Inversion of 3D Electromagnetic Data in frequency and time domain using an inexact all-at-once approach

Eldad Haber*
UBC-Geophysical Inversion Facility
Department of Earth and Ocean Sciences
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada.

Uri Ascher
Department of Computing Sciences
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada.

Douglas W. Oldenburg
UBC-Geophysical Inversion Facility
Department of Earth and Ocean Sciences
University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada.

June 13, 2002

Abstract

We present a general formulation for solving frequency or time domain electromagnetic data using an all-at-once approach. In this

*Currently at EMI Schlumberger, Richmond, California, USA 94804.

methodology the forward modelling equations are incorporated as constraints and thus we need to solve a constrained optimization problem where the parameters are the electromagnetic fields, the conductivity model and a set of Lagrange multipliers. This leads to a much larger problem than the traditional unconstrained formulation where only the conductivities are sought. Nevertheless, experience shows that the constrained problem can be solved faster than the unconstrained one. The primary reasons are that the forward problem does not have to be solved exactly until the very end of the optimization process, and permitting the fields to be away from their constraint values in the initial stages, introduces flexibility so that a stationary point of the objective function is found more quickly. In this paper we outline the all-at-once approach and apply it to EM problems in the frequency and time domain. This is facilitated by a unified representation for forward modelling for these two types of data. The optimization problem is solved by finding a stationary point of the Lagrangian. Numerically, this leads to a nonlinear system that is solved iteratively using a Gauss-Newton strategy. At each iteration a large matrix is inverted and we discuss how this can be accomplished. As a test, we invert frequency domain synthetic data from a grounded electrode system that emulates a field CSAMT survey. For the time domain, we invert borehole data obtained from a current loop on the surface.

1 Introduction

In this paper we develop an inversion methodology for 3D electromagnetic data in both frequency and time domains. This problem is of major interest in geophysics, medical imaging and non-destructive testing; see for example (Smith and Vozoff, 1984; Parker, 1994; Devaney, 1989; Borcea et al., 1996; Vogel, 1999; Haber and Ascher, 2001a; Cheney et al., 1999) and references therein. The forward model consists of Maxwell's equations in which the permeability is constant but electrical conductivity can be highly discontinuous. The parameter regimes considered give rise to highly stiff problems in the time domain, or alternatively low frequencies in the frequency domain. The goal of the inversion is to recover the conductivity given measurements of the electric and/or magnetic fields.

There are many practical challenges to solving the inverse problem. First, a fast, accurate and reliable algorithm for 3D forward modeling is required

in frequency and in time. Second, the sensitivities for such problems are too numerous to be formed or stored in a reasonable amount of time and space. Finally, finding the minimum of the objective function obtained by matching the data and incorporating a priori information on the conductivity field can be difficult due to the nonlinearity and sensitivity of the problem.

We use an inexact, all-at-once methodology (Biros and Ghattas, 2000; Haber and Ascher, 2001b; Ascher and Haber, 2001), solving the forward problem and the inverse problem simultaneously in one iterative process. This approach allows development of highly efficient algorithms. However, since it couples the solution of the forward problem with the solution of the inverse problem, the forward problem cannot be treated as a "black box". Care must be taken to properly fit the formulation and discretization of the forward problem within the inverse methodology.

To be more specific, assume that the forward problem (in continuous space) is written in the form

$$\mathcal{A}(m)u - q = 0 \tag{1.1}$$

where $\mathcal{A}(m)$ is a version of Maxwell's equations (including boundary conditions) either in time or in frequency, $m = \log(\sigma)$ is the log conductivity, u stands for the fields and q represents sources and boundary values. We assume for simplicity of exposition that $\mathcal{A}$ is invertible for any relevant m, i.e., there is a unique solution to the forward problem.

In the inverse problem we measure some function of the fields and want to recover the model m. Let us write the measured data as

$$b^{obs} = \mathcal{Q}u + \epsilon \tag{1.2}$$

where $\mathcal{Q}$ is a measurement operator which projects the fields (or their derivatives or integrals) onto the measurement locations in 3D space and possibly time, and ϵ is the measurement noise.

The data are finite and contaminated with noise. To obtain a unique model which depends stably on the data we use a-priori information and formulate the inverse problem (in continuous space) as a constrained optimization problem of the form

$$\begin{aligned} \min \quad & \frac{1}{2}\|\mathcal{Q}u - b^{obs}\|^2 + \beta\mathcal{R}(m) \\ \text{subject to} \quad & \mathcal{A}(m)u - q = 0. \end{aligned} \tag{1.3}$$

Here, $\beta > 0$ is the regularization parameter, and $\mathcal{R}(\cdot)$ is a regularization operator reflecting our *a-priori* information. Typically, we know that m is a piecewise smooth function over the spatial domain Ω in 3D, so we assume that $\mathcal{R}(\cdot)$ involves some norm of ∇m over Ω, e.g., weighted L_2 or L_1 or a Huber combination (Huber, 1964; Farquharson and Oldenburg, 1998). The data fitting term in (1.3) involves the L_2 norm over Ω.

Next, the problem (1.3) is discretized using some finite volume or finite element method over a finite grid representing the domain in space and time, yielding the finite dimensional optimization problem

$$\begin{aligned} \min \quad & \frac{1}{2}\|Qu - b^{obs}\|^2 + \beta R(m) \\ \text{subject to} \quad & A(m)u - q = 0, \end{aligned} \tag{1.4}$$

where u, m and q are grid functions ordered as vectors and correspond to their continuous counterparts above, and Q, A and $\frac{\partial R}{\partial m}$ are all large, sparse matrices. The matrix A depends on m and is nonsingular.

The common approach to solving this problem (Newman and Alumbaugh, 1995; Newman and Alumbaugh, 1997a; Newman and Alumbaugh, 1997b; Madden and Mackie, 1989; Tikhonov and Arsenin, 1977; Parker, 1994; Vogel, 1999), is to first eliminate the equality constraints, obtaining an unconstrained optimization problem of the form

$$\min \ \frac{1}{2}||QA(m)^{-1}q - b^{obs}||^2 + \beta R(m).$$

This approach is rooted in the vast amount of literature and methods which have been developed for unconstrained optimization and the positive definiteness of corresponding approximations to the Hessian matrix (Nocedal and Wright, 1999; Dennis and Schnabel, 1996; Kelley, 1999). However, each evaluation of the objective function requires a solution of the forward problem, and evaluating the gradients requires the solution of the adjoint problem. Evaluating the sensitivity matrix requires many more solutions of the forward and adjoint problems. The resulting procedure may therefore become very computing intensive.

For this reason a recent work using the unconstrained approach employs a nonlinear conjugate gradient (CG) method (Rodi and Mackie, 1998). The cost of the inversion is then proportional to the number of iterations taken in the nonlinear CG algorithm times twice the cost of solving one forward problem. For large scale problems this can still be prohibitively expensive. Even

if we replace the relatively slow nonlinear CG by a Newton or Gauss-Newton variant coupled with CG for the linearized problem (Nocedal and Wright, 1999; Vogel, 1999) the cost of carrying out each such iteration remains very high (Haber and Ascher, 2001b).

Here, instead, we consider the constrained optimization problem directly. This allows balancing the accuracy of the iterations for solving the forward and the inverse problems. The work is a follow on to that presented in Haber et al.(2000b).

Let us form the Lagrangian

$$\mathcal{L}(u,m,\lambda) \;=\; \frac{1}{2}\|Qu-b^{obs}\|^2+\beta R(m)+\lambda^T(A(m)u-q)$$

where λ is the vector, or grid function of the same form and size as u, of Lagrange multipliers.[1] The first order necessary condition of optimality for (1.4) is that the gradient of $\mathcal{L}$ vanish. This yields the nonlinear system of algebraic equations

$$\mathcal{L}_\lambda \;=\; A(m)u-q=0 \tag{1.5a}$$
$$\mathcal{L}_u \;=\; A(m)^*\lambda+Q^*(Qu-b^{obs})=0 \tag{1.5b}$$
$$\mathcal{L}_m \;=\; \beta\frac{\partial R}{\partial m}+G(m,u)^*\lambda=0 \tag{1.5c}$$

where $()^*$ is the adjoint operator and

$$G(m,u)=\frac{\partial[A(m)u]}{\partial m}.$$

The system (1.5) is solved by a Newton-type method. Within the outer, nonlinear iteration there are then iterative methods applied to solve the linearized problem approximately.

The system (1.5) is clearly a discretization of a system of boundary value partial differential equations (PDEs). The first PDE corresponding to (1.5a) is simply the forward problem, e.g. Maxwell's equations in our present case. The second PDE corresponding to (1.5b) can be viewed as the adjoint problem, i.e., we can view it as an equation for the Lagrange multiplier function

$$\mathcal{A}(m)^*\lambda=-\mathcal{Q}^*(\mathcal{Q}u-b^{obs})$$

[1]We are abusing notation slightly by using the same symbols for u, m and λ in the discrete and the continuous cases. The meaning should be clear from the context.

where the right hand side involves the noise. This view is tightly connected to the adjoint method (Chavent, 1989).

The last equation (1.5c) can be viewed as discretizing a diffusion equation with natural boundary conditions for the model of the form (Weickert, 1998),

$$-\beta\nabla\cdot\ (a\nabla m) = \mathcal{G}(m,u)^*\lambda.$$

Here, a could be a 3×3 matrix which depends on ∇m. However, in this paper we restrict attention to a constant, diagonal weight matrix (Ascher and Haber, 2001)

$$a = diag\{\alpha_1,\alpha_2,\alpha_3\}.$$

If we use the forward model to eliminate u and then the adjoint equation to eliminate λ then the above diffusion equation is expressed solely in terms of the model m. This again expresses the "eliminate first", or unconstrained approach. However, viewing (1.5) as discretizing the entire system of PDEs it is important to note that for realistically small values of the regularization parameter β the PDEs are tightly coupled (Haber and Ascher, 2001b; Ascher and Haber, 2001). This leads to complications in the solution of the problem and hints that the elimination route may not be the natural one.

We emphasize that the formulation and discretization of the forward and the inverse problems should generate "good", consistent forward adjoint and model equations in (1.5); otherwise, we may expect to have difficulties (Haber and Ascher, 2001b).

In the present article we apply the methodology described above (which we developed earlier mainly in a simpler context corresponding to DC-resistivity problems(Haber and Ascher, 2001b)) to Maxwell's equations, both in time and in frequency domains. In Section 2 we discuss the formulation and the discretization of the forward modeling in space and time. In Section 3 we reformulate the discrete inverse problem as a constrained optimization problem. In Section 4 we discuss the solution of the systems which evolve from our formulation. In Section 5 we discuss the optimization procedure. In all of these we concentrate on what is new or different from our previous expositions, and only briefly describe that which is similar. In Section 6 we give some geophysical examples.

2 The Forward Problem

In this section we present our forward problem, Maxwell's equations, and discuss solution procedures suitable for the parameter regimes of interest. Most, but not all of the present development follows our previous work (Haber et al., 2000a; Haber and Ascher, 2001a). Section 2.1, in particular, is completely new.

The time-dependent Maxwell equations can be written as

$$\begin{aligned} \nabla \times \mathbf{E} + \mu \mathbf{H}_t &= 0, && \text{(2.6a)} \\ \nabla \times \mathbf{H} - \sigma \mathbf{E} - \epsilon \mathbf{E}_t &= \mathbf{s}_r(t), && \text{(2.6b)} \end{aligned}$$

over a domain $\Omega \times [0, t_f]$, where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields, μ is the permeability, σ is the conductivity, ϵ is the permittivity and s_r is a source. The equations are given with some initial and boundary conditions which we discuss next. In the frequency domain, reusing the same symbols for $\mathbf{E}$, $\mathbf{H}$ and s_r (the context should make this unambiguous), the same equations in the spatial domain Ω read

$$\begin{aligned} \nabla \times \mathbf{E} - \imath\omega\mu \mathbf{H} &= 0 && \text{(2.7a)} \\ \nabla \times \mathbf{H} - (\sigma - \imath\omega\epsilon)\mathbf{E} &= \mathbf{s}_r(\omega), && \text{(2.7b)} \end{aligned}$$

where ω is some discrete frequency. The boundary conditions used for our experiments over the entire boundary of the spatial domain, $\partial\Omega$, are

$$\mathbf{n} \times \mathbf{H} = 0, \tag{2.8}$$

although other boundary conditions could be used.

To put the equations in frequency and time on equal footing we discuss first the discretization in time. We then treat the equations for both (2.6) and (2.7) in a similar way for the spatial discretization.

2.1 Discretization in time

In order to select a method for the time discretization of (2.6) we first note that given typical earth parameters, over very short time scales ($0 - 10^{-7}$ sec) Maxwell's equations represent wave phenomena, while over longer times the equations tend to have heavy dissipation. Considering also the range of conductivities (ground-air) there are many time scales in the system. Thus,

the equations are very stiff (Ascher and Petzold, 1998). If explicit methods are used, one must take extremely small steps in order to retain stability for such a problem. We therefore turn to implicit methods.

Common methods for very stiff equations are based on backward differentiation formulas (BDF) or on collocation at Radau points (Ascher and Petzold, 1998; Hairer and Wanner, 1991; Bastian, 1999; Heling, 1998; Turek, 1999). These methods have stiff decay (or, are L-stable) whereas conservative methods such as midpoint or trapezoidal are only A-stable. The latter exhibit oscillatory behavior unless the initial fast layer of the solution is resolved, while BDF or Radau methods strongly attenuate high frequencies of the error. Thus, even if the initial time layer is skipped (by taking a time step which is larger than its width) an accurate solution may be obtained away from this layer.

In our case, resolving the initial time layer where the fields change rapidly is not necessary for the inverse problem because the measurements are typically taken at late times. Therefore, it is natural to use a BDF or Radau type method for the solution of the forward problem, as this will result in more efficient computing procedures.

The simplest member of both these families of stiff integrators is the backward Euler method. It is only first order accurate; however, most geophysical systems produce sources which are merely continuous in time globally. This implies that the electric and magnetic fields are only once differentiable and no advantage is obtained by using a higher order discretization in time regions of lower smoothness.

Even more interesting is the effect of the discretization of the forward problem (which generates A) on the adjoint equation (1.5b). It is easy to show that using BDF for the forward problem yields a forward differentiation formula for the adjoint problem, but with a terminal end condition rather than an initial one. Thus, the adjoint equation is essentially integrated backwards in time. If the adjoint equation had a smooth right hand side then these methods would generate a faithful, stable discretization for the Lagrange multipliers. However, unfortunately, this is not the case. Note that the right hand side for the adjoint equation (1.5b) is the noise which is further sampled at discrete points in time. (The operator $\mathcal{Q}$ in (1.2) involves a combination of δ-functions in time.) As such, the right hand side of the adjoint equation is not smooth and the Lagrange multipliers are therefore generally discontinuous, although bounded, at the observation times. Whereas the value of recovering accurate Lagrange multipliers can be (and has been)

debated, there is hardly any incentive here to use a more complicated (and more expensive) method than backward Euler.

The above discussion may suggest using backward Euler at the data points, and then switching to a more accurate BDF or Radau method between the data time locations. In our case we consider measurements at most discretization times, and therefore we simply use backward Euler for the discretization of the problem. This leads to the following system of equations semi-discretizing (2.6), (2.8) over a time step $[t_{n-1}, t_n]$. Letting $\alpha_n = (t_n - t_{n-1})^{-1}$, the equations become

$$\nabla \times \mathbf{E}^n + \alpha_n \mu \mathbf{H}^n = \alpha_n \mathbf{H}^{n-1} \equiv \mathbf{s}_H \tag{2.9a}$$

$$\nabla \times \mathbf{H}^n - (\sigma + \alpha_n \epsilon)\mathbf{E}^n = \mathbf{s}_r^n - \alpha_n \epsilon \mathbf{E}^{n-1} \equiv \mathbf{s}_E \tag{2.9b}$$

$$\mathbf{n} \times \mathbf{H}^n = 0. \tag{2.9c}$$

The superscripts in (2.9) denote the time level, with solution quantities at n being unknown. Apparently, the systems (2.9) and (2.7) have a very similar form.

This system requires initial conditions for both $\mathbf{E}$ and $\mathbf{H}$. If we have a source which is zero before the initial simulation time then we set $\mathbf{E}^0 = \mathbf{H}^0 = 0$. However, if the source is assumed static before time zero then $\mathbf{E}^0 = \nabla\phi^0$ and we need to calculate $[\phi^0, \mathbf{H}^0]$ by solving the electro- and magnetostatic problems. This would yield a consistent initialization (Ascher and Petzold, 1998; Hairer and Wanner, 1991). Here we have used the method proposed in Haber(2000) for the solution of the static problems.

2.2 Reformulation

The semi-discrete system (2.9) and the system in frequency (2.7) apparently have the same form. Indeed, they can both be written as

$$\begin{aligned}
&\nabla \times \mathbf{E} + \alpha\mu\mathbf{H} = \mathbf{s}_H \quad \text{in } \Omega, \\
&\nabla \times \mathbf{H} - (\sigma + \alpha\epsilon)\mathbf{E} = \mathbf{s}_E \quad \text{in } \Omega, \\
&\mathbf{n} \times \mathbf{H} = 0 \quad \text{on } \partial\Omega,
\end{aligned}$$

where $\alpha = -\imath\omega$ in the frequency domain, and $\alpha = (t_n - t_{n-1})^{-1}$ for the time domain. Let us denote $\hat{\sigma} = \sigma + \alpha\epsilon$. As discussed in Haber et al (2000a) and Haber and Ascher(2001a), this form is not favorable for iterative solvers, especially when $|\alpha\hat{\sigma}|$ is small (for example in the air, with a large time step or

a low frequency). We therefore reformulated the problem prior to discretizing it further such that it is more amenable to applying standard iterative solvers.

A Helmholtz decomposition with Coulomb gauge is applied, decoupling the **curl** operator:

$$\begin{aligned} &\mathbf{E} = \mathbf{A} + \nabla\phi, \qquad \nabla \cdot \mathbf{A} = 0 \quad \text{in } \Omega \\ &\mathbf{A} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega. \end{aligned}$$

After adding a stabilization term and differentiating (Haber and Ascher, 2001a), this leads to the system

$$\nabla \times \mu^{-1} \nabla \times \mathbf{A} - \nabla \mu^{-1} \nabla \cdot \mathbf{A} + \alpha \hat{\sigma}(\mathbf{A} + \nabla\phi) = \alpha \mathbf{s} \tag{2.10a}$$

$$\nabla \cdot (\hat{\sigma}(\mathbf{A} + \nabla\phi)) = \nabla \cdot \mathbf{s} \tag{2.10b}$$

in Ω, subject to

$$\mathbf{n} \times \nabla \times \mathbf{A} = 0, \qquad \mathbf{n} \cdot \mathbf{A} = 0, \tag{2.10c}$$

$$\mathbf{n} \cdot \nabla\phi = 0, \tag{2.10d}$$

on the boundary $\partial\Omega$. This system is discretized next.

2.3 Discretization in space and solution of the discrete system

Following Haber and Ascher(2001a) and Haber et al.(2000a), we use a finite volume approach for the discretization of (2.10) on an orthogonal, staggered grid. We choose to discretize $\mathbf{A}$ on cell faces and ϕ at cell centers. This is closely related to mixed type finite element methods (Brezzi and Fortin, 1991; Bossavit, 1998; Haber, 2000). Note that the modified conductivity $\hat{\sigma}$ is averaged harmonically at cell faces, whereas the permeability is averaged arithmetically at edges (Haber and Ascher, 2001a).

We write the fully discretized system as

$$\begin{pmatrix} L_\mu + \alpha M_{\hat{\sigma}} & \alpha M_{\hat{\sigma}} \nabla_h \\ \nabla_h \cdot M_{\hat{\sigma}} & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h \end{pmatrix} \begin{pmatrix} \mathbf{A} \\ \phi \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{s} \\ \nabla_h \cdot s \end{pmatrix} \tag{2.11}$$

where $\nabla_h \cdot, \nabla_h \times$ and ∇_h are matrices arising from the discretization of the corresponding continuous operators, $M_{\hat{\sigma}}$ arises from the operator $\hat{\sigma}(\cdot)$ and L_μ is the discretization of the operator $\nabla \times \mu^{-1} \nabla \times \ - \nabla \mu^{-1} \nabla \cdot$.

This linear system can be solved using standard iterative methods (Saad, 1996) and effective preconditioners can be designed for it (Haber et al., 2000a; Aruliah and Ascher, 2002). Briefly, for small enough α the system is dominated by its diagonal blocks and therefore a good preconditioner can be obtained by using an approximation of the matrix

$$\begin{pmatrix} L_\mu & 0 \\ 0 & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h \end{pmatrix}. \tag{2.12}$$

It is possible to use one multigrid cycle (Aruliah and Ascher, 2002) or an Incomplete LU factorization (ILU) (Saad, 1996; Haber and Ascher, 2001a; Haber et al., 2000a) of (2.12) to obtain an effective preconditioner.

For larger α (i.e. high frequency), the above preconditioner may not suffice and a block preconditioner (or its approximation) of

$$\begin{pmatrix} L_\mu + \alpha M_{\hat{\sigma}} & \alpha M_{\hat{\sigma}} \nabla_h \\ 0 & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h \end{pmatrix} \tag{2.13}$$

is better. Here we have used the ILU for the approximation of the main blocks of (2.13). It must be understood, though, that our entire discretization is not suitable for high frequency parameter regimes or where wave phenomena dominate, and we do not propose to compensate for such inadequacy by manipulating preconditioners.

2.4 Formulating the forward problem

In many applications of our techniques we are concerned with multiple sources and with multiple frequencies or time steps. For solving the inverse problem it is useful to formulate the forward problem in a uniform way.

2.4.1 Formulation for the frequency domain

Assume we have a multiple source/frequency experiment. As explained in the introduction, we consider an optimization problem which involves the forward problem as a system of equality constraints. We therefore formulate the k^{th} experiment in real arithmetic, defining

$$A_k = \begin{pmatrix} L_\mu & 0 & \omega_k M_{\hat{\sigma}} & \omega_k M_{\hat{\sigma}} \nabla_h \\ \nabla_h \cdot M_{\hat{\sigma}} & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h & 0 & 0 \\ -\omega_k M_{\hat{\sigma}} & -\omega_k M_{\hat{\sigma}} \nabla_h & L_\mu & 0 \\ 0 & 0 & \nabla_h \cdot M_{\hat{\sigma}} & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h \end{pmatrix}, \; q_k = \begin{pmatrix} 0 \\ \nabla_h \cdot \mathbf{s}_r \\ -\omega_k \mathbf{s}_r \\ 0 \end{pmatrix}.$$

The entire frequency system is then

$$A(m)u = \begin{pmatrix} A_1(m) & & & & \\ & A_2(m) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & A_s(m) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_s \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ \vdots \\ q_s \end{pmatrix} = q. \quad (2.14)$$

2.4.2 Formulation for the time domain

We treat the time domain similarly to the multi-frequency domain experiment. However, the system is no longer quite block diagonal and the blocks get somewhat larger. The source term for the time domain problem is

$$\begin{aligned} \alpha_n \mathbf{s} &= \alpha_n \nabla_h^T \times \mathbf{H}_{n-1} + \epsilon \alpha_n^2 \mathbf{E}_{n-1} - \alpha_n \mathbf{s}_r^n = \\ & \alpha_n \nabla_h^T \times \mathbf{H}_{n-1} + \epsilon \alpha_n^2 (\mathbf{A}_{n-1} + \nabla \phi_{n-1}) - \alpha_n \mathbf{s}_r^n \end{aligned}$$

and, from Maxwell's equations, we have

$$\alpha_n^{-1} M_\mu^{-1} \nabla_h \times \mathbf{A}_n + \mathbf{H}_n - \mathbf{H}_{n-1} = 0.$$

Thus, we can write the problem in a block bidiagonal structure for $\mathbf{A}, \phi, \mathbf{H}$ as

$$\begin{pmatrix} A_1(m) & & & & \\ B_2 & A_2(m) & & & \\ & B_3 & A_3(m) & & \\ & & \ddots & \ddots & \\ & & & B_s & A_s(m) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_s \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ \vdots \\ q_s \end{pmatrix} \quad (2.15)$$

where

$$u_n = \begin{pmatrix} \mathbf{A}_n \\ \phi_n \\ \mathbf{H}_n \end{pmatrix}, \qquad A_n(m) = \begin{pmatrix} L_\mu + \alpha_n M_{\hat{\sigma}} & \alpha_n M_{\hat{\sigma}} \nabla_h & 0 \\ \nabla_h \cdot M_{\hat{\sigma}} & \nabla_h \cdot M_{\hat{\sigma}} \nabla_h & 0 \\ \alpha_n^{-1} M_\mu^{-1} \nabla_h \times & 0 & I \end{pmatrix}$$

$$q_n = \begin{pmatrix} -\alpha_n {\mathbf{s}_r}^n \\ -\nabla_h \cdot \mathbf{s}_r^n \\ 0 \end{pmatrix}, \quad B_n = \begin{pmatrix} -\epsilon \alpha_n^2 & -\epsilon \alpha_n^2 \nabla_h & -\alpha_n \nabla_h^T \times \\ -\nabla_h \cdot \epsilon \alpha_n & -\nabla_h \cdot \epsilon \alpha_n \nabla_h & 0 \\ 0 & 0 & -I \end{pmatrix}.$$

In the case of multiple sources we obtain a block diagonal system where each block has the same structure as (2.15). Note that only the diagonal blocks in (2.15) depend on the conductivity. Also, once we have an efficient solver for one block $A_k(m)$ as described in the previous subsection 2.3, solving the forward problem (2.15) or (2.14) at s times the cost is straightforward.

To test our forward solvers we made comparisons with other existing codes. For results see Appendix A.

3 The Discrete Inverse Problem

Having defined the discrete forward problem (1.5a) we next form the discrete constrained optimization problem (1.4). For this we discretize the regularization operator $\mathcal{R}$ on the same grid. Denote the result of discretizing the weighted gradient $(\sqrt{\alpha}_1 m_x, \sqrt{\alpha}_2 m_y, \sqrt{\alpha}_3 m_z)^T$ by Wm, i.e., W is a weighted difference matrix. Then

$$R(m) = m^T W^T W m. \tag{3.16}$$

The matrix $W^T W$ is a discretization of the weighted Laplacian with natural boundary conditions. The parameters α_j, which are hidden by the notation in (3.16), are chosen as per our a-priori information and are incorporated into the matrix. With this discretization we next form the nonlinear system (1.5).

In order to calculate the matrix G in (1.5c) we need to differentiate the forward modeling matrix times a vector with respect to m. This may look complicated at first; however, note that the matrix A in both frequency and time is made of blocks and each block depends on m only through the matrix $M_{\hat{\sigma}}$. Therefore, if we know how to calculate

$$N(m, v) = \frac{\partial [M_{\hat{\sigma}}(m) v]}{\partial m}$$

then we can differentiate any product involving M_σ. For example,

$$\frac{\partial}{\partial m}[\nabla_h \cdot M_\sigma \nabla_h w] = \nabla_h \cdot N(m, \nabla_h w).$$

To calculate this derivative, we recall that $M_{\hat{\sigma}}$ operates on the discrete $\mathbf{A}$ or $\nabla_h \phi$ which are cell face variables. The matrix is diagonal and each of its elements has the form

$$M_{\hat{\sigma}}^{(ii)} = 2(\hat{\sigma}_1^{-1} + \hat{\sigma}_2^{-1})^{-1}$$

where $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are the values of $\hat{\sigma}$ at the two sides of the face of the cell. From this form it is clear that $M_{\hat{\sigma}}^{-1}$ is *linear* with respect to $\hat{\sigma}^{-1}$ and therefore the matrix

$$N_r(v) = \frac{\partial[M_{\hat{\sigma}}^{-1} v]}{\partial[\hat{\sigma}^{-1}]}$$

is *independent of* $\hat{\sigma}$ and depends only on the vector field v at each cell. Using this observation and the chain rule we can easily calculate N,

$$\begin{aligned} N(m,v) &= \frac{\partial[M_{\hat{\sigma}}(m)v]}{\partial m} = \frac{\partial\left[((M_{\hat{\sigma}}(m))^{-1})^{-1}\, v\right]}{\partial m} = \frac{\partial\left[((M_{\hat{\sigma}}(m))^{-1})^{-1}\, v\right]}{\partial[\hat{\sigma}^{-1}]}\,\frac{\partial[\hat{\sigma}^{-1}]}{\partial m} = \\ & [M_{\hat{\sigma}}^{-1}]^{-2}\, N_r(v)\,\mathrm{diag}(\exp(-m)) = M_{\hat{\sigma}}^{2}\, N_r(v)\,\mathrm{diag}(\exp(-m)). \end{aligned}$$

We can now proceed and solve the discrete nonlinear system of equations (1.5) by some variant of Newton's method. Because of the chosen form of $\mathcal{R}$ we may use the Gauss-Newton method. Thus, in a typical iteration for given u, λ and m we differentiate (1.5) with respect to these variables and, dropping second order information, obtain the following linear system of equations for the corrections $\delta u, \delta\lambda$ and δm:

$$\begin{pmatrix} A & 0 & G \\ Q^T Q & A^T & 0 \\ 0 & G^T & \beta W^T W \end{pmatrix} \begin{pmatrix} \delta u \\ \delta\lambda \\ \delta m \end{pmatrix} = - \begin{pmatrix} \mathcal{L}_\lambda \\ \mathcal{L}_u \\ \mathcal{L}_m \end{pmatrix}. \tag{3.17}$$

See, e.g., (Nocedal and Wright, 1999; Dennis and Schnabel, 1996; Haber et al., 2000b).

4 Solution of the Linear system

The permuted KKT system (3.17) is very large with possibly millions of unknowns. It is strongly coupled, because it discretizes a strongly coupled PDE system, and it is indefinite. Therefore, special iterative linear algebra techniques are needed in order to solve it. As usual, the crux of the matter is designing an efficient preconditioner. One family of preconditioners for the solution of this system is obtained by approximating the block LU decomposition of its inverse (Haber and Ascher, 2001b; Biros and Ghattas, 2000). For a careful development of this method we point the reader to Haber and Ascher(2001b). Here is a synopsis.

It is easy to show that the system (3.17) can be decomposed into

$$\begin{pmatrix} A & 0 & G \\ Q^TQ & A^T & 0 \\ 0 & G^T & \beta W^TW \end{pmatrix}^{-1} = \tag{4.18}$$
$$\begin{pmatrix} A^{-1} & 0 & -A^{-1}GH_{red}^{-1} \\ 0 & A^{-T} & -A^{-T}Q^TJH_{red}^{-1} \\ 0 & 0 & H_{red}^{-1} \end{pmatrix} \cdot \begin{pmatrix} I & 0 & 0 \\ -Q^TQA^{-1} & I & 0 \\ -J^TQA^{-1} & -G^TA^{-T} & I \end{pmatrix}.$$

where $J = QA^{-1}G$ is the sensitivity matrix and $H_{red} = J^TJ + \beta W^TW$ is the reduced Hessian.

One need not actually calculate A^{-1}, A^{-T} and J but rather generate an approximation in order to precondition the system (3.17). If we have a matrix B such that, for any appropriate vector v, Bv approximates $A^{-1}v$, and a matrix M_{red} such that $M_{red}v$ approximates $H_{red}^{-1}v$ then we can calculate the action of the preconditioner M for the KKT system (3.17) as follows. For a vector v, obtaining $x = Mv$, we write v and x in their components form $v = [v_\lambda^T, v_u^T, v_m^T]^T$ and $x = [x_u^T, x_\lambda^T, x_m^T]^T$, and obtain the following algorithm.

Preconditioning Algorithm:

1. $w_1 = Bv_\lambda$;
2. $w_2 = B^T(v_u - Q^TQw_1)$;
3. $w_3 = v_m - G^Tw_2$;
4. $x_m = M_{red}w_3$;
5. $x_u = w_1 - BGx_m$;
6. $x_\lambda = B^T(v_u - Q^TQx_u)$.

One option to generate B is to use the preconditioners discussed in (2.3), that is, for the frequency domain we use the matrix $B^{-1} = \hat{A}$ which has the

structure

$$\begin{pmatrix} \hat{A}_1(m) & & & & \\ & \hat{A}_2(m) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \hat{A}_s(m) \end{pmatrix} \tag{4.19}$$

where $\hat{A}_k(m)^{-1}$ represents the preconditioner of the forward problem (2.12) or (2.13).

For the time domain (2.15) we use the approximation

$$\begin{pmatrix} \hat{A}_1(m) & & & & \\ B_2 & \hat{A}_2(m) & & & \\ & B_3 & \hat{A}_3(m) & & \\ & & \ddots & \ddots & \\ & & & B_s & \hat{A}_s(m) \end{pmatrix}. \tag{4.20}$$

As in Vogel(1999) and Haber and Ascher(2001a) we use the sparse matrix $\beta W^T W$ in order to approximate the reduced Hessian. Such an approximation works well for large enough β.

These preconditioners are stationary, that is they do not change with the iteration, and therefore standard iterative techniques can be used. We have used the Symmetric QMR (Freund and Jarre, 1996) with which we had good experience in (Haber and Ascher, 2001b).

A second option for the preconditioner, which is especially suited for parallel implementation in the time domain, is to use an inexact solver with a very rough tolerance (here we have used 10^{-2}) in order to approximate A^{-1} and A^{-T}. That is, we use another Krylov method (BICGSTAB) with the preconditioner (4.19) for both frequency and time. In this case the preconditioner is decoupled for both frequency and time. The problem with this type of preconditioner is that it is not stationary. That is, the matrices B which approximate A^{-1}, change at each iteration because they are an inexact solution of the forward problem using a non-stationary iterative technique. As such, this kind of preconditioner cannot be used in a straightforward manner in iterative Krylov space methods, and we turn to a special class of Krylov methods known as flexible methods, which allows the change of the preconditioner at each iteration with the cost of extra storage. This family includes

the flexible GMRES (FGMRES), FGCR (Saad, 1996) and FQMR. Here we have used a symmetric version of the FQMR and the FGCR for the solution of the system. For implementation issues of these algorithms the reader is referred to Saad(1996) and Freund and Jarre(1996).

5 Optimization Issues

In this brief section we return to the optimization problem. Within each nonlinear iteration the KKT system (3.17) is solved inaccurately and thus the SQP (Sequential Quadratic Programming) algorithm is not strictly followed(Nocedal and Wright, 1999). Therefore, care must be exercised so that at the end of the process we achieve both optimality, (i.e. have a rough solution to the optimization problem) and feasibility (i.e. solve Maxwell's equations sufficiently well).

In order to achieve this goal, we use two safeguards. First, we use the method of secondary correction. Thus, *after* each inexact Newton step we apply additional iterations to the solution of the forward problem to reduce the residual. Beginning with the updated m and current value of u we solve $A(m)u = q$ to reduce the misfit by a further order of magnitude. This has the effect that the computed solution converges towards feasibility (solving Maxwell's equations) faster than it converges to optimality. We use this property in our convergence criteria noting that, as in many inverse problems, we can take the optimization "less seriously" than we take the constraints. Thus, we can terminate the optimization process at a relatively large tolerance (for example 10^{-3}) while fitting the constraint to a much smaller tolerance (say 10^{-6}).

After each such iteration we test a decrease in a merit function which is a combination of the optimality and feasibility criteria,

$$\phi_1 = \|Qu - b^{obs}\|^2 + \beta R(m) + \mu_1 \|A(m)u - q\|_1.$$

This merit function was suggested in Nocedal and Wright(1999) and we have used it successfully for a simple 1D inverse problem(Haber et al., 2000b). The parameter μ_1 is chosen as in Haber et al.(2000b). If the merit function decreases then the step is accepted; however, if it does not then we use a line search on the updates of both u and m.

Unfortunately, there is no theory that guarantees decrease in the merit function for an inexact solution of the KKT system. We therefore add a

last safety mechanism. In the case where the line search fails we turn to an unconstrained Gauss-Newton iteration which is outlined in Appendix B. This updates m. The fields u and the Lagrange multipliers λ are upgraded according to procedure $N4$(Haber et al., 2000b).

Another basic issue is the selection of the regularization parameter. Here, we use the discrepancy principle; that is, we aim for a certain target misfit. To hit this target misfit, we use continuation in the regularization parameter. Thus, we start with a guess which is obviously larger than the true regularization parameter and solve the optimization problem. If β is large enough, then such a solution is achieved in 1-2 steps. We then decrease the regularization parameter and solve the problem again starting from the previously obtained solution. To guarantee that the first regularization parameter is large enough we use the estimate

$$\beta_0 = 100\|QBGv\|^2/\|Wv\|^2$$

where v is a random vector and B is our approximation to A^{-1}. This selection guarantees that at the initial step the model objective function dominates the optimization problem.

6 Examples

6.1 Inversion of Frequency domain data

As a first example we invert synthetic data from a grounded source. The transmitter and receiver geometry is the same as in an actual CSAMT field survey but the conductivity model is simplified compared to the true earth. Figure 1 shows the survey geometry and the 3-D model. The transmitter is a 1 km grounded wire that is a few kilometers west and north of the survey area. Within the survey area there are 11 east-west lines with a line spacing of 100 m. On each line there are 28 stations at intervals of 50 m. Five components $(E_x, E_y, H_x, H_y, H_z)$ represented as real and imaginary parts, at 3 frequencies (16 Hz, 64 Hz, and 512 Hz) result in 3080 data. We have used the code(Haber et al., 2000a) to generate the data. To simulate realistic noise, we added Gaussian noise that was 2% in the amplitude and 2 degrees in phase. Two representative field components at a frequency 512 Hz are shown in the top portion of Figure 2.

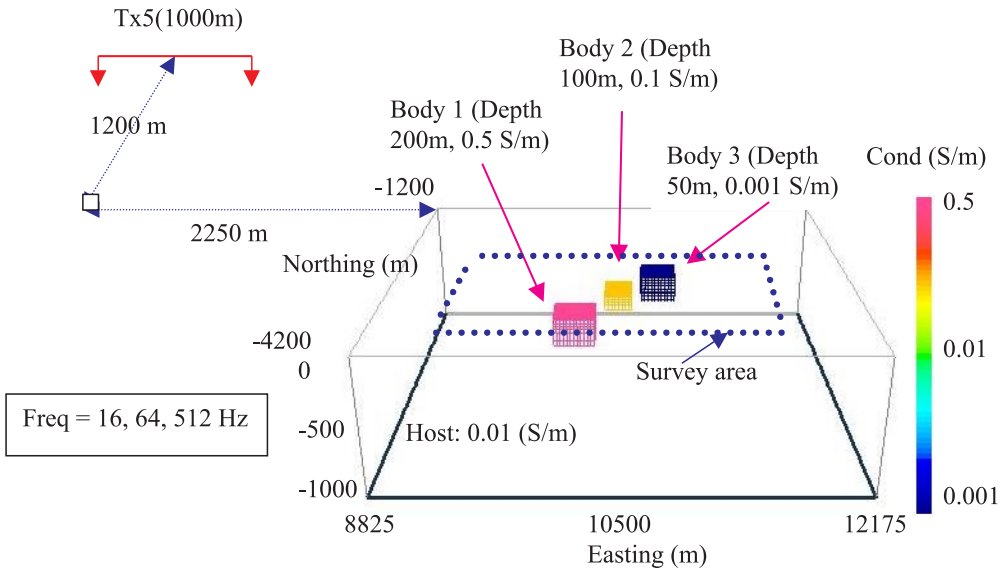

Figure 1: Experimental setting for frequency domain inversion. The cuboid indicates the earth volume in which the inversion is carried out. Data are acquired within the dotted rectangle at the surface.

The 3-D volume for inversion, $3350m \times 3000m \times 2000m$, was discretized into $64 \times 50 \times 30$ cells. The transmitter lies considerably outside this domain. To handle this, we have assumed that the fields at the edges of the model volume are equal to the primary fields in a homogeneous earth. We used the 1-D code (Routh and Oldenburg, 2001) to compute these fields. The inversion began with a homogeneous half-space that was equal to the true background conductivity. After 5 iterations, the final misfit was 3.2%. Three slices of the recovered model are shown in Figure 3. The resistive and two conductive targets are reasonably well recovered. The predicted data at 512 Hz are displayed if Figure 2, and they show good agreement with the true data.

The convergence results for the experiment are presented in Table 2. The table shows the regularization parameter β and the misfit which we got from solving the optimization problem with that β. For every β we record the number of nonlinear iterations, the number of iterations needed to solve the KKT matrix using the FGCR, the PDE residual $\|A(m)u - q\|/\|q\|$ and the

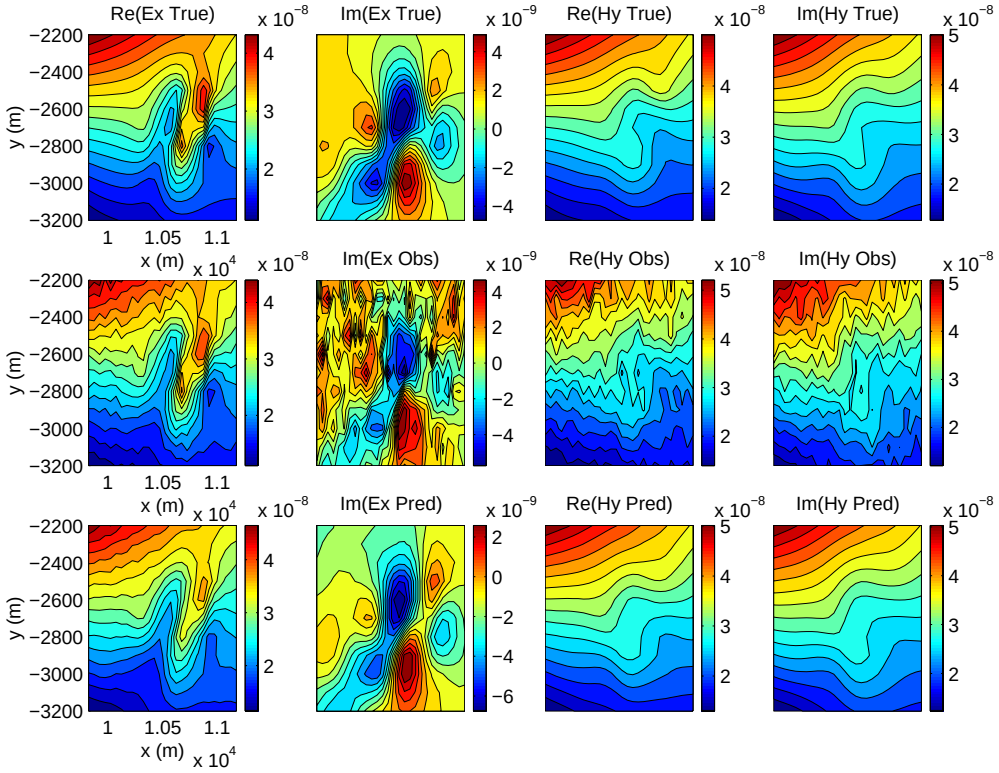

Figure 2: The accurate E_x, H_y data for 512 Hz are shown in the top row. The error contaminated data are shown in the middle row and the bottom row displays the data predicted from the inverted model.

relative gradient. For this example the starting value of the regularization parameter was $\beta = 100$. Three nonlinear iterations were required to achieve an adequate solution to the optimization problem and the final misfit was 6.0%. Within each nonlinear iteration, only three or four iterations were required to solve the large KKT system. The regularization parameter was then reduced to $\beta = 1.0$ and the process continued.

6.2 Inversion of time domain data

We consider the case of a square loop with dimensions of 50×50 meters located just above the earth's surface. The transmitter current is a step-

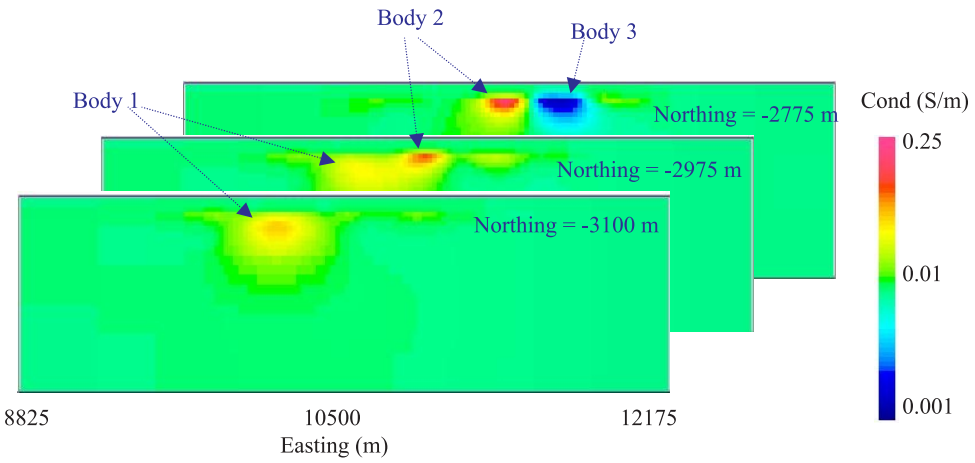

Figure 3: Three slices through the recovered 3D conductivity model obtained by inverting synthetic CSAMT data

off at time zero and responses are measured in 18 logarithmically spaced times between $10^{-4} - 10^{-1}$ sec. The earth model is a conductive sphere ($\sigma = 0.1 S/m$, radius-$15m$) buried in a uniform halfspace ($0.01 S/m$) and three components of the magnetic field are acquired at 20 depths in each of four boreholes that surround the conductor. The geometry is provided in Figure 4.

We use a grid of 64^3 cells in space. The grid is uniform around the loop area and stretched logarithmically at the boundary. For the discretization in time, we have used 32 time steps which are equally spaced on a log-grid from 10^{-7} to 10^{-1} sec.

The inverse problem is performed on a smaller grid, $40 \times 40 \times 32$ grid in space with the same grid in time. The inversion begins with the uniform halfspace equal to the true background conductivity. The convergence results for the experiment are presented in Table 2. Overall the data were fit to an average of about 2%, and a plot of the observed and predicted data from a representative station is shown in Figure 5. A cross-section through the inverted model is shown in Figure 6.

$\beta = 100$	Final misfit = 0.06		
Nonlinear iter	KKT iter	$\|\|Au - q\|\|/\|\|q\|\|$	Rel-grad
1	4	$3e-2$	2e-1
2	4	$2e-4$	3e-2
3	3	$2e-6$	5e-4
$\beta = 1e0$	Final misfit = 0.03		
Nonlinear iter	KKT iter	$\|\|Au - q\|\|/\|\|q\|\|$	Rel-rad
1	8	$1e-6$	3e-3
2	6	$8e-7$	9e-4

Table 1: Optimization path for the frequency domain inversion.. The number of iterations to solve the KKT system is listed in column 2. Column 3 indicates the relative accuracy to which the data constraints are solved. The fourth column tabulates the relative gradient to assess progress in solving the optimization problem. At the stationary point, the gradient is zero.

7 Discussion

We have shown how time and frequency domain electromagnetic data can be inverted with a procedure that simultaneously recovers the electrical conductivity model and the corresponding fields. The forward modelling equations are included as constraints but these are not satisfied until the optimization is complete (that is, a stationary solution of a Lagrangian has been found). Thus, in this procedure, it is not necessary to solve the forward problem exactly at intermediate iterations and effectively the forward problem is solved in tandom with the inverse problem. This has potential for a solution to be reached more quickly than in traditional unconstrained optimization approaches that are formulated to minimize a function of the conductivity only. The all-at-once methodology generates a large matrix that needs to be inverted but the numerical example shows that such computations are tractable.

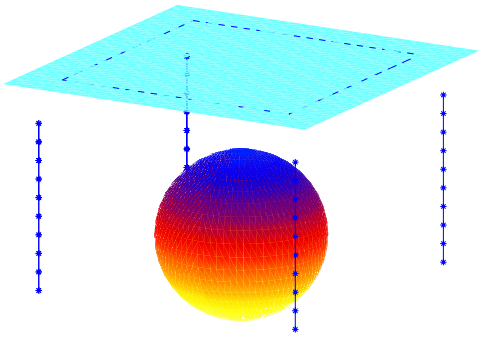

Figure 4: Borehole geometry for the time domain problem. The dashed line at the surface is the loop source. All three components of the magnetic field are acquired in the boreholes.

$\beta = 1e-1$	Final misfit = 0.1		
Nonlinear iter	KKT iter	$\|\|Au-q\|\|/\|\|q\|\|$	Rel-grad
1	2	$3e-3$	1e-2
2	3	$2e-4$	4e-3
3	2	$7e-6$	1e-3
4	2	$9e-7$	3e-4
$\beta = 1e-2$	Final misfit = 0.04		
Nonlinear iter	KKT iter	$\|\|Au-q\|\|/\|\|q\|\|$	Rel-grad
1	7	$4e-6$	2e-3
2	5	$6e-7$	7e-4
$\beta = 1e-3$	Final misfit = 0.02		
Nonlinear iter	KKT iter	$\|\|Au-q\|\|/\|\|q\|\|$	Rel-grad
1	8	$2e-6$	3e-3
2	7	$8e-7$	9e-4

Table 2: Optimization path for the time domain inversion. The number of iterations to solve the KKT system is listed in column 2. Column 3 indicates the relative accuracy to which the data constraints are solved. The fourth column tabulates the relative gradient to assess progress in solving the optimization problem.

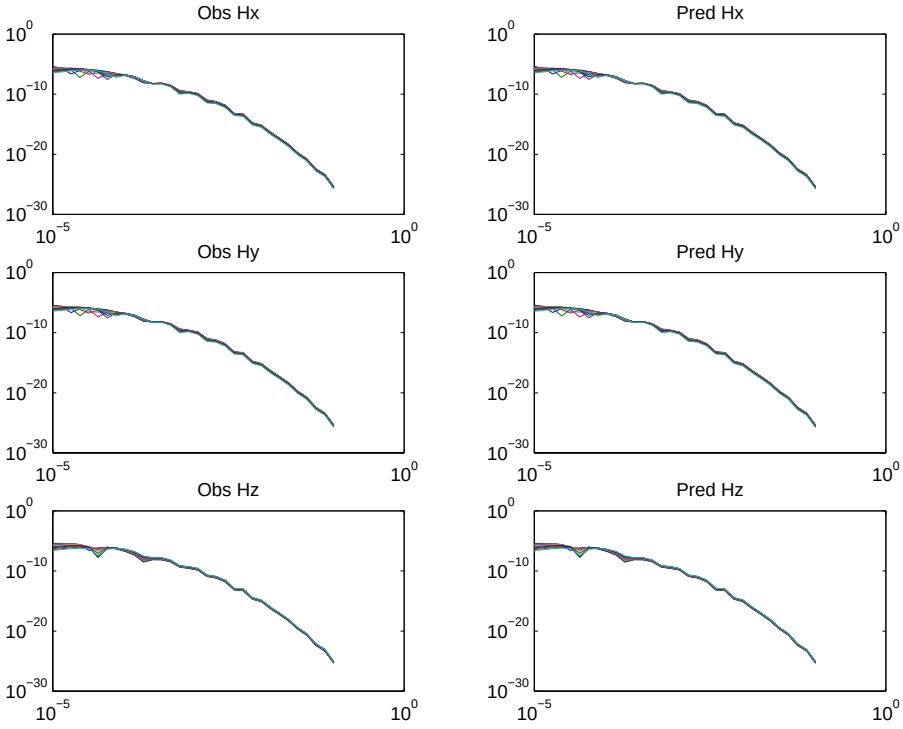

Figure 5: Observed and predicted magnetic field data for all receivers in the borehole survey.

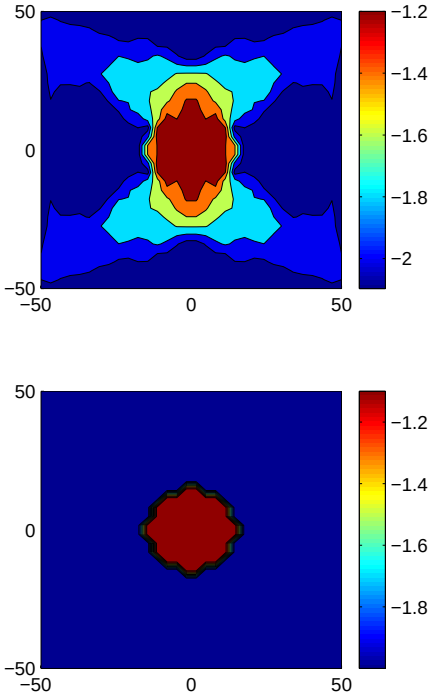

Figure 6: Result of inversion of time domain data. The top panel is a horizontal slice of the recovered conductivity at a depth corresponding to the center of the sphere. The true conductivity is shown in the bottom panel.

8 Appendix A - Testing the forward solution

8.1 Frequency domain

In order to verify the numerical solutions of the forward problem we have compared our codes with results from existing codes. We do not attempt to describe these comparisons in detail here as they are not the focus of this paper. However, for the frequency domain we have made comparisons with the MT code in (Madden and Mackie, 1989), the integral equation code (SYSEM) (Xiong, 1992) using both loop and wire sources, one dimensional solutions for electric sources obtained from the codes(Routh and Oldenburg, 2001) and one dimensional solutions for magnetic sources obtained from the codes in (Farquharson and Oldenburg, 1993). In all cases the results were acceptably close, and differences were most likely due to gridding issues and interpolation of fields to the same locations. As an explicit test we present the following.

We generate the electromagnetic responses due to a conductive block in a uniform host at a set of 31 frequencies. These responses are converted to the time domain using a digital filter (Christensen, 1990). The comparison is made with time domain fields generated from an integral equation code SYSEM (Xiong, 1992). Those fields were also generated in the frequency domain and transformed to time. The fact that our converted frequency results match those from SYSEM is validation of our frequency modelling code. The advantage of presenting results in the time domain is that the data plot is simple and the example can be compared with the direct computation in the time domain presented in the next section of the appendix.

The source is a $1km \times 1km$ loop that is $500m$ from the edge of the block. The conductivity of the block is 1 S/m and the halfspace is 0.005 S/m. The geometry is shown in Figure 7. The vertical component of the magnetic field, generated from a step-off current, is plotted at a sequence of times ranging from $20\mu sec$ to $200msec$. Figure 7 shows the model and magnetic field values along an east-west traverse over the conductive block. The results agree well with those obtained from an integral equation code.

8.2 Time domain

We do not have access to other direct solvers for the time-domain problem and thus for verification of our code we have used codes that compute responses in

the frequency domain and then we convert them to time. Here we use an integral equation code (SYSEM) (Xiong, 1992) for a 3D example and responses from a loop over a layered space obtained from the codes(Farquharson and Oldenburg, 1993) for our 1D example.

8.3 1D Conductive and Permeable Earth

As a first example we compute the vertical component of the magnetic field due to a step off current in a loop source on the surface of a conductive and magnetic permeable halfspace. The conductivity is $0.01 S/m$ and the magnetic susceptibility $1.0 SI$. In Figure 8 we show the vertical magnetic field responses at the center of the loop for the two cases of a magnetic and non-magnetic earth. Our time domain results are compared to those generated from a frequency domain 1D algorithm (Farquharson and Oldenburg, 1993) in which the conversion to time has been carried out by using a digital filter (Christensen, 1990). The agreement is good and we regard this as a solid test for our time domain code being able to handle both conductive and magnetic units.

8.4 3D Conductive Earth

In a second example we generate the time domain responses due to a conductive block buried in a homogeneous halfspace. The geometry is the same as that used to test the frequency domain algorithm. The source is a $1km \times 1km$ loop that is $500m$ from the edge of the block. The vertical component of the magnetic field, generated from a step-off current, is plotted at a sequence of times ranging from $20 \mu sec$ to $200 msec$. Figure 9 shows the magnetic field values along an east-west traverse over the conductive block. The results agree well with those obtained from an integral equation code SYSEM (Xiong, 1992).

9 Appendix B - Unconstrained optimization for parameter estimation in Maxwell's Equations in 3D

In this appendix we briefly discuss the solution of electromagnetic inverse problems using an unconstrained, inexact Gauss-Newton formulation.

As explained in §5, this approach is complementary to our constrained approach and it serves a few purposes. First, it is a straightforward procedure to implement and it allows us to examine methods for noise estimation, data weighting, model weighting and other practical aspects of the inversion procedure without addressing the more involved numerical issues of the constrained approach. Having decided upon these components of the inversion, the final large inversion is carried out using the constrained methodology. The second reason for implementing the unconstrained approach is that there is no proof of convergence for the constrained approach, and it is possible that the constrained step fails. In this case we may resort to taking an unconstrained step in the constrained inversion algorithm. The last reason for investigating the unconstrained methodology is that the code serves as a base for comparison with the constrained approach.

In the unconstrained approach the constraints are eliminated and the problem (1.4) is transformed into an unconstrained optimization problem of the form

$$\min_m \frac{1}{2}\|QA(m)^{-1}q - b\|^2 + \frac{1}{2}\beta\|W(m - m_{ref}\|^2. \tag{9.21}$$

Differentiating (9.21) with respect to m we obtain the nonlinear gradient system

$$g(m) \;=\; -G^T A^{-T} Q^T (QA^{-1}q - b) + \beta W^T W(m - m_{ref}) = 0. \tag{9.22}$$

Note that in order to evaluate the gradient one must solve the forward and the adjoint problems.

Also, comparing (9.22) with the usual unconstrained formulation we see that the sensitivity matrix can be expressed as

$$J(m) = -QA^{-1}G, \tag{9.23}$$

This is a key observation in the solution of the inverse problem using the unconstrained approach, because, upon applying an iterative method for the

linear system, we need not calculate the large and dense sensitivity matrix but merely evaluate its product with vectors in order to carry out a Gauss-Newton iteration (Haber et al., 2000b; Haber and Ascher, 2001b). The Gauss-Newton iteration can be written as

$$(J^T J + \beta W^T W)\delta m = -g(m). \tag{9.24}$$

In order to solve the system (9.24) we use a preconditioned conjugate gradient method, and thus only products of the form Jv and $J^T w$ are required. Using the decomposition (9.23) this can be achieved involving only sparse matrix operations.

As a preconditioner for our system we use $W^T W$. This works reasonably well as long as β is large enough. In order to further save computational time we use an inexact Gauss-Newton formulation (Kelley, 1999) and solve (9.24) to a rough tolerance (typically $10^{-1} - 10^{-2}$), which usually requires only few conjugate gradient iterations.

10 Acknowledgement

This work was supported by an NSERC IOR grant and an industry consortium Inversion and Modelling of Applied Geophysical Electromagnetic data (IMAGE) project. Participating companies are Newmont Gold Company, Falconbridge, Placer Dome, Anglo American, INCO Exploration & Technical Services, MIM, Cominco Exploration, AGIP, Muskox Minerals, Billiton, Kennecott Exploration Company.

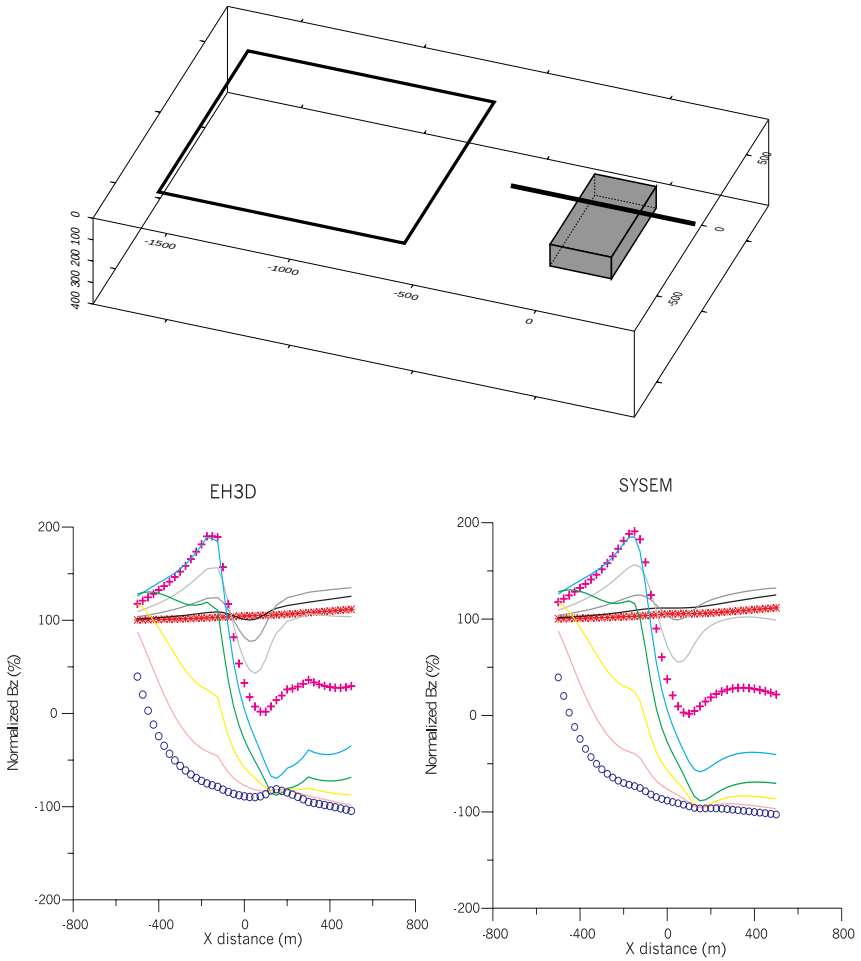

Figure 7: Validation of frequency domain forward modelling. The conductivity model is shown in the upper figure. The dark square at the surface is the source loop. Data are acquired along the solid line shown above the buried prism. Frequency domain data at 31 frequencies are generated and then converted to time. The earliest time channel $20\mu sec$ is shown by 'o' and the latest time channel at $200msec$ is indicated by the '*'.

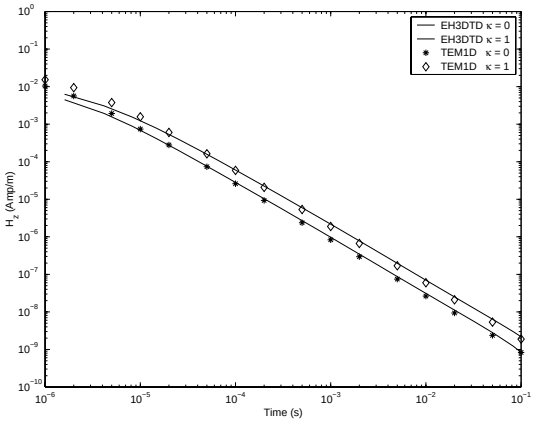

Figure 8: Validation of time domain forward modelling. A comparison between our 3D code and responses from a 1D code. Solid lines represent the results from the 3D code, symbols refer to the 1D code. The upper curve corresponds to a halfspace with unit susceptibility.

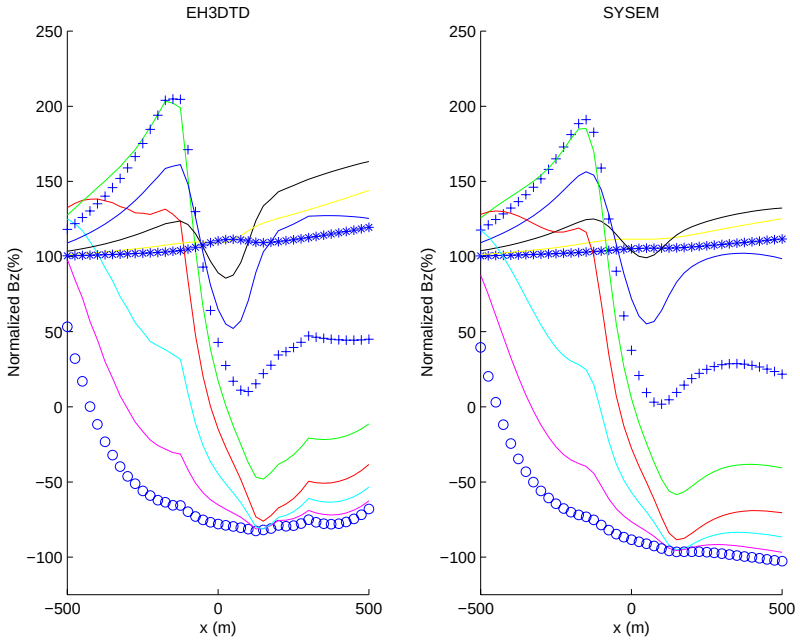

Figure 9: A comparison between our 3D time domain code and SYSEM over a 3D model. Data are acquired along the solid line shown in Figure 7. The earliest time channel $20\mu sec$ is shown by 'o' and the latest time channel at $200msec$ is indicated by the '*'.

11 References

Aruliah, D., and Ascher, U., 2002, Multigrid preconditioning for time-harmonic maxwell's equations in 3D: SIAM J. Scient. Comput.

Ascher, U., and Haber, E., 2001, A multigrid method for distributed parameter estimation problems: ETNA.

Ascher, U., and Petzold, L., 1998, Computer methods for ordinary differential equations and differential-algebraic equations: SIAM, Philadelphia.

Bastian, P., 1999, Numerical computation of multiphase flow in porous media: Ph.D. thesis, Universitat Kiel, Kiel.

Biros, G., and Ghattas, O., 2000, Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization Parts I,II: Preprints.

Borcea, L., Berryman, J. G., and Papanicolaou, G. C., 1996, High-contrast impedance tomography: Inverse Problems, **12**.

Bossavit, A., 1998, Computational electromagnetism. variational formulation, complementarity, edge elements: Academic Press.

Brezzi, F., and Fortin, M., 1991, Mixed and hybrid finite element methods: Springer-Verlag, New York.

Chavent, G., 1989, A non-linear least-square theory for inverse problems A non-linear least-square theory for inverse problems, Springer, Inverse Methods in Action.

Cheney, M., Isaacson, D., and Newell, J., 1999, Electrical impedance tomography: SIAM Review, **41**, 85–101.

Christensen, N. B., 1990, Optimized fast Hankel transform filters: Geophys. Prosp., **38**, 545–568.

Dennis, J. E., and Schnabel, R. B., 1996, Numerical methods for unconstrained optimization and nonlinear equations: SIAM, Philadelphia.

Devaney, A. J., 1989, The limited-view problem in diffraction tomography: Inverse Problems, **5**, 510–523.

Farquharson, C. G., and Oldenburg, D. W., 1993, Inversion of time domain electromagnetic data for a horizontally layered earth: GJI, **114**, 433–442.

Farquharson, C., and Oldenburg, D., 1998, Non-linear inversion using general measures of data misfit and model structure: Geophysics J., **134**, 213–227.

Freund, R. W., and Jarre, F., 1996, A QMR-based interior-point algorithm for solving linear programs: Mathematical Programming, Series B, **76**, 183–210.

Haber, E., and Ascher, U., 2001a, Fast finite volume simulation of 3D electromagnetic problems with highly discontinuous coefficients: SIAM J. Scient. Comput., **22**, 1943–1961.

——— 2001b, Preconditioned all-at-one methods for large, sparse parameter estimation problems: Inverse Problems.

Haber, E., Ascher, U., Aruliah, D., and Oldenburg, D., 2000a, Fast simulation of 3D electromagnetic using potentials: J. Comput. Phys., **163**, 150–171.

——— 2000b, On optimization techniques for solving nonlinear inverse problems: Inverse problems, **16**, 1263–1280.

Haber, E., 2000, A mixed finite element method for the solution of the magnetostatic problem in 3D: Computational Geosciences, **4**, 323–326.

Hairer, E., and Wanner, G., 1991, Solving ordinary differential equations ii: Stiff and differential-algebraic problems: Springer-Verlag.

Heling, R., 1998, Multiphase flow and transport process in the subsurface: Springer.

Huber, P. J., 1964, Robust estimation of a location parameter: Ann. Math. Stats., **35**, 73–101.

Kelley, C., 1999, Iterative methods for optimization: SIAM, Philadelphia.

Madden, T., and Mackie, R., 1989, Three-dimensional magnetotelluric modeling and inversion: Proceedings of the IEEE, **77**, 318–321.

Newman, G., and Alumbaugh, D., 1995, Frequency-domain modelling of airborne electromagnetic responses using staggered finite differences: Geophys. Prospecting, **43**, 1021–1042.

Newman, G., and Alumbaugh, D., 1997a, Three-dimensional massively parallel electromagnetic inversion–i. theory: Geophysical Journal International, **128**, 345–354.

——— 1997b, Three-dimensional massively parallel electromagnetic inversion–ii, analysis of a crosswell electromagnetic experiment: Geophysical Journal International, **128**, 355–367.

Nocedal, J., and Wright, S., 1999, Numerical optimization: New York: Springer.

Parker, R. L., 1994, Geophysical inverse theory: Princeton University Press, Princeton NJ.

Rodi, W., and Mackie, R., 1998, Nonlinear conjugate gradients algorithm for 2d magnetotelluric inversion: Submitted.

Routh, P., and Oldenburg, D., 2001, Electromagnetic coupling in frequency-domain induced polarization: a method for removal: Geophysical Journal International, **145**, 59–76.

Saad, Y., 1996, Iterative methods for sparse linear systems: PWS Publishing Company.

Smith, N., and Vozoff, K., 1984, Two dimensional DC resistivity inversion for dipole dipole data: IEEE Trans. on geoscience and remote sensing, **GE 22**, 21–28.

Tikhonov, A., and Arsenin, V., 1977, Methods for solving ill-posed problems: John Wiley and Sons, Inc.

Turek, S., 1999, Efficient solvers for incompressable flow problems: New York: Macmillan.

Vogel, C., 1999, Sparse matrix computation arising in distributed parameter identification: SIAM J. Matrix Anal. Appl., **20**, 1027–1037.

Weickert, J., 1998, Anisotropic diffusion in image processing: B.G Teubner Stuttgart.

Xiong, Z., 1992, Electromagnetic modeling of 3D structures by the method of system iteration using integral equations: Geophysics, **57**, 1556–1561.