JOURNAL OF COMPUTATIONAL PHYSICS **124,** 431–441 (1996)
ARTICLE NO. 0070

Homotopy, Polynomial Equations, Gross Boundary Data, and Small Helmholtz Systems

M. E. EVERETT*

Institute of Theoretical Geophysics, Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, United Kingdom

Received March 22, 1994; revised August 3, 1995

Inverse problems of the boundary measurement type appear in several geophysical contexts including DC resistivity, electromagnetic induction, and groundwater flow. The objective is to determine a spatially varying coefficient in a partial differential equation from incomplete knowledge of the dependent variable and its normal gradient at the boundary. Equivalent 2D discrete inverse problems based on the Helmholtz or modified Helmholtz equation reduce to systems of polynomial equations indicating that there are only a finite number of exact solutions, excluding certain pathological cases. A homotopy procedure decides whether real, positive solutions exist and, if so, generates the entire list. The computational complexity of the algorithm scales as $M^{M/2}$, where M is the number of model parameters to be found. Measurement errors are accommodated by oversampling the boundary data at additional frequencies. For test Helmholtz and modified Helmholtz inverse problems based on (i) perfect and (ii) noisy data I generate the full list of exact solutions. The homotopy approach applies to large scale, multidimensional geophysical inverse problems but at present is practical only for small systems, up to $M = 9$. Recent advances in homotopy theory should, however, reduce the complexity, making larger problems tractable in the future. © 1996 Academic Press, Inc.

1. INTRODUCTION

In this paper I consider the inverse problem of recovering a spatially varying coefficient in Helmholtz and modified Helmholtz equations from possibly imprecise measurements made at the boundary. There are several practical problems in geophysics which may be formulated in this way. The electromagnetic and DC resistivity methods of geophysical prospecting, for example, consist of inducing or directly injecting an electric current into the conducting ground and measuring the resulting surface electromagnetic fields [11]. The surface measurements constitute the data for the inverse problem and are influenced by the underlying electrical conductivity. The latter is a coefficient in Maxwell equations governing electromagnetism in the earth. Ore bodies, inorganic contaminant plumes at hazardous waste sites, and structural features such as faults may be mapped using the information on subsurface electrical conductivity provided by the data.

Inverse problems with boundary measurements also have a medical application. In a technique known as impedance tomography [8], current is injected into the human body and voltages are measured. Like the analogous geophysical data, the voltages are sensitive to the distribution of electrical conductivity within the body, which is an indicator of blood flow and the health of organs. Finally, there is a close relationship between the above problems and that of determining from spectral data the inhomogeneous mass distribution of a vibrating string or membrane. This is the classic inverse problem that once prompted the mathematician Mark Kac to ask [2] "Can one hear the shape of a drum?"

In this paper I show that a 2D discrete Helmholtz inverse problem with enough boundary measurements reduces to a well-determined system of polynomial equations. Such polynomial systems have only a finite number of geometrically isolated solutions. I show that a homotopy procedure can be used to decide whether physical solutions exist and, if so, to generate the entire solution list. The procedure applies to the solution of large scale problems but is tenable computationally at present only for small inverse problems. I present results from numerical studies of test Helmholtz and modified Helmholtz inverse problem with perfect data which confirm the theory. Finally, I describe a framework for accommodating noisy data and present additional test results. It is important to note that in this paper I seek only exact solutions to the inverse problem. This is in contrast to most practical geophysical inversion algorithms which try to find an optimal physical solution amongst all with a given, nonzero misfit to the data. There have been a few other applications of homotopy techniques to inverse problems [16, 17].

2. BACKGROUND

Inverse problems are ill-posed so that the unique recovery from imperfect boundary measurements of a spatially varying coefficient in a differential equation is generally

* Current address: Department of Geology and Geophysics, Texas A&M University, College Station, TX 77843.

0021-9991/96 $18.00

impossible. Uniqueness theorems, where available, do, however establish what amount of boundary data is necessary and sufficient for unique coefficient reconstruction. The theorems are of use primarily to indicate what inverse theory is able to achieve in principle, if not in practice. Uniqueness results have been derived for the modified Helmholtz equation governing electromagnetic induction in a 1D earth [15], the 3D direct current electrical prospecting equation [10], and the 3D, frequency-dependent reduced acoustic equation [6]. Note that if there are insufficient boundary data, the inverse problem can have any number of solutions.

Unique recovery requires perfect data because the stochastic measurement noise is not modeled by the underlying, deterministic partial differential equation. In fact, when noise is present, an exact solution to the inverse problem may not exist at all. Consider the example from DC resistivity. Suppose that the true subsurface electrical conductivity distribution is known exactly everywhere and, furthermore, that Ohm's law provides a precise physical description of current flow in conducting media. The voltage response predicted by Ohm's law would not match the observations simply because Ohm's law is independent of the inherent noise in the measurement apparatus. In this case the true conductivity distribution within the ground does not exactly solve the inverse problem.

Apart from a uniqueness theorem, what are the essential ingredients of a complete theory for solving multidimensional inverse problems of the boundary measurement type? My list follows. It reflects a personal bias towards the inverse of multidimensional elecromagnetic induction data.

(a) *Direct inversion to nearly machine precision.* Globally convergent algorithms which produce highly accurate solutions to the inverse problem by performing a finite number of well-defined mathematical operations directly on the boundary data.

(b) *Existence theory*. Algorithms which decide after a finite number of floating point operations whether a solution to the inverse problem exists. Algorithms to determine exactly how many solutions exist.

(c) *Full solution list.* Algorithms which generate the full list of solutions to the inverse problem in a finite number of steps. If there are infinitely many solutions, a guarantee that every solution can be constructed, although it would require an infinite amount of CPU time.

(d) *Noisy, sparse data.* A framework for accommodating sparse and noisy data.

(e) *Globally optimal solutions.* Globally convergent algorithms which examine the solution space and return the member which optimizes an arbitrary functional, such as least squares misfit or smoothness, provided by the user.

(f) *Regularity of data with respect to the model.* Algorithms to decide whether small changes in the model necessarily produce small changes in the measured response. The stability of the inverse problem is closely related to Frechet differentiability. A few results in this direction are available for mulidimensional nonlinear inverse problems but this is not of direct concern to this paper.

Slow but steady progress is being made on the above fronts [7, 12]. Practical algorithms for the solution of large-scale nonlinear inverse problems, however, tend by necessity to be linearized, iterative searches for isolated, locally optimal models [9]. These algorithms often involve repeated solutions to the forward problem and sometimes rely on unproven Frechet differentiability. Practical algorithms have produced results contributing to our understanding of the Earth's interior, yet the relationship between the true structure and a local minimum of an *ad hoc* functional is not clear.

In this paper, I describe discrete, 2D Helmholtz and modified Helmholtz inverse problems and show that goals (a)–(e) are readily attainable. The "catch," however, is that my method currently works only for small discrete inverse problems so that scaling up the results to the level of practical geophysical interest is not yet feasible. However, the theory described here is a significant step towards characterising the solution space of large scale, nonlinear, multidimensional inverse problems.

3. THE CONTINUOUS FORWARD THEORY

Consider the 2D differential equation

$$\nabla^2 u + q\sigma(x, z)u = 0, \quad (x, z) \text{ in } \Omega \in \mathscr{R}^2, \tag{1}$$

where Ω is some open subset of the (x, z)-plane. If the parameter q is real, Eq. (1) is the Helmholtz equation describing, for example, vibrating membranes. If q is imaginary, it is the modified Helmholtz equation describing monochromatic, steady state ($e^{i\omega t}$) electromagnetic induction in the earth. The function $\sigma(x, z)$ is real and positive, on physical grounds, and represents either the mass distribution of the membrane or the electrical conductivity of the earth. The function $u(x, z)$ corresponds to the membrane displacement or the electric field.

The forward problem, find u given σ, can be solved uniquely if Dirichlet (the function u) or Neumann (its normal derivative $\partial u/\partial n$) data are specified or measured everywhere along the boundary $\partial\Omega$. However, this is not enough information to solve the associated inverse problem: find σ given either $u|_{\partial\Omega}$ or $\partial_n u|_{\partial\Omega}$. To solve the inverse problem it is necessary to overspecify the boundary data. At some locations on the boundary *both* Dirichlet- and Neumann-type data must be known. At any location, there must be at least one or the other. For definiteness, I assume that Dirichlet data $u = f(x, z)$ are given everywhere along the boundary and that additional Neumann measurements $\partial_n u = g(x, z)$ are made along some portion $\Gamma \subset \partial\Omega$. I assume initially that the functions f and g are known precisely. Later I will admit the possibility that they might contain measurement error.

The forward problem (1) is related to a pair of modified Helmholtz problems of interest to the electromagnetic induction community, i.e., TE and TM mode electromagnetic induction in a 2D conducting earth. The associated TE and TM mode inverse problems bear directly on the interpretation of magnetotelluric (MT) data [13]. The TE mode forward problem, for example, consists of the Laplace and modified Helmholtz equations

$$\nabla^2 u = 0, \quad (x, z) \in \text{air} \tag{2}$$

$$\nabla^2 u + i\omega\mu_0\sigma(x, z)u = 0, \quad (x, z) \in \text{earth} \tag{3}$$

which are valid in the semiinfinite domain ($|x| \to \infty$, $z_1 \leq z \leq z_2$). The vertical coordinate z increases downwards and $z = 0$ is the surface of the earth. The boundary conditions are

$$\partial_z u(x, z_1) = i\omega\mu_0 H_0, \tag{4}$$

$$u(\pm L, z < 0) = 1 - i\omega\mu_0 H_0 z, \tag{5}$$

$$u(\pm L, z > 0) = u^{1\mathrm{D}}(z), \tag{6}$$

$$\partial_z u(x, z_2) + i\sqrt{i\omega\mu_0\sigma_T}\,u(x, z_2) = 0, \tag{7}$$

where μ_0 is the magnetic permeability of free space; ω is angular frequency; z_1 and z_2 are the z-coordinates of the top of the air layer and bottom of the earth region, respectively; σ_T is the conductivity of the terminating halfspace beneath the earth region; and H_0 is the strength of the magnetic field at the top of the air layer. The parameter L is the half-width of the solution domain, which is centered about $x = 0$. The electrical conductivity $\sigma(x, z)$ model in 2D MT numerical modeling is a confined, 2D heterogeneity located at or near the center of the solution domain, set into a background plane layered structure. The electric field $u(x = \pm L, z)$ at the domain edges is set to the 1D analytic solution $u^{1\mathrm{D}}(z)$ appropriate for the stratified structure there. In princple $L \to \infty$, but for numerical work L is simply chosen large enough that the electric field is not appreciably influenced by the presence of the central heterogeneity. The data for the inverse problem consist of geophysical measurements of u, $\partial_z u$, or the impedance ratio $u/\partial_z u$ on the surface $z = 0$. There is a similar formulation for the TM mode.

The inverse problem based on (2)–(7) is difficult to solve. I therefore consider a more tractable inverse problem based on (1). The problem that initially I consider describes the behaviour of nondissipative, vibrating membranes. It is the Helmholtz, or wave, Eq. (1) in a rectangle with Dirichlet data $u = f(x, z)$ prescribed on the periphery. Additional Neumann measurements $\partial_n u = g(x)$ are made along the top of the rectangle. Since this Helmholtz problem is still too difficult to solve when x, z are continuous variables, I consider a finite difference approximation (FDA). The resulting, discrete inverse problem proves tractable. I repeat the analysis for a modified Helmholtz inverse problem by letting the frequency parameter q be purely imaginary. The modified Helmholtz, or time-independent diffusion, equation describes electromagnetic induction in the earth. The study of its inverse problem should be viewed as a step towards resolving the more complicated MT inverse problem.

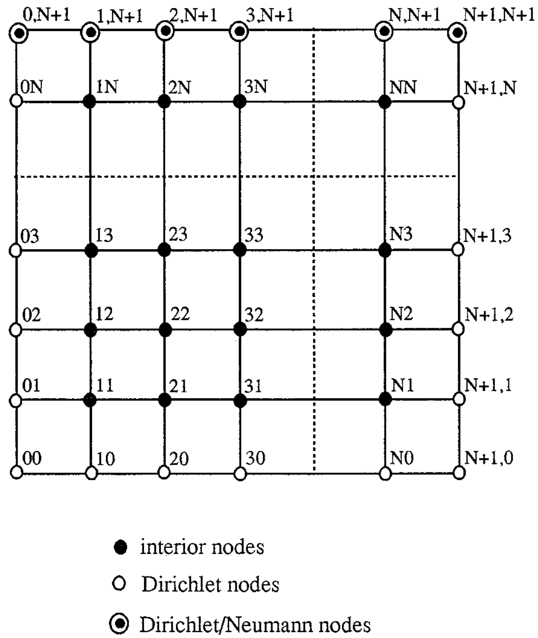

FIG. 1. The finite difference mesh with $N \times N$ interior nodes on which the Helmholtz equation is solved. Dirichlet nodes are those on which Dirichlet data are prescribed. Dirichlet/Neumann nodes are those on which Dirichlet data are prescribed and Neumann measurements have been made.

4. THE DISCRETE FORWARD THEORY

I now construct a standard FDA to the overspecified Helmholtz boundary value problem just described. The FDA is based on the uniform finite difference mesh shown in Fig. 1 and consists of the linear system

$$A\mathbf{u} = \mathbf{b}, \tag{8}$$

where the matrix A is given by the $N \times N$ block tridiagonal matrix

$$A = \begin{pmatrix} \mathbf{A}_1 & \mathbf{I} & 0 & 0 \\ \mathbf{I} & \mathbf{A}_2 & \mathbf{I} & 0 \\ & & \ddots & \\ 0 & 0 & \mathbf{I} & \mathbf{A}_N \end{pmatrix}. \tag{9}$$

In Eq. (9), $\mathbf{I}$ is simply the $N \times N$ identity matrix while $\mathbf{A}_i$ for $i = 1, ..., N$ are tridiagonal matrices given by

$$A_i = \begin{pmatrix} -4 + q\sigma_{1i}h^2 & 1 & 0 & \\ 1 & -4 + q\sigma_{2i}h^2 & 1 & \\ & & \ddots & \\ & 0 & 1 & -4 + q\sigma_{Ni}h^2 \end{pmatrix},$$
$$i = 1, ..., N. \tag{10}$$

In Eq. (10), σ_{jk} is the value of the unknown function $\sigma(x, z)$ on the interior node labelled jk (see Fig. 1), and h is the distance between adjacent nodes.

The right-hand side vector $\mathbf{b}$ in Eq. (8) has the block form $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_N)^T$, where the vectors $\mathbf{b}_i$ for $i = 1, ..., N$ are each of length N. In particular, for $i = 1$ and $i = N$ they are given by

$$\mathbf{b}_1 = \begin{pmatrix} -f_{10} - f_{01} \\ -f_{20} \\ -f_{30} \\ \vdots \\ -f_{N0} - f_{N+1,1} \end{pmatrix}, \quad \mathbf{b}_N = \begin{pmatrix} -f_{1,N+1} - f_{0N} \\ -f_{2,N+1} \\ -f_{3,N+1} \\ \vdots \\ -f_{N,N+1} - f_{N+1,N} \end{pmatrix}. \tag{11}$$

For the remaining values $i = 2, ..., N - 1$ they are $\mathbf{b}_i = (-f_{0i}, 0, ..., 0, -f_{N+1,i})^T$. The right-hand side vector $\mathbf{b}$ is the Helmholtz source vector and clearly it is determined by the prescribed Dirichlet function $f(x, z)$, which takes the value f_{jk} on the exterior node jk.

The solution vector $\mathbf{u}$ in (8) has a similar block form $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_N)^T$, but where $\mathbf{u}_i = (u_{1i}, u_{2i}, ..., u_{Ni})^T$ for $i = 1, ..., N - 1$ and

$$\begin{aligned} &\mathbf{u}_N \equiv \mathbf{m} = (m_1, m_2, ..., m_N)^T \\ &\text{with } m_i = (f_{i,N+1} + hg_i). \end{aligned} \tag{12}$$

The unknown electric field $u(x, z)$ on interior node jk is denoted by u_{jk}. From (12), it is clear that the last N entries of the solution vector $\mathbf{u}$ are determined by both Dirichlet and the auxiliary Neumann measurements $g(x)$ made along the top of the rectangular domain. Specifically, the Neumann measurement at exterior node $\{i, N + 1\}$, denoted by g_i for $i = 1, ..., N$, is related to the electric field at the adjacent interior node $\{i, N\}$ by the first differences formula given at the right-hand side of expression (12). The utility of the Neumann boundary data for solving the inverse problem is now apparent. The Neumann data provide some knowledge of the solution vector $\mathbf{u}$ to the linear system (8) governing the discrete inverse problem. Without partial knowledge of $\mathbf{u}$, the inverse problem cannot be solved.

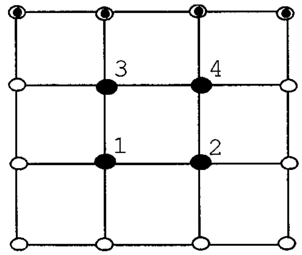

FIG. 2. The finite difference mesh with $N = 2$ showing the locations and revised numbering scheme for the model parameters $\{\sigma_i\}_{i=1}^4$.

5. POLYNOMIAL EQUATIONS

I now indicate the reduction of the discrete Helmholtz inverse problem with boundary measurements to a set of polynomial equations. To keep the exposition particularly simple I specialize to the case $N = 2$. The theory for general N follows later. The matrix A in the case $N = 2$ is the 2×2 block matrix

$$A = \begin{pmatrix} -4 + q\sigma_1 h^2 & 1 & 1 & 0 \\ 1 & -4 + q\sigma_2 h^2 & 0 & 1 \\ 1 & 0 & -4 + q\sigma_3 h^2 & 1 \\ 0 & 1 & 1 & -4 + q\sigma_4 h^2 \end{pmatrix}, \tag{13}$$

where the interior nodes of the mesh are re-labelled according to Fig. 2 so that singly subscripted indices suffice to denote the unknown model parameters, i.e., $\sigma = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$.

The solution vector in (8) for $N = 2$ is $\mathbf{u} = (u_1, u_2, m_1, m_2)^T$, where u_1, u_2 are discrete values of $u(x, z)$ on the interior nodes labelled 1 and 2 (see Fig. 2). The quantities m_1, m_2 are derived from the Neumann data, as described above. Equation (8) for $N = 2$ represents the four bilinear equations in six unknowns,

$$\begin{aligned} (-4 + q\sigma_1 h^2)u_1 + u_2 + m_1 &= b_1 \\ u_1 + (-4 + q\sigma_2 h^2)u_2 + m_2 &= b_2 \\ u_1 + (-4 + q\sigma_3 h^2)m_1 + m_2 &= b_3 \\ u_2 + m_1 + (-4 + q\sigma_4 h^2)m_2 &= b_4, \end{aligned} \tag{14}$$

which are obtained by explicit multiplication. The bilinearity is due to the product σu which appears in the original Helmholtz equation (1). The six unknowns are the four nodal "conductivities," or σ-variables, in addition to the two internal "electric field," or u-variables, u_1 and u_2.

The next step is to eliminate the u-variables which are of little physical interest. A little algebra (specifically, the last two equations in (14) are solved for u_1 and u_2, respectively, and the resulting expressions are plugged into the first two equations) yields the following pair of bilinear equations in the unknown conductivities

$$\begin{aligned} A\sigma_1 + B\sigma_3 + C\sigma_4 + D\sigma_1\sigma_3 + E = 0 \\ F\sigma_2 + G\sigma_3 + H\sigma_4 + I\sigma_2\sigma_4 + J = 0. \end{aligned} \tag{15}$$

These are two equations in four unknowns. The coefficients $A \cdots J$, given by

$$\begin{aligned} &A = qh^2(b_3 + 4m_1 - m_2), \quad B = 4qh^2m_1, \\ &C = -qh^2m_2, \quad D = -q^2h^4m_1, \\ &E = -4b_3 - 16m_1 + 8m_2 + b_4 - b_1 \\ &F = qh^2(b_4 + 4m_2 - m_1), \quad G = -qh^2m_1, \\ &H = 4qh^2m_2, \quad I = -q^2h^4m_2, \\ &J = -4b_4 - 16m_2 + 8m_1 + b_3 - b_2 \end{aligned}$$

depend on the boundary data (f, g), the mesh interval h, and the frequency parameter q. The system (15) is underdetermined since there are fewer equations than unknowns. To make the system well-determined, I supply additional Neumann data at a second frequency $q = q_2 \neq q_1$. This gives two further bilinear equations of the form (15) without introducing any new unknowns. Hence there are now four bilinear equations in as many unknowns.

In general, for a single frequency q, a mesh containing $N \times N$ interior nodes leads to a set of N polynomial equations in $M = N^2$ unknown σ-variables. The procedure for generating these polynomial equations for arbitrary N is as follows. First, the penultimate N equations in (8) are solved for $\mathbf{u}_{N-1}$ in terms of the known vectors $\mathbf{b}_N$ and $\mathbf{m}$. This is straightforward since these equations do not involve any other unknown $\mathbf{u}$-vectors. Then, the resulting expression for $\mathbf{u}_{N-1}$ is used to derive an expression for $\mathbf{u}_{N-2}$ in terms of $\mathbf{b}_{N-1}$, $\mathbf{b}_N$, and $\mathbf{m}$. This procedure, which is nothing but a vector equivalent of back substitution, is repeated until finally $\mathbf{u}_1$ is computed in terms of $\{\mathbf{b}_i\}_{i=2}^N$ and $\mathbf{m}$. Then, the expressions for $\mathbf{u}_1$ and $\mathbf{u}_2$ are inserted into the first N equations of (8), namely $\mathbf{A}_1\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{b}_1$. These equations are the desired system of N-linear polynomial equations in the σ-variables, with all the u-variables eliminated. Since a single frequency generates N polynomial equations, a total of N distinct frequencies are needed to construct a well-determined system. Thus, for $N = 3$, boundary data at three frequencies are required to generate nine cubics. For $N = 4$, data at four frequencies generate 16 quartics, and so on. For convenience, I denote the system of polynomial equations by $\mathbf{F}(\sigma) = \mathbf{0}$.

The degree of each polynomial in $\mathbf{F}$ is equal to the number of interior nodes in a column or row of the finite difference mesh. The total degree d of a polynomial system is defined as the product of these degrees, i.e.,

$$d = \prod_{j=1}^{M} d_j, \tag{16}$$

where d_j is the degree of the jth polynomial equation $F_j = 0$. In our case, $d = N^M = M^{M/2}$ since each polynomial is N-linear. It is well known (Bezout's theorem, [4]) that a polynomial system has at most d geometrically isolated solutions. A solution is geometrically isolated if there is a neighbourhood containing that and only that solution. For example, a system of two quadratics has at most four isolated solutions. The bilinear Helmholtz system for $N = 2$ therefore has at most $2^4 = 16$ isolated solutions. For general N, the discrete Helmholtz inverse problem clearly has a finite number of isolated solutions, although their number grows rapidly with N. Is there any possibility of generating them?

To summarize, the introduction of enough Neumann-type measurements on a portion of the boundary reduces the 2D discrete Helmholtz inverse problem to a well-determined system of polynomial equations in the unknown σ-variables. It now remains to find the isolated solutions to such systems.

6. HOMOTOPY

Homotopy is a globally convergent strategy for finding solutions to nonlinear equations, of which a system of polynomial equations is a special case. Consider a general zero finding problem in M-dimensional Euclidean space $\mathscr{E}^M$, namely,

$$\text{Find } \sigma \in \mathscr{E}^M \text{ such that } \mathbf{F}(\sigma) = \mathbf{0},$$

where $\mathbf{F}\colon \mathscr{E}^M \to \mathscr{E}^M$ is a smooth map. This problem is equivalent to the associated fixed point problem

$$\text{Find } \sigma \in \mathscr{E}^M \text{ such that } \mathbf{F}^*(\sigma) = \sigma,$$

if the jth equation $F_j(\sigma)$ is solved for σ_j and written as $F_j^*(\sigma) = \sigma_j$. In the discussion that follows, suppose that $\mathbf{F}^*$ is a smooth nonlinear mapping $B \to B$, where B is the closed unit ball in $\mathscr{E}^M$. Define a vector function

$$\rho(\mu, \sigma) = \mu[\sigma - \mathbf{F}^*(\sigma)] + (1 - \mu)[\sigma - \sigma_0]. \quad (17)$$

Note that $\rho: [0, 1] \times B \rightarrow \mathscr{E}^M$. The essential homotopy theorem [1] for the fixed-point problem is sketched below. The theorem for the zero-finding problem is slightly more complicated but the essence is the same. For rigorous proofs of both theorems and the precise meaning of "smooth," the interested reader should consult the reference. General background material on normed vector spaces is found in [3].

THEOREM. *For almost all $\sigma_0 \in B$, there exits a smooth zero curve $\gamma \subset [0, 1] \times B$ of the function $\rho(\mu, \sigma)$ emanating from $(0, \sigma_0)$ and reaching $(1, \sigma^*)$, where σ^* is a fixed point of $\mathbf{F}^*(\sigma)$.*

I do not use this general theorem, rather I use a more specialized version for polynomial systems; however, the general theorem provides the reader with a view of the salient features of homotopy theory. It follows from the theorem that the homotopy method for solving the fixed point (and zero finding) problems is globally convergent with probability one. The term "probability one" stems from that fact that for *almost all* choices of starting point σ_0, in the sense of Lebesgue measure, tracking the zero curve γ from $\mu = 0$ to $\mu = 1$ leads to a fixed point σ^* of $\mathbf{F}^*$. Global convergence of the homotopy method contrasts, for example, with the local convergence of Newton's iterative method, whose success depends on a good initial guess σ_0. The zero curve γ is the trajectory of an initial value problem whose independent variable is arc length s. The mechanics of tracking $\gamma[s]$ is performed by an appropriate ODE solver [14]. Note that when $\mu = 0$, the zero of $\rho(0, \sigma)$ occurs when $\sigma = \sigma_0$, the starting point, and when $\mu = 1$, the zero of $\rho(1, \sigma)$ occurs when $\sigma = \mathbf{F}^*(\sigma)$, the required fixed point.

I will now specialize to systems of polynomial equations. Let $\mathscr{C}^M$ be complex M-dimensional Euclidean space containing vectors $a = (a_1, ..., a_M)$ and $b = (b_1, ..., b_M)$. Define a function

$$\rho(\mu, \sigma) = (1 - \mu)\mathbf{G}(\sigma) + \mu\mathbf{F}(\sigma), \quad (18)$$

where $\mathbf{G}$ and $\mathbf{F}$ are systems of polynomials. The following basic theorem [4] applies to the zero finding problem for polynomial equations.

THEOREM. *Let $\mathbf{G}: \mathscr{C}^M \rightarrow \mathscr{C}^M$ be given by $G_j(\sigma) = b_j\sigma^{d_j} - a_j$, for $j = 1, ..., M$, where a_j and b_j are nonzero complex numbers and d_j is the degree of $F_j(\sigma)$. For almost all $a, b \in \mathscr{C}^M$, the set of zero curves of ρ in (14) is d smooth paths emanating from $\{0\} \times \mathscr{C}^M$, which either diverge to infinity as μ approaches 1 or converge to solutions to $\mathbf{F}(\sigma) = 0$ as μ approaches 1. For this choice of $\mathbf{G}(\sigma)$, each geometrically isolated solution of $\mathbf{F}(\sigma)$ has a path converging to it.*

TABLE I

Boundary Data for Test Inverse Problems

(a) *Helmholtz*		
Frequency	$q_1 = 1.0$	$q_2 = 2.0$
Neumann data	$m_1 = -2.0$ $m_2 = -0.1$	$m_1 = -2.8$ $m_2 = 3.1$
Dirichlet data	$b_1 = 1.0$ $b_2 = 11.4$ $b_3 = -297.5$ $b_4 = -7.45$	$b_1 = 5.5$ $b_2 = -3.1$ $b_3 = -838.2$ $b_4 = 557.6$
(b) *Modified Helmholtz*		
Frequency	$q = 1.0$	
Newmann data	$m_1 = 3.8 + i\ 7.0$ $m_2 = 1.6 - i\ 0.2$	
Dirichlet data	$b_1 = -17.2 + i\ 10.$ $b_2 = 30.6 + i\ 25.3$ $b_3 = -1078.1 + i\ 553.3$ $b_4 = 16.9 + i\ 151.8$	

The software package HOMPACK [14] available from the Association for Computing Machinery (ACM) finds all complex solutions to polynomial systems by tracking the smooth paths in the zero set of the function ρ. Note that HOMPACK transforms the polynomial zero finding problem $\mathbf{F}(\sigma) = \mathbf{0}$ into a new system which has no solutions at infinity so that every path is bounded.

7. HELMHOLTZ TEST RESULTS: EXACT DATA

I used HOMPACK to obtain all 16 complex solutions to the polynomial system

$$\begin{aligned}
&-7635 \cdot \sigma_1 - 200 \cdot \sigma_3 + 2 \cdot 5\sigma_4 \\
&\qquad + 1250 \cdot \sigma_1\sigma_3 + 1212 \cdot 75 = 0 \\
&146 \cdot 25\sigma_2 + 50 \cdot \sigma_3 - 10 \cdot \sigma_4 \\
&\qquad + 62 \cdot 5\sigma_2\sigma_4 - 293 \cdot 5 = 0 \quad (19) \\
&-42625 \cdot \sigma_1 - -560 \cdot \sigma_3 - 155 \cdot \sigma_4 \\
&\qquad + 7000 \cdot \sigma_1\sigma_3 + 3974 \cdot 5 = 0 \\
&28640 \cdot \sigma_2 + 140 \cdot \sigma_3 + 620 \cdot \sigma_4 \\
&\qquad - 7750 \cdot \sigma_2\sigma_4 - 3137 \cdot 5 = 0.
\end{aligned}$$

This system of bilinear equations is equivalent to the discrete Helmholtz inverse problem for $N = 2$. The input data which generated this system are listed in Table I(a). The full solution list is given in Table II(a). The system (19) has exactly three real positive solutions (including the "truth", $\sigma = \{0.2, 0.3, 6.1, 3.7\}$), a real negative solution

TABLE II

Full Solution Lists

	σ_1	σ_2	σ_3	σ_4
(a) *Helmholtz*				
1	0.200	0.300	6.10	3.70
2	0.126	0.289	6.12	3.70
3	0.153	0.0861	6.62	5.38
4	−4.56	0.0797	6.11	0.0339
+12 infinite solutions				
(b) *Modified Helmholtz*				
1	0.200	0.300	6.10	3.70
2	0.0963	0.228	6.11	3.72
3	0.207	0.395	6.10	3.68
4	0.0635	0.519	6.11	3.66
+12 infinite solutions				

and 12 complex solutions at infinity. There are no other solutions to this inverse problem.

8. HELMHOLTZ TEST RESULTS: NOISY DATA

The theory described in the previous sections permits generation of the entire solution list when the boundary data are perfect. It is guaranteed that at least one solution exists, namely the "true" solution which generated the data. In reality, however, the boundary data are never known precisely. This is due not only to the unavoidable, inherent error in the measurement apparatus but also to the fact that the FDA is not an exact representation of the underlying physics generating the observations.

My approach for solving the inverse problem based on noisy data is to model the measurement error as uncertainties in the Neumann measurements $\{g_i\}_{i=1}^N$, using normally distributed random deviates. This procedure is designed to mimic measurement errors that occur in actual geophysical experiments. The Gaussian errors I add to the Neumann data propagate into the polynomial coefficients, resulting in a set of "noisy" polynomial equations. The noisy polynomial equations may have no real, positive solutions. For instance, the solutions listed in Table II(a) do not satisfy the noisy polynomial equations. To make progress, I propose to add Neumann data at additional frequencies.

If a system of polynomial equations with a real solution has its coefficients perturbed, then the real solution may be perturbed into a complex solution (i.e., with non-zero imaginary parts). If the magnitude of the perturbation is reduced towards zero, the imaginary parts will converge to zero. The relative rates of convergence will be (grossly) represented by the condition number of the real solution, if it is not singular. One might in practice consider complex solutions with "small" imaginary parts as possibly being close to real solutions to the actual, noiseless system. However, in general, the behavior of solutions to polynomial systems whose coefficients are perturbed by noise is not well understood. The treatment here should be regarded as fairly preliminary.

The procedure for generating a synthetic noisy data set is as follows. Perfect Dirichlet and Neumann data are prescribed at N' distinct frequencies, where $N' > N$. The Neumann data are perturbed by adding to each of them a normally distributed random deviate with zero mean and a standard deviation equal to some fixed percentage of the value of the measurement. The noisy Neumann data are then used to compute the coefficients of the governing polynomial equations.

For the case $N = 2$ and $N' = 10$, I used HOMPACK to solve a total of $\binom{10}{2} = 45$ noisy polynomial systems, one for each pair of frequencies. The distributions of model parameters from the solutions (if any exist) for which *all* the model parameters in σ are real and positive are shown in Figures 3(a)–(c), for noise levels 0–5%. The frequencies were chosen logarithmically spaced in the range 0.001–1.0. The distributions of the model parameters from the physical (real, positive) solutions cluster about the true solution, indicating that this method of treating noisy data is quite effective. Another approach that might produce a good estimate of the true model parameter values is to solve polynomial systems that are obtained by averaging the noisy coefficients in some way.

9. MODIFIED HELMHOLTZ TEST RESULTS

Consider the same discrete inverse problem as above, but based now on the modified Helmholtz equation; i.e., the parameter q in Eq. (1) is pure imaginary. Solutions $u(x, z)$ of the modified Helmholtz equation are complex-valued functions, unlike solutions to the Helmholtz equation which are real functions. Therefore, the Dirichlet and Neumann data for the modified Helmholtz inverse problem are complex numbers. Suppose that data like those listed in Table I(b) are given or measured at the boundary of a mesh with $N = 2$. These data generate the following system of four bilinear equations in the unknown σ-variables

$$\begin{aligned} A_R\sigma_1 + B_R\sigma_3 + C_R\sigma_4 + D_R\sigma_1\sigma_3 + E_R = 0 \\ R_R\sigma_2 + G_R\sigma_3 + H_R\sigma_4 + I_R\sigma_2\sigma_4 + J_R = 0 \\ A_I\sigma_1 + B_I\sigma_3 + C_I\sigma_4 + D_I\sigma_1\sigma_3 + E_I = 0 \\ F_I\sigma_2 + G_I\sigma_3 + H\sigma_4 + I_I\sigma_2\sigma_4 + J_I = 0. \end{aligned} \tag{20}$$

The coefficients in the above equations are just the real and imaginary parts of the coefficients A, ..., J given in Section 5, for example, $A = A_R + iA_I$, and so forth. In Section 5, the coefficients A, ..., J were real.

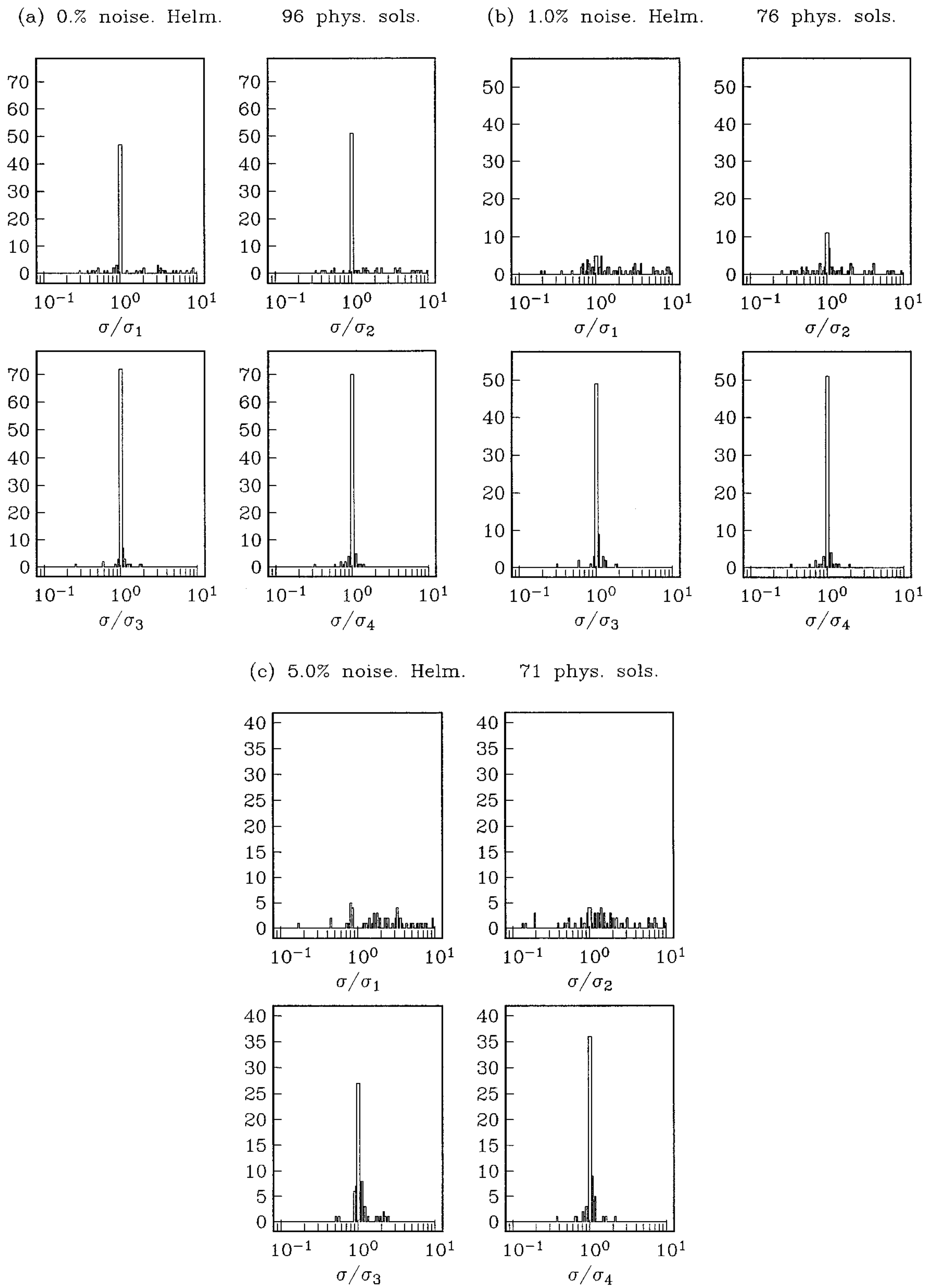

FIG. 3. The distribution of model parameters σ_1, ..., σ_4 from physical (real, positive) solutions to the noisy polynomial systems corresponding to the Helmholtz inverse problem, for different levels of measurement error in the Neumann data: (a) 0%; (b) 1%; (c) 5%. A peak in a histogram at the abscissa value of unity ($\sigma/\sigma_i = 1$) indicates that the model parameter clusters about its true value. For this run, reliable estimates of $\sigma_3 = 6.1$ and $\sigma_4 = 3.7$ therefore can be made in the presence of 5% noise in the Neumann measurements. Model parameter $\sigma_1 = 0.2$ cannot be extracted even in the presence of 1% noise. Model parameter $\sigma_2 = 0.3$ can be estimated reliably from data containing 1%, but not as much as 5%, noise. In Figs. 3 and 4, the vertical axis represents the number of physical solutions in a horizontal bin.

Note that the system (20), four equations in as many unknowns, were generated by boundary data at a single frequency. This stands in contrast to the Helmholtz equation which required data at a pair of frequencies to generate a well-determined system of bilinear equations. A single frequency suffices in this case because each modified Helmholtz boundary datum, see Table I(b), consists of a pair of independent components, the real and imaginary parts.

The data listed in Table I(b) generate the polynomial system

$$\begin{aligned}
&-14537 \cdot 5\sigma_1 - 700 \cdot \sigma_3 - 5 \cdot \sigma_4 \\
&\qquad + 2375 \cdot \sigma_1\sigma_3 + 4298 \cdot 5 = 0 \\
&-3600 \cdot \sigma_2 + 175 \cdot \sigma_3 + 20 \cdot \sigma_4 \\
&\qquad + 1000 \cdot \sigma_2\sigma_4 - 1171 \cdot 5 = 0 \qquad (21) \\
&-26612 \cdot 5\sigma_1 - +380 \cdot \sigma_3 - 40 \cdot \sigma_4 \\
&\qquad + 4375 \cdot \sigma_1\sigma_3 - 2185 \cdot = 0 \\
&487 \cdot 5\sigma_2 - 95 \cdot \sigma_3 + 160 \cdot \sigma_4 \\
&\qquad - 125 \cdot \sigma_2\sigma_4 - 20 \cdot = 0,
\end{aligned}$$

which I solved using HOMPACK. The full solution list is given in Table II(b). There are four real positive solutions (including the "truth") and 12 complex solutions at infinity. There are no other solutions to this inverse problem.

Next, I generated noisy modified Helmholtz data at $N' = 45$ different logarithmically spaced frequencies ranging from 0.001–1.0. Measurement noise was modeled as uncertainties in the real and imaginary components of the Neumann data using independent random deviates drawn from a normal distribution, as before. The deviates were scaled so that the real and imaginary parts contained, on average, the same relative error. Since boundary data at only a single frequency is required to generate a well-determined system of equations, there are again 45 noisy polynomial systems to solve; this time, however, there is only one system per frequency. After applying HOMPACK again, the resulting distributions of the model parameters for the physical solutions are presented, for noise levels 0%, 1%, and 5%, in Figs. 4(a)–(c). The total number of physical solutions is smaller for this problem than for the Helmholtz inverse problem. Also, for a given noise level, more Helmholtz than modified Helmholtz model parameters can be extracted reliably from the distributions. Based on this limited amount of statistical evidence, the inversion of the modified Helmholtz data appears to be more susceptible to noise than the corresponding Helmholtz inversion.

10. DISCUSSION

The method described in this paper can be used to provide complete solutions to discrete inverse problems based on the Helmholtz and modified Helmholtz equations. Many practical geophysical inverse problems can be formulated in this way. The method decides whether solutions to the inverse problem exist and, if so, generates the full solution list in a direct, i.e., non-iterative, manner. Globally optimal models are identified by inspecting the full solution list; e.g., the smoothest model can be determined readily. Noisy data are treated by adding additional Neumann data and adopting a statistical approach to the model parameter identification. Although the homotopy method can be used in principle for arbitrarily large scale inverse problems, the size of the solution space increases dramatically with the number of model parameters so that only small scale problems are at present tractable.

Penalty terms such as smoothness, or positivity, are sometimes imposed as side conditions in an inverse problem in order to stablize the solution. The penalty terms often can be represented by polynomial functions of the model parameters. Incorporating polynomial side conditions into the inverse problem leads to a polynomial programming problem [18]. The first-order necessary conditions for optimality in such problems are, once again, a polynomial system of equations [18]. The method described in this paper can therefore be used to solve the polynomially constrained inverse problem.

The discrete TE and TM modes magnetotelluric inverse problems are based on the modified Helmholtz equations and can therefore be reduced to systems of polynomial equations. The method described in this paper might soon enable the solution space of a magnetotelluric data set to be completely mapped. This would represent a significant breakthrough in geophysical inverse theory and would greatly improve the reliability of inferences about the subsurface electrical structure.

I considered test Helmholtz and modified Helmholtz problems on a small mesh with $N = 2$ ($M = 4$) and found in each case that there were 16 homotopy paths to track. For $N = 3$ ($M = 9$) it is easy to verify that there are 19,683 paths. Each path takes about one second to track on a 30 Mflop computer, so that as M increases, the ability to solve the inverse problem in a reasonable length of time decreases dramatically. There are at least two strategies for reducing the CPU requirements. The first is to exploit the recently developed multi-homogeneous polynomial continuation methods of grouping variables [5]. The second is to improve upon the naive homotopy $G_j(\sigma) = b_j\sigma^{d_j} - a_j$. A more efficient homotopy $\mathbf{G}$, chosen after closely examining the structure of the polynomial equations, can greatly reduce the arc length of the solution paths. Adopting the above considerations will make larger problems more tractable in the future. Another major consideration for the future is the coming availability of massively parallel computers. Each path can be tracked by a separate processor, enabling larger problems to be solved.

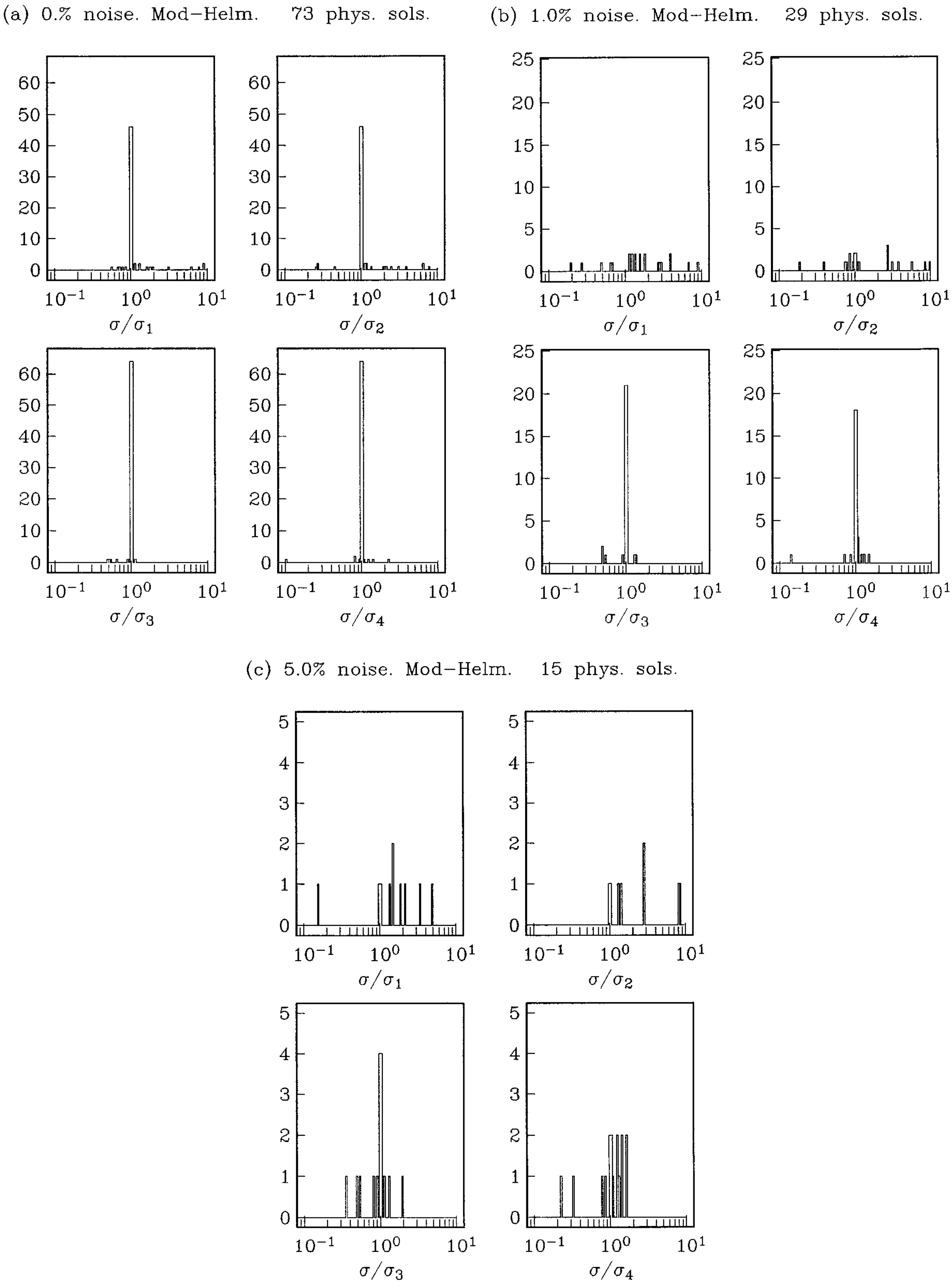

FIG. 4. The distribution of model parameters σ_1, ..., σ_4 from physical (real, positive) solutions to the noisy polynomial systems corresponding to the modified Helmholtz inverse problem, for different levels of measurement error in the Neumann data: (a) 0%; (b) 1%; (c) 5%. Model parameters σ_1 and σ_2 cannot be extracted from data containing 1% or greater noise. Only model parameter σ_3 can be estimated reliably up to the 5% noise level.

ACKNOWLEDGMENTS

The author has greatly benefited from discussions with Alexander Morgan, Adam Schultz, Torquil Smith, Geoff Pritchard, Gary Egbert, and Don Vasco. Financial support was provided by the Isaac Newton Trust.

REFERENCES

1. S. N. Chow, J. Mallet-Paret, and J. A. Yorke, *Math. Comput.* **32,** 887 (1987).
2. M. Kac, *Am. Math. Monthly* (1966).
3. D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1968).
4. A. P. Morgan, *Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems* (Prentice-Hall, Englewood Cliffs, New Jersey, 1987).
5. A. P. Morgan and A. Sommese, *Appl. Math. Comput.* **24,** 101 (1987).
6. A. I. Nachman, *Ann. Math.* **128,** 531 (1988).
7. R. L. Parker, *J. Geophys. Res.* **85,** 4421 (1980).
8. L. R. Price, *IEEE Trans. Biomed. Eng.* **19,** 418 (1979).
9. J. T. Smith and J. R. Booker, *J. Geophys. Res.* **96,** 3905 (1991).
10. J. Sylvester and G. Uhlmann, *Ann. Math.* **125,** 153 (1987).
11. W. M. Telford, L. P. Geldart, and R. E. Sheriff, *Applied Geophysics*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1990).
12. D. W. Vasco, L. R. Johnson, and E. L. Majer, *Geophys. J. Int.* **115,** 711 (1993).
13. K. Vozoff (Ed.), *Magnetotelluric Methods* (Soc. Explor. Geophys., Tulsa, 1985).
14. L. T. Watson, S. C. Billups, and A. P. Morgan, *ACM Trans. Math. Software* **13,** 281 (1987).
15. P. Weidelt, *Z. Geophys.* **38,** 257 (1972).
16. E. Wasserstrom, *SIAM Rev.* **15,** 89 (1973).
17. D. W. Vasco, *Geophys. J. Int.* **119,** 809 (1994).
18. L. T. Watson and A. P. Morgan, *J. Comput. Appl. Math.* **43,** 373 (1992).