GEOPHYSICS, VOL. 71, NO. 5 (SEPTEMBER-OCTOBER 2006); P. G191–G196, 3 FIGS., 1 TABLE.
10.1190/1.2236381

A limited-memory quasi-Newton inversion for 1D magnetotellurics

Anna Avdeeva[1] and Dmitry Avdeev[2]

ABSTRACT

We apply a limited-memory quasi-Newton (QN) method to the 1D magnetotelluric (MT) inverse problem. Using this method we invert a realistic synthetic MT impedance data set calculated for a layered earth model. The calculation of gradients based on the adjoint method speeds up the inverse problem solution many times. In addition, regularization stabilizes the QN inversion result and a few correction pairs are sufficient to produce reasonable results. Comparison with the L-BFGS-B algorithm shows similar convergence rates. This study is a first step towards the solution of large-scale electromagnetic problems, with a full treatment of the 3D conductivity structure of the earth.

INTRODUCTION

Quasi-Newton (QN) methods have become very popular tools for the numerical solution of electromagnetic (EM) inverse problems (Newman and Boggs, 2004; Haber, 2005). The reasoning behind it is the method requires calculation of gradients only, while at the same time avoiding calculations of second-derivative terms. However, even with the gradients only, the QN methods may require excessively large computation time if the gradients are calculated straightforwardly. An effective way to calculate the gradients is presented in Appendix A. Also, for large-scale inverse problems, the limited-memory QN methods must be applied because their requirements for storage are not as excessive as for other QN methods. In this paper, as a first step to solving the 3D EM case, we have applied a limited-memory QN method for constrained optimization to 1D magnetotelluric (MT) problems. This optimization method is an extension of previous work by Ni and Yuan (1997). As distinct from this earlier work, we implement Wolfe conditions to terminate the line search procedure, as was recommended by Byrd et al. (1995).

In the first section, we present our implementation of the limited-memory QN method for inversion of 1D synthetic MT data. The section includes a simple review of the method, the acceleration of the inverse problem solution, and our choice of regularization parameter. In the second section, we demonstrate the efficiency of our inversion on a synthetic, but realistic, numerical example, along with a comparison with the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm for bound constrained optimization (L-BFGS-B) by Byrd et al. (1995). The results presented are encouraging and suggest that the method has the potential to handle the more geophysically realistic 3D inverse problem. Therefore, our future attempts to solve the 3D MT inverse problem will be founded on the theory we describe here.

1D MT INVERSION

In the frame of 1D magnetotelluric (MT) inversion a layered-earth model is considered and conductivities of the layers are sought. This problem is usually solved by minimization $\min_{\mathbf{m}}\varphi(\mathbf{m},\lambda)$ of the following objective function (Tikhonov, 1963):

$$\varphi(\mathbf{m},\lambda) = \varphi_d(\mathbf{m}) + \lambda\varphi_s(\mathbf{m}), \tag{1}$$

where

$$\varphi_d(\mathbf{m}) = \frac{1}{2}\sum_{j=1}^{M} \beta_j |Z_j - d_j|^2 \tag{2}$$

is the data misfit. Here, $\mathbf{m} = (m_1,\ldots,m_N)^T$ is the vector consisting of the electrical conductivities of the layers; superscript T means transpose; N is the number of layers; $Z_j(\mathbf{m})$ and d_j are the complex-valued, modeled, and observed impedances at the jth period ($j = 1, \ldots, M$), respectively; $\beta_j = \frac{1}{M}\frac{2}{\varepsilon_j^2|d_j|^2}$ are the positive weights; ε_j is the relative error of the impedance $Z_j(\mathbf{m})$ and λ is the regularization parameter. As prescribed by the theory (Tikhonov, 1963), the function of equation 1 has a regularized part (stabilizer) $\varphi_s(\mathbf{m})$. This stabilizer can be chosen in many ways (see, for example, Farquharson and Oldenburg, 1998, and references therein), and moreover, the correct choice of $\varphi_s(\mathbf{m})$ is crucial for a reliable inversion. However,

Manuscript received by the Editor July 28, 2005; revised manuscript received January 27, 2006; published online August 28, 2006.
[1]Dublin Institute for Advanced Studies, School of Cosmic Physics, 5 Merrion Square, Dublin 2, Ireland. E-mail: aavdeeva@cp.dias.ie.
[2]Dublin Institute for Advanced Studies, School of Cosmic Physics, Dublin, Ireland and Russian Academy of Sciences, Institute of Terrestrial Magnetism, Ionosphere and Radiowave Propagation, 142190, Troitsk, Moscow region, Russian Federation. E-mail: d.avdeev@mtu-net.ru.

this aspect of the problem is outside the scope of this paper. Thus, we consider an Occam-type (Constable et al., 1987) stabilizer:

$$\varphi_s(\mathbf{m}) = \sum_{i=2}^{N} \left(\frac{m_i}{m_i^0} - \frac{m_{i-1}}{m_{i-1}^0} \right)^2, \tag{3}$$

where $\mathbf{m^0} = (m_1^0, ..., m_N^0)^T$ is an initial guess model. It is important because the conductivities $m_i (i = 1, ..., N)$ must be nonnegative and reasonable, that the optimization problem given in equation 1 is subject to bounds:

$$\mathbf{l} \le \mathbf{m} \le \mathbf{u}, \tag{4}$$

where $\mathbf{l} = (l_1, ..., l_N)^T$ and $\mathbf{u} = (u_1, ..., u_N)^T$ are the lower and upper bounds, respectively, $l_i \ge 0$ $(i = 1, ..., N)$. An alternative way to keep the conductivities positive is to define m_i as $m_i = \log(\sigma_i - l_i)$, or $m_i = \log\left(\frac{\sigma_i - l_i}{u_i - \sigma_i}\right)$. After such transformations, the bounds extend to infinity, and the constrained problem of equations 1–4 nominally turns to an easier unconstrained problem of equations 1–3. Although these transformations are commonly used, we believe, first, it is not practical that a constrained problem can ever be converted into an unconstrained one in such a simple way; and second, such transformations may slow down the convergence of the solution if a minimum of the objective function of equation 1 is located on or near the bounds.

A limited-memory quasi-Newton method

We notice that the problem posed in equations 1–4 is a typically constrained optimization problem with simple bounds. To solve this problem, Newton-type iterative methods are commonly applied. However, most of these methods are not applicable to large-scale optimization problems because the storage and computational requirements become excessive. To overcome this, a limited-memory quasi-Newton method has been developed [see Nocedal and Wright (1999) for a good introduction]. Let us now describe our implementation of such a technique.

At each iteration step k, the search direction vector $\mathbf{p}^{(k)}$ is calculated as

$$\mathbf{p}^{(k)} = -\mathbf{G}^{(k)}\mathbf{g}^{(k)}, \tag{5}$$

where the symmetric matrix $\mathbf{G}^{(k)}$ is an approximation to the inverse Hessian matrix, $\mathbf{H}^{(k)-1}$, and $\mathbf{g}^{(k)}$ is the gradient

$$\mathbf{g} = \frac{\partial \varphi}{\partial m} = \left(\frac{\partial \varphi}{\partial m_1}, ..., \frac{\partial \varphi}{\partial m_N} \right)^T, \tag{6}$$

calculated at $\mathbf{m} = \mathbf{m}^{(k)}$. The explicit expression for the matrix $\mathbf{G}^{(k)}$ is given in Nocedal and Wright [1999, p. 225, formulas (9.5)]. It is important that the matrix $\mathbf{G}^{(k)}$ is stored implicitly using n_{cp} correction pairs $\{\mathbf{s}^{(i)}, \mathbf{y}^{(i)}: i = k - n_{cp}, ..., k - 1\}$ previously computed as

$$\mathbf{s}^{(i)} = \mathbf{m}^{(i+1)} - \mathbf{m}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbf{g}^{(i+1)} - \mathbf{g}^{(i)}. \tag{7}$$

The main idea behind this approach is to use information from only the most recent iterations and the information from earlier iterations is discarded in the interests of saving storage. Nocedal and Wright (1999, p. 225) advocated n_{cp} between three and 20 to produce satisfactory results. The next iterate $\mathbf{m}^{(k+1)}$ is then found as

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \alpha^{(k)}\mathbf{p}^{(k)}, \tag{8}$$

where the step length $\alpha^{(k)}$ is computed by a line search procedure. This procedure finds a step length that delivers an adequate decrease in the objective function φ along the search direction $\mathbf{p}^{(k)}$. When $\mathbf{p}^{(k)}$ is defined by equation 5 and $\mathbf{G}^{(k)}$ is positive definite, one has

$$\mathbf{p}^{(k)T}\mathbf{g}^{(k)} = -\mathbf{g}^{(k)T}\mathbf{G}^{(k)}\mathbf{g}^{(k)} < 0, \tag{9}$$

and hence, $\mathbf{p}^{(k)}$ is a descent direction.

Let us demonstrate how in our implementation we provide the positive definiteness of the matrix $\mathbf{G}^{(k)}$, required to guarantee the descent direction $\mathbf{p}^{(k)}$. When the vectors $\mathbf{s}^{(k-1)}$ and $\mathbf{y}^{(k-1)}$ satisfy the curvature condition

$$\mathbf{s}^{(k-1)T}\mathbf{y}^{(k-1)} > 0, \tag{10}$$

it can be shown that the matrix $\mathbf{G}^{(k)}$ is positive definite. However, the condition of equation 10 does not always hold, and in such cases one needs to enforce this condition explicitly by imposing restrictions on the line search procedure. The condition of equation 10 is guaranteed to hold if we impose the following Wolfe conditions on the line search:

$$\varphi^{(k)} \le \varphi^{(k-1)} + c_1\alpha^{(k-1)}\mathbf{g}^{(k-1)T}\mathbf{p}^{(k-1)}, \tag{11a}$$

$$\mathbf{g}^{(k)T}\mathbf{p}^{(k-1)} \ge c_2\mathbf{g}^{(k-1)T}\mathbf{p}^{(k-1)}, \tag{11b}$$

with $0 < c_1 < c_2 < 1$, where $\varphi^{(k)} = \varphi(\mathbf{m}^{(k)}, \lambda)$. To verify this, we notice that from equations 7 and 11b it follows that $\mathbf{g}^{(k)T}\mathbf{s}^{(k-1)} \ge c_2\mathbf{g}^{(k-1)T}\mathbf{s}^{(k-1)}$, and therefore, $\mathbf{y}^{(k-1)T}\mathbf{s}^{(k-1)} \ge (c_2 - 1)\alpha^{(k-1)}\mathbf{g}^{(k-1)T}\mathbf{p}^{(k-1)}$. Because $c_2 < 1$ and $\mathbf{p}^{(k-1)}$ is a descent direction, the term on the right is positive, and the curvature condition of equation 10 holds.

Let us now recall that our problem is subject to the simple bounds defined by equation 4. As it was shown above, the Wolfe condition of equation 11b guarantees that the matrix $\mathbf{G}^{(k)}$ is positive definite and hence, $\mathbf{p}^{(k)} = -\mathbf{G}^{(k)}\mathbf{g}^{(k)}$ is a descent direction. But, the condition of equation 11b may not be reached inside the feasible region defined within the bounds set by equation 4. In this case one should modify $\mathbf{s}^{(k-1)}$ of equation 7 as follows

$$\theta \cdot \mathbf{s}^{(k-1)} + (1 - \theta)\mathbf{G}^{(k-1)}\mathbf{y}^{(k-1)} \to \mathbf{s}^{(k-1)}, \tag{12}$$

where

$$\theta = \begin{cases} 1, & \text{if } a \ge 0.2b \\ 0.8b/(b - a) & \text{otherwise} \end{cases},$$

$a = \mathbf{s}^{(k-1)T}\mathbf{y}^{(k-1)}$, $b = \mathbf{y}^{(k-1)T}\mathbf{G}^{(k-1)}\mathbf{y}^{(k-1)}$. Ni and Yuan (1997) have proved that transformation of equation 12 guarantees that matrix $\mathbf{G}^{(k)}$ is positive definite.

Additionally, the gradient projection method is used to determine a set of active constraints. Alternative ways to deal with the bound constrained QN optimization can be found in Byrd et al. (1995) and Kelley (1999).

The QN method described above does not depend on the dimensionality of the inverse problem. It can be applied equally to 1D, 2D, or 3D cases. For the 1D case, m_i is the electrical conductivity of the ith layer. For 2D and 3D cases, it is the electrical conductivity of the ith cell.

Choice of parameter λ

We know that the choice of parameter λ of equation 1 is crucial for finding a reliable solution to underdetermined $(M \ll N)$ inverse

problems (see Farquharson and Oldenburg, 1998). Too large a value of λ may lead to a model that is too smooth, whereas a value of λ that is too small may deliver a nonphysical model with unwanted artifacts. In this study, we applied a simple cooling-type approach that comprises two stages. First, we find two values λ_{low} and λ_{high} such that $\varphi_d(\mathbf{m}^{\lambda_{\text{low}}}) < 1$, and $\varphi_d(\mathbf{m}^{\lambda_{\text{high}}}) > 1$, where $\mathbf{m}^{\lambda}$ delivers the solution of $\min_{\mathbf{m}} \varphi(\mathbf{m}, \lambda)$ with the parameter λ. To do this we solve the inverse problem posed in equations 1–4 for different values of λ. As prescribed by the cooling approach, we start this process with a large value of λ. Note that weight coefficients $\beta_j (j = 1,\ldots,M)$ of the misfit φ_d are chosen so that the value $\varphi_d = 1$ exactly corresponds to the noise floor. In the second stage, we find the biggest possible λ_{opt} such that $\varphi_d(\mathbf{m}^{\lambda_{\text{opt}}}) \leq 1$, where $\mathbf{m}^{\lambda_{\text{opt}}}$ is a solution of $\min_{\mathbf{m}} \varphi(\mathbf{m}, \lambda_{\text{opt}})$. To find such a value λ, we iteratively apply the linear interpolation of $\phi(\lambda) = \varphi_d(\mathbf{m}^{\lambda})$ at the segment $[\lambda_{\text{low}}, \lambda_{\text{high}}]$. This approach is a simple variant of an algorithm by Haber and Oldenburg (1997).

In numerical examples we present later, we demonstrate that such a procedure works successfully.

Speeding up the solution

As presented above, at each iteration step k the inverse problem solution requires calculating the gradient $\mathbf{g}^{(k)} = [(\partial\varphi/\partial m_1), \ldots, (\partial\varphi/\partial m_N)]^T$. Straightforward numerical calculation of the gradient $\mathbf{g}^{(k)}$ involves $(N + 1)$ solutions of the forward problem: $\mathbf{m} \rightarrow \varphi(\mathbf{m})$. However, for large-scale problems (when N is large) such a straightforward calculation may be prohibitive in terms of computation time. One can significantly speed up the inverse problem solution by avoiding this straightforward calculation. In Appendix A we present such an approach for the 1D MT case, which exploits the Green's function technique. Numerically, calculation of the gradient with the use of this approach requires solving a single forward and adjoint problem, rather than $N + 1$ forward problems. Our approach resembles an implementation of the well-known adjoint method (see, for example, Rodi and Mackie, 2001; Chen et al., 2005) to the 1D MT inverse problem. It is important, however, to stress that our approach may be extended with some effort to the 3D case. It is also noteworthy that, for the 1D MT case, the gradient can also be calculated using the chain-rule (Constable et al., 1987).

Therefore, we have implemented the limited-memory QN method with simple bounds (hereinafter, referred to as LMQNB), which is described above. It should be noted here that our implementation differs from that of Ni and Yuan (1997) in that the LMQNB uses the Wolfe conditions of equations 11 to terminate the line search.

MODEL EXAMPLES

Let us demonstrate on a synthetic 1D MT example the extent to which the calculation of the gradient given by equation 6 presented in Appendix A accelerates the solution of the 1D MT inverse problem. A seven-layered earth model is given in Table 1. This model was compiled from the models (see Avdeev et al., 2004) derived from a seafloor MT and a global GDS long-period data set collected in the North Pacific Ocean. To complicate the inversion process, we subdivided the three upper layers in this model (to a depth of 394 km) into 197 sublayers of equal thickness. For this 201-layered ($N = 201$) model, we inverted the impedance $d_j = Z_j(\mathbf{m})$, calculated at $M = 30$ periods, from 10 s to 10,800 s. In addition, we added 0.5% random noise to the impedance data. The relative error ε_j $\left[\beta_j = \frac{1}{M}(2/\varepsilon_j^2|d_j|^2)\right]$ of the impedance was taken as 0.01. A 10-ohm-m uniform half-space was used as an initial guess $\mathbf{m} = \mathbf{m}^0$. In Figure 1a, we compare the convergence rates of two solutions obtained with straightforward calculation of the gradients and using the method presented in Appendix A, respectively. The curves are shown as a function of the number of evaluations of $\varphi(\mathbf{m}, \lambda)$ for two cases, $\lambda = 0$ and $\lambda = \lambda_{\text{opt}}$. From Figure 1a, it is seen that the calculation of the gradients given in Appendix A accelerates the solution by 130 times. The results of QN inversion are presented in Figure 1b. As might be

Table 1. A seven-layered earth model.

Conductivity (S/m)	Thickness (km)
0.01	64
0.05	180
0.1	150
0.12	126
0.28	130
1.1	150
1.5	

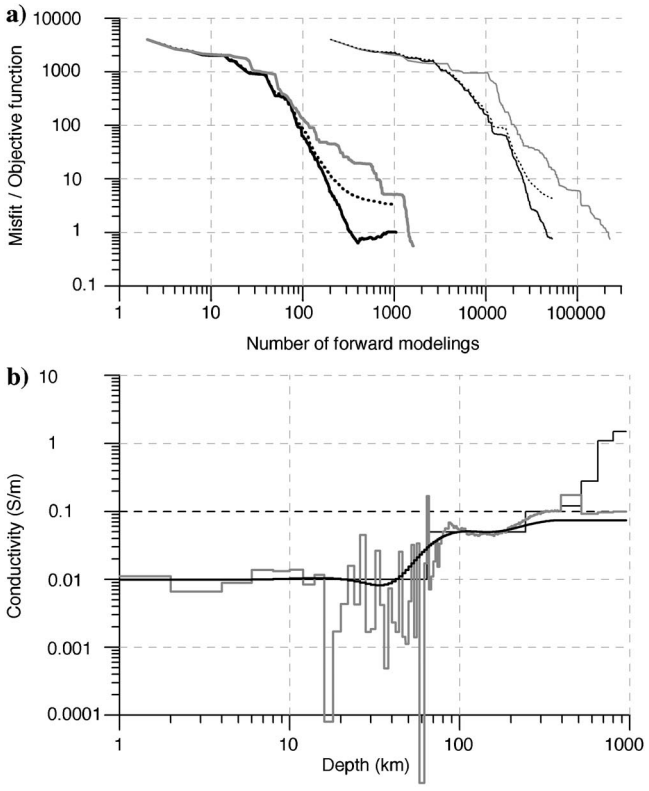

Figure 1. (a) Comparison of convergence rates: The curves present the misfit of equation 2 obtained using QN inversion with the straightforward calculation of the gradient of equation 6, shown for $\lambda = 0$ (thin grey line) and $\lambda_{opt} = 320$ (thin black line). The thin dotted line presents the objective function of equation 1 for $\lambda_{opt} = 320$. The same is shown with the calculation of the gradient that is presented in Appendix A, again for $\lambda = 0$ (thick grey line) and $\lambda_{opt} = 320$ (thick solid and dotted lines). (b) The conductivity models obtained from the QN inversion with the calculation of the gradients from Appendix A, again for $\lambda = 0$ (grey line) and $\lambda_{opt} = 320$ (black line). The thin solid and dashed lines show the true and initial guess models, respectively.

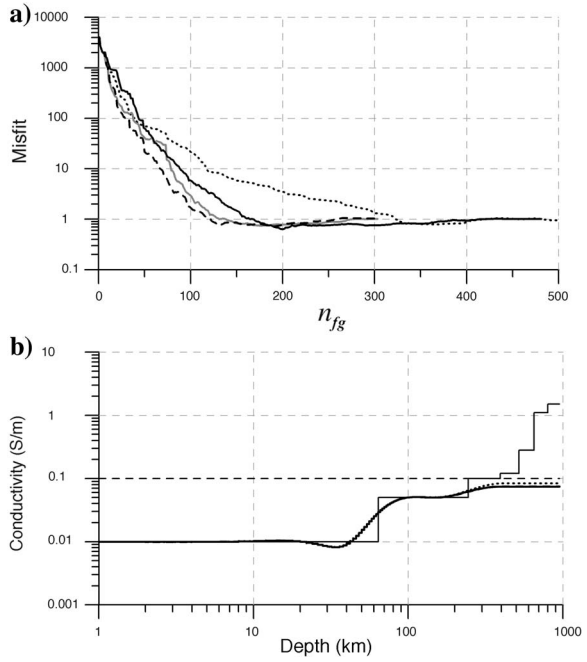

Figure 2. Comparison of the LMQNB inversions with different numbers of correction pairs (n_{cp}). (a) The convergence rate for n_{cp} = 2 (dotted line), n_{cp} = 5 (solid black line), n_{cp} = 20 (dashed line), and n_{cp} = 25 (grey line). (b) The corresponding conductivity models, which are plotted mostly on top of each other. The thin solid and dashed lines show the true and initial guess models, respectively.

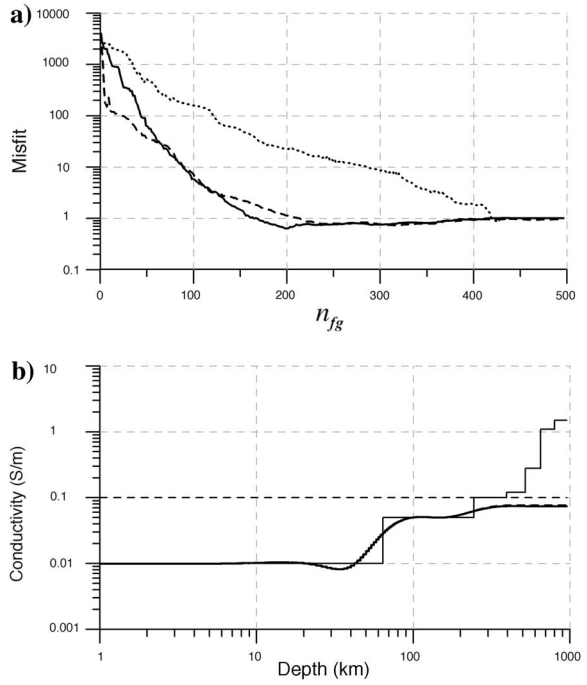

Figure 3. Comparison of three different QN inversions. (a) The convergence rate for LMQNB (solid line), L-BFGS-B (dashed line), and Gill et al.'s (1981) QN algorithms (dotted line). (b) The corresponding conductivity models, which are plotted on top of each other. The thin solid and dashed lines show the true and initial guess models, respectively.

expected, the conductivity models recovered do not fit well for depths greater than 400 km because the responses at the periods considered are almost insensitive to these depths.

Let us again consider the 201-layered model and the data set, described in the previous example. In our next example (Figure 2), we study the convergence rate of the LMQNB solution for a various number n_{cp} of correction pairs. The curves in Figure 2 are shown as a function of the number n_{fg} of evaluations of $\varphi(\mathbf{m},\lambda)$ and $\mathbf{g} = (\partial\varphi/\partial m)(\mathbf{m},\lambda)$ for $\lambda = \lambda_{\text{opt}}$. It is surprising that so small a number of the pairs ($n_{cp} = 5$) can be sufficient to get a relatively reasonable result.

In Figure 3, we present the third example — a comparison of two solutions of the 1D MT inverse problem. The first solution is based on the optimization method offered in this paper, and the second one uses the L-BFGS-B optimization code by Byrd et al. (1995). The comparison is presented for $n_{cp} = 5$ correction pairs and $\lambda_{\text{opt}} = 320$. These solutions converge in a similar way and produce similar models. In the same figure, we show the curves produced using the conventional QN algorithm of Gill et al. (1981), subject to simple bounds. To find the search direction $\mathbf{p}^{(k)}$, this algorithm solves the Newton system $\mathbf{H}^{(k)}\mathbf{p}^{(k)} = -\mathbf{g}^{(k)}$, rather than making use of equation 5, and it is valid for small-scale problems.

CONCLUSION

We have described a limited-memory QN method applied to solve the 1D MT inverse problem. We believe that the method may be applied equally for large-scale 1D, 2D, or 3D MT cases. In the numerical examples presented, we have demonstrated that the calculation of gradients with the use of the adjoint method dramatically accelerates 1D MT inversion. The nontrivial problem of such a calculation of gradients in the 3D MT case is the subject of ongoing research. Another finding of our numerical experiments is that the LMQNB solution converges similarly to the solution based on the L-BFGS-B method.

Further work will be concentrated on extension of this LMQNB method to the 3D MT case. The implementation of the method to the 3D case, if successful, may be a useful inversion tool in understanding a wide variety of electromagnetic earth phenomena.

ACKNOWLEDGMENTS

This work was carried out in the Dublin Institute for Advanced Studies and was supported by the Cosmogrid project, which is funded by the Programme for Research in Third Level Institutions under the National Development Plan and with assistance from the European Regional Development Fund. We acknowledge associate editor Randy Mackie, two anonymous referees, and Weerachai Siripunvaraporn for many valuable suggestions that led to a much better paper. Thanks to Brian O'Reilly for corrections to English. A. Avdeeva acknowledges the constructive support of her supervisors Alan Jones and Colin Brown.

APPENDIX A

CALCULATION OF GRADIENT

To derive derivatives $\partial\varphi_d/\partial\sigma_k$, we apply a popular adjoint method (cf., Rodi and Mackie, 2001). To do this, let us first rewrite equation 2 as

$$\varphi_d(\boldsymbol{\sigma}) = \tfrac{1}{2}\langle \mathbf{d} - \mathbf{Z}, \mathbf{d} - \mathbf{Z}\rangle_2, \tag{A-1}$$

where we introduce the vectors $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_N)^T$, $\mathbf{d} = (d_1, ..., d_M)^T$, and $\mathbf{Z} = (Z_1, ..., Z_M)^T$, and an inner product $\langle \cdot, \cdot \rangle_2$ of vectors $\mathbf{u}$ and $\mathbf{v}$ as

$$\langle \mathbf{u}, \mathbf{v}\rangle_2 = \sum_{j=1}^{M} \beta_j \overline{u}_j v_j. \tag{A-2}$$

From equations A-1 and A-2, it follows that

$$\frac{\partial\, \varphi_d}{\partial\, \sigma_k} = \mathrm{Re}\left\langle \mathbf{Z} - \mathbf{d}, \frac{\partial\, \mathbf{Z}}{\partial\, \sigma_k} \right\rangle_2. \tag{A-3}$$

In equation A-2 the over bar indicates complex conjugate, and in equation A-3 Re is the real part of a value.

The main problem now is how to calculate derivatives $\partial Z_j/\partial\sigma_k$. From Maxwell's equations written for a layered earth, it follows that at any depth z impedance can be derived as

$$Z_j(z) = \frac{i\omega_j\mu_0 v_j}{\partial_z v_j}, \tag{A-4}$$

where the potential $v_j(z)$ satisfies the Helmholtz equation

$$\partial_z^2 v_j + i\omega_j\mu_0\sigma(z)v_j = 0, \tag{A-5}$$

and the boundary condition $v_j \underset{z\to\infty}{\longrightarrow} 0$. Here, z is the depth, $\omega_j = 2\pi/T_j$ is the jth angular frequency, and μ_0 is the magnetic permeability of free space. Further, for a layered earth we easily may decompose $\sigma(z)$ as

$$\sigma(z) = \sum_{k=1}^{N} \sigma_k \chi_k(z), \tag{A-6}$$

where σ_k is the conductivity of the kth layer confined between boundaries z_k and z_{k+1}, and $\chi_k(z) = \theta(z - z_k) - \theta(z - z_{k+1})$, where $\theta(z)$ is the Heaviside function. Note that in equation A-6 we assume that $z_{N+1} = \infty$. From equations A-5 and A-6 it follows that

$$\partial_z^2 \frac{\partial\, v_j}{\partial\, \sigma_k} + i\omega_j\mu_0\sigma(z)\frac{\partial\, v_j}{\partial\, \sigma_k} = -\, i\omega_j\mu_0\chi_k(z)v_j. \tag{A-7}$$

From equation A-7, it follows that the derivative $\partial v_j/\partial\sigma_k$ is expressed as

$$\begin{aligned}\frac{\partial\, v_j}{\partial\, \sigma_k}(z) &= -\int G_j(z,\zeta)\chi_k(\zeta)v_j(\zeta)d\zeta \\ &= -\int_{z_k}^{z_{k+1}} G_j(z,\zeta)v_j(\zeta)d\zeta,\end{aligned} \tag{A-8}$$

where the Green's function $G_j(z,\zeta)$ is a solution to the following equation:

$$\partial_z^2 G_j(z,\zeta) + i\omega_j\mu_0\sigma(z)G_j(z,\zeta) = i\omega_j\mu_0\delta(z - \zeta), \tag{A-9}$$

and where $\delta(z)$ is the delta-function. From equation A-4 it follows that

$$\frac{\partial\, Z_j}{\partial\, \sigma_k}(z) = Z_j\left(\frac{1}{v_j}\frac{\partial\, v_j}{\partial\, \sigma_k} - \frac{1}{\partial_z v_j}\partial_z\frac{\partial\, v_j}{\partial\, \sigma_k}\right). \tag{A-10}$$

Substituting equation A-8 into equation A-10 (and assuming $z = 0$) yields

$$\begin{aligned}\frac{\partial\, Z_j}{\partial\, \sigma_k}(0) = Z_j(0)\Bigg(&-\frac{1}{v_j}\int_{z_k}^{z_{k+1}} G_j(0,\zeta)v_j(\zeta)d\zeta \\ &+ \frac{1}{\partial_z v_j}\int_{z_k}^{z_{k+1}} \partial_z G_j(z,\zeta)|_{z=0} v_j(\zeta)d\zeta\Bigg).\end{aligned} \tag{A-11}$$

Using the following properties of $G_j(z,\zeta)$ (at $z < \zeta$) (cf. Avdeev et al., 1997):

$$\partial_z G_j(z,\zeta) = -\, i\omega_j\mu_0\frac{1}{Z_j^*(z)}G_j(z,\zeta), \tag{A-12}$$

$$G_j(z,\zeta) = \frac{1}{Z_j^{-1}(z) - Z_j^{*-1}(z)}\gamma_j(z,\zeta), \tag{A-13}$$

$$\gamma_j(z,\zeta) = \exp\left(i\omega_j\mu_0\int_z^\zeta Z_j^{-1}(\zeta)d\zeta\right), \tag{A-14}$$

from equation A-11 we derive that

$$\frac{\partial\, Z_j}{\partial\, \sigma_k}(0) = -\, Z_j^2(0)\cdot\int_{z_k}^{z_{k+1}} \gamma_j^2(0,\zeta)d\zeta. \tag{A-15}$$

In equations A-12 and A-13, $Z_j^*(z)$ is the impedance of the layered earth model above the depth z. Substituting equation A-15 in equation A-3 yields the sought-for derivatives

$$\frac{\partial\, \varphi_d}{\partial\, \sigma_k} = -\,\mathrm{Re}\left(\sum_{j=1}^{M}\beta_j\Gamma_{kj}(\overline{Z}_j - \overline{d}_j)Z_j^2\right), \tag{A-16}$$

where for simplicity we denoted $Z_j = Z_j(0)$, and

$$\Gamma_{kj} = \int_{z_k}^{z_{k+1}} \gamma_j^2(0,\zeta)d\zeta \quad (k = 1, ..., N). \tag{A-17}$$

An explicit expression for the coefficients Γ_{kj} is given in Appendix B.

APPENDIX B

CALCULATION OF Γ_{KJ}

By definition (see equation A-17)

$$\Gamma_{kj} = \int_{z_k}^{z_{k+1}} \gamma_j^2(0,\zeta)d\zeta \quad (k = 1, ..., N), \tag{B-1}$$

where

$$\gamma_j(z,\zeta) = \exp\left(i\omega_j\mu_0\int_z^\zeta Z_j^{-1}(\zeta)d\zeta\right). \tag{B-2}$$

From equations B-1 and B-2 it follows that

$$\Gamma_{kj} = \prod_{l=1}^{k-1} \gamma_j^2(z_l, z_{l+1}) \cdot \gamma_{kj}, \tag{B-3}$$

where

$$\gamma_{kj} = \int_{z_k}^{z_{k+1}} \gamma_j^2(z_k, \zeta) d\zeta. \tag{B-4}$$

In the above expressions, we assumed that $z_1 = 0$ and $z_{N+1} = \infty$. From equation B-3 it follows that the coefficients Γ_{kj} can be calculated recursively as

$$\Gamma_{k+1j} = \gamma_j^2(z_k, z_{k+1}) \frac{\gamma_{k+1j}}{\gamma_{kj}} \Gamma_{kj}, \tag{B-5}$$

where $\Gamma_{1j} = \gamma_{1j}$ as can be seen from equations B-1 and B-4.

From equations A-4 and A-5 and the definition given by equation B-2, it follows that

$$\gamma_j(z_k, \zeta) = ch(\kappa_{kj}(\zeta - z_k)) + \lambda_{kj} sh(\kappa_{kj}(\zeta - z_k)) \quad (z_k \le \zeta \le z_{k+1}), \tag{B-6}$$

where $\kappa_{kj} = \sqrt{-i\omega_j \mu_0 \sigma_k}$ and

$$\lambda_{kj} = \frac{i\omega_j \mu_0}{\kappa_{kj}} Z_j^{-1}(z_k). \tag{B-7}$$

From equations A-4 and A-5 and the definition given by equation B-7, the following recursive formula follows

$$\lambda_{kj} = \begin{cases} \dfrac{-sh(\Delta_k \kappa_{kj}) + ch(\Delta_k \kappa_{kj}) \dfrac{\kappa_{k+1j}}{\kappa_{kj}} \lambda_{k+1j}}{ch(\Delta_k \kappa_{kj}) - sh(\Delta_k \kappa_{kj}) \dfrac{\kappa_{k+1j}}{\kappa_{kj}} \lambda_{k+1j}}, & k = N-1, \ldots, 1 \\ -1, & k = N \end{cases}. \tag{B-8}$$

From equation B-6 it also follows that

$$\gamma_j(z_k, z_{k+1}) = ch(\Delta_k \kappa_{kj}) + \lambda_{kj} sh(\Delta_k \kappa_{kj}). \tag{B-9}$$

Finally, substituting equation B-6 into equation B-4 and after some manipulations one obtains

$$\gamma_{kj} = \begin{cases} \Delta_k \left(\dfrac{1 - \lambda_{kj}^2}{2} + \dfrac{sh(\Delta_k \kappa_{kj})}{\Delta_k \kappa_{kj}} \left[\dfrac{(1 + \lambda_{kj})^2}{2} ch(\Delta_k \kappa_{kj}) + \lambda_{kj} sh(\Delta_k \kappa_{kj}) \right] \right), & k = N-1, \ldots, 1 \\ \dfrac{1}{2\kappa_{kj}}, & k = N \end{cases}. \tag{B-10}$$

Expressions presented in equations B-9 and B-10 permit the recursive calculation of the coefficients Γ_{kj} using equation B-5.

Comparing equations B-8–B-10, one can see that the coefficients λ_{kj}, γ_{kj}, and $\gamma_j(z_k, z_{k+1})$ are calculated simultaneously while moving from the bottom of the model to the surface. Such a calculation is somewhat equivalent to the forward problem solution. Once the coefficients γ_{kj} and $\gamma_j(z_k, z_{k+1})$ are found, the Γ_{kj} required by the derivatives defined in equation A-16 are calculated recursively from the surface to the bottom using equation B-5. Thus, as is predicted by the theory of the adjoint method, the time to calculate the derivatives of equation A-16 is not more than twice that required for the forward problem solution (calculating λ_{kj} alone).

REFERENCES

Avdeev, D. B., A. V. Kuvshinov, O. V. Pankratov, and G. A. Newman, 1997, High-performance three-dimensional electromagnetic modeling using modified Neumann series. Wide-band numerical solution and examples: Journal of Geomagnetism and Geoelectricity, **49**, 1519–1539.

Avdeev, D. B., H. Utada, A. V. Kuvshinov, and T. Koyama, 2004, Three-dimensional electromagnetic modeling of the Hawaiian swell: American Geophysical Union Fall Meeting, F630.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, A limited-memory algorithm for bound constrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208.

Chen, J., D. W. Oldenburg, and E. Haber, 2005, Reciprocity in electromagnetics: Application to modeling marine magnetometric resistivity data: Physics of the Earth and Planetary Interiors, **150**, 45–61.

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

Farquharson, C. G., and D. W. Oldenburg, 1998, Non-linear inversion using general measures of data misfit and model structure: Geophysical Journal International, **134**, 213–227.

Gill, P. E., W. Murray, and M. H. Wright, 1981, Practical optimization: Academic Press Inc.

Haber, E., 2005, Quasi-Newton methods for large-scale electromagnetic inverse problems: Inverse Problems, **21**, 305–323.

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

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

Newman, G. A., and P. T. Boggs, 2004, Solution accelerators for large-scale three-dimensional electromagnetic inverse problem: Inverse Problems, **20**, s151–s170.

Ni, Q., and Y. Yuan, 1997, A subspace limited-memory quasi-Newton algorithm for large-scale nonlinear bound constrained optimization: Mathematics of Computation, **66/220**, 1509–1520.

Nocedal, J., and S. J. Wright, 1999, Numerical optimization: Springer Publishing Company, Inc.

Rodi, W., and R. L. Mackie, 2001, Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion: Geophysics, **66**, 174–187.

Tikhonov, A. N., 1963, Regularization of incorrectly posed problems: Soviet Mathematics Doklady, **4**, 1624–1627.