction to speech coding.
- *Least-mean-square (LMS) family of adaptive filters.* The LMS filter is built around a transversal (i.e., tapped-delay-line) structure. In its most basic form, it is simple to design, yet highly effective in performance—two practical features that have made it highly popular in various applications. Chapter 4 presents the fundamentals of an old optimization technique known as the method of steepest descent, from which the LMS filter is readily derived. Chapter 5 presents a detailed treatment of the many facets of the LMS filter, its theory and practical applications. The two highlights of the chapter are:
  - **(i)** *Small-step-size statistical theory*, which provides a fairly accurate description of the transient behavior of the LMS filter and its learning curve when the step-size parameter is assigned a small value. This new theory, rooted in the Langevin equation of nonequilibrium thermodynamics, avoids the unrealistic assumptions made in the independence theory traditionally used in the study of LMS filters. Computer simulations are presented, demonstrating close agreement between the findings of the small-step-size theory and experimental results.
  - **(ii)** H^{∞} *theory*, which provides the mathematical basis for the deterministic robustness of the LMS filter.

  Chapters 6 and 7 expand on the LMS family of LMS filters by presenting detailed treatments of normalized LMS filters, affine projection adaptive filters, frequency-domain and subband adaptive LMS filters; the affine projection filter is an intermediate adaptive filter between the normalized LMS filter and recursive least-squares filter.
- *Recursive least-squares (RLS) adaptive filters.* The RLS filter overcomes some practical limitations of the LMS filter by providing a faster rate of convergence and an insensitive performance to variations in the eigenvalue spread of the correlation matrix of the input signal; the price paid for these improvements is increased computational complexity. Chapter 8 discusses the method of least squares, which may be viewed as the deterministic counterpart of the Wiener filter rooted in stochastic processes. In the method of least-squares, the input data are processed on a block-by-block basis; block methods, disregarded in the past because of their numerical complexity, are becoming increasingly attractive, thanks to continuing improvements in digital computer technology. Chapter 9 builds on the method of

least squares, and uses the matrix inversion lemma to derive the RLS filter. The highlights of this chapter are:

(i) Statistical theory of RLS filters.

(ii) H^∞ theory of RLS filters.

In reality, the RLS adaptive filter is a special case of the celebrated Kalman filter that emphasizes the notion of a state. It is therefore important that we have a good understanding of Kalman filter theory, which, in a stationary environment, also includes the Wiener filter as a special case. Derivations of the Kalman filter, its variants and extension, are discussed in Chapter 10. This chapter also establishes the one-to-one correspondences between Kalman filters and RLS filters, thereby providing the framework for a unified treatment of the whole family of RLS adaptive filters. Chapter 11, built on Givens rotations, derives the square-root information and covariance forms of square-root RLS filters, which overcome numerical difficulties encountered in the digital implementation of ordinary RLS filters. Chapter 12, on order-recursive filters, describes latticelike structures for the design of adaptive filters that are of the same order of computational complexity as LMS filters; the improvement in computational complexity is achieved by exploiting the time-shifting property inherent to temporal processing. Chapter 12 presents derivations of the gradient adaptive lattice (GAL) and QR-decomposition-based least-squares lattice (QRD–LSL) algorithms, which are respectively the simplest in computational terms and most powerful in performance terms.

- *Finite-precision effects.* The theory of linear adaptive filters presented in Chapters 5 through 12, is based on continuous mathematics (i.e., infinite precision). When, however, any adaptive filter is implemented in digital form, effects due to the use of finite-precision arithmetic arise. Chapter 13 discusses these effects in the digital implementation of LMS and RLS filters.
- *Tracking of time-varying systems.* Chapter 14 expands on the theory of LMS and RLS filters by evaluating and comparing their performances when they operate in a nonstationary environment, assuming a Markov model. The chapter uses the extended Kalman filter to derive modified forms of RLS filters that are assured of optimality over the LMS filter in a nonstationary environment. The chapter finishes by deriving new forms of LMS and RLS filters with adaptive memory which respectively permit recursive adjustments to the step-size parameter and exponential weighting factor of these filters.
- *Infinite-duration impulse response (IIR) adaptive filters.* The linear adaptive filters discussed in Chapters 5 through 14 are all built around transversal or lattice structures that are characterized by an impulse response of finite duration. Chapter 15 presents a brief treatment of the output-error method and equation-error method for the design of IIR adaptive filters, and discusses practical issues that arise in their use. The chapter also describes Laguerre adaptive filters that combine the desirable properties of FIR and IIR structures.
- *Blind deconvolution.* The linear adaptive filters studied in Chapters 5 through 15 assume the availability of a desired response, which is intended to guide the adap-

tation of free parameters of the filter during training. When there is no desired response available, we have to resort to the use of blind adaptation (i.e., unsupervised adaptive filtering). Chapter 15 discusses two primary blind deconvolution algorithms:

- Subspace decomposition algorithms based on second-order statistics, which exploit the cyclostationary nature of modulated signals in applications involving data transmission over communication channels.
- Bussgang algorithms, which exploit higher order statistics of the received signal at the output of a communication channel.
- The chapter finishes with a discussion of Bussgang fractionally spaced equalizers, where these two major themes of blind deconvolution are tied together.

- *Back-propagation learning.* Neural networks, built around nonlinear processing units, provide powerful tools for solving difficult nonlinear adaptive filtering problems. Multilayer perceptrons, with one or more hidden layers of processing units, constitute an important class of neural networks. Chapter 17 derives the back-propagation algorithm for the supervised training of multilayer perceptrons; the back-propagation algorithm may be viewed as a generalization of the LMS algorithm. The novel feature of the material presented in this chapter is emphasis on the complex form of the back-propagation algorithm, which makes it consistent with the rest of the book.

The book concludes with an Epilogue, where the following five topics are briefly discussed:

- Proportionate adaptation
- Robust statistics
- Blind source separation
- Dynamically driven recurrent neural networks
- Derivative-free state estimation for nonlinear dynamically systems

These discussions are included to provide the reader with a more complete picture of the ever-expanding subject of adaptive filters.

The book also includes appendices on the following topics:

- Complex variable theory
- Differentiation with respect to a vector
- Method of Lagrange multipliers
- Estimation theory
- Eigenanalysis
- Rotations and reflections
- Complex Wishart distribution

In different parts of the book, use is made of the fundamental ideas presented in these chapters.

Ancillary Material

- A Glossary is included, consisting of a list of definitions, notations and conventions, a list of abbreviations, and a list of principal symbols used in the book.
- All publications referred to in the text are compiled in the Bibliography. Each reference is identified in the text by the name(s) of the author(s) and the year of publication. The Bibliography also includes many other references that have been added for completeness.

Examples, Computer Experiments, and Problems

Many examples are included in different chapters of the book to illustrate concepts and theories under discussion.

The book also includes many computer experiments that have been developed to illustrate the underlying theory and applications of the LMS and RLS algorithms. These experiments help the reader to compare the performances of different members of these two families of linear adaptive filtering algorithms.

Each chapter of the book, except for the introductory chapter, ends with problems that are designed to do two things:

- Help the reader to develop a deeper understanding of the material covered in the chapter.
- Challenge the reader to extend some aspects of the theory discussed in the chapter.

Solutions Manual

The book has a companion solutions manual that presents detailed solutions to all the problems at the end of Chapters 1 through 17 of the book. A copy of the manual can be obtained by instructors who have adopted the book for classroom use by writing directly to the publisher.

The MATLAB codes for all the computer experiments can be accessed by going to the web site http://www.prenhall.com/haykin/

Use of the Book

The book is written at a level suitable for use in graduate courses on adaptive signal processing. In this context, it is noteworthy that the organization of the material covered in the book offers a great deal of flexibility in the selection of a suitable list of topics for such a graduate course.

It is hoped that the book will also be useful to researchers and engineers in industry as well as government establishments, working on problems relating to the theory and applications of adaptive filters.

Simon Haykin

Acknowledgments

I would like to record my gratitude to Dr. Hans Butterweck, Eindhoven University of Technology, The Netherlands, Dr. Babak Hassibi, California Institute of Technology, and Dr. Philip Regalia, Institut National des Telecommunications, France, for helping me in the writing of selected sections of the book. I am most grateful to Dr. James Zeidler, University of California, San Diego, for a critical review of an early version of the manuscript of the book and subsequently for giving me many helpful comments and suggestions for improving the book. Thanks are also due to Dr. Neil Bershad, University of California at Irvine, and Dr. Ali H. Sayed, University of California at Los Angeles, and an anonymous reviewer for useful comments and feedback on an early version of the manuscript.

Many others have been kind enough to critically read selected chapters/sections of the book; in alphabetical order, they are:

Dr. Sandro Bellini, Politechnico di Milano, Italy
Dr. Jacob Benesty, Lucent Technologies, Murray Hill, NJ
Dr. Zhi Ding, University of California, Davis
Dr. Lee Feldkamp, Ford Motor Company, Dearborn, MI
Dr. Steven Gay, Lucent Technologies, Murray Hill, NJ
Dr. Eberhard Hänsler, Darmstadt University of Technology, Germany
Dr. Thomas Kailath, Stanford University, California
Dr. Harold Kushner, Brown University, Rhode Island
Dr. John Makhoul, Bolt, Beranek & Newman, Cambridge, Mass.
Dr. John McWhirter, DERA, Malvern, England
Dr. Marc Moonen, Katholieke Universiteit Leuven, Belgium
Dr. George V. Moustakides, University of Patras, Greece
Dr. Ian Proudler, DERA, Malvern, England
Dr. Sanzheng Qiao, McMaster University, Canada
Dr. Ali H. Sayed, University of California at Los Angeles
Dr. John Shynk, University of California at Santa Barbara
Dr. Piet Sommen, Eindhoven University of Technology, The Netherlands
Dr. Stergiopoulos Theordoridis, University of Athens, Greece
Prof. P. P. Vaidyanathan, California Institute of Technology, Pasadena
Dr. Eric Wan, Oregon School of Graduate Studies, Portland, Oregon

I am grateful to the British Crown, the Institute of Electrical and Electronics Engineers, Dr. Eric Wan (Oregon School of Graduate Studies), and Dr. Jacob Benesty (Lucent Technologies), for giving me permission to reproduce certain figures, which are cited in the text.

I thank Dr. Hugh Pasika, Applied Biosystems Inc., San Francisco, CA, Dr. Mathini Sellathurai, McMaster University, Hamilton, Ont., Canada, Dr. Paul Yee, PMC Sierra, Richmond, BC, Canada, for developing the MATLAB codes for the experiments described in the book.

I wish to express my gratitude to Fran Daniele and the staff of Preparé Inc. for their meticulous effort in the copyediting and production of the book; it was a pleasure to work with them.

I thank my Publisher Tom Robbins, and Associate Editor Alice Dworkin at Prentice Hall for their encouragement and help in preparing the manuscript for the book and all their behind-the-scene effort in the selection of the book cover and the production of the book.

I am indebted to Regina Bendig, Shawn Leslie, and Brigitte Maier, at the Science and Engineering Library, McMaster University, for their help in checking the accuracy of many of the references listed in the Bibliography.

Last but by no means least, I am deeply grateful to my Technical Coordinator Lola Brooks for typing different versions of the manuscript for the book, and always doing it in a cheerful and meticulous way.

Simon Haykin

Contents