Highlights of Statistical Signal and Array Processing

The Statistical Signal and Array Processing Technical Committee

Edited by Alfred Hero

Unlike most other technical committees of the Signal Processing Society, which deal with signals of deterministic nature and process signals one at a time, the Statistical Signal and Array Processing (SSAP) Technical Committee deals with signals that are random and processes an array of signals simultaneously. This issue features the SSAP-TC's contribution to the Anniversary series, which covers this special field of random signals and array processing.

The field of SSAP represents both solid theory and practical applications. Starting with research in spectrum estimation and statistical modeling, study in this field is always full of elegant mathematical tools such as statistical analysis and matrix theory. The area of statistical signal processing expands into estimation and detection algorithms, time-frequency domain analysis, system identification, and channel modeling and equalization. The area of array signal processing also extends into multichannel filtering, source localization and separation, and so on. Work in SSAP areas has already made an impact in a large variety of applications, ranging from communication and radar/sonar processing to many medical imaging technologies, and even econometrics. This article represents an endeavor by the members of the SSAP-TC to review all these significant developments in the field of SSAP. To provide readers with pointers for further study of the field, this article includes a very impressive bibliography—close to 500 references are cited. This is just one of the indications that the field of statistical signals has been an extremely active one in the signal-processing community.

This article also introduces the recent reorganization of three technical committees of the Signal Processing Society. During the reorganization, the SSAP, Digital Signal Processing, and Underwater Acoustics Signal Processing technical committees were restructured to form three new committees: Signal Processing Theory and Methods, Signal Processing for Communications, and Sensor Arrays and Multichannel Signal Processing. After the reorganization, research topics that used to belong to the SSAP TC are now distributed to the three new TCs. Therefore, although the name "SSAP" does not exist anymore, the research activities related to it have been given a new life and will continue to thrive in the Signal Processing Society.

Now, let me invite you to enjoy this article, which will give you a quick but comprehensive tour of the highlights of statistical signal and array processing.

Tsuhan Chen, Guest Editor
Carnegie Mellon University

1053-5888/98/$10.00©1998IEEE

Many engineering applications require extraction of a signal or parameter of interest from degraded measurements. To accomplish this it is often useful to deploy fine-grained statistical models; diverse sensors that acquire extra spatial, temporal, or polarization information; or multidimensional signal representations, e.g., time-frequency or time scale. When applied in combination these approaches can be used to develop highly sensitive signal estimation, detection, or tracking algorithms that can exploit small but persistent differences between signals, interferences, and noise. Conversely, these approaches can be used to develop algorithms to identify a channel or system producing a signal in additive noise and interference, even when the channel input is unknown but has known statistical properties.

Broadly stated, the statistical signal and array processing (SSAP) area is concerned with reliable estimation, detection, and classification of signals that are subject to random fluctuations. Opening a recent issue of the *IEEE Transactions on Signal Processing* to a SSAP paper the reader will probably see one or more of the following: (1) description of a mathematical and statistical model for measured data, including models for sensor, signal, and noise; (2) careful statistical analysis of the fundamental limitations of the data including deriving benchmarks on performance, e.g., the Cramer-Rao, Ziv-Zakai, Barankin, Rate Distortion, Chernov, or other lower bounds on average estimator/detector error; (3) development of mathematically optimal or suboptimal estimation/detection algorithms; (4) asymptotic analysis of error performance establishing that the proposed algorithm comes close to reaching a benchmark derived in (2); and (5) simulations or experiments that compare algorithm performance to the lower bound and to other competing algorithms. Depending on the specific application, a SSAP algorithm may also have to be adaptive to changing signal and noise environments. This requires incorporating flexible statistical models, implementing low-complexity real-time estimation and filtering algorithms, and on-line performance monitoring.

Until recently the statistical signal and array processing area was covered by the SSAP Technical Committee, which grew out of the Spectrum Estimation and Modeling Technical Committee (discontinued in 1991). At ICASSP-98 in Seattle, an administrative restructuring took place that eliminated the SSAP, Digital Signal Processing (DSP), and Underwater Acoustics Signal Processing (UASP) Technical Committees, replacing them by three new committees: Signal Processing Theory and Methods (SPTM), Signal Processing for Communications (SPCOM), and Sensor Arrays and Multichannel signal processing (SAM). The SSAP areas described in this article have migrated to these new Technical Committees and remain very active within the Signal Processing Society. In particular, the following workshops sponsored or co-sponsored by SSAP will continue to provide forums for researchers in the area: the Workshop on Higher Order Statistics (to be held in Caesaria, Israel, in 1999 (http://sig.enst.fr/~hos99)), the Workshop on Statistical Signal and Array Processing (to be held in the Poconos, Pennsylvania, in 2000), and the Workshop on Signal Processing Advances in Communications (to be held in Annapolis, Maryland, in 2000).

Similar to other Technical Committees, SSAP ran workshops, recommended paper awards, and reviewed papers for ICASSP. To facilitate the paper review process and provide focus for award nominations, the scope of SSAP was divided into several subareas, called "SP EDICS" categories. These categories were spectral analysis; higher-order statistical analysis; cyclostationary signal analysis; statistical multichannel filtering; statistical modeling; parameter estimation; detection; performance analysis; system identification; computational algorithms; and applications. These categories are covered in this article and continue to be represented in the aggregated EDICS of the SPTM, SPCOM, and SAM Technical Committees.

SSAP TC Members (as of May 1998)

Alfred O. Hero III (SPTM), *University of Michigan, Ann Arbor (chairman)*
Georgios B. Giannakis (SPCOM), *University of Virginia (vice-chairman)*
Moeness Amin (SPTM), *Villanova University*
Kevin Buckley, *Villanova University*
Jean-François Cardoso (SPTM), *École Nationale Supérieure des Télécommunications*
Zhi Ding (SPCOM), *Auburn University*
Petar M. Djuric (SPTM), *State University of New York at Stony Brook*
Hamid Krim (SPTM), *North Carolina State University*
Jeffrey Krolik (SAM), *Duke University*
Fu Li, *Portland State University*
Hagit Messer-Yaron (SPTM), *Tel-Aviv University*
Eric Moulines (SPTM), *École Nationale Supérieure des Télécommunications*
Arye Nehorai (SAM), *Univ. of Illinois at Chicago*
Louis Scharf (SPTM), *University of Colorado, Boulder*
Ananthram Swami (SPCOM), *Army Research Laboratory*
A. Lee Swindlehurst (SPCOM), *Brigham Young University*
David J. Thomson (SPTM), *Lucent Technologies*
Jitendra Tugnait (SPCOM), *Auburn University*
Raghuveer M. Rao, *Rochester Inst. of Technology*
Lang Tong (SPCOM), *Cornell University*
Mats Viberg (current Vice-Chair SPTM), *Chalmers University of Technology*
Mati Wax, *Rafael, Israel* (SAM)
Guanghan Xu (SPCOM), *The University of Texas at Austin*
Michael Zoltowski (SPCOM), *Purdue University*

(new TC affiliations based on restructuring are indicated in parenthesis)

As the reader will see from this article, SSAP impacts a very wide range of applications. Among the applications mentioned in the sequel are: radar signal processing; sonar signal processing; geophysics and climate; radar and optical remote sensing; electrocardiography (ECG); electroencephalography (EEG); magnetoencephalography (MEG); nuclear magnetic resonance (NMR) imaging; radio-isotope imaging (PET and SPECT); chemical sensing of the environment; physical oceanography; fractal internet traffic modeling; astronomy; biology; econometrics; speech; and music analysis/synthesis.

Over the past several years the application of signal processing to communications has become a prevalent theme in SSAP. The pre-existence of many relevant core SSAP areas made communications a very ripe applications area. In particular, research in cyclostationarity, higher-order statistics, and system identification was a springboard to the development of novel methods for channel equalization in digital communications. Likewise, work in detection and estimation led naturally to iterative multiuser detection, source separation, and high-performance modulation classification algorithms. As another example, deployment of phased antenna arrays and the associated signal processing has spearheaded much recent activity in spatial diversity reception for wireless communications. The sections by Giannakis, Tong, and others highlight some of these communications applications of SSAP.

Our article begins with a group of two sections on recent developments in detection/estimation algorithms written by Alfred Hero and Petar Djuric, respectively. The section by Hero focuses on two areas of significant activity: constant-false-alarm-rate (CFAR) detection and iterative maximum-likelihood (ML) parameter estimation using the expectation-maximization (EM) algorithm. The section by Djuric describes the emerging area of Bayesian signal processing including estimation, detection, tracking and Monte Carlo Markov chain (MCMC) sampling, which is a technique that was largely impractical before the current generation of high-speed computers.

The article continues with a section on time-delay estimation written by Hagit Messer and Jason Goldberg and a section on multiwindow spectral estimation by David Thomson. From a historical perspective, time-delay estimation and spectral estimation are two of the oldest areas of statistical signal processing, dating back at least to the late 19th century (see [42]), yet they remain two of the most active areas today. Continuing along these lines are sections on the increasingly important problems of detection and estimation in the time-frequency domain, written by Moeness Amin, and the time-scale or multiresolution domain, written by Hamid Krim and Jean-Christophe Pesquet.

Next comes a section written by Georgios Giannakis on recent SSAP activity in channel estimation and equalization for digital communications. This is followed by two sections dealing with the critical problems of modeling, system identification, and the often overlooked area of data validation. Ananthram Swami starts off with a broad overview of non-Gaussian measurement models and higher-order statistical methods, followed by a section by Jitendra Tugnait on advances in multichannel system identification and testing random processes for non-Gaussian or nonlinear behavior. These are followed by a section written by Arye Nehorai on exciting opportunities in SSAP due to recent advances in sensor technology.

The fundamental theory behind detection, classification, and estimation has its home in mathematical statistics and decision theory.

Finally, the article turns to array signal processing with four sections written by Lee Swindlehurst, Jeff Krolik, Jean-François Cardoso, and Lang Tong, respectively. Swindlehurst provides a bird's-eye view of sensor-array processing and its applications to source localization, source separation, and channel estimation. Cardoso follows up with a section focusing on developments in blind-source-separation algorithms. Tong discusses the increasing importance of blind-source separation and diversity in multiuser communications systems design. The final section, written by Krolik, discusses the use of computational propagation models for processing sonar and radar array data.

It is essential to point out that, in a limited overview article such as this, one cannot possibly do justice to the large number of areas that comprise SSAP. Neither can we hope to cover but a fraction of the contributions of individuals who have had a role in the development of SSAP through the years. We offer our sincere apologies to any individuals who feel omitted from this overview.

WWW links relevant to the area of SSAP:

▲ The (old) SSAP home page:
http://www.eng.auburn.edu/~ding/SSAP/SSAP.html

▲ A database of "selected papers" that appeared in the *IEEE Transactions on Signal Processing* 1988-1995:
http://www.eng.auburn.edu/~ding/SSAP/Intp.html

▲ The SPTM, SPCOM and SAM Technical Committee home pages can be accessed through the IEEE Signal Processing Society home page:
http://www.ieee.org/society/sp/index.html

▲ A clearinghouse for information on many aspects of signal processing is the Signal Processing Information Base at:
http://spib.rice.edu/spib.html

▲ Some other web pages of interest to those working in SSAP:

—The IEEE Societies on Computers, Antennas and Propagation, Communications, Aerospace and Electronic Systems, Information Theory, and the IEEE Neural Network Council, all have SSAP related activities and links can be found on the IEEE page:

http://www.ieee.org/tab/cur_sub_soc_sub_hps.html

—The American Statistical Association:

http://www.amstat.org/

—The Institute of Mathematical Statistics:

http://www.imstat.org/

—The International Association for Statistical Computing:

http://www.stat.unipg.it/iasc.html

—The Royal Statistical Society:

http://maths.ntu.ac.uk/rss/index2.html

—The Acoustical Society of America:

http://asa.aip.org/

—The International Union of Radio Science:

http://www.intec.rug.ac.be/Research/Projects/ursi/welcome.html

—The Institution of Electrical Engineers (UK):

http://www.iee.org.uk/Welcome.html

Advances in Detection and Estimation Algorithms for Signal Processing

Alfred Hero, University of Michigan

The fundamental theory behind detection, classification, and estimation has its home in mathematical statistics and decision theory [109, 227]. In the context of statistical signal processing one must also contend with additional constraints: the exceedingly large size of signal-processing datasets; the absence of reliable and tractable signal models; the associated requirement of fast algorithms; and the requirement of real-time unsupervised algorithms. Two statistical signal-processing areas will be discussed in this section: algorithms for robust CFAR detection, and advances in iterative parameter estimation using the EM algorithm.

CFAR Detection

One of the most challenging problems in automated target detection and recognition is reliable detection of targets in high clutter backgrounds. When the clutter statistics are unknown or highly variable, the false-alarm rate of classical detection algorithms, e.g., the matched filter, cannot be controlled and target detection decisions become unreliable. The reason for this is lack of robustness of the test statistics to clutter variations.

The objective of CFAR detection is to produce a test statistic whose probability distribution does not depend on the unknown noise parameters, e.g., noise power or clutter spectrum, while ensuring a high probability of signal detection. Such a detector is also sometimes referred to as a noise-adaptive detector. For such a test statistic the detection threshold can be set to guarantee a prespecified false-alarm rate. There are a wide range of different strategies available for designing CFAR detectors including: min-max hypothesis testing [109], similar and unbiased hypothesis testing [227], invariant hypothesis testing [287], and the generalized likelihood ratio (GLR) test [205]. For lack of space we focus only on CFAR detection using the min-max, GLR, and invariant testing approaches. We regretfully must omit work in adaptive detection for assumed known noise backgrounds, nonparametric techniques, distributed detection, Huber robust detection, sequential detection, signal classification, and detection of number of signals.

Min-max CFAR hypothesis testing seeks to maximize detection probability subject to a constraint on maximum false-alarm rate. The min-max approach was recently adopted in [20] and [21] in the context of simultaneous detection and classification of multiple signals. This produced optimal detectors that took the form of a weighted likelihood ratio (LR) test. It was also shown in [20] that this min-max CFAR test implicitly implements a variant of Rissanen's maximum data length (MDL) signal selection criterion, establishing that MDL is min-max optimal. It is sometimes possible to arrive at min-max optimal detectors through the method of similar tests [366]. Finally, the min-max CFAR optimal detector can be viewed from the point of view of Bayesian detection implemented with a least favorable prior on the unknown noise density. Thus, in principle, the Bayesian methods developed in [119], [30], and more recently in [61], can be manipulated to provide CFAR tests.

In many cases direct min-max optimization is difficult, and simpler suboptimal CFAR alternatives are of interest. The conceptually simplest approach is the GLR "estimate and plug" procedure, which requires computing ML estimates for the unknown noise parameters. In [204] the GLR principle produced an adaptive detector for detecting spatio-temporal signals or targets in Gaussian noise with unknown spatial covariance. A different GLR adaptive target detector was derived in [57] for the case of optical images. The GLR for a general multichannel measurement was derived in [205], which specializes to the cases derived in [204] and [57] by applying suitable coordinate transformations. A related and important result was presented in [333] where exact confidence regions for the GLR-maximizing signal vector were derived for unknown spatial covariance. Additional applications of the GLR strategy to multispectral infrared images were presented in [330] and [464]. In [48] the GLR test was applied to arbitrary subspace projections of the data under similar assumptions as [205].

Other notable CFAR applications of GLR have appeared in the following areas: signal detection in noise of slowly fluctuating power [100]; transient signal detection in Gaussian noise of unknown power [316]; signal detection in unknown Gaussian-Gaussian mixture noise [39]; colored autoregressive noise [202, 367]; spatio-temporal signal detection in Gaussian noise with unknown spatial covariance [120, 266, 341]; signal detection in unknown impulsive noise [59]; multiwindow/GLR sinusoid detection [187, 296]; tests for presence of cyclostationary signals [82]; and detection of sampled signals having sampling jitter [371].

When the GLR is intractable, e.g., for non-Gaussian signals, and the noise covariance is known up to a scale factor, CFAR tests have been proposed based on maximizing alternative criteria such as: deflection or contrast [102, 310]; circular correlation coefficient for sinusoid detection [300]; and coherence [125] and generalized coherence [68] for multichannel signal detection. CFAR detectors have also been derived based on summary statistics such as: order statistical filter outputs [452], higher-order spectra [167, 209]; matched-filter multiple-correlation lag products [134]; weighted subspace fitting residuals [444]; and adaptive filtering followed by subspace projections [228]. Finally, when the noise covariance matrix is unknown, CFAR detection has been proposed using maximum signal-to-noise ratio (SNR) criteria and covariance estimates [101]; group delay statistics [231]; integrated-bispectrum non-Gaussianity tests [428] (corrections in [429]); and higher-order cumulants [140, 345, 346].

One of the main justifications of the GLR principle is its asymptotic optimality under broad conditions, e.g., [200, 201]. However, there are two factors that can make the GLR test unworkable in applications: 1) the GLR may not be of closed form when the clutter covariance has special structure, e.g., block diagonal; 2) use of the GLR principle entails a loss in efficiency [206, 331] which can severely impact finite sample performance. An alternative that can frequently lead to better finite sample performance is the application of the principle of invariance [355], also called exact robustness [192].

The method of invariance involves expressing uncertainty in the unknown clutter covariance as resulting from set of algebraic actions on the image by an appropriate group of transformations. Once the uncertainty has been mapped to group actions, one can often identify statistics whose statistical distributions are functionally invariant to unknown noise parameters yet entail minimum loss of target discrimination capability. On the basis of these statistics, optimal CFAR likelihood ratio tests can often be specified. These statistics are called "maximal invariants," and the resultant LR tests are called CFAR invariant tests. Many simple examples exist for which invariance principles give CFAR tests with higher power than the GLR test (for a simple but nontrivial example see [227, Ex. 6.18]). Despite the difficulty in finding maximal invariants and their statistical distributions the payoff for the extra effort in signal-processing applications can be high [35, 36, 354, 357, 356], where often the invariant LR test significantly outperforms the GLR or approximate GLR test.

The EM Algorithm for Parametric Estimation

The EM algorithm has generated much recent interest in the signal-processing community due to its ability to reliably compute iterative ML and penalized-ML estimates of signal parameters for cases where direct maximization is intractable. While the origins of the algorithm are decades older, it was only after the unified overview by Dempster, Laird, and Rubin (DLR) [89] that the wide applicability of the EM algorithm became recognized. Twenty years after DLR published their paper Meng and Van Dyk [265] published an update which cogently described the considerable recent advances in the EM algorithm. A review of some signal-processing applications of EM recently appeared in this magazine [273]. Here we will focus on important developments that were not covered in [273].

The intuition behind EM is simple to state. Based on an observation, $\mathbf{y}$, it is desired to maximize the log-likelihood $l_{\mathbf{y}}(\theta) = \ln f_{\mathbf{y}}(y;\theta)$, over an unknown parameter, θ. However, either due to missing data or to a complicated form of the log-likelihood, one would much rather maximize the simpler log-likelihood, $l_X(\theta) = \ln f_X(x;\theta)$, of a more informative data sample X, called the "complete data." As the complete data, X, is not available one strikes a compromise by iteratively maximizing the best estimate of the simpler log-likelihood given $\mathbf{y}$ and the previous estimate of θ. Here the "best estimate" is the one that minimizes mean-squared error: the conditional mean. Remarkably, this simple recipe leads to an algorithm that has many attractive properties such as stable convergence, monotone increasing likelihood, and flexible implementation.

One of the first signal-processing applications of EM after DLR appeared was to the problem of emission tomography [372] and shortly thereafter to transmission tomography [221]. Many follow-up papers appeared on this topic in the medical-imaging community (see, e.g. [110] for a partial list) before the EM algorithm was applied to other signal-processing problems such as parameter estimation for multiple superimposed signals [106, 107] and direction finding [271]. Wide adoption of the EM algorithm was hindered by its disappointingly slow convergence speed. Efforts to improve the convergence of EM include: Aitken's acceleration [247]; over-relaxation [229], conjugate gradient [180, 196]; Newton methods [38, 262]; quasi-Newton methods [220]; and ordered subsets EM (OSEM) [178]. Unfortunately, these methods do not automatically guarantee the monotone increasing likelihood property of standard EM, leading to the additional burden of monitoring the iterations for instability [222].

It has been established that the EM algorithm converges for bounded unimodal loglikelihood $l_{\mathbf{y}}(\theta) = \ln f_{\mathbf{y}}(y;\theta)$ [461]. When the likelihood function is twice differentiable the asymptotic speed of convergence of the EM algorithm is proportional to the maximum eigenvalue of $[F_X - F_{\mathbf{y}}]F_X^{-1}$ [89]. Here $F_{\mathbf{y}} = -\nabla^2 l_{\mathbf{y}}(\hat{\theta})$ and $F_X = E_{\hat{\theta}}[-\nabla^2 l_X(\hat{\theta}) \mid \mathbf{y}, \hat{\theta}]$ are the observed Fisher information matrices evaluated at the ML estimate $\hat{\theta}$. Furthermore, in [165] the *monotonic rate of convergence* of EM was shown to be equal to the matrix l_2 norm

In the statistical community the use of priors has been a controversial subject for many years.

$\|F_X^{-1/2}[F_X - F_y]F_X^{-1/2}\|$. Thus, the speed of convergence of the EM algorithm increases as the complete data, X, become less informative; i.e., as F_X approaches F_y and X gets closer to the actual measurements, y. However, there is a tradeoff between speed of convergence and implementation complexity: the M step of the standard EM algorithm usually becomes more difficult as X becomes less informative. It was discovered by Fessler and Hero [110] that this tradeoff can be eased by reformulation of the EM algorithm with "hidden data" sets, which are less informative than complete data sets and which can vary at each iteration. This led to the "space alternating expectation maximization" (SAGE) algorithm, its main feature being that it only updates small groups of the parameters, e.g., a single coordinate, at each iteration yet preserves monotonicity. In this respect, SAGE resembles the "expectation conditional maximization either" (ECME) [239] but SAGE generally has faster convergence [265].

In Meng and Van Dyk's paper [265] a generalization of SAGE and ECME was introduced called "alternating expectation conditional maximization" (AECM), which is a SAGE algorithm with a "design parameter" similar to SAGE-3 introduced in [111]. Another recent generalization, called parameter expansion EM (PX-EM), allows one to augment the parameter space, in addition to the data space, in order to obtain further convergence acceleration [240]. Interestingly, for the case of the superimposed signals problem PX-EM reduces to SAGE. In addition to the examples shown in [110], [111], and [265] the SAGE algorithm and its variants have been applied to angle of arrival estimation [115], multiuser detection [81, 289], estimation of constrained covariance matrices [364], and speckle interferometry [363].

For cases where the likelihood function is nonconvex neither the EM algorithm nor its variants are guaranteed to converge. Furthermore, grouped coordinate updating techniques like ECME, SAGE, and AECM may not converge even when the likelihood is convex but is nondifferentiable. In these cases it is still possible to improve on the plain-vanilla EM algorithm. Two recent advances are the method of Lavielle [51, 225] and the method of Chretien and Hero [63, 64]. The former allows implementation of the E step via stochastic approximation (or simulated annealing [226]) when a closed form for the E step is not available. This permits much less informative complete data sets to be used, for which the conditional expectation in the E step is intractable, thereby improving the asymptotic convergence speed. Furthermore, under appropriate conditions on the annealing schedule, this method guarantees (w.p.1) convergence to the global maximum for any initialization. On the other hand, the deterministic method of Chretien and Hero [63] was developed for nondifferentiable nonconvex likelihood functions and iteratively approximates the ML estimate via cutting plane methods, specifically proximal point iterations with Kullback-Liebler penalty. When the relaxation parameter in the proximal point algorithm is equal to one, this method reduces to the standard EM algorithm. By decreasing the relaxation parameter towards zero, a more rapidly convergent algorithm is obtained, all the while preserving the monotonic likelihood property.

A WWW link to the author of the above section:
http://www.eecs.umich.edu/~hero/hero.html

Bayesian Methods for Signal Processing

Petar M. Djurić, State University of New York at Stony Brook

A typical signal-processing task is to extract desired information about a signal from observed data. The sought information might be, for example, related to unknown signal parameters, number of signals in the data, distribution of signal power as a function of time and frequency, hidden states of a system producing a signal, or prediction of system and signal behavior. The Bayesian approach to making the required inference relies on the use of probability models for the observed data and the application of probability theory, where a key role is played by Bayes' theorem [28, 41, 123, 198, 339, 355, 379]. The inference is made in terms of probability statements, and the methodology, overall, is coherent and on conceptually sound and indisputable grounds. Its framework has considerable practical advantages including substantial flexibility and generality that allow coping with very complex models. The main distinction between Bayesian signal processing and non-Bayesian methods is in the use of prior densities that quantify uncertainty about the unknowns. Use of priors have important implications both in explanation of results and the repertoire of methods, and in the statistical community their use has been a controversial subject for many years. The controversy seems to have subsided lately, partially due to improved interpretations of the priors in many applications and also to the maturation of the theory of Bayes methods which has greatly clarified the impact of mismatched priors on the final results [28, 41].

Basics

To put things in perspective, consider first a problem where *estimation* is required. Let the observed data be denoted by y and the set of unknown parameters that have to be estimated by θ. The probability model is described by the joint probability density function, $f(\mathbf{y},\theta)$, which can be expressed as

$$f(\mathbf{y},\theta) = f(\mathbf{y}|\theta)f(\theta)$$

where $f(\mathbf{y}|\theta)$ is the conditional density of the data given the parameters θ, and $f(\theta)$ is the prior density of the parameters. Given the data, $\mathbf{y}$, and the model $f(\mathbf{y},\theta)$, where θ is unknown, all the information about θ is summarized in the (normalized) posterior density $f(\theta|\mathbf{y})$, which by Bayes' theorem is given by

$$f(\theta|\mathbf{y}) = \frac{f(\mathbf{y}|\theta)f(\theta)}{f(\mathbf{y})}. \tag{1}$$

Although the entire trajectory, $\{f(\theta|\mathbf{y})\}_\theta$, of the posterior is sometimes of interest, very often in practice a point estimate of θ is preferred. For example, the value of θ that maximizes $f(\theta|\mathbf{y})$, termed the maximum *a posteriori* (MAP) estimate, is routinely used. The search for the MAP estimate, clearly represents a multivariable optimization problem. Note that $f(\mathbf{y}|\theta)$, when viewed as a function of θ for given $\mathbf{y}$, is called a likelihood function. In the non-Bayesian literature, $f(\theta)$ is not available and the ML estimate, i.e., the value of θ that maximizes $f(\mathbf{y}|\theta)$, is often used.

Estimation, Detection, and Tracking

The best estimate of θ is specified as that value, $\hat{\theta}$, that minimizes a prespecified cost function. When the cost function is quadratic, the estimate $\hat{\theta}$ is the conditional mean, $E[\theta|\mathbf{y}]$, that is,

$$\hat{\theta} = \int \theta f(\theta|\mathbf{y})d\theta$$

and it represents the minimum mean-squared error (MMSE) estimate. To implement the MMSE estimate, two multidimensional integrations are required; one for computing $\hat{\theta}$ given the posterior, and one for computing the normalization factor $f(y) = \int f(\mathbf{y}|\theta)f(\theta)d\theta$ in Eq. (1). For an interesting application see [92], which treats the problem of Bayesian power spectral density estimation. Multidimensional integration is also necessary when some of the signal parameters are not of interest. These uninteresting parameters, referred to as nuisance parameters, are then integrated out of the posterior, which decreases the complexity of the problem in some cases.

In signal *detection*, the primary objective is to determine whether or not a signal is present in the observed data. The signal may be one of many hypothesized signals, so the goal is to decide which one is in the data. In the Bayesian literature this problem is known as model selection, and it is addressed by first defining a model, $\mathcal{M}_k$, for the k-th signal, followed by associating it with a joint probability distribution, $f(\mathbf{y},\theta_k,\mathcal{M}_k)$, where θ_k denotes the parameters of the model $\mathcal{M}_k$. Then one proceeds with the use of Bayes' theorem and evaluation of the posterior $f(\mathcal{M}_k|\mathbf{y})$, which again requires multidimensional integrations.

When it is of interest to *track* signal values that continuously change with time, the signal model is frequently described by a state-space representation. In the case of a linear model and additive Gaussian noise, the Bayesian solution is the well known Kalman filter [245]. If the model is nonlinear and/or the noise is non-Gaussian, the Bayesian formulation leads to a nonlinear filtering problem whose solution again requires multidimensional integration.

Some Difficulties

The conceptual simplicity of the Bayesian methodology notwithstanding, Bayesian methods have been underused by the signal-processing community mainly because of their high implementation complexity. Indeed, there are only a small number of scenarios where the needed optimizations or integrations can be carried out analytically. Generally, analytical or numerical approximations are required that often give discouragingly complicated mathematical functions. This picture has gradually changed with the emergence of increasingly powerful computers. Fast numerical techniques for implementation of the necessary operations have revolutionized the practice of Bayesian estimation, detection, and tracking over the past several years.

Monte Carlo Methods

The multidimensional integrations and optimizations involved in Bayesion methodology can be approximated accurately by Monte Carlo simulation if one is able to sample from the posterior distributions. In most cases of interest, however, such sampling is impossible. A more practical method is to generate samples from simpler distributions followed by approximation of the integrals by sample averages. One of these methods is importance sampling [28, 123, 297]. The generating distributions are called importance sampling functions, and their choice is crucial because the variance of the estimated integrals is critically dependent on them. A related procedure is sampling-importance resampling, which generates samples from the posterior distribution by repeated sampling from simpler distributions [342]. In this method, samples are first generated from a suitable approximation of $f(\theta|\mathbf{y})$, say $g(\theta)$ yielding samples $\theta_1, \theta_2, \cdots, \theta_L$. Then to each sample a probability mass proportional to the weight $w_l = v(\theta_l|\mathbf{y}) / g(\theta_l)$ is associated, where $v(\theta|\mathbf{y})$ is the un-normalized posterior density of θ. Finally, the posterior density is simulated by drawing samples from $\{\theta_1, \theta_2, \cdots, \theta_L\}$ with probabilities proportional to $(w_1, w_2, \ldots, w_L)$.

Another approach is to use MCMC methods [28, 123, 137, 297, 392]. These methods have extended the Bayesian methodology to many previously intractable applications. The key idea is to generate samples by running an ergodic Markov chain whose distribution after convergence is the desired posterior distribution. Similarly to importance sampling, samples are drawn from a simple posterior-approximating distribution and are subsequently corrected to improve the approximation. The samples are generated sequentially from distributions dependent on the samples most recently drawn, thereby forming a Markov chain.

The first MCMC method was proposed by Metropolis in the early fifties [268] and was used in computational physics. The Metropolis algorithm draws samples from a symmetric distribution and they are accepted or rejected according to a prescribed acceptance probability. This procedure is repeated a sufficiently large number of times. The Metropolis algorithm was later generalized by Hastings to nonsymmetric sampling distributions, which is known as the Metropolis-Hastings algorithm [155]. A third MCMC method is the Gibbs sampler, which employs conditional sampling and may be considered as a special case of the Metropolis-Hastings algorithm. The parameter vector is divided in subvectors, and each of them is drawn conditional on the remaining subvectors. It turns out that with this scheme the probability of acceptance is equal to one and there are no rejections. The Gibbs sampler has been widely and successfully applied to problems in image processing where the number of unknowns is very large [124].

The MCMC sampling methods were further generalized to allow for sampling from parameter spaces corresponding to different models [148]. The new method, called the reversible jump MCMC sampler, jumps from one parameter space to another based on transition probabilities of another Markov chain. Once the sampler has converged, the time it spent in a specific parameter space is proportional to the posterior probability of the associated model. Thus, the reversible-jump MCMC can implement simultaneous Bayesian detection and estimation.

There are several important issues related to the use of MCMC methods. First of all, these methods are iterative, so the question of convergence is critical. Typically, the first N iterations are thrown away, a period called burn-in, and determining N demands convergence diagnostics. Of practical importance, too, is the stopping time of the chain. One would like to run the chain long enough to obtain desired accuracy. How many parallel chains to run is another important question. When there are many chains, their comparison may allow for easier determination of their convergence.

MCMC Sampling for Signal Processing

MCMC methods have the potential to yield iterative solutions to many important but difficult signal-processing problems. An increasing large number of papers on the subject have appeared in signal processing conference proceedings and journals. The following is a short list of recent contributions; for a more detailed overview, see [8].

One standard problem in signal processing is blind deconvolution of noisy data that represent an output of a linear system excited by an unknown input. For instance, if a system is modeled as an FIR filter with unknown coefficients, and the input is a hidden Markov model with discrete and known state space but unknown initial state and transition probabilities, it is often important to determine the filter coefficients, the input, and its model parameters from a set of distorted observations. It was shown that the estimates of all the unknowns in this problem can be obtained straightforwardly by Gibbs sampling [58]. In a different setting, blind deconvolution of Bernoulli-Gaussian processes was implemented by an MCMC approach in [96].

Optimal filtering is another area of major activity in signal processing. When the signal model is nonlinear or the signal is non-Gaussian the Kalman or extended Kalman filters can give very poor performance. A powerful alternative is to use Bayesian filters based on sequential importance or Gibbs samplings [52, 146, 391]. A signal-processing application of the Metropolis algorithm is parameter estimation of damped sinusoids [16]. When joint detection of sinusoids and the estimation of their parameters are of interest, the reversible jump MCMC has been successfully applied [7, 91]. MCMC methods have been successfully applied to the selection of model order of a time series [15, 22, 145, 417]. In some of these papers nonstandard assumptions were made, which make the direct estimation problem quite intractable. Analysis of mixed spectra within a hierarchical Bayesian framework was proposed in [53] via an MCMC algorithm. MCMC methods have also been very useful in enhancing speech and music signals, which are degraded by non-Gaussian noise characterized by impulses superimposed on a Gaussian background [144]. Finally, in addition to its extensive application to static images, Gibbs sampling has been employed in enhancement of degraded video images [210].

With the ever growing power of modern computers, MCMC methods are becoming very useful tools for tackling even the most difficult signal-processing problems. In the years to come these methods will become more sophisticated, efficient and accurate, undoubtedly leading to many new and effective signal processing algorithms.

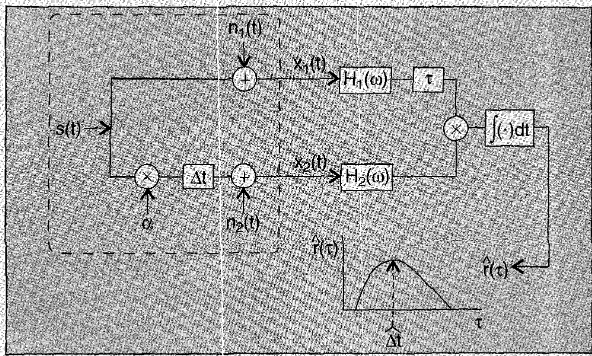
WWW links relevant to the above section:

- A WWW link to the author of the section: http://www.ee.sunysb.edu/~djuric/index.html
- A WWW link to topics in the area of MCMC: http://www.stats.bris.ac.uk/MCMC/

Time-Delay Estimation: Past, Present, and Future

Hagit Messer and Jason Goldberg, Tel Aviv University

Time-delay estimation (TDE), or time-of-arrival (TOA) estimation, is a basic tool in statistical signal processing. Applications of TDE follow from the simple relationship of $\Delta r = v \cdot \Delta t$, where Δr is the distance an object or a wavefield travels at some constant speed, v, over some time interval, Δt. For example, in range measurements for radar or sonar, v is assumed known, and the target's range is determined by measuring Δt, the time required for the transmitted signal to propagate to a target and be reflected back to point of transmission. Also, for velocity



▲ 1. *The generalized cross correlator for passive TDE. For active TDE (when s(t) is known and $n_1(t) = 0$), the matched filter is obtained by setting $H_1(\omega) = H_2(\omega) = 1$.*

measurements (e.g., biomedical [55, pp. 469-476], or nuclear engineering applications [55, pp. 363-366]), Δr is assumed known, and Δt, the time required for a signal to travel the distance, Δr, is measured.

TDE is also important for more complicated, nonlinear problems such as source direction of arrival (DOA) or bearing measurement. If (assuming free-space propagation conditions) the signal due to a source is received by two sensors separated by distance l, then the differential delay between the signal received by the sensors is given by $\Delta t = l \cdot \sin\theta / v$, where θ is the source DOA [55, pp. 403-409]. Determination of source DOA is often based on a measurement of differential time delay, Δt, or, for a narrow-band source, a measurement of differential phase shift, $\Delta\psi = \omega_o \Delta t = 2\pi l/\lambda \ \sin\theta$, where the source frequency is written in terms of the source wavelength as $\omega_o = 2\pi v/\lambda$.

In practice, one seeks to measure the delay between two *noisy* versions of a signal (which itself may even be unknown). Unfortunately, there is no single measurement procedure appropriate for all TDE scenarios. This fact, combined with the practical importance of measuring time delay in so many different applications, is why TDE has received so much attention over the last three decades, e.g., [55], [54].

Basic TDE

The most general TDE problem is now formulated. The noise-corrupted signals received by the two sensors over some time interval can be modeled as:

$$x_1(t) = s(t) + n_1(t)$$
$$x_2(t) = \alpha \cdot s(t - \Delta t) + n_2(t), \qquad t \in [0, T] \qquad (2)$$

The problem to be solved is that of using the measured data to determine $\Delta\hat{t}$, an estimate of the time-delay parameter, Δt. Depending on the application, different modeling assumptions are made on the signal waveform, $s(t)$; the noise waveforms, $n_1(t)$ and $n_2(t)$; and the parameters, Δt and α. It is convenient to distinguish between the cases of so-called "active" and "passive" TDE [55, pp. 442-448].

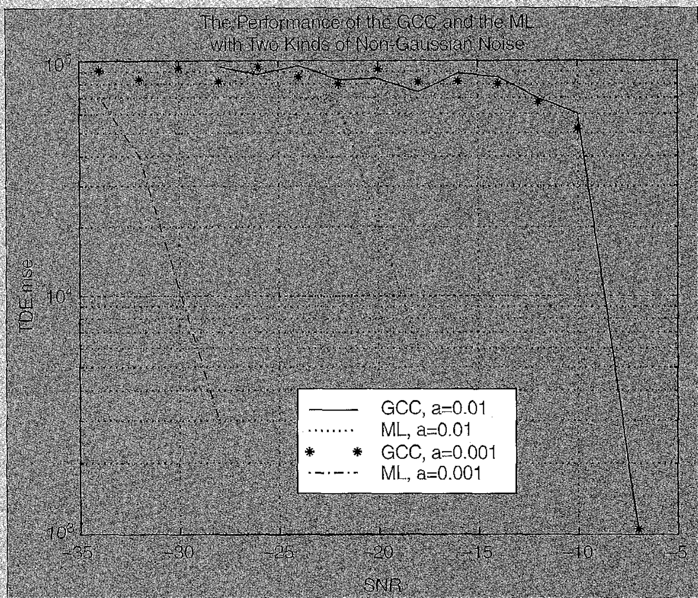
▲ In *active* TDE, where $x_1(t)$ and $x_2(t)$ correspond to the transmitted and received signals, respectively, it is appropriate to assume that $n_1(t) = 0$ and that $s(t)$ is a known, deterministic signal. It is well known that for the "nominal active scenario," where $n_2(t)$ is a realization of a white, Gaussian random process, the (asymptotically optimum) ML TDE processor is the matched filter, which cross-correlates $x_1(t)$ and $x_2(t)$. The estimate, $\Delta\hat{t}$, is the time that corresponds to the maximum of the matched filter output.

▲ In *passive* TDE, where $x_1(t)$ and $x_2(t)$ are two versions of the received signal, it is usually assumed that $\alpha = 1$, $s(t)$ is a realization of a stationary, Gaussian random process, and $n_1(t)$ and $n_2(t)$ are realizations of mutually uncorrelated, zero mean, white, stationary, Gaussian random processes that are also uncorrelated with the signal. This well known "nominal passive scenario" was addressed and "solved" some 20 years ago. In particular, it was shown that the (asymptotically optimum) ML processor for this scenario is the generalized cross-correlator (GCC) [55, pp. 138-144]. As shown in Fig. 1, the GCC cross-correlates appropriately prefiltered versions of the sensor outputs, forming $\Delta\hat{t}$ as the time corresponding to the maximum GCC output. Moreover, the GCC has been shown to be optimum even in nonasymptotic conditions [55, pp. 126-129].

Intuitively, an estimator for Δt in Eq. (1) should seek the best "match" between $x_2(t)$ and a delayed version of $x_1(t)$. In both active and passive TDE, some form of cross correlation has been proven to be the optimum measure for matching under Gaussian conditions. Another possible measure for matching could be the "error signal," $e(t) = x_2(t) - x_1(t - \tau)$. An appropriate estimate of time delay, $\Delta\hat{t}$, would be the τ, which minimizes this error in some sense. It can easily be shown that for the Gaussian scenarios described above, the optimum TDE processor minimizes the mean-square error (MSE), $E\left[|e(t)|^2\right]$. However, straightforward manipulation of the expression for the MSE shows that the minimum MSE and the correlator based processors are equivalent.

Advanced TDE

The intuitive basis on which both correlation and MSE-based processors match the two versions of the signal (Eq. (2)) leads one to believe that they may also function in "non-nominal" scenarios. However, if the assumed conditions under which these procedures were derived are violated, they are no longer optimum. This implies that better estimation performance can be achieved using other processors. Most of the TDE research carried out over the last 20 years has been devoted to the design and analysis of processors for non-nominal scenarios that are known to commonly arise in practice. Current and future TDE research focuses on both the theoretical and practical



▲ 2. *TDE performance vs. signal to noise ratio (SNR) in white, mixed-Gaussian noise (i.e., the noise probability density function is an average of two zero-mean Gaussian density functions of different variances). The ratio between the variances, a, indicates the deviation from Gaussianity (the noise becomes "more Gaussian" as a approaches unity). Note that the leveling off of estimator performance as SNR is decreased is in fact an artifact introduced by the limited range over which the search for Δt is carried out.*

issues associated with such scenarios. Some examples are given below:

▲ For a moving source and/or moving sensors, the delay to be estimated is time varying. Much research has been directed toward adaptive TDE, tracking, and the development of parametric techniques for simultaneously estimating delay and Doppler shift, e.g., [55, Part 4].

▲ The implementation of the GCC requires prior knowledge of the power spectra of the source signal and the noise processes. It is well known that the asymptotic delay estimation error does not increase if the spectra are unknown. The nonasymptotic error for the case where the source spectrum is known has been studied [267] using the Barankin bound. It has been shown that while the shape of the spectrum determines the threshold SNR (under which the estimation error is larger than the Cramer-Rao bound), prior knowledge of the shape yields negligible benefit in performance.

▲ Implementation of modern TDE processors requires the use of digital signal processing. The adaptation of classical TDE techniques to modern, software-based realizations raises both practical and theoretical problems. In particular, interpolation of the discrete delay estimate [55, pp. 343-350] and the effect of sampling on the achievable TOA error [18] have both been studied.

Perhaps the most important deviation from the nominal scenarios is that of *non-Gaussian* statistical scenarios. Some of the main results for non-Gaussian TDE are now reviewed.

Partially Correlated Gaussian Noise

Correlation-based TDE procedures are based on the assumption that the additive noise processes are uncorrelated. In cases where the signal, $s(t)$, is non-Gaussian while the additive noise processes are Gaussian (and possibly correlated), high-order statistics (HOS) techniques have been proposed for TDE [55, pp.168-171], [166], [427]. Since the Gaussian components of the received data are attenuated in the high- (i.e., greater than second-) order cumulants and spectra, processing the data in these domains can result in improved TDE performance when the noise processes are highly correlated.

Independent, Non-Gaussian Noise

While non-Gaussian signal and Gaussian noise processes are successfully handled by HOS techniques, a different approach is required when the noise processes are non-Gaussian. In [362] passive TDE has been studied where the additive, non-Gaussian noise processes are assumed to be statistically independent. Fig. 2 shows that under such an assumption the GCC continues to function but that procedures better matched to the distribution of the noise statistics (e.g., ML) can improve performance.

Multiple-Window Spectrum Estimates

David J. Thomson, Bell Labs

One of the basic problems in signal processing is to estimate the spectral density function, or power spectrum, from a finite data sample. Given N observations, $x(t)$ for $t = 0,1,\ldots, N-1$, equally spaced in time at $\Delta t = 1$, how does one estimate the spectrum? As the inventor of multitaper estimates [398], I may have a biased opinion on the answer to this question, but, taking a direct quote from [394]:

"Spectral analysis has recently undergone a revolution with the development at Bell Labs of sophisticated techniques in which the data are multiplied in turn by a set of tapers which are designed to maximize resolution and minimize bias [*Thomson* 1982]. In addition to minimizing the bias while maintaining a given resolution, the multitaper approach allows an estimate of the statistical significance of certain features (such as spectral lines) in the power spectrum by comparing the character of the DFT's of different data windows. These techniques are now in routine use..."

To understand the origins of this process, remember that the first commonly used estimate of the spectrum was Schuster's *periodogram*, introduced in 1894. It is the square of the discrete Fourier transform of the observations, scaled by $1/N$ and, as an estimate of the spectrum, is both biased and inconsistent. In practice this means it gets an unstable wrong answer. There are statements in the statistical spectrum estimation literature that, as a function of sample size, the periodogram is "asymptotically unbiased." Ignore these. Engineers learn that the only reason anyone goes to the trouble and expense of collecting more data is because they are going to ask more difficult questions, so one is always working with small-sample problems, not with asymptotics. In engineering data this bias can overwhelm the signal of interest; in [397] I showed data where the periodogram was in error by more than a factor of 10^{10} over most of the frequency range. The periodogram is *inconsistent* because its variance, $\mathbf{E}\{\mathbf{P}(\mathbf{f})\}^2+$, does not decrease with sample size and, in the cited example, was too large by a factor of more than 10^{20}. Although these problems with the periodogram were known before World War II, many researchers persist in using periodograms. (Be cautioned that estimates that are based on the periodogram or raw DFTs have similar bias problems. Sample autocorrelations are just the Fourier transform of the periodogram so Blackman-Tukey, autoregressive, maximum-entropy, and other spectrum estimates that depend *directly* on sample autocorrelations should not be used. Similarly, estimates of the analytic signal derived from Hilbert transforms using unwindowed DFTs have periodogram bias.) A glance at the current literature on spectrum estimation theory and practice confirms that evolution is a slow process.

Fortunately, most signal processors follow the recipe in Tukey's 1966 paper [436] for computing an estimate of the spectrum: choose a suitable *data window, $D(t)$*, compute

$$S_D(f)=\left|\sum_{t=0}^{N-1} x(t)D(t)e^{-i2\pi ft}\right|^2 \tag{3}$$

and smooth, often by convolving $S_D(f)$, with a second window. Good data windows give a much less biased estimate of the spectrum than the periodogram. Because $S_D(f)$ is the sum of two squares (the real and imaginary parts of the DFT at frequency f) it has a *chi*-squared distribution with two degrees of freedom. Thus $S_D(f)$ is still inconsistent and the smoothing part of the recipe is necessary to obtain a useful estimate. This estimate, however, poses a new problem: where, apart from John Tukey saying it was a good idea, did D, the data window, come from?

Multiple-window, or *multiple-taper*, spectrum estimates were introduced in [398] in an attempt to correct many of the shortcomings with "standard" spectrum estimation procedures. Here one chooses an analysis bandwidth, W, for the estimate $0 < W \le \frac{1}{4}$ with $NW \approx 4$ or 5 a typical choice. The dimensionality of a signal with bandwidth W and a time duration of N samples is $K = 2NW$. Because our goal is to estimate the energy in a frequency band $(f-W, f+W)$ as accurately as possible, we must choose the K sequences of duration N whose energy concentration in this band is the best possible. These sequences are the *discrete prolate spheroidal sequences*, or *Slepian sequences*, [375]. Now compute the *eigencoefficients*:

$$y_k(f)=\sum_{t=0}^{N-1} x(t)v_t^{(k)}(N,W)e^{-i2\pi ft} \tag{4}$$

where $v_t^{(k)}(N,W)$ is the k^{th} Slepian sequence and parameters N and W for $k=0,1,\ldots,K-1$. From these the crudest multiple window spectrum estimate is

$$\hat{S}(f)=\frac{1}{K}\sum_{k=0}^{K-1}|y_k(f)|^2, \tag{5}$$

which is simply an average of K estimates of the form (Eq. (3)) made using the same data but with different tapers. Because the different tapers are orthogonal, the different terms in Eq. (5) are approximately uncorrelated. Each contributes two degrees-of-freedom, so Eq. (5) has a χ^2 distribution with about $4NW$ df, so the estimate is consistent. In practice, an adaptively weighted average of the $|y_k(f)|^2$'s is preferred to Eq. (5), see [398], [400], and [403].

The multiple-window theory explained the origins of the data window, or taper. One is finding an approximate solution of the integral equation connecting the observations and the spectral representation of the process. The windows are the eigenfunctions of the kernel. From this perspective, Tukey's estimate (Eq. (3), was approximately the first term of the series solution (Eq. (5)).

As part of this theoretical development an effort was made to separate the deterministic (commonly periodic) components of the process from the nondeterministic background, which the older estimates had lumped together. With multiple windows, detection of sinusoids is commonly done with an F-test, which is a ratio of the energy explained by a periodic component at frequency f, to the remainder of the energy in the frequency band $(f-W, f+W)$; see [398], [401], [400], [187], and [408].

In the original multitaper estimate, an approximate *linear* inversion of the integral equation was used, and the spectrum was obtained by local averages of its magnitude; *quadratic-inverse* theory [400] gives minimum-variance unbiased expansions of the spectrum and represents a step in the process of eliminating the dependence on the choice of the bandwidth W. This has been extended [402] to nonstationary problems.

When Do You Use Multitaper Estimates?

Generally, multitaper methods have become the estimate of choice for serious spectrum estimation problems, are becoming "routine" in geophysics, [393], and, with an excellent text on the subject [307] and availability in *MATLAB*, appear to be becoming so in other fields, [208, 238]. A quick survey of papers describing work using multitaper methods shows that most of the early applications were scientific. Special windows [305], and combined time and space F-tests [235, 236] were developed for normal-mode seismology, estimates of polarization [303], attenuation [186, 468] and other geophysical quantities [169, 191, 304, 393, 394]. These, augmented with coincidence tests, have been applied to processes with *many* line components in [340] and [408]. Theory and examples of coherence estimation and some multivariate applications are in [218], [398], [407], and [442]. Several papers ([398], [235], [237], [47], [369], [337], [448], [153], and [259]) find multitaper methods outperform "classical" alternatives. Because sample autocorrelations are just the Fourier transform of the periodogram (and hence undesirable), multitaper correlation estimates, the Fourier transform of Eq. (5), have been studied in [440] and [261].

For "traditional" stationary, Gaussian processes spectrum estimates have chi-square distributions. In practice, however, confidence intervals based on these were often wildly optimistic, and resampling estimates based on the jackknife [407] and bootstrap [472] are becoming common. For a comparison of resampling methods, see [116]. For explicitly, non-Gaussian data, multitaper *bispectrum* estimates have been developed, [279, 399], and work well. (A few of the difficulties encountered with nonstationary and non-Gaussian data are discussed elsewhere in this article.)

Variations on the Multiple-Window Theme

Using a different error norm for solving the integral equation, Riedel and Sidorenko [334, 335] developed a multitaper "sine" estimate. While the 60 dB range of these windows lack the crushing sidelobe performance of the Slepian tapers, they are adequate for many applications and easier to compute. Other choices for tapers are discussed in [280] and [17], and efficient methods for computing the Slepian sequences are given in [400] and [150] and elsewhere.

The theory has been extended to arbitrarily-sampled spatial data [46, 243] but, for irregularly sampled time series, interpolation may be a viable alternative [409, 410].

Turning to nonstationary processes, multitaper spectrograms have been in use since shortly after the invention of multitaper estimates, see [401], [224], [191], [336], [312], and [164]. Multitaper wavelets [80, 85, 234] have also been used. For weakly nonstationary processes, quadratic-inverse theory [402, 403, 406] works well (the first few coefficients are nearly the *time-derivatives* of the spectrum) while, in violently nonstationary examples, estimates of the Loeve, or two-frequency, spectrum [164, 263, 359, 398] often give more insight. Taking a singular value decomposition of a log multitaper spectrogram is often useful, [401]. Narrowband tracking filters, [223] projection filters [404, 405], and inverse-theory reconstructions [302] are in use. These have been used in a series of papers applying signal-processing methods to study the relationship between atmospheric CO_2 and global warming [193, 218, 405, 406]; the warming is mostly CO_2 and, more disturbing, the seasonal cycle is also being disrupted by human use of fossil fuels.

To summarize: if you are estimating spectra, you should be using multitaper estimates. If the data is expensive, of limited duration, or if difficult questions are being asked, multitaper estimates are mandatory.

Time-Frequency Distributions in Statistical Signal and Array Processing

Moeness G. Amin, Villanova University

The 50th anniversary of the IEEE Signal Processing Society also marks 50 years since Ville [446] applied Wigner distribution (WD) to signal analysis. The WD, which was the first distribution introduced in the context of quantum mechanics [451], has paved the way to several key contributions to advances in the area of time-frequency distributions (TFDs) as well as representations of signals with time-varying characteristics. These contributions have aimed at overcoming the drawbacks of the WD and sought new, more effective tools for nonstationary signal analysis, synthesis, and processing.

In the limited space provided, we will highlight only some of the advances made in the time-frequency distributions over the past half a century with more emphasis on immediate than distant past contributions and on the articles that are relevant to themes embraced by the SSAP technical community.

Sixteen years after Wigner had introduced his distribution, Cohen [69] provided a consistent set of definitions for a desirable class of TFDs, often referred as Cohen's class. This class has been of great value in guiding efforts in this area of research. Cohen's class of time-frequency (t,ω) distributions for the signal $x(t)$ may be presented in different forms, including

$$\rho(t,\omega;\phi)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(t-u,\tau)x(u+\tau/2)x(u-\tau/2)e^{-j\omega\tau}\,du\,d\tau \qquad (6)$$

Different distributions are obtained by selecting different kernels, $\phi(t,\tau)$. Both WD (also known as Wigner-Ville (WV) distribution) and the spectrogram are prominent members of Cohen's class. Extensive investigation of the desirable properties of a distribution and associated kernel requirements was done by Claasen and Mecklenbrauker [67]. Their three-part paper published in 1980 drew much attention to the limitations and offerings of Wigner distribution and marked the first comprehensive treatment of the subject using familiar continuous and discrete signal-analysis methods. Other important contributions that have given a timely overall "big picture" view of the state of this dynamically changing and rapidly growing area in signal processing are the review articles by Cohen [70], Boashash [32], and

Hlawatsch and Boudreaux-Bartels [170]. In some respects, these articles have elaborated on a single-channel deterministic aspect of TFDs and did not fully address the problem from a statistical signal-processing perspective. The focus on deterministic signals stemmed from the fact that TFDs have clear and well understood properties when dealing with noiseless and nonstochastic environments. Further, TFDs have been successfully applied to areas where signals are localizable in the time-frequency domain and have fixed distinct signatures that permit their classification and separation. Many of these applications are discussed in the book by Cohen [71] and also in the book by Qian and Chen [325].

The paper by Boudreaux-Bartels and Parks [37] was the first to recognize that by devising a method to synthesize the signal from the time-frequency domain, the WD may be cast as a tool for signal enhancement and noise suppression. In the case of signal in additive white noise, the WD of the noise is scattered over the entire time-frequency domain whereas that of the signal is confined to a much smaller region. If only the signal in that region is synthesized, the desired signal can be retrieved with reduced noise contamination and improved SNR. Several papers have since appeared in the literature as alternatives to the least-squares approximation technique employed by the synthesis method.

For multicomponent signals, the cross-terms (also referred to as interference terms), which are introduced from the bilinear nature of the TFDs (Eq. (6)), intrude into the time-frequency regions containing true signal power concentrations, known as "auto-term regions." In such a case, and also for low SNR environment, the signal auto-terms may not be identifiable using WD, which renders signal classifications and synthesis difficult and sometimes impossible. Choi and Williams [62] and Zhao, Atlas, and Marks [465] have proposed t-f kernels, which make such identification much more feasible than attainable using the WD. The distributions corresponding to these two kernels have come to be known as the Choi-Williams (CW) and the ZAM TFDs. In both distributions, the kernel is characterized by one parameter whose value may be adjusted to achieve a tradeoff between resolution and cross-term suppression. A fully signal-dependent kernel was proposed by Baraniuk and Jones [12], where the kernel self-adapts its shape based on the underlying signal characteristics. In this respect, unlike both the CW and the ZAM TFDs, the signal-dependent TFD is a nonlinear smoothing of the WD. This distribution, which remains attractive for a wide class of signals, is based on the same foundation as that of the earlier CW distribution: they both make use of the fact that the cross-terms are oscillatory in nature and therefore lie away from the origin when examined in the ambiguity domain. A byproduct of mitigating cross-terms in a distribution is smoothing the noise level in the time-frequency domain. The receding of the noise fluctuations brings about a clear manifestation of the signal signature in the time-frequency domain. Step-by-step design for kernels leading to reduce interference distributions (RIDs) was given by Jeong and Williams [183]. Indeed, the above two influential papers ([62], [183]) on the reduction of cross-terms through low-pass filtering in the ambiguity domain have made the technical community more attentive to the flexibility and generalization underlying Cohen's class of TFDs and has set the stage for a surge of activities in this area in the last decade. Other papers that have given valuable insights and important perspectives to TFDs include the maximum entropy approach to positive TFDs by Loughlin, Pitton, and Atlas [246], the polynomial WV distributions by Boashash and O'Shea [33], and the L-Wigner distribution by L. Stankovic and S. Stankovic [380].

Martin and Flandrin [256] considered the WD for random processes by carrying over its desirable properties to stochastic environments. The expected value of the WV defines the Wigner-Ville spectrum (WVS), where the desirable properties of WD are now satisfied in the mean sense and given in terms of moments and power spectrum. The ensemble average of the distribution in Eq. (6) is the time-frequency spectrum, which simplifies to the WV spectrum when the kernel is an impulse function over time. Estimating the WVS using time-averages and assuming quasi-stationarity brings the WVS to the same form as that of the distribution in Eq. (6). It is noteworthy that for smoothed Pseudo Wigner estimators, the t-f kernel is separable, leading to independent smoothing in time and frequency. In addition to satisfying the TFD properties, reducing cross-terms, providing positive spectrum, etc., the kernels in the statistical context must further lead to an unbiased and reliable estimate. Therefore, the difference between the t-f kernels chosen for deterministic signals and those for stochastic processes is that the latter are subject to an increased number of constraints in order to serve the statistical aspects of the problem. The compatible and conflicting requirements on the t-f kernel along with a discussion of nonstationary processes where the kernel statistical constraints are compatible with their deterministic counterparts were presented by Amin [4]. Amin has also derived the optimum kernels for reducing spectral variance when dealing with nonstationary signal in additive white Gaussian noise [5]. Sayeed and Jones [350] proposed a minimum MSE of the WVS. Their method is based on the optimization of the variance-bias tradeoff using knowledge of certain second- and fourth-order moments. The results of this optimization are kernels very different from those considered by invoking the quasi-stationary assumption.

Kayhan, El-Jaroudi, and Chaparro have revived in [203], and also in follow-up publications, the evolutionary spectrum (ES) considered by Priestley [322]. This spectrum is based on the modeling of a nonstationary signal, $x(n)$, as a collection of uncorrelated sinusoids with random time-varying amplitudes, namely:

$$x(n) = \int_{-\pi}^{\pi} A(n,\omega) e^{j\omega n} dZ(\omega)$$

where $A(n,\omega)$ is the time-dependent amplitude and $Z(\omega)$ is an orthogonal increments process. For a given signal that admits this representation, the ES is defined as the magnitude squared of $A(n,\omega)$. The work in this area has successfully led to the generalization, estimation, and the linkage of ES to TFDs. For processes with restricted time-frequency correlation, referred to as underspread nonstationary random processes, it has been shown by Matz, Hlawatsch, and Kozek in [258] that the most popular definitions of time-varying spectra, such as the generalized WV spectrum and generalized ES, yield effectively equivalent results. The concept of underspread processes proves useful in analyzing Doppler shifts and fading communication channels using time-frequency structures.

TFDs have also been examined for detection of nonstationary signals. Almost one decade ago, Flandrin [113] provided a coherent framework for WV time-frequency receivers. It was shown that classical receiver structures designed for optimum detection of Gaussian signals in Gaussian noise admit an equivalent formulation in the time-frequency domain. The work by Sayeed and Jones [350] has gone past the mere equivalence of classical optimum detectors to exploit the time-frequency structures in composite hypothesis tests where an optimum quadratic detector is implemented at each time-frequency point.

Time-frequency distributions have recently found applications to radar and sensor-array processing. Barabarossa and Farina [13] have combined conventional space-time processing with TFDs and demonstrated the advantage of joint processing for clutter suppression and target detection. Belouchrani and Amin [23] have introduced the spatial time-frequency distribution and used it to solve blind source separation and direction finding problems. They derived the fundamental equation

$$\mathbf{D}_{xx}(t,f) = \mathbf{A}\mathbf{D}_{ss}(t,f)\mathbf{A}^{H}, \tag{7}$$

which relates the TFDs of the sensors to those of the sources. In Eq. (7) $\mathbf{A}$ is the spatial signature matrix. The elements of $\mathbf{D}_{xx}(t,f)$ and $\mathbf{D}_{ss}(t,f)$ are not the commonly used matrix correlation functions, but rather the self- and cross-TFDs of the sensors and sources, respectively.

With their role in advancing knowledge, theory, and applications in the statistical signal and array processing area, TFDs have been established as an integral part of this area, and will remain so for many years to come.

A WWW link to the author of the above section:
http://www.ece.vill.edu/user/moeness

Multiresolution Analysis

Hamid Krim, North Carolina State University, and Jean-Christophe Pesquet, UPS/LSS

Solutions to many engineering problems traditionally invoke (at least partly) spectral domain techniques. The simplicity often provided by the Fourier domain in analyzing a process is in large part due to an implied linearity and time-invariance (whether by assumption for tractability or otherwise), which yields complex exponentials as the eigenfunctions of the process. Equivalently, in a stochastic setting, stationary and linear processes have led to high-performance techniques for estimation and detection. The demand for increased performance in statistical signal applications and the emergence of a whole class of nonstationary problems have resulted in a new active area of research, namely time-frequency methodology (described in the section of this article by Amin) and more recently time-scale (multiresolution) techniques.

Complementing Fourier Analysis

Multiresolution Analysis

It is well known that many physical phenomena may be distinguished by characteristics present at different scales [254]. The statistics associated with these characteristics and/or their evolution across scales may provide unique and powerful signatures. The fractal structure of many natural phenomena, e.g., the coast line of a continent and the patterns on a tree leaf, may be viewed as the result of statistically self-similar patterns [254, 460]. The ubiquity of such phenomena in physical processes has called for a systematic and efficient methodology to capture the multiscale trend and preserve the statistics at the various scales for further processing.

The wavelet theory [86] provided a powerful framework to meet this challenge and gained further prominence upon the ingenious connection made by Mallat [253], with the then well known filter-bank theory [437, 443]. The orthonormal wavelet analysis afforded, in many cases, analytical tractability and, equally important, an efficient tracking of the statistical properties at different scales.

Scale Refinement

An additional scale refinement may be obtained by iterating the basic wavelet analysis on the coarse scales as well as the fine scales of the signals [72]. Referred to as the wavelet packet dictionary, this representation consists of an overcomplete set of functions, out of which an orthonormal basis is selected via an efficient dynamic programming procedure and an additive criterion [72]. Other extensions of the wavelet/wavelet packet decompositions are obtained by imposing some shift-invariance properties [308].

In the interest of space, we classify the multiresolution statistical developments into two major thrusts, the first of which is centered around the analysis with an impact on compression and nonlinear estimation, and the second focusing on the multiresolution modeling aimed more at large-scale estimation and classification problems.

Multiscale Statistical Signal Analysis and Modeling

On the analysis front the good localization properties of wavelets have played a key role in the development of various applications such as compression [253, 328] and signal reconstruction [94, 211, 253, 275, 347] (also referred to as denoising). The tree-like structure of the wavelet analysis framework has also led to efficient multiresolution stochastic modeling techniques with a remarkable impact on large scale physics-based estimation and classification problems [19, 249].

Fractal Analysis

Fractal analysis has been a focus of interest in signal processing for almost two decades and continues to play an important role in applications fields such as compression, analysis of turbulence, and communications.

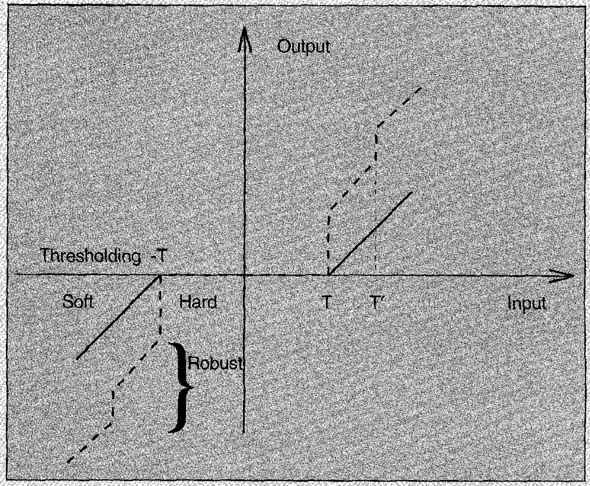
The characterizing scale-invariance of fractal signals, as noted earlier, makes wavelet analysis the tool of choice. The orthogonal wavelet representation of fractal processes entirely reflect the properties (statistical or others) residing at different scales. Such analysis led to significant advances in fully and accurately identifying fractal and multifractal signals [179], and in other related synthesis problems of fractional Brownian motion [114, 460, 395, 257]. Using the parsimony-achieving potential of wavelets, other well adapted analyses were proposed [84] for identification/synthesis of turbulent velocity signals. In particular, these were shown to provide a considerable simplification and usefulness for turbulence signals.

Other applications where fractal signals have shown promise include biomedical, biochemistry [1], and communications signals [460].

Other multiscale analyses of related processes, such as nonstationary processes with stationary increments [49, 212], have resulted in stationarizing properties, thus allowing the application of classical statistical techniques and, equally important, a deeper understanding of other nonstationary parametric processes [212]. (Fractional Brownian motion is a special case of a nonstationary process with stationary increments.) A particularly interesting extension to two-dimensional signals also resulted in analyzing and synthesizing fields such as the ocean floor [314].

Signal Reconstruction

In an attempt to filter the observation noise from a signal, and using the ability of wavelets to concentrate signal energy along a few directions, researchers have tried to separate the directions containing signal energy and the orthogonal directions containing mostly noise energy. These noise directions are then discarded and the former used in the reconstruction. The development of this thrust is marked by three phases that took place in sequence.

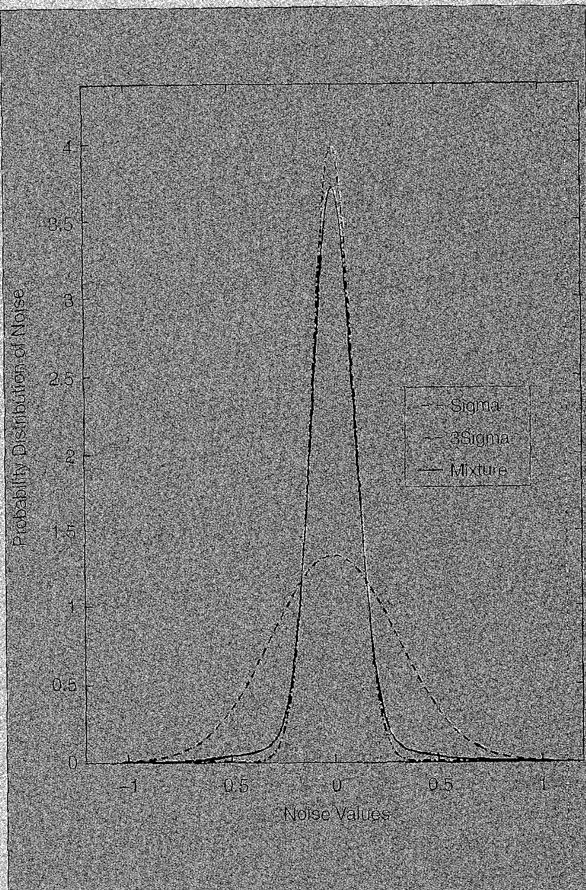


▲ *3. A hard, soft, as well as robust thresholding.*

Signal in Gaussian Noise: Using to advantage the energy compaction property of wavelets, Mallat and Hwang [253] first showed that effective noise suppression may be achieved by transforming the noisy signal into the wavelet domain, and preserving only the local maxima of the transform for subsequent reconstruction. While performing amazingly well, a sound statistical formalism to various deterministic thresholding techniques was first proposed by Donoho and Johnston [94] who considered the observed signal

$$x(t) = s(t) + n(t) \qquad (8)$$

with $t = 1, \cdots, K$, $n(t) \sim N(0, \sigma^2)$ and $s(t)$ unknown. They showed that a certain optimality was achieved by thresholding all wavelet coefficients of $x(t)$ below $T = \sqrt{2\sigma^2 \log K}$. Specifically, given that a wavelet basis is an unconditional basis for a great many smoothness spaces [253], they showed that the reconstruction error was within a scalar multiple of the minimum worst case error over these signal spaces. Further developments ensued as other interpretations of signal enhancement or denoising were adopted. In [275], [347], and [211] the notion of coding was independently used to lead to algorithms with more or less the same type of nonlinear thresholding. The information-theoretic criterion minimum description length (MDL) developed in the late 1970s [338, 365] proved to be very useful in not only providing the threshold, T, upon its minimization in the wavelet domain, but also in clarifying the relationships between the various wavelet properties. In Fig. 3, a summary of the nonlinear thresholding technique (otherwise referred to as a hard-thresholding) together with its various modifications are presented, which for the purpose of this overview are only worth noting. Other extensions in-



▲ 4. *A least favorable noise distribution: Gaussian in the middle and Laplacian on the tails.*

clude techniques that account for a simple correlation structure among the coefficients, which, if it agrees with the signal structure, should lead to better performance [79]. This, however, also results in a higher computational cost. A statistical approach to optimize a process representation in a wavelet packet set that also led to denoising techniques has been proposed in [211], [213], [309], and [95].

Beyond Normality

While the Gaussian noise assumption may be valid in a number of applications, it certainly does not hold in certain environments; for instance, those encountered in an industrial setting. The Gaussian assumption is thus generally limiting, particularly when we have no knowledge about the prevailing noise. An approach inspired by Huber's early work [358] is due to Schick and Krim [358], who, by obtaining the *minimum maximal* description length (DL) of a data sequence in the wavelet domain, derived a *robust* nonlinear filter to cope with non-Gaussian scenarios. The *minimax* DL filter resulted from a mixture noise as shown in Fig. 4 and led to a thresholding rule shown in Fig. 3.

Bayesian Approach

The above thresholding approaches have been demonstrated to lead to good results in relatively moderate noise scenarios and have been successfully applied in a variety of settings. They are, however, based upon threshold values, which present two drawbacks:

▲ They are directly dependent upon the noise variance without regard to the signal characteristics

▲ They grow without bound with the data record length

To address these limitations and particularly when prior information about the underlying signal is available (quite reasonable in practice), a purely Bayesian approach was adopted in [309]. The prior knowledge about the signal in essence regularized the estimation problem. This clearly led to MAP estimates of the signal coefficients.

In contrast to the previous approaches, this approach takes a more elaborate form, allowing one to account for probabilistic prior information one may have about the signal of interest. The Bayesian thresholding filter that results is independent of the data length, K, and is only dependent on the signal and noise variance. Interestingly, it has been shown that many thresholding rules may be included within this framework [445]. For instance, if the noise components are i.i.d. Gaussian and the signal components are i.i.d., zero-mean and have a Laplacian distribution, a soft thresholding policy allows us to recover the signal.

To better account for the expected sparsity of the components of the signal of interest (parsimonious wavelet representation), a prior equal to a Bernoulli-Gaussian distribution (which is a degenerate Gaussian mixture) in presence of i.i.d. Gaussian noise, leads to an estimate that is a tradeoff between a Wiener and a thresholding estimator [309].

Multiresolution Modeling

Multiresolution modeling first appeared in the context of data compression [253]. Its recent increased prominence is largely due to the computational efficiency of filter-bank implementations. This led to a multitude of applications, and one of particular interest herein is the clever recasting of time-recursive filtering into *scale recursive* filtering [19]. This led to modeling algorithms with high efficiency as a result of the tree-like structure of the analysis.

The generality of this modeling technique comes from the fact that the nodes may be interpreted differently depending on the application [249] and the appeal of easily implementable algorithms. Other extensions included the introduction of time dynamics [174] with a variety of applications. One interesting applications is to ground water hydrology where measurements are in fact multiresolution over space, resulting in very efficient data-fusion algorithms and subsequent estimation procedures [83]. Another application is to ocean height estimation with sparse satellite measurement data where multiresolution methods could be implemented with unprecedented efficiency and accuracy, creating a ground-breaking milestone in computational oceanography [112]. Other applications include synthetic aperture radar (SAR) image classification, compression [207], as well as the analysis of self-similar processes [112].

A WWW link to the author of the above section:
http://www.ece.ncsu.edu/people/faculty/bios/ahk.html

Channel Estimation and Equalization

Georgios Giannakis, University of Virginia

Communications and, in particular, channel estimation and equalization are areas that offer a fertile ground for statistical signal-processing tools and algorithms. Information sources are mapped to (generally complex) symbols that take values from a finite alphabet and are thus non-Gaussian signals. They undergo distortions that introduce intersymbol or interchannel interference (ISI or ICI) as they propagate through channels before being received by single or multiple sensors in noise. Receiver noise is narrowband and hence is modeled well as additive Gaussian, although impulsive models appear also with ambient and atmospheric noise sources in underwater acoustic and radio communications. ISI and ICI arise due to bandlimited transmit- and receive-filters, amplifiers, delay- and multipath-propagation, relative transmitter-receiver motion, coupling effects, and multiple access interference (MAI) [323]. Depending on the transmission rate, the propagation conditions, the number of transmitters and receivers, the complex discrete-time equivalent baseband channels can be: (1) deterministic or random constants over one or more information symbols (modeling flat fading effects); (2) single-input single-output (SISO) linear time-invariant (LTI) FIR filters (accounting for frequency-selectivity); (3) multiplicative sequences (modeling time-selectivity); (4) linear time-varying (LTV) filters with random or deterministic coefficients (capturing fast fading effects); (5) nonlinear FIR filters of the Volterra type (modeling saturation nonlinearities of power amplifiers); (6) multi-input multi-output (MIMO) filters (for multiuser scenarios); or, possible combinations of (1)-(6). Equalizers, on the other hand, undo channel effects to recover the transmitted sequence and, depending on complexity versus performance tradeoffs, they can be: (1) linear or nonlinear; (2) FIR or IIR, and (3) batch for block-by-block equalization, or adaptive for efficient online processing and tracking of slowly varying channels.

Channel estimation, equalization, and symbol recovery algorithms are founded on detection-estimation and system identification principles, and their advances parallel and cross-fertilize ideas in diverse signal-processing applications including seismic deconvolution, sonar de-reverberation, image restoration, signal reconstruction, time-series modeling, and various inverse problems involving dispersive media [162]. Estimation of channels using the received (output) samples can be viewed as an input-output system identification problem if a known *training* sequence (input) is transmitted during the acquisition mode. In the operational stage, receivers usually switch to a *decision-feedback* mode where previously equalized and quantized (according to the alphabet) symbols are used together with the received data to update channel estimates [24, 321, 323]. Training sequences consume bandwidth and consumption increases if frequent re-training is required to avoid erroneous convergence of decision-feedback equalizers (DFEs), which occurs when the propagation channel is time-varying; see, e.g., [326] and [90]. To obviate training and thus utilize bandwidth efficiently, self-recovering (a.k.a. *blind*) algorithms have received attention over the last dozen years for identifying the channel or estimating the equalizer directly using output data only—a feature also important when information transmission cannot be interrupted for training, as, for example, in broadcasting and multicasting scenarios [26, 117, 143]. The success of blind methods in a communications context depends on maximum exploitation of input features such as whiteness, non-Gaussianity, finite alphabet, constant modulus, cyclostationarity—properties that equalizers are often designed to restore at their outputs by optimizing pertinent criteria. Some of these input properties can be imposed by the transmitter design (e.g., periodically inserted guard times or constant modulus [3, 143, 416]) and from this point of view blind schemes can be thought of as semi-blind identification approaches because the input may be unknown to the receiver but its characteristics, which can be imposed at the transmitter, are known—a distinct difference with time-series modeling applications where the input is inaccessible and thus cannot be affected by design.

Equalization of linear FIR channels with training is commonly used in practice, especially with wired-line transmissions over telephone lines, cable television, and asymmetric digital subscriber loops (ADSLs); see, e.g., [248], [326], [195], and [343]. Blind approaches show more promise in future wireless and mobile communications, high-frequency modems, and digital audio and video broadcasting systems where rapid channel variations render the overhead for training prohibitive; see, e.g., [117], [103], and [90]. Although receiver-based (or post-) equalization is predominant, when a reverse link is available from the receiver, transmitter-based (or pre-) equalization is also possible when channel estimates are available as shown by [411] and [154]; see also [121] for recent advances in pre-equalization and its relationship with DFEs.

The tools used for channel estimation and equalization include second- and *higher-order statistics* (SOS and HOS). SOS are sample correlations and power spectral densities of the received (symbol-rate sampled) stationary data, whereas HOS refer to, e.g., sample fourth-order cumulants and polyspectral densities that can extract additional information from the non-Gaussian data about the underlying channel [233]. They are useful for mitigation of nonlinear effects [24] and blind estimation of LTI channels because they complement channel magnitude response information (conveyed by SOS) with complete phase response information, which allows equalization of nonminimum phase channels [131, 157, 370, 430]. However, HOS exhibit high-variance, and channel variations may violate the stationarity assumption as the receiver collects long records required for reliable HOS estimation.

With sufficient excess bandwidth at the transmitter, most FIR channels can be estimated and equalized blindly using *second-order cyclic statistics* (SOCS) that become available when one oversamples (or fractionally samples) the continuous-time received signal at a rate higher than the symbol rate [122, 129, 130, 413, 414]. The resulting time series is cyclostationary and the redundancy introduced renders the LTI SISO model equivalent to either a linear periodically time-varying (LPTV) SISO model or to an LTI single-input multi-output (SIMO) model, which can be characterized by multivariate stationary SOS. SIMO models are also appicable, even when symbol-rate samples are collected by multiple receive-antennas [276], and the excess-bandwidth condition adopted to introduce diversity now translates to sufficient sensor separation in order to guarantee channel disparity (co-primeness of the SIMO transfer functions). An important consequence of the LPTV-SISO and LTI-SIMO structures under such time- or space-diversity conditions is that LTI FIR channels can be equalized exactly by FIR equalizers of the same or greater order; see, e.g., [376]. Recall that LTI-SISO channels with zeros on the unit circle cannot be inverted, and even with MMSE (Wiener) equalization performance drops due to noise amplification at the frequencies of the channel nulls. The LTI-SIMO equalization property is analogous to the perfect reconstruction encountered with multirate analysis-synthesis FIR filterbanks [437], and it has practical implications in communications even when zeros are close to the unit circle: (1) truncated FIR equalizers of FIR channels need not be excessively long to approximate ideal IIR behavior; (2) with appropriate initialization, adaptive equalization algorithms can be globally convergent (at least in high-SNR environments) [232]; and (3) so long as the input sequence is persistently exciting (a minimal condition for identifiability) it can be colored or even deterministic, which is important for coded sequences that are nonwhite; see e.g., [276] and [462].

Recent work focuses on inducing cyclostationarity or diversity at the transmitter by means of periodic modulating sequences or redundant filterbank precoders [60, 128, 353, 368, 424]. Equalization exploiting transmit-diversity is very promising because, unlike fractionally sampled equalization (FSE), it imposes no channel disparity conditions, is applicable to nonwhite and deterministic inputs, and shows robustness to channel order overestimation and additive stationary colored noise [368, 424]. At the expense of added complexity and possibility of divergence, DFEs offer a practical alternative to equalizing channels with unit-circle zeros and their asymptotic performance has been shown recently to approach that of *maximum likelihood sequence estimation* (MLSE) [66, 323]. Given a channel estimate, the latter is implemented using Viterbi's algorithm [24] and is the optimum means of recovering the finite-alphabet input, although its practical use is limited due to its complexity, which, especially with multiuser communications, increases exponentially in the channel order and the number of users [441].

Simple nonfrequency-selective FIR channels introduce constant amplitude and phase distortions that can be compensated with automatic gain control (AGC) and differential encoding, respectively (see, e.g., [323]). But recently, even polynomial phase distortions arising due to relative transmitter/receiver motion can be mitigated with generalized differential encoding [142]. Carrier and timing synchronization are, in principle, frequency and time-delay estimation problems, respectively, and can be solved using ML [323] or cyclic methods with fractional sampling [141]. Correct timing acquisition affects performance considerably, especially for code-division multiple access (CDMA) systems entailing asynchronous users, and a variety of methods (including subspace approaches) have been proposed recently (see e.g., [383] and [25]).

Blind equalizers of FIR frequency-selective channels do not need to acquire timing because they absorb it into the channel itself. In addition, methods that rely on the *constant-modulus* (restoral) algorithm (CMA) do not require frequency offset estimation [143, 416]. CMA is basically a HOS-based technique equivalent to the Shalvi-Weinstein algorithm (SWA) [90, 370]. By constraining the linear equalizer's output, CMA reduces the variance of HOS-based techniques and estimates the equalizer directly, as opposed to most parametric and nonparametric HOS based approaches that estimate the channel first (via linear equation or nonlinear matching methods [233, 430]) and next equalize using either: (1) the computationally complex Viterbi; or, (2) the so-called zero-forcing equalizer (ZFE), which is nothing but a truncated version of the inverse channel [248]; or (3) the MMSE (a.k.a. Wiener) equalizer that assumes knowledge of the SNR to obtain a regularized inverse [317, 323]. Direct adaptive linear equalizers of the CMA/SWA type are particularly attractive computationally, but convergence and speed may be problematic, especially with channel roots on (or close to) the unit circle—a concern that is alleviated with fractionally sampled versions of the originally developed symbol-spaced CMA/SWA [90, 232]. MMSE or ZFEs can be used to initialize the CMA or other nonlinear schemes whose convergence depends crucially on initialization. CMA can also be invoked in a DFE mode to improve performance, especially with channel nulls. Although interesting preliminary results have appeared, the convergence of FIR CMA-FSEs and their DFE versions in the presence of noise is not fully understood.

Fading channels appear with mobile cellular telephony, temperature, and salinity variations in underwater environments, and ionospheric fluctuations in microwave links, where variations of short coherence time cause runaway effects in adaptive tracking algorithms. They are modeled as LTV FIR filters with the average extent (delay-spread) of the multipath defining channel memory and degree of frequency-selectivity, and with the so called Doppler-spread accounting for the average channel variation and measuring time selectivity [323]. The latter can also arise due to oscillator drifts and relative motion that manifest themselves as multiplicative noise when frequency selectivity is negligible. In general, LTV channel taps are modeled as uncorrelated stationary random processes that are assumed to be low-pass, Gaussian, with zero mean (Rayleigh fading) or nonzero mean (Rician fading) depending on whether line-of-sight propagation is absent or present [323]. Correlations of the unknown taps capture average propagation characteristics and are used to track the channel's time evolution using Kalman Filtering estimators. The challenging task of estimating random channel parameters using training data or decision-feedback has been addressed in [425]. Blind approaches for *random coefficient* fading channels are yet to be developed.

However, blind methods adopting *deterministic finitely parameterized LTV models* have been proposed recently using a basis expansion [133, 421]. They turn LTV-SISO models to LTI-MIMO structures with inputs formed by modulating the transmitted sequence with the bases. Fourier bases are well motivated for modeling rapidly fading mobile communication channels when multipath propagation caused by a few dominant reflectors gives rise to (Doppler induced) linearly varying path delays. Doppler frequencies can be estimated blindly using cyclic statistics, and channel orders can be determined from rank properties of a received data matrix [422]. When *channel* (or Doppler) *diversity* is complemented by temporal or spatial diversity (available with oversampling or multiple antennas) blind estimators of LTV channels along with direct equalizers become available even with minimal (persistence-of-excitation) assumptions about the input and the bases [133]. Multivariate LTI ZFEs lend themselves to adaptive algorithms that provide fine tuning for possible model mismatch of the bases that capture only the nominal part of the rapidly fading channel (see [133] for a tutorial treatment and [422] for blind HOS LTV channel estimators and DFEs).

MIMO channel equalization is a major challenge in multiple-access wireless communications because multipath introduces multiple-access interference (MAI), which limits system capacity and bit-error-rate (BER) performance. Low-complexity CDMA systems able to cope with (perhaps unknown and time-varying) multipath and equipped with self-recovering capabilities

Unfortunately, there is no single measurement procedure appropriate for all TDE scenarios.

are most desirable because they are versatile in variable data rates and fading (e.g., mobile) environments. They also relieve the need for power-control and bandwidth-consuming training sequences. Linear suboptimum equalizers with training that either suppress MAI completely (a.k.a. zero-forcing (ZF) or decorrelating receivers) [250], or their MMSE [177] and minimum-output energy [423] counterparts, offer a compromise between the high-complexity ML solution [441] (that assumes knowledge of all system parameters) and the matched-filter (MF) multiuser demodulators that are not only known to suffer from near-far effects but also exhibit an error floor in their BER due to MAI. Additional approaches include multistage adaptive demodulators, DFEs [98], and spatial combiners (RAKE receivers that in fact are nothing but inverses of multivariate channels that are assumed to have been estimated) [323]. Frequency-selective multipath induces interchip interference. Simultaneous incorporation and mitigation of asynchronism and multipath effects were reported recently in [420] and [423] using a multirate equivalent discrete-time model (see also [458] and [459]). Blind approaches are well motivated when high-rate communication protocols entail small data packets (e.g., in distributed networks and wireless PCS prototypes) or when the propagation medium is rapidly varying (e.g., in large cells with considerable delays and high data rate time-varying wireless environments). Such self-recovering CDMA receivers were proposed recently in [420], [473], [383], [241], [25] and [419]. They capitalize on code diversity offered by the user code(s) and the received data but have relatively high complexity especially because they adopt subspace decomposition (via the singular-value decomposition (SVD)) of large matrices for signature waveform estimation. Inverse filtering criteria and recursive least-squares (RLS) and least mean-square (LMS) algorithms that include multipath were reported recently in [419] and [426] by viewing self-recovering CDMA demodulation as a blind beamforming problem. Interesting directions for mitigating MAI, asynchronism, and multipath by judicious code design at the transmitter include the blind Lagrange-Vandermonde CDMA transceivers of [352] and the emerging multicarrier CDMA systems, both of which turn frequency-selective multipath into flat fading [105, 471] (see also [349] for wavelet-based codes that produce graceful degradation, especially with oversaturated CDMA systems).

The plethora of propagation conditions and requirements for transmitter-receiver constraints in terms of complexity, transmit-energy, bandwidth, SNR, and performance specifications (MMSE or BER), offer numerous possibilities for statistical signal-processing algorithms. Digital communication systems generally entail man-made components and are a "paradise" for signal-processing research and development because they provide considerable flexibility to the SP designer. At the same time, SP algorithms must be tuned to the often strict specifications of communication systems and standards. A number of interesting directions open up for future research: (1) maximum exploitation of available information and communication constraints (*semi-blind* approaches along the lines of [56] and [147] offer promising directions for linear equalization); (2) if there is a choice for inducing diversity in the input, the channel, or the received output, it appears that input (i.e., transmit-) diversity in the form of short training sequences, modulation, codes, or filterbanks is to be preferred and optimal transceivers should be designed along the lines of [463] (see also [351]); (3) with increasing interest toward low--power communications and nonconstant modulus transmissions (e.g., OFDM or downlink CDMA in general) pre- or postmitigation of power-amplifier nonlinear distortions is necessary using training or self-recovering receivers (see [132] and references therein for steps in this direction); (4) convergence studies of adaptive CMA and DFE algorithms in realistic noise environments [90]; (5) performance analysis of channel estimators, especially when model perturbations due to synchronization effects and Doppler frequency drifts are present; (6) BER evaluation of ZF equalizers and experimental comparisons with the MSE equalizers in SISO, SIMO, and MIMO structures; (7) diversity techniques for blind identification of random coefficient models and performance comparisons with the basis expansion models using real data; (8) development and performance analysis of low-complexity blind multiuser equalizers; (9) joint design of equalizers with channel encoders and interleavers; (10) exploitation of network protocol structures from higher layers (e.g., ATM) for designing equalizers at the physical layer (see [11] along these lines); (11) balanced combination of the various possible diversity-inducing factors (e.g., codes, channels, fractional sampling, antennas) for blind channel estimation and equalization of general MIMO channels under minimal but realistic identifiability assumptions (see [373] for preliminary steps in this direction).

A WWW link to the author of the above section:
http://watt.seas.virginia.edu/~gg4w/

Non-Gaussian Processes

Ananthram Swami, ARL

The Gaussian distribution enjoys a central place in statistical signal processing; the Gaussian assumption is often justified by appealing to the Central Limit Theorem. The pdf is tractable, and algorithms derived under the Gaussian assumption are usually simple (linear/closed-form). In contrast, non-Gaussian signal processing typically involves nonlinear processing. In this section, we will look at recent trends in the modeling of non-Gaussian processes: finite-variance and infinite-variance processes, fractal point processes, and multiplicative noise models.

Finite Variance Models

Non-Gaussian data are encountered in various fields, such as agriculture (the Fisher Iris data) and economics (unemployment data) [6]; astronomy (sunspot data) and biology (Canadian lynx data) [322], music (average kurtosis of music has been steadily increasing!) [149], exploration seismology, radar, sonar, speech, and image and communication signals. Experimental measurements show that ambient noise is often significantly non-Gaussian, particularly in urban and radio channels [270] and underwater acoustic channels [45, 269]. In communication channels, multiple user interference is highly structured and non-Gaussian. Of course, nonlinearities usually lead to non-Gaussian outputs.

In general, the multivariate pdfs of non-Gaussian processes are intractable, and with few exceptions there are no general models. In the non-Gaussian context, linear techniques and second-order statistics are not merely suboptimal but are, in some cases, incapable of providing acceptable performance (e.g., multichannel processes and source separation). Since for many problems the Gaussian environment is the least favorable, exploitation of non-Gaussian features can lead to significantly improved performance, although this usually involves nonlinear signal processing [199].

Optimal estimation/detection may be feasible if the multivariate pdfs are analytically known (and tractable). If we have access *only to training data* (or if the signal is weak), the noise pdf can be estimated using various approaches (kernel density approaches, stabilized histogram estimators, type-based estimators, etc). Alternatively, one may use a parametric model, such as the Gaussian mixture model, whose parameters can be efficiently estimated via the EM algorithm [260, 273]. Depending upon the application, the (multivariate) pdf estimators may not be satisfactory unless sufficient data are available, and the resulting signal estimators may be highly nonlinear.

Since non-Gaussian processes are not completely characterized by their first and second-order moments, higher-order statistical descriptors, such as the higher-order moments and cumulants, and their Fourier transforms, the moment and cumulant spectra are required. Consequently, in the last two decades, a lot of attention has centered around HOS [264, 291, 292, 315, 320, 387]. HOS provide a parsimonious (but generally incomplete) characterization of the non-Gaussian process and are multidimensional statistics; for example, the fourth-order moment of a stationary random process is a function of three indices: $m_{4_x}(i,j,k) := E\{x(n)x(n+i)x(n+j)x(n+k)\}$. HOS have been used to provide tractable solutions to various "nonproblems" in signals and systems: non-Gaussianity, nonminimum phase, noncausality, nonlinearity, nonreversibility, nonadditivity, and nonstationarity. Sampling a continuous-time minimum-phase linear process usually renders the discrete-time process nonminimum phase (NMP) and NMP signals are encountered in frequency selective communication channels. Nonlinearities are encountered, for example, in high-power amplifiers in communication satellites operating near the saturation point, interactions in ocean waves, magnetic recording channels, and scattering phenomena in radar and sonar.

The history of HOS can be traced back to Fisher and the seminal work of statisticians in North America and Eastern Europe: Brillinger, Kolmogorov, Leonov, Rosenblatt, Shiryaev, Sinai, and Tukey. The interest of the signal processing and systems communities perked up in the early 1980s, due partly to a US-ONR funded initiative in non-Gaussian processes [449]. Subsequently, there have been five biannual international workshops and several special issues of the *Transactions on Signal Processing* devoted to the subject; a comprehensive bibliography may be found in [387]. The success of HOS-based methods depends clearly on the amount of non-Gaussianity and nonlinearity of the underlying processes and models; hence, tests of Gaussianity and linearity are important. HOS-based methods typically entail an increase in dimensionality, computational load, and statistical variance of the sample estimators. The potential need for longer data records also cautions one to test for stationarity.

The cumulants of a stationary-random process (or of a random vector) can be represented either as tensors or as m-D matrices. Several results in the scalar case can be generalized to the vector case by replacing scalar multiplications with Kronecker products [386]. But, notions related to rank, eigenvector decomposition, diagonalization etc are largely unsolved, but some interesting results may be found in [75, 76] where relationships with the theory of homogeneous multivariable polynomials are established. Even when slices or projections are used, the resulting 2D matrices are generally not symmetric, so that the general nonsymmetric eigenvector problem is involved.

Other relatively virgin areas of HOS research include: extensions of HOS to nonstationary processes, truly adaptive algorithms, efficient algorithms for nonlinear system analysis, robust estimation algorithms, applications in point processes, synchronization and multiuser problems in communications, and performance analyses of algorithms.

Infinite Variance Models

Heavy-tailed non-Gaussian processes, particularly the Gaussian mixture model and the Cauchy r.v., have long been used to develop and test signal-processing techniques that are robust to *impulsive* noise [194]. The Cauchy r.v. has infinite variance and is a special case of an *alpha-stable* r.v. It is easy to create a Cauchy r.v. as the ratio of two (possibly correlated) Gaussian random variables, and Feller [108] shows how the Cauchy r.v. arises in an example with rotating mirrors. Stable r.v.'s result from generalized central limit theorems [469], and are characterized by four parameters: scale, location, the characteristic exponent (α), and a skewness index (β) [293, 348, 469]. The exponent satisfies $0 < \alpha \leq 2$, where $\alpha = 2$ corresponds to the Gaussian and $\alpha = 1$ corresponds to the Cauchy. The skewness index satisfies $-1 \leq \beta \leq 1$, with $\beta = 0$ indicating symmetry. The non-Gaussian stable processes are characterized by their infinite variance; indeed, $E|x|^r = \infty$ if $r \geq \alpha$. Hence, one must use lower-order ($r < \alpha$) rather than higher-order moments to study these processes [293]. The non-Gaussian alpha-stable process may be considered impulsive, and at least in the symmetric case, there are connections with the Middleton B model [270] that remain to be fully explored. A survey of parameter estimation techniques may be found in [293]. Although the lack of a closed-form pdf, except in special cases, has made analysis difficult, numerically stable ML estimators are discussed in [294].

Despite the "infinite variance" of these processes, they have found applications in areas such as astronomy (gravitational fields) [469], econometrics (income distributions) [255], modeling of radar data [294, 418], and Ethernet traffic data [447]; several applications in the physical sciences are discussed in [181].

Good data windows give a much less biased estimate of the spectrum than the periodogram.

In the context of linear symmetric stable processes, identifiability of mixed-phase ARMA models was established in [252]; in [389], it was shown that normalized HOS could be used to estimate these parameters, extending the earlier work of Davis-Resnick, and Mikosch et al. Alpha-stable processes may be used to model very impulsive noise, and they can be suppressed (to some extent) using the usual tools of ranks, weighted medians, and order statistics. Locally optimum detectors may also be used [293]; another alternative is to use data-adaptive nonlinearities [388]. Impulsive noise suppression is particularly important for compression and dynamic range reduction. Model validation is important, and several ideas are discussed in [294] and [295]. Differentiating between nonlinearity and nonstationarity may be more difficult than in the finite variance case [87].

Fractal-Point Processes

Recent papers by Willinger et al. [447] show that Ethernet traffic data are self-similar and alpha-stable, so that the standard Poisson model and conventional analysis [44] are no longer adequate. The self-similarity naturally leads to wavelet-based analysis, see, e.g., [2]. Semi-parametric techniques, based on log-periodogram regression, are proposed for estimating the fractal dimension in [277], and fractal-point process models have been proposed in [14, 219, 344], but several interesting analysis and synthesis problems remain open.

Time-frequency distributions have recently found applications to radar and sensor-array processing.

Multiplicative Noise

Just as nonlinearities give rise to non-Gaussianity, they also lead to nonadditive noise. For example, if the received signal, $x(k) = s(k) + w(k)$, passes through a zero-memory nonlinearity (ZMNL) (the receiver amplifier), and if the signal, $s(k)$, is weak relative to the noise, $w(k)$, the observed signal can be written as $y(k) = s(k)[\theta + g(k)] + v(k)$, where θ is a parameter of the ZMNL. The noises $v(k)$ and $g(k)$ are now, in general, correlated, non-Gaussian, and have nonzero means. Multiplicative noise is encountered in speckle imagery [40], fading channels [141, 324], underwater acoustics [99], and lidar and radar [29].

An interesting approach is taken in [318], where the signal is a harmonic (i.e., line components), the noises are modeled as being *nonrandom*, and error bounds on the frequency estimates are developed. In the case of a harmonic signal, and stationary noises, conventional spectral analysis of the data ("cyclic mean") or of the squared data ("cyclic variance") have been proposed and analyzed in [467]; fourth-order cumulants were used in [99]. Cramer-Rao bounds were established in [118], [466], and [385]. Some parameter estimation problems are discussed in [384]. Note that all these papers assume that the additive and multiplicative noises are independent.

The detection of the weak signal, $s(k)$, from the observed data, $y(k) = \theta s(k)[1 + g(k)] + v(k)$, has been considered both for the random and nonrandom cases in [378], [31], [9],and [126], but much more work remains to be done (e.g., temporally correlated noise, and "nonweak" signals).

A WWW link to more information on non-Gaussian processes and on higher order statistics:

http://www.comm.uni-bremen.de/HOSHOME

System Identification and Tests for Non-Gaussianity and Linearity

Jitendra K. Tugnait, Auburn University

System Identification

System identification is the field of mathematical modeling of systems and signals from experimental data [377]. In signal-processing applications, system-identification methods are used for linear prediction, adaptive filtering, noise and interference cancellation, parametric spectral estimation, inverse and forward modeling, and numerous other objectives. In systems and control applications, models obtained by system-identification approaches are used for controller design, system simulation, and prediction. The system-identification methods are applicable to both cases, when input-output data of the system under investigation are available as well as when one only has the system output measurements (time-series analysis).

The area of system identification has been an active research area for past 30 to 40 years. Most of the focus has been on SISO linear models and on scalar stationary time series, mainly because of its wide applicability and partly because of its analytical tractability. The available methods have been well analyzed in excellent texts such as [377], [244], and [245]. The text by Widrow and Stearns contains a wide range of applications [450]. A typical approach is to choose a model structure with fixed (order) order, e.g., state-space models with known state dimension, autoregressive moving average (ARMA) models with known AR and MA model orders, MA models, etc., and then turn the system-identification problem into one of parameter estimation. The choice of the model structure is dictated by the intended application. For instance, [450] favors using MA models for adaptive signal processing in applications where the stability of the fitted model is paramount. A large number of methods exist for parameter estimation including least-squares, prediction-error minimization, ML, instrumental variable methods, output error minimization and others [244, 245, 377]. A complete system-identification methodology should, however, include an iterative process of model structure determination, parameter estimation, and model validation [244, 377].

The problem of MIMO system identification has proved to be more complex. Unlike the scalar (SISO) case, for MIMO ARMA models with known orders the representation of system output measurements in terms of AR and MA matrix coefficients is not necessarily unique. This lack of uniqueness can lead to ill-conditioning in parameter estimation. A possible solution is to use a canonical parameterization requiring knowledge of certain structure indices that are difficult to determine in practice [152]. An interesting solution to MIMO model identification (including ARMA parameter estimation) via a subspace-based realization approach using state-space formulation has recently been proposed [438]. It is applicable to both input-output ("deterministic") models with noisy output measurements as well as to multivariate time series ("stochastic" models).

Systems and techniques not captured by the above formulations (stationary linear time series, linear systems with noisy output measurements but noise-free input measurements, time-domain approaches) have also received considerable attention in recent years. Some representative examples include:

▲ Errors-in-variables models: These are models where both input and output measurements are noisy. Higher-order statistics have been used in [135], [431], and [434]; second-order statistics and subspace instrumental variable methods in [381]; and cyclic and/or higher-order spectral analysis in [135] and 415].

▲ Nonlinear systems: Volterra systems have attracted considerable attention [281]. More general Hammerstein systems have been considered in [327]. Time-series models have been treated in detail in [412].

▲ Frequency-domain approaches have been investigated in [432], [433],and [361]. Approaches of [432] and [433] do not require explicit noise modeling.

▲ For time-series models higher-order statistics-based approaches have attracted considerable attention due to their ability to (blindly) identify nonminimum phase and/or noncausal models [292, 435].

Gaussianity and Linearity Tests

Linear parametric models of stationary random processes, whether signal or noise, have been found to be useful in a wide variety of signal-processing tasks such as signal detection, estimation, filtering, and classification, and in a wide variety of applications such as digital communications, automatic control, radar and sonar, and other engineering disciplines and sciences. Parsimonious parametric models such as AR, MA, ARMA or state-space, as opposed to impulse-response modeling, have been popular together with the assumption of Gaussianity of the data. Linear Gaussian models have long been dominant both for signals as well as for noise processes. Assumption of Gaussianity allows implementation of statistically efficient parameter estimators such as ML estimators. A stationary Gaussian process is completely characterized by its second-order statistics (autocorrelation function or equivalently, its power spectral density—PSD) and it can always be represented by a linear process. Since the PSD depends only on the magnitude of the underlying transfer function, it does not yield information about the phase of the transfer function. Determination of the true phase characteristic is crucial in several applications such as seismic deconvolution and blind equalization of digital communications channels. Use of higher-order statistics allows one to uniquely identify nonminimum-phase parametric models. Higher-order cumulants of Gaussian processes vanish; hence, if the data are stationary Gaussian, a minimum-phase (or maximum-phase) model is the "best" that one can estimate. Given these facts, it has been of some interest to investigate the nature of the given signal: whether it is a Gaussian process and, if it is non-Gaussian, whether it is a linear process.

Gaussianity Tests

Several tests have been devised on the basis of the fact that the higher-order cumulant spectra [43] of Gaussian processes vanish. One of the earliest tests based upon testing of the signal bispectrum is given in [329]. Hinich [168] has simplified the test of [329] by using the known expression for the asymptotic covariance of the bispectrum estimators. Notice that a vanishing bispectrum does not necessarily imply that the underlying signal is Gaussian; it may result from the fact that the signal is non-Gaussian with zero bispectrum. Therefore, a next logical step would be to test for vanishing trispectrum of the record. This has been done in [272] using the approach of [168]; extensions of [329] are too complicated. Computationally simpler tests using "integrated polyspectrum" of the data have been proposed in [429]. The integrated polyspectrum (bispectrum or trispectrum) is computed as cross-power spectrum and it is zero for Gaussian processes. Alternatively, one may test higher-order cumulant functions of the data in time-domain. This has been done in [136].

Other tests that do not rely on higher-order cumulant spectra of the data may be found in [412].

Linearity Tests

For a stationary time series one can define a normalized bispectrum, called bicoherence or skewness function [292], which turns out to be a (nonzero) constant for all bifrequencies if the signal is linear non-Gaussian with nonvanishing bispectrum. This property has been used by Subba Rao and Gabr [329] to design a statistical test for linearity. Hinich [168] has "simplified" the test of [329]. Notice that this test is useless if the signal is non-Gaussian with zero bispectrum. Therefore, a next logical step would be to test a normalized trispectrum (tricoherence function). This has been done in [272] using the approach of [168]; extensions of [329] are too complicated. The approaches of [329] and [168] will fail if the data are noisy. A modification to [329] is presented in [428] when additive Gaussian noise is present. Finally, other tests that do not rely on higher-order cumulant spectra of the record may be found in [412].

Advanced-Sensor Signal Processing

Arye Nehorai, The University of Illinois at Chicago

Over the last several years, advanced sensors have been introduced and combined with statistical sensor-array-processing methods. These sensors exploit more physical information and are orders of magnitude more sensitive than traditional sensors. Their use has improved the performance of current systems, increased

Just as nonlinearities give rise to non-Gaussianity, they also lead to nonadditive noise.

the scope of signal-processing methods, and created entirely new applications. Four examples are presented in the following.

Electromagnetic Vector Sensors

Each element in an electromagnetic vector sensor, introduced in [284] and [287] for estimating the direction and polarization of electromagnetic sources, measures all six electromagnetic field components at a single point. In contrast, conventional antennas measure a single component of the electric field. Vector sensors are commercially available and actively researched. EMC Baden Ltd. in Baden, Switzerland, manufactures them for a 75 Hz-30 MHz frequency range, and Flam and Russell, Inc., in Horsham, Pennsylvania, for 2 MHz-30 MHz. Lincoln Labs at MIT has performed preliminary localization tests with the vector sensors of Flam and Russell, Inc. [156]. Other research on sensor development is reported in [189] and [190].

Electromagnetic vector sensors are sensitive to both the direction and polarization information in the incoming waves. The polarization provides a crucial criterion for distinguishing and isolating signals that may otherwise overlap in conventional scalar-sensor arrays. When a single vector sensor is used for direction finding, it has the following advantages and capabilities:

▲ Direction estimation in 3D is possible with sensors occupying very little space

▲ Estimation of the directions and polarization ellipses of up to three sources [172, 175] is possible

▲ Resolution of very closely spaced (even co-incident) sources based on polarization differences is possible

▲ One is able to process wideband signals in the same way as narrowband signals

▲ One can handle sources with either single or dual-message signals

▲ Vector sensors have isotropic response

▲ There is no need for location calibration and time synchronization as there is among different sensor elements of a conventional spatial array

Some of these advantages result from the fact that no time delays are used. In contrast, conventional scalar-sensor methods require a 2D array for direction finding in a 3D space, need accurate location calibration and time synchronization, and require much higher computational cost to process wideband rather than narrowband signals.

The optimum accuracy of source parameter estimation for vector-sensor arrays is analyzed in [284] and [287] in terms of Cramer-Rao bounds (absolute limits on the accuracy of the class of unbiased parameter estimators). Quality measures are defined for estimating direction and orientation in 3D space, including mean-square angular error and covariance of vector-angular error. Lower bounds on these measures give concrete results on expected performance.

A fast algorithm for direction finding using an electromagnetic vector sensor has been proposed in [284] and [287]. Inspired by the Poynting theorem, it forms the cross-product of the electric field vector with the complex conjugate of the magnetic vector and averages over time. The asymptotic performance under general conditions is shown to be close to optimum.

A minimum-noise-variance beamformer for interference cancellation employing a single electromagnetic vector sensor has been proposed in [283]. It assumes that the direction and polarization of the source are known. This enables suppression of uncorrelated interference, even if it comes from the same direction as the source, based on polarization as well as location differences. Analysis of the signal to interference-plus-noise ratio showed the beamformer to be very effective, particularly when the signal and interference are differently polarized.

An array of spatially distributed vector sensors can additionally exploit time delays among the sensors. In general, arrays of vector sensors can achieve better performance than scalar-sensor arrays, while occupying less space. Alternatively, they can be spaced further apart to increase aperture and hence performance without introducing ambiguities [390, 453]. Some high-resolution direction-finding algorithms have been developed for electromagnetic vector-sensor arrays [171, 173, 230]. Preliminary results on the application of vector sensors to communication problems can be found in [470] and [65].

Acoustic Vector Sensors

Acoustic vector sensors measure the acoustic pressure and all three components of the acoustic particle velocity vector at a given point; standard methods measure only the pressure. The use of acoustic vector sensors for array processing has been proposed in [285] and [286]. These references introduced measurement models, derived fast direction-estimation algorithms, and presented general Cramer-Rao bounds on direction parameters. These developments have coincided with a surge of interest in particle-velocity sensors [27] and improvements in fabrication techniques [97] to make vector-sensor arrays a practical reality [290].

Since a vector sensor extracts directional information directly from the structure of the velocity field, it can, for example, locate up to two sources in 3D space [175]. By making use of this extra information, arrays of vector sensors improve source localization accuracy while using smaller array apertures.

Beamforming and Capon direction-estimation procedures with acoustic vector-sensor arrays have been developed in [160]. It was shown that the extra velocity-field information removes all ambiguities such as grating lobes. This allows the use of simple structures for which fast direction estimation algorithms exist, e.g., a uniform linear array, to determine both the azimuth and elevation of a source. It also means that spatially undersampled arrays of vector sensors can be used to increase aperture and hence performance. A fast estimation algorithm that makes uses of this property was developed in [455] and another algorithm for arbitrary array shapes appears in [454].

In [158] the effect of sensor placement on the direction estimation performance of an array of acoustic vector sensors has been considered. Using the Cramer-Rao bound on the parameters of a single source, conditions were derived that minimize the lower bound on the asymptotic mean-square angular error, and that it is isotropic. The increase in estimation accuracy obtained by vector sensors is greatest for linear or planar arrays (as opposed to 3D), small number of sensors, and low SNRs. By exploiting velocity and pressure information, any vector-sensor array, a popular linear array, can be used to unambiguously determine both the azimuth and elevation of a single source.

Vector sensors have been successfully applied in other areas such as hull-mounted applications, where they overcome serious problems in detecting low-frequency emitting targets [159]. (At low frequency the vessel's hull is acoustically flexible, leading to a very low pressure signal but a strong velocity signal.)

Chemical Sensors

Chemical sensors are useful for detecting explosives, drugs, and leakage of hazardous chemicals, and for monitoring the environment. They are manufactured by companies such as Cyrano Sciences, Inc.; Science Applications International Corporation (SAIC); and Jaycor in San Diego, California. Array-processing techniques using chemical sensors have been proposed in [184], [288], and [319]. Compared with animal chemoreception, these techniques have the advantage that they share information and can be optimized.

Methods for detecting and localizing vapor-emitting sources were developed in [288]. Based on the diffusion equation, distributions of vapor concentration in time and space were derived for various environments. The results were used to develop statistical models of the array measurements. Maximum-likelihood estimates and general-likelihood ratio tests were derived to estimate the unknown source parameters and detect the existence of a source. The performance was analyzed using Cramer-Rao bounds and probabilities of detection and false alarm.

Employment of moving sensors was proposed in [319]. A single moving sensor can achieve the task of multiple stationary sensors by taking measurements at different locations and times. Its path can be planned in real time to optimize a performance criterion. The criterion used in [319] was to reduce the expected location error, which was accomplished by moving the sensor in directions opposite to the gradients of the Cramer-Rao bound.

Monitoring of disposal sites on the ocean floor using chemical sensor arrays was considered in [184]. Such sites have been suggested as suitable for the relocation of dredge materials from harbors and shipping channels, where their buildup has a detrimental impact upon economy and military security. Algorithms for detecting possible release of pollutants near these sites were developed, and their performance was analyzed. The results were used to optimally design arrays with respect to the numbers of sensors and time samples as well as sensor locations.

Superconducting Quantum Interference Devices (SQUIDs)

SQUIDS are the most sensitive detectors of magnetic flux currently available. They find broad application, from measurements of magnetic fields induced by brain activity to nondestructive evaluation of materials and the location of underground objects and structures. Their most important commercial use is in magnetoencephalography (MEG). MEG is concerned with mapping electrical activity in the brain by measuring the induced external magnetic field [151]. MEG sensor arrays measure extracranial magnetic fields of only a few hundred femtoTesla—a billion times smaller than the Earth's steady magnetic field. Together with electroencephalography (EEG), which measures electric potentials on the scalp, MEG has emerged as a powerful noninvasive tool for the localization and tracking of electrical sources in the brain. The solution to this problem is of great importance in the diagnosis and evaluation of various brain disorders such as epilepsy, and for furthering understanding of brain function.

In [274] the MUSIC algorithm was applied for localizing brain sources modeled as current dipoles. Maximum-likelihood techniques have been developed in [93] to account for unknown spatially correlated noise, predominantly due to sporadic background activity in neurons. The optimization of MEG sensor arrays to minimize the mean-square error of dipole location estimates has been proposed in [176]. SQUIDs have opened other new applications of signal processing, such as detecting the wake of a ship using an airborne system [282].

In conclusion, it is expected that the use of novel sensors will continue to be a source of new applications and further developments in signal processing.

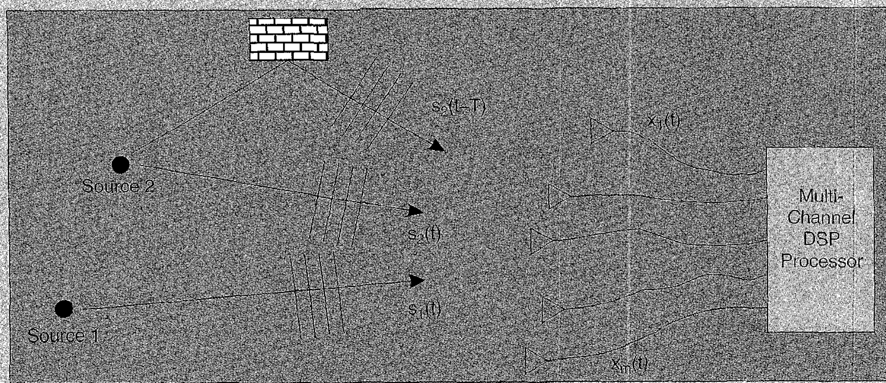
A WWW link to the author of the above section: http://www.eecs.uic.edu/~nehorai/

Sensor-Array Signal Processing

A. Lee Swindlehurst, Brigham Young University

The processing and manipulation of data received by a spatially distributed array of sensors has been an active area of research in the signal-processing community for well over

30 years. The long-lasting attention devoted to this area can be traced to the large number of applications where data is collected in both space and time. Figure 5 depicts a generic scenario in which energy (possibly acoustic, electromagnetic, seismic, etc.) from two sources is received by a sensor array. There are a number of possible objectives of such a system, with the most important being:



▲ 5. A Generic array signal-processing scenario.

▲ *Source Localization*—determine the azimuth and elevation angles to the sources, and possibly the range to the sources as well (if they are located in the near-field of the array); information on source velocity can be obtained by measuring frequency shifts, or angle and range rates of change.

▲ *Source Separation*—determine the signal waveforms transmitted by each source; the fact that the energy from each source arrives from different directions allows these waveforms to be separated even if they overlap in time and frequency.

▲ *Channel Estimation*—determine the space-time propagation effects between the sources and the array; estimate where reflections occur or how much the transmitted signal is spread in time and angle.

Which of the above three objectives is most important depends on the application. In active radar and sonar, the received waveforms are approximately scaled and delayed versions of a known signal, so it is the location (and motion) of the sources that is most important. In a communications system, it is the information-bearing waveform and not the location of the sources that is critical. For seismic applications, the source signals arise from explosive charges. The received energy is used to characterize the propagation channel, which in this case provides information about the structure of the ground.

To be more precise, and to aid the discussion that follows, we introduce some simple mathematical notation. Referring to Fig. 5, in the simplest case the sources and array lie in the same plane, and the sources are far enough from the array so that the arriving signals have planar wavefronts. For this case, if the signals are assumed to be "narrowband" and there are a total of d signals, then the output of array element p is given by

$$x_p(t) = \sum_{k=1}^{d} a_p(\theta_k) s_k(t) \tag{9}$$

where θ_k represents the DOA of the wavefront from source k, and $a_p(\theta_k)$ is the (complex) response of element p to a signal arriving from that direction. In the general case, we would treat the reflection of source 2's signal in Fig. 5 as a separate term in the above sum; i.e., we would let $s_3(t) = s_2(t-T)$. The outputs of an array of m elements can be stacked in a vector, as follows:

$$\mathbf{x}(t) \begin{bmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{bmatrix} = \sum_{k=1}^{d} \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) \tag{10}$$

where $\mathbf{a}(\theta_k) = [a_1(\theta_k) \cdots a_m(\theta_k)]^T$ denotes the vector array response, and we have added a term, $\mathbf{n}(t)$, to account for unmodeled measurement noise and interference. Using this notation, we can give concrete examples of the three objectives listed above. For example, in source localization, we would use samples of the array output $\mathbf{x}(t_1), \cdots, \mathbf{x}(t_N)$ to estimate the DOAs $\theta_1, \cdots, \theta_d$ of the sources. Source separation involves extracting samples of one (or more) of the signals, $s_k(t_1), \cdots, s_k(t_N)$, from the array data. In channel estimation, we may have $s_k(t) = \alpha_k s(t - T_k)$, in which case we are interested in the amplitudes, α_k, and delays, T_k, of the various arrivals (perhaps as well as the DOAs).

Early research in sensor-array signal processing, conducted mainly in the 1960s and 70s, was based on the observation that, if the array is composed of identical uniformly spaced elements (i.e., a uniform linear array, or ULA for short), then a direct analogy exists with temporal sampling. The array elements perform a uniform (one-dimensional) spatial sampling of the wavefield, and the spacing between elements determines what spatial frequencies can be uniquely represented. Signals whose wavefronts are nearly parallel to the ULA ($\theta \approx 0°$) have low spatial frequencies; as $|\theta| \to 90°$ increases, the spatial frequency of the signal increases and reaches a maximum at $\theta = \pm 90°$. A spatial version of the Nyquist criterion states that the array response vector $\mathbf{a}(\theta)$ of a ULA is unique provided that the elements of the array are separated by no more than one-half the wavelength of the signal. Using this analogy with temporal sampling, it is possible to design spatial filters that pass signals with certain spatial frequencies (i.e., that arrive from certain directions) and attenuate others. However, unlike temporal filtering, it is usually not known *a priori* what spatial frequencies are occupied by signals of interest. Thus, the emphasis was on data *adaptive* methods that estimated the directions to the desired user(s) prior to computing the appropriate spatial filter coefficients. Such algorithms are usually referred to as adaptive *beamformers*, since the desired spatial frequency response is a narrow beam "steered" towards the source of interest.

The problem of estimating the source directions can also be addressed using this analogy with temporal sampling. A point source located in the far field of a ULA at DOA θ produces the following output at sensor k:

$$x_k(t) = e^{-2\pi\delta(k-1)\sin\theta/\lambda} + n_k(t) \quad (11)$$

where δ is the distance between adjacent sensors and λ is the wavelength of the signal. Viewed as a function of k, the vector of samples from the array is seen to be a complex exponential in noise. With multiple sources, the problem of DOA estimation is seen to be equivalent to the classical spectral analysis problem of determining the frequencies of a collection of sinusoids in noise. This connection has led to numerous "cross-over" techniques, the most popular of which are those based on simple Fourier analysis with windows (Bartlett, Hamming, Chebyshev, etc.) used to control resolution and sidelobe levels. Unlike their counterparts in temporal frequency estimation, the resolution of these methods cannot be increased by collecting more data from the array; the ability to resolve sources spaced closely in θ is limited by the aperture of the array, which is typically fixed. In addition, the fixed aperture can produce large sidelobes in the spatial frequency spectrum that lead to inconsistent DOA estimates when more than one source is present. To overcome some of these deficiencies, techniques based on maximum entropy, autoregressive modeling, and linear prediction were proposed, with mixed success.

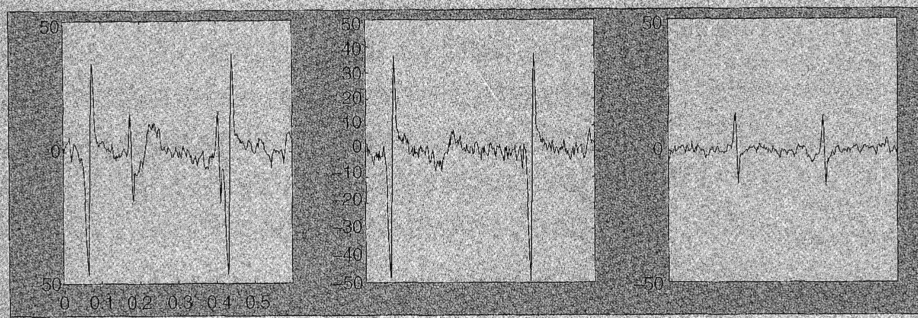
A dramatic shift in emphasis in sensor-array signal processing occurred during the 1980s with the introduction of the so-called *subspace*-based techniques. These methods are based on the observation that, if the number of sensors is strictly greater than the number of sources, the signal component of the array data lies in a low-rank subspace. Under certain conditions, this subspace uniquely identifies the DOAs of the signals and can be determined quite accurately using, for example, a numerically stable singular value decomposition. A large number of parametric estimators have been developed based on the subspace idea. These techniques enjoy a number of advantages over earlier methods, including statistical consistency and very high resolution. Numerous extensions to the simple model outlined above have been considered, including generalizations to wideband or diversely polarized signals, techniques for handling correlated or perfectly coherent signals, or arrays whose response, $\mathbf{a}(\theta)$, is not precisely known or calibrated.

The main driving force behind research in array signal processing in the 1980s was provided by military applications, primarily in radar and sonar. In these applications, the sources are often noncooperative and little may be known about the signals they generate. In such situations, discrimination based on only the spatial properties of the received signals is necessary, which in turn requires that the response of the array be accurately calibrated with respect to θ. The sensitivity of subspace based methods to array calibration errors limited their usefulness to some degree, particularly in underwater environments where the propagation medium is severely nonuniform. As military funding has waned, and as the field of personal wireless communications has emerged, interest in applications of array signal processing to communications systems has blossomed in the past few years. The use of multiple antennas at the base station of a wireless network offers a processing gain that can increase base station range and improve coverage. By exploiting the spatial selectivity of an antenna array, co-channel interference may be reduced, which in turn can be traded for increased system capacity. In addition, communication channels can be multiplexed in the spatial dimension just as in the frequency and time dimensions. This is often referred to as spatial-division multiple access (SDMA).

The further integration of computational wave-equation solutions and signal-processing techniques will pose many challenges and rewards in the future.

A distinguishing aspect of using antenna arrays in communications applications is that, due to the cooperative nature of such systems, significant information about the source signals is available and can be exploited for spatial processing. For example, it may be known that training sequences are present in the data, or that the signal is digitally modulated with a known symbol constellation and pulse-shaping filter, or that the signal has a constant amplitude envelope, etc. Each of these properties can be used by a system employing multiple antennas to achieve source separation without the need for explicit array calibration data. Algorithms that use this approach are referred to as "blind" source separation methods (see the sections by Cardoso and Tong). Such techniques can also be extended to perform blind equalization of propagation channels with significant delay spread. A breakthrough in this area came in the early 1990s, when it was shown that if a pulse-amplitude modulated signal is received by an array of antennas, then the channel can be identified using only second-order cyclostationary statistics.

Although blind methods can eliminate the need for calibration information, significant performance improvement can be achieved if reasonably accurate calibration is available. Techniques that exploit both the spatial



▲ 6. Output of one of eight EKG electrodes (left); reconstruction of the mother's EKG contribution (middle), and of the fetus' EKG contribution (right) using all eight electrodes.

and temporal properties of the received signals are thus of high interest at the present time. Joint space-time processing in radar and sonar applications is also receiving added attention, as the speed and throughput of multichannel DSP processors continues to increase. Further advances in computing power will bring other difficult array signal-processing problems to the forefront, such as source localization and separation for wideband signals, parameter estimation for sources with distributed spatial spectra, and matched-field processing for sonar applications.

There is a large body of published literature in the area of sensor-array signal processing available to the interested reader. The references listed below are good starting points because of their tutorial nature and their extensive bibliographies:

▲ General books: [161], [311], [185]
▲ Connections with spectral analysis: [197], [382]
▲ Adaptive beamforming [77], [439]
▲ Applications to radar systems [104], [163]
▲ Subspace methods [299], [34], [214]
▲ Applications in communications [298], [34], [306]

A WWW link to the author of the above section:
http://www.ee.byu.edu/~swindle

Blind Separation of Sources

Jean-François Cardoso, École Nationale Supérieure des Télécommunications (ENST)

Objectives

Source Separation

Source separation consists of recovering a set of "source signals" from the observation of several mixtures of these signals. This problem typically arises when the available signals are obtained at the output of an *array* of sensors that temporally and spatially samples signals emitted at different locations in space. In general, each sensor receives a mixture of all the source signals: if there are fewer sources than sensors, the received mixture of signals is (in general) linearly invertible: this is the case of *spatial diversity* discussed in the following section by L. Tong.

Figure 6 shows an example of separation of electrocardiography (EKG) signals. The left panel shows the output of one EKG electrode located on the abdomen of a pregnant woman: the fetus heart beat cannot be easily distinguished. The data set [88] contains the outputs of seven other sensors placed on the mothers chest and abdomen. A source-separation technique allows the contributions from the mother (middle) and from the fetus (right panel) to be separated.

Blind Source Separation

Exploiting an array of sensors to focus on a particular source signal while rejecting other "interferers" is a standard task in array processing (see the sections by Swindlehurst and Krolik in this article). The *blind source separation* (BSS) problem consists in recovering all the sources without using prior information about the channels; i.e., about the transfer function between the sources and the sensors. BSS is an "output-only" technique in the sense that neither source signals nor training sequences are available: all of the available information is contained in the observed data themselves. Two seminal papers on this topic are those by Jutten et al. [188] and Comon [74].

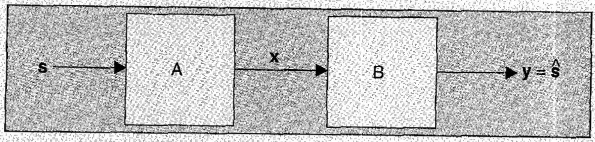
The major strength of the blind approach to source separation stems precisely from the fact that a precise model of the underlying physical phenomena, e.g., wave generation, propagation, and transduction, is not required. Thus, for example, BSS can be applied to uncalibrated arrays in situations where calibration is difficult or impossible or when physical modeling is overly complicated or unreliable.

The basic idea of BSS is that one makes up for the lack of information about the channels by assuming that the source signals are (statistically) *independent*. Statistical independence is a relatively strong assumption but it is plausible in many contexts because it arises from a lack of physical relationship between the various sources.

The simplest source separation model assumes an n-sensor array receiving signals, $x_1(t),\ldots,x_n(t)$, from as many sources, $s_1(t),\ldots,s_n(t)$, and an instantaneous and noise-free mixture:

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & a_{ij} & \vdots \\ a_{1n} & \cdots & a_{nn} \end{bmatrix} \times \begin{bmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{bmatrix}.$$

The mixture coefficients, a_{ij}, can be collected in an $n \times n$ "mixing matrix," $\mathbf{A}$. Collecting the n source signals and the n array outputs into $n \times 1$ column vectors, the BSS model reads more concisely as $\mathbf{x(t)} = \mathbf{As}$.



▲ *7. Mixing and separating. Source signals: **s**; array output: **x**, estimated sources: **y**.*

If the mixing matrix, **A**, were known, source signals could easily be recovered by direct inversion: $\mathbf{s}(\mathbf{t}) = \mathbf{A}^{-1}\mathbf{x}(\mathbf{t})$. Conversely, if the source signals were available, the mixing matrix could be easily estimated by a simple input-output identification procedure. The challenge of BSS is that of "double blindness": neither $\mathbf{s}(\mathbf{t})$ nor **A** is available in the model $\mathbf{x} = \mathbf{A}\mathbf{s}(\mathbf{t})$.

Algorithms for the blind separation of sources try to determine a *separating matrix*, **B**, to un-mix the observation vector into $\mathbf{y} = \mathbf{B}\mathbf{x}$ (see Fig. 7). Ideally, the separating matrix should approximate the inverse of the mixing matrix. The next section outlines some statistical ideas for accomplishing BSS.

Principles

Before listing a few of the ideas behind BSS algorithms, it is instructive to explain why the simplest idea—finding a separating matrix that makes the outputs uncorrelated—does *not* work. The reason is that decorrelation is a symmetric property: $\text{Corr}(y_i, y_j) = 0$ also implies that $\text{Corr}(y_j, y_i) = 0$. Therefore, there are only as many decorrelation conditions as pairs of sources, namely $n(n-1)/2$, which is about half of the constraints needed to determine the n^2 entries of a separating matrix. Thus, pairwise correlations (second-order information) are not sufficient to solve the BSS problem: it is necessary to express statistical independence in a stronger sense.

That decorrelation is not sufficient has an important consequence. Recall that for Gaussian variables, decorrelation implies independence; it follows that Gaussian sources cannot be blindly separated because their independence boils down to pairwise decorrelation. Therefore, some non-Gaussianity is needed to achieve BSS. One may see BSS as the art of exploiting the non-Gaussianity of the signals and measurements.

Some ideas for deriving source separation algorithms are to adjust **B** in such a way that:

▲ The probability distribution of **y** is as close as possible to some prespecified distribution of independent components, or

▲ The outputs, $y_1, \ldots, y_n$, are as independent as possible, or

▲ The outputs are decorrelated *and* as non-Gaussian as possible or as low-entropic as possible.

All these objective functions can be derived from the application of the ML principle [50] under various assumptions. The first objective function results from fitting **x** to the linear model $\mathbf{x} = \mathbf{A}\mathbf{s}$ where **A** is unknown and **s** has a fixed (hypothetical) distribution of non-Gaussian independent components; the second objective function results from fitting to the model $\mathbf{x} = \mathbf{A}\mathbf{s}$ with respect to both the unknown system, **A**, *and* to the distribution of the independent sources, **s**; the third objective function arises when one imposes the addition constraint that the recovered signals should be uncorrelated. This can be understood as follows: summing independent random variables "tends" to produce a more Gaussian result (think of the central limit theorem) so that driving **y** away from Gaussianity may be thought of as a way of recovering the original source signals.

The ML principle gives rise to measures of independence, non-Gaussianity and entropy (as listed above), which are based on information-theoretic quantities. Because they may be difficult to manipulate, one often re-

Estimating signals under various diversity receptions is currently an active research area.

sorts in practice to more tractable approximations, e.g., high-order correlations (triple, quadruple correlations), high-order cumulants, or pair-wise correlations between *nonlinear* functions of $y_1, \ldots, y_n$.

Perspectives

Beyond the simple cases described here, blind source separation has been applied to much more general models such as noisy observations, complex signals, nonsquare mixtures, and convolutive mixtures [50]. The latter extension often brings the BSS problem close to the problem of channel equalization and of system identification (see the sections by Giannakis and Tugnait in this article).

Another interesting extension is to consider the case where the BSS model does *not* hold. In this case BSS may be implemented as a data exploration technique for which one is interested in finding the linear transformation of a random vector, **x**, into $\mathbf{y} = \mathbf{B}\mathbf{x}$ such that the components of **y** are as independent as possible. BSS is then seen as a device for independent component analysis (ICA), which can complement principal component analysis (PCA).

WWW links relevant to the above section:

▲ The author's web page:
http://sig.enst.fr/~cardoso/stuff.html

▲ WWW links to independent components analysis (ICA):

—ICA research group at the Helsinki University of Technology: http://www.cis.hut.fi/projects/ica/

—Laboratory for Open Information Systems at the Riken Institute:
http://www.bip.riken.go.jp/open/Welcome.html

—Computational Neurobiology Lab at the Salk Institute: http://www.cnl.salk.edu/~tewon/ica sub cnl.html

▲ Miscellaneous WWW links of relevance:
http://sound.media.mit.edu/~paris/ica.html
http://www.bmc.riken.go.jp/sensor/Allan/ICA/

Source Separation and Diversity for Communications

Lang Tong, Cornell University

Source Separation and Diversity

Separating multiple sources is a fundamental signal processing problem that arises in many applications. An easily understood scenario is the so-called "cocktail party" problem where the objective is to separate one voice from others in the room. To a large degree, humans perform this task remarkably well. We are able to distinguish different voices, and may even follow the conversation. The cocktail-party problem is germane to many similar applications in biomedical signal processing, geophysical signal processing and, most notably, in communication system designs.

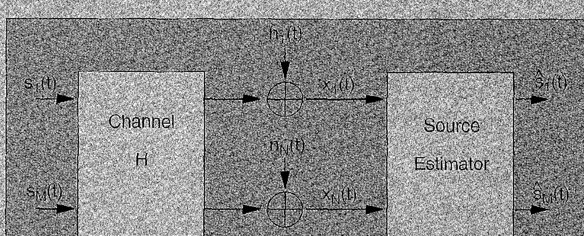
In wireless communications, for example, when the signal is transmitted over a nonideal channel, sophisticated signal processing is often required at the receiver to counter noise and interference from various sources. One of the major channel impairments is multipath fading. The problem here is analogous to calling someone a hundred yards away in the Grand Canyon. Your friend will hear your voice and its echoes. He can separate your voice and its echoes unless, of course, you speak too fast. It will be much harder for him if someone else tries to talk to him at the same time. But this is exactly the problem in today's digital cellular system where the transmitted signal is interfered not only by its echoes (multipath interference) but also by other users in the neighborhood (cochannel interference). We face the source separation problem similar to that in a cocktail party.

What makes it possible for us to separate and track different voices in a cocktail party? What features do we use? Can this process be made automatic? These questions touch upon some of the fundamental issues in source separation. Perhaps it is easier to see what makes it more difficult to separate sources. Can you still separate different voices with only one ear? Perhaps, but it is more difficult. What if you close your eyes? The loss of visual signals will definitely make it much harder. As it turns out, the key to source separation is *diversity*. To separate different sources, it is important to have *different* receptions of the signal. This can be done in different ways by exploiting the characteristics of the signal and its propagation medium.

▲ *Spatial Diversity*. Using sensor arrays is an effective way to gain spatial diversity. You can see the antenna array on the cell-phone tower along major highways. (Imagine how much more you could hear at the cocktail party if you had an array of ears!) Recently, the so-called smart antenna technology has attracted considerable interest from both academia and industry (see a recent survey by Paulraj and Papadias [306]).

▲ *Spectral Diversity*. Spectral diversity can be obtained in many ways. To counter frequency-selective fading caused by echoes, the transmission band can be divided into a number of smaller frequency bands through which modulated signals of different power are transmitted. This kind of multicarrier transmission is the basis for the OFDM (orthogonal frequency division multiplexing) system for digital audio broadcasting (DAB). Frequency-hopping (FH) spread spectrum [374] is another way of achieving spectral diversity. Invented in the 1940s, this technique avoids hostile interference by changing the transmission band in a way known to the receiver but unpredictable to the jammer.

▲ *Temporal Diversity*. Diversity can also be achieved in the time-domain by transmitting multiple copies of the signal at different times or, in a more intelligent way, by transmitting the signal in some special waveform known to the receiver. A direct-sequence spread-spectrum signal gains diversity by modulating the source with a special code sequence. When the received signal is sampled at a rate higher than the bit rate, we effectively obtain the transmitted symbol at a different time, which is analogous to spatial diversity where the transmitted signal is obtained at different locations. Different users can be separated if they use different codes. The exploitation of temporal correlation of the signal waveform provides a crucial property used in many signal-processing algorithms. See again [306] for discussions about space-time processing.



▲ *8. Source separation with receiver diversity.*

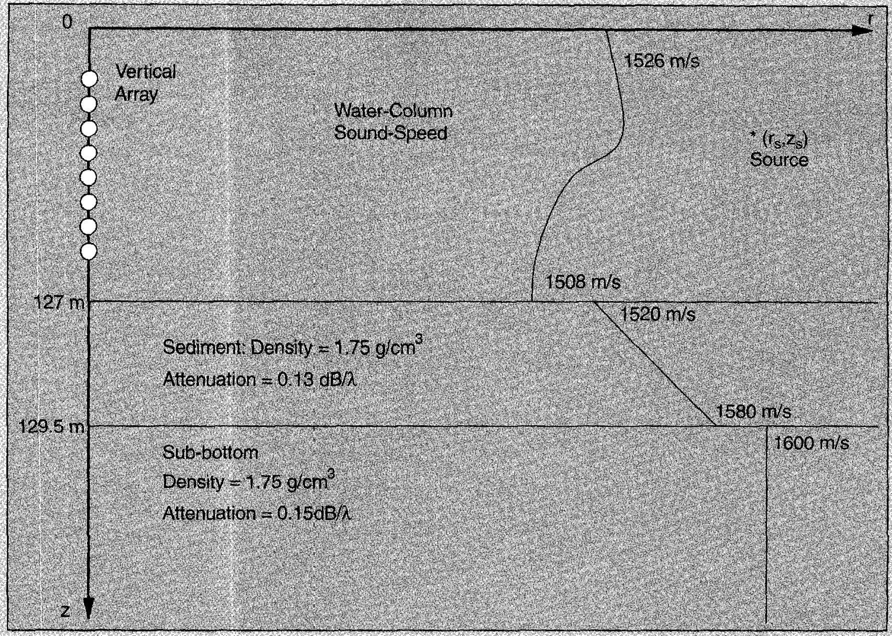
Statistical Signal Processing for Source Separation

Estimating signals under various diversity receptions is currently an active research area. A general model for source separation is shown in Fig. 8 where sources $\{s_i(t)\}$ are propagated through a channel $\mathbf{H}$, contaminated by noise $\{s_i(t)\}$, and received with diversity as $\{x_i(t)\}$. The goal of source separation is to estimate one or all of the source signals from $\{x_i(t)\}$.

Classical approaches to source separation are based on either the knowledge of the channel or the ability of having access to the channel input so that the channel (or its inverse) can be estimated by sending "training" signals. Depending on the application, different criteria (such as minimizing the detection error probability or minimizing the mean-square error) can be used to design the signal estimator.

In recent years, there has been considerable interest in the so-called blind-source- separation problem. Here, neither the channel is known a priori, nor is it possible to have access to the channel input so that training can be made (recall again the cocktail-party problem). The merit of blind signal separation is twofold. First, there are cases when channel estimation by training is very difficult. Second, the transmission of training inevitably reduces the transmission rate of information. Such reduction can be significant when training has to be performed repeatedly.

The key to blind signal separation is to exploit *qualitative* information about the structure of the channel and characteristics of the input sources. For example, source signals in communications often can only have a finite number of alphabets, which enable the separation of multiple sources. The statistical independency among sources is another condition that leads to a number of effective source separation algorithms [50], among them is the widely applied CMA [143, 416]. In a special issue of *IEEE Proceedings* [242] to appear later this year, tutorials of blind-source-separation techniques and their applications are presented. For related topics on this growing field of research, readers are also referred to a special issue (edited by G. Giannakis and G. Xu) on signal processing for advanced communication [127] and the recent book by Poor and Wornell [313].



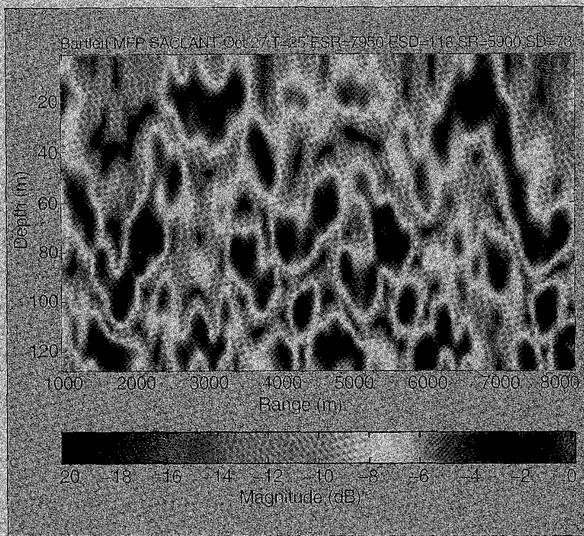
▲ *9. A typical scenario for passive sonar matched-field processing.*

SSAP with Computational Acoustic and Electromagnetic Propagation Models

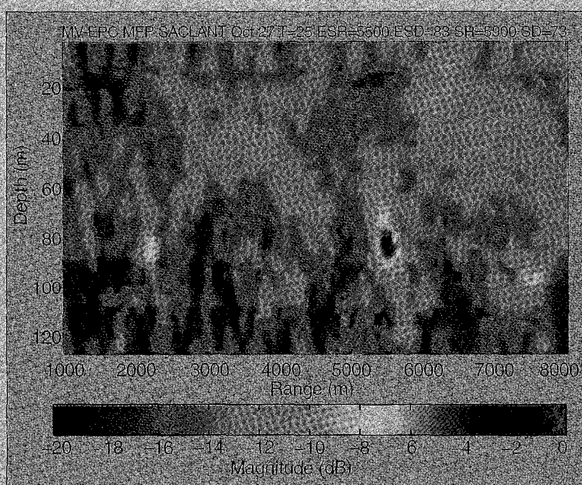
Jeffrey Krolik, Duke University

Statistical signal and array processing of signals carried by propagating waves has traditionally been developed assuming simple plane-wave acoustic or electromagnetic propagation models. This is despite the fact that in many problems associated with sonar, radar, wireless communications, and geophysics, complex coherent multipath propagation between the source and receiver is a dominant feature. The historical focus on signal processing using plane-wave models has been a result of: 1) their analytic simplicity, 2) elegant analogies between familiar time-domain filtering/spectral analysis and plane-wave beamforming/field-directionality mapping, and 3) the fact that accurate numerical models for complex multipath propagation were too computationally intensive for signal-processing applications. Although difficulties with plane-wave approximations in coherent multipath environments are often dealt with by a variety of mitigation techniques, the performance of such methods is inevitably upper bounded by the case where multipath is absent. The notion that instead of trying to undo the effects of coherent multipath one could actually exploit them to achieve dramatically improved performance with the assistance of a numerical propagation model is the essence what is now known as matched-field processing (MFP). The availability of inexpensive, high-power computing to rapidly calculate numerical solutions of the wave equation is what has really driven the development of MFP methods over the last decade.

Matched-field processing was first developed as a simple generalization of narrowband plane-wave beamforming wherein conventional array weights based on plane-wave "steering vectors" were replaced by "replica vectors" derived from the full-field solution of the wave equation in a ducted channel. Historically, these techniques have been applied to the problem of underwater passive source localization [10]. A common set-up for passive sonar MFP is illustrated in Fig. 9, where acoustic signals in a shallow-water waveguide are received at a vertical array of hydrophones from a distant source in the presence of interference from surface shipping plus diffuse ambient noise. The objective is to detect and localize the source in range and depth. The medium is described by a sound-speed profile within the water column together with the bathymetry and geoacoustic properties of the bottom. Because distant signals have numerous interactions with the ocean boundaries, single-path plane-wave propagation is clearly not an appropriate model here. However, given sufficiently accurate environmental information, the coherent sum of multipaths between source and receivers for different hypothesized source locations can be predicted by numerical solution of the wave equation and its boundary conditions. The most basic MFP approach, known as Bartlett matched-field beamforming, is then to simply correlate each of these replicas with the field measured at the array.



▲ *10. A Bartlett matched-field ambiguity surface for SACLANT Mediterranean data.*



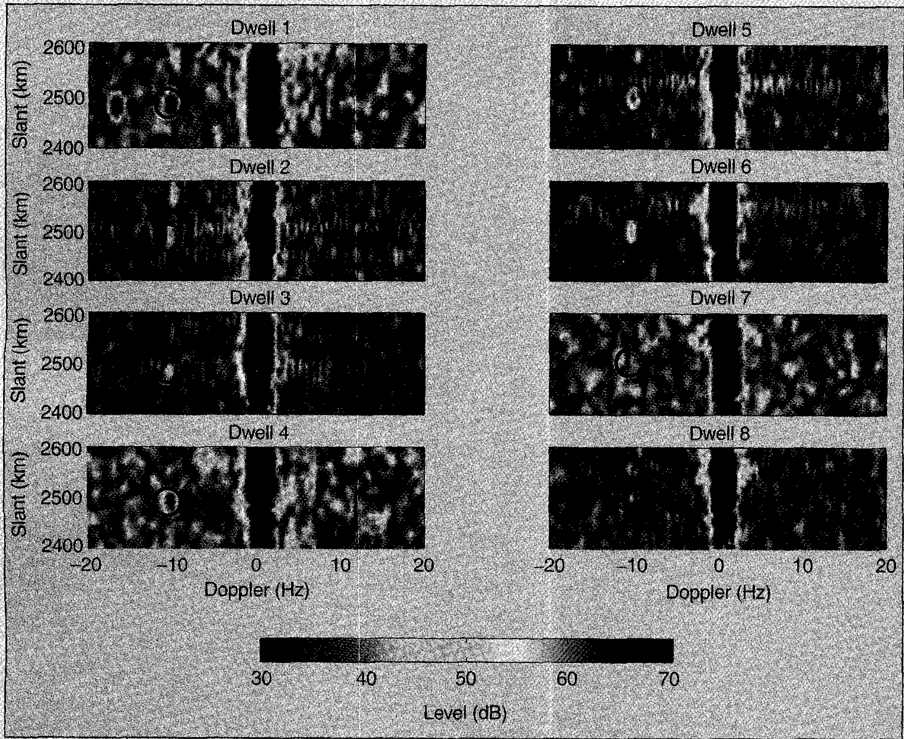
▲ *11. An MV-EPC matched-field ambiguity surface for SACLANT Mediterranean Data.*

The estimated source location is then the hypothesized range and depth which maximizes the power at the output of the beamformer. The so-called ambiguity surface of the Bartlett beamformer is its output power versus hypothesized range and depth. A typical Bartlett ambiguity surface obtained using actual Mediterranean data from a 48-sensor vertical array is shown in Fig. 10. Note that there is a local maximum of this ambiguity surface near the true source location at a range of 5900 m and depth of 70 m. This data set was collected by the NATO SACLANT Centre [139] and is currently available to the public on the IEEE SP database web-site at Rice University (http://spib.rice.edu).

Two major difficulties facing matched-field processors are: 1) high sidelobe levels, as seen in Fig. 10, which result in ambiguous source location estimates, and 2) the sensitivity of these methods to errors in the assumed environmental conditions. In order to suppress ambiguous sidelobes, several variations of minimum variance (MV) adaptive beamforming have been proposed [78, 215, 360]. The basic MV matched-field beamformer selects weights for each hypothesized source location that minimize output power subject to the constraint of unity gain for signals emanating from the desired range-depth point under the assumed environmental model. MV matched-field beamformers generally do provide lower sidelobe levels than the Bartlett processor, but often this comes at the price of even greater sensitivity to mismatch between the assumed and actual environmental conditions. To achieve matched-field localization performance that is robust to environmental mismatch, a number of approaches have been proposed, not only in the form of beamforming methods, but also by jointly estimating the source location and environmental parameters [73, 332]. One robust beamforming approach that has been demonstrated to provide lower sidelobes and higher probability of correct source localization in the presence of environmental uncertainty is the minimum variance beamformer with environmental perturbation constraints (MV-EPC) [215, 216]. Essentially, the MV-EPC method minimizes output power subject to a set of linear constraints designed to maintain array gain for signals emanating from a hypothesized range-depth point but over an ensemble of perturbed channel models derived from the statistics of uncertain ocean parameters.

A typical MV-EPC ambiguity surface for the Mediterranean SACLANT dataset as shown in Fig. 11. Observe that the sidelobe levels are substantially lower than those of the Bartlett processor in Fig. 10 with an unambiguous peak at a range of 5500 m and depth of 82 m, which is close to the true source position. Among parameter-estimation approaches to robust matched-field processing, Bayesian methods that involve numerical or Monte Carlo integration over a priori distributions for the uncertain environmental parameters have been developed in [332] and global optimization methods for jointly estimating the source and environmental parameters are proposed in [139] and [216]. In addition, it is worth noting that wideband matched-field processing, usually achieved by incoherently averaging ambiguity surfaces across frequency, often significantly increases robustness to mismatch. Wideband averaging increases robustness to mismatch under conditions where perturbations of the environment cause only a frequency-independent shift in the source location peak, while simultaneously decorrelating the sidelobe structure across the receiver band.

Although an exhaustive literature now exists on matched-field processing for underwater acoustics, there are still several important open problems. These include: 1) reducing the threshold SNR required to detect submerged targets, 2) achieving computationally efficient MFP methods with large multidimensional arrays, and 3) handling more dynamic targets that seriously limit avail-



▲ *12. Log-amplitude delay (a.k.a. slant range) versus Doppler frequency surfaces for multiple radar dwells.*

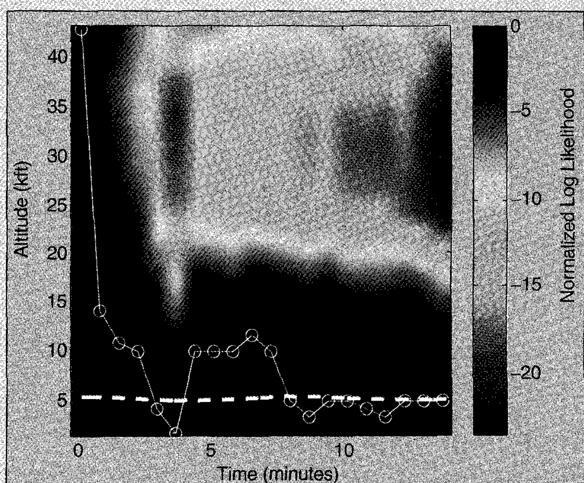
unresolved multipaths from ground reflections local to the aircraft [301]. These predictions are essentially correlated with complex delay-Doppler data in a ML altitude estimation method that uses multiple radar dwells on the aircraft target as illustrated in Fig. 12. The sub-plots in this figure represent consecutive delay-Doppler surfaces from an OTH radar. Note that the dominant vertical band in each sub-plot is due to ground clutter and the peak encircled in each plot is a small twin-engine aircraft approximately 2,500 km away from the radar. Matched-field altitude estimation (MFAE) exploits the fact that the complex fading characteristic of the target peak is highly dependent on aircraft altitude. In MFAE, a time-evolving log-likelihood function of aircraft altitude is updated with each radar revisit of the target. An example of this function is shown in Fig. 13 for this real-data example. Observe that the matched-field estimate of altitude is remarkably close to its true height of 5,000 ft.

In summary, statistical signal and array processing with computational propagation models permit the exploitation of complex multipath conditions to achieve significantly enhanced performance over traditional methods.

able integration times. Those interested in pursuing these problems can find much of the necessary software and data at NJIT's Ocean Acoustics Library web-site (http://oalib.njit.edu).

Beyond underwater acoustic applications, matched-field techniques are also currently being explored for problems involving multipath electromagnetic propagation. Although this work is still in its early stages, electromagnetic matched-field processing (EM-MFP) has already been proposed in several applications including: 1) aircraft height-finding using a low-angle microwave radar that can exploit multipath due to the target's direct-path and specular reflection off the ground [182]; 2) inversion of tropospheric refractivity parameters that characterize ducted radio-wave propagation conditions over the sea surface using point-to-point microwave transmissions [138]; and 3) target localization in high-frequency skywave over-the-horizon (OTH) radar [217, 301]. One application where EM-MFP provides an existing radar with an entirely new capability is that of aircraft altitude estimation for OTH radar. Over-the-horizon radars use the refractive properties of the ionosphere for wide-area surveillance of targets at megameter ranges. And although target localization in latitude and longitude is typically achieved by tracing the paths of rays refracted through the ionosphere, determination of target altitude has never been reliably achieved.

In recent work, however, a form of EM-MFP has been developed that uses an ionospheric propagation model to predict the signal in complex delay-Doppler space due to



▲ *13. Time-evolving log-likelihood of aircraft altitude for a small twin-engine plane at an altitude of 5,000 feet and range of 2,500 kilometers.*

Such approaches involve a tight coupling between the physics of wave propagation and signal processing. They also rely on the availability of sufficiently accurate estimates of the environmental parameters. Numerous results obtained with real data in very different settings, however, suggest that "sufficiently accurate" should by no means be interpreted as "perfect knowledge" of the environment. Indeed, in some situations, robust signal processing methods have been developed that facilitate matched-field processing with almost "common knowledge" of the environment. Clearly, the further integration of computational wave-equation solutions and signal-processing techniques will pose many challenges and rewards in the future.

A WWW link to the author of the above section: http://www.ee.duke.edu/people/jk.html

References

1. A. Arneodo, Y. D. Carafa, B. Audit, E. Bacry, J. Muzy, and C. Thermes, "What can we learn with wavelets about dna sequences," *Physica A*, vol. 249, pp. 439-448, 1998.
2. P. Abry and P. Flandrin, "Point processes, LRD and wavelets," in *Wavelets in Biology and Medicine*, A. Aldroubi and M. Unser, editors, CRC Press, 1996.
3. B. G. Agee, S. V. Schell, and W. A. Gardner, "Spectral self-coherence restoral: a new approach to blind adaptive signal extraction using antenna arrays," *Proceedings of the IEEE*, vol. 78, pp. 753-767, 1990.
4. M. Amin, "Time-frequency spectrum analysis and estimation for nonstationary random processes," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash, editor, pp. 208-232, Wiley Halsted Press, Australia, 1992.
5. M. Amin, "Minimum variance time-frequency distribution kernels for signal in additive noise," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2352-2356, 1996.
6. D. Andrews and A. Herzberg, *Data: A collection of problems from many fields for the student and research worker*, Springer-Verlag, 1985.
7. C. Andrieu, A. Doucet, and P. Duvaut, "Joint Bayesian detection and estimation of sinusoids embedded in noise," in *ICASSP*, pp. 2245-2248, 1998.
8. C. Andrieu, A. Doucet, W. J. Fitzgerald, and S. J. Godsill, "An introduction to the theory and applications of simulation based computational methods in signal processing," Technical Report CUED/F-INFENG/TR. 324, University of Cambridge, UK, 1998.
9. J. Bae and I. Song, "Rank-based detection of weak random signals in a multiplicative noise model," *Signal Processing*, vol. 63, pp. 121-131, 1997.
10. A. Baggeroer, W. Kuperman, and P. N. Mikhalevsky, "An overview of matched-field methods in ocean acoustics," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 401-424, 1993.
11. J. Q. Bao and L. Tong, "Applications of blind equalization in wireless atm networks," in *Proc. ICASSP'98*, 1998.
12. R. Baraniuk and D. Jones, "A signal dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1589-1602, 1993.
13. S. Barbarossa and A. Farina, "Space-time-frequency processing of synthetic aperture radar signals," *IEEE Transactions on Aerospace Electr. Syst.*, vol. 30, pp. 341-358, 1994.
14. S. Barbarossa, A. Scaglione, A. Baiocchi, and G. Colletti, "Modeling network traffic data by doubly stochastic point processes with self-similar intensity process and fractal renewal point process," in *Proc. Thirty-First Asilomar Conf on Signals, Systems and Computers*, pp. 1112-16, Monterey, CA, Nov 1997.
15. G. Barnett, R. Kohn, and S. Sheather, "Bayesian estimation of an autoregressive model using markov chain monte carlo," *Journal of Econometrics*, vol. 74, pp. 237-254, 1996.
16. P. Barone and R. Ragona, "Bayesian estimation of parameters of a damped sinusoidal model by a markov chain monte carlo method," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1806-1814, 1997.
17. D. L. Bartholomew and J. A. Tague, "Quadratic power spectrum estimation with orthogonal frequency division multiple windows," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1279-1282, 1995.
18. A. Bartov and H. Messer, "Lower bound on the achievable DSP performance for localizing step-like continuous signals in noise," *IEEE Transactions on Signal Processing*, vol. SP-46, 1988.
19. M. Basseville, M. Benveniste, A. Chou, K. Golden, R. Nikoukhah, and A. S. Willsky, "Modeling and estimation of multiresolution stochastic processes," *IEEE Trans. on IT*, vol. IT-38, pp. 529-532, Mar. 1992.
20. B. Baygun and A. O. Hero, "Optimal simultaneous detection and estimation under a false alarm constraint," *IEEE Trans. on Inform. Theory*, vol. 41, no. 3, pp. 688-703, 1995.
21. B. Baygun and A. O. Hero, "An iterative solution to the min-max simultaneous detection and estimation problem," in *Proc. of the IEEE Workshop on Statistical Signal and Array Processing*, pp. 8-11, Corfu, Greece, June 1996.
22. E. Beadle and P. M. Djuric, "Parameter estimation for non-Gaussian autoregressive processes," in *ICASSP*, pp. 3557-3560, 1997.
23. A. Belouchrani and M. Amin, "Blind source separation based on time-frequency signal representations," *IEEE Transactions on Signal Processing*, p. Submitted.
24. S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*, Prentice-Hall Inc., New Jersey, 1987.
25. S. E. Bensley and B. Aazhang, "Subspace based channel estimation for cdma communication systems," *IEEE Trans. on Communications*, pp. 1009-1020, August 1996.
26. A. Benveniste, M. Goursat, and G. Ruget, "Blind equalizers," *IEEE Trans. on Communications*, pp. 871-882, August 1982.
27. M. J. Berliner and J. F. L. (Eds.), *Acoustic Particle Velocity Sensors: Design, Performance and Applications*, AIP Press, Woodbury, NY, 1996.
28. J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, John Wiley, New York, 1994.
29. O. Besson and P. Stoica, "Sinusoidal signals with random amplitudes: least-squares estimators and their statistical analysis," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2733-44, Nov 1995.
30. T. G. Birdsall and J. O. Gobien, "Sufficient statistics and reproducing densities in simultaneous sequential detection and estimation," *IEEE Trans. on Inform. Theory*, vol. 19, pp. 760-768, 1973.
31. R. Blum, "Asymptotically robust detection for known signals in contaminated multiplicative noise," *IEEE Transactions on Signal Processing*, vol. 38, pp. 197-210, 1994.
32. B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, editor, pp. 418-517, Prentice-Hall, 1990.
33. B. Boashash and P. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Transactions on Signal Processing*, vol. 42, 1995.
34. N. Bose and C. Rao, editors, *Handbook of Statistics, Volume 10, Signal Processing and its Applications*, Elsevier Science Publishers, Amsterdam, 1993.
35. S. Bose and A. O. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2801-2809, Nov 1996.
36. S. Bose and A. O. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2164-2175, Sept. 1995.

37. G. F. Boudreaux and T. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 442-451, 1986.

38. C. Bouman and K. Sauer, "Fast numerical methods for emission and transmission tomographic reconstruction," in *Proc. Conf. on Inform. Sciences and Systems*, Johns Hopkins, 1993.

39. M. Bouvet and S. C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 5, pp. 621-626, 1989.

40. A. Bovik, "On detecting edges in speckle imagery," *IEEE Trans ASSP*, vol. 36, no. 10, pp. 1618-27, Oct 1988.

41. G. E. P. Box and G. C. Tiao, *Bayesian Inference*, John Wiley, New York, 1973.

42. D. R. Brillinger, *Time Series: Data Analysis and Theory*, Springer-Verlag, New York, 1981.

43. D. Brillinger, "An introduction to polyspectra," *Annals Math. Statistics*, vol. 36, pp. 1351-1374, 1965.

44. D. Brillinger, "Comparative aspects of the study of ordinary time series and point processes," *Developments in Statistics*, vol. 1, pp. 33-133, 1978.

45. P. L. Brockett, M. Hinich, and G. R. Wilson, "Nonlinear and non-Gaussian ocean noise," *J. Acoustical Society of America*, vol. 82, pp. 1386-1394, 1987.

46. T. P. Bronez, "Spectral estimation of irregularly sampled multidimensional processes by generalized prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 36, pp. 862-873, 1988.

47. T. P. Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2941-2946, 1992.

48. K. Burgess and B. D. Van Veen, "Subspace-based adaptive generalized likelihood ratio detection," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 912-927, 1996.

49. S. Cambanis and C. Houdré, "On the continuous wavelet transform of second order random processes," *IEEE Trans. on IT*, vol. 41, no. 3, pp. 628-633, May 1995.

50. J.-F. Cardoso, "Blind signal separation: statistical principles," *Proc. of the IEEE. Special issue on blind identification and estimation*, 1998. To appear.

51. J.-F. Cardoso, M. Lavielle, and E. Moulines, "Un algorithme d'identification par maximum de vraisemblance pour des donnees incompletes," *Comptes Rendus de l'Academie des Sciences, Series I*, vol. 320, no. 3, pp. 363-368, 1995.

52. B. P. Carlin, N. G. Polson, and D. S. Stoffer, "A monte carlo approach to nonnormal and nonlinear state-space modeling," *Journal of the American Statistical Association*, vol. 87, pp. 493-500, 1992.

53. C. K. Carter and R. Kohn, "Semiparametric Bayesian inference for time series with mixed spectra," *Journal of the Royal Statistical Society B*, vol. 59, pp. 255-268, 1997.

54. G. Carter, editor, *Special Issue on Time Delay Estimation*, volume ASSP-29, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981.

55. G. Carter, editor, *Coherence and Time Delay Estimation-An Applied Tutorial for Research, Development, Test, and Evaluation Engineers*, IEEE-Press, Piscataway, NJ, 1993.

56. E. D. Carvalho and D. T. M. Slock, "Cramer-rao bounds for semi-blind and training sequence based channel estimation," in *Proc. 1st IEEE Signal Proc. Work. on Wireless Comm.*, pp. 129-132, 1997.

57. J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Trans. on Aerosp. Electron. and Systems*, vol. AES-23, no. 1, pp. 46-59, 1987.

58. R. Chen and T.-H. Li, "Blind restoration of linearly degraded discrete signals by gibbs sampling," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2410-2413, 1995.

59. J. Cheung and L. Kurz, "A generalized M-interval partition detcetoir with applications to signal detection in impulsive noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 41, no. 1, pp. 213-221, 1993.

60. A. Chevreuil and P. Loubaton, "Blind second-order identification of fir channels: Forced cyclo-stationarity and structured subspace method," *IEEE Signal Proc. Letters*, vol. 4, no. 7, pp. 204-206, July 1997.

61. C.-M. Cho and P. Djuric, "Bayesian detection and estimation of cisoids in colored noise," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2943-2952, 1995.

62. H. Choi and W. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 862-871, 1989.

63. S. Chretien and A. Hero, "Acceleration of the EM algorithm via proximal point iterations," in *Proc. of IEEE Symposium on Information Theory*, MIT, Cambridge, August 1998.

64. S. Chretien and A. Hero, "Generalized proximal point algorithms and bundle implementations," Technical Report CSPL-314, Comm. and Sig. Proc. Lab. (CSPL), Dept. EECS, University of Michigan, Ann Arbor, Mar. 1998.

65. P.-H. Chua, C.-M. S. See, and A. Nehorai, "Vector-sensor array processing for estimating angles and times of arrival of multipath communication signals," in *Proc. IEEE Intl Conf. on Acoust., Speech, and Sig. Proc. (ICASSP98)*, pp. 3325-3328, Seattle, WA, May 1998.

66. J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision feedback equalization and coding - parts i and ii," *IEEE Trans. on Communications*, vol. 43, no. 10, pp. 2582-2604, Oct. 1995.

67. T. Claasen and W. Mecklenbrauker, "The Wigner-distribution - a tool for time-frequency signal analysis- part I: Continuous-time signals, - part II: Discrete time signals, - part III: Relations with other time-frequency signal transformations," *Philips J. Research*, vol. 35, 1980.

68. D. Cochran, H. Gish, and D. Sinno, "A geometric approach to multiple-channel signal detection," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2049-2057, 1995.

69. L. Cohen, "Generalized phase space distribution functions," *Journal of Math. Physics.*, vol. 7, pp. 781-786, 1966.

70. L. Cohen, "Time-frequency distributions - a review," *Proceedings of the IEEE*, vol. 77, pp. 941-981, 1989.

71. L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, New Jersey, 1995.

72. R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. on IT*, vol. IT-38, pp. 713-718, Mar. 1992.

73. M. Collins and W. Kuperman, "Focalization: Environmental focusing and source localization," *Journal of the Acoustical Society of America*, vol. 90, no. 3, pp. 1410-1422, 1991.

74. P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, Elsevier, vol. 36, no. 3, pp. 287-314, 1994.

75. P. Comon and J. F. Cardoso, "Eigenvalue decomposition of a tensor with applications," in *Proc. SPIE Conf. on Advanced Signal Processing Algorithms, Architectures and Implementations*, pp. 361-372, San Diego, CA, July 1990.

76. P. Comon and B. Mourrain, "Decomposition of quantics in powers of linear forms," *Signal Processing*, vol. 53, no. 2-3, pp. 93-108, Sep 1996.

77. R. T. Compton, *Adaptive Antennas - Concepts and Performance*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1988.

78. H. Cox, R. Zeskind, and M. Myers, "A subarray approach to matched-field processing," *Journal of the Acoustical Society of America*, vol. 85, no. 3, pp. 1158-1166, 1990.

79. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Signal estimation using wavelet-markov models," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing–ICASSP '97*, Munich, Germany, 1997.

80. Z. Cvetkovic, "Short-time fourier analysis—a novel window design procedure," in *Proc. ICASSP*, volume 3, pp. 1773-1776, 1998.

81. D. Dahlhaus, A. Jarosch, B. Fleury, and R. Heddergott, "Joint demodulation in ds/cdma systems exploiting the space and time diversity of the mobile radio channel," in *Proc. IEEE Int. Symp. on Personal Indoor and Mobile radio Communications*, pp. 47-52, Helsinki, Finland, 1997.

82. A. Dandawate and G. Giannakis, "Statistical tests for presence of cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2355-2369, 1994.

83. M. Daniel and A. Willsky, "A multiresolution methodology for signal-level fusion and data assimilation with applications to remote sensing," *IEEE Proceedings*, vol. 85, no. 1, 1997.

84. K. Daoudi, "Multifractal representation of turbulence signals: A wavelet based approach," in *INRIA International Wavelets Conference*, INRIA, editor, Tanger, Morocco, April 1998.

85. I. Daubechies, "The wavelet transform, time-frequency localization, and signal analysis," *IEEE Trans Inf. Theor.*, vol. 36, pp. 961-1005, 1990.

86. I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF, SIAM, Philadelphia, 1992.

87. R. Davis and S. Resnick, "Limit theory for bilinear processes in heavy-tailed noise," *Annals of Probability*, vol. 6, no. 4, pp. 1191-1210, 1996.

88. De Moor B.L.R. (ed.). *DAISY: Database for the Identification of Systems*. http://www.esat.kuleuven.ac.be/sista/daisy, October 1997.

89. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statistical Society, Ser. B*, vol. 39, no. 1, pp. 1-38, 1977.

90. Z. Ding, "Adaptive filters for blind equalization," in *Digital Signal Processing Handbook*, V. K. Madisetti and D. Williams, editors, chapter 24, CRC Press, 1998.

91. P. M. Djuric, S. J. Godsill, W. J. Fitzgerald, and P. J. W. Rayner, "Detection and estimation of signals by reversible jump markov chain monte carlo computations," in *ICASSP*, pp. 2269-2272, 1998.

92. P. M. Djuric and H.-T. Li, "Bayesian spectrum estimation," *Signal Processing Letters*, vol. 2, pp. 213-215, 1995.

93. A. Dogandzic and A. Nehorai, "Estimating evoked dipole responses by MEG/EEG for unknown noise covariance," in *Proc. 19th Ann. Intl Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1224-1227, Chicago, IL, October 1997.

94. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," preprint Dept. Stat., Stanford Univ., Jun. 1992.

95. D. Donoho and I. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *JASA*, vol. 90, pp. 1200-1223, Dec. 1995.

96. A. Doucet and P. Duvaut, "Bayesian estimation of state space models applied to deconvolution of Bernoulli-Gaussian processes," *Signal Processing*, vol. 57, pp. 147-161, 1997.

97. G. D'Spain and W. S. Hodgkiss, "The simultaneous measurement of infrasonic acoustic particle velocity and acoustic pressure in the ocean by freely drifting swallow floats," *IEEE J. Oceanic Eng.*, vol. 16, pp. 195-207, April 1991.

98. A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous CDMA channel," *IEEE Trans. on Communications*, pp. 285-290, February 1993.

99. R. F. Dwyer, "Fourth-order spectra of Gaussian amplitude-modulated sinusoids," *J. Acoustical Society of America*, vol. 90, no. 2, pp. 918-926, August 1991.

100. M. H. El Ayadi, "Generalized likelihood adaptive detection of signals deformed by unknown linear filtering in noise with slowly fluctuating power," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-33, no. 2, pp. 401-405, 1985.

101. M. H. El Ayadi, "NAR estimators of spatial covariance matrices for adaptive array detection," *IEEE Transactions on Signal Processing*, vol. 39, no. 7, pp. 1682-1690, 1991.

102. M. H. El Ayadi and B. Picinbono, "NAR AGC adaptive detection of nonoverlapping signals in noise with fluctuating power," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 5, pp. 952-962, 1981.

103. European Telecommunications Standards Institute, editor, *European Telecommunications Standard, Radio broadcast systems: Digital audio broadcasting (DAB) to mobile, portable, and fixed receivers*, Standard ETS 300 401, 1994.

104. A. Farina, *Antenna Based Signal Processing Techniques for Radar Systems*, Artech House, Norwood, MA, 1992.

105. K. Fazel and G. Fettweis, *Multi-Carrier Spread Spectrum*, Kluwer Academic Publ., 1997.

106. M. Feder and E. Weinstein, "Optimal multiple source location estimation via the EM algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, pp. 1762-1765, 1985.

107. M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. 36, no. 4, pp. 477-489, April 1988.

108. W. Feller, *An Introduction to Probability Theory and Its Applications - Volume II*, John Wiley and Sons, second edition, 1971.

109. T. S. Ferguson, *Mathematical Statistics - A Decision Theoretic Approach*, Academic Press, Orlando FL, 1967.

110. J. A. Fessler and A. O. Hero, "Space-alternating generalized EM algorithm," *IEEE Transactions on Signal Processing*, vol. SP-42, no. 10, pp. 2664-2677, Oct. 1994.

111. J. A. Fessler and A. O. Hero, "Penalized maximum-likelihood image reconstruction using space alternating generalized EM algorithm," *IEEE Trans. on Image Processing*, vol. IP-4, no. 10, pp. 1417-1429, Oct. 1995.

112. P. Fieguth, W. Karl, A. Willsky, and C. Wunsch, "Multiresolution optimal interpolation and statistical analysis of topex/poseidon satellite altimetry," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 280-292, March 1995.

113. P. Flandrin, "A time-frequency formulation of optimum detectors," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1377-1384, 1988.

114. P. Flandrin, "On the spectrum of fractional Brownian motion," *IEEE Trans. on IT*, vol. IT-35, pp. 197-199, Jan. 1989.

115. B. Fleury, D. Dahlhaus, R. Heddergott, and M. Tschudin, "Wideband angle of arrival estimation using the SAGE algorithm," in *Proc. of IEEE Int. Symp. on Spread Spectrum Tech. and Apps*, pp. 79-85, Mainz, DL, 1996.

116. I. K. Fodor and P. B. Stark, "Multitaper spectrum estimates," in *Proc. SOHO 6 / GONG 98 workshop on helioseismology*, 1998, in press.

117. G. J. Foschini, "Equalization without altering or detect data," *AT&T Tech. Journal*, pp. 1885-1911, October 1985.

118. J. Francos and B. Friedlander, "Bounds for estimation of multicomponent signals with random amplitude and deterministic phase," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1161-72, May 1995.

119. A. Fredricksen, D. Middleton, and D. VandeLinde, "Simultaneous signal detection and estimation under multiple hypotheses," *IEEE Trans. on Inform. Theory*, vol. 18, pp. 607-614, 1972.

120. D. R. Fuhrmann, "Application of toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Transactions on Signal Processing*, vol. 39, no. 10, pp. 2194-2198, 1991.

121. M. V. E. G. D. Forney, Jr., "Combined equalization and coding using precoding," *IEEE Communications Magazine*, pp. 25-34, Dec 1991.

122. W. A. Gardner, editor, *Cyclostationarity in Communications and Signal Processing*, IEEE Press, 1994.

123. A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman and Hall, New York, 1995.

124. S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis, and Machine Intelligence*, vol. 6, pp. 721-741, 1984.

125. A. A. Gerlach, "Optimum detection performance of passive coherence estimators," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-33, no. 1, pp. 240-247, 1985.

126. M. Ghogho and A. Nandi, "Locally optimum detectors for deterministic signals in multiplicative noise," in *Proc. ICASSP'98*, volume IV, pp. 2137-40, Seattle, May 1998.

127. G. Giannakis and G. Xu, editors, *Special Issue on Signal Processing For Advanced Communications*, 1997.

128. G. B. Giannakis, "Filterbanks for blind channel identification and equalization," *IEEE Signal Processing Letters*, vol. 4, pp. 184-187, June 1997.

129. G. B. Giannakis, "Cyclostationary signal analysis," in *Digital Signal Processing Handbook*, V. K. Madisetti and D. Williams, editors, chapter the Statistical Signal Processing Section, CRC Press, 1998.

130. G. B. Giannakis and S. Halford, "Blind fractionally-spaced equalization of noisy fir channels: direct and adaptive solutions," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2277-2292, September 1997.

131. G. B. Giannakis and J. M. Mendel, "Identification of non-minimum phase systems using higher-order statistics," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 360-377, March 1989.

132. G. B. Giannakis and E. Serpedin, "Linear multichannel blind equalizers of nonlinear fir volterra channels," *IEEE Transactions on Signal Processing*, vol. 45, pp. 67-81, January 1997.

133. G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind equalization of time-varying channels," *Proceedings of the IEEE*, September 1998 (to appear).

134. G. B. Giannakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, pp. 1284-1296, 1990.

135. G. Giannakis, "Polyspectral and cyclostationary approaches for identification of closed-loop systems," *IEEE Trans. Automatic Control*, vol. 40, pp. 882-885, 1995.

136. G. Giannakis and M. Tstatsanis, "Time-domain tests for Gaussianity and time-reversibility," *IEEE Transactions on Signal Processing*, vol. 42, pp. 3460-3472, 1994.

137. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, New York, 1996.

138. D. Gingras, P. Gerstoft, and N. Gerr, "Electromagnetic matched-field processing: Basic concepts and tropospheric simulations," *IEEE Transactions on Antennas and Propagation*, vol. 45, pp. 1536-1545, 1997.

139. D. Gingras and P. Gerstoft, "Inversion for geometric and geoacoustic parameters in shallow water: Experimental results," *Journal of the Acoustical Society of America*, vol. 97, no. 6, pp. 3589-3598, 1995.

140. F. Gini, "A cumulant-based adaptive technique for coherent radar detection in a mixture of K-distributed clutter and a Gaussian disturbance," *IEEE Transactions on Signal Processing*, vol. 45, no. 6, pp. 1507-1519, 1997.

141. F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat fading channels: A cyclostationary approach," *IEEE Transactions on Communications*, vol. 46, pp. 400-411, March 1998.

142. F. Gini and G. B. Giannakis, "Generalized differential encoding: A nonlinear signal processing framework," *IEEE Transactions on Signal Processing*, vol. 46, 1998 (to appear).

143. D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. on Communications*, pp. 1867-1875, 1980.

144. S. Godsill, "Bayesian enhancement of speech and audio signals which can be modelled as ARMA processes," *International Statistical Review*, vol. 65, pp. 1-21, 1997.

145. S. J. Godsill, "Robust modelling of noisy ARMA signals," in *ICASSP*, pp. 3797-3800, 1997.

146. N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings*, vol. 140, pp. 107-113, 1993.

147. A. Gorokhov and P. Loubaton, "Semi-blind second-order identification of convolutive channels," in *Proc. ICASSP'97, vol. V*, 1997.

148. P. J. Green, "Reversible jump markov chain monte carlo computation and Bayesian model determination," *Biometrika*, vol. 82, pp. 711-732, 1995.

149. M. Grigoriu, *Applied Non-Gaussian Processes*, Prentice-Hall, 1995.

150. D. M. Gruenbacher and D. R. Hummels, "A simple algorithm for generating discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 42, pp. 3276-3278, 1994.

151. M. Hämäläinen, R. Hari, R. Ilmoniemi, J. Knuutila, and O. Lounasmaa, "Magnetoencephalography—Theory, Instrumentation, and Applications to Noninvasive Studies of Signal Processing of the Human Brain," *Rev. Mod. Phys.*, vol. 65, pp. 413-497, April 1993.

152. E. Hannan and M. Deistler, *The Statistical Theory of Linear Systems*, Wiley, New York, 1988.

153. M. Hansson and G. Salomonsson, "A multiple window method for estimation of peaked spectra," *IEEE Transactions on Signal Processing*, vol. 45, pp. 778-781, 1997.

154. H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. on Commun.*, pp. 774-780, Aug 1972.

155. W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, pp. 97-109, 1970.

156. G. Hatke, "Performance analysis of the SuperCART antenna array," Technical Report #AST-22, MIT Lincoln Laboratory, Lexington, MA, March 1992.

157. D. Hatzinakos and C. L. Nikias, "Blind equalization using a tricepstrum based algorithm," *IEEE Trans. on Communications*, vol. 39, pp. 669-682, May 1991.

158. M. Hawkes and A. Nehorai, "Effects of sensor placement on acoustic vector-sensor array performance," Submitted to *IEEE J. Oceanic Eng.*

159. M. Hawkes and A. Nehorai, "Surface-mounted acoustic vector-sensor array processing," in *Proc. IEEE Intl Conf. on Acoust., Speech, and Sig. Proc. (ICASSP96)*, pp. 3710-3713, Atlanta, GA, May 1996.

160. M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and Capon direction estimation," *IEEE Transactions on Signal Processing*, vol. 46, September 1998.

161. S. Haykin, *Array Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ., 1984.

162. S. Haykin, editor, *Blind Deconvolution*, Prentice-Hall, 1994.

163. S. Haykin, J. Litva, and T. Shepherd, editors, *Radar Array Processing*, Springer-Verlag, Berlin, 1993.

164. S. Haykin and D. J. Thomson, "Signal detection in a nonstationary environment reformulated as an adaptive pattern classification problem," *Proc. IEEE*, vol. 86, pp. in press, 1998.

165. A. O. Hero and J. A. Fessler, "Convergence in norm for alternating expectation-maximization (EM) type algorithms," *Statistica Sinica*, vol. 5, no. 1, pp. 41-54, 1995.

166. M. J. Hinich and C. R. Wilson, "Time delay estimation using the cross bispectrum," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-40, 1992.

167. M. J. Hinich and C. R. Wilson, "Detection of non-Gaussian signals in non-Gaussian noise using the bispectrum," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, pp. 1126-1131, 1990.

168. M. Hinich, "Testing for Gaussianity and linearity of a stationary time series," *J. Time Series Analysis*, vol. 3, pp. 169-176, 1982.

169. L. Hinnov and J. Park, "Multi-windowed spectrum estimates of the ILS polar motion," in *The Earth's Rotation and Reference Frames for Geodesy and*

Geodynamics, A. Babcock and G. Wilkins, editors, pp. 221-226, Kluwer Academic, Dordrecht, 1988.

170. F. Hlawatsch and G. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21-68, 1992.

171. K.-C. Ho, K.-C. Tan, and A. Nehorai, "Estimation of Directions-of-arrival of Completely Polarized and Incompletely Polarized Signals with Electromagnetic Vector Sensors," in *11th IFAC Symp. on Syst. Ident.*, volume 2, pp. 523-528, Kitakyushu City, Japan, July 1997.

172. K.-C. Ho, K.-C. Tan, and W. Ser, "An investigation on number of signals whose directions-of-arrival are uniquely determinable with an electromagnetic vector sensor," *Signal Processing*, vol. 47, pp. 41-54, November 1995.

173. K.-C. Ho, K.-C. Tan, and B. Tan, "Efficient method for estimating directions-of-arrival of partially polarized signals with electromagnetic vector sensors," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2485-2498, October 1997.

174. T. Ho, P. Fieguth, and A. Willsky, "Multiresolution stochastic models for efficient solution of large-scale space-time estimation problems," in *ICASSP'96*, volume VI, pp. 3098-4001, Atlanta, GA, 1996.

175. B. Hochwald and A. Nehorai, "Identifiability in array processing models with vector-sensor applications," *IEEE Transactions on Signal Processing*, vol. 44, pp. 83-95, January 1996.

176. B. Hochwald and A. Nehorai, "Magnetoencephalography with diversely-oriented and multi-component sensors," *IEEE Trans. Biomed. Eng.*, vol. 34, pp. 40-50, January 1997.

177. M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. on Info. Theory*, pp. 944-96, July 1995.

178. H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 12, pp. 601-609, 1994.

179. S. Jaffard, "Exposants de Hölder en des Points Donnes et Coefficients d'Ondelettes," *C.R.A.S.Paris*, vol. 1, no. 308, pp. 79-81, 1989.

180. M. Jamshidian and R. I. Jennrich, "Conjugate gradient acceleration of the EM algorithm," *J. Am. Statist. Assoc.*, vol. 88, no. 421, pp. 221-228, 1993.

181. A. Janicki and A. Weron, "Can one see α-stable variables and processes?," *Statistical Science*, vol. 9, no. 1, pp. 109-126, 1994.

182. J. Jao, "A matched array beamforming technique for low angle radar tracking in multipath," in *IEEE National Radar Conference*, pp. 171-176, 1994.

183. J. Jeong and W. Williams, "Kernel design for reduced interference distributions," *IEEE Transactions on Signal Processing*, vol. 40, pp. 402-412, 1992.

184. A. Jeremic and A. Nehorai, "Design of chemical sensor arrays for monitoring disposal sites on the ocean floor," To appear in *IEEE J. Oceanic Eng.*

185. D. Johnson and D. Dudgeon, *Array Signal Processing—Concepts and Techniques*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1993.

186. J. Johnson, D. J. Thomson, E. X. Wu, and S. C. R. Williams, "Multiple-window spectrum estimation applied to in vivo NMR spectroscopy," *J. Mag. Resonance*, vol. B 110, pp. 138-149, 1996.

187. J. O. Jonsson and A. O. Steinhardt, "The total p_{fa} of the multi-window harmonic detector and its application to real data," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1702-1705, 1993.

188. C. Jutten and J. Herault, "Blind separation of sources I. An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, no. 1, pp. 1-10, July 1991.

189. M. Kanda, "An electromagnetic near-field sensor for simultaneous electric and magnetic-field measurements," *IEEE Trans. Electromag. Compat.*, vol. 26, pp. 102-110, August 1984.

190. M. Kanda and D. Hill, "A three-loop method for determining the radiation characteristics of an electrically small source," *IEEE Trans. Electromag. Compat.*, vol. 34, pp. 1-3, February 1992.

191. M. E. Kappus and F. L. V. III, "Acoustic signature of thunder from seismic records," *J. Geophys. Res.*, vol. 96, pp. 10,989-11,006, 1991.

192. T. Kariya and B. K. Sinha, *Robustness of Statistical Tests*, Academic Press, San Diego, 1989.

193. T. R. Karl, P. D. Jones, R. W. Knight, O. R. White, W. Mende, J. Beer, and D. J. Thomson, "Testing for bias in the climate record," *Science*, vol. 271, pp. 1879-1883, 1996.

194. S. Kassam and H. Poor, "Robust techniques for signal processing: a survey," *Proc. IEEE*, vol. 73, pp. 433-481, 1985.

195. S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Trans. on Info. Theory*, vol. 36, pp. 741-761, July 1990.

196. L. Kaufman, "Implementing and accelerating the EM algorithm for positron emission tomography," *IEEE Trans. on Medical Imaging*, vol. MI-6, no. 1, pp. 37-51, 1987.

197. S. Kay, *Spectral Estimation: Theory and Application*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.

198. S. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, PTR Prentice-Hall, Englewood Cliffs, NJ, 1993.

199. S. Kay and D. Sengupta, "Recent advances in non-Gaussian autoregressive processes," in *Advances in spectrum analysis and array processing*, S. Haykin, editor, volume 1, pp. 141-210, Prentice-Hall, 1991.

200. S. M. Kay, "Asymptotically optimal detection in unknown colored noise via autoregressive modeling," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-31, no. 5, pp. 927-933, 1983.

201. S. M. Kay, "Asymptotically optimal detection in incompletely characterized non-Gaussian noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 4, pp. 627, 1989.

202. S. M. Kay and D. Sengupta, "Detection in incompletely characterized colored non-Gaussian noise via parametric modeling," *IEEE Transactions on Signal Processing*, vol. 41, no. 10, pp. 3066-3069, 1993.

203. S. Kayhan, A. El-Jaroudi, and L. Chaparro, "Evolutionary periodogram for nonstationary signals," *IEEE Transactions on Signal Processing*, vol. 42, pp. 527-1536, 1994.

204. E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. on Aerosp. Electron. and Systems*, vol. AES-22, no. 3, pp. 115-128, 1986.

205. E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Technical Report 848, M.I.T. Lincoln Laboratory, April, 1989.

206. C. G. Khatri and C. R. Rao, "Effects of estimated noise covariance matrix in optimal signal detection," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 5, pp. 671-679, 1987.

207. A. Kim and H. Krim, "Hierarchical stochastic modeling of SAR imagery for segmentation/compression," To appear in *IEEE Transactions on Signal Processing*.

208. I. P. Kirsteins, S. K. Mehta, and J. Fay, "Adaptive separation of unknown narrowband and broadband time series," in *Proc. ICASSP*, vol. 4, pp. 2525-2529, 1998.

209. D. Kletter and H. Messer, "Suboptimal detection of non-Gaussian signals by third-order analysis," *IEEE* Transactions on Acoustics, Speech, and Signal Processing, vol. 38, no. 6, pp. 901-909, 1990.

210. A. Kokaram, *Motion Picture Restoration*, Springer Verlag, New York, 1998.

211. H. Krim, S. Mallat, D. Donoho, and A. Willsky, "Best basis algorithm for signal enhancement," in *ICASSP'95*, Detroit, MI, May 1995, IEEE.

212. H. Krim and J.-C. Pesquet, "Multiresolution analysis of a class of nonstationary processes," *IEEE Trans. of Inf. Theory*, vol. 41, no. 4, pp. 1010-1020, July 1995.

213. H. Krim, D. Tucker, S. Mallat, and D. Donoho, "Near-optimal risk for best basis search," *submitted to IEEE Trans. on IT*, 1997.

214. H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research," *IEEE SP Magazine*, vol. 13, no. 4, pp. 67-94, July 1996.

215. J. Krolik, "Matched-field minimum variance beamforming in a random ocean channel," *Journal of the Acoustical Society of America*, vol. 92, no. 3, pp. 1408-1419, 1992.

216. J. Krolik, "The performance of matched-field beamformers with mediterranean vertical array data," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2605-2611, 1996.

217. J. Krolik and R. Anderson, "Maximum likelihood coordinate registration for over-the-horizon radar," *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 945-959, 1997.

218. C. Kuo, C. Lindberg, and D. J. Thomson, "Coherence established between atmospheric carbon dioxide and global temperature," *Nature*, vol. 343, pp. 709-714, 1990. (Reprinted in pp. 395-400 of *Coherence and Time Delay Estimation*, G. C. Carter, Ed., IEEE Press, 1993.).

219. W. Lam and G. Wornell, "Multiscale representation and estimation of fractal point processes," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2606-17, Nov 1995.

220. K. Lange, "A quasi-newtonian acceleration of the EM algorithm," *Statistica Sinica*, vol. 5, no. 1, pp. 1-18, 1995.

221. K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comp. Assisted Tomography*, vol. 8, no. 2, pp. 306-316, April 1984.

222. D. Lansky and G. Casella, "Improving the EM algorithm," in *Computing and Statistics: Proc. Symp. on the Interface*, C. Page and R. LePage, editors, pp. 420-424, Springer-Verlag, 1990.

223. L. J. Lanzerotti, T. P. Armstrong, R. E. Gold, C. G. Maclennan, E. C. Roelof, G. M. Simnett, D. J. Thomson, K. A. Anderson, S. E. H. III, S. M. Krimigis, R. P. Lin, M. Pick, E. T. Sarris, and S. J. Tappin, "Over the southern solar pole: low-energy interplanetary charged particles," *Science*, vol. 268, pp. 1010-1013, 1995.

224. L. J. Lanzerotti, R. E. Gold, D. J. Thomson, R. E. Decker, C. G. Maclennan, and S. M. Krimigis, "Statistical properties of shock-accelerated ions in the outer heliosphere," *Ap. J.*, vol. 380, pp. L93-L96, 1991.

225. M. Lavielle, "Stochastic algorithm for parametric and non-parametric estimation in the case of incomplete data," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 3-17, 1995.

226. M. Lavielle and E. Moulines, "A simulated annealing version of the EM algorithm for non-Gaussian deconvolution," *Statist. Comput*, to appear, 1998.

227. E. L. Lehmann, *Testing Statistical Hypotheses*, Wiley, New York, 1959.

228. H. Leung and S. Haykin, "Detection and estimation using an adaptive rational function filter," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3366-3376, 1994.

229. R. Lewitt and G. Muehllehner, "Accelerated iterative reconstruction for positron emission tomography," *IEEE Trans. on Medical Imaging*, vol. MI-5, no. 1, pp. 16-22, 1986.

230. J. Li, "Direction and polarization estimation using arrays with small loops and short dipoles," *IEEE Trans. Antennas and Prop.*, vol. 41, pp. 379-387, March 1993.

231. X. Li and M. Bilgutay, "Wiener filter realization for target detection using group delay statistics," *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2067-2074, 1993.

232. Y. Li and Z. Ding, "Global convergence of fractionally spaced godard adaptive equalizers," *IEEE Transactions on Signal Processing*, pp. 818-826, April 1996.

233. K. S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *The Annals of Statistics*, vol. 10, pp. 1195-1208, 1982.

234. J. Lilly and J. Park, "Multiwavelet spectral and polarization analysis of seismic records," *Geophys. J. Intl.*, vol. 122, pp. 1001-1021, 1995.

235. C. R. Lindberg, *Multiple taper spectral analysis of terrestrial free oscillations*, PhD thesis, Univ. Calif., San Diego, 1986.

236. C. R. Lindberg and J. Park, "Multiple-taper spectral analysis of terrestrial free oscillations: part II," *Geophys. J. Royal Astr. Soc.*, vol. 91, pp. 795-836, 1987.

237. C. R. Lindberg and D. J. Thomson, "Comment on "a new method of spectral analysis and its application to the earth's free oscillations: the 'sompi' method," by S. Hori et al.," *J. Geophys. Res.*, vol. 95, pp. 12, 785-12, 788, 1990.

238. C. R. Lindberg and D. J. Thomson, *Method and Apparatus for Detecting Control Signals*, 1995. U.S. Patent 5,442,696.

239. C. Liu and D. B. Rubin, "The ECME algorithm: a simple extension of EM and ECM with fast monotone convergence," *Biometrika*, vol. 81, pp. 633-648, 1994.

240. C. H. Liu, D. B. Rubin, and Y. N. Wu, "Parameter expansion for EM acceleration - the PX-EM algorithm," *Biometrika*, to appear, 1998.

241. H. Liu and G. Xu, "A subspace method for signature waveform estimation in synchronous CDMA systems," *IEEE Trans. Communications*, pp. 1346-1354, Nov. 1996.

242. R. Liu and L. Tong, editors, *Special Issue on Blind Channel Identification and Signal Estimation*, 1998.

243. T.-C. Liu and B. D. Van Veen, "Multiple window based minimum variance spectrum estimation for multidimensional random fields," *IEEE Trans. SP*, vol. 40, pp. 578-589, 1992.

244. L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

245. L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1987.

246. Loughlin, J. Pitton, and L. Atlas, "Construction of positive time-frequency distributions," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2697-2705, 1994.

247. T. A. Louis, "Finding the observed information matrix when using the EM algorithm," *J. Royal Statistical Society, Ser. B*, vol. 44, no. 2, pp. 226-233, 1982.

248. R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell Syst. Tech. Journal*, pp. 255-286, February 1966.

249. M. Luettgen and A. Willsky, "Multiscale smoothing error models," *IEEE Trans. on Automatic Control*, 1994.

250. R. Lupas and S. Verdú Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, pp. 123-136, 1989.

251. R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. on Comm.*, pp. 496-508, April 1990.

252. X. Ma and C. Nikias, "Parameter estimation and blind channel identification in impulsive signal environments," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2884-2897, Dec 1995.

253. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, Boston, 1998.

254. B. Mandelbrot, *The Fractal Geometry Of Nature*, W.H. Freeman and company, 1977.

255. B. Mandelbrot, *The fractal geometry of nature (2nd Edition)*, W.H. Freeman and Co., 1983.

256. W. Martin and P. Flandrin, "Wigner-Ville spectral analysis for nonstationary processes," *IEEE Transactions on Acoustics, Speech, and Signal, Processing*, vol. 33, pp. 1461-1471, 1995.

257. E. Masry, "The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion," *IEEE Trans. on IT*, vol. IT-39, pp. 260-264, Jan. 1993.

258. G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1520-1534, 1997.

259. E. J. McCoy, A. T. Walden, and D. B. Percival, "Multitaper spectral estimation of power law processes," *IEEE Transactions on Signal Processing*, vol. 46, pp. 655-668, 1998.

260. G. McLachlan and T. Krishnan, *The EM algorithm and extensions*, Wiley, 1997.

261. L. T. McWhorter and L. L. Scharf, "Multiwindow estimators of correlation," *IEEE Transactions on Signal Processing*, vol. 46, pp. 440-448, 1998.

262. I. Meilijson, "A fast improvement to the EM algorithm on its own terms," *J. Royal Statistical Society, Ser. B*, vol. 51, no. 1, pp. 127-138, 1989.

263. R. Mellors, F. L. Vernon, and D. J. Thomson, "Detection of dispersive signals using multitaper double frequency coherence," *Geophys. J. Intl.*, to appear, 1998.

264. J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, pp. 278-305, March 1991.

265. X. L. Meng and D. Van Dyk, "The EM algorithm - an old folk-song sung to a fast new tune," *J. Royal Statistical Society, Ser. B*, vol. 59, no. 3, pp. 511-567, 1997.

266. H. Messer, "Potential performance gain in using spectral information in passive detection/localization of wideband sources," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2964-2974, 1995.

267. H. Messer and S. Tsruya, "Performance analysis of time delay estimation of a signal with unknown spectral parameters," in *Proceedings of the Fifth* ASSP Workshop on Spectrum Estimation and Modeling, IEEE, 1990.

268. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equations of state calculations by fast computing machines," *Journal of Chemical Physics*, vol. 21, pp. 1087-1092, 1953.

269. D. Middleton, "Channel modeling and threshold signal processing in underwater acoustics: an analytic overview," *IEEE J Oceanic Engg*, vol. 12, pp. 4-28, 1987.

270. D. Middleton and A. Spaulding, "Elements of weak-signal detection in non-Gaussian noise," in *Advances in statistical signal processing - vol 2: Signal Detection*, H. Poor and J. Thomas, editors, pp. 137-215, JAI press, Greenwich, CT, 1993.

271. M. I. Miller and D. R. Fuhrmann, "Maximum-likelihood narrow-band direction finding and the EM algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 9, pp. 1560, 1990.

272. J. D. Molle and M. Hinich, "Tripsectral analysis of stationary time series," *J. Acoustical Soc. America*, vol. 97, 1995.

273. T. Moon, "The expectation maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47-60, 1996.

274. J. Mosher, P. Lewis, and R. Leahy, "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 541-557, June 1992.

275. P. Moulin, "A wavelet regularization method for diffuse radar target imaging and speckle noise reduction," *Journ. Math. Imaging and Vision*, vol. 3, pp. 123-134, 1993.

276. E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel fir filters," *IEEE Transactions on Signal Processing*, vol. 43, pp. 516-525, 1995.

277. E. Moulines and P. Soulier, "Fractional exponential model for fractal point processes," in *Proc. Thirty-First Asilomar Conf on Signals, Systems and Computers*, pp. 1107-11, Monterey, CA, Nov 1997.

278. R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, Wiley, New York, 1982.

279. C. F. Mullins, "Multiple window cumulant estimation," in *Proc. IEEE SP workshop on Higher-Order statistics*, p. M1.4, South Lake Tahoe, CA, 1993.

280. T. C. Mullis and L. L. Scharf, "Quadratic estimators of the power spectrum," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, editor, volume 1, pp. 1-57, Prentice-Hall, 1991.

281. S. Nam and E. Powers, "Applications of higher order spectral analysis to cubically nonlinear system identification," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2124-2135, 1994.

282. Z. Nan and A. Nehorai, "Detection of ship wake using an airborne magnetic transducer," in *Proc. 32nd Asilomar Conf. on Signals, Syst. and Comput.*, Pacific Grove, CA, November 1998.

283. A. Nehorai, K.-C. Ho, and B. Tan, "Minimum-noise-variance beamformer with an electromagnetic vector sensor," in *Proc. IEEE Intl Conf. on Acoust., Speech, and Sig. Proc. (ICASSP98)*, pp. 2021-2024, Seattle, WA, May 1998.

284. A. Nehorai and E. Paldi, "Vector-sensor processing for electromagnetic source localization," in *Proc. 25th Asilomar Conf. on Signals, Syst. and Comput.*, pp. 566-572, Pacific Grove, CA, November 1991.

285. A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," in *Proc. 26th Asilomar Conf. on Signals, Syst. and Comput.*, pp. 192-198, Pacific Grove, CA, October 1992.

286. A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2481-2491, September 1994.

287. A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Transactions on Signal Processing*, vol. 42, pp. 376-398, February 1994.

288. A. Nehorai, B. Porat, and E. Paldi, "Detection and localization of vapor-emitting sources," *IEEE Transactions on Signal Processing*, vol. 43, pp. 243-253, January 1995.

289. L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: an EM-based approach," *IEEE Trans. on Communications*, vol. 44, no. 12, pp. 1700-1710, 1996.

290. J. Nickles, G. Edmonds, R. Harriss, F. Fisher, W. Hodgkiss, J. Giles, and G. D'Spain, "A vertical array of directional acoustic sensors," in *Proc. Mast. Oceans thru Tech. (Oceans 92)*, pp. 340-345, Newport, RI, October 1992.

291. C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, pp. 10-37, July 1993.

292. C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

293. C. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*, John Wiley & Sons, 1995.

294. J. Nolan, "Parameter estimation and data analysis for stable distributions," in *Proc. Thirty-First Asilomar Conf on Signals, Systems and Computers*, pp. 443-47, Monterey, CA, Nov 1997.

295. J. Nolan, "Multivariate stable distributions: approximation, estimation, simulation and identification," in *A practical guide to heavy tails: statistical techniques for analyzing heavy tailed distributions*, R. F. R.J. Adler and M. Taqqu, editors, Birkhauser, 1998.

296. R. Onn and A. Steinhardt, "A multiwindow method for spectrum estimation and sinusoid detection in an array environment," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3006-3015, 1994.

297. J. J. K. O'Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Springer Verlag, New York, 1996.

298. B. Ottersten, "Array Processing for Wireless Communications," in *Proc. 8th Workshop on Stat. Sig. and Array Proc.*, pp. 466-473, Corfu, Greece, June 1996.

299. B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing," in *Radar Array Processing*, Haykin, Litva, and Shepherd, editors, pp. 99-151, Springer-Verlag, Berlin, 1993.

300. L. Pakula and S. M. Kay, "Detection performance of the circular correlation coefficient receiver," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-34, no. 3, pp. 399-404, 1986.

301. M. Papazoglou and J. Krolik, "Multi-dwell matched-field altitude estimation for over the-horizon radar," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, 1998.

302. J. Park, "Envelope estimation for quasi-periodic geophysical signals in noise: a multitaper approach," in *Statistics in the Environmental and Earth Sciences*, A. Walden and P. Guttorp, editors, pp. 189-219, Edwin Arnold, London, 1992.

303. J. Park, F. V. III, and C. Lindberg, "Frequency dependent polarization analysis of high-frequency seismograms," *J. Geophys. Res.*, vol. 92, pp. 12,664-12,674, 1987.

304. J. Park, C. Lindberg, and F. V. III, "Multitaper spectral analysis of high-frequency seismograms," *J. Geophys. Res.*, vol. 92, pp. 12,675-12,684, 1987.

305. J. Park, C. Lindberg, and D. Thomson, "Multiple-taper spectral analysis of terrestial free oscillations: part I," *Geophys. J. Royal Astr. Soc.*, vol. 91, pp. 755-794, 1987.

306. A. Paulraj and C. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, pp. 49-83, 1997.

307. D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications; Multitaper and Conventional Univariate Techniques*, Cambridge Univ. Press, 1993.

308. J.-C. Pesquet, H. Krim, and H. Carfantan, "Time invariant orthonormal wavelet representations," *IEEE Transactions on Signal Processing*, Aug. 1996.

309. J. Pesquet, H. Krim, H. Leporini, and E. Hamman, "Bayesian approach to the best basis selection," in *ICASSP'96*, Atlanta, GA, May 1996.

310. B. Picinbono and P. Duvaut, "Optimal linear quadratic systems for detection and estimation," *IEEE Trans. on Inform. Theory*, vol. IT-34, no. 2, pp. 304-311, 1988.

311. S. U. Pillai, *Array Signal Processing*, Springer Verlag, 1989.

312. J. W. Pitton, "Nonstationary spectrum estimation and time-frequency concentration," in *Proc. ICASSP*, volume 4, pp. 2425-2428, 1998.

313. V. Poor and G. Wornell, *Wireless Communication: Signal Processing Perspectives*, Prentice Hall, New Jersey, 1998.

314. B. Popescu, P. Costa, P. Larzabal, and H. Clergeot, "Extraction of parameters from a wavelet packet analysis for terrain matching," in *Proc. of the 3-rd IEEE Int. Workshop on Time-Scale, Time-Frequency*, pp. 423-426, Paris, 18-21 Juin, 1996.

315. B. Porat, *Digital Processing of Random Signals, Theory & Methods*, Prentice Hall, Englewood Cliffs, NJ, 1994.

316. B. Porat and B. Friedlander, "Adaptive detection of transient signals," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-34, no. 6, pp. 1410-1418, 1986.

317. B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Transactions on Signal Processing*, vol. 39, pp. 522-526, 1991.

318. B. Porat and B. Friedlander, "Estimation of frequency in the presence of non-random interference," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 640-651, Mar 1996.

319. B. Porat and A. Nehorai, "Localizing vapor-emmiting sources by moving sensors," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1018-1021, April 1996.

320. E. Powers, B. Boashash, and A. M. Zoubir, editors, *Higher Order Statistical Signal Processing and Applications*, Longman Cheshire, Melbourne, Australia, 1995.

321. R. Price, "Nonlinear feedback equalized PAM versus capacity for noisy filter channels," in *Proc. of Intl. Conf. on Communications*, pp. 22.12-22.17, 1972.

322. M. B. Priestley, *Non-Linear and Non-Stationary Time Series Analysis*, Academic Press, San Diego CA, 1988.

323. J. G. Proakis, *Digital Communications*, McGraw-Hill Book Company, NY, 2nd edition, 1989.

324. J. G. Proakis, *Digital Communications, Third edition*, McGraw Hill, 1995.

325. S. Qian and D. Chen, *Joint Time-Frequency Analysis - Methods and Applications*, Prentice-Hall, New Jersey, 1996.

326. S. U. H. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, pp. 1349-1387, 1985.

327. J. Ralston, A. Zoubir, and B. Boashash, "Identification of a class of nonlinear systems under stationary non-Gaussian excitation," *IEEE Transactions on Signal Processing*, vol. 45, pp. 719-735, 1997.

328. K. Ramchandran and M. Vetterli, "Best Wavelet Packets in a Rate Distortion Sense," *IEEE Trans. on Image Proc.*, pp. 160-175, 1993.

329. T. S. Rao and M. Gabr, "A test for linearity of stationary time series," *J. Time Series Analysis*, vol. 1, pp. 145-158, 1980.

330. I. S. Reed and X. Yu, "Adaptive multi-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. 38, no. 10, pp. 1760-1771, 1990.

331. I. Reed, J. Mallet, and L. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. on Aerosp. Electron. and Systems*, vol. AES-10, pp. 853-863, 1974.

332. A. Richardson and L. Nolte, "A posteriori probability source localization in an uncertain sound speed deep ocean environment," *Journal of the Acoustical Society of America*, vol. 89, no. 5, pp. 2280-2284, 1991.

333. C. D. Richmond, "Derived PDF of maximum likelihood signal estimator which employs an estimated noise covariance," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 305-315, 1996.

334. K. S. Riedel and A. Sidorenko, "Minimum bias multiple taper spectral estimation," *IEEE Transactions on Signal Processing*, vol. 43, pp. 188-195, 1995.

335. K. S. Riedel and A. Sidorenko, "Adaptive smoothing of the log-spectrum with multiple tapering," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1794-1800, 1996.

336. K. S. Riedel, A. Sidorenko, N. Bretz, and D. J. Thomson, "Spectral estimation of plasma fluctuations. II. nonstationary analysis of edge localized mode spectra," *Phys. Plasma*, vol. 1, pp. 501-514, 1994.

337. K. S. Riedel, A. Sidorenko, and D. J. Thomson, "Spectral estimation of plasma fluctuations. 1. comparison of methods," *Phys. Plasma*, vol. 1, pp. 485-500, 1994.

338. J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465-471, 1978.

339. C. P. Robert, *The Bayesian Choice*, Springer Verlag, New York, 1994.

340. D. A. Roberts., K. W. Ogilvie, M. L. Goldstein, D. J. Thomson, C. G. Maclennan, and L. J. Lanzerotti, "The nature of the solar wind," *Nature*, vol. 381, pp. 31-32, 1996.

341. F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 12, pp. 2964-2974, 1992.

342. D. B. Rubin, "Using the SIR algorithm to simulate posterior distributions," in *Bayesian Statistics 3*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith, editors, pp. 395-402, University Press, Oxford, UK, 1988.

343. A. Ruiz, J. M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. on Communications*, pp. 1012-1029, 1992.

344. B. Ryu and S. Lowen, "Point-process approaches to the modeling and analysis of self-similar traffic," in *Proc IEEE Conf Infocom*, pp. 1468-75, 1996.

345. B. M. Sadler, "Detection in correlated impulsive noise using fourth-order cumulants," *Signal Processing*, vol. 44, no. 11, pp. 2793-2800, 1996.

346. B. M. Sadler, G. Giannakis, and K. Li, "Estimation and detection in non-Gaussian noise using higher order statistics," *IEEE Transactions on Signal Processing*, vol. 42, no. 4, pp. 2729-2741, 1994.

347. N. Saito, *Local feature extraction and its applications using a library of bases*, PhD thesis, Yale University, Dec. 1994.

348. G. Samorodnitsky and M. Taqqu, *Stable non-Gaussian random processes*, Chapman & Hall, 1994.

349. S. D. Sandberg and M. A. Tzannes, "Overlapped discrete multitone modulation for high speed copper wire communications," *IEEE Journal on Selected Areas in Communications*, pp. 1571-1585, 1995.

350. A. Sayeed and D. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2872-2883, 1995.

351. A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Transactions on Information Theory*, 1999.

352. A. Scaglione and G. B. Giannakis, "Code-only dependent asynchronous CDMA receivers for mui elimination and mitigation of unknown multipath," in *Proc. of 31st Asilomar Conf. on Signals, Systems, and Computers*, pp. 950-954, 1997.

353. A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers, parts I & II," to appear in *IEEE Transactions on Signal Processing*, 1999.

354. L. L. Scharf, "Invariant Gauss-Gauss detection," *IEEE Trans. on Inform. Theory*, vol. IT-19, no. 3, pp. 422-427, 1973.

355. L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Reading, MA, 1991.

356. L. L. Scharf and D. W. Lytle, "Signal detection in Gaussian noise of unknown level: an invariance application," *IEEE Trans. on Inform. Theory*, vol. IT-17, no. 3, pp. 404-411, 1971.

357. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146-2157, 1994.

358. I. Schick and H. Krim, "Robust wavelet denoising," in *ICASSP'97*, Munich, Germany, May 1997, IEEE.

359. R. E. Schild and D. J. Thomson, "The Q0957+561 time delay, quasar structure, and microlensing," in *Astronomical time series*, D. M. et al., editor, pp. 73-84, Kluwer Academic Publishers, Dordrecht, 1997.

360. H. Schmidt, A. Baggeroer, W. Kuperman, and E. Sheer, "Environmentally-tolerant beamforming for high-resolution matched-field processing: Deterministic mismatch," *Journal of the Acoustical Society of America*, vol. 88, pp. 1802-1810, 1990.

361. J. Schoukens and R. Pintelon, *Identification of Linear Systems*, Pergamon Press, Oxford, UK, 1991.

362. P. Schultheiss, H. Messer, and G. Shor, "Maximum likelihood time delay estimation in non-Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 45, 1997.

363. T. J. Schultz, "Fast algorithm for maximum-likelihood imaging with coherent speckle measurements," in *Proc. IEEE Conf on Image Proc.*, pp. 679-682, Santa Barbara, CA, 1997.

364. T. J. Schultz, "Penalized maximum-likelihood estimation of covariance matrices with linear structure," *IEEE Transactions on Signal Processing*, vol. 45, no. 12, pp. 3027-3038, 1997.

365. G. Schwartz, "Estimating the dimension of a model," *Ann. Stats.*, vol. 6, pp. 4611-4640, 1978.

366. R. E. Schwartz, "Minimax CFAR detection in additive Gaussian noise of unknown covariance," *IEEE Trans. on Inform. Theory*, vol. 15, no. 4, pp. 722-725, July 1969.

367. D. Sengupta and S. M. Kay, "Parameter estimation and GLRT detection in colored non-Gaussian autoregressive processes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1661-1676, 1990.

368. E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization using modulation induced cyclostationarity," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1930-1944, 1998.

369. M. S. Seymour and S. Haykin, "ISAR using Thomson's multiwindow adaptive spectrum estimation method," *IEEE Trans. Aero. and Elec. Sys.*, vol. 29, pp. 1065-1070, 1993.

370. O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. on Information Theory*, vol. 36, pp. 312-321, 1990.

371. I. Sharfer and H. Messer, "The bispectrum of sampled data: part II - Monte Carlo simulations of detection and estimation of the sampling jitter," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2706-2714, 1994.

372. L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. on Medical Imaging*, vol. MI-1, No. 2, pp. 113-122, Oct. 1982.

373. N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Deterministic waveform-preserving blind separation of DS-CDMA signals using an antenna array," in *Proc. SSAP'98, Portland, Oregon*, 1998.

374. M. Simon, J. Omura, R. Scholtz, and B. Levitt, *Spread Spectrum Communication Handbook*, McGraw-Hill, New York, 1994.

375. D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: the discrete case," *Bell System Tech. J.*, vol. 57, pp. 1371-1429, 1978.

376. D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. ICASSP'94*, vol. IV, pp. 585-588, 1994.

377. T. Soderstrom and P. Stoica, *System Identification*, Prentice Hall Intern., London, 1989.

378. I. Song and S. Kassam, "LOD of signals in a generalized observation model," *IEEE Trans Info Theory*, vol. 36, no. 3, pp. 502-15;516-30, May 1990.

379. H. W. Sorenson, *Parameter Estimation*, Marcel Dekker, New York, 1980.

380. L. Stankovic and S. Stankovic, "An analysis of instantaneous frequency representation using time-frequency distributions-generalized Wigner distribution," *IEEE Transactions on Signal Processing*, vol. 43, 1995.

381. P. Stoica, M. Cedervall, and A. Eriksson, "Combined instrumental variable and subspace fitting approach to parameter estimation of noisy input-output systems," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2386-2397, 1995.

382. P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, Inc., Upper Saddle River, NJ, 1997.

383. E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. on Comm.*, pp. 84-93, Jan. 1996.

384. A. Swami, "Multiplicative noise models: parameter estimation using cumulants," *Signal Processing*, vol. 36, no. 3, pp. 355-373, April 1994.

385. A. Swami, "Cramer-Rao bounds for deterministic signals in additive and multiplicative noise," *Signal Processing*, vol. 53, no. 2-3, pp. 231-244, Sep 1996.

386. A. Swami, G. B. Giannakis, and S. Shamsunder, "Multichannel ARMA processes," *IEEE Transactions on Signal Processing*, vol. 42, pp. 898-913, 4 1994.

387. A. Swami, G. Giannakis, and G. Zhou, "Bibliography on Higher-Order Statistics," *Signal Processing*, vol. 60, no. 1, pp. 65-126, July 1997.

388. A. Swami and B. M. Sadler, "TDE, DOA, and related parameter estimation problems in impulsive noise," in *Proc. IEEE SP Workshop on HOS*, pp. 273-77, Banff, Canada, 1997.

389. A. Swami and B. M. Sadler, "Parameter estimation for linear alpha-stable processes," *IEEE Signal Processing Letters*, vol. 5, no. 2, pp. 48-50, Feb 1998.

390. K.-C. Tan, K.-C. Ho, and A. Nehorai, "Uniqueness study of measurements obtainable with arrays of electromagnetic vector sensors," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1036-1039, April 1996.

391. H. Tanizaki, *Nonlinear filters: Estimation and Applications*, Springer Verlag, New York, 1996.

392. M. A. Tanner, *Tools for Statistical Inference*, Springer Verlag, New York, 1996.

393. L. Tauxe, "Sedimentary records of relative paleointensity of the geomagnetic field: theory and practice," *Rev. Geophysics*, vol. 31, pp. 319-354, 1993.

394. L. Tauxe and G. Wu, "Normalized remanence in sediments of the western equatorial Pacific: relative paleointensity of the geomagnetic field?," *J. Geophys Res.*, vol. 95, pp. 12,337-12,350, 1990.

395. A. Tewfik and A. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Trans. on IT*, vol. IT-38, pp. 904-909, 1992.

396. J. Thompson, P. Grant, and B. Mulgrew, "Smart Antenna Arrays for CDMA Systems," *IEEE Personal Comm. Magazine*, vol. 3, no. 5, pp. 16-25, Oct. 1996.

397. D. J. Thomson, "Spectrum estimation techniques for characterization and development of WT4 waveguide," *Bell System Tech. J.*, vol. 56, pp. *Part I*, 1769-1815, *Part II*, 1983-2005, 1977.

398. D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, pp. 1055-1096, 1982.

399. D. J. Thomson, "Multi-window bispectrum estimates," in *Proc. IEEE Workshop on Higher-order spectral analysis*, pp. 19-23, Vail, Colorado, 1989.

400. D. J. Thomson, "Quadratic-inverse spectrum estimates: applications to paleoclimatology," *Phil. Trans. R. Soc. Lond.*, vol. A 332, pp. 539-597, 1990.

401. D. J. Thomson, "Time series analysis of holocene climate data," *Phil. Trans. R. Soc. Lond.*, vol. A 330, pp. 601-616, 1990.

402. D. J. Thomson, "Non-stationary fluctuations in "stationary" time series," *Proc. SPIE*, vol. 2027, pp. 236-244, 1993.

403. D. J. Thomson, "An overview of multiple-window and quadratic-inverse spectrum estimation methods," in *Proc. ICASSP*, volume VI, pp. 185-94, 1994.

404. D. J. Thomson, "Projection filters for data analysis," in *Proc. Seventh IEEE SP Workshop on Stat. Sig. and Array Proc.*, pp. 39-42, Quebec, 1994.

405. D. J. Thomson, "The seasons, global temperature, and precession," *Science*, vol. 268, pp. 59-68, 1995.

406. D. J. Thomson, "Dependence of global temperatures on atmospheric CO_2 and solar irradiance," *Proc. Natl. Acad. Sci. USA*, vol. 94, pp. 8370-8377, 1997.

407. D. J. Thomson and A. D. Chave, "Jackknifed error estimates for spectra, coherences, and transfer functions," in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, editor, volume 1, pp. 58-113, Prentice-Hall, 1991.

408. D. J. Thomson, C. G. Maclennan, and L. J. Lanzerotti, "Propagation of solar oscillations through the interplanetary medium," *Nature*, vol. 376, pp. 139-144, 1995.

409. D. J. Thomson and R. Schild, "Processes with level-dependent delay," in *Proc. IEEE SP workshop on Higher-Order statistics*, pp. 374-378, South Lake Tahoe, CA, 1993.

410. D. J. Thomson and R. Schild, "Time delay estimates for Q0957+561 A, B," in *Applications of time series analysis in astronomy and meteorology*, T. S. Rao, M. B. Priestley, and O. Lessi, editors, pp. 187-204, Chapman and Hall, London, 1997.

411. M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electron. Lett.*, vol. 7, pp. 138-139, 1971.

412. H. Tong, *Non-linear Time Series*, Oxford University Press, New York, 1990.

413. L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: A frequency-domain approach," *IEEE Trans. on Information Theory*, vol. 41, pp. 329-334, 1995.

414. L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Trans. on Information Theory*, pp. 340-349, 1994.

415. C. Tontiruttananon and J. Tugnait, "Identification of closed-loop linear systems via cyclic spectral analysis: An equation-error formulation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, 1998.

416. J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-31, pp. 459-472, April 1983.

417. P. Troughton and S. J. Godsill, "Bayesian model selection for time series using markov chain monte carlo," in *ICASSP*, pp. 3733-3736, 1997.

418. P. Tsakalides and C. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Transactions on Signal Processing*, pp. 1613-1622, Jul 1996.

419. M. K. Tsatsanis, "Inverse filtering criteria for CDMA systems," *IEEE Transactions on Signal Processingg*, pp. 102-112, Jan. 1997.

420. M. K. Tsatsanis and G. B. Giannakis, "Multirate filter banks for code-division multiple access systems," in *Proc. ICASSP'95, vol. II*, pp. 1484-1487, 1995.

421. M. K. Tsatsanis and G. B. Giannakis, "Equalization of rapidly fading channels: Self recovering methods," *IEEE Transactions on Communications*, vol. 44, pp. 619-630, 1996.

422. M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *International Journal of Adaptive Control and Signal Processing*, vol. 10, pp. 159-176, 1996.

423. M. K. Tsatsanis and G. B. Giannakis, "Optimal linear receivers for DS-CDMA systems: a signal processing approach," *IEEE Transactions on Signal Processing*, vol. 44, pp. 3044-3055, 1996.

424. M. K. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1785-1794, 1997.

425. M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Processing*, vol. 53, pp. 211-229, 1996.

426. M. K. Tsatsanis and Z. Xu, "On minimum output energy cdma receivers in the presence of multipath," in *Proc. 31st Conf. on Info. Sciences and Systems*, pp. 724-729, Johns Hopkins Univ., Baltimore MD, March 19-21, 1997.

427. J. Tugnait, "On time delay estimation with unknown spatially correlated Gaussian noise using fourth-order cumulants and cross cumulants," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-39, 1991.

428. J. Tugnait, "Detection of non-Gaussian signals using integrated polyspectrum," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3137-3149, 1994.

429. J. Tugnait, "Corrections to "detection of non-Gaussian signals using integrated polyspectrum," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2792-2793, 1995.

430. J. K. Tugnait, "Identification of linear stochastic systems via second- and fourth-order cumulant matching," *IEEE Trans. on Information Theory*, vol. 33, pp. 393-407, 1987.

431. J. K. Tugnait, "Identification of multivariable stochastic linear systems via polyspectral analysis given noisy input-output time-domain data," *IEEE Trans. Automatic Control*, vol. 43, pp. to appear, 1998.

432. J. K. Tugnait, "Identification of multivariable stochastic linear systems via spectral analysis given time-domain data," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1458-1463, 1998.

433. J. K. Tugnait and C. Tontiruttananon, "Identification of linear systems via spectral analysis given time-domain data: consistency, reduced-order approximation and performance analysis," *IEEE Trans. Automatic Control*, vol. 43, pp. to appear, 1998.

434. J. K. Tugnait and Y. Ye, "Stochastic system identification with noisy input-output measurements using polyspectra," *IEEE Trans. Automatic Control*, vol. 40, pp. 670-683, 1995.

435. J. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher-order statistics and inverse filter criteria," *IEEE Transactions on Signal Processing*, vol. 45, pp. 658-672, 1997.

436. J. W. Tukey, "An introduction to the calculations of numerical spectrum analysis," in *Spectral Analysis of Time Series*, B. Harris, editor, pp. 25-46, J. Wiley and Sons, 1967.

437. P. P. Vaidyanathan, *Multirate systems and filter banks*, Prentice Hall, New Jersey, 1992.

438. P. van Overschee and B. de Moor, *Subspace Identification for Linear Systems: Theory - Implementation-Methods*, Kluwer Academic, Boston, MA, 1996.

439. B. D. Van Veen and K. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4-24, April 1988.

440. B. D. Van Veen and L. L. Scharf, "Estimation of structured covariance matrices and multiple window spectrum analysis," *IEEE Transactions on Signal Processing*, vol. 38, pp. 1467-1471, 1990.

441. S. Verdú, "Multiuser detection," in *Advances in Statistical Signal Processing*, V. P., editor, pp. 369-409, JAI, NY, 1997.

442. F. L. Vernon III, *Analysis of data recorded on the ANZA seismic network*, PhD thesis, Univ. Calif., San Diego, 1989.

443. M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, Englewood Cliffs, NJ, 1995.

444. M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436-2449, 1991.

445. B. Vidakovic, "Nonlinear wavelet shrinkage with Bayes rules and Bayes," *Journal of the American Statistical Association*, vol. 93, pp. 173-179, 1998.

446. J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, pp. 61-74, 1948.

447. W. L. W. Willinger, M.S. Taqqu and D. Wilson, "On the self-similar nature of ethernet traffic," *IEEE/ACM Trans on Networking*, vol. 2, pp. 1-15, 1994.

448. A. T. Walden, "Multitaper estimation of the innovation variance of a stationary time series," *IEEE Transactions on Signal Processing*, vol. 43, pp. 181-187, 1995.

449. E. Wegman, S. Schwartz, and J. Thomas, editors, *Topics in Non-Gaussian Signal Processing*, Springer-Verlag, New York, 1989.

450. B. Widrow and S. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1985.

451. E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physics Review*, vol. 40, pp. 749-759, 1932.

452. K. M. Wong and S. Chen, "Detection of narrow-band sonar signals using order statistical filters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 5, pp. 597-613, 1987.

453. K. Wong and M. Zoltowski, "High accuracy 2D angle estimation with extended aperture vector sensor arrays," in *Proc. IEEE Intl Conf. on Acoust., Speech, and Sig. Proc. (ICASSP96)*, pp. 2789-2792, Atlanta, GA, 1996.

454. K. Wong and M. Zoltowski, "Closed-form underwater acoustic direction-finding with arbitrarily spaced vector-hydrophones at unknown locations," *IEEE J. Oceanic Eng.*, vol. 22, pp. 649-658, October 1997.

455. K. Wong and M. Zoltowski, "Extended-aperture underwater acoustic multisource azimuth/elevation direction-finding using uniformly but sparsely spaced vector hydrophones," *IEEE J. Oceanic Eng.*, vol. 22, pp. 659-672, October 1997.

456. K. Wong and M. Zoltowski, "Uni-vector-sensor ESPRIT for multisource azimuth, elevation, and polarization estimation," *IEEE Trans. Antennas and Prop.*, vol. 45, pp. 1467-1474, October 1997.

457. K. Wong and M. Zoltowski, "Closed-form direction-finding with arbitrarily spaced electromagnetic vector-sensors at unknown locations," in *Proc. IEEE Intl Conf. on Acoust., Speech, and Sig. Proc. (ICASSP98)*, pp. 1949-1952, Seattle, WA, May 1998.

458. G. Wornell, "Spread-signature CDMA: Efficient multiuser communications in the presence of fading," *IEEE Trans. on Info. Theory*, pp. 1418-1438, Sept. 1995.

459. G. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proceedings of the IEEE*, pp. 586-603, 1996.

460. G. Wornell, *Signal Processing with Fractals: A Wavelet-Based Approach*, Prentice-Hall, Upper Saddle River, NJ, 1995.

461. C. F. J. Wu, "On the convergence properties of the EM algorithm," *Annals of Statistics*, vol. 11, pp. 95-103, 1983.

462. G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2982-2993, 1995.

463. J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. on Communications*, pp. 3221-3231, 1994.

464. X. Yu, I. S. Reed, and A. D. Stocker, "Comparative performance analysis of adaptive multi-spectral detectors," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2639-2655, 1993.

465. Y. Zhao, L. Atlas, and R. Marks, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1084-1091, 1990.

466. G. Zhou and G. Giannakis, "Harmonics in Gaussian multiplicative and additive noise: CRBs," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1217-31, May 1995.

467. G. Zhou and G. Giannakis, "Harmonics in Gaussian multiplicative and additive noise: performance analysis of cyclic statistics," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1445-60, Jun 1996.

468. T. Zhu, K.-P. Chun, and G. F. West, "High-frequency *p*-wave attenuation determination using multiple-window spectral analysis method," *Bull. Seism. Soc. Am.*, vol. 79, pp. 1054-1069, 1989.

469. V. Zolotarev, *One-dimensional stable distributions*, American Mathematical Society, 1986.

470. M. Zoltowski and K. Wong, "Polarization diversity and extended-aperture spatial diversity to mitigate fading-channel effects with a sparse array of electric dipoles or magnetic loops," in *Proc. IEEE Int. Vehic. Tech. Conf.*, pp. 1163-1167, 1997.

471. W. Y. Zou and Y. Wu, "OFDM: An overview," *IEEE Trans. on Broadcasting*, vol. 41, pp. 1-8, 1995.

472. A. M. Zoubir, "The bootstrap and its applications," in *Proc. ICASSP*, volume VI, pp. 65-100, 1994.

473. Z. Zvonar and D. Brady, "Linear multipath-decorrelating receivers for CDMA frequency-selective fading channels," *IEEE Trans. on Communications*, pp. 650-663, Jan. 1996.