TITLE: A Center on Communications, Control and Computation

TYPE OF REPORT: Final

AUTHOR: Sanjoy K. Mitter

DATE: April 6, 1988

U.S. Army Research Office

CONTRACT OR GRANT NUMBER: DAAG29-84-K-0005

INSTITUTION: Massachusetts Institute of Technology

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
LIST OF MANUSCRIPTS SUBMITTED OR PUBLISHED UNDER ARO SPONSORSHIP INCLUDING JOURNAL REFERENCES: (*indicates manuscripts or papers in the last six months of the contract; abstracts appended, if submitted or published)


A Center on Communications, Control and Computation (unclassified)

Professor Sanjoy K. Mitter

During the grant period research was conducted on many fundamental aspects of communication, control and computation. The report summarizes some of the results achieved in the following areas:

i) Parallel and Distributed Computation

ii) Communication Aspects of Parallel and Distributed Computation.

iii) Convergence Analysis for Partially Asynchronous Algorithms

iv) Communication

v) Efficient Algorithms for Solving Elliptic PDEs

vi) Multisource Array Processing

vii) Estimation and Modeling of Spatially-Distributed Data

viii) Singularly Perturbed Systems

ix) Discrete Event Systems (continued on reverse)
19. abstract (continued)
  x) Decentralized Detection
  xi) Markov Chains and Simulated Annealing
  xii) Simulated Annealing
  xiii) Stochastic Quantization
  xiv) Sensitivity Optimization for Delay Systems


SCIENTIFIC PERSONNEL SUPPORTED BY THIS PROJECT AND DEGREES AWARDED:

Faculty

Dimitri P. Bertsekas
Pierre Humblet
Robert Gallager
Bernard C. Levy
Sanjoy K. Mitter
Daniel Stroock
Robert Tenney
John N. Tsitsiklis
George C. Verghese
Alan S. Willsky

Visitors

Vivek S. Borkar
Ioannis Karatzas
P.R. Kumar
Lennart Ljung
Jose M.F. Moura
Charles Rockland
Steven E. Shreve

Research Staff

Robert D. Bruen
Jose Marroquin
Research Assistants

Talissa Altes
Yehuda Avniel
Carey Bunks
Isidro Castineyra
Kenneth C. Chou
Saul Gelfand
Chung-Chieh Kuo
Richard Lamb
Whay Chiou Lee
Xi-C. Lou
Zhi-Quan Luo
J. Marroquin
P.T. Ng
Ali Ozbek
Jerry Prince
Thomas Richardson
A. Isaac Saias
Ahmed Tewfik
Paul Tseng
Albert Kai-Sun Wong
Students awarded degrees with ARO sponsorship:

Talissa Altes (Ph.D.)
Yehuda Avniel (Ph.D.)
Carey Bunks (Ph.D.)
Isidro Castineyra (Ph.D.)
Kenneth Chou (S.M.)
Saul Gelfand (Ph.D.)
Chung-Chieh Kuo (Ph.D.)
Richard Lamb (Ph.D.)
Xi-C. Lou (Ph.D.)
Jose Marroquin (Ph.D.)
P.T. Ng (Ph.D.)
Jerry Prince (Ph.D.)
Ahmed Tewfik (Ph.D)
Paul Tseng (Ph.D.)
Albert Kai-Sun Wong (Ph.D)
Andrew Yagle (Ph.D.)
BRIEF OUTLINE OF RESEARCH FINDINGS

During the grant period research was conducted on many fundamental aspects of communication, control and computation. The following sections summarize some of the results.

(i) Parallel and Distributed Computation

(ii) Communication Aspects of Parallel and Distributed Computation

(iii) Convergence Analysis for Partially Asynchronous Algorithms

(iv) Communication

(v) Efficient Algorithms for Solving Elliptic PDEs

(vi) Multisource Array Processing

(vii) Estimation and Modeling of Spatially-Distributed Data

(viii) Singularly Perturbed Systems

(ix) Discrete Event Systems

(x) Decentralized Detection

(xi) Markov Chains and Simulated Annealing

(xii) Simulated Annealing

(xiii) Stochastic Quantization

(xiv) Sensitivity Optimization for Delay Systems
1.1 Parallel and Distributed Computation

A variety of problems have been studied in this area:

1. We have completed a study of the communications requirements of distributed convex optimization. More specifically, we considered two processors, each one having access to a different convex cost function $f_i$, who communicated with the objective of computing some vector $x$ which minimizes $f_1(x) + f_2(x)$, within a prescribed accuracy $\epsilon$. Lower bounds were derived on the number of bits that have to be transmitted. An algorithm with optimal (respectively, almost optimal), communication requirements was found for the case where $x$ is one-dimensional (respectively, multi-dimensional). (Tsitsiklis and Luo, 1987).

2. We have completed a study of asynchronous iterations. Here an operator $T_i(t)$ (chosen from a finite set $\{T_2, ..., T_k\}$ of possible operators is applied at each time step, thus yielding the iteration $x(t+1) = T_i(t)(x(t))$. We have considered the convergence of this iteration for the case where every operator is applied an infinite number of times but in an arbitrary order. This situation is of interest in revealing distributed iterative algorithms as well as in other contexts such as the stability of time varying dynamical systems. General necessary and sufficient conditions for convergence have been derived. These results bear interesting similarities with direct and converse theorems in Lyapunov stability theory.

1.2 Communication Aspects of Parallel and Distributed Computation

We have conducted a study of some generic communication problems that arise in many types of parallel numerical algorithms. We have developed optimal or nearly optimal algorithms for each of these problems for some of the most popular processor interconnection networks including the hypercube. This work will be part of the forthcoming book Parallel and Distributed Computation: Numerical Methods by Professors Dimitri Bertsekas and John Tsitsiklis.

1.3 Convergence Analysis for Partially Asynchronous Algorithms

We have developed a convergence theory for partially asynchronous distributed algorithms that applies to special classes of problems including strictly convex network optimization problems. An interesting feature of this theory is that, for convergence, there is no restriction on the size of the interprocessor communication delays as long as they are bounded. This work will be part of the forthcoming book on parallel computing by Professors Dimitri Bertsekas and John Tsitsiklis and will appear in a joint paper by Paul Tseng, D. Bertsekas and J. Tsitsiklis.
1.4 Communication

There is no central unifying theory for data networks of the type that have been developed over the last twenty years. Most important network problems are intractable when modelled in a highly realistic way and we have studied a number of simple models, each focusing on an interesting aspect.

For wire networks, we have particularly studied issues in routing and flow control. The approach has been to set up an optimization problem on the flow of data traffic (including rejected traffic in the case of flow control) and to develop efficient centralized and decentralized algorithms to find the best operating point. These models have relied on making stationarity assumptions on the traffic statistics and on ignoring dynamic effects. Our research efforts attempted to understand those dynamic effects, particularly in the case of flow control techniques using windowing algorithms.

The study of routing issues has also led us to investigate distributed algorithms to solve graph oriented problems (shortest path, spanning trees, max flow) that are needed as building blocks in the solution of larger problems. This research is in addition to asynchronous distributed algorithms outlined elsewhere.

For networks involving broadcast media such as cable, radio and satellite, there is a major problem of contention between nodes attempting to access the channel. For the case where all nodes can hear each other we have developed a good theoretical understanding of the problem and we have designed and analyzed high performance access methods.

One can also view channel access problems within the framework of Information Theory, albeit with unusual assumptions relative timing at different transmitters and even on the composition of the active transmitter set. With these assumptions we have succeeded in reconciling results from random access contention and information theory. In the same vein, but on a more practical level, we examined the links between spread spectrum, error correction coding and random access.

Finally, we worked in the area of fiber based, very high bandwidth networks. They differ in such a dramatic fashion from traditional transmission media that the way they are used is bound to be very different from a straightforward extension of the current techniques.

1.5 Efficient Algorithms for Solving Elliptic PDEs

Professor Bernard Levy and his students have pursued work on the development of parallel and efficient algorithms for the solution of elliptic PDEs. A complete account of the results that we have obtained have appeared in Jay Kuo's Ph.D. thesis. The main contributions of this work are as follows:

(i) A local relaxation algorithm for the solution of elliptic PDEs with space dependent coefficients. This algorithm selects different relaxation parameters \( \omega_{ij} \) for different discretization points. The resulting numerical procedure is parallelizable, efficient and robust. It is faster than the standard SOR method for solving space-dependent elliptic PDEs.

(ii) A two-level four-color SOR method for solving higher-order discretizations of elliptic PDEs such as the 9-point stencil discretization of the Poisson equation. This method relies on a two-level iterative scheme and is based on a formulation of the SOR acceleration procedure in the Fourier domain. Both theoretical analysis and numerical simulations indicate that the resulting iterative procedure has a convergence rate similar to that of the standard 5-point SOR method.

(iii) Mode-dependent finite-difference discretization schemes for linear homogeneous PDEs. These discretization schemes are considerably more accurate than standard finite difference schemes and take into account the structure of the PDE of interest (such as the convection-diffusion equation, the Helmhotz equation, etc.) to approximate accurately the most significant modes appearing in its solution.

(iv) Finally, we have developed and analyzed multigrid PDE solvers which use a red/black SOR smoother with a value of the relaxation parameter \( W \neq 1 \), as the smoother on each grid of the multigrid solver. By performing a two-grid two-color Fourier analysis of the resulting multigrid solver it has been shown theoretically, and verified experimentally that the optimum convergence rate is obtained for \( w = 1.15 \) (when two smoothing iterations are used at each grid level), instead of \( w = 1 \) (The Gauss-Seidel smoother) as was previously believed. These results will be described in a paper which is currently in preparation.
1.6 Multisource Array Processing

Professor George Verghese and Mariam Motamed are continuing the study of signal processing for sensor array data in chemometrics, using methods of numerical linear algebra (subspace methods, eigenstructure methods, and so on). Of particular interest is the use of "total least squares" to obtain greatly improved results in the case of low signal-to-noise ratios.

1.7 Estimation and Modeling of Spatially-Distributed Data

We have continued our work on stochastic models of the geometry of random fields. The Ph.D. thesis of Jerry L. Prince is now complete. In this work, we have developed a number of estimation-based results in the context of tomographic reconstruction. The most significant of these are:

(i) The development of local relaxation algorithms for tomographic reconstruction based on sparse data together with geometric constraints arising from estimation of object geometry and mathematical constraints required of all Radon transforms. These algorithms are extremely fast, parallelizable, and robust.

(ii) The development of algorithms for the estimation of object geometry (more precisely estimation of the support function of the object). This has led to new results and a continuing investigation into probabilistic models of shape that (a) capture desired geometric features; and (b) lead to efficient and robust algorithms.

(iii) The development of local relaxation algorithms that incorporate the fundamental constraints required of all Radon transforms.

At least four papers will result from all of this research.

We have pursued the development of estimation-based approaches to inverse problems. In this work we formulate the inverse problems (at present an inverse resistivity problem) as an estimation problem at multiple spatial scales. This approach (a) allows one to overcome well-posedness problems and to identify a resolution consistent with the available data; (b) overcomes much of the computational burden in solving the forward problem at each iteration by performing the bulk of iterations at coarse scales; and (c) leads to a highly modular and parallel inversion scheme in which identical problems are solved over subregions of the medium being imaged.
One stage of our work in this area, including experimental and analytical performance studies has now been completed and is documented in the S.M. thesis of Mr. K.C. Chou and in a forthcoming paper. Further work is continuing, focusing now on the investigation of multi-scale representations of stochastic processes and random fields. In addition a project on the development of probabilistic model-based algorithms for recovering 3-D structure from X-ray crystallographic measurements has recently been initiated.

We have also continued our research on problems of computational vision based on the theories of spatial estimation we have developed over the past few years. Our present research focuses on problems of motion tracking over time from image sequences, depth estimation, and the development of algorithms that avoid the use of the so-called “brightness constraint.”

Clem Karl has continued his work with Professor George Verghese on reconstruction of objects from shadows. A paper on this work is in preparation. Most recently, they have been extending earlier results on ellipsoids and convex objects to the case of star-shaped objects such as those represented via spherical harmonic surfaces. Some progress has been made within this framework towards the objective of representing and estimating dynamically evolving shapes.

1.8 Singularly Perturbed Systems

Our work on multiple time scale analysis during the past six months has focused on documenting previous results of J.R. Rohlicek, X.-C. Lou, A.S. Willsky and G.C. Verghese and on one new research direction developed in the recently completed S.M. thesis of C.A. Caromicoli. In this work we (a) apply our methods to the analysis of complex, flexible manufacturing systems; and (b) develop new theoretical results on the computation of asymptotic rates of particular events in perturbed Markov processes. In addition we have recently initiated an investigation of estimation for such processes and have had some initial success in developing asymptotic optimality results.

The Master’s thesis of Cuneyt Ozveren, supervised by Professors Alan Willsky and George Verghese, has been extended. A paper describing these extensions, treating asymptotic orders of reachability in perturbed systems over the ring of asymptotic series in some small parameter is to be published.
1.9 Discrete Event Systems

Professor Alan Willsky and his students have recently initiated a research project aimed at developing a higher-level, linguistic control theory for systems characterized by the occurrence of discrete events. There has been a recent burst of activity in this area, sparked by the work of W.M. Wonham who has investigated control concepts in contexts more usually used by computer scientists. Our work has as its aim the development of important concepts and results lacking in the present theory. Specific results to date are:

(i) The definition and development of criteria for the stability, or more precisely, the resiliency, of discrete event systems.

(ii) The development of the fundamentals of a hierarchical theory for such systems based on the notion of "tasks."

(iii) The development of a theory of system interconnections and conditions under which analysis of such interconnections are computationally tractable.

(iv) The development of a class of Petri net models describing the coordination in distributed models of the heart.

Professor Willsky, Mr. Ramine Nikoukhah and Professor B.C. Levy of U.C. Davis have continued their research on developing a system theory for discrete-time noncausal models described by boundary-value equations. Several papers are in progress dealing with reachability and observability, stability, stochastic and deterministic stationarity, realization theory, and optimal smoothing. Work is continued on these subjects and on the problems of stochastic realization and identification.

Our work from this point will focus on (a) developing methods of aggregation and decomposition that allow us to overcome the complexity associated with analyzing or designing these systems and (b) the development of a regulator theory.
1.10 Decentralized Detection

We have completed a study of decentralized detection problems in which a large number $N$ of independent sensors receive measurements from the environment and transmit a finite valued function of their measurements to a fusion center. Then, the fusion center uses the messages received to decide which one of $M$ hypotheses on the environment is true. The problem consists of finding the optimal messages to be transmitted by the sensors, so as to maximize the probability of a correct final decision by the fusion center. This problem is computationally intractable when $N$ is large but finite. Still, we have succeeded in obtaining a simple, complete and computationally easy solution, for the limiting case where $N$ tends to infinity. Both Bayesian and Neymann-Pearson formulations have been considered.

1.11 Markov Chains and Simulated Annealing

We undertook a theoretical study of the structure of finite state, discrete time Markov chains in which transitions between different states have probabilities of occurrence which are of different orders of magnitude, as a function of a small parameter. While singular perturbation theory has addressed this problem, our results were of a different flavour because we considered non-stationary Markov chains and because we were able to obtain a graph-theoretic algorithm which provides information on the order of magnitude of the probability that a certain transition occurs over a long enough time interval, even if this probability was vanishingly small. (Singular perturbation theory usually determines only whether a certain probability goes to zero or not; we were able to determine the rate at which it goes to zero.)

The above outlined theory was then used to analyze the asymptotic behaviour of simulated annealing algorithms. In particular, given any Markov chain whose transition probabilities are controlled by a decreasing and asymptotically vanishing small parameter (this the temperature in the simulated annealing context) we can determine the set of states which will have positive probability, in the limit as time goes to infinity. In particular, these results provide a different perspective on previously available results or this subject.

In more technical detail, the results were the following: A discrete time, nonstationary, finite state Markov chain $\{x(t)\}$ is considered whose one-step transition probabilities satisfy inequalities of the form $C_1\epsilon^{a(i,j)} \leq p_{ij}(t) \leq C_2\epsilon^{a(i,j)}$, where $\epsilon$ is a small positive parameter, $C_1$, $C_2$ are constants and $A = \{a(i,j)\}$ is a set of nonnegative integer coefficients, which determine the order of magnitude of the probabilities of different transitions. Subject to a minor aperiodicity assumption, it is shown that, for any nonnegative $d$ there exists a set of nonnegative integer coefficients $\{A(d;i,j)\}$ (depending only on $A$) such that $C_1\epsilon^{a(d;i,j)} \leq P(x(1/\epsilon^d) = j | x(0) = i) \leq C_2\epsilon^{a(d;i,j)}$, where $C_1$, $C_2$, are constants independent of $\epsilon$. An
algorithm for computing \( V(d; i, j) \) is also provided. It is then assumed that \( \epsilon \) is actually a nonincreasing function \( \epsilon(t) \) of time, as is the case in simulated annealing. It is shown that, if \( \sum_{i=0}^{\infty} \epsilon^d(t) = \infty \) and \( \sum_{i=0}^{\infty} \epsilon^{D+1}(t) < \infty \), then the smallest subset of the state space to which \( x(t) \) converges (in probability) equals the set \( R^d \) defined by \( R^d = \{ i : V(d; i, j) = 0 \} \implies V(d; j, i) = 0 \). The proof consists of partitioning the infinite time horizon into a sequence of intervals such that \( \epsilon(t) \) is approximately constant during each such interval and then applying the previously mentioned estimates.

1.12 Simulated Annealing

In the thesis of Saul Gelfand, he has given an analysis of the relationship between the Langevin equation method for finding a global minimum of a function and the method of simulated annealing. He has also proposed a new algorithm which exploits the desirable properties of the two above algorithms.

1.13 Stochastic Quantization

Sanjoy K. Mitter, in joint work with V.S. Borkar and R.T. Chari, is continuing work in this area as follows. The study of stochastic quantization leads to a study of a distribution-valued stochastic differential equations. A stochastic differential equation describing a symmetric Markov process taking values in the space of distributions on a planar rectangle or the whole plane has been studied. Main result include the existence and uniqueness of the solution, ergodicity, a finite to infinite volume limit theorem and partial results on large deviations. The problem originates from quantum field theory and has implications in image analysis. This work is reported in “Stochastic Quantization of Field Theory in Finite and Infinite Volume,” to appear in the Journal of Functional Analysis in 1988.

1.14 Sensitivity Minimization for Delay Systems

In the work of Flamm and Mitter we solved the \( H^\infty \) sensitivity optimization for plants with an input delay and general rational transfer function. We regard this as a major contribution and this research has important implications in modelling unmodelled dynamics using variable time delays.
2. Interaction with Army Research Office and Army Laboratories

Professor Daniel Stroock gave a mini-course organized by the Army Research Office on "Theory of Large Deviations" at the University of Maryland, October 25-27, 1987.

Professor Sanjoy Mitter visited the U.S. Army's Human Engineering Laboratory on December 21, 1987. His host was Dr. Benjamin Cummings. On December 22, Professor Mitter was invited to visit the U.S. Army's Engineer Topographic Laboratory at Ft. Belvoir, Virginia, for a briefing for personnel associated with the ARO.

3. Honors, Awards and Invited Lectures

Professor Sanjoy Mitter was elected to the National Academy of Engineering "for outstanding contributions to the theory and applications of automatic control and nonlinear filtering."

Professor Mitter was also invited to give the lecture "On Stochastic Quantization," at the First International Conference on Industrial and Applied Mathematics, Paris, in July 1987. He was also invited to lecture on "Recursive Maximum Likelihood" at the IFAC '87 Conference, Munich, in July 1987. In addition, Professor Mitter lecture in the $H^\infty$ Workshop at the IEEE CDC Conference, Los Angeles, in December 1987.

Professor Robert Gallager was invited to speak in the Robert T. Chien Distinguished Lecture Series, Coordinated Science Laboratory, University of Illinois in October 1987. He was also invited to speak in the Distinguished Lecture Series at Concordia, Montreal, Canada.


Professor George Verghese was invited to give the talk "Subspace Methods in Chemometrics" at the International Conference on Linear Algebra and its Applications, Valencia, Spain, in September 1987. He was also invited to give the talk "Array Processing in Chemometrics" at the Indo-US Workshop on Systems and Signal Processing, India, in January 1988.

Professor Alan Willsky was invited to speak on "Generalized Riccati Equations for Two-Point Boundary-Value Descriptor Systems," at the IEEE Conference on Decision and Control, Los Angeles, in December 1987.
Upper Bounds for symmetric Markov transition functions

by

E. A. Carlen (*)
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA 02139, U.S.A.

S. Kusuoka (*) (**)
Department of Mathematics,
Faculty of Science, University of Tokyo,
Hongo, Tokyo 113, Japan

and

D. W. Stroock (*) (**)
Department of Mathematics,
Massachusetts Institute of Technology,
Cambridge, MA 02139, U.S.A.

ABSTRACT. — Based on simple ideas introduced by J. Nash, it is shown that certain estimates on the transition function for a symmetric Markov process are equivalent to certain coercivity conditions involving the associated Dirichlet form.

Key words. Symmetric transition functions, Dirichlet forms.

RESUME. — En s’inspirant de quelques idées simples introduites par J. Nash, il est démontré que certaines estimations de probabilités de transition pour un processus de Markov symétrique sont équivalentes à des conditions de coercivité sur la forme de Dirichlet associée.
Logarithmic Sobolev Inequalities and Stochastic Ising Models

Richard Holley and Daniel Stroock

Received March 24, 1986

We use logarithmic Sobolev inequalities to study the ergodic properties of stochastic Ising models both in terms of large deviations and in terms of convergence in distribution.

KEY WORDS: Logarithmic Sobolev inequalities, stochastic Ising models; large deviations; rates of convergence.

1. INTRODUCTION

The theme of this article is the interplay between logarithmic Sobolev inequalities and ergodic properties of stochastic Ising models.

To be more precise, let $g$ be a Gibbs state for some potential and suppose $\{P_t : t > 0\}$ is the semigroup of an associated stochastic Ising model. Then $\{P_t : t > 0\}$ determines on $L^2(g)$ a Dirichlet form $\mathcal{E}$. A logarithmic Sobolev inequality is a relation of the form

$$ (\text{L.S.) } \int f^2 \log \frac{f}{\mathbb{E} f} \, dg \leq z \mathcal{E}(f, f), \quad f \in L^2(g) $$

for some positive $z$ (known as the logarithmic Sobolev constant). In this article we discuss some of the implications that (L.S.) has for the ergodic theory of the stochastic Ising model.

In Section 2 we discuss ergodic properties from the standpoint of large-deviation theory. In particular, we introduce and compare rate
Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator

by S. Kusuoka* and D. Stroock*

0. Introduction

By now it is abundantly clear that precise local regularity properties of second order subelliptic operators depend heavily on the structure of the degeneracy being considered. Indeed, when we start with the paper [R-S] by L. Rothschild and E. Stein, in which the sharp form of Hörmander's famous subellipticity theorem is proved, and continue through the work of C. Fefferman and D. Phong [F] and A. Sanchez-Calle [S], it becomes increasingly clear that local regularity estimates for these operators reflect the geometry associated with the operator under consideration.

For example, if the operator $\mathcal{L} = \sum_{k=1}^{d} V_k^* V_k$, where the $V_k$'s are smooth vector fields on $\mathbb{R}^n$ (thought of as directional derivatives) and $V_k^*$ denotes the formal adjoint of the operator $V_k$, and if one defines $d(x,y)$ to be the $(V_1,\ldots,V_d)$-control distance between $x$ and $y$ (cf. Section 1), then, under a suitably uniform version of Hörmander's condition (cf. (3.14) in Section 3), one can show that the fundamental solution $p(t,x,y)$ to the Cauchy initial value problem for $\partial_t u = \mathcal{L} u$ satisfies an estimate of the form:

\[
\frac{1}{M_B(x,t^{1/2})} \exp\left[ - M d(x,y)^2 / t \right] \leq p(t,x,y) \leq \frac{M}{|B_d(x,t^{1/2})|} \exp\left[ - d(x,y)^2 / M t \right]
\]

for all $(t,x,y) \in (0,1] \times \mathbb{R}^n \times \mathbb{R}^n$, where $B_d(x,r) = \{ y \in \mathbb{R}^n : d(x,y) < r \}$. (This estimate was first derived by Sanchez [S] for $t \in (0,1]$ and $x$ and $y$ satisfying $d(x,y) \leq t^{1/2}$. More recently, it was extended to $(t,x,y) \in (0,1] \times \mathbb{R}^n \times \mathbb{R}^n$ with $d(x,y) \leq 1$ by D. Jerison and Sanchez [J-S]; and, at about the same time, it was proved for general $x$ and $y$ by the present authors [K-S. III].)

*During the period of this research, both authors were partially supported by NSF DMS-6415211 and ARO DAAG29-84-K-0005.
PARAMETRIC NON-LINEAR FILTERING

by

Richard H. Lamb, Jr.

This report is based on the unaltered thesis of Richard H. Lamb, submitted in partial fulfillment of the requirements for the degree of Doctor of Science at the Massachusetts Institute of Technology in May 1987. The research was conducted at M.I.T. Laboratory for Information and Decision Systems with support provided by the Office of Naval Research under ONR Contract N00014-77-C-0532 and by ARO Contract DAAG29-84-K-0005.
FAILURE DETECTION AND IDENTIFICATION

Mohammad-Ali Massoumnia, George C. Verghese, Alan S. Willsky

Massachusetts Institute of Technology
Cambridge, MA 02139, USA

13 June, 1987

Abstract

Using the geometric concept of an unobservability subspace, a solution is given to the problem of detecting and identifying control system component failures in linear, time-invariant systems. Conditions are developed for the existence of a causal, linear, time-invariant processor that can detect and uniquely identify a component failure, first for the case where components can fail simultaneously, and second for the case where they fail only one at a time. Explicit design algorithms are provided when these conditions are satisfied. In addition to time domain solvability conditions, frequency domain interpretations of the results are given, and connections are drawn with results already available in the literature.

* Work of the first author was supported by NASA Langley Research Center under Grant NAG1-128. The other two authors were supported in part by the Air Force Office of Scientific Research under Grant AFOSR-82-0258 and in part by the Army Research Office under Grant DAAG-29-84-K-0005.
1 Formerly with the Space Systems Laboratory at MIT, now with the Sharif University of Technology, Tehran, Iran.
2 Laboratory for Electromagnetic and Electronic Systems.
3 Laboratory for Information and Decision Systems.
STABILITY, STOCHASTIC STATIONARITY AND GENERALIZED LYAPUNOV EQUATIONS
FOR TWO-POINT BOUNDARY-VALUE DESCRIPTOR SYSTEMS

Ramine Nikoukhah 2, Bernard C. Levy 3, Alan S. Willsky 4

ABSTRACT

In this paper, we introduce the concept of internal stability for two-point boundary-value descriptor systems (TPBVDSs). Since TPBVDSs are defined only over a finite interval, the concept of stability is not easy to formulate for these systems. The definition which is used here consists in requiring that as the length of the interval of definition increases, the effect of boundary conditions on states located close to the center of the interval should go to zero. Stochastic TPBVDSs are studied, and the property of stochastic stationarity is characterized in terms of a generalized Lyapunov equation satisfied by the variance of the boundary vector. A second generalized Lyapunov equation satisfied by the state variance of a stochastically stationary TPBVDSs is also introduced, and the existence and uniqueness of positive definite solutions to this equation is then used to characterize the property of internal stability.

1. This research partially supported by the U.S. Army Research Office, contract DAAL03-86-K-0171 (Center for Intelligent Control Systems), the Air Force Office of Scientific Research, contract AFOSR-88-0032, and the National Science Foundation, grant ECS-8700903.

2. Laboratory for Information and Decision Systems, Room 35-437, Massachusetts Institute of Technology, Cambridge, MA 02139.

3. Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

ON STOCHASTIC SCHEDULING WITH IN-TREE PRECEDENCE CONSTRAINTS

CHRISTOS H PAPADIMITRIOU AND JOHN N TSITSIKlis:

Abstract. We consider the problem of optimal scheduling of a set of jobs obeying in-tree precedence constraints, when a number M of processors is available. It is assumed that the service times of different jobs are independent identically distributed random variables. Subject to a minor assumption on the service time distribution, we show that policies of the "Highest Level First" type are optimal asymptotically, as the number of jobs tends to infinity.

Key words. scheduling, stochastic, trees

AMS(MOS) subject classification. 90B35

1. Introduction. Scheduling jobs with equal execution times and in-tree precedence constraints to minimize makespan (latest finishing time) is a classical problem in scheduling [Co]. It is known that, for any number of processors, the most natural policy, namely the one that assigns to the next available processor a leaf that has the largest height (distance from the root), is optimal [Hu]. This policy is called highest level first.

An interesting twist on this theme is to allow the execution times of the tasks to be independent and identically distributed exponential random variables [CR]. The results here are less complete: It is known that the highest level first policy is optimal only in the two-processor case [CR]. In fact, for two processors this policy is optimal under a wide variety of criteria [BR]. Unfortunately, for three processors the highest level first policy is not optimal, and no tractable way to approach this problem is known. It was shown in [Pa] that a generalization of the two-processor problem in another direction (and-or precedences) is PSPACE-complete.

In this paper we show that for any number of processors the highest level first policy, although suboptimal, is not much worse than the optimum. In particular, we show that, for any in-tree, the cost associated with a highest level first policy is no larger than the optimal cost times a factor that goes to one as N increases to infinity.

Moreover, this result is shown to be true for a fairly wide class of service time distributions, exponential distributions being a special case.

2. Problem definition. We are given a set of M processors and an in-tree G with N nodes. Each node represents a job that may be processed by any of the processors. We assume that the jobs have service times that are independent and identically distributed random variables, with known distributions. We assume that these random variables are positive, with probability one.

A scheduling policy is a rule that, at time i = 0, assigns L of the processors to L leaves of the tree, where L is the minimum of M and the number of leaves of G. If at time t some processor terminates the processing of a job, we delete the corresponding
RECONSTRUCTING CONVEX SETS FROM SUPPORT LINE MEASUREMENTS

Jerry L. Prince, Alan S. Willsky

ABSTRACT

This paper proposes algorithms for reconstructing convex sets given noisy support line measurements. We begin by observing that a set of measured support lines may not be consistent with any set in the plane. We then develop a theory of consistent support lines which serves as a basis for reconstruction algorithms that take the form of constrained optimization algorithms. The formal statement of the problem and constraints reveals a rich geometry which allows us to include prior information about object position and boundary smoothness. The algorithms, which use explicit noise models and prior knowledge, are based on ML and MAP estimation principles, and are implemented using efficient linear and quadratic programming codes. Experimental results are presented. This research sets the stage for a more general approach to the incorporation of prior information concerning the estimation of object shape.

1. This research was partly supported by the U.S. Army Research Office grant DAAL03-86-K-0171 (Center for Intelligent Control Systems) and grant DAAG29-84-K-0005, and by the National Science Foundation, grant ECS-8312921. In addition, the first author was partially supported by a U.S. Army Research Office Fellowship.

2. Laboratory for Information and Decision Systems, Room 35-439, Massachusetts Institute of Technology, Cambridge, MA 02139.

3. Center for Intelligent Control Systems, Room 35-437, Massachusetts Institute of Technology, Cambridge, MA 02139.
A PROJECTION SPACE MAP METHOD FOR
LIMITED ANGLE RECONSTRUCTION¹

Jerry L. Prince ², Alan S. Willsky ³

ABSTRACT

We present a method to reconstruct images from finite sets of noisy projections which are available only over limited or sparse angles. The method solves a constrained optimization problem to find a maximum a posteriori (MAP) estimate of the full 2-D Radon transform of the object, using prior knowledge of object mass, center of mass, and convex support, and information about fundamental constraints and smoothness of the Radon transform. This efficient primal-dual algorithm consists of an iterative local relaxation stage which solves a partial differential equation in Radon-space, followed by a simple Lagrange multiplier update stage. The object is reconstructed using convolution backprojection applied to the Radon transform estimate.

1. This research was partly supported by the U.S. Army Research Office, contracts DAAL03-86-K-0171 (Center for Intelligent Control Systems) and DAAG29-84-K-0005, and by the National Science Foundation, grant ECS-8312221. In addition, the first author was partially supported by a U.S. Army Research Office Fellowship

2. TASC, 55 Walker’s Brook Drive, Reading, MA 01867. Formerly of the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

3. Center for Intelligent Control Systems, Room 35-437, Massachusetts Institute of Technology, Cambridge, MA 02139.
GEOMETRIC MODEL-BASED ESTIMATION FROM PROJECTIONS

by

Jerry Ladd Prince

Submitted to the Department of Electrical Engineering and Computer Science on January 2, 1988 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

This thesis addresses the problem of image reconstruction from noisy and limited- or sparse-angle projections. An algorithm is presented for the estimation of the maximum a posteriori (MAP) estimate of the full sinogram (which is an image of the 2-D Radon transform of the object) from the available data. It is implemented using a primal-dual constrained optimization procedure, which solves a partial differential equation in the primal phase using an efficient local relaxation algorithm, followed by a simple Lagrange multiplier update in the dual phase. The sinogram prior probability is given by a Markov random field (MRF) which includes information about the mass, center of mass, and convex hull of the object, and about the smoothness, fundamental constraints, and periodicity of the 2-D Radon transform. The geometric information reflected in the MRF formulation is estimated hierarchically, in part by new set reconstruction algorithms developed herein. These algorithms in turn are based on probabilistic estimation formulations which incorporate prior information about the size, position, and shape of the object. In particular, knowledge of the eccentricity, orientation, and boundary curvature may be used. This thesis contributes to the state of knowledge in the field of computed tomography, particularly in those disciplines in which noise and limited-data is a problem, and in the field of computational geometry, where our probabilistic formulations provide new and interesting insights on the problem of convex set reconstruction from support line measurements.

Thesis Supervisor: Alan S. Willsky
Title: Professor of Electrical Engineering
STRUCTURAL DECOMPOSITION OF MULTIPLE TIME SCALE MARKOV PROCESSES

J. R. ROHLICEK
BBN Laboratories Incorporated,
10 Moulton St., Cambridge, MA 02238.

A. S. WILLSKY
Laboratory for Information and Decision Systems,
Massachusetts Institute of Technology,
Cambridge, MA 02139.

Abstract
A straightforward algorithm for the multiple time scale decomposition of singularly perturbed Markov processes has been presented in [1,2]. That algorithm provides a uniform approximation of the probability transition function over the interval 0 ≤ t ≤ T through the construction of a sequence of aggregate models each valid at successively slower time scales. When only the structure of these models is desired, the algorithm can be expressed simply in terms of graphs associated with each of the aggregated models. The major computation then becomes computing shortest paths in these graphs. This representation of the algorithm furthermore allows analysis of more complex systems where there are multiple perturbation parameters with unknown relative orders of magnitude.

Keywords: Markov processes, singular perturbation, multiple time scales, graph theory.

1 INTRODUCTION
In this paper, an algorithm for the analysis of the multiple time scale structure of a perturbed Markov process is developed which builds directly on the multiple time scale decomposition algorithm presented in [1,2]. Structure of a perturbed Markov process refers to the complete multiple time scale decomposition of the type performed in [1,2,3,4] where the evolution of the state probabilities is approximated using a sequence of aggregated models, each valid at successively slower time scales. In this case, however, only the position of the nonzero transition rates of each of the unperturbed, aggregated time scale models, and the sets of states which constitute the aggregate classes, are determined. Although the detailed behavior of the system cannot be recovered from these descriptions, much useful information is retained. It will be shown that: this subset of information about the decomposition can be obtained using a very simple graph-theoretic algorithm which is implicit in the algorithm presented in [1,2].

Use of these structural aspects of a Markov chain can have many applications. For instance, an algorithm similar to that presented in this paper has been applied to the analysis of the behavior of "simulated annealing" optimization methods [5]. Other applications include computing order of magnitudes for the time of events in systems with rare transitions. Specific applications include analysis of the failure time of a fault-tolerant system or the error rate of a communication protocol.

The nature of the decomposition algorithms presented in this paper are very different from those which attempt to approximate the detailed behavior of system. In those algorithms, numerical calculations of dominant eigenvectors of unperturbed systems, and of dependent "trapping" probabilities were required. The algorithms in this paper basically consist of computing various types of connectivity in labeled graphs. Since the computations involve only integer quantities, issues of numerical stability are not relevant. Also, by introducing the graph-theoretic formalism, it is possible to employ standard graphical algorithms for computing quantities such as shortest paths between vertices of a graph.

This paper is organized as follows. In the next section, the decomposition algorithm and uniform approximation result present in [1,2] is stated and a simple example is provided. In the next section, the graphical formalism is introduced and the new graphical decomposition algorithm is presented. The final section includes conclusions and a discussion of the relationship of this approach to other work. Also, the application of this algorithm to models with multiple perturbation parameters is discussed with application to a fault-tolerant system model.

*This research was conducted under the support of the Air Force Office of Scientific Research under grant AFOSR-82-0256 and the Army Research Office under grant DAAG-29-84-K-008.
A SURVEY OF LARGE TIME ASYMPTOTICS OF SIMULATED ANNEALING ALGORITHMS

John N. Tsitsiklis
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

Simulated annealing is a probabilistic algorithm for minimizing a general cost function which may have multiple local minima. The amount of randomness in this algorithm is controlled by the "temperature", a scalar parameter which is decreased to zero as the algorithm progresses. We consider the case where the minimization is carried out over a finite domain and we present a survey of several results and analytical tools for studying the asymptotic behavior of the simulated annealing algorithm, as time goes to infinity and temperature approaches zero.

I. Introduction.

Simulated annealing is a probabilistic algorithm for minimizing a general cost function which may have multiple local minima. It has been introduced in [1] and [2] and was motivated by the Metropolis algorithm [3] in statistical mechanics. Since then, it has been applied to a variety of problems, the main ones arising in the context of combinatorial optimization [1,4,5,6] and in the context of image restoration [7].

Let $S = \{1, \ldots, N\}$ be a finite state space, let $\mathbb{Z}$ be the set of nonnegative integers and let $J : S \to \mathbb{Z}$ be a cost function to be minimized. We assume that, for each $i \in S$, we are also given a set $S(i) \subseteq S$, to be called the set of neighbors of $i$. Let us assume that

$$j \in S(i) \quad \text{if and only if} \quad i \in S(j). \quad (1.1)$$

The neighborhood structure of $S$ may be also described by a graph $G = (S, E)$, where $E$, the set of edges is defined by $E = \{(i,j) : j \in S(i)\}$. Given a neighborhood structure, a natural method for trying to optimize $J$ is the "descent" method: given the present state $i \in S$, one examines the neighbors of $i$ and lets the state become some $j \in S(i)$ such that $J(j) \leq J(i)$. The descent method, in general, cannot find a global optimum; it is possible that $J(i) < J(j)$, $\forall j \in S(i)$, without $i$ being a global minimizer of $J$. Multistart algorithms provide a popular modification of the descent method in which the above described procedure is repeated, starting from random and independently chosen initial states. This guarantees that eventually a global minimum will be reached, but depending on the structure of the problem, this may take too long.

* Research supported by the Army Research Office under contract DAAAG-29-64-K-0005.
Stable, robust tracking by sliding mode control

George C. VERGHESE **, Benito FERNANDEZ R. ***, and J. Karl HEDRICK ***
Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 12 June 1987

Abstract: Sliding mode control is examined from the perspective of obtaining stable and robust tracking of an arbitrary time-varying reference by a multi-input-output linear, time-invariant system driven by a certain class of bounded errors, nonlinearities and disturbances. Most existing sliding mode schemes for such systems are subsumed by the one presented here. The results are developed via the use of inverse models and make clear the constraints imposed by the finite and infinite zero structure of the system. In particular, stable and robust tracking is shown to be obtained by the scheme in this paper if and only if the system is minimum phase. The blending of state-space and polynomial matrix descriptions allows our results to be derived cleanly and generally.

1. Stable, robust tracking

Our starting point is the system
\[ \dot{x}(t) = Ax(t) + Bu(t) + v(x,t), \]  
\[ y(t) = Cx(t), \]

where \( x \) is the \( n \)-dimensional state vector and \( u, y \) are \( m \)-dimensional control input and tracking output vectors respectively. The vector \( v \) represents model errors, nonlinearities, and disturbances. (The time argument \( t \) will be dropped for notational simplicity wherever it is not needed for clarity or emphasis.) One can obviously not ask for more independently tracking outputs than there are control inputs. On the other hand, if the number of desired tracking outputs is fewer than the number of control inputs, then there are additional degrees of freedom that may be exploited, and some remarks on the possibilities are made in the Conclusion of this paper.

The control objective is to pick \( u \) as a function of the measured \( x \) so as to have
\[ e(t) = y(t) - y_d(t) \]
go to zero, where \( y_d \) is the desired output, i.e., the tracked reference, and is assumed to be sufficiently differentiable – the precise degree of differentiability will become explicit later. We also require that the dynamics of \( x \) be independent of \( u \) – which gives rise to the label 'robust' – and that when \( y_d = 0 \) identically then \( x \) goes asymptotically to zero – hence the label 'stable'.

1.2. Assumptions

Assume that the system (1) is controllable by \( u \) and observable from \( y \). Also impose the standard
EN D DATED FILM 8-88 DTIC