LBNL-46959

ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Optimization and Geophysical Inverse Problems

J. Barhen, J.G. Berryman, L. Borcea, J. Dennis, C. de Groot-Hedlin, F. Gilbert, P. Gill, M. Heinkenschloss, L. Johnson, T. McEvilly, J. Moré, G. Newman, D. Oldenburg, P. Parker, B. Porto, M. Sen, V. Torczon, D. Vasco, and N.B. Woodward

Earth Sciences Division

October 2000

Optimization and Geophysical Inverse Problems

Report of a Workshop at
San Jose, California, February 5-6, 1999

Sponsored by
U.S. Department of Energy, Office of Basic Energy Sciences
Division of Chemical Sciences, Geosciences and Biosciences

Participants

Jacob Barhen	Tom McEvilly
James G. Berryman	Jorge Moré
Liliana Borcea	Gregory Newman
John Dennis	Douglas Oldenburg
Catherine de Groot-Hedlin	Paul Parker
Freeman Gilbert	Bill Porto
Philip Gill	Mrinal Sen
Matthias Heinkenschloss	Virginia Torczon
Lane Johnson	Don Vasco
	Nicholas B. Woodward

This report was supported by the Director, Office of Energy Research, Office of Basic Energy Science, Division of Chemical Sciences, Geosciences and Biosciences of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Preface

This report summarizes the results of a two-day meeting that was held during February 1999 in San Jose, California, for the purpose of exploring the interdisciplinary area between geophysical inverse problems and mathematical optimization methods. The meeting was sponsored by the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Science, U.S. Department of Energy. The participants in the meeting are listed below:

Jacob Barhen
Center for Engineering Science Advanced Research
Oak Ridge National Laboratory
1000 Bethel Valley Road
Building 6010, MS-6355
Oak Ridge, TN 37831-6355
barhenj@ornl.gov

James G. Berryman
Lawrence Livermore National Laboratory
P. O. Box 808, L-200
Livermore, CA 94551-9900
berryman1@llnl.gov

Liliana Borcea
Department of Computational and Applied Mathematics
Rice University
6100 Main Street
Houston, TX 77005-1892
borcea@caam.rice.edu

John Dennis
Department of Computational and Applied Mathematics
Rice University
6100 Main Street
Houston, TX 77005-1892
dennis@rice.edu

Tom McEvilly
Department of Geology and Geophysics
307 McCone Hall
University of California
Berkeley, CA 94720
mcevilly@ccs.lbl.gov

Jorge Moré
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439
more@mcs.anl.gov

Gregory Newman
Sandia National Laboratories
Geophysical Technology Department
Organization 6116
1515 Eubank SE
Albuquerque, NM 87185-0750
ganewma@sandia.gov

Douglas Oldenburg
UBC-Geophysical Inversion Facility
Department of Earth and Ocean Sciences
Geophysics Building
University of British Columbia
Vancouver, BC, V6T 1Z4, Canada
doug@geop.ubc.ca

Catherine de Groot-Hedlin
Institute of Geophysics and Planetary Physics
Scripps Institution of Oceanography
Mail Code 0225
9500 Gilman Drive
La Jolla, CA 92093-0225
cdh@eos.ucsd.edu

Freeman Gilbert
Institute of Geophysics and Planetary Physics
Scripps Institution of Oceanography
Mail Code 0225
9500 Gilman Drive
La Jolla, CA 92093-0225
fgilbert@ucsd.edu

Philip Gill
Department of Mathematics
University of California
9500 Gilman Drive, Dept. 0112
La Jolla, CA 92093-0112
pgill@ucsd.edu

Matthias Heinkenschloss
Department of Computational and
Applied Mathematics, MS134
Rice University
6100 Main Street
Houston, TX 77005-1892
heinken@caam.rice.edu

Lane Johnson
Department of Geology and Geophysics
307 McCone Hall
University of California
Berkeley, CA 94720
lrj@ccs.lbl.gov

Paul Parker
10323 Woodbine St.
Apt. 410
Los Angeles, CA 90034
parker@itginc.com

Bill Porto
Natural Selection, Inc.
La Jolla, CA 92037
bporto@natural-selection.com

Mrinal Sen
Institute for Geophysics
University of Texas
4412 Spicewood Spgs., Bldg. 600
Austin, TX 78759-8500
mrinal@bullen.ig.utexas.edu

Virginia Torczon
Department of Computer Science
Room 126
McGlothlin-Street Hill
College of William and Mary
Williamsburg, VA 23187-8795
va@cs.wm.edu

Don Vasco
Lawrence Berkeley National Laboratory
Earth Sciences Division, Building 90
1 Cyclotron Road
Berkeley, CA 94720
dwvasco@lbl.gov

Nicholas B. Woodward
Division of Chemical Sciences,
Geosciences and Biosciences
Office of Basic Energy Science
U.S. Department of Energy
nick.woodward@science.doe.gov

Optimization and Geophysical Inverse Problems

1 Introduction

A fundamental part of geophysics is to make inferences about the interior of the earth on the basis of data collected at or near the surface of the earth. In almost all cases these measured data are only indirectly related to the properties of the earth that are of interest, so an inverse problem must be solved in order to obtain estimates of the physical properties within the earth.

In February of 1999 the U.S. Department of Energy sponsored a workshop that was intended to examine the methods currently being used to solve geophysical inverse problems and to consider what new approaches should be explored in the future. The interdisciplinary area between inverse problems in geophysics and optimization methods in mathematics was specifically targeted as one where an interchange of ideas was likely to be fruitful. Thus about half of the participants were actively involved in solving geophysical inverse problems and about half were actively involved in research on general optimization methods. This report presents some of the topics that were explored at the workshop and the conclusions that were reached.

In general, the objective of a geophysical inverse problem is to find an earth model, described by a set of physical parameters, that is consistent with the observational data. It is usually assumed that the forward problem, that of calculating simulated data for an earth model, is well enough understood so that reasonably accurate synthetic data can be generated for an arbitrary model. The inverse problem is then posed as an optimization problem, where the function to be optimized is variously called the objective function, misfit function, or fitness function. The objective function is typically some measure of the difference between observational data and synthetic data calculated for a trial model. However, because of incomplete and inaccurate data, the objective function often incorporates some additional form of regularization, such as a measure of smoothness or distance from a prior model. Various other constraints may also be imposed upon the process.

Inverse problems are not restricted to geophysics, but can be found in a wide variety of disciplines where inferences must be made on the basis of indirect measurements. For instance, most imaging problems, whether in the field of medicine or non-destructive evaluation, require the solution of an inverse problem. In this report, however, the examples used for illustration are taken exclusively from the field of geophysics. The generalization of these examples to other disciplines should be straightforward, as all are based on standard second-order partial differential equations of physics. In fact, sometimes the non–geophysical inverse problems are significantly easier to treat (as in medical imaging) because the limitations on data collection, and in particular on multiple views, are not so severe as they generally are in geophysics.

This report begins with an introduction to geophysical inverse problems by briefly describing four canonical problems that are typical of those commonly encountered in geophysics. Next the connection with optimization methods is made by presenting a general formulation of geophysical inverse problems. This leads into the main subject of this report, a discussion of methods for solving such problems with an emphasis upon newer approaches that have not yet become prominent in geophysics. A separate section is devoted to a subject that is not encountered in all optimization problems but is particularly important in geophysics, the need for a careful appraisal of the results in terms of their resolution and uncertainty. The impact on geophysical inverse problems of continuously improving computational resources is then discussed. The main results are then brought together in a final summary and conclusions section.

2 Canonical Problems

To provide specific examples that can be used to illustrate the basic properties of geophysical inverse problems and methods of solution, a small set of canonical problems have been selected. The basic equations are stated and formulations as both a forward problem and an inverse problem are given. Examples that involve elliptic, parabolic, and hyperbolic differential equations are all included. Also, both linear and non-linear inverse problems are represented.

Problem 1 - Gravity

The first problem, which involves the solution of an elliptic second-order differential equation, represents a type of geophysical inverse problem that contains a fundamental nonuniqueness. Consider the force of gravity measured at any point within or on the surface of the earth. Because of variations in density within the earth, this gravitational acceleration deviates from the value expected for a radially symmetric earth. Newton's theory of gravitation says that

$$\Delta g(\mathbf{x}) = G \int_D \frac{(\mathbf{x} - \xi) \cdot \hat{\mathbf{x}}_3}{|\mathbf{x} - \xi|^3} \, \Delta\rho(\xi) \, d\xi^3 \,, \tag{1}$$

where

$\Delta g(\mathbf{x})$ is the anomaly in the vertical $\hat{\mathbf{x}}_3$ component of gravitational acceleration at the location $\mathbf{x}$ in units of $m\ s^{-2}$,

G is the gravitational constant $6.67 \cdot 10^{-11}$ *newton* $m^2\ kg^{-2}$,

D is the spatial domain of interest, the half space $x_3 \geq 0$,

$\hat{\mathbf{x}}_3$ is the vertical unit vector normal to the surface of the half space,

$\Delta\rho(\xi)$ is the anomaly in density within D in units of $kg\ m^{-3}$.

Forward problem:

Given:

- $\Delta\rho(\xi)$ at all points ξ within D,

Determine:

- $\Delta g(\mathbf{x})$ at arbitrary points $\mathbf{x}$ either within or outside D.

Inverse problem:

Given:

- $\Delta g(\mathbf{x})$ at the specified N points $\mathbf{x}_n$ $(n = 1, \ldots, N)$ located either within or outside D,

Determine:

- $\Delta\rho(\xi)$ at all points ξ within D.

Problem 2 - Electrical Sounding

The second problem represents the class of parabolic second-order differential equations that are quite common in the study of thermal and electrical properties of the earth (Oldenburg, 1979). Consider an earth in which the electrical conductivity $\sigma(z)$ varies only as a function of depth z. Then Maxwell's equations in the frequency domain reduce to the ordinary differential equation

$$\frac{d^2}{dz^2}E(z,\omega) + i\omega\mu\sigma(z)E(z,\omega) = 0 \ , \qquad (2)$$

where

> z is depth below the surface of the earth in units of m,
>
> ω is angular frequency in units of *radians* s^{-1},
>
> $E(z,\omega)$ is the electrical field strength in units of *volt* m^{-1},
>
> μ is the magnetic permeability in units of *henry* m^{-1},
>
> $\sigma(z)$ is the electrical conductivity in units of ohm^{-1} m^{-1}.

It has been assumed here that the frequency ω is low enough so that displacement currents can be ignored in comparison to conduction currents. The boundary conditions are that

> $\frac{d}{dz}E(0,\omega) = 0,$
>
> $E(Z,\omega) = 0$ where Z is large compared to the skin depth $d = \sqrt{\frac{2}{\mu\sigma\omega}}$.

Forward problem:

Given:

- $\sigma(z)$ at all depths $z \geq 0$,

Determine:

- $E(z,\omega)$ at arbitrary depth z either within or on the surface of the half space and for arbitrary frequency ω.

Inverse problem:

Given:

- $E(0,\omega_n)$ at the specified frequencies ω_n $(n = 1,\ldots,N)$,

Determine:

- $\sigma(z)$ for all depths $0 \leq z \leq Z$.

Notes

Because σ can vary over extremely large ranges in the earth, it is common to write $\sigma(z) = e^{m(z)}$ and let $m(z)$ represent the model parameter. The differential equation then becomes

$$\frac{d^2}{dz^2}E(z,\omega) + i\omega\mu e^{m(z)}E(z,\omega) = 0 \ , \tag{3}$$

and the inverse problem is to solve for $m(z)$ at all depths $0 \leq z \leq Z$.

Problem 3 - Seismic Sounding

The third problem is one version of the type of hyperbolic second-order differential equations that are common in seismology. Consider an earth in which the seismic velocities and density vary only as a function of depth z. Then the elastodynamic equations of motion in the frequency-slowness domain can be put in the form

$$\frac{d}{dz}\mathbf{v}(\omega,p,z) = \omega \ \mathbf{A}(p,z)\,\mathbf{v}(\omega,p,z) + \mathbf{f}(\omega,p) \ , \tag{4}$$

where

ω is angular frequency in units of *radians* s^{-1},

p is horizontal slowness of the wave in units of $s\ m^{-1}$,

z is depth below the surface of the earth in units of m,

$\mathbf{v}(\omega,p,z)$ is the displacement-stress vector $\mathbf{v} = [u_z, u_x, \tau_{zz}/\omega, \tau_{zx}/\omega]^T$

$\mathbf{A}(p,z)$ is a matrix depending upon the material properties of the medium,

$$\mathbf{A}(p,z) = \begin{bmatrix} 0 & p\frac{\lambda(z)}{\lambda(z)+2\mu(z)} & \frac{1}{\lambda(z)+2\mu(z)} & 0 \\ -p & 0 & 0 & \frac{1}{\mu(z)} \\ -\rho(z) & 0 & 0 & p \\ 0 & p^2\gamma(z)-\rho(z) & -p\frac{\lambda(z)}{\lambda(z)+2\mu(z)} & 0 \end{bmatrix} , \tag{5}$$

where

$$\gamma(z) = 4\mu(z)\frac{\lambda(z)+\mu(z)}{\lambda(z)+2\mu(z)} \ ,$$

$\lambda(z)$ and $\mu(z)$ are elastic constants in units of $kg\ m^{-1}\ s^{-2}$,

$\rho(z)$ is the density in units of $kg\ m^{-3}$,

$\mathbf{f}(\omega,p)$ is the source vector that acts at the depth z_s,

$$\mathbf{f}(\omega,p) = \begin{bmatrix} 0 \\ 0 \\ f_z(\omega,p)/\omega \\ f_x(\omega,p)/\omega \end{bmatrix} . \tag{6}$$

The boundary conditions are that

u_z and u_x are continuous everywhere,

τ_{zz} and τ_{zx} are continuous everywhere except at the depth z_s where $\mathbf{f}$ is non-zero,

$\tau_{zz} = \tau_{zx} = 0$ at $z = 0$,

$\mathbf{v}$ consists of downward propagating waves as $z \to \infty$.

Forward problem:

Given:

- $\lambda(z)$, $\mu(z)$, and $\rho(z)$ at all depths $z \geq 0$,
- $\mathbf{f}(\omega, p)$ at z_s,

Determine:

- $\mathbf{v}(\omega, p, z)$ at arbitrary depths z either within or on the surface of the half space and for arbitrary frequency ω and arbitrary slowness p.

Inverse problem:

Given:

- $u_z(\omega_n, p_m, 0)$ and $u_x(\omega_n, p_m, 0)$ at $z = 0$ for the specified frequencies ω_n $(n = 1, \ldots, N)$ and for the specified slownesses p_m $(m = 1, \ldots, M)$,

Determine:

- $\mathbf{f}(p, \omega)$ at the specified source depth z_s,
- $\lambda(z)$, $\mu(z)$ and $\rho(z)$ for all depths $z \geq 0$.

Notes

The observational data are actually acquired in the time-space domain and are of the form $u_z(t, x, z)$ and $u_x(t, x, z)$. A Legendre transformation of the form $\tau = t - px$ is then applied to obtain data in the tau-slowness domain, $u_z(\tau, p, z)$ and $u_x(\tau, p, z)$. Finally, a Fourier transformation of the data leads to $u_z(\omega, p, z)$ and $u_x(\omega, p, z)$.

Problem 4 - Travel Time Tomography

The fourth problem represents the general class of tomography inversions that have become the preferred approach to the study of three-dimensional earth structure using seismic body waves. The problem contains a nonlinearity that enters indirectly through the dependence upon the ray path. A general discussion of one version of this problem can be found in Trampert (1998). The travel time along a ray between points $\mathbf{x}_s$ and $\mathbf{x}_r$ is given by

$$t(\mathbf{x}_r, \mathbf{x}_s) = \int_{\Sigma(\mathbf{x}_r, \mathbf{x}_s)} \frac{d\sigma}{v(\mathbf{x})}, \tag{7}$$

where

t is the travel time in units of s,

$\mathbf{x}_s$ is the souce point where the ray originates,

$\mathbf{x}_r$ is the receiver point where the ray terminates,

$\Sigma(\mathbf{x}_r, \mathbf{x}_s)$ is the ray path between $\mathbf{x}_s$ and $\mathbf{x}_r$ that is obtained by finding a solution to the two-point boundary value problem for the equation

$$\frac{d^2}{d\sigma^2}\mathbf{x}(\sigma) + v(\mathbf{x})\frac{d}{d\sigma}[\frac{1}{v(\mathbf{x})}]\frac{d}{d\sigma}\mathbf{x}(\sigma) = v(\mathbf{x})\nabla[\frac{1}{v(\mathbf{x})}] , \tag{8}$$

σ is distance measured along the ray path,

$v(\mathbf{x})$ is the velocity in units of $m\ s^{-1}$.

Forward problem:

Given:

- $v(\mathbf{x})$ in a domain D,

Determine:

- $t(\mathbf{x}_r, \mathbf{x}_s)$ for arbitrary values of $\mathbf{x}_s$ and $\mathbf{x}_r$ on the boundary of D.

Inverse problem:

Given:

- $t_{mn}(\mathbf{x}_m, \mathbf{x}_n)$ for specified values of $\mathbf{x}_s = \mathbf{x}_n$ ($n = 1, \ldots, N$) and specified values of $\mathbf{x}_r = \mathbf{x}_m$ ($m = 1, \ldots, M$) on the boundary of D,

Determine:

- $v(\mathbf{x})$ within D.

Notes

A more convenient form of the ray equations is the system

$$\frac{d}{d\sigma}\mathbf{x}(\sigma) = v(\mathbf{x})\mathbf{p}(\sigma) ,$$

$$\frac{d}{d\sigma}\mathbf{p}(\sigma) = \nabla[\frac{1}{v(\mathbf{x})}] ,$$

$$\mathbf{p}(\sigma) \cdot \mathbf{p}(\sigma) = \frac{1}{v^2(\mathbf{x})} .$$

3 Formulation of Geophysical Inverse Problems

In preparing to solve a geophysical inverse problem, a number of critical decisions have to be made regarding the manner in which the problem is formulated. The success or failure of the entire process can often be traced back to these initial choices. First, methods of parameterizing both the data space and the model space have to be selected. Next, the solution of the inverse problem is regularized by imposing constraints upon the model space, which may involve choosing criteria for measuring the "goodness" of a model. The net result of this process in practically all cases is a problem in mathematical optimization.

3.1 Parameterization

The parameterization of the data space is determined in most cases by what is feasible in the data collection experiment. The field that is being measured is generally continuous in space and time, but the data are almost always given in a discrete form, representing sampled values at discrete points in space and time. In principle, this is not a restriction, as sampling theory tells us that the sampling can be performed with no loss of information at wavelengths greater than some prescribed lower limit. In practice, this criterion is usually met for the sampling in time (i.e., no aliasing), but, because of constraints on cost and accessibility, the sampling in space is often much less dense than optimum. For instance, in many cases it is practical to collect data only on the surface of the earth. Finally, there is the problem of noise. The geophysical fields of interest are often quite weak and must be observed in an environment that contains extraneous fields that contribute noise to the observational data. In some cases it is possible to characterize the statistical properties of this noise, particularly if it is stationary in time, but in many cases such a characterization is not possible. So the typical situation in geophysical inverse problems is that we must deal with data that have finite dimension, are insufficient, and are inaccurate. In the discussion that follows these data will be represented by the symbol $\mathbf{d}$.

The parameterization of the model space is much more under the control of the individual solving the inverse problem. The proper domain for describing properties within the earth is in most cases that of piece-wise continuous functions. This is the starting point for one of the first and most complete approaches to modern geophysical inverse problems, known as the Backus-Gilbert method (see Parker, 1994, for details). It was developed for linear inverse problems and requires that a Frechet derivative of the functional relating the model to the data be available. This approach is in a mature state, with the method fully developed and its properties well understood, so it will not be considered further in this report.

The need to work in an infinite dimension functional space, as required in the Backus-Gilbert method, is commonly avoided in geophysics by expanding the model in a set of basis functions, with the expansion coefficients then becoming the model parameters. Truncating the expansion then results in a discrete finite-dimension model space, which is similar to the data space, and will be represented as $\mathbf{m}$. This discretization of the model space is one of the most critical steps in the formulation of the problem, as later steps in the process, such as regularization and optimization, are highly dependent upon the scale of the discretization. An example of this type of approach is matched field processing, which can be very effective for some applications. Matched field processing uses just a small number of functions, which must themselves be determined by the data, to characterize the model of interest (Tolstoy et al., 1991).

3.2 Constraints

Once methods of parameterizing the data space and the model space have been selected, the basic idea of the inverse problem is to determine what constraints can be placed upon the model space so as to specify a model or group of models that are compatible with a particular set of observations drawn from the data space. A number of different types of constraints are possible. A theoretical constraint is a mapping from the model space to the data space that allows a direct relationship to be established. An objective function is some measure of distance in the data space that allows an evaluation of a model based on how close its simulated data are to the observed data. Regularization is generally a measure of some property of the model that is deemed to be desirable. Each of these will be discussed below.

It is assumed here that there exists some physical theory that relates the model to the observational data that can be expressed in the form

$$f(\mathbf{d}, \mathbf{m}) = 0 \, . \tag{9}$$

The usual situation is that the observational data $\mathbf{d}$ represent the solution of the theoretical problem, while the model $\mathbf{m}$ represents parameters of the equations. In many cases this becomes more explicit because the theory can be expressed in the reduced form

$$\mathbf{d} = \mathbf{a}(\mathbf{m}) \, . \tag{10}$$

This theory is typically in the form of an integral equation (canonical problems 1 and 4) or a differential equation (canonical problems 2 and 3). Thus solving the forward problem consists of specifying the model $\mathbf{m}$ and then obtaining a solution $\mathbf{d}$ that represents the data. An inverse problem arises when it is the data $\mathbf{d}$ that are given and the task is to find a model $\mathbf{m}$ that is compatible with these data.

There exist a few problems in geophysics, for example the one-dimensional travel time problem, where the inverse problem can be solved analytically to obtain a solution of the inverse problem in the form

$$\mathbf{m} = \mathbf{a}^{-1}(\mathbf{d}) \, . \tag{11}$$

Such solutions, which are rare and need to be treated as special cases, will not be discussed further here.

The usual situation is that only the forward problem can be solved in an analytical or semi-analytical sense. The solution of the inverse problem then proceeds by solving the forward problem employing a candidate model $\mathbf{m}$ in order to obtain simulated data $\mathbf{a}(\mathbf{m})$. A comparison between the simulated data and the observed data can then be used to make improvements to the candidate model. This requires the use of some measure of distance in the data space, i.e. a norm, that can be represented as

$$N(\mathbf{d}, \mathbf{m}) = ||\mathbf{d} - \mathbf{a}(\mathbf{m})|| \, . \tag{12}$$

The requirement that $N(\mathbf{d}, \mathbf{m})$ be a minimum represents a constraint upon the model that incorporates both a theoretical model of the forward problem and observational data.

A variety of other types of constraints can be placed upon the model $\mathbf{m}$ by specifying certain required or desirable properties that it should have, such as positivity, nearness to a particular model, or some measure of smoothness. These constraints can usually be expressed in the form of equality constraints

$$\mathbf{c}_e(\mathbf{m}) = 0 \ , \tag{13}$$

inequality constraints

$$\mathbf{c}_i(\mathbf{m}) \geq 0 \ , \tag{14}$$

or a regularization condition

$$S(\mathbf{m}) = minimum \ . \tag{15}$$

A typical equality constraint might be the total mass or moment of inertia of the whole Earth. A typical example of inequality constraints is to place upper and lower bounds on the permitted values of seismic velocity. Regularization conditions typically involve minimizing the deviation from a known model or minimizing fluctuations in the model. Recently a method has been introduced to minimize the total variation in the model (Rudin et al., 1992; Vogel and Oman, 1996; Dobson and Santosa, 1996).

3.3 Distinguishing Features

Most geophysical inverse problems have a number of distinguishing features that require special attention when methods of solution are being considered. The first feature is the fact that the observational data are usually incomplete in the sense that they do not contain enough information to resolve all features of the model. Because of this, solving a geophysical inverse problem generally consists of two separate stages, finding an optimum solution and appraising the validity of that solution. The appraisal stage includes an analysis of resolution, which is a determination of what features of the solution are necessary in order to explain the data. Invariably the optimum solution is non-unique in the sense that some of its features could be changed without changing the fit to the data.

The second distinguishing feature is that the data generally contain a noise component. The relationship in equation (10) between data and model should thus be given as

$$\tilde{\mathbf{d}} = \mathbf{a}(\mathbf{m}) + \mathbf{n} \ , \tag{16}$$

where $\mathbf{n}$ is the noise. This noise comes from two primary sources, a random component in the observational data and approximations or errors contained in the theory that connects the data and model. In either case, the model is not, and should not be, capable of completely explaining the data. The presence of noise means that the appraisal stage of the inverse problem should also include an analysis of uncertainty, which is a determination of how much the optimum solution would change if a different realization of the noise were to be used.

An important consequence of the incompleteness and inaccuracy found in geophysical inverse problems is that there are many possible solutions to the problem. The basic task, in its most complete form, is to describe all of these possible solutions. There are a number of methods that attempt this task by performing a general search of the model space, including grid searchs, random searchs, and pseudo-random searchs. There are also methods that estimate a relative probability density for the model space. All of these methods, which will be discussed more fully in a later section, are not very efficient and so far have been used only for small and moderate sized geophysical inverse problems.

The most common method of addressing the fundamental non-uniqueness of geophysical inverse problems is to impose additional constraints on the solution and in this way reduce the number of acceptable solutions (Parker, 1994; Oldenburg et al., 1998). This general process of introducing constraints that restrict the size of the solution space, in many cases reducing it to a single solution, is known as regularization and was introduced earlier as equation (15). These constraints that are imposed upon the model space are generally an attempt to reward certain properties that are deemed to be desirable, and in this sense they can be quite subjective, relying on information and prejudices that are independent of the data.

3.4 Optimization

With the definitions given above, it is now possible to define a typical geophysical inverse problem. An objective function is defined as

$$\Omega(\tilde{\mathbf{d}}, \mathbf{m}) = ||\tilde{\mathbf{d}} - \mathbf{a}(\mathbf{m})|| + \beta\, S(\mathbf{m})\,, \tag{17}$$

where β is a parameter that weights the relative importance of fitting the data and satisfying the regularization condition and where $\tilde{\mathbf{d}}$ are the observed data. Then the inverse problem is:
Given:

observational data $\tilde{\mathbf{d}}$, of finite dimension and possibly containing noise,

Determine:

a model $\mathbf{m}^*$ as the solution of

$$\begin{array}{ll} \min & ||\tilde{\mathbf{d}} - \mathbf{a}(\mathbf{m})|| + \beta\, S(\mathbf{m})\,, \\ \text{s.t.} & \mathbf{c}_e(\mathbf{m}) = 0\,, \\ & \mathbf{c}_i(\mathbf{m}) \geq 0\,. \end{array} \tag{18}$$

It is clear from this description that a typical geophysical inverse problem reduces to a problem in numerical optimization. The optimization problem usually involves equality and inequality constraints. The mathematical formulation of the optimization problem, however, is not unique. The formulation can have significant impact on the optimization techniques applicable and on the efficiency of the solution approach. A variety of objective functions are encountered, linear, quadratic, and more complicated forms. The objective function is generally smooth, although there is increasing interest in problems where it is not smooth, as in the total variation method (Rudin et al., 1992; Vogel and Oman, 1996) or the piecewise polynomial modification of truncated singular value decomposition (Hansen et al., 2000). The constraints, both equality and inequality, can be linear or nonlinear.

3.5 Probabilistic approach

The optimization methods described above can be postulated directly as we have done here so far, or the entire inverse problem can be viewed probabilistically from the outset and the resulting optimization methods derived in that context. The least squares type of optimization method described above follows naturally from some assumptions about Gaussian statistics for the data errors and maximum likelihood estimators (Aki and Richards, 1980; Tarantola, 1987; 1990). The advantage of the probabilistic approach is that it helps to clarify the assumptions that have gone into the formulation of a particular optimization scheme, which are not always obvious, and it also suggests alternative optimization schemes for different classes of problems.

3.6 Combined inversions

One of the most effective methods of reducing the fundamental non-uniqueness of geophysical inverse problems is to combine the results from several different data sets. The inverse problems for these different data sets can be solved separately and then the results compared, or the data can be combined and a single inversion performed. So long as the separate data sets all have similar relationships to the model parameters, this process of combined inversion presents no new problems, except that the size of the inverse problem grows with the size of the combined data set. The combined inversion should have advantages over the separate inversions in the areas of improved resolution and decreased uncertainty, particularly if the different data sets are distinct in the manner in which they sample the model. The combination of data sets that sample different types of model parameters is also possible. Here it is necessary to assume that the different model parameters, such as density and velocity, share a common structure and then a combined inversion is possible (Haber and Oldenburg, 1997).

There are also situations in geophysics where what appears to be a single inverse problem can be separated into different inverse problems with improved results. For instance, Xia et al. (1998) show how the long wavelength and short wavelength parts of a velocity model can be separately estimated. The primary advantage of such a separation is that different optimization procedures can be used for the different parts of the inverse problem, allowing the optimization procedure to be tailored to the particular attribute of the model that is being estimated. Another example is that of Gritto et al. (1999), where it is shown that a strongly nonlinear inverse scattering problem can be separated into a linear numerical optimization part and a nonlinear part that has an analytical solution.

4 Obtaining a Solution

The general outline of a geophysical inverse problem presented in the preceding section has the form of a mathematical optimization problem. Such problems have been thoroughly studied and a large selection of methods for solution are available (see for instance Nocedal and Wright, 1999; Dennis and Schnabel, 1996; Fletcher, 1987), depending upon the type of parameterization, objective function, and constraints. This is indeed fortunate, as, given the variety of geophysical inverse problems that are encountered, it is unrealistic to think that any one approach would be optimum for all of them. Thus one of the tasks of solving an inverse problem is to chose the optimization method that is most appropriate. In this section some of the considerations that help determine this choice will be discussed.

It is important to point out that considerable resources are already available within the numerical optimization community for this task of choosing the most appropriate optimization algorithm. Compilations such as the *Optimization Software Guide* (Moré and Wright, 1993) present outlines of software available for various types of optimization problems and provide guidance in making a choice. It is actually possible to do some of this optimization over the network through the Network-Enabled Optimization Server (NEOS) (Cryzyk et al., 1996)(http://www.mcs.anl.gov/home/otc/).

A general finding of the workshop was that those solving geophysical inverse problems could benefit from more familiarity and better access to the various optimization methods that are available. Resources such as those described in the previous paragraph have so far received very little use in the geophysical community. The concept that no single algorithm is likely to be best for all inverse problems is formalized in the "no free lunch" theorems (Wolpert and Macready, 1997), which demonstrate how an algorithm that performs well for one class of problems may perform poorly for another class. What is clearly needed is more collaboration between geophysicists working on inverse problems and mathematicians working on optimization methods.

A related issue is the design of software interfaces between application codes that provide objective and constraint function information and the optimization software. Currently, many implementations of optimization algorithms, especially those for constrained optimization, require objective and constraint functions and their derivatives in formats that are not suitable for complex, large-scale geophysical inverse problems. As a result, optimization algorithms are often laboriously re-implemented for specific applications, numerical comparisons of different optimization approaches for the application at hand are strongly discouraged, and the dissemination of new optimization techniques into the inverse problem community is severely hindered. This issue and remedies are discussed in Heinkenschloss and Vicente (1999a) and in Gockenbach et al. (1999). The design of suitable interfaces between application and optimization requires a careful software design, but also involves all phases of the problem solution from the mathematical statement and parameterization of the problem to the design and analysis of optimization algorithms.

4.1 Formulation of the optimization problem, implicit versus explicit constraints

Choices of the parametrization, of the regularization term $S(\mathbf{m})$, and of the constraints $\mathbf{c}_e(\mathbf{m})$, $\mathbf{c}_i(\mathbf{m})$ clearly all affect the formulation of the optimization problem and are crucial for the solution of the inverse problem. However, even after these choices have been made, there are several ways to formulate the resulting optimization problem mathematically. A specific mathematical formulation of the optimization problem excludes certain optimization approaches and therefore has an impact on the solution efficiency and possibly even robustness.

For example, in (18) we have stated a formulation of the geophysical inverse problem as an optimization problem. There we have assumed that a solution $\mathbf{d}$ of the forward problem (9) is available in the reduced form (10) and we have inserted this into our optimization problem (18). Thus, whenever we have to evaluate the objective function (17) or, possibly, the constraints $\mathbf{c}_e$ or $\mathbf{c}_i$, we have to evaluate $\mathbf{a}(\mathbf{m})$, i.e., we have to solve the forward problem (9) for $\mathbf{d}$. Alternatively, we may include the equation $f(\mathbf{d},\mathbf{m}) = 0$ that relates the model $\mathbf{m}$ to the data $\mathbf{d}$ as an explicit constraint into the optimization problem. The problem (18) is then reformulated as

$$\begin{aligned} \min \quad & ||\tilde{\mathbf{d}} - \mathbf{d}|| + \beta\, S(\mathbf{m}) \,, \\ \text{s.t.} \quad & f(\mathbf{d},\mathbf{m}) = 0 \,, \\ & \mathbf{C}_e(\mathbf{d},\mathbf{m}) = 0 \,, \\ & \mathbf{C}_i(\mathbf{d},\mathbf{m}) \geq 0 \,, \end{aligned} \tag{19}$$

where, as before, $\tilde{\mathbf{d}}$ denotes the observed data and $\mathbf{d}$ denotes the simulated data. In (19) both $\mathbf{d}$ and $\mathbf{m}$ are optimization variables. Here we have assumed that the constraint functions $\mathbf{c}_e(\mathbf{m})$ and $\mathbf{c}_i(\mathbf{m})$ in (18) may even depend on the simulated data, i.e., are of the form $\mathbf{c}_e(\mathbf{m}) = \mathbf{C}_e(\mathbf{a}(\mathbf{m}),\mathbf{m})$ and $\mathbf{c}_i(\mathbf{m}) = \mathbf{C}_i(\mathbf{a}(\mathbf{m}),\mathbf{m})$.

Often, the two formulations (18) and (19) of the inverse problem are equivalent. Solution approaches for (18) and (19) require similar problem information (see, e.g., Dennis et al. (1998)).

However, the formulation (19) can be tackled by a broader class of optimization algorithms, which includes those that do not enforce the forward problem $f(\mathbf{d}, \mathbf{m}) = 0$ at every step of the optimization. Currently, many large–scale geophysical inverse problems are formulated as (18) and, if there are no constraints, they are solved using nonlinear conjugate gradient methods. Line search procedures within the conjugate gradient method require repeated solves of the expensive forward problem. The formulation (19) of the inverse problem on the other hand, can be tackled by sequential quadratic programming (SQP) methods (see, e.g., Nocedal and Wright, 1999), which only require that the forward problem be solved in the limit as the iterates approach the solution, which can result in significant computational savings. For a comparison on an optimization problem with a nonlinear forward problem see, e.g., Ghattas and Bark (1997). Moreover, (19) reveals how the forward problem enters the optimization formulation, which offers ways to control the accuracy with which the forward problem has to be solved within the optimization. Inexact, inexpensive solutions of the forward problem can be admitted in the early stages of the optimization (Heinkenschloss and Vicente, 1999b). Trust region strategies guarantee the convergence from a bad initial solution estimate for nonconvex optimization. The theory and implementation of these methods has progressed to the point that the local convergence behavior of most SQP type algorithms is much more difficult to analyze than the global behavior (see Nocedal and Wright, 1999, or Conn et al., 2000). Trust region methods seem particularly appealing in the context of inverse problems because of the regularizing effect of the trust region constraint on the optimization step.

Formulation of the geophysical inverse problem and choice of optimization algorithms for their solution were areas identified during the workshop where usual practices in geophysics need to be re–evaluated.

4.2 Complexity of the forward problem

Solution of the forward problem may represent a major time factor in the overall optimization process. One of the findings of the workshop was that one of the best methods of improving the solution of geophysical inverse problems is to develop better methods of solving forward problems.

In addition to the development of better forward problem solvers, optimization methods that allow greater flexibility in the integration of forward problem solves or linearized forward problem solves should be investigated. Current optimization approaches for (18) often require rather accurate forward solves because inaccuracies in the simulated data is a source of noise in function and derivative evaluations. Ideally one would like to adapt the accuracy in forward or linearized forward solves to the progress of the optimization algorithm and allow coarse, relatively inexpensive forward solves away from the minimum and only tighten the accuracy requirements in the forward solve as one approaches the minimum.

An approach that appears to have obtained little notice in the field of geophysics is the use of surrogates for the forward problem. This approach has been quite successful in engineering design where there may be at most a few dozen optimization variables (Torczon and Trosset, 1998), and Booker, et al. (1999) have put the approach into a rigorous mathematical framework. For many optimization methods, precise and detailed solutions to the forward problem are really not required in all stages of the process. The point is that much of the search for a solution to the inverse problem can be performed with a surrogate solution to the forward problem, based either on interpolatory surfaces or simplified physical principles. The idea is that one approximates the essential features of the complete solution with a surrogate simulation that can be executed much more efficiently. An example of this approach is the use of straight rays in travel time tomography. Even though rays clearly bend in the Earth, the straight ray approximation may nevertheless be very useful in low contrast or anisotropic media, and it is very quick to compute.

Software (C++) for incorporating surrogates into a filter pattern search method for generally constrained problems can be found at www.caam.rice.edu/~dougm. It is assumed that the user will prefer to furnish their own specific application-specific surrogates. The work of Booker, et al. (1999), and other work by the same group, is based on the kriging surrogates from geophysics.

Still another approach that has been shown to work in difficult highly non-linear problems is to obtain solutions to an approximate problem that behaves like the real problem in an asymptotic limit. Approaches of this type appear to be effective for various transport problems in high contrast media (Borcea et al., 1996; Borcea and Papanicolaou, 1998; Borcea, 1999; Borcea et al., 1999; Dorn, 1998; Dorn et al., 1999). Work should be done to ascertain whether these approximate solutions can be used in the surrogate management framework of Booker et al. (1999).

4.3 Discretization

An important step in many geophysical inverse problems is the part of the parameterization process where a continuous model space is converted to a discrete set of parameters through an expansion in a set of basis functions. A common form of this discretization process is to divide the model space into a set of non-overlapping cells with the parameters being the mean values for the cells. A critical question arises in regard to the best choice for the scale of the discretization, which often reduces to a choice for the dimensions of the cells. Awareness of the following two points can be useful in making this choice. First, it is important to understand that discretization is really a form of regularization, as the smoothness of the model and its ability to fit the data are directly related to the scale of the discretization. For instance, whether a problem is under-determined or over-determined is directly related to the scale of the discretization in most problems. Second, the scale of the discretization can be included in the inversion process as a parameter to be optimized. This is an opportunity to remove one type of subjectivity from the inversion process, and it also has benefits in the appraisal stage.

Multiple scales can also be used to increase the efficiency of the optimization. For example, optimization on coarse scales can give good starting values for the optimization on fine scales, coarse discretizations can be used to obtain less expensive second derivative approximations for the problems on fine scales, and coarse scale information can be used to design preconditioners for optimization subproblems on the fine grids. Finally, properties of the underlying infinite dimensional problem and the choice of discretization provides important information about the 'scaling' (as used in the optimization language, see, e.g., Dennis and Schnabel, 1996; Heinkenschloss and Vicente, 1999a) of the optimization problem.

4.4 Feasibility constraints

For some inverse problems the theory constraint can be stated in the form of a variational principle (Berryman, 1990; Berryman and Kohn, 1990; Berryman, 1993, 1997), which has several attractive implications for the inverse problem. It may be possible to demonstrate that the set of feasible solutions is convex, and nonlinear programming methods are well suited for these types of problems. As an example, in canonical problem 4 for travel time tomography it is possible to show that the ray path Σ that satisfies the ray equations (8) also satisfies

$$t(\mathbf{x}_r, \mathbf{x}_s) = min \int_{\Sigma(\mathbf{x}_r, \mathbf{x}_s)} \frac{d\sigma}{v(\mathbf{x})} , \tag{20}$$

where the minimization is taken over all possible paths Σ. Thus, the only velocity distributions that are feasible are those for which

$$t(\mathbf{x}_r, \mathbf{x}_s) \geq \int_{\Sigma(\mathbf{x}_r, \mathbf{x}_s)} \frac{d\sigma}{v(\mathbf{x})} . \tag{21}$$

It has been shown that the feasible set for this problem is convex and that the solution lies on the boundary of the feasible region. Thus, many optimization methods can be used to take advantage of the variational structure of the travel time tomography problem.

It is also known that for some problems, such as electrical resistance tomography (Berryman and Kohn, 1990), there are dual variational principles available, which means that data can be used to provide rigorous bounds on the reconstructed electrical conductivity model. One unique feature of the feasibility constraints is that their form does not change in a significant way whether the inversion problem is linear or nonlinear, making this approach one of the few permitting rigorous statements about nonlinear inversion problems. To date, these features have not been very well exploited in inverting geophysical data.

4.5 Derivatives

Closely related to the complexity of the forward problem is the question of whether partial derivatives of the data with respect to the model parameters are required by the optimization process. Methods that require such derivatives, such as Newton methods and quasi-Newton methods, typically are efficient in that they converge to a solution with a minimum number of evaluations of the forward problem. However, evaluating the derivatives may in itself be a major task that is even more complicated than the forward problem. A recent development in this area that appears to be under-utilized in the field of geophysics is the availability of automatic differentiation tools.

Automatic differentiation (AD) tools make computer models more useful by augmenting them to provide sensitivity information in addition to the model outputs. Furthermore, the augmentation process requires little or no user intervention (hence the "automatic"). The capability of deriving accurate sensitivities from a computer model with little additional development cost enables users of computer models to develop sophisticated applications in pleasingly short amounts of time.

Current state-of-the-art automatic differentiation tools employ a variety of techniques to compute derivatives. For example, a user of a modern AD tool could employ forward mode (standard chain rule), reverse mode (adjoints), sparse vectors and matrices, univariate Taylor interpolation, and fixed-point iteration to compute the desired sensitivities.

Current AD tools, however, are language specific. To augment a computer model with derivatives, the user must employ an AD tool that is specifically designed for the source language of the model. The primary language addressed by a current high-quality AD tool will almost certainly be Fortran 77 or C. At present, this is not a serious limitation, as an overwhelming majority of "legacy" computer models are written in Fortran or C.

Advancing the state-of-the-art in AD will come from considering improved techniques, additional languages, and additional augmentation. The term "improved techniques" refers to ongoing algorithm and implementation research. The AD methods that are currently considered best may be supplanted by better methods. In particular, the memory requirements for AD adjoint techniques should certainly be improved. Furthermore, parallel computing may be better exploited by improved AD techniques.

The term "additional languages" addresses the known limitation of current AD tools to Fortran and C. More recent (i.e. non-legacy) computer models are being implemented in more modern languages such as C++, Fortran 90 and Java. The sophisticated semantic features of these languages pose some interesting challenges for AD tools.

Finally, "additional augmentation" generalizes the augmentation aspect of AD. An AD tool augments a computer model to "propagate" derivatives according to well-known rules. There are, however, additional mathematical objects besides derivatives that might be propagated. For examples, the "verified computing" community already constructs programs to propagate "intervals". Augmenting a computer model to propagate intervals could certainly be potential benefit when validating a computer model, or using it for robust (i.e. minimum variance) design applications. A second kind of interesting augmentation could be "probability measures". For some given probability distribution of input values, computing (or approximating) the probability distribution of model outputs could be useful for reliability modeling. A third possible augmentation could be Fourier coefficient propagation. For a given frequency component, Fourier propagation might enable users of a computer model to assess certain stability properties. Generalizing Fourier propagation to wavelet propagation might also produce some interesting additional information.

4.6 Linear problems

For geophysical inverse problems where the relationship between the data and the model is linear, the methods of solution are well developed and well understood (Menke, 1989; Parker, 1994). For model parameterizations that are either continuous or discrete, methods of solution have been developed for a number of different optimization criteria. For instance, when noise is present and its distribution is known, solutions that are maximum likelihood are available. Measures of resolution in both the data space and model space can also be calculated (Berryman, 2000), thus allowing quantitative estimates of the fitting of the data and the uniqueness of the model. An important feature of linear problems is that it is generally possible to map noise in the data directly into a measure of uncertainty in the model.

One feature of linear problems that needs improvement is a more optimum handling of regularization. Most geophysical inverse problems are not well posed as originally formulated, and the usual method of alleviating this situation is to impose some form of regularization. However, the regularization is most often imposed in an ad hoc manner and it is difficult to find a balance between the amount of regularization necessary to make the problem well posed and that which grossly distorts the solution. What is often possible but rarely done in geophysical problems is to make the degree of regularization a variable parameter that is optimized as part of obtaining the solution. This appears to be one area where the handling of linear problems could be improved.

4.7 Linearized problems

Given the considerable machinery that exists for solving linear inverse problems, there is a tendency to formulate problems so that linear methods can be used whenever possible. For problems that are not too strongly non-linear, this can be done by a process of linearization. Consider the reduced form of the theory constraint (equation 10) and write

$$\mathbf{d} = \mathbf{a}(\mathbf{m}_o + \delta\mathbf{m}) \approx \mathbf{a}(\mathbf{m}_o) + \frac{\partial \mathbf{a}(\mathbf{m}_o)}{\partial \mathbf{m}}\, \delta\mathbf{m}\ , \tag{22}$$

where it has been assumed that $\mathbf{m}_o$ is a reference model and that higher order derivative terms are small enough to be ignored. Then, defining $\delta\mathbf{d} = \mathbf{d} - \mathbf{a}(\mathbf{m}_o)$, we have the linearized problem

$$\delta\mathbf{d} = \frac{\partial \mathbf{a}(\mathbf{m}_o)}{\partial \mathbf{m}}\, \delta\mathbf{m}\ . \tag{23}$$

So long as the the solution does not stray too far from the reference model $\mathbf{m}_o$, this problem can be solved with standard linear methods, which also includes the standard linear estimates of resolution and uncertainty. The situation where the reference model is unknown is handled by an iterated linearization procedure in which a new reference model is taken to be $\mathbf{m}_o + \delta\mathbf{m}$ and the entire linearization and solution process repeated. This type of linearized approach to the solution of an inverse problem is commonly used in the location of earthquakes where it is known as Geiger's method (Lee and Stewart, 1981).

The process of solving a non-linear inverse problem by solving a series of linearized problems is in principle no different from some of the standard iterative methods developed for solving non-linear problems, such as the line search and trust region methods (Dennis and Schnabel, 1996). An advantage of using these established non-linear methods for problems of this type is that convergence proofs exist and well-tested algorithms are available. Thus, in the case of many geophysical problems, it is difficult to justify the linearization of a problem when efficient methods of solving the non-linear problem are available.

4.8 Nonlinear problems

Many geophysical inverse problems fall into the category where both the objective function and the constraints are significantly nonlinear in the model parameters. The choice of whether to use methods that do or do not require derivatives is especially important in this case because both first and second derivatives may be required by the methods that do use derivatives.

There is a variety of methods for this type of problem that use derivatives and global and local convergence proofs are available for most of these methods. For a discussion of these methods it is important to distinguish between the two related formulations (18) and (19).

For problems (18) without constraints one could apply Newton's method. This requires both first and second derivatives, i.e. the Hessian matrix. Obtaining the second derivatives may be difficult, either because the mathematical expressions are difficult to derive or because they are computationally expensive. There is a class of geophysical inverse problems, such as global seismic tomography, where the problem is so large that second derivatives are out of the question, as are any matrix inversions. The nonlinear conjugate gradient (CG) methods and variants such as LSQR (Paige and Saunders, 1982) have turned out to be the methods of choice for these types of problems and their performance seems to be satisfactory (Nolet, 1984, 1985; Newman and Alumbaugh, 1997). Alternatively, inexact CG–Newton methods or limited memory quasi–Newton methods should be be considered (Nocedal and Wright, 1999). Inexact CG–Newton methods do not require the explicit computation of the Hessian matrix, but only require the calculation of Hessian–matrix–vector products. The latter is often feasible even for very large–scale problems and the relations between the problem formulations (18) and (19) reveal structure in the gradient and Hessian computations for (18) that can be used to implement Hessian–matrix–vector products (see, e.g., Dennis et al., 1997; Dennis and Vicente, 1997; Heinkenschloss and Vicente, 1999b). Limited memory quasi–Newton methods replace the true Hessian matrix with low rank matrix whose storage requirement can be limited by the user. For both methods, inexact CG–Newton and limited memory quasi–Newton, convergence of the iterations from arbitrary starting values can be obtained by using line–search and trust–region globalization strategies (Nocedal and Wright, 1999). For both methods the added line–search or trust–region globalization strategy will typically become inactive near the minimum, which eliminates the need for additional function evaluations. Nonlinear CG methods require a line–search globalization and, even near the minimum, a line–search involving expensive function evaluations at trial steps is usually necessary. Trust–region globalizations have a regularizing effect on the optimization step computation, since the trust–region may be viewed as a regularization constraint on the step with the regularization parameter adapted by the optimization method.

Problems (19) can be solved using sequential quadratic programming (SQP) methods. Inexact step computations, quasi–Newton approximations to Hessians, and line–search and trust–region globalizations are available for this class of methods. As pointed out in Section 4.1, one can formulate most geophysical inverse problems as (18) or (19). The formulation (19) promises significant advantages and should be further evaluated in the context of geophysical inverse problems.

The efficient solution of problems with many variables and many inequality constraints is a very active area in optimization research. Optimization techniques such as projection methods or interior–point methods need to be investigated for large–scale inverse problems. Since some inequality constraints might be rather 'soft' regularization constraints, other application dependent techniques for handling inequality constraints could be envisioned.

There has been progress in recent years on search methods that do not require derivatives. Such methods are sometimes chosen because they are less likely to converge to a nearby local optimum than derivative based methods. Some choices are genetic algorithms, pattern search algorithms, and an interpolation method, DFO, of Conn et al. (1997). DFO seems very promising, but there is a random component to the algorithm that makes it difficult to predict how it will behave in a given run.

Genetic algorithms will be discussed more below, but they are thought generally to obtain excellent initial decrease, but then to stagnate around a local optimum. A representative example of this behavior can be found in Booker, et al. (1999), on a problem in helicopter rotor design. Results are given there for DFO, for a surrogate management method based on kriging, and for the parallel direct search (PDS) of Dennis and Torczon (1991). DFO and the surrogate management method are far better for this example than the alternatives, but all the results are given for a sequential implementation. That is important because, though all the methods have a great deal of potential for parallelism, one can not be sure of how each would do on a particular class of problems (see Hough et al., 2000).

As implemented, PDS fails in a completely trivial way to satisfy the general convergence theory given by Torczon (1997), and so the PDS results can be viewed as representative of the results to be obtained from the generalized pattern search (GPS) class of algorithms defined and analyzed by Torczon (1997). Torczon and Trosset (1997), Lewis and Torczon (1996, 1998a, 1998b, 1999), and Lewis et al. (1998) give useful and interesting extensions of the algorithms and the supporting theory, especially their work on problems with a finite number of linear constraints, which could be bounds on the parameters in geophysical inverse problems. They also give an interesting approach related to GPS with a very satisfying convergence theory for general constraints. Audet and Dennis (2000a, 2000b, 2000c) extend this class of algorithms to handle mixed continuous and discrete variables and general constraints, and they give a new analysis showing convergence for discontinuous problems with appropriate optimality conditions for limit points at which the problem is locally smooth. Thus, the GPS algorithms are especially attractive in that they are broadly applicable, simple to implement and supported by a strong proof of convergence to local optima. When used in conjunction with the feasibility constraint methods mentioned previously (Berryman, 1997), such methods can take advantage of the variational structure of the fully non-linear inversion problem.

The problem of multiple solutions is particularly important for nonlinear inverse problems, as most optimization methods only provide a local extremum and separate procedures must be used to find a more global extremum. Methods are available that are designed to find the global extremum. For example, one such method uses a combination of smoothing and continuation to find global solutions (Moré and Wu, 1997). Another method, the terminal repeller unconstrained subenergy tunneling (TRUST) algorithm, has been used to solve a fairly difficult geophysical inverse problem, the estimation of residual seismic static corrections (Barhen et al., 1997). Another approach deals directly with the nonlinear nature of the problem and uses recent advances in computational algebra to handle the polynomial equations that must be solved (Everett, 1996; Vasco, 1999, 2000).

Grid search and stochastic search methods are also designed to find global extrema. For small problems it may be possible to perform a grid search in which all members of the model space are examined and either accepted or rejected. For somewhat larger problems a Monte Carlo search may be possible and it has the advantage of being simple to implement and easy to check (Mosegaard and Tarantola, 1995; Mosegaard, 1998). However, for most geophysical inverse problems the number of model parameters and the required accuracy are such that a complete Monte-Carlo search is unfeasible simply because of the number of times the forward problem would have to be calculated to achieve sufficient sampling of the model space. The search of the model space can also be guided by a statistical Bayesian approach in which a combination of prior information and the information contained in the observational data are used to construct some measure of relative probability for the model space (see for example Tarantola, 1987, or Sen and Stoffa, 1995). While these methods that attempt a general search of the model space are appealing because of their simplicity and completeness, they are not yet practical for most geophysical inverse problems. This is because the size of the model space, which is of order 10^M where M is the number of model parameters, is generally much too large to allow a general search in finite computational time. Nevertheless, there continues to be considerable effort devoted to the task of improving the efficiency of stochastic search methods. Bosch et al. (2000) obtain promising results using a combination of importance sampling and multi-step sampling.

While neither enumerative nor completely random searchs of the model space have proven to be effective methods of solving most large geophysical inverse problems, there are some directed search methods, also called pseudo-random search, that have been more successful. Two examples are simulated annealing and genetic algorithms. Both of these approaches retain some aspects of a random statistical search of the model space but use the gradually accumulating information about acceptable models to direct the search into those parts of the model space where good models are most likely to be found. These approaches appear to be feasible for moderately sized problems where a full Monte Carlo approach would be prohibitive (Scales et al., 1992).

Simulated annealing is based upon an analogy with a natural optimization process in thermodynamics and uses a directed stochastic search of the model space. It requires no derivative information. Its use in numerical optimization problems began with Kirkpatrick et al. (1983) and its first use in geophysical problems appears to be Rothman (1985, 1986). A review of the method and its application to geophysical problems can be found in Sen and Stoffa (1995) and examples of its use in Sen and Stoffa (1991), Mosegaard and Vestergaard (1991), and Varela et al. (1998).

Another class of directed search methods are the evolutionary methods that make use of analogies with the natural optimization processes found in the evolution of biological systems (Holland, 1975; Goldberg, 1989). One class of such methods, genetic algorithms, applies the operators of coding, selection, crossover, and mutation to a finite population of models and allows the principle of "survival of the fittest" to guide the population toward a composition that contains the optimum model. This approach has been applied to a number of geophysical problems (see for instance Stoffa and Sen, 1991; Sen and Stoffa, 1992, 1995; Sambridge and Drijkoningen, 1992; Kennett and Sambridge, 1992; Everett and Schultz, 1993; Sambridge and Gallagher, 1993; Nolte and Frazer, 1994; Boschetti et al., 1996; Parker, 1999). Another class of evolutionary methods, evolutionary programming (Fogel, 1962; Fogel, 1995; Bäck, 1996), uses only the operators of selection and mutation and has only recently been applied to geophysical problems (Minster et al., 1995; de Groot-Hedlin and Vernon, 1998).

Approaches that attempt some combination of stochastic and deterministic search methods would appear to hold considerable promise. The general idea is to combine the global search property of the stochastic methods with the efficiency of the deterministic methods. Of course, one must always keep in mind that global optimization of a general function is computationally intractable and so no method is sure to work. This is understood for most general nonlinear problems, but global optimization has a further difficulty. Specifically, even if one has found a global optimizer, it is impossible to recognize it for a general problem (Stephens and Baritompa, 1998). This does diminish the importance of work on global optimization methods, but, on the contrary, it just shows how difficult the problem is and what we can hope to accomplish.

4.9 Very large problems

There exist some geophysical inverse problems that are so large that special methods of solution have to be used. Tomography problems such as canonical problem 4 often fall into this class, where it is not unusual to have on the order of 10^7 data and 10^5 model parameters. Linearization about a reference model $\mathbf{m}_o$ is almost always performed in these problems, and the reference model is usually held fixed. The choice of a method of solving such a linear system is restricted by the fact that it is not possible to fit the entire coefficient matrix into primary storage of most computer systems, even though this matrix is usually sparse. This restriction eliminates many standard solution methods, but there do exist approaches that require access to only one row of the coefficient matrix at a time. These are iterative methods, but in this case the iterations are needed just to solve a linear system of equations. Two general approaches of this type have been commonly used in geophysics, the class of algebraic reconstruction methods and the class of projection methods (van der Sluis and van der Vorst, 1987). A more recent approach to the solution of linear large-scale discrete ill-posed problems that only requires matrix–vector products is described in Rojas and Sorensen (1999).

5 Appraisal of the Results

An important characteristic of geophysical inverse problems is that the solution to the optimization problem is most likely not the true model of the earth that generated the data. Thus, it is generally recognized in geophysics that a complete solution should include a description of the optimum model and an evaluation of how this optimum model may be related to the true model. This latter evaluation typically includes two separate aspects, the sensitivity of the optimum model to incompleteness in the data and the sensitivity to noise in the data.

5.1 Resolution

For most geophysical inverse problems the amount of information in the data is insufficient to independently determine all parameters of the model (Jackson, 1972; Alumbaugh and Newman, 1997). This can be expressed as

$$\mathbf{m}_{opt} = \mathbf{R}(\mathbf{m}) \tag{24}$$

where the resolution operator $\mathbf{R}$ maps the true model $\mathbf{m}$ into the model $\mathbf{m}_{opt}$ produced by the optimization procedure. Departure of $\mathbf{R}$ from an identity operator signifies imperfect resolution. Typically it describes a smoothing operation because fine details of the true model can not be resolved by the available data. Stated in another way, an imperfect resolution operator says that the solution is non-unique, a characteristic of most geophysical inverse problems.

For linear problems the construction of the resolution operator is straightforward (Jackson, 1972; Wiggins, 1972; Menke, 1989). However, for very large problems this task may represent a prohibitive computational burden. Recent advances have shown how to compute approximations to the resolution operator iteratively for such large systems (Nolet, 1985; Zhang and McMechan, 1995; Minkoff, 1997; Berryman, 2000). Alternatively, resolution can be approximated by showing how well the features of a synthetic model can be reproduced by the inversion method. The use of such approximate measures of resolution can be misleading (see for instance Lévêque et al., 1993) unless the synthetic model contains a complete range of features.

For nonlinear problems the concept of resolution is still important but general methods for its estimation are not available. The Occam's razor approach (Constable et al., 1987; deGroot-Hedlin and Constable, 1990) proceeds by over-parameterizing the model, including a measure of smoothness in the objective function, and then solving for the smoothest possible model consistent with the data. Kennett and Nolet (1978) suggested an approach that can be used with stochastic search methods (Sen and Stoffa, 1995) to produce a suite of successful models.

5.2 Uncertainty

When noise is present in the data the inverse problem acquires a random nature and the task is to determine how uncertainty in the data is propagated into uncertainty in the optimum model.

This uncertainty is typically represented in terms of a covariance matrix and then

$$cov[\mathbf{m}_{opt}] = \mathbf{U}(cov[\mathbf{d}]) \tag{25}$$

where $\mathbf{U}$ represents the uncertainty operator.

For linear problems explicit expressions for $\mathbf{U}$ are possible (Menke, 1989) and the major difficulty lies in estimating the statistical properties of the noise. The critical element of the analysis is that there is a linear mapping between the probability distribution of the noise and that of the model, which means that meaningful statistics can be derived.

For nonlinear problems a measure of uncertainty is much more problematical. Analytical expressions for the shape of the objective surface in the vicinity of the optimum model are generally not available. When sufficient data redundancy is present, a direct exploration of this surface with resampling methods, such as the jackknife and bootstrap, is possible (Efron, 1982).

5.3 Trade-offs

In almost all geophysical inverse problems there is a relative weighting between the objectives of fitting the data and satisfying the regularization condition (the parameter β in equations 18 and 19). When the emphasis is on fitting the data the solution is likely to be unstable and the uncertainty large. When the emphasis is on satisfying the regularization condition the solution is likely to be inaccurate and the resolution poor. Thus there is a trade-off between two incompatible objectives and some method of choosing the trade-off parameter is required. This choice is often rather subjective, depending upon estimates of the accuracy of the data and expected properties of the solution. It would be a useful contribution to geophysical inverse problems if methods of optimizing the choice of this trade-off parameter could be included in the solution of the problem. Lenhart et al. (1997) suggest one way of achieving this using optimal control methods.

5.4 Posterior probability

In statistical approaches to the inverse problem, such as Bayesian inference, the concepts of resolution and uncertainty are lumped together into a posterior probability density function. For realistic geophysical inverse problems the calculation and display of this probability density function can represent a major numerical task, particularly for nonlinear problems where analytical approximations are not valid. Sen and Stoffa (1996) have considered several different methods of making this process more efficient.

6 Computational Needs

Throughout the history of geophysical inverse methods the improvement in computational resources has been just as important as the improvement in methods of analysis for advancement in the quality and quantity of the solutions that can be obtained. This is likely to be true for future advancement also. The situation still exists in geophysics where there are significant inverse problems that are not being solved primarily because of the lack of the necessary computational resources.

6.1 Infrastructure

The availability of the necessary computational facilities continues to be an important consideration in the choice of geophysical inverse problems that are attempted. Fast work stations have greatly expanded the convenience of solving small to moderate sized problems. However, for large problems there is still a need for the resources that can only be found at specialized facilities. This is because the large problems need not only fast computers but also large amounts of memory and storage.

The solution of geophysical inverse problems can place demands upon a computational system that are different from those of normal usage and often in conflict with normal administrative methods. For instance, some problems require the analysis of massive amounts of observational data both prior to and during the actual solution of the optimization problem. These data are typically too massive to be stored in active memory and must be continuously migrated between memory, cache, and disk. Furthermore, with allowance for monitoring of the process and inspection of intermediate steps, such analysis of the data could go on for weeks or even months, even though the amount of CPU time used during that period could be quite moderate.

6.2 Massively parallel systems

There is a class of large and difficult geophysical inverse problems that at present can only be attempted on massively parallel computer systems (Newman and Alumbaugh, 1997). This is true of many problems that attempt a complete analysis of three-dimensional properties of the earth, with the appraisal stage of the analysis often being more computer intensive than the solution stage. Solving such problems on massively parallel systems usually requires special organization of the problem and the use of special methods that take advantage of the parallel architecture.

6.3 Visualization

Improvements in visualization equipment and software could contribute significantly to the solution of many geophysical inverse problems. Visualization is needed not only for the display of final results, such as the the three-dimensional distribution of some property within the earth, but also for the display of intermediate results, such as a depiction of the progress being made by the search algorithm.

7 Summary and Conclusions

The workshop was successful in identifying a number of areas where improvements are needed in our ability to solve geophysical inverse problems and in suggesting some directions of research that might possibly achieve these improvements. A summary of these targets for future study is listed below. It should be pointed out that this list is not complete, as only a few specific topics of geophysical inverse practice were discussed at the workshop, with the choice controlled mainly by the interests of the participants. The list is heavily weighted toward those areas where there were obvious connections between the fields of numerical optimization and geophysical inverse methods.

Two general needs that always have and always will be controlling factors in the progress of geophysical inverse methods are:

- Better solutions to the forward problem. It is a common situation in geophysical inverse problems that the solution of the forward problem is extremely difficult and very time consuming on a computer, which imposes a severe limitation upon the solution of the inverse problem. Many problems, particularly those that involve three-dimensional distributions of material properties within the earth, are simply not being done because of this limitation. In some cases it is possible to use simplified and approximate solutions, but this can introduce additional uncertainty into the inverse problem. Thus there is a continuing need for more efficient and more accurate methods of solving geophysical forward problems.
- More computational resources. The types and sizes of geophysical problems that are being solved today is limited by the available computational resources, with a latent list of additional problems that await improvements in those resources. In addition to improvements in speed, the computational needs for geophysical inverse problems include:
  - Large amounts of memory, cache, and disk space.
  - Management policy that allows long residency of data.
  - More access to massively parallel systems.
  - Better visualization capabilities.

It is obvious that these two needs, better solutions to the forward problem and more computational resources, are closely related, as the solution of the forward problem is often the most computationally intensive part of the inverse problem and it is here that additional computational resources would be most effective.

The traditional approach of dividing a geophysical inverse problem into the separate stages of formulation, solution, and appraisal, with optimization included primarily in the solution stage, should be re-thought. A more general and more effective paradigm may be:

- Optimization should be included in all stages of the inverse problem. This would allow a number of improvements:
  - The scale of the discretization, which is often part of the formulation stage, could be chosen in an optimum manner.
  - The role of regularization in the determination of the solution would become more evident and could be selected in a more optimum manner.
  - The trade-off between fitting the data and satisfying the regularization condition could be made more objective.

Certain general tendencies have developed in the formulation and solution of geophysical inverse problems that need to be re-examined. Other possibilities that are available and should be considered include:

- Include the theory connecting the data and the model as a constraint rather than as part of the objective function. In principle, this change does not really change the formulation of the problem, but it can have a significant effect upon the ability of optimization algorithms to effectively find a solution. The main effects of this approach are:
  - Couplings between the variables that makes the problem appear more nonlinear to the optimizer are avoided.

- The optimizer performs more efficiently in finding a local minimum.
- The number of parameters that have to be optimized is increased.

- Invoke the theory part of the inverse problem as a feasibility constraint.
- Take advantage of automatic differentiation tools for problems where optimization methods requiring derivatives have been avoided because of the difficulty involved in obtaining analytical derivatives.
- Use well developed nonlinear optimization methods instead of linearizing the inverse problem.
- Consider the possibility that non-smooth objective functions and regularization conditions might be valid descriptions of the situation within the earth.

While very large problems that push against the limits of the available computational resources are likely to remain as one of the main challenges of geophysical inverse problems, there are some approaches that might alleviate at least part of this difficulty:

- The use of surrogates to solve the forward problem may greatly increase the efficiency of the calculations without significantly affecting the results of the optimization process.
- Asymptotic solutions to strongly nonlinear forward problems may lead to inverse solutions that retain the main features of solutions to the exact problem.
- Depending upon the nature of the problem, there may be advantages in combining several inverse problems into one or in separating a single inverse problem into several.

Due to the effects of nonlinearity and noise, optimization in geophysical inverse problems often involves a function that has numerous local extrema. Thus there is a need for optimization methods that search for a global extremum. Important aspects of this global optimization problem are:

- There are deterministic search methods for finding more global solutions to the optimization problem, but they need further testing on geophysical problems. Examples are:
  - smoothing and continuation methods.
  - TRUST.
  - pattern search methods.
  - computational algebra methods.
- The dimension of the model space for typical geophysical inverse problems is usually sufficiently large so that grid-search and Monte-Carlo methods are not practical.
- Directed stochastic search methods, such as simulated annealing and evolutionary methods, are being successfully used for some problems of this type, but they are not very efficient.
- A combination of stochastic and deterministic search methods might achieve the dual objectives of being both global and efficient.

- Finding a global extremum does not necessarily mean the problem is solved, because, primarily due to the effects of noise, the global solution may not be the best solution.

The fundamental non-uniqueness of geophysical inverse problems means that appraisal of the results of the optimization process is an essential part of a complete solution. Relevant aspects of this stage of the process are:

- Appraisal can be the most computer-intensive part of the inversion process and is often neglected or approximated because of this.
- The proper treatment of both resolution and uncertainty remain as essentially unsolved issues for nonlinear inverse problems.
- Methods of incorporating resolution and uncertainty more directly into the optimization process are desirable.

All of the issues mentioned above would benefit from more interaction between the optimization community and the geophysical community. Ways of facilitating this include:

- Information about geophysical inverse problems should be more readily available, preferably on the internet. Efforts that would help in this area include:
  - Availability of canonical geophysical inverse problems, including trial sets of data, which could be used to test and benchmark various optimization algorithms.
  - Availability of well documented codes, both for forward and inverse problems.
  - Construction of an optimization guide for geophysical inverse problems.
- Geophysicists should become more familiar with the broad range of optimization algorithms that are available. Keeping abreast of developments in both geophysics and optimization is not a simple task, but it could be helped by:
  - More use of the optimization software and software guides that are available on the internet.
  - More collaboration between geophysicists and mathematicians working on the same inverse problem.

References

1. Aki, K., and P. G. Richards, *Quantitative Seismology*, W. H. Freeman, San Francisco, 932 pp., 1980.

2. Alumbaugh, D. L., and G. A. Newman, Three-dimensional massively parallel electromagnetic inversion - II. Analysis of a crosswell electromagnetic experiment, *Geophys. J. Int.*, **128**, 355–363, 1997.

3. Audet, C., and J. E. Dennis, Analysis of generalized pattern searches, *TR00-07* Department of Computational & Applied Mathematics, Rice University, Houston TX, 2000a.

4. Audet, C., and J. E. Dennis, A pattern search filter method for nonlinear programming without derivatives, *TR00-09*, Department of Computational & Applied Mathematics, Rice University, Houston TX, 2000b.

5. Audet, C., and J. E. Dennis, Pattern search algorithms for mixed variable programming, *TR99-02*, Department of Computational & Applied Mathematics, Rice University, Houston TX, 2000c.

6. Bäck, T., *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, 1996.

7. Barhen, J., V. Protopopescu, and D. Reister, TRUST: A deterministic algorithm for global optimization, *Science*, **276**, 1094–1097, 1997.

8. Berryman, J. G., Stable iterative reconstruction algorithm for nonlinear traveltime tomography, *Inverse Problems*, **6**, 21–42, 1990.

9. Berryman, J. G., Global extrema in traveltime tomography, p. 45–62 in *Computational Acoustics - Volume 2*, D. Lee, A. R. Robinson, and R. Vichnevetsky (eds.), Elsevier, 1993.

10. Berryman, J. G., Variational structure of inverse problems in wave propagation and vibration, p. 13–44 in *Inverse Problems in Wave Propagation*, G. Chavent, G. Papanicolaou, P. Sacks, and W. W. Symes (eds), Springer, New York, 1997.

11. Berryman, J. G., Analysis of approximate inverses in tomography I. Resolution analysis of common inverses, *Optimization and Engineering*, **1**, 87–115, 2000.

12. Berryman, J. G., and R. V. Kohn, Variational constraints for electrical impedance tomography, *Phys. Rev. Lett.*, **65**, 325–328, 1990.

13. Booker, A., P. Frank, D. Serafini, V. Torczon, and M. Trosset, A rigorous framework for optimization of expensive functions by surrogates, *Structural Optimization*, 17(1), 1–13, 1999.

14. Borcea, L., Asymptotic analysis of quasi-static transport in high contrast conductive media, *SIAM J. Appl. Math.*, **59**, 597–635, 1999.

15. Borcea, L., J. G. Berryman, and G. C. Papanicolaou, High contrast impedance tomography, *Inverse Problems*, **12**, 835–858, 1996.

16. Borcea, L., J. G. Berryman, and G. C. Papanicolaou, Matching pursuit for imaging high contrast conductive media, *Inverse Problems*, **15**, 811–849, 1999.

17. Borcea, L., and G. C. Papanicolaou, Network approximation for transport properties of high contrast materials, *SIAM J. Appl. Math.*, **58**, 501–539, 1998.

18. Bosch, M., C. Barnes, and K. Mosegaard, Multi-step samplers for improving efficiency in probabilistic geophysical inference, p. 50 – 67 in *Methods and Applications of Inversion*, P. Hansen, B. H. Jacobsen, and K. Mosegaard (eds), Springer, Berlin, 2000.

19. Boschetti, F., M. C. Dentith, and R. D. List, Inversion of seismic refraction data using genetic algorithms, *Geophysics*, **61**, 1715–1727, 1996.

20. Conn, A. R., K Scheinberg, and P. L. Toint, On the convergence of derivative-free methods for unconstrained optimization, In M. D. Buhmann and A. Iserles, editors, *Approximation Theory and Optimization*, pages 83–108. Cambridge University Press, Cambridge, 1997.

21. Conn, A. R., N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, MPS-SIAM Series on Optimization, Vol 1&2 , Philadelphia, 2000.

22. Constable, S. C., R. L. Parker, and C. G. Constable, Occam's inversion: a practical algorithm for generating smooth models from EM sounding data, *Geophysics*, **52**, 289–300, 1987.

23. Cryzyk, M., P. Mesnier, and J. J. Moré, *The Network-Enabled Optimization System (NEOS) Server*, Preprint MCS-P615-0996, Argonne National Laboratory, Argonne, Illinois, 1996.

24. deGroot-Hedlin, C., and S. Constable, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, **55**, 1613–1624, 1990.

25. deGroot-Hedlin, C. D., and F. L. Vernon, An evolutionary programming method for estimating layered velocity structure, *Bull. Seism. Soc. Am.*, **88**, 1023–1035, 1998.

26. Dennis, J. E., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 378 pp., 1996.

27. Dennis, J. E., M. El-Alem, and M. C. Maciel, A global convergence theory for general trust-region-based algorithms for equality constrained optimization, *SIAM J. Optim.*, **7**, 177–207, 1997.

28. Dennis, J. E., M. Heinkenschloss, and L. N. Vicente, Trust–region interior–point algorithms for a class of nonlinear programming problems, *SIAM J. Control and Optimization*, **36**, 1750–1794, 1998.

29. Dennis, J. E., and V. Torczon, Direct search on parallel machines, *SIAM J. Optim.*, **1**, 448–474, 1991.

30. Dennis, J. E., and L. N. Vicente, On the convergence theory of trust-region-based algorithms for equality-constrained optimization, *SIAM J. Optim.*, **7**, 927–950, 1997.

31. Dobson, D. C., and F. Santosa, Recovery of blocky images from noisy and blurred data, *SIAM J. Appl. Math.* **56**, 1181–1198, 1996.

32. Dorn, O., A transport-backtransport method for optical tomography, *Inverse Problems* **14**, 1107–1130, 1998.

33. Dorn, O., H. Bertete-Aguirre, J. G. Berryman, and G. C. Papanicolaou, A nonlinear inversion method for 3D-electromagnetic imaging using adjoint fields, *Inverse Problems* **15**, 1523–1558, 1999.

34. Efron, B., *The Jackknife, the Bootstrap, and Other Resampling Plans*, SIAM, Philadelphia, 1982.

35. Everett, M. E., Homotopy, polynomial equations, gross boundary data, and small Helmholtz systems, *J. Comp. Physics*, **124**, 431–441, 1996.

36. Everett, M. E., and A. Schultz, Two-dimensional nonlinear magnetotelluric inversion using a genetic algorithm, *J. Geomag. Geoelectr.*, **45**, 1013–1026, 1993.

37. Fletcher, R., *Practical Methods of Optimization*, Wiley, New York, 436 pp., 1987.

38. Fogel, L. J., Autonomous automata, *Industrial Res.*, **4**, 14–19, 1962.

39. Fogel, D. B., *Evolutionary Computation*, IEEE Press, 1995.

40. Ghattas, O., and J.-H. Bark, Optimal control of two- and three-dimensional Navier–Stokes flow, *J. Comp. Physics*, **136**, 231–244, 1997.

41. Gockenbach, M. S., M. J. Petro, and W. W. Symes, C++ classes for linking optimization with complex simulation, *ACM Transaction on Mathematical Software*, **25**, 191–212, 1999.

42. Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

43. Gritto, R., V. A. Korneev, and L. R. Johnson, Nonlinear three-dimensional inversion of low-frequency scattered elastic waves, *Pure. Appl. Geophys.*, **156**, 557–589, 1999.

44. Haber, E., and D. Oldenburg, Joint inversion: a structural approach, *Inverse Problems*, **13**, 63–77, 1997.

45. Hansen, P. C., M. Jacobsen, J. M. Rasmussen, and H. Sorensen, The PP-TSVD algorithm for image restoration problems, p. 171 – 186 in *Methods and Applications of Inversion*, P. Hansen, B. H. Jacobsen, and K. Mosegaard (eds), Springer, Berlin, 2000.

46. Heinkenschloss, M., Mesh independence for nonlinear least squares problems with norm constraints, *SIAM J. Optimization*, **3**, 81–117, 1993.

47. Heinkenschloss, M., and L. N. Vicente, An interface between optimization and application for the numerical solution of optimal control problems, *ACM Transactions on Mathematical Software*, **25**, 157–190, 1999a.

48. Heinkenschloss, M., and L. N. Vicente, Analysis of inexact trust–region SQP algorithm, *TR99–18*, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, http://www.caam.rice.edu/, 1999b.

49. Holland, J., *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, Ann Arbor, 1975.

50. Hough P. D., T. G. Kolda, and V. Torczon, Asynchronous parallel pattern search for nonlinear optimization, Sandia National Laboratories, Report SAND2000-8213, 2000.

51. Jackson, D. D., Interpretation of inaccurate, insufficient, and inconsistent data, *Geophys. J. Roy. AstronSoc.*, **28**, 97–109, 1972.

52. Kennett, B. L. N., and G. Nolet, Resolution analysis for discrete systems, *Geophys. J. R. Astr. Soc.*, **53**, 413–425, 1978.

53. Kennett, B. L. N., and M. S. Sambridge, Earthquake location – genetic algorithms for teleseisms, *Phys. Earth Planet. Interiors*, **75**, 103–110, 1992.

54. Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science*, **220**, 671–680, 1983.

55. Lee, W. H. K., and S. W. Stewart, *Principles and Applications of Microearthquake Networks*, Academic Press, New York, 293 pp., 1981.

56. Lenhart, S., V. Protopopescu, and J. Yong, Solving inverse problems of identification type by optimal control methods, p. 87–94 in *Applied Nonlinear Dynamics and Stochastic Systems Near the Millennium*, J. B. Kadtke and A. Bulsara (eds), AIP Conference Proceedings 411, 1997.

57. Lévêque, J.-J., L. Rivera, and G. Wittlinger, On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, **115**, 313–318, 1993.

58. Lewis R. M., and V. Torczon, Rank ordering and positive basis in pattern search algorithms, *ICASE NASA Langley Research Center*, TR 96-71, 1996.

59. Lewis R. M., and V. Torczon 1998, Pattern search methods for linearly constrained minimization, *ICASE NASA Langley Research Center*, TR 98-3, 1998a, to appear in *SIAM Journal on Optimization.*

60. Lewis R. M., and V. Torczon, A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds, *ICASE NASA Langley Research Center*, 1998b.

61. Lewis R. M., and V. Torczon, Pattern search algorithms for bound constrained minimization, *SIAM Journal on Optimization*, Vol.9 No.4, 1082-1099, 1999.

62. Lewis, R. M., V. Torczon, and M. W. Trosset, Why pattern search works, *OPTIMA*, **59**, 1–7, October, 1998.

63. Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, New York, 1989.

64. Minkoff, S. E., A computationally feasible approximate resolution matrix for seismic inverse problems, *Geophys. J. Int.*, **126**, 345–359, 1997.

65. Minster, J.-B. H., N. P. Williams, T. G. Masters, J. F. Gilbert, and J. S. Haase, Applications of evolutionary programming to earthquake hypocenter determination, *Proc. of the Fourth Annual Conference of Evolutionary Programming*, J. R. McDonnell, R. G. Reynolds, and D. B. Fogel (eds), 3–17, 1995.

66. Moré, J. J., and S. J. Wright, *Optimization Software Guide*, SIAM, Philadelphia, 154 pp., 1993.

67. Moré, J. J., and Z. Wu, Global continuation for distance geometry problems, *SIAM J. Optim.*, **7**, 814–836, 1997.

68. Mosegaard, K., Resolution analysis of general inverse problems through inverse Monte Carlo sampling, *Inverse Problems*, **14**, 405–426, 1998.

69. Mosegaard, K., and A. Tarantola, Monte-Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, **100**, 12431–12447, 1995.

70. Mosegaard, K., and P. D. Vestergaard, A simulated annealing approach to seismic model optimization with sparse prior information, *Geophys. Prospecting*, **39**, 599–611, 1991.

71. Newman, G. A., and D. L. Alumbaugh, Three-dimensional massively parallel electromagnetic inversion - I. Theory, *Geophys. J. Int.*, **128**, 345–354, 1997.

72. Nocedal, J., and S. J. Wright, *Numerical Optimization*, Springer, New York, 1999

73. Nolet, G., Damped least squares methods to solve large tomographic systems and determine the resolving power, *Terra Cognita*, **4**, 285, 1984.

74. Nolet, G., Solving or resolving inadequate and noisy tomographic systems, *J. Comp. Phys.*, **61**, 463–482, 1985.

75. Nolte, B., and L. N. Frazer, Vertical seismic profile inversion with genetic algorithms, *Geophys. J. Int.*, **117**, 162–178, 1994.

76. Oldenburg, D. W., One-dimensional inversion of natural source magnetotelluric observations, *Geophysics*, **44**, 1218–1244, 1979.

77. Oldenburg, D. W., Y. Li, C. F. Farquharson, P. Kowalczyk, T. Aravanis, A. King, P. Zhang, and A. Watts, Applications of geophysical inversions in mineral exploration, *The Leading Edge*, 461–465, 1998.

78. Paige, C. C., and M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.*, **8**, 43–71, 1982.

79. Parker, P. B., *Genetic Algorithms and Their Use in Geophysical Problems*, Ph.D. Thesis, University of California, Berkeley, 202 pp., 1999.

80. Parker, R. L., *Geophysical Inverse Theory*, Princeton University Press, 386 pp., 1994.

81. Rojas, M., and D. Sorensen, A trust–region approach to the regularization of large–scale discrete ill-posed problems, *TR99–26*, Department of Computational and Applied Mathematics, Rice University, 1999.

82. Rothman, D. H., Nonlinear inversion, statistical mechanics, and residual statics estimation, *Geophysics*, **50**, 2784–2796, 1985.

83. Rothman, D. H., Automatic estimation of large residual statics correction, *Geophysics*, **51**, 332–346, 1986.

84. Rudin, L. I., S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, **60**, 259–268, 1992.

85. Sambridge, M., and G. Drijkoningen, Genetic algorithms in seismic waveform inversion, *Geophys, J. Int.*, **109**, 323–342, 1992.

86. Sambridge, M., and K. Gallagher, Earthquake hypocenter location using genetic algorithms, *Bull. Seism. Soc. Am.*, **83**, 1467–1491, 1993.

87. Scales, J. A., M. L. Smith, and T. L. Fischer, Global optimization methods for multimodal inverse problems, *J. Comput. Phys.*, **103**, 258–268, 1992.

88. Sen, M. K., and P. L. Stoffa, Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics*, **56**, 1624–1638, 1991.

89. Sen, M. K., and P. L. Stoffa, Rapid sampling of model space using genetic algorithms: examples from seismic waveform inversion, *Geophys, J. Int.*, **108**, 281–292, 1992.

90. Sen, M. K., and P. L. Stoffa, *Global Optimization Methods in Geophysical Inversion*, Elsevier, Amsterdam, 281 pp., 1995.

91. Sen, M. K., and P. L. Stoffa, Bayesian inference, Gibbs sampler and uncertainty estimation in geophysical inversion, *Geophys. Prospect.*, **44**, 313–350, 1996.

92. Sluis, A. van der, and H. A. van der Vorst, Numerical solution of large, sparse linear algebraic systems arising from tomographic problems, p. 49–83 in *Seismic Tomography*, G. Nolet (ed.), D. Reidel, Dordrecht, 1987.

93. Stephens C. P. and W. Baritompa, Global optimization requires global information, *Journal of Optimization Theory and Applications*, **96**, 575–588, 1998.

94. Stoffa, P. L., and M. K. Sen, Nonlinear multiparameter optimization using genetic algorithms: inversion of plane-wave seismograms, *Geophysics*, **56**, 1794–1810, 1991.

95. Tarantola, A., *Inverse Problem Theory*, Elsevier, 1987.

96. Tarantola, A., Probabilistic foundations of inverse theory, in *Geophysical Tomography*, Y. Desaubies, A. Tarantola, and J. Zinn-Justin (eds.), North Holland, 1990.

97. Tolstoy, A., O. Diachok, and L. N. Frazer, Acoustic tomography via matched field processing, *J. Acoust. Soc. Am.*, **89**, 1119–1127, 1991.

98. Torczon, V., On the convergence of pattern search algorithms, *SIAM J. Optim.*, **7**, 1–25, 1997.

99. Torczon, V., and M. W. Trosset, Using approximations to accelerate engineering design optimization, AIAA Paper 98-4800, Proc. 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, St. Louis, Sept. 2–4, 1998.

100. Torczon, V., and M. W. Trosset, From evolutionary operation to parallel direct search: pattern search algorithms for numerical optimization, Computing Science and Statistics, 29(1), Proceedings of the 29th Symposium on the Interface, D. W. Scott (ed), Houston, TX, May, 1997.

101. Trampert, J. Global seismic tomography: the inverse problem and beyond, *Inverse Problems*, **14**, 371–385, 1998.

102. Varela, C., P. L. Stoffa, and M. K. Sen, Background velocity estimation using nonlinear optimization for reflection tomography and migration misfit, *Geophys. Prospect.*, **46**, 51–78, 1998.

103. Vasco, D. W., Intersections, ideals, and inversion, *Inverse Problems*, **15**, 1573–1602, 1999.

104. Vasco, D. W., An algebraic formulation of geophysical inverse problems, *Geophys. J. Int.*, **142**, 970–990, 2000.

105. Vogel, C. R., and M. E. Oman, Iterative methods for total variation denoising, *SIAM J. Sci. Computing*, **17**, 227–238, 1996.

106. Wiggins, R. A., The general linear inverse problem: Implications of surface waves and free oscillations of Earth structure, *Rev. Geophys. Space Phys.*, **10**, 251–285, 1972.

107. Wolpert, D. H., and W. G. Macready, No free lunch theorems for optimization, *IEEE Trans. Evol. Prog.*, 1, 67–82, 1997.

108. Xia, G., M. K. Sen, and P. L. Stoffa, 1D elastic waveform inversion: A divide and conquer approach, *Geophysics*, **63**, 1670–1684, 1998.

109. Zhang, J., and G. A. McMechan, Estimation of resolution and covariance for large matrix inversions, *Geophys. J. Int.*, **121**, 409–426, 1995.