Preface

I am indebted to the many readers and colleagues who have written to me on several occasions with encouraging feedback on my earlier textbook *Fundamentals of Adaptive Filtering* (Wiley, NJ, 2003). Their enthusiastic comments encouraged me to pursue this second project in an effort to create a revised version for teaching purposes. During this exercise, I decided to remove some advanced material and move select topics to the problems. I also opted to fundamentally restructure the entire text into eleven consecutive *parts* with each part consisting of a series of focused lectures and ending with bibliographic comments, problems, and computer projects. I believe this restructuring into a sequence of lectures will provide readers and instructors with more flexibility in designing and managing their courses. I also collected most background material on random variables and linear algebra into three chapters at the beginning of the book. Students and readers have found this material of independent interest in its own right. At the same time, I decided to maintain the same general style and features of the earlier publication in terms of presentation and exposition, motivation, problems, computer projects, summary, and bibliographic notes. These features have been well received by our readers.

AREA OF STUDY

Adaptive filtering is a topic of immense practical relevance and deep theoretical challenges that persist even to this date. There are several notable texts on the subject that describe many of the features that have marveled students and researchers over the years. In this textbook, we choose to step back and to take a broad look at the field. In so doing, we feel that we are able to bring forth, to the benefit of the reader, extensive commonalities that exist among different classes of adaptive algorithms and even among different filtering theories. We are also able to provide a uniform treatment of the subject in a manner that addresses some existing limitations, provides additional insights, and allows for extensions of current theory.

We do not have any illusions about the difficulties that arise in any attempt at understanding adaptive filters more fully. This is because adaptive filters are, by design, time-variant, nonlinear, and stochastic systems. Any one of these qualifications alone would have resulted in a formidable system to study. Put them together and you face an almost impossible task. It is no wonder then that current practice tends to study different adaptive schemes separately, with techniques and assumptions that are usually more suitable for one adaptation form over another. It is also no surprise that most treatments of adaptive filters, including the one adopted in this textbook, need to rely on some simplifying assumptions in order to make filter analysis and design a more tractable objective.

Still, in our view, three desirable features of any study of adaptive filters would be (1) to attempt to keep the number of simplifying assumptions to a minimum, (2) to delay their use until necessary, and (3) to apply similar assumptions uniformly across different classes of adaptive algorithms. This last feature enables us to evaluate and compare the performance of adaptive schemes under similar assumptions on the data, while delaying the use of assumptions enables us to extract the most information possible about actual filter performance. In our discussions in this book we pay particular attention to these three features throughout the presentation.

In addition, we share the conviction that a thorough understanding of the performance and limitations of adaptive filters requires a solid grasp of the fundamentals of least-mean-squares estimation theory. These fundamentals help the designer understand what it is that an adaptive filter is trying to accomplish and how well it performs in this regard. For this reason, Parts I (*Optimal Estimation*) and II (*Linear Estimation*) of the book are designed to provide the reader with a self-contained and easy-to-follow exposition of estimation theory, with a focus on topics that are relevant to the subject matter of the book. In these initial parts, special emphasis is placed on geometric interpretations of several fundamental results. The reader is advised to pay close attention to these interpretations since it will become clear, time and again, that cumbersome algebraic manipulations can often be simplified by recourse to geometric constructions. These constructions not only provide a more lasting appreciation for the results of the book, but they also expose the reader to powerful tools that can be useful in other contexts as well, other than adaptive filtering and estimation theory.

The reader is further advised to master the convenience of the vector notation, which is used extensively throughout this book. Besides allowing a compact exposition of ideas and a compact representation of results, the vector notation also allows us to exploit to great effect several important results from linear algebra and matrix theory and to capture, in elegant ways, many revealing characteristics of adaptive filters. We cannot emphasize strongly enough the importance of linear algebraic and matrix tools in our presentation, as well as the elegance that they bring to the subject. The combined power of the geometric point of view and the vector notation is perhaps best exemplified by our detailed treatment later in this book of least-squares theory and its algorithmic variants. Of course, the reader is exposed to geometric and vector formulations in the early chapters of the book.

STRUCTURE OF THE BOOK

The book is divided into eleven core parts, in addition to a leading part on *Background Material* and a trailing part on *References and Indices*. Table P.1 lists the various parts. Each of the core parts, numbered I through XI, consists of four distinctive elements in the following order: (i) a series of lectures where the concepts are introduced, (ii) a summary of all lectures combined, (iii) bibliographic commentary, and (iv) problems and computer projects.

Lectures and Concepts. In the early parts of the book, each concept is motivated from first principles; starting from the obvious and ending with the more advanced. We follow this route of presentation until the reader develops enough maturity in the field. As the book progresses, we expect the reader to become more sophisticated and, therefore, we cut back on the "obvious."

Summaries. For ease of reference, at the end of each part, we collect a summary of the key concepts and results introduced in the respective lectures.

Bibliographic Commentaries. In the remarks at the end of each part we provide a wealth of references on the main contributors to the results discussed in the respective lectures. Rather than scatter references throughout the lectures, we find it useful to collect all references at the end of the part in the form of a narrative. We believe that this way of presentation gives the reader a more focused perspective on how the references and the contributions relate to each other both in time and context.

Problems. The book contains a significant number of problems, some more challenging than others and some more applied than others. The problems should be viewed as an *integral* part of the text, especially since additional results appear in them. It is for this reason, and also for the benefit of the reader, that we have chosen to formulate and design most problems in a guided manner. Usually, and especially in the more challenging cases, a problem starts by stating its objective followed by a sequence of guided steps until the final answer is attained. In most cases, the answer to each step appears stated in the body of the problem. In this way, a reader would know what the answer should be, even if the reader fails to solve the problem. Thus rather than ask the reader to "find an expression for x,", we would generally ask instead to "show that x is given by $x = \ldots$" and then give the expression for x.

All instructors can request copies of a free solutions manual from the publisher.

Moreover, several problems in the book have been designed to introduce readers to useful topics from related fields, such as multi-antenna receivers, cyclic-prefixing, maximal ratio combining, OFDM receivers, and so forth. Students are usually surprised to learn how classical concepts and ideas form the underpinnings of seemingly advanced techniques.

Computer Projects. We have included several computer projects (see the listing in Table P.2) to show students, and also practitioners, how the results developed in the book can be useful in situations of practical interest (e.g., linear equalization, decision feedback equalization, channel estimation, beamforming, tracking fading channels, line echo cancellation, acoustic echo cancellation, active noise control, OFDM receivers). In designing these projects, we have made an effort at choosing topics that are relevant to practitioners. We have also made an effort at illustrating to students how a solid theoretical understanding can guide them in challenging situations. MATLAB[1] programs are available for solving all computer projects in the book, in addition to a solutions manual. The programs are offered without any guarantees. While we have found them to be effective for the instructional purposes of this textbook, the programs are not intended to be examples of full-blown or optimized designs; practitioners should use them at their own risk. For example, in order to keep the codes at a level that is easy to understand by students, we have often decided to sacrifice performance in lieu of simplicity.

> MATLAB programs that solve all computer projects in the book, in addition to a solutions manual for the projects with extensive commentary and typical performance plots, can be downloaded for free by all readers (including students and instructors) from either the publisher's website or the author's website.

Background Material. We provide three self-contained chapters that explain all the required background material on random variables and linear algebra for the purposes of this book. Actually, after progressing sufficiently enough in the book, students will be able to master many useful concepts from linear algebra and matrix theory, in addition to adaptive filtering.

COVERAGE AND TOPICS

The material in the book can be categorized into five broad areas of study (A through E), as listed in Table P.3. Area A covers the fundamentals of least-mean-squares estimation theory with several application examples. Areas B and C deal mainly with LMS-type adaptive filters, while areas D and E deal with least-squares-type adaptive filters. If an instructor wishes to focus mostly on LMS-type filters, then the instructor can do so by covering only material from within areas B and C. Even in this case, students will still be exposed to the recursive-least-squares (RLS) algorithm and its performance results from the discussions in Chapter 14 and Area C. However, for a more-in-depth treatment of RLS and its many variants, instructors will need to select chapters from within Area D as well.

DEPENDENCIES AMONG THE CORE PARTS

Figure P.1 illustrates the dependencies among the eleven core parts in the book. In the figure, the material in a part that is at the receiving end of an arrow requires some (but not necessarily all) of the material from the part at the origin of the arrow. A dashed arrow indicates that the dependency between the respective parts is weak and, if desired, the parts can be covered independently of each other. For example, in order to cover Part III (*Stochastic Gradient Methods*), the instructor would need to cover Part II (*Linear Estimation*). The material in Part I (*Optimal Estimation*) is not necessary for Part II (*Linear Estimation*) but it is useful for a better understanding of it. Figure P.1 can be

[1] MATLAB is a registered trademark of the MathWorks Inc., 24 Prime Park Way, Natick, MA 01760-1500, http://www.mathworks.com.

TABLE P.1 A breakdown of the book structure into eleven core parts.

Parts	Chapters
Background Material	A. Random Variables B. Linear Algebra C. Complex Gradients
I. Optimal Estimation	1. Scalar-Valued Data 2. Vector-Valued Data
II. Linear Estimation	3. Normal Equations 4. Orthogonality Principle 5. Linear Models 6. Constrained Estimation 7. Kalman Filter
III. Stochastic Gradient Methods	8. Steepest-Descent Technique 9. Transient Behavior 10. LMS Algorithm 11. Normalized LMS Algorithm 12. Other LMS-Type Algorithms 13. Affine Projection Algorithm 14. RLS Algorithm
IV. Mean-Square Performance	15. Energy Conservation 16. Performance of LMS 17. Performance of NLMS 18. Performance of Sign-Error LMS 19. Performance of RLS and Other Filters 20. Nonstationary Environments 21. Tracking Performance
V. Transient Performance	22. Weighted Energy Conservation 23. LMS with Gaussian Regressors 24. LMS with Non-Gaussian Regressors 25. Data-Normalized Filters
VI. Block Adaptive Filters	26. Transform-Domain Adaptive Filters 27. Efficient Block Convolution 28. Block and Subband Adaptive Filters
VII. Least-Squares Methods	29. Least-Squares Criterion 30. Recursive Least-Squares 31. Kalman Filtering and RLS 32. Order and Time-Update Relations
VIII. Array Algorithms	33. Norm and Angle Preservation 34. Unitary Transformations 35. QR and Inverse QR Algorithms
IX. Fast RLS Algorithms	36. Hyperbolic Rotations 37. Fast Array Algorithm 38. Regularized Prediction Problems 39. Fast Fixed-Order Filters
X. Lattice Filters	40. Three Basic Estimation Problems 41. Lattice Filter Algorithms 42. Error-Feedback Lattice Filters 43. Array Lattice Filters
XI. Robust Filters	44. Indefinite Least-Squares 45. Robust Adaptive Filters 46. Robustness Properties
References and Indices	

TABLE P.2 A listing of all computer projects in the book. MATLAB programs that solve these projects can be downloaded by all readers from the publisher's or author's websites, in addition to a solutions manual.

Computer project	Topic
I.1	Comparing optimal and suboptimal estimators
II.1	Linear equalization and decision devices
II.2	Beamforming.
II.3	Decision-feedback equalization
III.1	Constant-modulus criterion
III.2	Constant-modulus algorithm
III.3	Adaptive channel equalization
III.4	Blind adaptive equalization
IV.1	Line echo cancellation
IV.2	Tracking Rayleigh fading channels
V.1	Transient behavior of LMS
VI.1	Acoustic echo cancellation
VII.1	OFDM receiver
VII.2	Tracking Rayleigh fading channels
VIII.1	Performance of array implementations in finite precision
IX.1	Stability issues in fast least-squares
X.1	Performance of lattice filters in finite precision
XI.1	Active noise control

TABLE P.3 A breakdown of the book structure into five broad topic areas.

Category	Parts
A. Introduction and Foundations	Part I: Optimal Estimation Part II: Linear Estimation
B. Stochastic–Gradient Methods	Part III: Stochastic–Gradient Methods Part VI: Block Adaptive Filters
C. Performance Analyses	Part IV: Mean-Square Performance Part V: Transient Performance
D. Least-Squares Methods	Part VII: Least-Squares Methods Part VIII: Array Algorithms Part IX: Fast RLS Algorithms Part X: Lattice Filters
E. Indefinite Least–Squares	Part XI: Robust Filters

used by instructors to design different course sequences according to their needs and interests. For example, if the instructor is interested in covering only LMS-type adaptive filters and in studying their performance, then one possibility is to cover material from within Parts II, III, IV, and V.

AUDIENCE

The book is intended for a graduate-level course on adaptive filtering. Although it is beneficial that students have some familiarity with basic concepts from matrix theory, linear algebra, and random variables, the book includes three chapters on background material in these areas. The review is done in a motivated manner and is tailored to the needs of the presentation. From our experience,

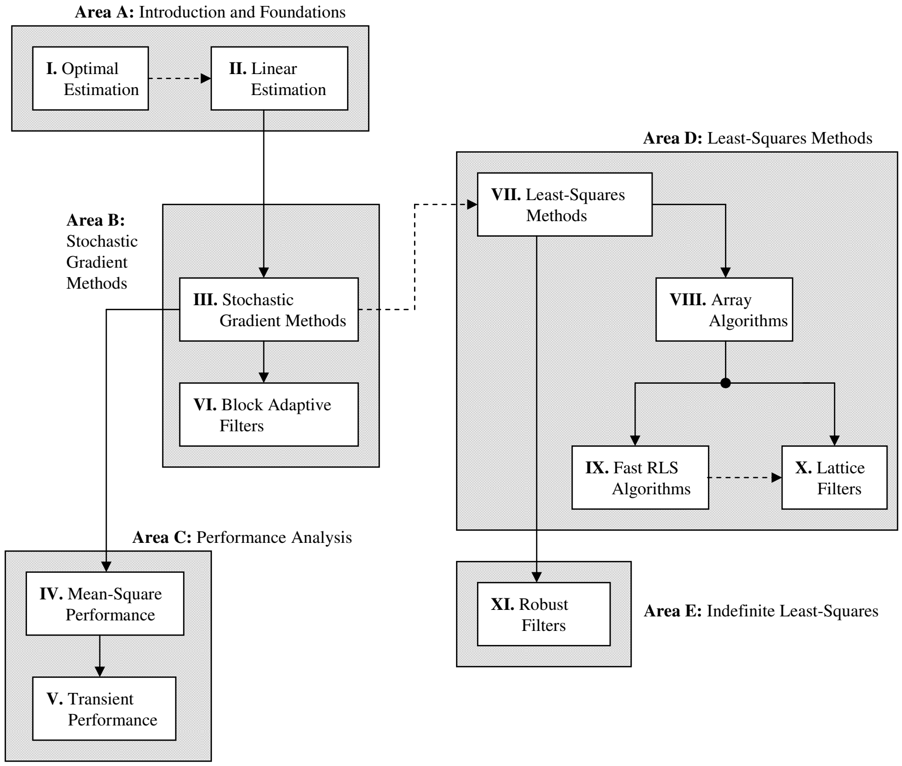

FIGURE P.1 Dependencies among the chapters. Instructors can design different course sequences in accordance with their needs and interests.

these reviews are sufficient for a thorough understanding of the discussions in the book. In addition, several of the problems reinforce the linear algebraic and matrix concepts, so much so that students will get valuable training in linear algebra and matrix theory, in addition to adaptive filtering, from reading (and understanding) this book.

The book is also intended to be a reference for researchers, which explains why we have chosen to include some advanced topics in a handful of places. As a result, the book contains ample material for instructors to design courses according to their interests. Clearly, we do not expect instructors to cover all the material in the book in a typical course offering; such an objective would be counter-productive. In our own teaching of the material, we instead *focus on some key sections and chapters and request that students complement the discussions by means of reading and problem solving*. As explained below, several key sections have been designed to convey the main concepts; while the remaining sections tend to include more advanced material and also illustrative examples. Once students understand the basic principles, you will be amazed at how well they can follow the other lectures on their own and even solve the pertinent problems.

To facilitate course planning, Table P.4 lists the key chapters or sections from the various core parts of the book for both lecturing and reading purposes. For example, Part V (*Transient Performance*) studies the transient behavior of a large family of adaptive filters in a uniform manner. The main idea is captured by the transient analysis of the LMS algorithm in Chapters 23 and 24; these chapters rely on the machinery developed in Chapter 22. Once students understand the framework as applied to LMS, they will be able to study the transient analysis of other filters on their own. This is one key advantage of adopting and emphasizing a uniform treatment of adaptive filter performance throughout our presentation. Similar remarks hold for the steady-state and tracking analyses of Part IV (*Mean-Square Performance*). It is sufficient to illustrate how the methodology applies to the special case of LMS, for example, by covering Chapters 15 and 16, as well as Sec. 21.1. Students would then do well in studying the extensions on their own if desired.

TABLE P.4 A suggested list of key chapters and sections.

Part	Key chapters for lecturing	Key chapters for reading
Part I	Chapters 1 and 2	
Part II	Chapters 3, 4	Sections 5.2, 5.3, 5.5, 6.3, 6.5
	Sections 5.1, 5.4, 6.1, 6.2, 6.4	Chapter 7
Part III	Chapters 8, 9, 10, 11, 14	Chapters 12, 13
Part IV	Chapters 15, 16, 20, 21.1	Chapters 17, 18, 19, 21.2–21.8
Part V	Chapters 22, 23, 24	Chapter 25
Part VI	Chapters 26, 27, 28	
Part VII	Chapters 29, 30	Chapters 31, 32
Part VIII	Chapters 33, 34, 35	
Part IX	Chapters 36, 37	Chapters 38, 39
Part X	Chapters 40, 41	Chapters 42, 43
Part XI	Chapters 44, 45	Chapter 46

SOME FEATURES OF OUR TREATMENT

There are some distinctive features in our treatment of adaptive filtering. Among other elements, experts will be able to notice the following contributions:

(a) We treat a large variety of adaptive algorithms.

(b) Parts IV and V study the mean-square performance of adaptive filters by resorting to energy-conservation arguments. While the performance of different adaptive filters is usually studied separately in the literature, the framework adopted in this book applies uniformly across different classes of adaptive filters. In addition, the same framework is used for steady-state analysis, transient analysis, tracking analysis, and robustness analysis (in Part XI).

(c) Part VI studies block adaptive filters, and the related class of subband adaptive filters, in a manner that clarifies the connections between these two families more directly than prior treatments. Our presentation also indicates how to move beyond DFT-based transforms and how to use other classes of orthogonal transforms for block adaptive filtering (as explained in Chapter 10 of Sayed (2003)).

(d) Parts VII-IX provide a detailed treatment of least-squares adaptive filters that is distinct from prevailing approaches in a handful of respects. First, we focus on regularized least-squares problems from the onset and take the regularization factor into account in all derivations. Second, we insist on deriving time- and order-update relations independent of any structure in the regression data (e.g., we do not require the regressors to arise from a tapped-delay-line implementation). In this way, one can pursue efficient least-squares filtering even for some non-FIR structures (as explained in Chapter 16 of Sayed (2003)). Third, we emphasize the role and benefits of array-based schemes. And, finally, we highlight the role of geometric constructions and the insights they bring into least-squares theory.

(e) Part XI develops the theory of robust adaptive filters by studying indefinite least-squares problems and by relying on energy arguments as well. In the process, the robustness and optimality properties of several adaptive filters are clarified. The presentation in this part is developed in a manner that parallels our treatment of least-squares problems in Chapters 29 and 30 so that readers can appreciate the similarities and distinctions between both theories (classical least-squares versus indefinite least-squares).