

## A MULTI-RESOLUTION, PROBABILISTIC APPROACH TO TWO-DIMENSIONAL INVERSE CONDUCTIVITY PROBLEMS\*

Kenneth C. CHOU and Alan S. WILLSKY

*Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science  
Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

Received 12 April 1988

Revised 6 March 1989

**Abstract.** In this paper a method of estimating the conductivity in a bounded 2-D domain at multiple spatial resolutions given boundary excitations and measurements is presented. The problem is formulated as a maximum-likelihood estimation problem and an algorithm that consists of a sequence of estimates at successively finer scales is presented. By exploiting the structure of the physics, the problem at each scale is divided into two linear subproblems, each of which is solved using parallelizable relaxation schemes. The success of our algorithm on synthetic data is demonstrated and numerical results based on using the algorithm as a tool for exploring estimation performance is presented, as well as results based on using the algorithm to study the well-posedness of the problem, and the effects of fine-scale variations on coarse scale estimates. Examples based on analytical results that further the understanding of these issues are also presented. The results suggest the use of inhomogeneous spatial scales as a possible way of overcoming ill-posedness at points far away from the boundary.

**Zusammenfassung.** In dieser Arbeit stellen wir ein Verfahren zum Schätzen der Leitfähigkeit in einem abgeschlossenen zweidimensionalen Gebiet mit verschiedenen räumlichen Auflösungen bei gegebenen Randerregungen und -messungen vor. Wir formulieren das Problem als Maximum-Likelihood-Schätzproblem und geben einen Algorithmus an, der aus einer Folge von Schätzern auf immer kleiner werdenden Gittern besteht. Unter Ausnutzung der physikalischen Struktur haben wir das Problem auf jedem Gitter in zwei lineare Teilprobleme aufgeteilt, die jeweils mit Hilfe von parallelisierbaren Relaxationsverfahren gelöst werden. Wir wenden unseren Algorithmus erfolgreich auf synthetische Daten an und präsentieren numerische Ergebnisse, die auf dem Einsatz des Algorithmus als Hilfsmittel zur Untersuchung der Güte der Schätzung, der Gutartigkeit des Problems und der Auswirkungen von Variationen auf einem feineren Gitter auf den Schätzer eines gröbereren Gitters beruhen. Wir stellen auch Beispiele vor, die auf analytischen Ergebnissen beruhen und das Verständnis dieser Dinge fördern. Unsere Ergebnisse legen die Verwendung inhomogener räumlicher Gitter als einen möglichen Weg zur Überwindung der schlechten Kondition an Punkten, die weit vom Rand entfernt sind, nahe.

**Résumé.** Dans ce papier on présente une méthode espace-échelle pour l'estimation de la conductivité dans un domaine 2-D borne, dans le cas d'une excitation et de mesures bornées. On formule le problème comme une estimation par maximum de vraisemblance et on présente un algorithme forme d'une suite d'estimateurs aux échelles de plus en plus fines. En exploitant la structure des équations physiques, le problème est divisé à chaque échelle en deux sous-problèmes lineaires dont chacun est résolu en utilisant les méthodes de relaxation qui sont parallélisables. On démontre le succès de notre algorithme sur des données simulées. Cet algorithme est aussi utilisé pour étudier les performances de l'estimation et le caractère bien ou mal posé du problème. Les effets de variations de conductivité à une échelle fine sur les estimateurs obtenus à une échelle moins fine sont regardés. On présente aussi des exemples basés sur les résultats analytiques pour une meilleure compréhension. Nos résultats suggerent l'usage des échelles spatiales non-homogènes comme une possibilité pour résoudre l'estimation en des points éloignés de la frontière bien que ce problème soit mal posé.

**Keywords.** Multi-resolution, maximum likelihood estimation, partial differential equations, Gauss-Seidel relaxation schemes, Cramer-Rao bound.

\* The research described in this paper was supported in part by the National Science Foundation under Grant ECS-8700903, in part by the Army Research Office under Grant DAAL03-86-K-0171, and in part by the Institut de Recherche en Informatique et Systemes Aleatoires, Rennes, France.

## 1. Introduction

Imaging the electrical conductivity in the cross-section of an object by numerical inversion of low frequency, electromagnetic boundary data has applications in various fields of engineering, e.g. exploratory geophysics [4]. The problem involves determining the conductivity in a closed 2-D domain by first applying DC voltage excitations along the boundary and then inverting the resulting DC current measurements normal to the boundary.

Two major difficulties with this 2-D signal processing problem are its well-posedness and the computational complexity of the algorithms that provide solutions. With regard to the former, it is generally not possible to resolve arbitrarily fine spatial fluctuations in an object, especially at points distant from the boundary. Indeed one would expect that as finer resolution inversions are sought, i.e., as the number of degrees of freedom in the model increases, performance would deteriorate. Indeed one danger in seeking an inversion at too fine a resolution is that one may corrupt the estimation of lower resolution features. Also, as the number of degrees of freedom increases, the apparent computational complexity of the required algorithms increases dramatically.

In this paper we present an approach to this problem that was directly motivated by a desire to overcome these difficulties. Specifically, we consider the estimation of the conductivity at a *sequence* of increasingly fine resolutions. One motivation for such a structure is the desire to preserve the quality of lower resolution estimates while still allowing the possibility of attempting higher resolution reconstructions. A second is that by tying this sequence of inversions together, we may be able to achieve substantial computational savings. Specifically, lower resolution reconstructions require far less computation than those at higher resolution. Thus, if we can use coarse-scale inversions to “guide” those at finer scales, we may obtain an algorithmic structure in which most of the work at any particular scale has actually been done at coarser (and computationally simpler)

scales. Note that this philosophy bears some similarity in spirit to so-called multi-grid methods for solving partial differential equations [2, 6], although full multi-grid methods have both coarse-to-fine *and* fine-to-coarse processing.

The framework and method we develop here is based on the maximum likelihood (ML) estimation of a pixelated version of the conductivity profile  $\sigma$ , where a series of pixelation scales is considered. This framework provides us with an extremely useful set of tools to quantify how performance varies with scale. Furthermore, by exploiting the structure of the physical equations we find that not only do we achieve the computational savings hoped for by using a coarse-to-fine estimation structure but we actually obtain much more. Specifically, although the overall inversion problem is highly nonlinear, we develop an iterative relaxation algorithm that alternates between two *linear* inverse problems. Furthermore, both of these problems are highly parallelizable and in fact the parallel pieces of these look the *same* at all scales, i.e., at each scale we have the same, parallel problems to solve. Only the number of these problems increases with resolution.

In the next section we present the physical equations for our problem, a 2-D piecewise constant model for  $\sigma$ , and the resulting ML equation. We also discuss how we exploit the structure of the problem in developing an iterative, highly parallelizable algorithm for computing the ML estimate of  $\sigma$  at any particular scale and then present our overall algorithm for estimating  $\sigma$  at multiple scales. In Section 3 we present both analytical and numerical results on estimation performance as a function of spatial scale. In particular the Cramer–Rao bound provides us with one tool with which to investigate how performance degrades with scale. Also, an important question concerns the effect of fine-scale variations on coarse-scale estimation. We analyze this problem as well and demonstrate that coarse estimates are robust in the presence of such variations. Finally, in Section 4 we present numerical results of applying our algorithm to simulated, noisy data. Our

results show that coarse scale information can be used to improve the computational efficiency of the algorithm at a finer scale. Our results also suggest the use of inhomogeneous spatial scales for  $\sigma$  in order to accommodate for the fact that estimation performance deteriorates at points inward from the boundary.

**2. Problem formulation and multi-resolution estimation**

*2.1. Physical equations*

The problem we consider is the estimation of the conductivity  $\sigma(x, y)$  of an object confined to the unit square given the results of a set of experiments. Each such experiment consists of the application of a known potential on the boundary of the square and the (noise-corrupted) measurements of current around the boundary. If we use the subscript  $i$  to index the number of the experiment, the physics of the problem reduces to the following set of equations [3]

$$\nabla\sigma(x, y)\nabla\phi_i(x, y) = 0, \tag{2.1}$$

for  $0 \leq x \leq 1, 0 \leq y \leq 1,$

where  $\phi_i$  is the potential for the  $i$ th experiment with known, applied boundary conditions

$$B_i(s) \triangleq \phi_i(s), \quad s \in \Gamma, \tag{2.2}$$

where  $\Gamma$  is the boundary of the unit square. If we assume for the moment that current measurements are made continuously around the boundary, our measurements take the form

$$r_i(s) = \sigma(s) \frac{\partial\phi_i(s)}{\partial n} + v_i(s), \tag{2.3}$$

where  $\partial/\partial n$  denotes normal derivative and  $v_i(s)$  denotes the measurement noise which, for simplicity, we model as being independent from experiment to experiment, Gaussian, and white with intensity  $\gamma^{-1}$  for each experiment.

We begin our discretization of the problem by examining the equations when  $\sigma(x, y)$  is piecewise-

constant. Specifically, with an eye toward our eventual multi-scale algorithm, consider the case in which the unit square is divided into  $N \times N$  smaller squares or *pixels* and suppose that  $\sigma$  is constant on each of these. Thus,  $\sigma$  is represented by a finite vector indexed lexicographically as illustrated in Fig. 1 (the same indexing will also be used for other quantities associated with each small

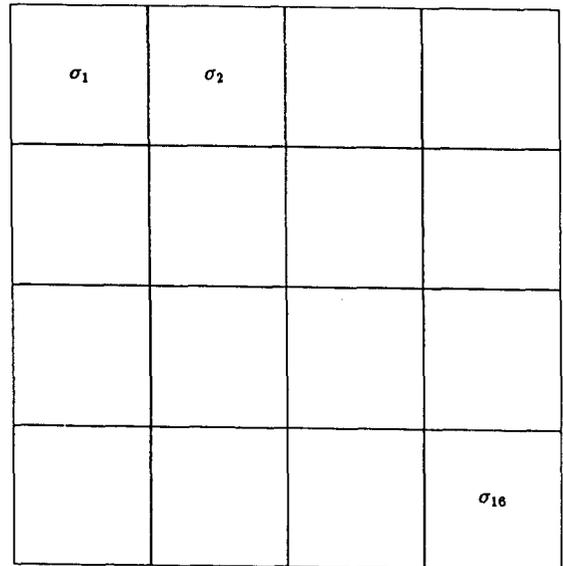


Fig. 1. Example of  $4 \times 4$  pixelation of  $\sigma$ .

square). Note now that *within* each pixel the differential equation (2.1) simplifies to Laplace's equation,

$$\nabla^2\phi = 0, \tag{2.4}$$

which does not depend on  $\sigma$ . The dependence on  $\sigma$  in this case comes from the integral form of Gauss' Law which provides constraints on the boundary conditions for (2.4) along the interior edges of the pixels. Namely, the normal current must be continuous across each interior edge. The normal current along an edge is simply equal to the derivative of  $\phi_i$  in the direction normal to that edge multiplied by the value of  $\sigma$  in the square with which the edge is associated. Note that since  $\sigma$  is discontinuous across an edge,  $\nabla\phi_i$  is also discontinuous across that edge, making what one

calls the normal derivative of  $\phi_i$  along an edge dependent on the side of the edge to which one is referring. Fig. 2 illustrates this for a particular vertical edge. These continuity conditions serve in effect to couple the solutions in the various pixels and to introduce the dependence on  $\sigma$ .

## 2.2. Maximum-likelihood estimation

The ML estimate of the conductivity profile given the measurements (2.3) is obtained by minimizing

$$J = \gamma \sum_{i=1}^M \int_{\Gamma} (r_i(s) - \sigma(s) \frac{\partial \phi_i(s)}{\partial n})^2 ds, \quad (2.5)$$

of auxiliary parameters to be estimated along with  $\sigma$ . Specifically, we consider the problem of optimizing *jointly* with respect to the values of  $\sigma$  in all of the pixels and the values of the potential  $\phi^i$  along each interior edge for each experiment. The criterion to be minimized is the sum of (2.5) and a term penalizing violations in the continuity conditions across internal edges. More precisely, let  $b_{m,j}^i$  denote the potential function on the  $j$ th edge of the  $m$ th square for the  $i$ th experiment. Here  $j = 1, 2, 3, 4$  denote the 4 edges, numbered clockwise starting from the top edge (see Fig. 3). Similarly, let  $z_{m,j}^i$  denote the corresponding normal derivative of the potential. Consider then the minimization over  $\{\sigma_m\}$  and  $\{b_{m,j}^i\}$ ,  $m = 1, \dots, N^2$ ,  $j = 1, \dots, 4$ ,  $i = 1, \dots, M$  of the following

$$J_1 = J + \lambda \sum_{i=1}^M \left\{ \sum_{t=0}^{N-1} \sum_{m=tN+1}^{N(t+1)-1} \|\sigma_m z_{m,2}^i - \sigma_{m+1} z_{m+1,4}^i\|^2 + \sum_{t=0}^{N-2} \sum_{m=tN+1}^{N(t+1)} \|\sigma_m z_{m,3}^i - \sigma_{m+N} z_{m+N,1}^i\|^2 \right\}. \quad (2.6)$$

(here  $\sigma$  could be either piecewise-constant or continuously varying) where  $M$  is the number of experiments. This formula is deceptively simple. Indeed what makes the optimization of (2.5) non-trivial in this case is the fact that the normal derivative function,  $\partial \phi_i(s)/\partial n$ , is an extremely complicated function of  $\sigma$ . That is, to evaluate (2.5) for one candidate profile  $\sigma(x, y)$  requires the complete forward solution of (2.1) for  $\sigma(x, y)$  continuously varying or the set of coupled solutions to (2.4) in all pixels with the continuity constraint imposed across all interior edges. Evaluating the gradient of (2.5) with respect to  $\sigma$  is obviously of at least equal complexity.

Thus, the direct optimization of (2.5) runs into the problem of computational intractability. To overcome this, we focus on the case of a pixelated version of  $\sigma$  and consider the relaxation of the continuity constraints and the introduction of a set

Here the second term is the penalty term on discrepancies in currents across interior edges and  $\lambda$  is a parameter controlling the weight placed on this penalty term. As  $\lambda \rightarrow \infty$ , the solution to (2.6) approaches our original ML solution.

Consider now the structure of (2.6). Note first that, thanks to the linearity of Laplace's equation, each  $z_{m,j}^i$  is a linear function of the corresponding set of potentials  $b_{m,1}^i, b_{m,2}^i, b_{m,3}^i, b_{m,4}^i$ . Also  $\partial \phi_i(s)/\partial n$  in (2.5) corresponds to nothing more than some of the  $z_{m,j}^i$ , i.e., those along the exterior edges, and  $\sigma(s)$  is piecewise-constant with values equal to  $\sigma_m$  for  $m$  corresponding to the pixels that touch  $\Gamma$ . Thus, we see that with all of the  $b_{m,j}^i$  fixed,  $J_1$  is a *quadratic* function of  $\{\sigma_m\}$  while for a fixed set of  $\sigma_m$  values,  $J_1$  is quadratic in  $\{b_{m,j}^i\}$ . This suggests an algorithmic structure with *two* levels of iterations: The outer level consists of the successive minimization of  $J_1$  for an increasing set of

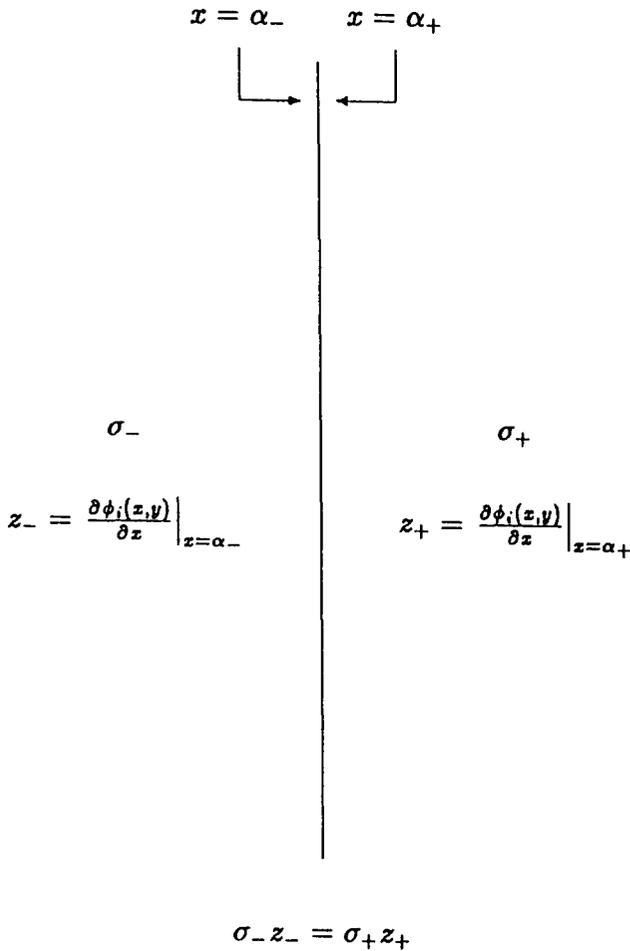


Fig. 2. Continuity of normal current across an edge.

values of  $\lambda$ ; the inner level minimizes  $J_1$  for a given value of  $\lambda$  by alternately solving quadratic minimization problems for the  $\{\sigma_m\}$  and the  $\{b_{m,j}^i\}$  with the other set of variables fixed at the values obtained at the previous step of the iteration.

A deeper examination of (2.6) shows that we can take the algorithmic structure described, one very important step further. Specifically, even though the core of the algorithm described consists of unconstrained quadratic optimization problems, the apparent computational complexity is still daunting given the dimensionality of these problems. This suggests that there may be benefit in using a third level of iteration, e.g., Gauss-Seidel iterations in which we optimize a single  $\sigma_m$  or  $b_{m,j}^i$

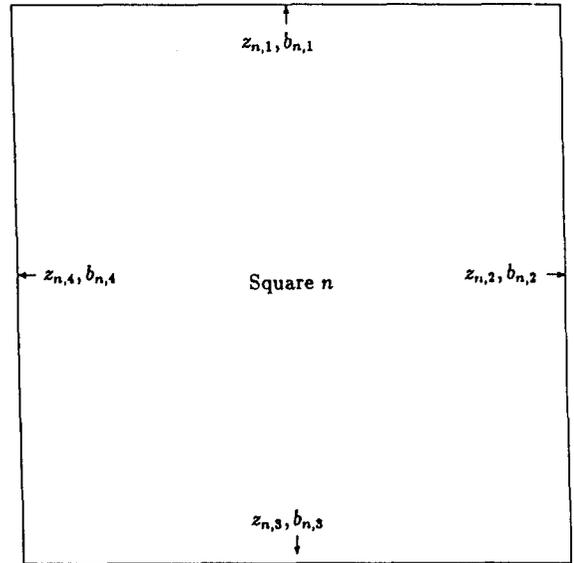


Fig. 3. Normal derivative potential functions defined on edges of a square (for simplicity superscript experiment index  $i$  is not indicated).

keeping all other variables fixed. Not only is this feasible but the structure of (2.6) points to some appealing symmetry that can be exploited. Specifically, from the viewpoint of any particular pixel, we wish to estimate the constant conductivity in the pixel and the potentials on its edges for all experiments by minimizing quadratic terms measuring the difference between the currents at the boundaries (which are functions of the  $\sigma$ 's and  $b$ 's of the pixel) and the currents predicted at these boundaries by neighboring pixels. That is, in a Gauss-Seidel iteration with all variables fixed other than those associated with one pixel, we have an extremely small optimization problem to choose  $\sigma$  and the  $b$ 's for the pixel to minimize the squared error between the resulting currents and the "pseudo-measurements" of current from neighboring pixels. Furthermore, not only is this a small problem, but it is essentially identical in structure for all squares and in fact this decomposition suggests also a highly parallel, distributed algorithm in which processors associated with each pixel interact only through the exchanging of present estimates of the pseudo-measurements needed by

neighboring processors. This suggests a simple architecture for a parallel processing implementation of the algorithm consisting of a network of processors, each of which is connected to each of its four nearest neighbors. Note that the only differences in the problems as seen by each pixel comes from the fact that some pixels are adjacent to the exterior boundary  $\Gamma$ . For such a pixel, the  $b$ 's along edges coincident with  $\Gamma$  are known and the "currents predicted by neighboring pixels" are the actual measurements.

In the next subsection we discuss the details of algorithms having the structure just described. As we will see, there is still substantial flexibility in how one performs the various levels of iteration, and as we show later, certain variations produce considerably superior algorithms.

### 2.3. Algorithm at a single scale

To begin, we assume that we use a regular discretization scheme to solve (2.4) within each pixel.

$$\begin{pmatrix} z_{m,1}^i \\ \vdots \\ z_{m,4}^i \end{pmatrix} = \begin{pmatrix} T_{1,1} & \cdots & T_{1,4} \\ \vdots & & \vdots \\ T_{4,1} & \cdots & T_{4,4} \end{pmatrix} \begin{pmatrix} b_{m,1}^i \\ \vdots \\ b_{m,4}^i \end{pmatrix}, \quad (2.7)$$

where (2.7) represents the approximate solution of (2.4). Here each  $T_{ij}$  is an  $L \times L$  matrix (see Appendix A).

We also assume now that our measurements are also samples along  $\Gamma$ , and we break up this measurement set into subvectors corresponding to the exterior edges of our pixels. Specifically, let  $r_{1,m}^i$  denote the vector of measurements along that portion of the top of the unit square coincident with pixel  $m$  (here  $m = 1, \dots, N$  are the top-most pixels). Similarly,  $r_{2,m}^i$  is a vector corresponding to pixel edges along the right side of the unit square ( $m = tN; t = 1, \dots, N$ ). In this framework the discretized version of (2.6) is

$$J_1 = \sum_{i=1}^M (\gamma \Psi_i + \lambda \Delta_i), \quad (2.8)$$

where

$$\begin{aligned} \Psi_i = & \sum_{m=1}^N \left\| r_{1,m}^i - \sigma_m \sum_{l=1}^4 T_{1,l} b_{m,l}^i \right\|^2 \\ & + \sum_{t=1}^N \left\| r_{2,tN}^i - \sigma_{tN} \sum_{l=1}^4 T_{2,l} b_{tN,l}^i \right\|^2 \\ & + \sum_{m=N(N-1)+1}^{N^2} \left\| r_{3,m}^i - \sigma_m \sum_{l=1}^4 T_{3,l} b_{m,l}^i \right\|^2 \\ & + \sum_{t=1}^N \left\| r_{4,1+(t-1)N}^i - \sigma_{1+(t-1)N} \sum_{l=1}^4 T_{4,l} b_{1+(t-1)N,l}^i \right\|^2 \end{aligned}$$

and

$$\begin{aligned} \Delta_i = & \sum_{t=0}^{N-1} \sum_{m=tN+1}^{N(t+1)-1} \left\| \sigma_m \sum_{l=1}^4 T_{2,l} b_{m,l}^i - \sigma_{m+1} \sum_{l=1}^4 T_{4,l} b_{m+1,l}^i \right\|^2 \\ & + \sum_{t=0}^{N-2} \sum_{m=tN+1}^{N(t+1)} \left\| \sigma_m \sum_{l=1}^4 T_{3,l} b_{m,l}^i - \sigma_{m+N} \sum_{l=1}^4 T_{1,l} b_{m+N,l}^i \right\|^2. \quad (2.9) \end{aligned}$$

That is, we now assume that  $z_{m,j}^i$  and  $b_{m,j}^i$  are  $L$ -dimensional vectors, where  $L$  is the number of sample points per pixel edge. In this case the discretization of (2.4) yields a linear relationship of the following form

Detailed examination of (2.8) and (2.9) shows that it has the structure we have indicated. Each pixel contributes terms to  $J_1$  corresponding to the current along each of its edges. If an edge is exterior, its term appears in  $\Psi_i$ , while it appears in  $\Delta_i$  if it is an interior edge. Note from (2.8) that  $\gamma$  and  $\lambda$

enter in similar ways, representing the relative weights placed on matching the actual external measurements or on matching currents across internal boundaries. If we then consider optimization with respect to any individual  $\sigma_m$  or  $b_{m,j}^i$  with all other variables fixed, we obtain a local and relatively simple problem. For example the optimization for the scalar  $\sigma_n$  depends only on  $\sigma_{n-1}$ ,  $\sigma_{n+1}$ ,  $\sigma_{n-N}$ , and  $\sigma_{n+N}$ . Similar structure is present for the  $b_{m,j}^i$ -optimization, although in this case we must solve an  $M$ -dimensional set of linear equations<sup>1</sup>.

There are a wide variety of ways in which one can imagine iterating among all of these variables. One of the most obvious of these, which we refer to as "Algorithm 1", has exactly the structure described in the previous subsection. That is, we fix the values of all of the  $\sigma$ 's or all of the  $b$ 's and optimize with respect to the other. We then reverse the roles of fixed and free variables and continue in this alternating manner until convergence is achieved. In any one of these steps, the optimization with respect to all the  $\sigma$ 's or all the  $b$ 's is also achieved in an iterative manner, i.e., we sweep through the set of variables optimizing with respect to one variable with all others fixed. This sweep is then iterated until convergence is achieved. The convergence of each of these multi-sweep iterations (for the  $\sigma$ 's and for the  $b$ 's) can be readily established (see [3]).

An alternative algorithm, which we refer to as "Algorithm 2", alternates between single sweeps of  $\sigma_m$ ,  $m = 1, \dots, N^2$  and  $\{b_{m,2}^i, b_{m,3}^i\}$ , i.e., we do not actually take to completion (or even near completion) the minimization with respect to either set of variables with the other fixed. As we will see, this algorithm appears to have superior performance in terms of number of operations required until overall convergence is achieved. Note also that the locality of the individual optimizations would also allow one to develop highly parallel

<sup>1</sup> Note that there are actually far fewer free  $b_{m,j}^i$  than might be apparent at first glance. Specifically, those on exterior edges are given, while on interior edges,  $\phi$  is continuous, i.e.,  $b_{m,2}^i = b_{m+1,4}^i$ ,  $b_{m,3}^i = b_{m+N,1}^i$ .

versions of these algorithms using appropriate coloring schemes.

#### 2.4. Algorithm at multiple scales

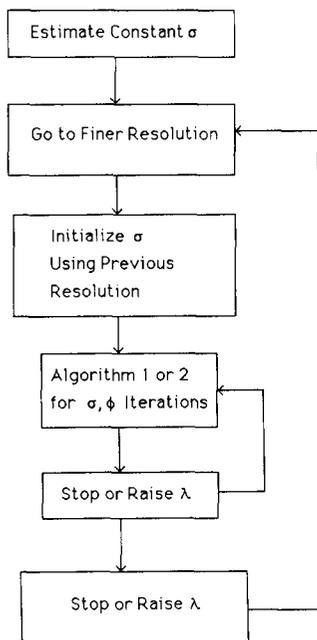
The main idea of our multi-resolution scheme is to compute efficiently the ML estimate of  $\sigma$  at a reasonably fine scale by computing estimates at a sequence of scales starting with a very coarse scale then moving to successively finer scales. The algorithm to accomplish this essentially consists of the following sequence of steps.

- Step 1. Compute  $\hat{\sigma}$  assuming constant  $\sigma$  throughout the unit square.
- Step 2. Subdivide each existing square into four equal squares. Initialize the value of  $\sigma$  in each of these four squares with the value of  $\hat{\sigma}$  computed for the larger square at the previous scale. Given these initial  $\sigma$ , initialize the edge potential vectors by optimizing with respect to  $\phi$ .
- Step 3. Compute  $\hat{\sigma}$ ,  $\hat{\phi}$  using iterative and distributed algorithms described in the previous subsection.
- Step 4. If finer resolution is desired, repeat Steps 2 and 3. Otherwise, stop.

Let us comment on the several aspects of this algorithm. Note first that at the coarsest scale, i.e., where  $\sigma$  is assumed to be constant throughout the unit square, there are no internal potentials to be estimated, i.e., at this scale the estimate of  $\sigma$  is the solution of a single, non-iterative scalar least-squares problem. Note also that the iterations at any subsequent scale are initialized using values of  $\sigma$  at the preceding scale. This offers several potential advantages. First, assuming we have done a good job at one scale, the initial guess for the next scale should be reasonably good, so that the total number of iterations required at the higher resolution, where computations are more complex, should be reduced. Also, the overall function to be optimized at the ultimate resolution will in general have local minima. By "guiding" the solution using estimates at coarser levels, one may be

able to avoid these, assuming of course that they have been avoided at the coarser level. Also, even if finer-scale estimates degrade, this will not disrupt the quality of the lower-resolution reconstructions.

In addition to the recursion in scale and the iterations required for ML estimation at any particular scale, there is also the question of increasing  $\lambda$ . As we discuss later, there appear to be good reasons for keeping  $\lambda$  fixed and not increasing it to arbitrarily large values, i.e., for never actually enforcing the current constraints exactly. However, if increases in  $\lambda$  are desired, one can implement this in two different ways. In particular, as we originally described, one can incorporate a loop in which  $\lambda$  is increased from a small to a large value during the course of iterations at one scale. In this case we increase  $\lambda$  gradually in order to force the constraint of the PDE to be satisfied more and more with each iteration. Alternatively (or in addition), we can also imagine increasing  $\lambda$  with each change in resolution, forcing current constraints to be more closely satisfied as we get to finer resolutions. Fig. 4 is a flow chart illustrating the main structure of the algorithm.



### 3. Analytical estimates of performance

In this section we explore the issue of estimation performance. Our first set of results pertain to a fundamental question arising in our multi-scale formulation. Specifically, we examine the influence of fine-scale variations in  $\sigma$  on the estimate at a coarser scale. By linearizing the PDE, we derive an analytical approximation to the bias and the mean-square error due to fine-scale fluctuations. Numerical evaluation of this shows a significant level of robustness.

Our second set of results uses a similar linearization technique to evaluate the Cramer-Rao bound for estimation at a particular scale. Numerical evaluation again shows a relative insensitivity of performance to actual conductivity values.

#### 3.1. The effect of fine-level fluctuations

For the sake of simplicity we focus here on estimation at the coarsest scale, i.e.,  $\sigma$  is assumed to be constant throughout the unit square when  $\sigma$  actually varies at the next scale. As indicated in the previous section, estimation at the coarsest scale is in fact a linear problem. Specifically, in this case the potential  $\phi_i$  for each set of boundary conditions satisfies Laplace's equation (2.4) throughout the unit square. Thus in this case the normal derivatives for each experiment

$$z_i(s) = \frac{\partial \phi_i(s)}{\partial n}, \quad s \in \Gamma, \tag{3.1}$$

are independent of  $\sigma$  and can be computed a priori. With  $\sigma$  assumed to be constant in (2.3), the ML estimate is seen to be

$$\hat{\sigma} = \gamma P \sum_{i=1}^M \int_{\Gamma} z_i(s) r_i(s) ds, \tag{3.2}$$

where  $P$  is the mean-square estimation error

$$P = E[(\sigma - \hat{\sigma})^2] = \left[ \gamma \sum_{i=1}^M \int_{\Gamma} |z_i(s)|^2 ds \right]^{-1}. \tag{3.3}$$

Suppose now that we use the estimator of (3.2), where the  $z_i(s)$  are computed assuming  $\sigma$  constant, but where in reality  $\sigma$  varies at the next scale, i.e., it takes on different values in each of four smaller squares:

$$\sigma_j = \sigma_0 + \delta\sigma_j, \quad j = 1, \dots, 4, \quad (3.4)$$

where  $\sigma_0$  is the background value that we would like to estimate at the coarsest scale.

In this case the actual normal derivatives of the potential are highly nontrivial and nonlinear functions of the  $\sigma_i$ . However, as outlined in Appendix B, by linearizing the current continuity constraints about the background  $\sigma_0$ , we can compute a first-order approximation to the true normal derivatives as functions of the  $\delta\sigma_j$  and from this approximation to the bias and the mean-square error in estimating  $\sigma_0$  due to the unmodeled fluctuations  $\{\delta\sigma_j\}$ . Specifically, if we condition on  $\sigma_0, \delta\sigma_1, \dots, \delta\sigma_4$  and take the expectation over the noise, we find that an approximation to the bias in our estimate is

$$E[(\sigma_0 - \hat{\sigma}) | \sigma_0, \delta\sigma_1, \dots, \delta\sigma_4] \approx \beta, \quad (3.5)$$

where  $\beta$  is given in (B.12). Also, since the estimator (3.2) is linear, the variance in the estimate is unaffected by the  $\delta\sigma$  fluctuations. Thus, the mean-square error, conditioned on a particular set of  $\delta\sigma$ 's, is approximately given by

$$E[(\sigma_0 - \hat{\sigma})^2 | \sigma_0, \delta\sigma_1, \dots, \delta\sigma_4] \approx \beta^2 + P. \quad (3.6)$$

The quantity  $\beta$  in (3.5) and (3.6) is a quadratic function of the  $\delta\sigma_i$ 's. Because of this simple form it is conceptually straightforward to take an expectation over the  $\delta\sigma_i$ 's as well, now treating them as random variables, to obtain approximations to the unconditional bias and mean-square error. For example, suppose that the  $\delta\sigma_i$ 's are independent, zero-mean and Gaussian with variance  $q$ . Then from (3.5) and (B.12) we obtain

$$E[(\sigma_0 - \hat{\sigma}) | \sigma_0] \approx E[\beta | \sigma_0] = \frac{\alpha_1 q}{\sigma_0}, \quad (3.7)$$

where  $\alpha_1$ , given in (B.16), is a precomputable function of the experimental conditions and is

independent of the background conductivity  $\sigma_0$ . Also, from (3.6), (B.12), and the Gaussian moment factoring formula, we see that

$$E[(\sigma_0 - \hat{\sigma})^2 | \sigma_0] \approx \varepsilon + P, \quad (3.8)$$

where

$$\varepsilon \triangleq E[\beta^2 | \sigma_0] = \alpha_2 q + \alpha_3 q^2. \quad (3.9)$$

The precomputable coefficients  $\alpha_2$  and  $\alpha_3$  are given in (B.17) and (B.18).

Numerical experimentation with this problem indicates a significant level of robustness for the coarse-scale estimator (see [3] for details). In particular, consider a single experiment with  $\gamma = 1$  and boundary conditions

$$\phi(x, y) = \begin{cases} \sin(2\pi x), & y = 1, \\ & 0 \leq x \leq 1, \\ -\sin(2\pi y), & 0 \leq y \leq 1, \\ & x = 1, \\ -\sin(2\pi x), & y = 0, \\ & 0 \leq x \leq 1, \\ -\sin(2\pi y), & 0 \leq y \leq 1, \\ & x = 0. \end{cases} \quad (3.10)$$

In this case we find that  $\alpha_1 \approx -4.57 \times 10^{-12}$  indicating that at least to first order the bias in our estimate is extremely small. Indeed Monte Carlo simulation results in [3] indicate that the true bias is only a few tenths of one percent of  $\sigma_0$  even for rather large fluctuations, e.g., for  $\rho/\sigma_0$  as large as 30% where

$$\rho = \sqrt{q}. \quad (3.11)$$

Since the bias effect of fine-level fluctuations is small, we focus attention on their effect on the mean-square error. Note in particular that the expression in (3.8) decomposes the mean-square error into two terms—one  $\varepsilon$ , due entirely to the fine-scale fluctuations and one  $P$  due solely to the presence of additive measurement noise. One meaningful comparison of the sizes of these two terms can be accomplished by comparing the plots of each versus some appropriate measure of the "size" of the effect causing the error.

Consider first the term due solely to measurement noise. It is convenient to plot the percentage root mean-square (rms) error, i.e.,  $\sqrt{P}/\sigma_0$  versus the signal-to-noise ratio (SNR), or its reciprocal, where

$$\text{SNR} = \frac{\text{signal rms}}{\text{noise strength}}. \tag{3.12}$$

From the measurement model (2.3) we have, for the general case of  $M$  experiments, the definition

$$\text{SNR} = \left( \frac{\sigma_0^2}{4M} \sum_{i=1}^M \int_{\Gamma} |z_i^0(s)|^2 ds \right)^{1/2} \sqrt{\gamma}, \tag{3.13}$$

where we have used the fact that we are assuming  $\sigma$  to be constant; here  $z_i^0$  denotes the boundary normal derivative for the  $i$ th experiment assuming that  $\sigma$  is constant. Then, combining (3.3), (3.12), and (3.13), we see that

$$\frac{\sqrt{P}}{\sigma_0} = \left( \frac{1}{2\sqrt{M}} \right) \text{SNR}^{-1}. \tag{3.14}$$

In the case of the term due to finer-level fluctuations, we can plot  $\sqrt{\epsilon}/\sigma_0$  versus the percentage of fine-level rms fluctuation,  $\rho/\sigma_0$ .

Fig. 5 displays the results of such a comparison. To obtain these results  $P$  was computed using (3.3) but  $\epsilon$  was estimated via Monte Carlo simulation rather than via use of our linearized analysis. This plot allows us to determine the ranges of parameter values over which each of the two terms is dominant. For example at an SNR of 5 the fine-level distortion must be nearly 40% in order to cause an error of size equal to the noise effect. On the other hand for an SNR of 20, the fine-level effect becomes dominant for  $\rho/\sigma_0$  less than 10%. Note also that this plot suggests that in assessing coarse-level estimator performance we can interpret the effect of these finer-level fluctuations as a reduction in the effective SNR.

### 3.2. The Cramer-Rao bound at a single scale

In the case of additive Gaussian noise, as in our problem, the Cramer-Rao lower bound is based

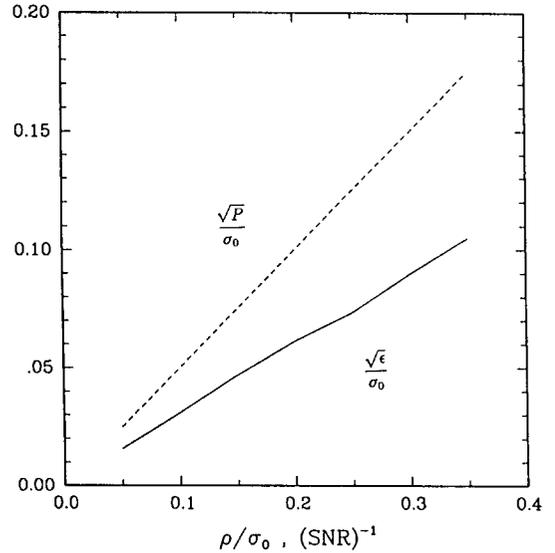


Fig. 5. Plots of  $\sqrt{P}/\sigma_0$  versus  $(\text{SNR})^{-1}$  and of  $\sqrt{\epsilon}/\sigma_0$  versus  $\rho/\sigma_0$ , the fine-level distortion.

on a linearized analysis of estimator performance. It is possible in our case to compute this bound by solving a linearized version of the original PDE similar to that arising in the analysis of the preceding section. Specifically, consider the analysis of the performance in estimating  $\sigma$  when it is assumed to take on different values in each of four sub-squares. Let  $\sigma$  denote the vector of the four values of  $\sigma$ . Then our measurements (2.3) are of the form

$$r_i(s) = h_i(\sigma, s) + v_i(s), \quad i = 1, \dots, M, \tag{3.15}$$

where the measurement function  $h_i$  is defined implicitly through (2.3) and the solution of Laplace's equation in each square with the required current continuity constraints between squares. The Cramer-Rao bound at any hypothesized value  $\sigma^0$  in this case is given by

$$E[(\sigma^0 - \hat{\sigma})(\sigma^0 - \hat{\sigma})'] \geq J^{-1}, \tag{3.16}$$

where  $\hat{\sigma}$  is any unbiased estimate and ' denotes the transpose, and

$$J = \gamma \sum_{i=1}^M \int_{\Gamma} [\nabla h_i(\sigma^0, s)][\nabla h_i(\sigma^0, s)]' ds, \tag{3.17}$$

where

$$\nabla h_i(\boldsymbol{\sigma}^0, s) = \begin{pmatrix} \partial \nabla h_i(\boldsymbol{\sigma}^0, s) / \partial \sigma_1 \\ \vdots \\ \partial \nabla h_i(\boldsymbol{\sigma}^0, s) / \partial \sigma_4 \end{pmatrix}. \quad (3.18)$$

Appendix C outlines how the computation of this gradient can be accomplished by solving PDEs for the individual sensitivities in (3.18).

In [3] the Cramer-Rao bound is displayed for a number of different examples and a number of conclusions can be drawn from these numerical results. The first is that, based on a variety of tests, the bound appears to be relatively insensitive to the particular choice of the nominal conductivity vector  $\boldsymbol{\sigma}^0$ . A second is that for experiment sets that are symmetric in that they probe all parts of the

unit square equally, the errors in estimating the four values of  $\sigma$  are essentially uncorrelated. For example, the bound for the case of a constant nominal,  $\sigma_1^0 = \sigma_2^0 = \sigma_3^0 = \sigma_4^0 = 1$ , and SNR = 1, is given by equation (3.19) in Table 1, while the corresponding matrix of correlation coefficients is given by (3.20).

As we would expect the correlation between the errors in squares one and four is less than that of squares one and two due to the proximities of the squares; however, all of these correlations are small. Similarly, consider the case of 16 independent experiments each consisting of a single impulse in the boundary potential with the location of the impulse being one of 16 points symmetrically distributed about the boundary. In this case, with

Table 1

Cramer-Rao bounds with their corresponding matrix of correlation coefficients

$$J^{-1} = \begin{pmatrix} 3.213024 \times 10^{-2} & -4.489673 \times 10^{-4} & -4.489673 \times 10^{-4} & 1.769225 \times 10^{-5} \\ -4.489673 \times 10^{-4} & 3.213024 \times 10^{-2} & 1.769223 \times 10^{-5} & -4.489673 \times 10^{-4} \\ -4.489673 \times 10^{-4} & 1.769223 \times 10^{-5} & 3.213024 \times 10^{-2} & -4.489673 \times 10^{-4} \\ 1.769225 \times 10^{-5} & -4.489673 \times 10^{-4} & -4.489673 \times 10^{-4} & 3.213024 \times 10^{-2} \end{pmatrix} \quad (3.19)$$

$$\begin{pmatrix} 1 & -1.397336 \times 10^{-2} & -1.397336 \times 10^{-2} & 5.50641 \times 10^{-4} \\ -1.397335 \times 10^{-2} & 1 & 5.50641 \times 10^{-4} & -1.397336 \times 10^{-2} \\ -1.397336 \times 10^{-2} & 5.50641 \times 10^{-4} & 1 & -1.397336 \times 10^{-2} \\ 5.506416 \times 10^{-4} & -1.397336 \times 10^{-2} & -1.397336 \times 10^{-2} & 1 \end{pmatrix} \quad (3.20)$$

$$J^{-1} = \begin{pmatrix} 2.023046 \times 10^{-3} & -3.557312 \times 10^{-5} & -3.557306 \times 10^{-5} & 1.224941 \times 10^{-6} \\ -3.557312 \times 10^{-5} & 2.023046 \times 10^{-3} & 1.22494 \times 10^{-6} & -3.557306 \times 10^{-5} \\ -3.557306 \times 10^{-5} & 1.22494 \times 10^{-6} & 2.023046 \times 10^{-3} & -3.557312 \times 10^{-5} \\ 1.224941 \times 10^{-6} & -3.557306 \times 10^{-5} & -3.557312 \times 10^{-5} & 2.023046 \times 10^{-3} \end{pmatrix} \quad (3.21)$$

$$\begin{pmatrix} 1 & -1.758394 \times 10^{-2} & -1.758391 \times 10^{-2} & 6.054933 \times 10^{-4} \\ -1.758394 \times 10^{-2} & 1 & 6.054928 \times 10^{-4} & -1.758391 \times 10^{-2} \\ -1.758391 \times 10^{-2} & 6.054928 \times 10^{-4} & 1 & -1.758394 \times 10^{-2} \\ 6.054933 \times 10^{-4} & -1.758391 \times 10^{-2} & -1.758394 \times 10^{-2} & 1 \end{pmatrix} \quad (3.22)$$

$$J^{-1} = \begin{pmatrix} 7.813642 \times 10^{-3} & 0.15047 & 0.15047 & 0.26502 \\ 0.15047 & 38631.4 & 265.971 & 12691.5 \\ 0.15047 & 265.971 & 38631.4 & 12691.5 \\ 0.26502 & 12691.5 & 12691.5 & 2.897401 \times 10^{+5} \end{pmatrix} \quad (3.23)$$

$$\begin{pmatrix} 1 & 8.660616 \times 10^{-3} & 8.660616 \times 10^{-3} & 5.569841 \times 10^{-3} \\ 8.660616 \times 10^{-3} & 1 & 6.884837 \times 10^{-3} & 0.11996 \\ 8.660616 \times 10^{-3} & 6.884837 \times 10^{-3} & 1 & 0.11996 \\ 5.569841 \times 10^{-3} & 0.11996 & 0.11996 & 1 \end{pmatrix} \quad (3.24)$$

same constant nominal value of 1 and  $\text{SNR} = 1$ , we find the bound may be represented by (3.21) with a corresponding matrix of correlation coefficients given by (3.22).

Comparing (3.19) and (3.20) with (3.21) and (3.22) we see that the variances in the 16 experiment case are roughly 1/16 those of the 1 experiment example, which is what one would expect from simple linear analysis and our definition of SNR (which is actually an average SNR per experiment). Also, the correlations in the two examples are quite similar and small.

Another conclusion one can draw from our analysis is illustrated in the example in which we have only a single potential impulse in the upper-left corner. The corresponding bound for the same  $\sigma^0$  and SNR of 1 is given by (3.23) and has a corresponding matrix of correlation coefficients given by (3.24).

Notice from Table 1,  $J_{11}^{-1} \leq J_{22}^{-1} \leq J_{44}^{-1}$ , where  $J_{ij}^{-1}$  is the  $ij$ th entry of the matrix (3.23). This indicates that the performance degrades in rather dramatic fashion away from the location in which experimental energy is concentrated. Note also, the larger correlation between errors in squares away from the one in which the impulse is located. This suggests some difficulty in separately estimating these conductivities. We will see a much more dramatic example of this in the next section.

#### 4. Numerical results of algorithm performance

In this section we exercise the algorithms described in Section 2 in order to illustrate several important points. In these examples we use synthetic data generated using a 16-square parametrization of  $\sigma$  and a set of 16 experiments consisting of individual impulse excitations at 16 locations distributed uniformly around the boundary (see [3] for discussion of the generation of this data). Note that at this scale there are four interior squares which are not in contact with the outer boundary on which measurements are taken.

Finally, recall that in describing our algorithms there was the issue of the choice or method of

adjustment for the penalty-term weighting  $\lambda$ . Aspects of this problem are addressed in Section 4.3. In the first two subsections, however, we use noise-free data in order to investigate two other issues, and in these cases we fix  $\lambda$  to be equal to  $\gamma$ .

##### 4.1. Comparison of Algorithms 1 and 2

The core portion of the iterative solution to the estimation problem at any particular scale consists of two types of minimization steps—minimizing with respect to  $\sigma$  given fixed values of internal edge potentials and minimizing with respect to these edge potentials with the  $\sigma$  values held fixed. While each of these is an unconstrained quadratic minimization problem which in principle could be solved without iteration, the algorithms described in Section 2 use Gauss–Seidel sweeps through the various components to overcome dimensionality problems. Algorithm 1, as described in Section 2, carries each such Gauss–Seidel iteration to completion before switching to the optimization with respect to the other variable set. Algorithm 2, in contrast, makes a single Gauss–Seidel sweep through each set of variables and then switches to the other. A fundamental question is: Which is better? i.e., does it pay to carry each  $\sigma$ - or  $\phi$ -iteration to completion or only to perform each partially?

Our method for answering this question involved the comparison of convergence rates using a noiseless data set run at the finest 16-square scale. The specific set of true conductivity values used is given in Table 2. In order to initialize the iterations at

Table 2

True conductivity values ( $\sigma$ ) used to compare Algorithms 1 and 2 (16-square scale)

|         |         |         |         |
|---------|---------|---------|---------|
| 56.7142 | 92.8996 | 102.77  | 100     |
| 117.513 | 68.5311 | 110.48  | 133     |
| 140.345 | 64.9578 | 122.151 | 86.9013 |
| 134.194 | 106.852 | 100     | 135.18  |

this scale for both Algorithms 1 and 2 we use as the starting point for  $\sigma$  the average of the true  $\sigma$  at a four-square scale, e.g., the initial value of  $\sigma$  in each of the four upper-left-hand squares was taken to be the average of the corresponding four values in Table 2. Note that this represents the ideal initialization from the preceding scale. For each algorithm we iterate until the average percentage error of the inner squares reaches a certain level<sup>2</sup>, and then compare the number of iterations required for each algorithm to meet this percentage error criterion. We do this for several levels of the percentage error at which iterations are terminated.

For the case of Algorithm 2 a single iteration is well defined, i.e., an iteration consists of one sweep with respect to  $\sigma$  and one sweep with respect to  $\phi$ . For the case of Algorithm 1 a single iteration consists of two sub-iterations, one for  $\sigma$  and one for  $\phi$ . The number of iterations involved in each of these sub-iterations is the number required for convergence in each case, and the condition for such convergence must be specified. With respect to  $\sigma$  we take adequate convergence to be the point in the iteration at which the percentage change of  $\sigma$  in the inner squares falls below a certain threshold. With respect to  $\phi$  we take adequate convergence to be the point in the iteration at which the percentage change in  $\phi$  along the edges of the inner squares falls below a certain threshold. In our examples we take the value of the threshold to be 0.0001 for both cases.

Fig. 6 is a log-log plot of the total number of  $\sigma$  iterations required by each algorithm in order to meet the following 4 percentage error criteria: 0.5, 0.05, 0.005 and 0.0005.<sup>3</sup> Fig. 7 is a log-log plot of the total number of  $\phi$  iterations required by each for the same set of percentage error criteria.

These plots indicate that for percentage errors down to approximately 0.001 Algorithm 2 per-

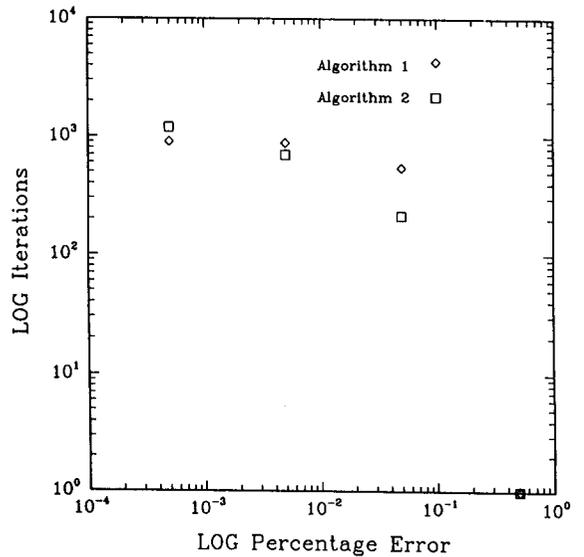


Fig. 6. Log-log plot of total number of  $\sigma$  iterations performed for four different percentage error criteria.

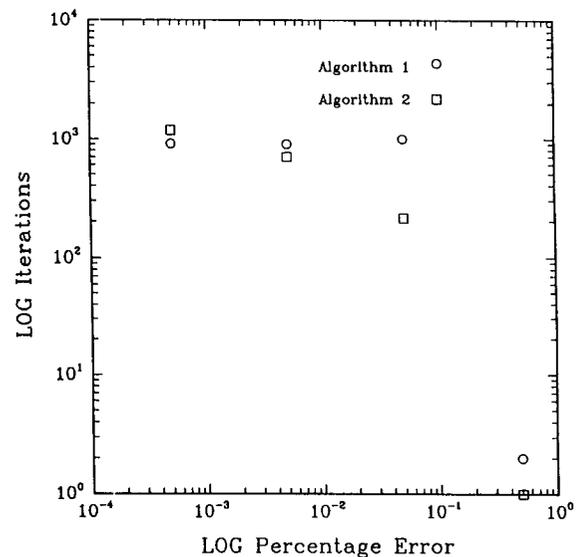


Fig. 7. Log-log plot of total number of  $\phi$  iterations performed for four different percentage error criteria.

forms better than Algorithm 1. For percentage errors lower than this, however, Algorithm 1 performs better than Algorithm 2, but the difference in performance at this level error is small. What this indicates is that if noise levels are sufficient so that such fine a level of accuracy is unachievable, Algorithm 2 is superior. In extremely low noise,

<sup>2</sup> We use the performance of the inner squares as our criterion since the errors there dominate the errors of the outer squares.

<sup>3</sup> These percentage errors are expressed in absolute units rather than in units of percent. We will henceforth adhere to this convention.

however, where extremely accurate optimization is desired, one would do better with a hybrid scheme, i.e., beginning with Algorithm 2-type iterations with a gradual increase in the number of sweeps performed in each individual  $\sigma$ - or  $\phi$ -iteration.

4.2. Initialization using coarse scale information

An important aspect of our algorithm is the use of coarse-scale information to assist finer-level estimation. In this section we investigate the value of this information by examining its influence on the number of iterations required to compute estimates at the 16-square scale. Specifically, we consider the application of Algorithm 2 (with  $\gamma/\lambda$  fixed at 1) to noise-free data generated using the true  $\sigma$  image of Table 3. The different cases looked at were:

- (1) Initialization with information at the preceding scale. In this case the initial values of  $\sigma$  were taken to be constant over  $2 \times 2$  sets of squares of the 16-square image, with values on each  $2 \times 2$  set equal to the average of the true  $\sigma$  values over that region. Specifically, the initial condition corresponding to Table 3 consisted of a value of 10 in the upper-left  $2 \times 2$  set of squares and in the lower-right set and a value of 100 in the two other  $2 \times 2$  sets of squares.
- (2) Initialization using information at the next coarsest scale. In this case the initial value of  $\sigma$  in each of the 16 squares was set to 55, the average of the values in Table 3.
- (3) Random initialization. In the two examples illustrated here in Fig. 8, the random initial  $\sigma$  was

Table 3

True conductivity values ( $\sigma$ ) used to study the influence of coarse scale information on fine scale estimates (16-square scale)

|      |      |     |      |
|------|------|-----|------|
| 10.5 | 12.5 | 105 | 75   |
| 9.5  | 7.5  | 95  | 125  |
| 95   | 80   | 15  | 10.6 |
| 105  | 120  | 5   | 9.4  |

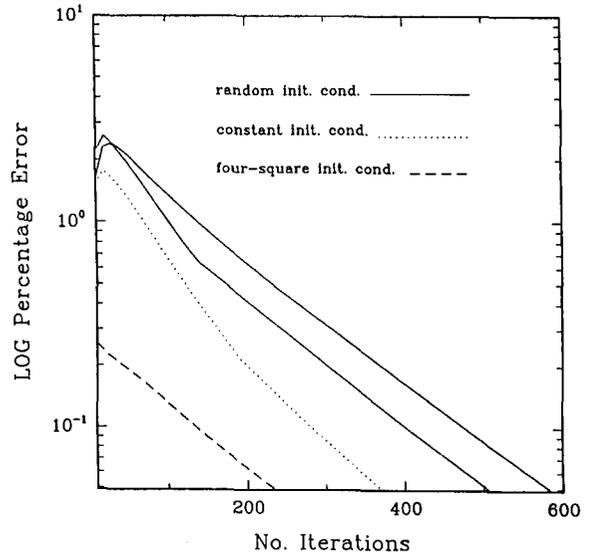


Fig. 8. Semi-log plot of percentage error versus number of iterations for four different initial conditions (the two solid lines represent two different choices of randomly chosen initial conditions).

obtained by choosing the 16 values to be independent and identically distributed with mean 100 and standard deviation of 40.

Fig. 8 represents the results of this comparison. Here we compare the average percentage error in the four inner squares at the end of each iteration. The iterations were terminated once this average fell below 5%. These examples support the claim that the coarse-to-fine algorithmic structure offers potential computational savings in terms of the number of iterations that need to be performed at computationally expensive fine scales.

4.3. Performance of the overall algorithm and the issue of well-posedness

In [3] a variety of tests on the entire hierarchical algorithm<sup>4</sup>, estimating in sequence at constant, 4-square, and 16-square scales, are described. Some of the major observations and conclusions from this study are described here. The specific set of true  $\sigma$ -values used is given in Table 4 where we

<sup>4</sup> In all these tests Algorithm 2 was used throughout.

Table 4

Specific set of true conductivity values used in testing the overall algorithm. (16-square, 4-square and constant value scale)

Constant scale:

|         |
|---------|
| 104.531 |
|---------|

4-square scale:

|         |         |
|---------|---------|
| 83.9145 | 111.563 |
| 111.587 | 111.058 |

16-square scale:

|         |         |         |         |
|---------|---------|---------|---------|
| 56.7142 | 92.8996 | 102.77  | 100     |
| 117.513 | 68.5311 | 110.48  | 133     |
| 140.345 | 64.9578 | 122.151 | 86.9013 |
| 134.194 | 106.852 | 100     | 135.18  |

show the true 16-square set of  $\sigma$ -values and the true 4-square and constant values obtained by averaging the true values at the corresponding scale.

A first set of experiments performed involved the use of noise-free data with  $\gamma/\lambda$  fixed at 1. In this case, one would expect perfect estimation (up to computer accuracy) at the 16-square scale but not at coarser scales, thanks to the analysis in Section 3.1. In this example a fractional error of 0.25 was obtained at the coarsest scale, while errors ranging in magnitude from 0.05 to 0.17 were observed at the 4-square scale, with larger errors occurring in the two left-hand  $2 \times 2$  blocks in which there is greater fluctuation at the finest scale. An additional point that was observed was that convergence at the finest scale required a considerable number of iterations. While values of  $\sigma$  in the outer 12 squares converge quickly, convergence is much slower in the four inner squares, indicative of the conditioning of the inversion problem.

A number of experiments were also run using noise-corrupted measurements. The two major

points at issue were the inclusion of iterations to enforce the current constraints, i.e., an iteration loop in which  $\lambda$  is increased, and the accuracy in estimating  $\sigma$  in the inner squares. The first major conclusion is that increasing  $\lambda$ , i.e., decreasing  $\gamma/\lambda$  from a value of 1, has little effect on estimation performance. For example, at an SNR of 50, an average fractional error of 0.08 is achieved in the four inner squares with  $\gamma/\lambda = 1$ , and the gradual increasing of  $\lambda$  by four orders of magnitude changes this by something less than 0.01. At an SNR of 10, for which we would expect rather poor performance in the inner squares, an average fractional error of 0.8 is achieved in the four inner squares with  $\gamma/\lambda = 1$ , and the gradual increasing of  $\lambda$  by four orders of magnitude changes this by something on the order of 0.1. Thus, a change of four orders of magnitude in  $\lambda$  leads to performance changes in the order of 10%.

A second and not terribly surprising conclusion is that estimation performance in the four inner squares is considerably worse than that in the 12 outer squares that are in direct contact with the measurements. This effect is illustrated in Fig. 9. Here we have plotted three distinct fractional errors as a function of SNR. In particular  $\epsilon_1$  and

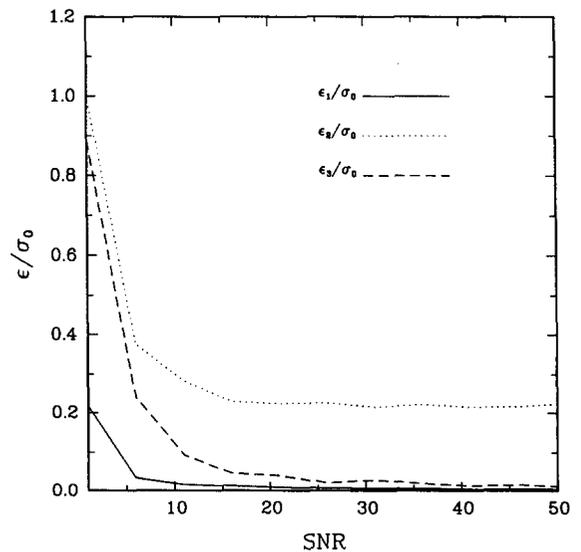


Fig. 9. Plot of the statistics  $\epsilon_1/\sigma_0$ ,  $\epsilon_2/\sigma_0$ , and  $\epsilon_3/\sigma_0$  versus SNR.

$\varepsilon_2$  are, respectively, the average errors in the outer and inner squares, respectively:

$$\varepsilon_1 = \left\{ \frac{1}{12} \sum_o E[(\sigma_o - \hat{\sigma}_o)^2] \right\}^{1/2}, \quad (4.1)$$

$$\varepsilon_2 = \left\{ \frac{1}{4} \sum_i E[(\sigma_i - \hat{\sigma}_i)^2] \right\}^{1/2}, \quad (4.2)$$

where

$$o \in \{1, 2, 3, 4, 5, 8, 9, 12, 13, 14, 15, 16\}, \quad (4.3)$$

and

$$i \in \{6, 7, 10, 11\}. \quad (4.4)$$

The plots of  $\varepsilon_1/\sigma_0$  and  $\varepsilon_2/\sigma_0$  in Fig. 9 indicate the difficulty in resolving the estimate of  $\sigma$  in the inner squares. This suggests that it may be better to estimate these interior values at a coarser scale, since fine-level variations are essentially unobservable from the measurements. To examine this, we have considered taking the average of the estimates in the four inner squares as an estimate of the corresponding average of the true  $\sigma$ -values. The statistic  $\varepsilon_3$  defined by

$$\varepsilon_3 = \left\{ E \left[ \left( \frac{1}{4} \sum_i \sigma_i - \frac{1}{4} \sum_i \hat{\sigma}_i \right)^2 \right] \right\}^{1/2}, \quad (4.5)$$

represents the root-mean-square value of the resulting estimation error. As can be seen in Fig. 9 performance in estimating this coarser-scale quantity is superior to  $\varepsilon_2$  and approaches  $\varepsilon_1$  at high enough SNR!

## 5. Conclusions

In this paper we have presented a way of controlling the large number of degrees of freedom in an inverse conductivity problem by estimating the conductivity at various spatial scales, starting from coarse scales then going to fine. By taking advantage of the structure of the physical equations we have developed an algorithm that, at each scale, consists of a sequence of highly parallelizable relaxation schemes. We have demonstrated the

success of this algorithm on synthetic data as well as investigated various algorithmic and analytic issues that relate to the performance of our method. Two variations of our algorithmic paradigm suggest themselves. The first, suggested by the results in Section 4, is the development of algorithms for inversion in which different resolutions are used in different spatial regions. The second is more speculative and concerns the development of true multi-grid-style algorithms in which there are fine-to-coarse as well as coarse-to-fine iterations. In particular the algorithm as we have presented it involves the solution and inversion of a very accurate approximation of Laplace's equation in each sub-square at each scale. It would seem plausible that one might use a much coarser approximate solution within each square, relying on finer scales to correct coarse-scale approximation errors, as is typically done in multi-grid solutions of forward problems [2, 6].

## Appendix A. A discretized Dirichlet to Neumann map

A critical component of all of our analysis is the use of a mapping from boundary potentials (Dirichlet conditions) on a square to the normal derivatives (Neumann conditions) of the potential on the boundary, when the potential satisfies Laplace's equation within the square. In this appendix we describe a discretized version of this mapping which yields explicitly the block matrix defined in (2.7).

Consider a square with edges of length  $\alpha$  and with boundary potential given by

$$\phi_s = \begin{cases} b_1(x), & 0 \leq x \leq \alpha, \\ & y = \alpha, \\ b_2(x), & 0 \leq y \leq \alpha, \\ & x = \alpha, \\ b_3(y), & 0 \leq x \leq \alpha, \\ & y = 0, \\ b_4(y), & 0 \leq y \leq \alpha, \\ & x = 0. \end{cases} \quad (A.1)$$

The solution to Laplace's equation within the square is then given by [1]

$$\begin{aligned} \phi(x, y) = & \sum_{m=1}^{\infty} \hat{b}_1(m) \sin\left(\frac{m\pi x}{\alpha}\right) \\ & \times \sinh\left(\frac{m\pi y}{\alpha}\right) \\ & + \sum_{m=1}^{\infty} \hat{b}_2(m) \sinh\left(\frac{m\pi x}{\alpha}\right) \\ & \times \sin\left(\frac{m\pi y}{\alpha}\right) \\ & + \sum_{m=1}^{\infty} \hat{b}_3(m) \sin\left(\frac{m\pi x}{\alpha}\right) \\ & \times \sinh\left(\frac{m\pi}{\alpha}(\alpha - y)\right) \\ & + \sum_{m=1}^{\infty} \hat{b}_4(m) \sinh\left(\frac{m\pi}{\alpha}(\alpha - x)\right) \\ & \times \sin\left(\frac{m\pi y}{\alpha}\right), \end{aligned} \tag{A.2}$$

where

$$\begin{aligned} \hat{b}_k(m) = & \frac{2}{\alpha \sinh(m\pi)} \int_0^{\alpha} b_k(s) \\ & \times \sin\left(\frac{m\pi s}{\alpha}\right) ds, \quad k = 1, 2, 3, 4. \end{aligned} \tag{A.3}$$

The corresponding normal derivative functions along each of the four edges are given by

$$\begin{aligned} z_1(x) = & \frac{\partial \phi}{\partial y}(x, \alpha), & z_2(y) = & \frac{\partial \phi}{\partial x}(\alpha, y), \\ z_3(x) = & \frac{\partial \phi}{\partial y}(x, 0), & z_4(y) = & \frac{\partial \phi}{\partial x}(0, y). \end{aligned} \tag{A.4}$$

Note that while  $z_1$  and  $z_2$  are outward pointing normal vectors,  $z_3$  and  $z_4$  point inward. This convention was chosen so that currents across common internal boundaries of adjacent squares would be defined in the *same* direction.

Using (A.2) and (A.4), we can obtain explicit expressions for the normal derivatives. To obtain a discretized version of this relationship between

boundary potentials and normal derivatives, we (a) sample each  $z_i$  expression at  $N$  equally-spaced points along each edge and (b) replace the infinite sums in (A.2) by finite sums and the continuous transforms of (A.3) by  $N$ -point discrete sine transforms of  $N$  equally-spaced samples of each  $b_k$  (see [3] for details). The result is a relationship between the sampled  $b$ 's and  $z$ 's of the form of (2.7), which we can now give explicitly as

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \approx \begin{pmatrix} D & \hat{H}_0 & -S_0 & \hat{H}_{i0} \\ \hat{H}_0 & D & \hat{H}_{i0} & -S_0 \\ S_0 & H_0 & -D & H_{i0} \\ H_0 & S_0 & H_{i0} & -D \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \tag{A.5}$$

where the  $N \times 1$  vectors  $b_k$  and  $z_k$  are the samples of the corresponding boundary potential and boundary normal derivative functions. The  $N \times N$  blocks appearing in (A.5) are defined by the following:

$$S_{i,j} = \sin\left(\frac{ij\pi}{N+1}\right); \tag{A.6}$$

$$D = \mathcal{D}S, \tag{A.7}$$

$$\begin{aligned} \mathcal{D}(i, j) = & \left\{ 2j\pi \cosh(j\pi) \sin\left(\frac{ij\pi}{N+1}\right) \right\} \\ & \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}; \\ S_0 = & \mathcal{S}_0S, \end{aligned} \tag{A.8}$$

$$\begin{aligned} \mathcal{S}_0(i, j) = & \left\{ 2j\pi \sin\left(\frac{ij\pi}{N+1}\right) \right\} \\ & \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}; \\ H_0 = & \mathcal{H}_0S, \end{aligned} \tag{A.9}$$

$$\begin{aligned} \mathcal{H}_0(i, j) = & \left\{ 2j\pi \sinh\left(\frac{ij\pi}{N+1}\right) \right\} \\ & \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}; \\ H_{i0} = & \mathcal{H}_{i0}S, \end{aligned} \tag{A.10}$$

$$\begin{aligned} \mathcal{H}_{i0}(i, j) = & \left\{ 2j\pi \sinh\left(j\pi \left(1 - \frac{i}{N+1}\right)\right) \right\} \\ & \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}; \\ \hat{H}_0 = & \hat{\mathcal{H}}_0S, \end{aligned} \tag{A.11}$$

$$\begin{aligned} \hat{\mathcal{H}}_0(i, j) &= \left\{ 2j\pi \cos(j\pi) \sinh\left(\frac{ij\pi}{N+1}\right) \right\} \\ &\quad \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}; \\ \hat{H}_{i0} &= \hat{\mathcal{H}}_{i0} S, \\ \hat{\mathcal{H}}_{i0}(i, j) &= \left\{ 2j\pi \cos(j\pi) \right. \\ &\quad \times \left. \sinh\left(j\pi\left(1 - \frac{i}{N+1}\right)\right) \right\} \\ &\quad \times \{\alpha(N+1) \sinh(j\pi)\}^{-1}. \end{aligned} \quad (\text{A.12})$$

## Appendix B. Effect of fine level fluctuations on coarse estimates

In this appendix we outline the steps in a linearized analysis to determine the effect of fine-level fluctuations on coarse-level estimates. For simplicity in the discussion we assume that we have a single experiment so that we can omit the subscript  $i$  in our development.

Suppose then that we consider the estimator described in Section 3.1 that assumes that  $\sigma$  is constant throughout the unit square. To emphasize that this is an idealized assumption we rewrite this estimate (see (3.2)) as

$$\hat{\sigma} = \gamma P \int_{\Gamma} z^0(s) r_i(s) ds, \quad (\text{B.1})$$

where the superscript "0" denotes that this is a quantity computed assuming that  $\sigma$  is constant throughout, i.e., in this case  $z^0$  depends only on the applied boundary condition and not on  $\sigma$ . Suppose, however, that  $\sigma$  takes on different values, as given in (3.4), in each of the four subsquares of the unit square. In this case the actual measurements are

$$\begin{aligned} r(s) &= (\sigma_0 + \delta\sigma(s)) \\ &\quad \times (z^0(s) + \delta z(s)) + v(s), \end{aligned} \quad (\text{B.2})$$

where  $\delta\sigma(s)$  takes on one of the four  $\delta\sigma_i$  values corresponding to the sub-square in which the boundary point  $s$  is located. Here  $\delta z(s)$  represents

the perturbation in the normal derivative  $z(s)$  caused by the varying  $\sigma$ .

Substituting (B.2) into (B.1) and using the expression for  $P$  (see (3.3)), we find that

$$\begin{aligned} \sigma_0 - \hat{\sigma} &= -\gamma P \int_{\Gamma} z^0(s) [z^0(s) \delta\sigma(s) \\ &\quad + \sigma_0 \delta z(s) \\ &\quad + \delta\sigma(s) \delta z(s)] ds \\ &\quad - \gamma P \int_{\Gamma} z^0(s) v(s) ds, \end{aligned} \quad (\text{B.3})$$

the second term in (B.3) is the term due to the presence of measurement noise which is the same whether or not there are fine-level fluctuations, while it is the first term on which we wish to concentrate now. In order to do this, we must obtain an expression for  $\delta z(s)$  in terms of the  $\delta\sigma$ 's, and it is here that we perform our linearization.

Recall that in the case in which  $\sigma$  is piecewise-constant, the physics of the problem reduces to Laplace's equation within each sub-square and current continuity at internal boundaries. Note that in the 4-square case we have four separate current constraints for the four internal edges. For example

$$\sigma_1 z_{1,2} = \sigma_2 z_{2,4}. \quad (\text{B.4})$$

However,

$$\sigma_1 = \sigma_0 + \delta\sigma_1, \quad \sigma_2 = \sigma_0 + \delta\sigma_2 \quad (\text{B.5})$$

and

$$z_{1,2} = z_1^0 + \delta z_{1,2}, \quad z_{2,4} = z_1^0 + \delta z_{2,4}. \quad (\text{B.6})$$

Note that we have the *same*  $z_1^0$  in  $z_{1,2}$  and  $z_{2,4}$  since for the nominal condition of  $\sigma$  constant,  $z$  is continuous across this edge. Then, substituting (B.5), (B.6) into (B.4) and linearizing by neglecting the second-order  $(\delta\sigma)(\delta z)$ -terms, we obtain

$$\delta z_{1,2} - \delta z_{2,4} \approx \frac{z_1^0}{\sigma_0^0} (\delta\sigma_2 - \delta\sigma_1). \quad (\text{B.7})$$

Similarly, we find that

$$\delta z_{2,3} - \delta z_{4,1} \approx \frac{z_2^0}{\sigma_0^0} (\delta\sigma_4 - \delta\sigma_2), \quad (\text{B.8})$$

$$\delta z_{3,1} - \delta z_{1,3} \approx \frac{z_3^0}{\sigma_0} (\delta \sigma_1 - \delta \sigma_3), \tag{B.9}$$

$$\delta z_{4,4} - \delta z_{3,2} \approx \frac{z_4^0}{\sigma_0} (\delta \sigma_3 - \delta \sigma_4). \tag{B.10}$$

What we now have is the following *linear* problem to solve:

- Laplace’s equation holds for the perturbation to the potential in each sub-square.
- The perturbation to the potential,  $\delta b_{i,j}$ , on the outer boundary  $\Gamma$  is zero (since the potential on this boundary is the fixed experimental boundary condition).
- We are given the *difference* in the perturbed normal derivative across the internal boundaries (B.7)–(B.10).
- We wish to compute the perturbation  $\delta z(s)$  along the exterior boundary, i.e.,  $s \in \Gamma$ .

It is not difficult to check that this is a well-posed problem, and, thanks to its linearity and the linearity of (B.7)–(B.10) in the  $\delta \sigma_i$ , that we can then express the desired  $\delta z(s)$  in the form

$$\delta z(s) \approx \frac{1}{\sigma_0} \sum_{k=1}^4 f_k(s) \delta \sigma_k, \tag{B.11}$$

where the  $f_k(s)$  are functions of  $z^0$  but not  $\sigma_0$  (see (B.7)–(B.10)). The method used in [3] to evaluate this numerically is as follows. From (2.7) we know that when we use our discretized numerical solution to Laplace’s equation, there is a linear relationship (namely (2.7)) between the discretized versions of the  $\delta \sigma$ ’s and  $\delta b$ ’s on the boundaries of each sub-square. However, many of the  $\delta b$ ’s are zero; indeed only the four  $\delta b$  vectors corresponding to the four internal edges are nonzero. Also, we have four sets of constraints in (B.7)–(B.10). This allows us to solve for the  $\delta b$ ’s in terms of the right-hand sides of (B.7)–(B.10) and then we can use (2.7) to solve for the  $\delta z$ ’s on the outer edges. See [3] for details in which these steps are carried out explicitly using a particular numerical scheme.

Given (B.3) and (B.11) we can now obtain an approximation to the conditional bias in our esti-

mate due to fine-level fluctuations. Specifically,

$$\begin{aligned} E[(\sigma_0 - \hat{\sigma}) | \sigma_0, \delta \sigma_1, \dots, \delta \sigma_4] & \\ & \triangleq \beta \\ & \approx -\gamma P \int_{\Gamma} z^0(s) \\ & \quad \times \left[ z^0(s) \delta \sigma(s) + \sum_{k=1}^4 f_k(s) \delta \sigma_k \right. \\ & \quad \left. + \frac{\delta \sigma(s)}{\sigma_0} \sum_{k=1}^4 f_k(s) \delta \sigma_k \right] ds, \end{aligned} \tag{B.12}$$

Extending this to the case of  $M$  experiments indexed by  $i$  and introducing some simplifying notation we have

$$\begin{aligned} -\frac{\beta}{\gamma P} &= \sum_{i=1}^M \int_{\Gamma} z_i^0(s) \\ & \quad \times \left[ z_i^0(s) \delta \sigma(s) + \sum_{k=1}^4 f_{i,k}(s) \delta \sigma_k \right. \\ & \quad \left. + \frac{\delta \sigma(s)}{\sigma_0} \sum_{k=1}^4 f_{i,k}(s) \delta \sigma_k \right] ds \\ &= \sum_{j=1}^4 \mu_j \delta \sigma_j + \sum_{k=1}^4 \sum_{l=1}^4 \nu_{k,l} \delta \sigma_k \delta \sigma_l, \end{aligned} \tag{B.13}$$

where  $P$  is computed from (3.3) with  $z_i^0$  replacing  $z_i$  and

$$\begin{aligned} \mu_j &= \sum_{i=1}^M \left[ \int_{\Gamma_j} |z_i^0(s)|^2 ds \right. \\ & \quad \left. + \int_{\Gamma} z_i^0(s) f_{i,j}(s) ds \right], \end{aligned} \tag{B.14}$$

$$\nu_{k,l} = \frac{1}{\sigma_0} \sum_{i=1}^M \int_{\Gamma_i} z_i^0(s) f_{i,k}(s) ds, \tag{B.15}$$

where  $\Gamma_j$  is the portion of  $\Gamma$  contained in the  $j$ th square. Note also that the linearization coefficient functions  $f_{i,k}(s)$  are generally different for different experiments as they are a function of the boundary conditions.

Suppose now that the  $\delta \sigma_i$  are independent, zero-mean Gaussian random variables with variance  $q$ .

Then we immediately see that  $E[\beta]$  has the form given in (3.7), with

$$\alpha_1 = -\gamma P \sum_{k=1}^4 \nu_{k,k}. \quad (\text{B.16})$$

Furthermore, using Gaussian moment properties, we find that  $\varepsilon$  in (3.8), (3.9) does have the indicated form, with

$$\alpha_2 = \gamma^2 P^2 \sum_{j=1}^4 \mu_j^2, \quad (\text{B.17})$$

$$\alpha_3 = \gamma^2 P^2 \left\{ \left[ \sum_{k=1}^4 \nu_{k,k} \right]^2 + \sum_{k=1}^4 \sum_{l=1}^4 \nu_{k,l} (\nu_{k,l} + \nu_{l,k}) \right\}. \quad (\text{B.18})$$

### Appendix C. Measurement sensitivity computations

In this appendix we outline the steps involved in calculating the sensitivities required in computing the Cramer-Rao bound described in Section 3.2. These sensitivities are needed for each individual experiment and thus we focus on a single experiment and drop the experiment index  $i$ . From (2.3) and (3.25) we have that

$$\begin{aligned} h(\boldsymbol{\sigma}, s) &= \sigma(s)z(s), \quad s \in \Gamma \\ &= \sigma_j z(s), \quad s \in \Gamma_j, \\ & \quad j = 1, \dots, 4, \end{aligned} \quad (\text{C.1})$$

where the normal derivative  $z(s)$  is an implicit function of  $\boldsymbol{\sigma}$  and  $\Gamma_j$  is the portion of  $\Gamma$  in the  $j$ th of the four sub-squares.

For simplicity we focus on the sensitivity with respect to  $\sigma_1$ , as the others are obtained in an analogous fashion. From (C.1) we then have that

$$\frac{\partial h(\boldsymbol{\sigma}^0, s)}{\partial \sigma_1} = \begin{cases} z^0(s) + \sigma_1 \frac{\partial z^0(s)}{\partial \sigma_1}, & s \in \Gamma_1, \\ \sigma_j \frac{\partial z^0(s)}{\partial \sigma_1}, & s \in \Gamma_j, \\ & j = 2, 3, 4. \end{cases} \quad (\text{C.2})$$

Here  $z^0(s)$  is the normal derivative function when the conductivity profile is given by  $\boldsymbol{\sigma}^0$ , and  $\partial z^0/\partial \sigma_1$  is the sensitivity of this profile with respect to  $\sigma_1$ . The first of these  $z^0(s)$  can be computed directly by solving Laplace's equation within each of the four sub-squares with the given outer boundary potential and with the following continuity constraints on internal boundaries:

$$\begin{aligned} \sigma_1^0 z_{1,2}^0 &= \sigma_2^0 z_{2,4}^0, & \sigma_2^0 z_{2,3}^0 &= \sigma_4^0 z_{4,1}^0, \\ \sigma_3^0 z_{3,1}^0 &= \sigma_1^0 z_{1,3}^0, & \sigma_4^0 z_{4,4}^0 &= \sigma_3^0 z_{3,2}^0. \end{aligned} \quad (\text{C.3})$$

Differentiating with respect to  $\sigma_1^0$  then yields the PDE to be solved in order to compute  $\partial z^0/\partial \sigma_1$ : We solve Laplace's equation in each square with zero outer boundary conditions and the following constraints on the internal boundaries:

$$z_{1,2}^0 + \sigma_1^0 \frac{\partial z_{1,2}^0}{\partial \sigma_1} = \sigma_2^0 \frac{\partial z_{2,4}^0}{\partial \sigma_1}, \quad (\text{C.4})$$

$$\sigma_2^0 \frac{\partial z_{2,3}^0}{\partial \sigma_1} = \sigma_4^0 \frac{\partial z_{4,1}^0}{\partial \sigma_1},$$

$$\sigma_3^0 \frac{\partial z_{3,1}^0}{\partial \sigma_1} = z_{1,3}^0 + \sigma_1^0 \frac{\partial z_{1,3}^0}{\partial \sigma_1},$$

$$\sigma_4^0 \frac{\partial z_{4,4}^0}{\partial \sigma_1} = \sigma_3^0 \frac{\partial z_{3,2}^0}{\partial \sigma_1}.$$

Thus the only driving terms in this PDE are the  $z_{1,2}^0$  and  $z_{1,3}^0$  terms in (C.4), which are obtained from the solution associated with (C.3).

We refer the reader to [3] for a detailed description of explicit computations (again involving the discrete mapping (2.7)) that allow one to determine the desired sensitivities numerically.

### References

- [1] W.F. Ames, *Numerical Methods for Partial Differential Equations*, Academic Press, New York, NY, 1977.
- [2] A. Brandt, "Multi-level adaptive solutions to boundary value problems", *Math. Comp.*, Vol. 13, 1977, pp. 333-390.
- [3] K.C. Chou, "A multi-resolution approach to an inverse conductivity problem", SM Thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, December 1987.

- [4] K. Dines and R. Lytle, "Analysis of electrical conductivity imaging", *Geophysics*, Vol. 46, No. 7, 1981, pp. 1025-1036.
- [5] G. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MS, 1983.
- [6] W. Hackbusch and U. Trottenberg, eds., *Multigrid Methods*, Springer-Verlag, New York, NY, 1982.
- [7] A.K. Jain and E. Angel, "Image restoration modelling and reduction of dimensionality", *IEEE Trans. Comput.*, Vol. C-23, No. 5, May 1974, pp. 470-476.
- [8] D. Luenberg, *Linear and Nonlinear Programming*, Addison Wesley, Reading, MA, 1984.
- [9] D. Paddon and H. Holstein, eds., *Multigrid Methods for Integral and Differential Equations*, Clarendon Press, Oxford, England, 1985.
- [10] G. Strang, *Linear Algebra and its Applications*, Academic Press, New York, NY, 1980.
- [11] A. Tarantola and B. Valette, "Generalized nonlinear inverse problems solved using the least squares criterion", *Rev. Geophys. Space Phys.*, Vol. 20, No. 2, 1982, pp. 219-232.
- [12] A. Tarantola and B. Valette, "Inverse problems—Quest for information", *J. Geophys.*, Vol. 50, 1982, pp. 159-170.
- [13] D. Terzopoulos, "Image analysis using multigrid relaxation methods", *IEEE Trans. Pattern Analy. Mach. Intell.*, Vol. PAMI-8, No. 2, March 1986, pp. 129-139.
- [14] H.L. Van Trees, *Detection, Estimation and Modulation Theory*, Vol. 1, John Wiley, New York, NY, 1968.