FINAL REPORT

STATISTICAL INFERENCE
FROM SAMPLED DATA

ONR Contract N00014-84-K-0042

(April 1, 1984 - December 31, 1989)

Elias Masry

Department of Electrical & Computer Engineering

University of California at San Diego

La Jolla, California 92093

February 28, 1990

TABLE OF CONTENTS

I. RESEARCH ACCOMPLISHMENTS

Research for ONR Contract "Statistical Inference From Sampled Data" has been concerned in the past 6 years (1984-1989) with studies in the general area of statistical signal processing. Contributions were made in a wide range of topics many of which were motivated by practical problems in communication systems and digital signal processing. The following is a list of the main topics on which the research was focused:

A. Frequency Wavenumber Spectral Estimation Using Random Acoustic Arrays.

B. Spread Spectrum Communication Systems.

C. Probability Density and Regression Estimation for Dependent Data.

D. Parametric and Nonparametric Spectral Estimation.

E. Adaptive Linear Estimation Algorithms for Dependent Processes

F. Inference for Continuous-Time Processes from Sampled Data.

The research under this contract resulted in the publication of 19 papers in mathematical, statistical, and engineering journals. In addition, several papers were presented and subsequently published in proceedings of conferences. Copies of these works were routinely sent to the Office of Naval Research, Mathematical Sciences Division. The following is a list of the journal publications under this contract.

1. E. Masry, "The Estimation of the Frequency-Wavenumber Spectrum Using Random Acoustic Arrays - Part I: Performance of Beampower Pattern Estimators," *J. Acoustical Society of America,* Vol. 76 (July 1984), pp. 139-149.

2. E. Masry, "Closed-Form Analytical Results for the Rejection of Narrow-Band Interference in PN Spread-Spectrum Systems - Part I: Linear Prediction Filters," *IEEE Trans. Communications,* COM-32 (August 1984), pp. 888-896.

3. E. Masry, "The Estimation of the Frequency-Wavenumber Spectrum Using Random Acoustic Arrays - Part II: A Class of Consistent Estimators," *J. Acoustical Society of America,* Vol. 76 (October 1984), pp. 1123-1131.

4. E. Masry, "A Class of Generalized Szasz Operators and their Convergence Properties," *Mathematica,* Vol. 13 (1984), pp. 45-56.

5. E. Masry, "Closed-Form Analytical Results for the Rejection of Narrow-Band Interference in PN Spread-Spectrum Systems - Part II: Linear Interpolation Filters," *IEEE Trans. Communications,* COM-33 (January 1985), pp. 10-19.

6. D.F. Gingras and E. Masry, "Spectral Density Estimation from Nonlinearly Observed Data," *J. Time Series Analysis,* Vol. 6 (1985), pp. 63-80.

7. E. Masry and L. Milstein, "Performance of DS Spread-Spectrum Receiver Employing Interference-Suppression Filters under a Worst-Case Jamming Condition," *IEEE Trans. Communications,* COM-34(January 1986), pp. 13-21.

8. E. Masry, "Recursive Probability Density Estimation for Weakly Dependent Stationary Processes," *IEEE Trans. Information Theory,* IT-32(March 1986), pp.

9. E. Masry and R. Rao, "On the Choice of the Timeout Distribution in Decentralized Random Access Systems," *IEEE Trans. Communications,* COM-34(May 1986), pp.501-504.

10. E. Masry and B. Picinbono, "Linear/Nonlinear Forms and the Normal Law: Characterization by High Order Correlations," *Annals Institute Statist Math.,* Vol. 39(1987), Part A, pp. 417-428.

11. E. Masry and L. Gyorfi, "Strong Consistency and Rates for Recursive Probability Density Estimators of Stationary Processes," *J. Multivariate Analysis,* Vol. 22(June 1987), pp. 79-93.

12. E. Masry, "Almost Sure Convergence of Recursive Density Estimators for Stationary Mixing Processes," *Statistics & Probability Letters*, Vol. 5(1987), pp. 249-254.

13. D.F. Gingras and E. Masry, "Autoregressive Spectral Estimation in Additive Noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-36(April 1988), pp. 490-501.

14. S. Cambanis and E. Masry, "The Performance of Discrete-Time Predictors of Continuous-Time Stationary Processes," *IEEE Trans. Information Theory*, IT-34(July 1988), pp. 655-668.

15. A. Krieger and E. Masry, "Convergence Analysis of Adaptive Linear Estimation for Dependent Stationary Processes," *IEEE Trans. Information Theory*, IT-34(July 1988), pp. 642-654.

16. E. Masry, "Random Sampling of Continuous-Parameter Stationary Processes: Statistical Properties of Joint Density Estimators," *J. Multivariate Analysis*, Vol. 26(August 1988), pp. 133-165.

17. E. Masry, "Nonparametric Estimation of Conditional Probability Densities and Expectations of Stationary Processes: Strong Consistency and Rates," *J. Stochastic Processes & Applics.*, Vol. 32(1989), pp. 109-127.

18 A. Krieger and E. Masry, "Constrained Adaptive Filtering Algorithms: Asymptotic Convergence Properties for Dependent Data," *IEEE Trans. Information Theory*, IT-35(November 1989), pp. 1166-1176.

19. E. Masry and S. Cambanis, "Trapezoidal Monte Carlo Integration," *SIAM J. Numerical Analysis*, Vol. 27(February 1990).

II. DESCRIPTIVE SUMMARY OF PROBLEMS AND RESULTS

A descriptive summary of the research problems studied under this contract and the nature of the results obtained is now presented for each of the principal areas listed in Section I.

A. Frequency-Wavenumber Spectral Estimation Using Random Acoustic Arrays

The use of an array of sensors for underwater acoustic measurement is standard. In an array of equally-spaced elements, the use of spacing greater than half-wavelength leads to wavenumber aliasing [1]; aliasing can be reduced by using aperiodic arrays. There has been increasing interest in random arrays [2][3] which are aperiodic arrays whose elements' positions are selected at random. The interest is particularly motivated by the idea that a field of randomly deployed, freely drifting, sonobuoys ca be used to form a high gain array for underwater acoustic measurements. An array is totally random if the sensors' locations are realizations of independent identically distributed random vectors [2]. Thorn et al [2] derived fundamental properties of random arrays such as the mean and variance of the pattern function and its power.

Among the generic signal processing tasks of an array system is the estimation of the frequency-wavenumber spectrum of the ambient noise field. When the array system is random, the estimation of the frequency-wavenumber spectrum from data collected at its output is quite complex. Hinich [3] proposed a conventional (periodogram) approach and showed that the estimator is asymptotically unbiased provided that the density of the number of sensors per unit area tends to infinity; thus the sensors must cover the plane densely. The practical significance of this approach is clearly doubtful.

In [4][5] we provided a quadratic-mean convergence analysis of appropriate estimates of the frequency-wavenumber spectrum of the ambient noise field, in n-dimensional Euclidean space, from data collected by a totally random n-dimensional array. In contrast to [3] the density of the number of sensors per unit volume was allowed to be arbitrary. For a class of frequency-wavenumber spectral estimates, we provided closed-form expression for the asymptotic bias and covariance of these estimates. The estimates were shown to be consistent in quadratic-mean as the observation interval length T and the number of sensors M tend to infinity (the array can be arbitrarily sparse). The rate of quadratic mean convergence was determined. The dependence of the mean-square estimation error on the parameters of the system, such as the observation length T, the number of sensors M, the probability density of the sensors' positions, and the number of sensors per unit volume was investigated.

B. Spread Spectrum Communication Systems

Spread Spectrum communication systems offer immunity against narrow-band interference. This can be further improved by employing linear mean-square estimation techniques[6]-[8] for the estimation and subsequent suppression of the interference. These techniques exploit the distinct spectral characteristics of the interference, which is generally narrow-band, and that of the pseudo-noise (PN) or direct sequence (DS) signals which are broad-band. When the interference is modeled as a multiple tone, the improvement in the performance of the system due to the employment of linear interference-suppression filters has been thoroughly studied in [6]-[8].

In [9][10] we provided an analytical study of the performance of PN spread spectrum communication systems, using interference-rejection filters, when the interference is modeled as a stationary process with *arbitrary* spectral density. In [9] linear predictive filters were employed and in [10] linear interpolative filters were used. Specialized results were obtained for interferences which are either bandlimited or have rational spectral densities. In each case, the additional improvement in the performance of the receiver was investigated particularly its dependence on such factors as the interference power and the number of taps of the filter.

In [11] we investigated the spectral characteristics of the worst-case jammer, assumed to be a narrow-band Gaussian process, under a joint power and bandwidth constraints. An expression for the average probability of error is derived for a DS spread-spectrum communication system. The effectiveness of the suppression filter under these conditions was studied and illustrated for various combinations of system parameters.

C. Probability Density and Regression Estimation for Dependent Observations

Probability density and regression estimation play an important role in communication theory, pattern recognition and classification [12]-[18].

Nonparametric density estimation for stationary random processes (dependent observations) has been receiving increasing attention in recent years. Rosenblatt [19] considered nonrecursive kernel-type density estimation for Markov processes; Takahata [20], Robinson [21][22], and Castellana and Leadbetter [23] dealt with similar estimators for mixing processes. These works are primarily concerned with establishing quadratic-mean convergence and asymptotic normality of nonrecursive density estimators.

Under the current contract we initiated an investigation into the the statistical properties of recursive kernel density estimators for stationary processes. In [24] we established the quadratic-mean convergence, along with the asymptotic expressions for the bias and covariance of such estimators, for

mixing processes. Also. asymptotic normality was established. In [25] [26] we derived sharp almost sure convergence rates for the marginal density estimators of vector-valued stationary mixing processes. In [27] we investigated the almost sure convergence of recursive kernel-type estimators for the joint probability densities, conditional densities, and conditional expectations of scalar-valued stationary mixing processes. Specifically, For each integer $m \geq 1$ and integers $0 = i_1 < i_2 < \cdots < i_m$ let $f(\underline{x}; \underline{i}_m) = f(x_1, \ldots, x_m; i_1, \ldots, i_m)$ be the joint probability density function of the random variables $X_{i_1}, \ldots, X_{i_m}$, which is assumed to exist. For any integer p, $1 \leq p < m$, put $\underline{i}'_p = (i_1, \ldots, i_p)$ and $\underline{i}''_{m-p} = (i_{p+1}, \ldots, i_m)$. The conditional probability density function of $\underline{X}''_j = (X_{j+i_{p+1}}, \ldots, X_{j+i_m})$ given $\underline{X}'_j = (X_{j+i_1}, \ldots, X_{j+i_p})$ is denoted by

$$f(\underline{x}_2 | \underline{x}_1) \triangleq f(\underline{x}_2; \underline{i}''_{m-p} | \underline{x}_1; \underline{i}'_p) = \frac{f(\underline{x}; \underline{i}_m)}{f(\underline{x}_1; \underline{i}'_p)}$$

where $\underline{x}_1 \in R^p$, $\underline{x}_2 \in R^{m-p}$ and $\underline{x} = (\underline{x}_1, \underline{x}_2) \in R^m$. Let q be a Borel measurable function on R^{m-p} such that $E|q(\underline{X}''_j)| < \infty$. The conditional expectation of $q(\underline{X}''_j)$ given $\underline{X}'_j = \underline{u}$ is denoted by

$$Q(\underline{u}) = E[q(\underline{X}''_j) | \underline{X}'_j = \underline{u}].$$

A special case of interest is when $m = p+1$ and $q(y) = y$ for which $Q(\underline{u}) = E[X_{j+i_{p+1}} | \underline{X}'_j = \underline{u}]$ is the usual regression function. In [27] recursive kernel estimators $\hat{f}_n(\underline{x}; \underline{i}_m)$, $\hat{f}_n(\underline{x}_2 | \underline{x}_1)$, and $\hat{Q}_n(\underline{u})$ are defined and their strong consistency and almost sure convergence rates are established.

D. Parametric and Nonparametric Spectral Estimation

In the area of nonparametric spectral estimation, we considered in [28] the problem of estimating the spectral density function $\phi(\lambda)$ of a stationary Gaussian process $\{X_k\}_{k=-\infty}^{\infty}$ from observations that have been distorted by a known zero-memory nonlinearity. Such distortions may be due to the presence of a nonlinear device in the data communication system; an n-bit quantizer is an example. Certain aspects of this problem were considered earlier by Rodemich [29], McNeil [30] and Brillinger [31]. In [28] we considered the case where the nonlinearity f is real, bounded, odd, and nondecreasing function on the real line. The output process is $\{Y_k = f(X_k)\}_{k=-\infty}^{\infty}$ and the problem is the estimation of the spectral density $\phi(\lambda)$ of the input process X from the knowledge of f and a finite set of observations $\{Y_k\}_{k=1}^N$ of the output process Y. In [28] we introduced a class of spectral estimates and its quadratic-mean consistency is established along with the asymptotic expressions for the bias and covariance. In addition, the asymptotic normality of the spectral estimates is derived. Our results extend those of Brillinger [31], where a signum nonlinearity $f(x) = \text{sign}[x]$ was considered, and refine and correct those of [29] [30].

In the area of parametric spectral estimation we considered in [32] the problem of estimating the spectral density of an autoregressive (AR) stationary process from a finite set of *noisy* observations. A modified spectral estimator based on the high-order Yule-Walker equations was considered. Joint

asymptotic normality of the spectral estimator is established along with a precise expression for its covariance matrix. The work extends the results of Akaike [33] for the noise-free case, and complements the work of Pagano [34] where nonlinear regression method is used. The advantage of using the (linear) high-order Yule-Walker equations is the simplicity of implementation.

E. Adaptive Linear Estimation Algoriths for Dependent Data

In recent years adaptive linear estimation based on the principle of steepest descent and its variations have been applied in a wide range of problems such as filtering [35], pattern recognition [36], line enhancement [37], antenna processing [38], and interference suppression in spread-spectrum communication systems [39]. Probably the most widely-used algorithm is the so called "Widrow LMS" algorithm, which uses the gradient method to find the direction of the steepest descent and at each iteration replaces the true gradient vector by its instantaneous estimate. The main advantage of the LMS algorithm is the simplicity and the relative low complexity of its implementation. The algorithm is defined as follows. Suppose $\{a_j\}$ and $\{x_j\}$ are two jointly stationary second-order processes representing the desired signal and observation process, respectively. Suppose we wish to estimate a_k from n observations x_j, $j=k-n+1,...,k$ in a linear fashion

$$\hat{a}_k = \sum_{i=1}^{n} h_i\, x_{k-i+1} \equiv \underline{h}^T \underline{x}_k$$

where

$$\underline{h} = (h_1, ..., h_n)^T$$

and

$$\underline{x}_k = (x_{k-n+1}, ..., x_k)^T . \tag{3}$$

The optimal linear mean-square filter $\underline{h}*$ which minimizes the error $E[(a_k - \hat{a}_k)^2]$ is given by

$$\underline{h}* = \underline{R}^{-1} \underline{b}$$

where $\underline{R}$ is the covariance matrix of the data vector $\underline{x}_k$,

$$\underline{R} = E[\underline{x}_k\, \underline{x}_k^T]$$

and $\underline{b}$ is given by

$$\underline{b} = E[a_k\, \underline{x}_k] .$$

When the second-order moments of $\{a_j\}$ and $\{x_j\}$ are unknown, $\underline{R}$ and $\underline{b}$ are not available to determine the optimal filter $\underline{h}*$. The LMS algorithm adapts the filter's coefficients to the incoming data: at the $(k+1)^{st}$ "instant", we have an estimate $\hat{\underline{h}}_{k+1}$ of $\underline{h}*$ given recursively by

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k + \mu\, \underline{x}_k\, (a_k - \underline{x}_k^T\, \hat{\underline{h}}_k) , \quad k=1,2,... \tag{4}$$

$$\hat{\underline{h}}_1 = \underline{0}$$

so that the value of $\underline{h}$ is updated with the kth incoming block of data $\underline{x}_k$. The constant μ is the adaptation step size of the algorithm. The signal estimator at the kth instant is then

$$\hat{a}_k = \hat{\underline{h}}_k \, \underline{x}_k \ .$$

The LMS algorithm (4) is simple to implement and requires little storage. Furthermore it is applicable not only to stationary inputs but also to slowly time-varying processes by tracking system parameters in a nearly optimum way. In terms of the convergence properties of the LMS algorithm, one is usually concerned with establishing asymptotic expressions for the filter's coefficients estimation error

$$E \parallel \hat{\underline{h}}_k - \underline{h}* \parallel^2$$

and for the signal estimation error

$$E[(\hat{a}_k - a_k)^2]$$

when the number of adaptation steps k becomes large.

Under the assumption of independent input vectors $\{\underline{x}_j\}$, the performance of LMS algorithm has been extensively studied in the literature (see [35], [40] and the references therein). The assumption of independent input vectors makes the analysis more tractable but in many applications it is not a realistic model. Furthermore, for scalar-valued processes $\{x_j\}$, the assumption of independent input vectors $\{\underline{x}_j\}$ - see (3) - cannot hold since successive vectors share all but one of their components. Analysis of the LMS algorithms when the input vectors $\{\underline{x}_j\}$ are correlated is considerably more complex due to the highly nonlinear nature of the algorithm. Some results are available in the literature under fairly restrictive conditions: an assumption of M-dependence on the observation process $\{x_j\}$ is made in [41]; an assumption of bounded conditional moments is made in [42] which excludes the important Gaussian data case.

During the the contract period, we initiated a comprehensive study of the convergence properties of the LMS algorithm and its variants under general dependence structure on the signal and observation processes $\{a_j\}$ and $\{x_j\}$. In particular we investigated the following aspects.

i) In [43] we provided a convergence analysis of a modified LMS algorithm under the fairly weak assumption that the input processes satisfy a mixing condition (strong mixing or ρ mixing). In this analysis it was assumed that the optimal Wiener-Hopf coefficients are known to lie in a bounded subset of R^n; however, it was shown that for broad classes of random processes such sets can be determined without actually solving the Wiener-Hopf equations. The results of this work provide sharper bounds on the mean-square error of the coefficients' and signal estimates than those obtained in [44].

ii) In [45] we provided a convergence analysis of a constrained LMS algorithm for dependent processes satisfying various mixing conditions. The need for such algorithms arises from the fact that in

many applications the optimal finite impulse response filter of order n is restricted such that its coefficients lie in a subset of the n-dimensional vector space R^n. For example, in antenna beam-forming the need to control the magnitude of the sidelobes imposes such a constraint [46][47][48]. Similar situations arise in parametric spectral estimation such as the Pisarenko harmonic retrieval [49]. In these references, the performance of the constrained LMS algorithm was evaluated either by simulation or under the assumption of independent input vectors. In [45] we considered two distinct constraint sets: A bounded hypercube in R^n, which leads to a magnitude-constrained algorithm, and a bounded hypersphere, which leads to a quadratically-constrained algorithm. The purpose of the algorithm is to provide a linear estimate $\hat{a}_k$ of a_k from the observation vector $\underline{x}_k = (x_{k,1}, \ldots, x_{k,n})^T$ where $x_{k,i} = x_{k-i+1}$. This estimate is of the form

$$\hat{a}_k = \underline{w}^T \underline{x}_k$$

where $\underline{w}$ is the vector of filter coefficients, with a fixed dimensionality n, constrained to a bounded set S in R^n. When the second-order statistics of the processes $\{a_j\}$ and $\{x_j\}$ are known, the optimal mean-square filter coefficients $\underline{w}_*$ is obtained by solving the following problem:

$$\min_{\underline{w} \in S} E[a_k - \hat{a}_k]^2$$

where, for a magnitude-constrained algorithm, the set S is defined by

$$S = \{\underline{w}: |w_i| \le B, \ i=1,\ldots,n\} \quad,$$

w_i is the ith component of $\underline{w}$, B is a finite constant; for a quadratic-constrained algorithm, the set S is given by

$$S = \{\underline{w} : \|\underline{w}\| \le B\} \quad,$$

where $\| \ \|$ is the Euclidean norm in R^n. We show in [45] that the adaptive algorithm, when the second order statistics of $\{a_j\}_{j=-\infty}^{\infty}$ and $\{x_j\}_{j=-\infty}^{\infty}$ are unknown, takes the form

$$\hat{\underline{w}}_{k+1} = P[(I - \underline{x}_k \underline{x}_k^T)\hat{\underline{w}}_k + \mu a_k \underline{x}_k], \quad k=0,1,\ldots ,$$

$$\hat{a}_k = \underline{x}_k^T \hat{\underline{w}}_k, \quad k=0,1,\ldots .$$

where the operator P is a projection from R^n to S. We show in [45] that for input processes obeying fairly weak mixing conditions, we have

$$\lim_{k \to \infty} \sup E\|\hat{\underline{w}}_k - \underline{w}_*\|^2 \le \mu \, C_1 \tag{5}$$

and

$$E[a_k - \hat{a}_k]^2 = \varepsilon_* + \varepsilon_k$$

with

$$\lim_{k \to \infty} \sup \varepsilon_k \leq \mu \, C_2 . \tag{6}$$

Here ε_* is the minimum MSE corresponding to the optimal constrained Wiener-Hopf filter $\underline{w}_*$, ε_k is the excess MSE of the algorithm, and C_i, $i = 1,2$ are positive constants. Note that the bounds in (5)(6) are proportional to the adaptation step size μ and thus have the same form as the bounds established in the literature for the unconstrained LMS algorithm with dependent data (e.g. [41] [42]).

F. Inference for Continuous-Time Processes from Sampled Data

Let $X = \{X(t), -\infty < t < \infty\}$ be a continuous-time stationary process, $\{t_k\}$ be the sampling instants, and $\{X(t_k)\}$ be the discrete-time observation processes. We are interested in estimating the statistical structure of the process $\{X(t), -\infty < t < \infty\}$ from a finite set of discrete-time observations $\{X(t_k)\}_{k=1}^{n}$. Particular functions of interest are the family of finite-dimensional distributions and densities of the process X, the correlation function $C(t)$ and spectral density $\phi(\lambda)$ of the process X. Clearly if the the sampling instants $\{t_k\}$ are equally-spaced, consistent estimates of the joint densities of the process X from the observations $\{X(t_k)\}$ is not feasible; similarly, if the process X is not bandlimited, consistent estimates of $C(t)$ and $\phi(\lambda)$ from equally- spaced observations is not possible due to aliasing.

During the contract period we considered in [50] the problem of consistently estimating the family of finite-dimensional densities of the process X from discrete-time observations. Let $f(\underline{x};\underline{\tau}) = f(x_1,...,x_{m+1};\tau_1,\tau_2,...,\tau_m)$ be the joint probability density of the random variables $X(0), X(\tau_1), X(\tau_2),...,X(\tau_m)$, $0 < \tau_1 < \tau_2 < ... < \tau_m$, which is assumed to exist. Let $K(\underline{x})$ be a bounded nonnegative function on R^{m+1} satisfying

$$\int_{R^{m+1}} K(\underline{x}) d\underline{x} = 1, \; \lim_{\|\underline{x}\| \to \infty} \|\underline{x}\|^{m+1} K(\underline{x}) = 0,$$

and let $K_n(\underline{x}) = (1/b_n^{m+1}) K(\underline{x}/b_n)$. Similarly, let $W(\underline{u})$ be a bounded nonnegative function on R^m satisfying

$$\int_{R^m} W(\underline{u}) d\underline{u} = 1, \; \lim_{\|\underline{u}\| \to \infty} \|\underline{u}\|^m W(\underline{u}) = 0,$$

and put $W_n(\underline{u}) = (1/b_n^m) W(\underline{u}/b_n)$. Given the observations $\{X(t_j), t_j\}_{j=1}^{n+m}$ we estimate $f(\underline{x};\underline{\tau})$ by

$$\hat{f}_n(\underline{x};\underline{\tau}) = 1/(na(\underline{\tau})) \sum_{j=1}^{n} W_n(\underline{\tau} - \underline{D}_j) K_n(\underline{x} - \underline{X}_j)$$

where

$$a(\underline{\tau}) = p(\tau_1) \prod_{i=1}^{m-1} p(\tau_{i+1} - \tau_i),$$

$\underline{X}_j = (X(t_j), X(t_{j+1}),...,X(t_{j+m}))$ and $\underline{D}_j = (t_{j+1} - t_j, t_{j+2} - t_j, ..., t_{j+m} - t_j)$. Here the sampling instants $\{t_i\}_{i=1}^{\infty}$ are assumed to be a renewal process on $[0,\infty)$; $p(t)$ is the probability density function of the interarrival

times which is assumed to be positive on $[0,\infty)$ with finite mean $\frac{1}{\beta}$. We allow the average sampling rate β to take any positive value and consequently, no minimum sampling rate β is required. We established in [50] the quadratic-mean convergence of $\hat{f}_n(\underline{x};\tau)$ as the sample size n tends to infinity along with precise asymptotic expressions for the bias and variance/covariance. In addition, we established the asymptotic normality of the estimators $\hat{f}_n(\underline{x};\tau)$. These results in [50] are shown to hold for continuous-time processes X which are either strong mixing or ρ mixing.

During the contract period we were also concerned with sampling design problems of continuous-time processes [51] [52]: Let $\{X(t), -\infty < t < \infty\}$ be a second-order continuous-time process with zero mean and correlation function $R(t,s)$. Many problems in detection and estimation involve the evaluation of the integral

$$I = \int_0^T f(t)X(t)dt$$

where f is a given function. In practice, the evaluation of I is carried out in a digital fashion; the observation process X is sampled at n instants $T_n = \{t_{n,1}, \cdots, t_{n,n}\}$ satisfying

$$0 \le t_{n,1} < t_{n,2} < \cdots < t_{n,n} \le T,$$

and based on the samples $\{X(t_{n,i})\}_{i=1}^n$, an approximation of the integral I of the form

$$I_n = \sum_{i=1}^n c_{n,i} X(t_{n,i})$$

is used. One is then concerned with the performance of such approximation measured in terms of the mean-square error $e_n^2 = E[I - I_n]^2$. Clearly this depends on the choice of the coefficients $\{c_{n,i}^*\}_{i=1}^n$ which may be optimal or suboptimal, and on the choice of the sampling instants T_n. The sampling design problem is to find the best location of the sampling points for fixed sample size n or, asymptotically, as the sample size tends to infinity. For a given choice of the coefficients and the sampling points, optimal or not, one also desires to find the rate of convergence to zero of the error e_n^2 as $n \to \infty$ which can be used to determine the required sample size n for an acceptable level of accuracy of the discrete-time approximation. We remark that in the context of regression problems, sampling designs were considered earlier by Sacks and Ylvisaker [53]-[55]. Sampling designs for suboptimal integral approximations were considered by Schoenfelder [56] for the case of processes with no quadratic-mean derivatives.

In [51] the asymptotic performance of linear predictors of continuous-time stationary processes from observations at n sampling instants on a fixed interval was studied. Both optimal and simpler choices of predictor coefficients were considered using uniform and nonuniform sampling schemes. The focus was on processes with rational spectral density. When the process has no quadratic-mean derivatives, it was shown that both the optimal and suboptimal predictors have the same rate of quadratic-mean convergence n^{-2} and the same asymptotic constant which depends on the sampling

design. The performance of the asymptotically optimal design is compared analytically and computationally to that of uniform sampling and the improvement factor is determined. When the process has exactly one quadratic-mean derivative, and a suboptimal coefficients are used, it is shown that uniform sampling has a rate of quadratic-mean convergence of only n^{-1} while nonuniform sampling can be designed to achieve a rate of n^{-4}.

In [52] we considered the problem of Monte Carlo approximation of the integral of weighted random processes. The simplest Monte Carlo method for approximating the integral

$$I(fX) = \int_0^1 f(t) X(t)\, dt$$

of a function f over a finite interval, uses n independent samples $U_1, ..., U_n$ from a uniform distribution over the unit interval, and forms the average estimate

$$J_n(fX) = \frac{1}{n} \sum_{i=1}^n f(U_i) X(U_i).$$

The mean-square error of $J_n(f)$ is

$$E[I(fX) - J_n(fX)]^2 = \frac{1}{n} \{ \int_0^1 f^2(x) R(x,x) dx - \int_0^1\int_0^1 f(x) R(x,y) f(y) dx dy \}$$

and thus tends to 0 at the rate of n^{-1}. No improvement in the rate of convergence is generally expected from any additional smoothness assumptions on f or on X.

Motivated by the work of Yakowitz, Krimmel and Szidarovszky [57] for the integral of nonrandom functions, we considered in [52] the convergence properties of a trapezoidal Monte Carlo approximation based on the ordered sample $t_{n,0} \triangleq 0 < t_{n,1} < t_{n,2} < ... < t_{n,n} < 1 \triangleq t_{n,n+1}$ obtained from the independent, uniformly distributed samples $U_1, U_2, ..., U_n$. The approximation of $I(fX)$ is now given by

$$I_n(fX) = \frac{1}{2} \sum_{i=0}^n [f(t_{n,i}) X(t_{n,i}) + f(t_{n,i+1}) X(t_{n,i+1})] (t_{n,i+1} - t_{n,i}).$$

We show in [52] that when f has two continuous derivatives on [0, 1] and the processes X has one quadratic-mean derivative, the rate of mean-square convergence of the approximation $I_n(fX)$ is precisely n^{-4}, i.e.,

$$\lim_{n \to \infty} n^4 E[I(fX) - I_n(fX)]^2 = C(f,R)$$

where the asymptotic constant $C(f,R)$, which depends on f and on correlation function R, is determined explicitly.

References

[1]   B.D. Steinberg, *Principles of Aperture and Array System Design.* New York: Wiley, 1976.

[2]   J. V. Thorn, N. O. Booth, and J. C. Lockwood, "Random and partially random acoustic arrays," *J. Acoust. Soc. Amer.*, Vol. 67(1980), pp. 1277-1285.

[3]   M. J. Hinnich, "Estimating signal and noise using a random array," *J. Acoust. Soc. Amer.*, Vol. 71(1982), pp. 97-99.

[4]   E. Masry, "The estimation of the frequency-wavenumber spectrum using random acoustic arrays - Part I: Performance of beampower pattern estimators," *J. Acoustical Society of America,* Vol. 76 (July 1984), pp. 139-149.

[5]   E. Masry, "The estimation of the frequency-wavenumber spectrum using random acoustic arrays - Part II: A class of consistent estimators," *J. Acoustical Society of America,* Vol. 76 (October 1984), pp. 1123-1131.

[6]   F. M. Hsu and A. A. Giordino, "Digital whitening techniques for improving spread spectrum communication performance in the presence of narrowband jamming and interference," *IEEE Trans. Commun.*, Vol. COM-26(1978), pp. 209-216.

[7]   L. M. Li and L. B. Milstein, "Rejection of narrow-band interference in PN spread-spectrum systems using transversal filters," *IEEE Trans. Commun.*, Vol. COM-30(1982), pp. 925-928.

[8]   J. W. Ketchum and J. G. Proakis, "Adaptive algorithm for estimating and suppressing narrow-band interference in PN spread-spectrum systems," *IEEE Trans. Commun.*, Vol. COM-30(1982), pp. 913-927.

[9]   E. Masry, "Closed-form analytical results for the rejection of narrow-band interference in PN spread-spectrum systems - Part I: Linear prediction filters," *IEEE Trans. Communications*, COM-32 (August 1984), pp. 888-896.

[10] E. Masry, "Closed-form analytical results for the rejection of narrow-band interference in PN spread-spectrum systems - Part II: Linear interpolation filters," *IEEE Trans. Communications*, COM-33 (January 1985), pp. 10-19.

[11] E. Masry and L. Milstein, "Performance of DS spread-spectrum receiver employing interference-suppression filters under a worst-case jamming condition," *IEEE Trans. Communications*, COM-34(January 1986), pp. 13-21.

[12] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic Press, 1972.

[13] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Information Theory*, Vol. IT-13(1967), 21-27.

[14] C.T. Wolverton and T.J. Wagner, "Asymptotically optimal discriminant functions for pattern classification," *IEEE Trans. Information Theory*, Vol. IT-15(1969), 258-265.

[15] K. Fukunaga and L.D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Information Theory*, Vol. IT-21(1975), 32-40.

[16] M. Goldstein, "Comparison of some density estimates classification procedures," *J. Amer. Statist. Assoc.*, Vol. 70(1975), 666-669.

[17] A. Krzyzak and M. Pawlak, "Distribution-free consistency of nonparametric kernel regression estimates and classification," *IEEE Trans. Inform. Theory*, vol. IT-30(1984), pp. 78-81.

[18] A. Krzyzak, "The rates of convergence of kernel regression estimates and classification rules," *IEEE Trans. Inforn Theory*, vol. IT-32(1986), pp. 668-679.

[19] M. Rosenblatt, "Density estimates and Markov sequences," in *Nonparametric Techniques in Statistical Inference*, M. Puri, Ed., 199-210, 1970.

[20] H. Takahata, "Limiting behavior of density estimates for stationary asymptotically uncorrelated processes," *Bull. Tokyo Gakugei Univ*, Ser. 4, vol. 29(1977), pp.1-9.

[21] P.M. Robinson, "Nonparametric estimators for time series," *J. Time Series Analysis*, vol. 4(1983), pp. 185-207.

[22] P.M. Robinson, "On the consistency and finite sample properties of nonparametric kernel time series regression, autoregression and density estimators," *Annals Inst. Statist. Math.*, vol. 38(1986), pp. 539-549.

[23] J.V. Castellana and M.R. Leadbetter, "On smoothed probability density estimation for stationary processes," *Stoch. Processes and their Applics.*, vol. 21(1986), pp. 179-193.

[24] E. Masry, "Recursive probability density estimation for weakly dependent stationary processes," *IEEE Trans. Information Theory*, IT-32(March 1986), pp.

[25] E. Masry and L. Gyorfi, "Strong consistency and rates for recursive probability density estimators of stationary processes," *J. Multivariate Analysis*, Vol. 22(June 1987), pp. 79-93.

[26] E. Masry, "Almost sure convergence of recursive density estimators for stationary mixing processes," *Statistics & Probability Letters*, Vol. 5(1987), pp. 249-254.

[27] E. Masry, "Nonparametric estimation of conditional probability densities and expectations of stationary processes: Strong consistency and rates," *J. Stochastic Processes & Applics.*, Vol. 32(1989), pp. 109-127.

[28] D.F. Gingras and E. Masry, "Spectral density estimation from nonlinearly observed data," *J. Time Series Analysis*, Vol. 6 (1985), pp. 63-80.

[29] E. R. Rodemich, "Spectral estimates using nonlinear functions," *Ann. Math.*, Vol. 37(1966), pp. 1237-1256.

[30] D. R. McNeil, "Estimating the covariance and spectral density functions from clipped stationary time series," *J. R. Statist. Soc. Ser. B*, Vol. 29(1967), pp. 180-195.

[31] D. R. Brillinger, "Estimation of the cross-spectrum of a stationary bivariate Gaussian process from its zeros," *J. R. Statist. Soc. Ser. B*, Vol. 30(1968), pp. 145-159.

[32] D.F. Gingras and E. Masry, "Autoregressive spectral estimation in additive noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, ASSP-36(April 1988), pp. 490-501.

[33] E. H. Akaike, "Power spectrum estimation through autoregressive model fitting," *Ann. Inst. Statist. Math.*, Vol. 21(1969), pp. 407-419.

[34] M. Pagano, "Estimation of models of autoregressive signal plus white noise," *Ann. Statist.*, Vol. 2(1974), pp. 99-108.

[35] B. Widrow, "Adaptive filters," in *Aspects of Network and Systems*, R.E. Kalman and N. Declaris, Eds. New York: Holt, Rinehart and Winston, 1971.

[36] J.S. Koford and G.F. Groner, "The use of an adaptive threshold element to design a linear optimal pattern classifier," *IEEE Trans. Information Theory*, Vol. IT-12(1966), 42-50.

[37] C.M. Anderson, E.H. Satorius and J.R. Zeidler, "Adaptive enhancement of finite bandwidth signals in white Gaussian noise," *IEEE Trans. Acoust., Speech. Signal Processing*, Vol. ASSP-31(1983), 17-28.

[38] B. Widrow et al., "Adaptive antenna systems," *Proc. IEEE*, Vol. 55(1967), 2143-2159.

[39] J.W. Ketchum and J.G. Proakis, "Adaptive algorithms for estimating and suppressing narrow-band interference in PN spread-spectrum," *IEEE Trans. Com.*, Vol. COM-30(1982), 913-924.

[40] A. Feuer and E. Weinstein, "Convergence analysis of LMS filter with uncorrelated Gaussian data," *IEEE Trans. Acoust. Speech, Signal Processing*, Vol. ASSP-33(1985), 220-230.

[41] O. Macchi and E. Eweda, "Second-order convergence analysis of stochastic adaptive linear filtering," *IEEE Trans. Automat. Contr.*, Vol. AC-28(1983), 76-85.

[42] T.P. Daniell, "Adaptive estimation with mutually correlated training sequences," *IEEE Trans. Syst., Sci., Cybern.*, Vol. SSC-6(1970), 12-19.

[43] A. Krieger and E. Masry "Convergence analysis of adaptive linear estimation for dependent stationary processes," *IEEE Trans. Information Theory*, Vol. 34(1988), pp. 642- 654.

[44] D.C. Farden, "Tracking properties of adaptive signal processing algorithms," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. ASSP-29 (1981), pp. 439-446.

[45] A. Krieger and E. Masry, "Constrained adaptive filtering algorithms: Asymptotic convergence properties for dependent data," *IEEE Trans. Information Theory*, Vol. IT-35(November 1989).

[46] O.L. Frost, III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, Vol. 60(1972), 926-935.

[47] W.F. Gabriel, "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Trans. on Antenna and Prop.*, Vol. 34(1986), pp. 291-300.

[48] N.L. Owsley, "A recent trend in adaptive spatial processing for sensor arrays: Constrained adaptation," in *Array Processing Applications to Radar*, S. Haykin ed., Academic Press, 1980.

[49] D.R. Fuhrman, "An approximation subspace method for eigenfilter computation," *IEEE Trans. Acoust., Speech, and Signal Processing*, Vol. 34(1986), pp. 205-207.

[50] E. Masry, "Random sampling of continuous-parameter stationary processes: Statistical properties of joint density estimators," *J. Multivariate Analysis*, Vol. 26(August 1988), pp. 133-165.

[51] S. Cambanis and E. Masry, "The performance of discrete-time predictors of continuous-time stationary processes," *IEEE Trans. Information Theory*, IT-34(July 1988), pp. 655-668.

[52] E. Masry and S. Cambanis, "Trapezoidal Monte Carlo integration," *SIAM J. Numerical Analysis*, Vol. 27(February 1990).

[53] J. Sacks and D. Ylvisaker, "Designs for regression problems with correlated errors," *Ann. Math. Statist.*, Vol. 37(1966), 66-89.

[54] _____, "Designs for regression problems with correlated errors III," *Ann. Math. Statist.*, Vol. 41(1970), 2057-2074.

[55] _____, "Statistical designs and integral approximation," in *Proc. Twelfth Biennial Seminar of the Canadian Math. Congress*, 115-136, Canadian Math. Congress, Montreal, 1970.

[56] C. Schoenfelder, "Random designs for estimating integrals of stochastic processes," *Institute of Statistics Mimeo Series* 1201, University of North Carolina at Chapel Hill, NC, 1978.

[57] S. Yakowitz, J.E. Krimmel and F. Szidarovszky, "Weighted Monte Carlo integration", *SIAM J. Numer. Anal.*, 15 (1978), pp. 1289-1300.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Statistical Inference From Sampled Data		5. TYPE OF REPORT & PERIOD COVERED Final Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) E. Masry		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-0042
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Electrical & Computer Engineering University of California at San Diego La Jolla, CA 92093		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-661-002
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical Sciences Division Office of Naval Research Arlington, VA 22217		12. REPORT DATE February 28, 1990
		13. NUMBER OF PAGES 19
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEYWORDS (Continue on reverse side if necessary and identify by block number)

Frequency-wavenumber spectral estimation, spread-spectrum communication systems, parametric spectral estimation in additive noise, adaptive linear estimation algorithms, inference for continuous-time processes from sampled data.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This final report of research Contract N00014-84-K-0042 provides a summary of research accomplishments during the past 6 years. A descriptive outline of the main subjects studied and the results obtained is presented.

DD FORM 1473
1 JAN 73
EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified, February 28, 1990
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)