Applied Mathematics: Methods

Nonlinear Partial Differential Equations in Image Processing and Computer Vision

Alfred Carasso

Over the last ten years, the use of nonlinear partial differential equations (PDE) methods in image processing has mushroomed into a major new topic in applied mathematics, with hundreds of publications and numerous symposia and workshops. A recent *SIAM Conference on Imaging Science* held in Boston in March 2002, attracted large numbers of participants and featured numerous talks devoted to this subject. Significant contributions to this area continue to be made by researchers in several disciplines, including mathematics, electrical engineering, computer science, and optics. Three major textbooks have appeared: *Geometric PDEs and Image Analysis*, by Guillermo Sapiro (2001), *Level Set Methods*, by James Sethian (1998), and *Anisotropic Diffusion and Image Processing*, by Joachim Weickert (1998). In addition, a special issue of the *IEEE Transactions on Image Processing*, (March 1998), was entirely devoted to this topic. Such nonlinear processing is widely believed to produce better imagery than had previously been possible. Critical study of this developing methodology and evaluation of its potential application to NIST imaging problems is an ongoing MCSD activity. This involves extensive computational exploration of these new procedures, using real or synthetically degraded images. Expertise in numerical PDE computations, coupled with long research experience in ill-posed deconvolution problems is an important MCSD asset in this work.

Image denoising and image deblurring are two basic tasks in computer vision. Given a noisy image, smoothing that image by convolution with a Gaussian distribution is commonly used to remove noise. This is mathematically equivalent to solving the linear isotropic diffusion equation forward in time, with the noisy image as initial data. However, such isotropic processing also blurs out important edge information along with removing noise. In an influential 1990 paper, Perona and Malik advocated the use of nonlinear anisotropic diffusion equations, where the diffusion coefficient is a function of the image gradient and is designed to become vanishingly small at points where the gradient is large. This results in an adaptive smoothing process that removes noise while preserving edges. Impressive results have been obtained using variants of these ideas in denoising medical, fingerprint, and document images. This is illustrated below with the Marilyn Monroe image in Figure 1. Here, a well-defined sharp image is artificially corrupted by the addition of a large amount of noise. Linear PDE denoising noticeably blurs the image, while nonlinear PDE denoising results in a sharper image.

Image deblurring is a more delicate problem in which seriously attenuated features need to be restored, usually in the presence of small amounts of noise. That problem is notoriously ill conditioned and must be treated by some type of regularization in order to prevent overwhelming noise amplification. This modified penalized problem is then recast as the problem of minimizing an appropriate cost functional that includes the regularization terms. Numerous such cost functionals have been proposed and used over the last 25 years. In order to accurately recover edge information, Rudin and Osher (1994) proposed penalizing the *total variation* (TV) norm of the image, an approach that continues to stimulate a great deal of research. The Euler-Lagrange problem for minimizing the total variation can be written as a nonlinear anisotropic diffusion equation, with a forcing term that describes convolution of the unknown image with the known point spread function. This is supplemented by homogeneous Neumann boundary conditions together with the blurred image as initial data. Deblurring the image is equivalent to stepwise numerical computation of this nonlinear

initial value problem until a steady state is reached. In fact, several other recently proposed cost functionals, such as the Perona-Malik, Geman-McClure, and Herbert-Leahy formulations, also lead to similar anisotropic diffusion problems for image deblurring. The particular form for the nonlinear diffusion coefficient in each case is dictated by the type of penalty term that is included in the cost functional. All of these nonlinear PDE deblurring methods require sophisticated numerical differencing for their successful implementation. These methods are also computationally intensive, generally requiring several hours of CPU time for 512x512 imagery. In particular, they are not well suited for *real-time* applications.

Figure 1. Advantages of nonlinear PDE approach in denoising. (A) Noisy Marilyn Monroe image. (B) Denoising with linear isotropic heat equation blurs the image. (C) Denoising with nonlinear anisotropic diffusion equation preserves life like features.

Loss of Texture in Nonlinear PDE Deblurring. It develops that the deblurring problem displays a fundamentally different behavior in the nonlinear PDE approach than does the denoising problem. An important question is how well *fine scale* information is recovered by these nonlinear methods. Recent work by Morel, Yves Meyer, and others in the French school indicates that because of their fine texture, most natural images *are not of bounded variation*. However, the Rudin-Osher theory is based on minimizing image total variation. For this reason, total variation deblurring often displays the so-called *staircase effect*, whereby spurious piecewise constant regions appear in the deblurred image. In some cases, this results in an 'oil-painting' appearance that impairs recognition. Carasso has obtained several examples of this phenomenon, and is continuing to accumulate valuable experience in the behavior of such methods. Other researchers, including Nikolova, Lions, Santosa, Weickert, Scherzer and Sapiro, have obtained similar results. This indicates that nonlinear PDE methods must be used with caution when the desired true image is known to be finely textured. In Figure 2, a synthetically blurred noisy US Air Force resolution test chart (A), is deblurred using the total variation method (B), and the Perona-Malik method (C). Important fine-scale information has been eliminated in (B), while being reasonably well recovered in (C). However, the conventional Lucy-Richardson method (D), which is not based on PDEs, produces substantially better results.

NIST Strategic Focus Areas. *Homeland Security. Emerging Industries*: Information and Knowledge Management (dynamic data infrastructure).
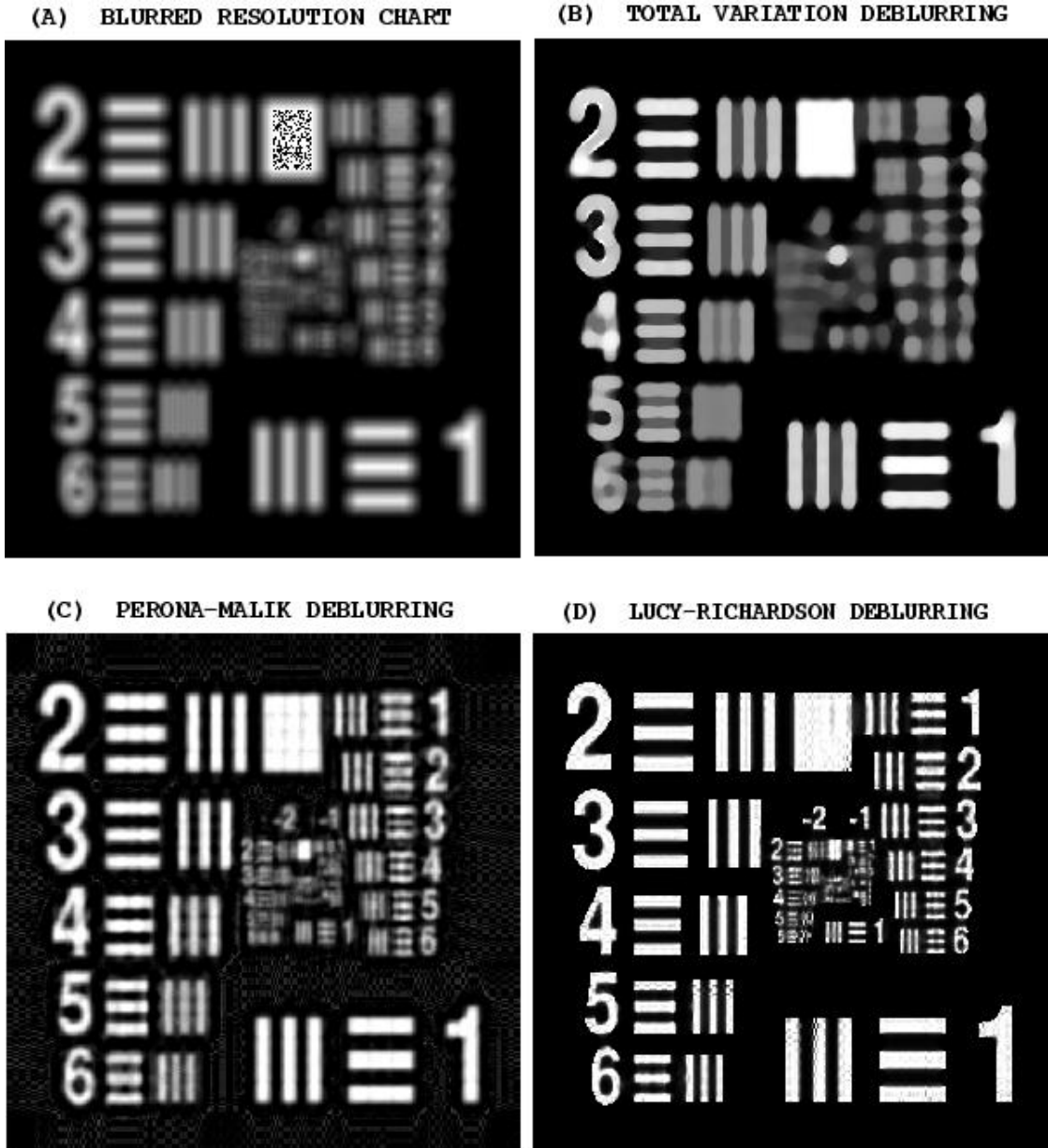
Figure 2. Nonlinear PDE approach in deblurring may eliminate texture and must be used with caution. (A) Synthetically blurred US Air Force resolution chart. (B) Total variation deblurring cannot recover small-scale features. (C) Perona-Malik deblurring is better behaved than total variation method. (D) Conventional Lucy-Richardson method, based on probabilistic approach, outperforms nonlinear PDE methods in this example.

Time-Domain Algorithms for Computational Electromagnetics

Bradley Alpert
Andrew Dienstfrey
Leslie Greengard (New York University)
Thomas Hagstrom (University of New Mexico)

http://www.boulder.nist.gov/itl/div891/alpert/td-em/index.html

Acoustic and electromagnetic waves, including radiation and scattering phenomena, are increasingly modeled using time-domain computational methods, due to their flexibility in handling wide-band signals, material inhomogeneities, and nonlinearities. For many applications, particularly those arising at NIST, the accuracy of the computed models is essential. Existing methods, however, typically permit only limited control over accuracy; high accuracy generally cannot be achieved for reasonable computational cost.

Applications that require modeling of electromagnetic (and acoustic) wave propagation are extremely broad, ranging over device design, for antennas and waveguides, microcircuits and transducers, and low-observable aircraft; nondestructive testing, for turbines, jet engines, and railroad wheels; and imaging, in geophysics, medicine, and target identification. At NIST, applications include the modeling of antennas (including those on integrated circuits), waveguides (microwave and photonic), transducers, and in nondestructive testing.

The objective of this project is to advance the state of the art in electromagnetic computations by eliminating three existing weaknesses with time-domain algorithms for computational electromagnetics to yield: (1) accurate nonreflecting boundary conditions (that reduce an infinite physical domain to a finite computational domain), (2) suitable geometric representation of scattering objects, and (3) high-order convergent, stable spatial and temporal discretizations for realistic scatterer geometries. The project is developing software to verify the accuracy of new algorithms and reporting these developments in publications and at professional conferences.

Earlier work of the project on nonreflecting boundary conditions developed an efficient and accurate approach for certain artificial (computational) boundaries, in particular planar, circular, and spherical boundaries. Acceptance of the method in the broad computational community has been slow, however, due to the complication of combining a Cartesian discretization with a circular or spherical boundary. More fundamentally, such a boundary implies excessive interior grid work for a source or scatterer that is much longer in one dimension than another. For these reasons, it is desirable to extend the work on nonreflecting boundary conditions to handle rectangular domains. Since the earlier work explicitly depends on a separation-of-variables formulation, this extension is far from straightforward. In the process of studying the relevant operators, Alpert has been joined by Yu Chen (New York University) in discovering a new metric for waves in unbounded domains. This inner product, which is based on energy considerations, appears to offer a new, possibly much more efficient, way of representing scattered waves in a variety of situations, including for waveguides, layered media, and beam construction. Its implications for inverse scattering problems are being evaluated. The work is detailed in the paper, "A Representation of Acoustic Waves in Unbounded Domains," NISTIR 6623 (2002).

Another spinoff of the work on nonreflecting boundary conditions, in which compressed representations of kernels of certain convolutions were derived, is the generalization of this work to other problems of model reduction. Model reduction for linear, stationary systems is a classical problem in control theory, with methods that assume a state-space problem formulation and exploit developments arising from fundamental mathematical work circa 1970 by Adamjan, Arov, and Krein on Hankel operators. In principle, these methods are also applicable to many problems that arise in applied mathematics, including rational approximation of mathematical or experimentally measured

functions and accurate, stable discretization of linear PDE. Nevertheless, the control theory methods have significant limitations, including the state-space assumption and lack of numerical robustness, and are rarely used in applied mathematics. Alpert and G. Beylkin (University of Colorado) have pursued generalizations to the rational approximation procedure developed for the nonreflecting boundary kernels, and have recently developed a robust procedure that avoids any state-space requirement. In preparation for reporting the new method, its connections to filter design, control theory, and stable discretization are being explored.

This year the papers, "Nonreflecting Boundary Conditions for the Time-Dependent Wave Equation," B. Alpert, L. Greengard, and T. Hagstrom, and "Adaptive Solution of Partial Differential Equations in Multiwavelet Bases," B. Alpert, G. Beylkin, D. Gines, and L. Vozovoi, appeared in *Journal of Computational Physics*.

The work of the project is supported in part by the Defense Advanced Research Projects Agency (DARPA). The work has been recognized by researchers developing methods for computational electromagnetics (CEM) and has influenced work on these problems at Boeing and HRL (formerly Hughes Research Laboratories). It has also influenced researchers at Yale University and University of Illinois. In each of these cases, new research in time-domain CEM is exploiting discoveries of the project. In particular, some efforts for the new DARPA program on Virtual Electromagnetic Testrange (VET) are incorporating these developments. We expect that design tools for the microelectronics industry and photonics industry, which increasingly require accurate electromagnetics modeling, will also follow.

NIST Strategic Focus Areas. *Homeland Security. Existing Industries*: Electronics.

Hierarchical Control of Some Advection Diffusion Equations

Anthony Kearsley
Paul Boggs (Sandia National Laboratory)
Jon Tolle (University of North Carolina)

Optimal control problems constitute an interesting case of optimization problems based on partial differential equations (PDEs). There is a rich history of work in this area, beginning with the development of the calculus of variations and work in the control of ordinary differential equations. More recently, researchers have begun to investigate the control of PDEs. Instances of these types of problems abound in applications. Thus, the development of efficient numerical methods for the solution of these problems has also been the subject of significant recent research.

In this project, we have examined a particular instance of optimal control problems where multiple controls seek to force behavior of multiple "targets" simultaneously. These problems belong to the class of problems called *multicriteria optimization*. There is no unique mathematical formulation of these types of problems. In practice they are often solved by minimizing a weighted sum of the deviations from the targets with the weights corresponding to preferences. Another formulation sometimes referred to as goal programming insists that the preferred targets must be satisfied to within certain tolerances and the others are reduced as much as possible within these constraints.

Both of these approaches involve the choice of a set of weights or tolerances for which there can be little theoretical guidance. Yet another approach, called bilevel optimization, nests optimization problems, optimizing with respect to one target and using the solutions as parameters in a larger problem.

Motivated by specific applications, especially those arising in optimal well placement in reservoir engineering and optimal control of combustion systems, we investigated means of

formulating a special class of optimal control problems in which there is a partition of the targets into categories of increasing importance. For example, if one might be concerned with contaminant transport (in subsurface flow) being controlled so as to reduce the contamination in two specific areas very differently. Likewise, one can imagine controlling a combustion system where a portion of the domain must be permitted to change less than the rest of the domain. This is the case controlling weapons systems – in the event of fire there are portions of the computational domain in which fire control systems must at all costs keep temperature controlled even at the risk of the remaining domain. These problems are of obvious interest to both the Department of Energy and the Department of Defense both of which have enthusiastically encouraged this work.

NIST Strategic Focus Areas. *Existing Industries*: Petroleum; Combustion. *Emerging Industries:* Information and Knowledge Management.

Mathematical Methods for Austine-Martensite

Anthony Kearsley
Geoffrey McFadden
Luis Melara
Daniel Anderson (George Mason University)

Materials play an important and vital role in today's society, with uses ranging from shields in orbiting satellites, to tubular stents inserted into blood vessels to prevent clogging. Their applications continue to multiply as we learn more about their properties and behavior. Many important material properties are attributed to the structures observed at the microscopic level. These microstructures are the result of thermal or mechanical deformation of the alloy and are directly related to the total stored energy in the material. This is important because energy can be used to measure the work a material can do, consequently, providing us with insight into the material's behavior. In recent decades, mathematics has been applied to the simulation of microstructures by approximating the amount of total stored energy in a material. The energy is represented using partial differential equations, with microstructures simulated by minimizing the total stored energy.

The focus of our work has been to simulate the microstructures that occur in Martensite, near the Austenite-Martensite interface. An example of this interface is shown in the Figure below. The domain on the top is Austenite. It has a simpler crystallographic structure than Martensite, the region on the bottom. Martensite is characterized by two variants that are differentiated by the red and white alternating colors. This variant of Martensite is commonly found in shape memory alloys and results in its "memory" behavior. We are developing new computational methods to simulate the microstuctures in Martensite.
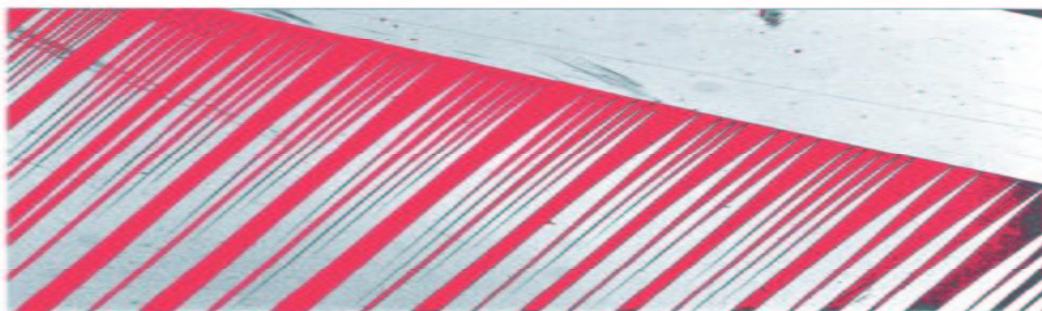
Our approach has been to use a new finite element space and new optimization techniques for minimizing the total stored energy near the Austenite-Martensite interface. We simulate the microstructures observed in the Martensite region by searching for the function that minimizes the total material stored energy. There are a variety of energy functions that have been suggested in the vast literature of this research area. Based on conversations with experimentalists we have developed our own energy functions and are currently using them to predict microstructure. We have implemented a novel conforming finite element scheme and are in the process of comparing results to standard discretizations.

In addition to a new finite element space, we use new optimization techniques for minimizing the total stored energy. We have successfully employed a new hybrid limited memory quasi-Newton method for minimizing the energy (discretized using our new conformal finite element method or more traditional Galerkin finite element methods). Use of this requires an expensive calculation of the gradient of the energy in either form.

This work is of interest to both the Departments of Energy and Defense. We have been invited to give a series of talks at two government laboratories in California.

NIST Strategic Focus Areas. *Existing Industries*: Materials processing. *Emerging Industries:* Information and Knowledge Management.

Modeling and Computational Techniques for Bioinformatics Based Data Mining

Fern Hunt
Anthony Kearsley
Agnes O'Gallagher
Antti Pesonen (VTT, Helsinki, Finland)

http://math.nist.gov/~FHunt/GenPatterns/

As the sequencing of the human genome and other genomes of medical and commercial importance near completion, attention is turning to the interpretation of the massive amounts of genome sequence data that have been gathered and that are projected to accumulate in the near future. Understanding the biological function of the genes that have been discovered and the mechanisms that coordinate and regulate them is critical for further scientific progress and applications. To accomplish this, high throughput experimentation is needed, but it very clearly must be supplemented, and in some cases guided, by the results of knowledge obtained from mining the sequence data itself. Biologists infer the biological function of an unknown sequence by finding a nearby (evolutionarily speaking) sequence in the existing sequence database that has a known biological function. The searching techniques are based on the comparison and alignment algorithms of bioinformatics. The purpose of this project is to develop generic bioinformatics algorithms that can utilize large-scale and fast computational methods to deal with the comparison and alignment of a large number of long biological sequences. Almost all methods currently in use are based on an optimization problem that is solved using the methods of dynamic programming. Starting from a Markov decision theory approach, we instead solve a linear programming problem. Our goal is to avoid the exponential increase in computation as the number of sequences increases and to exploit the high performance algorithms available for solving this type of problem. Potential applications for this work are in drug discovery where large numbers of sequence sites must be evaluated for use as drug targets, in the development of gene based therapies, and the characterization of biological threats.

During fiscal year 2002, we developed software implementing the algorithm. The software comes in several parts: firstly, *training* which receives sequence alignment data as input, computes

parameter values of the Markov decision model, and outputs the matrix and vector of the constraints in the linear programming problem, secondly the *linear programming* part where the data are put into MPS format and PCx is used to obtain the solution, and finally *alignment*, where the solution to the linear programming problem is used to produce a suggested alignment of given input sequences. A draft package containing all these parts has been completed, and testing and comparison with other alignment methods is underway.

There were several theoretical developments this year as well. Since our approach is based on a random process, the alignment cost that is minimized is an expected or average cost. In the course of investigating the linear programming problem that is dual to the original we were able to show that the suggested alignment produces an alignment cost that is asymptotically pathwise optimal with probability one. In this sense then, the average cost can be replaced by the actual cost. We were also able to investigate the stability or robustness of the algorithm to changes in the cost functions and the parameters of the model itself. All these results will be detailed in a paper that is in preparation.

Our future work will center on developing a package that can be used to test data and linear programming algorithms. We also want to develop effective but computationally efficient training algorithms for obtaining model parameters.

In related work, GenPatterns, the software developed by Fern Hunt and Antti Pesonen for analyzing patterns in DNA sequences, is being used in a course in Computational Biology at Capital University. The software is being tested for use as a standard module for college courses in an effort that is being funded by the National Science Foundation, the W.M. Keck Foundation and the Batelle Institute.

NIST Strategic Focus Areas. *Emerging Industries:* Health Care (clinical diagnostics); Information and Knowledge Management (dynamic data infrastructure).

A Novel Algorithm for Solving Binary Regression Problems

Anthony Kearsley

The need to rapidly identify patterns in large data sets surfaces regularly in many applications of communications. This task is complicated tremendously in the presence of corrupted or noisy data. An interesting subclass of these problems arises when examining large binary data sets corrupted by Bernoulli-like processes, the type of data encountered when analyzing raw wireless communications. Typically one would like to rapidly scan these large data sets for prescribed recursion patterns significantly smaller (e.g., many orders of magnitude) than the size of the original raw data set. This regression has a much higher probability of success (identifying targeted recursions) if the scanning algorithm is capable of anticipating the effect of the noise and estimating the amount of corruption present.

A myriad of numerical approaches for solving this problem have been suggested and are currently under investigation. A vast majority of these pattern-recognizing algorithms depend on discrete and combinatorial heuristics for successively relaxing and enforcing binary constraints in an attempt to enumerate and eliminate poorly matched patterns and, in turn, increase the probability of locating well-matched patterns. This is, in large part, due to the fact that the relaxed problem is a linear programming problem and linear programming is solvable in polynomial time, if the dimension is fixed. These so-called *branch-and-bound* and *branch-and-cut* methods have been and still are the mainstay of raw voice data mining algorithms. However, these methods begin to fall short as the amount of raw voice data to be processed continues to grow drastically.

In this project we have been examining the possibility of using higher-order approximation methods for pattern discernment. In particular, one could allow *non-linear* definitions of wellness-of-

fit that incorporate noise models (among other factors) into the pattern-fitting objective. This idea has advantages over currently employed methods. The newer methods have the potential to be faster than currently employed ones. They may be able to recognize patterns that are simply impossible for current methods to locate, and in the unlikely event that one is searching for a particular pattern then the algorithm can take that into account. Of course, very little is known about the performance of these newer methods — there is no theory predicting performance, there are no complexity estimates addressing run-time or problem difficulty, and there is no computer implementation from which intuition about numerical behavior can be formed.

Our algorithmic approach was, at least in part, motivated by a query from the Department of Defense (DoD) which was interested in the performance of these nonlinear techniques, specifically our formulation, on some large scale structured binary regression problems. We have developed our algorithm and are currently working on a large-scale implementation.

NIST Strategic Focus Areas. *Homeland Security. Emerging Industries:* Information and Knowledge Management (dynamic data infrastructure).

Optimal Signal Sets for Non-Gaussian Detectors

Anthony Kearsley

In order to send and receive communications through a wireless device, such as a modem, one must be able to choose signal sets that are relatively resilient to noise corruption. Optimal signal sets are those that minimize the probability that a wireless communication as is incorrectly received a result of noise. One major difficulty with designing such a signal set is that one is never certain of what type of noise will be present in the system. Clearly the noise encountered in urban settings will be quite different from that of a rural setting. Geographic conditions like trees, water, and even weather conditions like wind and lightning can play a major role in producing noise in a wireless communications device.

Substantial effort is exerted yearly by US industry to build noise models for wireless devices. These models can then be used to construct signal sets that are optimal with respect to various noise estimates. We have been developing an algorithm for doing exactly this, leaving the estimation of the noise model general. To date, our algorithm has yielded better solutions that any off-the-shelf software package, and continues to be used by industry. We have distributed 216 copies of the code to researchers in government, universities and US companies (including Hewlett Packard, Lucent, and AT&T).

Although this project is nearing completion, we are still working on a few aspects of the algorithm that can be improved. Our industrial colleagues have suggested a numerical comparison between some slightly varying noise models, as well as a sensitivity study. If one changes certain aspects of the noise model, only slightly, can one say anything about the resulting optimal signal set? We plan to examine this and other remaining open questions in our future work.

NIST Strategic Focus Areas. *Established Industries:* Information Technology.

Monte Carlo Methods for Discrete Problems

Isabel Beichl
Dianne P. O'Leary
April Andreas
Francis Sullivan (IDA Center for Computing Sciences)

Combinatorial problems arise in a wide variety of applications, from nanotechnology to computer network analysis. Fundamental models in these fields are often based on quantities that are extremely difficult (i.e., exponentially hard) to compute. We have devised methods to compute such quantities approximately (with known error bars) using Monte Carlo methods. Traditional Monte Carlo methods can be slow to converge, but we have made progress in significantly speeding up these computations using importance sampling. In the past few years I. Beichl, D. O'Leary and F. Sullivan have made progress in evaluating the partition function for describing the probability distribution of states of a system. In the past year I. Beichl and a student, A. Andreas, developed a new algorithm to estimate the work of the standard methods for integer partitioning. Number partitioning is the problem of dividing a set of numbers into subsets so that the sums of the subsets are as close as possible. Another version is to have the sums and the cardinality of the two sets close to being equal. This is a provably hard problem that is important in multiprocessor load balancing and scheduling. Our new method can estimate quickly the amount of processing a given instance of either version of partitioning will require. Such a capability is useful in deciding how much computing effort to put into load balancing, for example.

In the past year, a stratified sampling method for estimating the number of independent sets of a graph was also developed. Stratified sampling is an extension of importance sampling and an independent set of a graph is a subset of the nodes where there are no connections between any of the elements of the subset. This is a problem with applications in thermodynamics and in descriptions of networks.

A novel combination of two fundamental ideas was made in order to estimate probability distributions of certain physical systems. The two ideas were aggregation in Markov chain Monte Carlo and importance sampling. Our plans are to apply this to estimating the reliability of computer network by approximating the number of cut sets.

NIST Strategic Focus Areas. *Emerging Industries:* Nanotechnology*;* Information and Knowledge Management.

Surface Reconstruction by Tetrahedralization

Javier Bernal
Zachary H. Levine (NIST PL)

Many applications in computer-aided design involve reconstructing the surface of a 3D object from scattered sample points on the surface of the object. In general the surface reconstruction problem asks for a triangular mesh of the sample points that faithfully represents the surface. Very often a tetrahedralization such as the Delaunay tetrahedralization of the sample points is first computed and from it the triangular mesh is extracted. However in practice this triangular mesh is usually topologically incorrect, i. e. at least one vertex in the mesh can be identified for which the union of the triangles in the mesh that have the vertex in common either does not contain the vertex in its relative interior or is not homeomorphic to a 2D closed disk. This is usually a consequence of noise, undersampling, and sharp surface features. Research is currently under way for developing a method

for reconstructing surfaces by tetrahedralization that produces topologically correct triangular meshes while taking into account noise, undersampling, and sharp surface features.

Zahary Levine of the NIST Physics Lab is currently using our techniques in tomographic applications. Tomography, a diagnostic technique, was first implemented with X-rays. In the case of X-rays, a law called Beer's Law is used to relate the observed intensity to a projection of a material property. Tomographic algorithms are then used to reconstruct the projections. Tomography is also performed using scanning transmission electron microscopy (STEM). In the case of STEM, multiple scattering can occur, leading to deviations from Beer's Law. Nevertheless, for thin samples, even for STEM, Beer's Law is a good approximation. Tetrahedralizations are currently being used for modeling tomographic reconstructions using STEM. The aim of this work is to demonstrate STEM-based tomography for thick samples.

NIST Strategic Focus Areas. *Emerging Industries:* Information and Knowledge Management.

Parameter Selections for Constrained Estimates for Ill-Posed Problems

Bert W. Rust
Dianne P. O'Leary

Many measurement processes can be modeled by linear first kind integral equations. Discretizing them almost always produces an ill-conditioned linear regression problem for which the standard least squares estimate is hopelessly unstable. Stable estimates are usually gotten by one of two general methods: (1) truncating a singular value decomposition (SVD) of the model matrix to remove components of the solution thought to be corrupted by measurement errors, or (2) appending additional linear equations to the model which describe a priori constraints that the unknown solution must satisfy. The latter technique, which is usually called regularization, requires the user to choose a free parameter which controls the relative weighting of the measurement equations and the additional constraint equations. The truncated SVD strategy also requires the choice of a free parameter, namely the threshold for the truncation. The only quantitative diagnostics available for making such choices are the residuals for the resulting estimate. The goal of this project is to develop methods for using those residuals to make the selections and to compare the results when the selections are applied to the various estimation techniques.

The guiding principle for this work is to seek estimates that produce residuals that are as much like the measurement errors as can be determined. For example, if the measurement errors are normally distributed, and the regression model is normalized to reduce the variance-covariance matrix to an identity matrix, then the residuals should be independently, identically distributed with a standard normal distribution, and the sum of squared residuals (SSR) should be a sample from a chi-square distribution with m degrees of freedom, where m is the row dimension of the model matrix. This fact gives a statistical variant of the "discrepancy principle" which was originally developed for problems with bounded errors, but which has thus far not been very useful because real measurement errors are never bounded. The fact that the mean value of the above chi-square distribution is m suggests that one should seek an estimate that gives an SSR as close as possible to m. Figure 1 compares the root mean square average error in the elements of the estimates obtained when this discrepancy principle is applied to the conventional truncated SVD method and to Rust's new truncated singular component method for 1000 realizations of the measurement error vector in a 150 x 121 variant of the well known Phillips problem. The x* represents the true solution and the plain x represents the estimate. The new method produced a smaller rms average error for 954 of the 1000 problems. Figure 2 shows a similar comparison for the new method and the simplest variant of Tikhonov regularization.
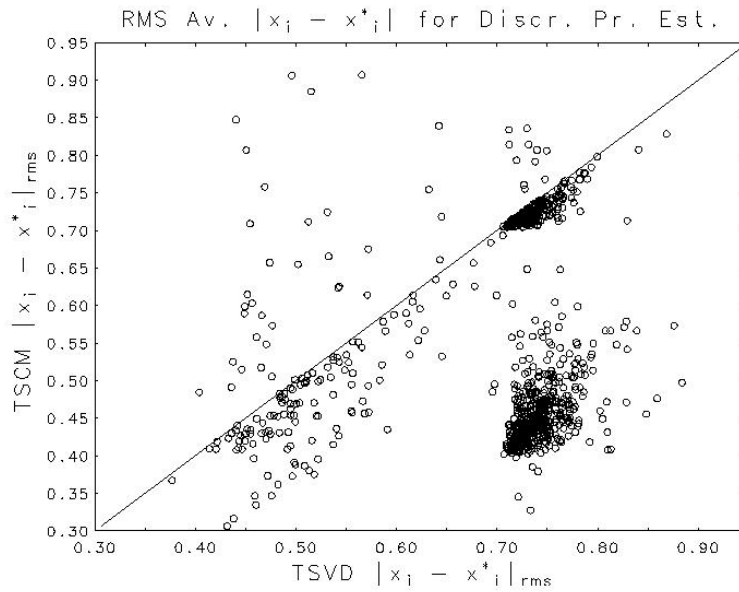
RMS Av. $|x_i - x^*_i|$ for Discr. Pr. Est.

Figure 1

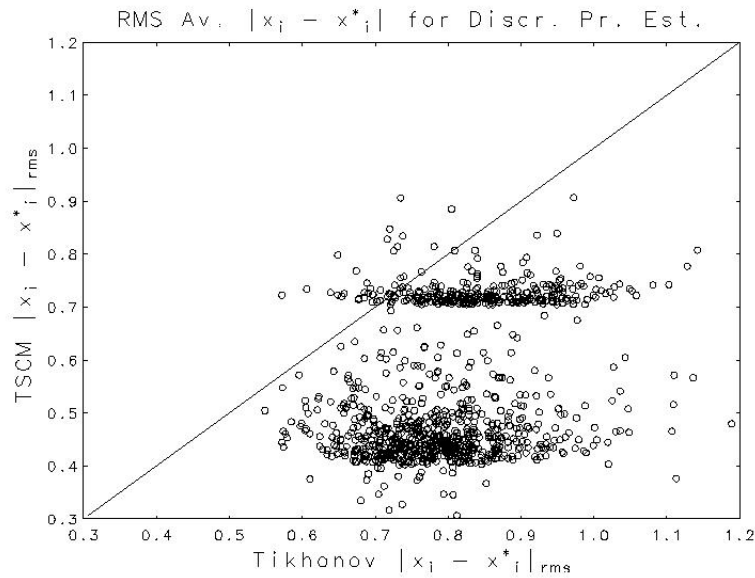RMS Av. $|x_i - x^*_i|$ for Discr. Pr. Est.

Figure 2

Future work on this project will not only compare other methods using the discrepancy principle, but also make similar comparisons using other parameter selection criteria including generalized cross validation, the L-curve method, the length of the cumulative periodogram of the residuals, and other measures of the randomness of the residuals.

NIST Strategic Focus Areas. *Emerging Industries:* Information and Knowledge Management.
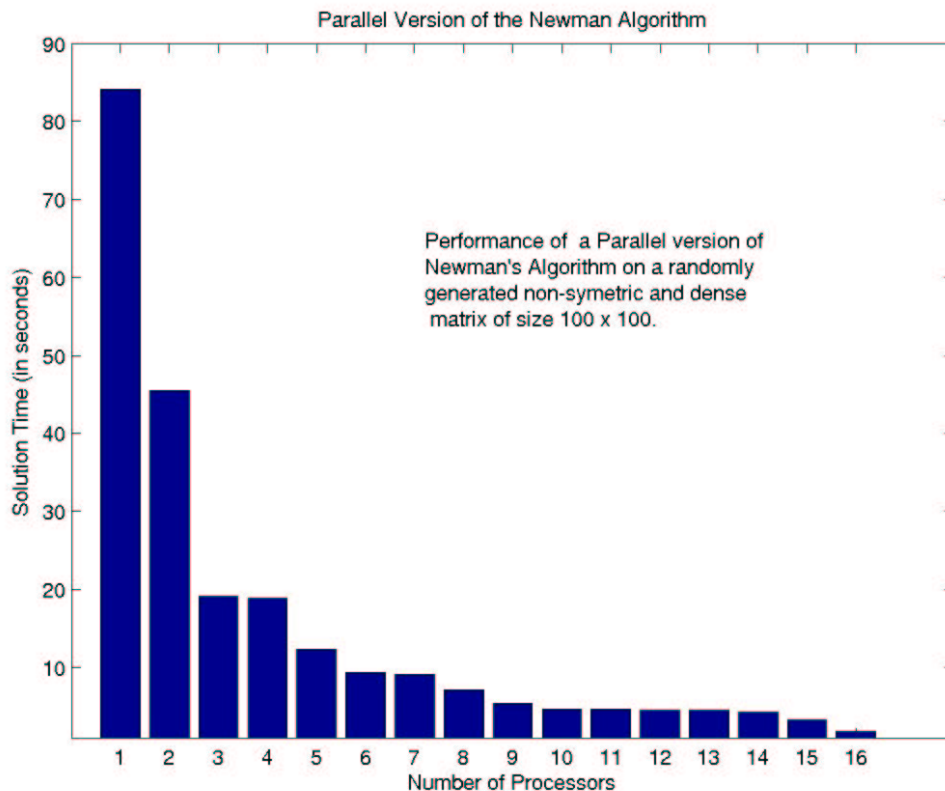
Exact Parallel Methods for Linear Systems

Anthony Kearsley
Eric Baer (Carnegie Mellon University)

In 1969 Morris Newman (of NBS) published a paper describing an algorithm for solving integral linear systems exactly using modular arithmetic. The fundamental idea involved solving the linear system modulo a collection of (relatively) prime numbers and storing the residues. Using the Chinese Remainder Theorem (CRT) one can then construct a unique and exact solution to the linear system. The arithmetic needed to apply this algorithm is purely integer, save the final construction of the solution where some finite precision scheme must be selected to represent the solution.

This algorithm requires each integral linear system to be solved multiple times modulo a potentially large collection of numbers. These solutions are independent from each other and can, therefore, be produced separately on different computer processors (i.e., one can construct an embarrassingly parallel version of this algorithm). We have developed a parallel extension of Newman's algorithm and tested it on some small-medium sized linear systems. The algorithm appears to work well (see figure).

It is interesting to note that this algorithm fails for reasons very different from most currently employed solution techniques, such as iterative methods and matrix decompositions. The classical condition number, the ratio of largest to smallest singular value, does not predict performance of this algorithm. Instead, it performs well or poorly as a function of how quickly one identifies a unique element using a priori CRT bounds.

Parallel Version of the Newman Algorithm

Performance of a Parallel version of Newman's Algorithm on a randomly generated non-symetric and dense matrix of size 100 x 100.

This algorithm does not require any particular matrix structure (e.g., symmetry or sparsity), and hence there are many applications in coding and cryptographic applications. Those currently interested in this work include branches of the Department of Defense (DoD) and smaller contract companies working on integer based security schemes. We are finalizing a manuscript describing our algorithm together with numerical results.

NIST Strategic Focus Areas. *Emerging Industries:* Information and Knowledge Management.