Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Final Report on
ESTIMATION AND STATISTICAL ANALYSIS OF SPATIALLY
DISTRIBUTED RANDOM PROCESSES

NSF Grant ECS-8312921

Covering the Period
February 1, 1984 to July 31, 1987

Prepared by
Professor Alan S. Willsky
Professor Bernard C. Levy
Professor George C. Verghese

Submitted to: Dr. Michael Polis, Program Director
Systems Theory and Operations Research
Division of Electrical, Computer and
Systems Engineering
National Science Foundation
Washington, D.C. 20550
1. Introduction

A brief description of the research carried out by faculty, staff, and students of the M.I.T. Department of Electrical Engineering and Computer Science under NSF Grant ECS-8312921 is presented. The principal investigator for this research was Prof. Alan S. Willsky, and the co-principal investigator was Prof. Bernard C. Levy. The period covered by this grant was February 1, 1984 to July 31, 1987.

The basic scope of this grant was to carry out fundamental research on several interrelated classes of problems involving spatially-distributed random processes. The general philosophy behind our research was to use the structure of the models used to describe spatial processes to obtain insight into the form of solutions and to derive efficient algorithms for their implementation. Significant results have been obtained in several research areas which are described in the remaining sections of this report. A list of personnel involved in this project and a list of publications supported in whole or in part by this grant are also included.

Our work in this area has produced algorithms for decentralized linear estimation and for the updating of smoothed estimates or maps. While most of these results focus on 1-D processes (i.e. processes with a single independent variable), our more recent discrete-time results [3] have some application to 2-D map updating as well as to other applications such as model validation. In addition these results, provide theoretical insight into the structure of smoothing algorithms and the construction of models for smoothing errors. Our most recent work in this area [24]-[25] has focused on the development of parallel processing algorithms for smoothing. In contrast to our earlier work and other efforts in this area, in which a vector of observations is separated into subvectors each defined over the full range of the independent variable and each processed by a separate estimator, we have focused on decompositions of the range of the independent variable over which each local processor operates. Such decompositions are of particular importance in 2-D applications, since (1) algorithm complexity in 2-D is a strong function of data array size; and (2) array processing architectures make such decompositions attractive. There are numerous alternatives for organizing the computations and important theoretical questions related to the way in which neighboring estimators interact to produce globally optimal estimates. Our work has provided many answers in the 1-D case but much remains to be done in several dimensions.

Specifically, it would be of interest to develop map updating techniques which would rely on multidimensional noncausal boundary value models such as those discussed in the next section. The development of parallel processing algorithms in several dimensions is also a very challenging problem which is closely related to our work on boundary value processes discussed below.


This portion of our research was motivated by M. Adams's Ph.D. thesis, in which we extend the concept of complementary processes in order to develop a general approach to constructing a generalized Hamiltonian boundary-value description of the optimal smoother for noncausal processes in several dimensions. The Hamiltonian system then serves as the starting point for devising efficient smoothing algorithms. In the 1-D case this leads to Riccati-equation-based Hamiltonian diagonalization procedures, which are somewhat more complex in the noncausal case because of the boundary conditions. In [6]-[7] we use our machinery to develop efficient

¹Numbers refer to the list of publications supported by NSF.
smoothing algorithms for a particular class of 2-D processes described by nearest neighbor difference equations. In addition this work makes use of descriptor-form difference equations which provide a natural framework for describing noncausal systems. This has led us to focus attention on developing a system theory for boundary-value descriptor systems \[4\]-\[5\], \[8\]-\[15\]. In particular, in addition to providing contributions to existing lines of investigation on descriptor systems (such as the introduction of a standard form that greatly simplifies the generalized Cayley-Hamilton theorem as well as results on reachability and observability), our work investigates the new notions of recursion (namely inward from the boundaries and outward from the center) that Krener first introduced in the nondescriptor continuous-time case. These concepts lead to new methods for the efficient solution of boundary-value equations, and in addition led us to develop completely new notions of stability for such systems and a Lyapunov theory both for stability and for the investigation of the existence and nature of stationary statistics for random processes described by such models (a topic that is more complex than the causal case, again because of the presence of boundary conditions). In \[4\]-\[5\] we have investigated the optimal smoother for these processes and have introduced a new type of generalized Riccati equation that is useful for Hamiltonian diagonalization in far more interesting cases than previously-used descriptor Riccati equations and that points the way to a number of questions and results relating our system-theoretic results to properties, such as stability, of the smoother.

Our work on boundary value stochastic processes can be used as the starting point for a large number of promising new research directions. These include for example the completion of the system theory for two-point boundary descriptor systems described in \[4\]-\[5\], \[8\]-\[15\]. Such a completion would include the development of deterministic and stochastic realization theories, the development of new recursive algorithms based on inward/outward recursions, and extensions of these results to two dimensions. In addition, our work on the estimation of boundary value processes has already had some impact on several other areas. These include the use of our estimation formalism in problems of motion estimation and robot vision (see our discussion in Section 6), as well as for the development of an identification-based approach to inverse signal processing problems, which is discussed in Section 5.

4. Isotropic Random Fields \[16\] - \[23\], \[26\]

Isotropic fields, in which the covariance between the field at two points depends only on the distance between the points, are a natural generalization of stationary random processes or, more precisely, reversible processes. The additional rotational symmetry of these fields, as compared to only the translational symmetry of homogeneous random fields, can be exploited with great success in developing results and algorithms that are not available or are significantly less efficient for homogeneous fields.
In our earliest work [16] we used this structure to develop efficient Levinson-type algorithms for computing the weighting pattern of the optimal estimator for a random field observed in additive noise. In our more recent research we have used our results on boundary-value processes to develop recursive, Kalman filter-like implementations of the optimal estimators for isotropic fields described by certain partial differential equations [22]-[23] and have developed counterparts of two 1-D methods, namely the MUSIC algorithm and the maximum entropy method (MEM), for spectral estimation of isotropic random fields [17]-[20], [26]. These results, which have application in a variety of signal processing contexts (oceanography, seismics, etc.), illustrate the considerable level to which isotropy can be exploited to develop algorithms that are far more efficient than their homogeneous counterparts. One important aspect of our work in this area is the representation of an isotropic random field $x(r)$ in a Fourier series with respect to $\theta$, where $(r,\theta)$ are polar coordinates for the plane. This allows us to focus on the Fourier coefficients which are functions of the scalar radius $r$, and this leads directly to algorithms that have radially outward and inward directions of recursion.

Although the results that we have obtained were derived for scalar isotropic fields, we believe that they can be extended to a wider class of multidimensional random fields. An interesting such extension would be to isotropic vector fields, since these fields arise naturally in the study of physical (electromagnetic, acoustic) phenomena in several dimensions. Other extensions include the development of estimation techniques for random fields defined as hexagonal lattices, as well as the development of spectral estimation methods for anisotropic fields capable of estimating the principal directions in such fields.

5. Inverse Problems [27] - [44], [48] - [53]

The main accomplishment of our research in this area has been to develop efficient signal processing algorithms for solving inverse problems of the type arising in exploration geophysics, remote sensing and ultrasonic imaging. For 1-D inverse problems, which correspond to media whose properties vary in only one direction, we have obtained in [27]-[30] some efficient layer-stripping algorithms which reconstruct the medium layer by layer, in a recursive fashion. These algorithms are extremely fast and are also stable numerically. In [31] this approach was then extended to a multidimensional inverse scattering problem associated to the time-invariant Schrodinger equation. The solution that we obtained is exact, i.e. no approximations such as the Born or Rytov approximation are introduced, and it is considerably simpler than previous solutions of this problem. It also relies on the layer stripping principle, whereby the medium is reconstructed layer by layer along a plane wavefront. More recently, we have focused on developing more practical approximate multidimensional inversion techniques. In [32]-[34] we have developed an entirely new approach to the Born inversion problem which relies on the concept of backpropagated field. In
this approach, the scattered field observed along a receiver array is propagated backwards in time into the scattering medium and is imaged and filtered to give the reconstructed velocity profile of the medium. This procedure, which can be viewed as an improvement on the existing migration methods of exploration geophysics, has already received a significant amount of attention (it has been implemented on real data at Schlumberger), and is therefore likely to have an important impact on the field of inversion. We have also developed in [35]-[38] an alternative approach which formulates the inversion problem as a generalized tomographic problem. In this approach, it is shown that, depending on the experiment geometry, we are given projections of the velocity function along curves such as circles, parabolas or ellipses. Then the reconstructed velocities are obtained by a two step backprojection and filtering technique which is analogous to the procedure used in X-ray tomography. In this procedure, the backprojection operation sums the contributions of all the projections passing through a given point, giving a reconstructed image to which a 2-D shift invariant filter is then applied to obtain the function we seek to recover. The case of a constant density medium probed by an incident plane wave was considered in [36]-[37], whereas in [38] we examine the problem of reconstructing both the velocity and density functions of an acoustic medium from several plane wave experiments. We feel that although both our backpropagated field and tomographic multidimensional Born inversion methods have already had a significant impact, they are applicable to a far wider context that would include for example the study of inverse problems for EM or elastic wave phenomena.

More recently, in [39] we have developed an identification and estimation based approach to an inverse resistivity problem. In this problem we are given several experiments where a potential distribution is applied on the boundary of an object, and where the normal current is observed on this boundary, and we seek to recover the resistivity function inside the object. The identification-based approach that we have developed leads to the iterative solution of a nonlinear least-squares problem. The solution of this type of problem requires usually a large amount of computation, but the approach that we propose consists of solving this problem at various resolution levels, going from coarse to finer resolution levels. The procedure that we plan to use to improve the resolution has some similarities with multigrid solution techniques for partial differential equations, and it has therefore the potential to provide a very efficient approach for solving iteratively multidimensional inverse problems.

In parallel with the above work, we have also examined recently in [40]-[44] several problems of source location. Specifically, in addition to the traditional problem of estimating the location of radiating sources, we have also examined the problem of estimating parameters related to the radiation patterns of the sources and receivers. The techniques that we studied in this context are a direct iterative maximum likelihood method, the EM (Estimation-Maximization) algorithm, the eigenstructure (MUSIC) approach and the polynomial approach. The resulting methods, such as the
improved MUSIC method described in [42], present a number of advantages, since they do not require an exact calibration, of the receiver array, and can handle the case where the receiver array is not uniform [41].

Our original proposal noted our intention to investigate a wide range of inverse problems, in the expectation that techniques of identification and estimation theory would lead to significant new insights in our understanding of inverse problems. One application area with which we have had considerable success recently is analytical chemistry.

A common situation in analytical chemistry is that coupled experimental procedures are carried out on a sample containing a mixture of unknown chemical components. The objective is to determine the qualitative and quantitative composition of the sample from resulting data. The multidimensional flavor of the inversion is a consequence of the coupling of the separate procedures. Our studies have been restricted to the case of bilinearly coupled procedures, which lead to 2D data arrays with special structure.

In a typical bilinear case, the data array has the form

$$M = \sum_{i=1}^{r} c_i N_i + E,$$

where $N_i = x_i y_i^T$ and where:

- $M$ is the $p \times q$ data matrix; the matrix $E$ represents modeling and measurement error;
- the scalar $c_i$ represents the amount of the $i$th component;
- the vector $x_i$ represents the response pattern of a unit of the $i$th component of the first of the two coupled procedures; and the vector $y_i$ represents the response of a unit of the $i$th component in the second of the two coupled procedures.

Several questions may be posed in the context of the above model, all aimed at "inversion" of the measured, noisy data matrix $M$ to determine the $c_i$, $x_i$ and $y_i$ as well as possible. How many components are there, i.e. what is $r$? Given calibration data for a target component, e.g. given $N_1$, what is the amount of that component in the mixture, i.e. what is $c_1$? Given a parametric model of responses in one of the procedures, e.g. given that $x_i = f(p_i)$ for some (low-dimensional) vector of parameters $p_i$, what are the values of $p_i$ and $c_1 y_1$? These questions are of great practical interest in such fields as high-performance liquid chromatography (HPLC), which is one of the mainstays of analytical chemistry (and has several professional journals devoted almost exclusively to it).

Our work in [48]-[53] has provided answers, in the form of efficient
and numerically sound algorithms, to all the specific questions listed above, with very encouraging results in preliminary tests on real data from HPLC. Just as with our geophysical inversion results, the key to our progress has been the application and development of a modern estimation and signal processing perspective in the context of a physical problem.

Methods of numerical linear algebra have played a key role. We have, for instance, found that a so-called "total least squares" problem arises naturally in treating the case of very noisy data. Application of an elegant algorithm, recently developed in the numerical analysis literature and based on a singular value decomposition, allows us to obtain striking results in this noisy case. Other strong connections have been made with subspace methods for sensor array data currently being pursued in the signal processing literature.

Current work, initiated in [51], is aimed at extending the algorithms to the case of models displaying quadratic nonlinearities. This is partly in response to the observation that our bilinear algorithms can now resolve data that is too difficult for currently used approaches, so our algorithms start to "see" nonlinearities that do not bother the much coarser methods in current use.

6. Motion Estimation [45] - [47]

Our work in this area has focused on the problem of estimating optical flow, that is the velocity vector field at an instant in time in a sequence of images. In standard approaches to this problem the edges are first located on each of a sequence of images, and the movement of the edge from one frame to the next provides a measurement of the component of velocity normal to edge; the tangential component is then estimated by solving an optimization problem for a best, smooth fit. The best-known methods of this type in computer vision employ iterative optimization methods. In [45]-[47] we use the framework of boundary value estimation to investigate the problem of estimating the velocity vector along a boundary given measurements of its normal component and then of estimating the velocity field on the interior of a closed contour. We find that the problems being solved by previous algorithms are precisely boundary-value estimation problems. This not only allows us to use our results on efficient estimation to derive non-iterative optical flow reconstruction algorithms which are considerably faster than previous methods but also allows us to use the flexibility of model-based estimation to investigate both various modifications to the basic problem (such as spatially-varying weighting on the fit criterion) with no modification in the algorithm or its complexity and a variety of other problems in vision. The problems of computer vision for which we believe that our general estimation-based methodology is applicable include for example the problem of tracking the velocity field over time (i.e. viewing it as a 3-D stochastic process, with time as the third variable), the estimation of the depth field (i.e. the distance to objects in an image)
from an image sequence, as well as the shape from shading and stereo vision problems.

7. Estimation of Geometrical Features and Computational Geometry [56] - [70]

In our earliest work in this area [56]-[57], we focused on the estimation of simple parameters of object geometry -- location, size, eccentricity, orientation -- from tomographic measurements. The principal purpose of this work was the demonstration that a nonlinear estimation formulation of this geometric estimation problem could lead to efficient and high performance algorithms when the quality and quantity of data fell significantly below the level at which the far more ambitious goal of reconstructing an entire image becomes unachievable. More recently, in [62], [66], [69] we have investigated the use of Markov random field (MRF) models and simulated annealing to identify directional features in images (with the problem of identifying layering structures from geophysical data as motivation). This work not only has led to some efficient algorithms but has also produced a useful class of anisotropic fields obtained by directionally-dependent changes of scale on an isotropic random field. In addition in [65] we have had success in applying MRF/simulated annealing techniques to the problem of identifying regions in cross-sectional images of the brain. These results are presently being used at the Dana-Farber Cancer Institute in Boston.

In [59], [60], [63], [67], [68] we investigate several problems that fall within the domain of computational geometry. In particular we have developed new algorithms for the reconstruction of simplexes and convex sets given knowledge of subsets of their interior, exterior and boundary. Problems of this type arise in robot vision, medical imaging, and in the component analysis problem of chromatography. The system-theoretic perspective we have brought to such problems has resulted in algorithmic structures quite different from ones usually found in this field. In contrast to combinatorial/search approaches, our algorithms are iterative. This also allows us to investigate algorithm efficiency in terms of convergence properties, where, instead of focusing on fixed points of mappings we are led to study fixed figures of geometrical construction.

In some very recent work [61] we have begun to consider estimation theoretic versions of problems in computational geometry. In particular, we have investigated the problem of optimal estimation of a convex polygon given noisy measurements of its support function. The constraints on support functions together with the noisy nature of the data lead to a quadratic programming problem with considerable structure that can be exploited in developing efficient algorithms. In addition, this problem is a natural one for the problem of tomography in which there are additional constraints to exploit. In [69] we describe the first step in developing iterative tomographic reconstruction methods that blend estimation of geometric features such as object support and the full reconstruction of the object.
given a support estimate and an MRF model of the object profile.

The problem of estimating the curvature and other shape parameters of objects, from partial measurements, continues to be of interest. In particular, we have obtained extensions, [54], [55], of results of Van Hove that relate the curvatures of shadows to the curvature of the projected object. Our derivation is considerably more streamlined, and exposes more clearly the facts that lead to the results. Because of this, we have been able to generalize to arbitrary dimensions some of the major results that tell us what combinations of projections (how many, and of what dimensions) will permit determination of curvature of the projected object at a point on its surface.

The problem of estimating dynamically evolving geometric objects was posed in our proposal, and has been pursued. The thesis proposal [55] describes more fully our goal of applying the methods of estimation and signal processing to the problem of dynamic shape reconstruction. Certain prototype geometric reconstruction problems are described. Approaches to developing recursive and iterative algorithms for them are outlined, and illustrative results from numerical experiments are presented.

For instance, the structure-from-motion problem for a rigid assembly of points has been studied by several people. The focus has, however, always been on batch methods and on combinatorial results. We formulate the problem as one of bilinear estimation. This naturally suggests certain iterative algorithms, whose performance in simulations has been encouraging. The algorithms are less sensitive to noise than results based on combinatorial methods, and have the advantage of always providing an estimate of the object, which can be easily updated as new measurements are taken. We are currently analyzing the behavior of these algorithms to elucidate how the special features of the problem may be further exploited. Extensions to the case of non-rigid assemblies are also being studied, with the points evolving in accordance with some dynamic model.

The study of dynamically evolving ellipsoids that was outlined in our proposal has been pursued too, and preliminary results are summarized in [55]. Taking the ellipsoid specified by the condition $x^T H x = 1$, where $H$ is positive definite, we can induce dynamic behavior by imposing a dynamic evolution equation on $H$, for instance:

$$H_{k+1}^{-1} = A^T H_k^{-1} A + Q_k$$

where the symmetric matrix $Q_k$ represents a driving term, which may be noise. If the ellipsoid is observed in projection, we obtain an observation equation of the form

$$Y_k = C^T H_k^{-1} C + W_k$$
where $Y_k^{-1}$ is the matrix that defines the projected ellipsoid in the same way that $H$ defines the original ellipsoid. This provides us with a linear model for which estimation may be carried out by simple modifications of well known results, and such estimation is illustrated in [55]. The interesting part of this work is to see what the geometric interpretations and implications of these results are, and this effort is in progress.
The following is a list of individuals who took part in the research supported by NSF Grant ECS-8312921:

Prof. Alan S. Willsky, principal investigator
Prof. Bernard C. Levy, co-principal investigator
Prof. George Verghese
Dr. Martin G. Bello
Dr. David A. Castanon
Mr. Ramine Nikoukhah
Dr. Milton B. Adams
Prof. John N. Tsitsiklis
Dr. Ahmed H. Tewfik
Dr. Andrew E. Yagle
Dr. Cenzig Esmer soy
Mr. Ali Ozbek
Dr. Anne Rougee
Dr. David J. Rossi
Mr. Daniel Spielman
Dr. John P. Greschak
Mr. Jerry L. Prince
Dr. Carey Bunks
Mr. Dwayne Mann
Mr. Young-Jo Kim
Mr. Kenneth Chou
Dr. A.J. Weiss
Dr. R.R. Kim
Mr. J.M. Hummel
Mr. M.J. Boyd
Mr. S. Friedman
Mr. R. King
Mr. W.C. Karl
The publications listed below represent papers and reports supported in whole or in part by NSF Grant ECS-8312921.


11. R. Nikoukhah, "System Theory for Two Point Boundary Value Descriptor


65. Y.J. Kim, "Identification of Fluid Regions in Brain Computerized


