PREFACE TO SECOND EDITION

The second edition of *Bayesian Signal Processing* incorporates a chapter on "Sequential Bayesian Detection" (Chapter 10) and a section on "Ensemble Kalman Filters" (Section 6.7), as well as an expansion of case studies in the final chapter (Chapter 11). These new "physics-based" studies detail Bayesian approaches to problem solving in real-world applications incorporating detailed particle filter designs, adaptive particle filters, and sequential Bayesian detection. In addition to these major developments, a variety of sections are expanded to "fill in the gaps" of the first edition. Here, metrics for particle filter (PF) designs with emphasis on classical "sanity tests," introduced earlier in model-based processors, led to *ensemble techniques* as a basic requirement for performance analysis. Next, the expansion of *information theory metrics* (Kullback–Leibler divergence (KD), Hellinger distance (HD)) and their application to PF designs is discussed. These "fill-in-the-gap" expansions provide a more cohesive discussion with examples and applications enabling the comprehension of these alternative approaches to solving estimation/detection problems.

Detection theory, and more specifically sequential detection theory, is closely coupled to sequential estimation techniques presented in this text and is often the primary reason for constructing the estimators in the first place [1]–[14]. Sequential techniques find application in many technical application areas such as radar, sonar (detection/tracking), biomedical (anomaly detection/localization), speech (recognition/tracking), communications (real-time/obstructed environments), the sciences (e.g., seismology (earthquakes), structures (vibrations), materials (additive manufacturing/threat detection), radiation (threat detection, etc.), and of course, a huge variety of military applications [3], [7]. By incorporating a new chapter on sequential detection techniques primarily aimed at the binary decision problem, we enable the extension of these estimation methods to an entire class of problems especially when a physical model is available that can be incorporated into the algorithm [4], [6]. This new chapter, in itself, will provide wider application, since sequential detection is such a natural extension to sequential estimation and vice versa.

The ensemble Kalman Filter (EnKF) addition to the second edition is an area that has been neglected in most non-specialized texts. The EnKF is basically a little known hybrid in the engineering area, but well-known in the sciences. It is a hybrid of a regression-based processor (e.g., unscented Kalman filter (UKF)) and a particle-like (PF) "sampling" filter. The EnKF is well known in science areas where large-scale computations are required such as seismology, energy systems (wind, ocean waves, etc.), weather prediction, climatology (global warming), computational biology, large structures (vibrations), and more because of its computational efficiency for very large-scale computational problems (super-computer applications). Here, the coupling of model-based Bayesian techniques to these large-scale problems is unique.

With this in mind, let us consider the construct of the new chapter entitled "Sequential Bayesian Detection." Here, we develop the Bayesian approach to decision theory primarily aimed at a coupling of sequential Bayesian estimation to sequential decision-making. We start with the binary decision problem for multi-channel measurements and develop the usual Bayesian solutions based on probability-of-error minimization leading to the well-known Bayes' risk criterion. Next, the Neyman–Pearson detection approach (maximize detection probability for fixed false-alarm probability) is developed and compared to the classical Bayesian schemes illustrating their similarity and differences. Once these "batch schemes" are developed, we introduce the Wald sequential approach to solving these problems in pseudo real time [3], [7]. Once developed, we then investigate a variety of performance criteria based on the receiver operating characteristic (ROC) curve and its variants that provide the foundations for classical analysis [9], [10]. Other metrics (e.g., area-under-curve, and so on) associated with the ROC curve are introduced and applied as well. With the sequential detection theory developed, we investigate the basic linear Gaussian case and demonstrate that a sequential scheme easily follows when coupled to the model-based (Kalman) processor. Next, we generalize this approach to nonlinear models and again under Gaussian-like approximations develop the sequential detection scheme [7]. Finally, we remove the Gaussian assumptions and show how, using an MCMC (particle filter), sequential detection schemes can be developed and applied. A variety of applications are included in case studies on anomaly/change detection.

Finally, sequential detection enables the inclusion of more relevant case studies (Chapter 11) in ocean acoustics and physics-based radiation detection as well as X-ray threat material detection offering a completely different perspective on classical problem solving incorporating these physics-based approaches from the sequential Bayesian framework.

JAMES V. CANDY

Danville, CA

REFERENCES

1. S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory* (Englewood Cliffs, NJ: Prentice-Hall, 1998).
2. C. Therrien, *Decision, Estimation, and Classification: An Introduction to Pattern Recognition and Related Topics* (New York: John Wiley & Sons, 1989).
3. J. Melsa and D. Cohn, *Detection and Estimation Theory* (New York: McGraw-Hill, 1978).
4. M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Application* (Englewood Cliffs, NJ: Prentice-Hall, 1993).
5. F. Gustafasson, *Adaptive Filtering and Change Detection* (Hoboken, NJ: John Wiley & Sons, 2000).
6. K. Burnham and D. Anderson, *Model Selection and Multimodal Inference* (New York: Springer, 1998).
7. A. Sage and J. Melsa, *Estimation Theory with Applications to Communications and Control* (New York: McGraw-Hill, 1971).
8. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis* (Reading, MA: Addison-Wesley, 1990).
9. H. Van Trees, *Detection, Estimation and Modulation Theory, Part 1* (New York: John Wiley & Sons, 1968).
10. R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd Ed. (Hoboken, NJ: John Wiley & Sons, 2001).
11. A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis* (New York: Chapman & Hall, 2004).
12. J. Hancock and P. Wintz, *Signal Detection Theory* (New York: McGraw-Hill, 1966).
13. D. Middleton, *Introduction to Statistical Communication Theory* (New York: McGraw-Hill, 1960).
14. S. Kassam, *Signal Detection in Non-Gaussian Noise* (New York: Springer-Verlag, 1988).

PREFACE TO FIRST EDITION

In the real world, systems designed to extract signals from noisy measurements are plagued by errors evolving from constraints of the sensors employed, random disturbances and noise, and probably, most common, the lack of precise knowledge of the underlying physical phenomenology generating the process in the first place! Methods capable of extracting the desired signal from hostile environments require approaches that capture all of the a priori information available and incorporate them into a processing scheme. This approach is typically model-based [1], employing mathematical representations of the component processes involved. However, the actual implementation providing the algorithm evolves from the realm of statistical signal processing using a Bayesian approach based on Bayes' rule. Statistical signal processing is focused on the development of processors capable of extracting the desired information from noisy, uncertain measurement data. This is a text that develops the "Bayesian approach" to statistical signal processing for a variety of useful model sets. It features the next generation of processors which have recently been enabled with the advent of high-speed/high-throughput computers. The emphasis is on nonlinear/non-Gaussian problems, but classical techniques are included as special cases to enable the reader familiar with such methods to draw a parallel between the approaches. The common ground is the model sets. Here, the state–space approach is emphasized because of its inherent applicability to a wide variety of problems both linear and nonlinear as well as time invariant and time-varying problems including what has become popularly termed "physics-based" models. This text brings the reader from the classical methods of model-based signal processing including Kalman filtering for linear, linearized and approximate nonlinear processors as well as the recently developed unscented or sigma-point filters to the next generation of processors that will clearly dominate the future of model-based signal processing for years to come. It presents a unique viewpoint of signal processing from the Bayesian perspective in contrast to the pure statistical approach found in many textbooks. Although designed primarily as a graduate textbook, it will prove very useful to the practicing signal processing professional or scientist, since a wide variety of applications are included to demonstrate the applicability of the Bayesian approach to real-world problems. The prerequisites for such a text is a melding of undergraduate work in linear algebra, random processes, linear systems, and digital signal processing as well as a minimal background in model-based signal processing illustrated in the recent text [1]. It is unique in the sense that few texts cover the breadth of its topics, whereas, the underlying theme of this text is the Bayesian approach that is uniformly developed and followed throughout in the algorithms, examples, applications, and case studies. It is this theme coupled with the hierarchy of physics-based models developed that contribute to its uniqueness. This text has evolved from three previous texts, Candy [1–3] coupled with a wealth of practical applications to real-world Bayesian problems.

The Bayesian approach has existed in statistical physics for a long time and can be traced back to the 1940s with the evolution of the Manhattan project and the work of such prominent scientists as Ulam, von Neumann, Metropolis, Fermi, Feynman, and Teller. Here the idea of Monte Carlo (MC) techniques to solve complex integrals evolved [4]. Since its birth, Monte Carlo related methods have been the mainstay of many complex statistical computations. Many applications have evolved from this method in such areas as physics, biology, chemistry, computer science, economics/finance, material science, statistics and more recently in engineering. Thus, statisticians have known for a long time about these methods, but their practicalities have not really evolved as a working tool until the advent of high-speed super computers around the 1980s. In signal processing, it is hard to pinpoint the actual initial starting point but clearly the work of Handschin and Mayne in the late 1960s and early 1970s [5, 6] was the initial evolution of Monte Carlo techniques for signal processing and control. However, from the real-time perspective, it is probably the development of the sequential Bayesian processor made practical by the work of Gordon, Salmond, and Smith in 1993 [7] enabling the evolution and the explosion of the Bayesian sequential processor that is currently being researched today. To put this text in perspective, we must discuss the current signal processing texts available on Bayesian processing. Since its evolution much has been published in the statistical literature on Bayesian techniques for statistical estimation; however, the earliest texts are probably those of Harvey [8], Kitigawa and Gersch [9], and West [10] which emphasize the Bayesian model-based approach incorporating dynamic linear or nonlinear models into the processing scheme for additive Gaussian noise sources leading to the classical approximate (Kalman) filtering solutions. These works extend those results to nonGaussian problems using Monte Carlo techniques for eventual solution laying the foundation for works to follow. Statistical MC techniques were also available, but not as accessible to the signal processor due to statistical jargon and abstractness of the discussions. Many of these texts have evolved during the 1990s such as Gilks [11], Robert [12], Tanner [13], Tanizaki [14], with the more up-to-date expositions evolving in the late 1990s and currently such as Liu [4], Ruanaidh [15], Haykin [16], Doucet [17], Ristic [18], and Cappe [19]. Also during the last period a sequence of tutorials and special IEEE issues evolved exposing the MC methods to the signal processing community such as Godsill [20], Arulampalam [21], Djuric [22], Haykin [23], Doucet [24], Candy [25], as well as a wealth of signal processing papers (see references for details). Perhaps the most complete textbook from the statistical researcher's perspective is that of Cappe [19]. In this text, much of the statistical MC sampling theory is developed along with all of the detailed mathematics—ideal for an evolving researcher. But what about the entry level person—the engineer, the experimentalist, and the practitioner? This is what is lacking in all of this literature. Questions like, how do the MC methods relate to the usual approximate Kalman methods? How does one incorporate models (model-based methods) into a Bayesian processor? How does one judge performance compared with classical methods? These are all basically pragmatic questions that the proposed text will answer in a lucid manner through coupling the theory to real-world examples and applications. Thus, the goal of this text is to provide a bridge for the practitioners with enough theory and applications to provide the basic background to comprehend the Bayesian framework and enable the application of these powerful techniques to real-world problem solving. Next, let us discuss the structure of the proposed text in more detail to understand its composition and approach.

We first introduce the basic ideas and motivate the need for such processing while showing that they clearly represent the next generation of processors. We discuss potential application areas and motivate the requirement for such a generalization. That is, we discuss how the simulation-based approach to Bayesian processor design provides a much needed capability, while well known in the statistical community, not very well known (until recently) in the signal processing community. After introducing the basic concepts in Chapter 1, we begin with the basic Bayesian processors in Chapter 2. We start with the Bayesian "batch" processor and establish its construction by developing the fundamental mathematics required. Next we discuss the well-known maximum likelihood (ML) and minimum (error) variance (MV) or equivalently minimum mean-squared error (MMSE) processors. We illustrate the similarity and differences between the schemes. Next, we launch into sequential Bayesian processing schemes which forms the foundation of the text. By examining the "full" posterior distribution in both dynamic variables of interest as well as the full data set, we are able to construct the sequential Bayesian approach and focus on the usual *filtered* or *filtering* distribution case of highest interest demonstrating the fundamental prediction/update recursions inherent in the sequential Bayesian structure. Once establishing the general Bayesian sequential processor (BSP), the schemes that follow are detailed depending on the assumed distribution with a variety of model sets.

We briefly review simulation-based methods starting with sampling methods, progressing to Monte Carlo approaches leading to the basic iterative methods of sampling using the Metropolis, Metropolis-Hastings, Gibb's, and slice samplers. Since one of the major motivations of recursive or sequential Bayesian processing is to provide a real-time or pseudo real-time processor, we investigate the idea of *importance sampling* as well as *sequential importance sampling* techniques leading to the generic Bayesian sequential importance sampling algorithm. Here we show the solution can be applied, once the importance sampling distribution is defined.

In order to be useful, Bayesian processing techniques must be specified through a set of models that represent the underlying phenomenology driving the particular application. For example, in radar processing, we must investigate the propagation models, tracking models, geometric models, and so forth. In Chapter 4, we develop the state–space approach to signal modeling which forms the basis of many applications such as speech, radar, sonar, acoustics, geophysics, communications, control, etc. Here, we investigate continuous, sampled-data and discrete state–space signals and systems. We also discuss the underlying systems theory and extend the model-set to include the stochastic case with noise driving both process and measurements leading the well-known Gauss–Markov (GM) representation which forms the starting point for the *classical* Bayesian processors to follow. We also discuss the equivalence of the state–space model to a variety of time series (ARMA, AR, MA, etc.) representations as well as the common engineering model sets (transfer functions, all-pole, all-zero, pole-zero, etc.). This discussion clearly demonstrates why the state–space model with its inherent generality is capable of capturing the essence of a broad variety of signal processing representations. Finally, we extend these ideas to *nonlinear* state–space models leading to "approximate" Gauss-Markov representation evolving from nonlinear, perturbed and linearized systems.

In the next chapter, we develop *classical* Bayesian processors by first motivating the Bayesian approach to the state–space where the required conditional distributions use the embedded state–space representation. Starting with the linear, time-varying, state–space models, we show that the "optimum" classical Bayesian processor under multivariate Gaussian assumptions leads to minimum (error) variance (MV) or equivalently minimum mean-squared error (MMSE), which is the much heralded *Kalman filter* of control theory [1]. That is, simply substituting the underlying Gauss-Markov model into the required conditional distributions leads directly to the BSP or Kalman filter in this case. These results are then extended to the *nonlinear* state–space representation which are linearized using a known reference trajectory through perturbation theory and Taylor-series expansions. Starting with the linearized or approximate GM model of Chapter 4, we again calculate the required Bayesian sequential processor from the conditionals which lead to the "linearized" BSP (or linearized Kalman filter) algorithm. Once this processor is developed, it is shown that the "extended" Bayesian processor follows directly by linearizing about the most currently available estimate rather than the reference trajectory. The extended Bayesian processor (XBP) or equivalently extended Kalman filter (EKF) of nonlinear processing theory evolves quite naturally from the Bayesian perspective, again following the usual development by defining the required conditionals, making nonlinear approximations and developing the posterior distributions under multivariate Gaussian assumptions. Next, we briefly investigate an iterative version of the XBP processor, again from the Bayesian perspective which leads directly to the iterative version of the extended Bayesian processor (IX-BP) algorithm—an effective tool when nonlinear measurements dominate the uncertain measurements required.

Chapter 6 focuses on *statistical linearization* methods leading to the *modern* unscented Bayesian processor (UBP) or equivalently sigma-point Bayesian processor (SPBP). Here we show how statistical linearization techniques can be used to transform the underlying probability distribution using the sigma-point or unscented nonlinear transformation technique (linear regression) leading to the *unscented* Bayesian processor or equivalently the unscented Kalman filter (UKF). Besides developing the fundamental theory and algorithm, we demonstrate its performance on a variety of example problems. We also briefly discuss the Gaussian–Hermite quadrature (G-H) and Gaussian sum (G-S) techniques for completeness.

We reach the heart of the *particle filtering* methods in Chapter 7, where we discuss the Bayesian approach to the state–space. Here the ideas of Bayesian and model-based processors are combined through the development of Bayesian state–space particle filters. Initially, it is shown how the state–space models of Chapter 4 are incorporated into the conditional probability distributions required to construct the sequential Bayesian processors through importance sampling constructs. After investigating a variety of importance proposal distributions, the basic set of state-space particle filters (SSPF) are developed and illustrated through a set of example problems and simulations. The techniques including the Bootstrap, auxiliary, regularized MCMC and linearized particle filters are developed and investigated when applied to the set of example problems used to evaluate algorithm performance.

In Chapter 8, the important joint Bayesian SSPF are investigated by first developing the joint filter popularly known as the parametrically adaptive processor [1]. Here both states and static as well as dynamic parameters are developed as solutions to this joint estimation problem. The performance of these processors are compared to classical and modern processors through example problems.

In Chapter 9, the hidden Markov models (HMM) are developed for event-related problems (e.g., Poisson point processes). This chapter is important in order to place purely discrete processes into perspective. HMM evolve for any type of memoryless, counting processes and become important in financial applications, communications, biometrics, as well as radiation detection. Here we briefly develop the fundamental ideas and discuss them in depth to develop a set of techniques used by the practitioner while applying them to engineering problems of interest.

In the final chapter, we investigate a set of physics-based applications focusing on the Bayesian approach to solving real-world problems. By progressing through a step-by-step development of the processors, we see explicitly how to develop and analyze the performance of such Bayesian processors. We start with a practical laser alignment problem followed by a broadband estimation problem in ocean acoustics. Next, the solid-state microelectromechanical (MEM) sensor problem for biothreat detection is investigated followed by a discrete radiation detection problem based on counting statistics. All of these methods invoke Bayesian techniques to solve the particular problems of interest enabling the practitioner the opportunity to track "real-world" Bayesian model-based solutions.

The place of such a text in the signal processing textbook community can best be explained by tracing the technical ingredients that comprise its contents. It can be argued that it evolves from the digital signal processing area primarily from those texts that deal with random or statistical signal processing or possibly more succinctly "signals contaminated with noise." The texts by Kay [26–28], Therrien [29], and Brown [30] all provide the basic background information in much more detail than this text, so there is little overlap at the detailed level with them.

This text, however, possesses enough theory for the graduate or advanced graduate student to develop a fundamental basis to go onto more rigorous texts like Jazwinski [31], Sage [32], Gelb [33], Anderson [34], Maybeck [35], Bozic [36], Kailath [37, 38], and more recently, Mendel [39], Grewel [40], Bar-Shalom [41], and Simon [42]. These texts are rigorous and tend to focus on Kalman filtering techniques ranging from continuous to discrete with a wealth of detail on all of their variations. The Bayesian approach discussed in this text certainly includes the state–space models as one of its model classes (probably the most versatile), but the emphasis is on various classes of models and how they may be used to solve a wide variety of signal processing problems. Some of the more recent texts about the same technical level, but again, with a different focus are Widrow [43], Orfanidis [44], Sharf [45], Haykin [46], Hayes [47], Brown [30], and Stoica [48]. Again, the focus of these texts is not the Bayesian approach but the narrow set of specific models and the development of a variety of algorithms to estimate these sets. The system identification literature and texts therein also provide some overlap with this text, but again the approach is focused on estimating a model from noisy data sets and not really aimed at developing a Bayesian solution to a particular signal processing problem. The texts in this area are Ljung [49, 50], Goodwin [51], Norton [52], and Soderstrom [53].

The recent particle filtering texts of Ristic [18] and Cappe [19] are useful as references to accompany this text, especially if more details are required on the tracking problem and the fundamental theorems governing statistical properties and convergence proofs. That is, Ristic's text provides an introduction that closely follows the 2002 tutorial paper by Arulampalam [21] but provides little of the foundational material necessary to comprehend this approach. It focuses primarily on the tracking problem. Cappe's text is at a much more detailed technical level and is written for researchers in this area not specifically aimed at the practitioner's viewpoint. The proposed text combines the foundational material, some theory along with the practice and application of PF to real-world applications and examples.

The approach we take is to introduce the basic idea of Bayesian signal processing and show where it fits in terms of signal processing. It is argued that BSP is a natural way to solve basic processing problems. The more a priori information we know about data and its evolution, the more information we can incorporate into the processor in the form of mathematical models to improve its overall performance. This is the theme and structure that echoes throughout the text. Current applications (e.g., structures, tracking, equalization, biomedical) and simple examples are incorporated to motivate the signal processor. Examples are discussed to motivate all of the models and prepare the reader for further developments in subsequent chapters. In each case, the processor, along with accompanying simulations, is discussed and applied to various data sets demonstrating the applicability and power of the Bayesian approach. The proposed text is linked to the MATLAB (signal processing standard software) software package providing notes at the end of each chapter.

In summary, this Bayesian signal processing text will provide a much needed "down-to-earth" exposition of modern MC techniques. It is coupled with well-known signal processing model sets along with examples and problems that can be used to solve many real-world problems by practicing engineers and scientists along with entry-level graduate students as well as advanced undergraduates and post-doctorates requiring a solid introduction to the "next generation" of model-based signal processing techniques.

JAMES V. CANDY

Danville, CA