5th International Symposium on Three-Dimensional Electromagnetics
Sapporo, Japan, May 7–9, 2013

Preparation for a 3D magnetotelluric inversion – specific characteristics of the all-at-once approach

Wenke Wilhelms, Ralph-Uwe Börner, and Klaus Spitzer
TU Bergakademie Freiberg

SUMMARY

The all-at-once inversion approach requires no explicit forward calculation, because the forward modelling equations are incorporated in the objective function as constraints. This leads to a huge, so-called Karush-Kuhn-Tucker (KKT) system, which is solved in each step of the iteration procedure to update model parameters, Lagrangian multipliers, and data - all at once. Still, the forward problem needs detailed consideration, because the system matrix of the forward problem is necessary to calculate first and second derivatives for setting up the KKT matrix. Our all-at-once inversion results for the 1D magnetotelluric problem are promising. As necessary prerequisite for a 3D inversion we successfully implemented and validated a 3D forward operator. Forming the Hessian, i.e., the matrix containing second derivatives required in a Newton step, is key in the all-at-once approach. Besides regularization, we emphasize the chosen method of calculating the residual. Since the 3D magnetotelluric forward response is represented by two linear independent solutions of equation systems, the data has to be formed as a linear combination of electric field values and their spatial derivatives, the magnetic fields. We reformulate the problem by scaling the measured impedances with the magnetic fields obtained from the forward step, thus eliminating the non-trivial step of forming impedances from numerical results alone. The resulting KKT system can be solved using Krylov subspace projection techniques.

Keywords: magnetotellurics, 3D inversion, finite differences, all-at-once

INTRODUCTION

As a first step in the process of developing a 3D all-at-once inversion code, we have investigated the 1D problem (Figure 1). The resulting models are comparable with those of a traditional Gauß-Newton inversion. However, as expected, the Gauss-Newton method converges more slowly than the all-at-once approach, because the latter, as a Newton type scheme, takes into account second derivative information. As a second step, a 3D finite difference forward operator could already be tested successfully using the 3D-2 COMMEMI model.

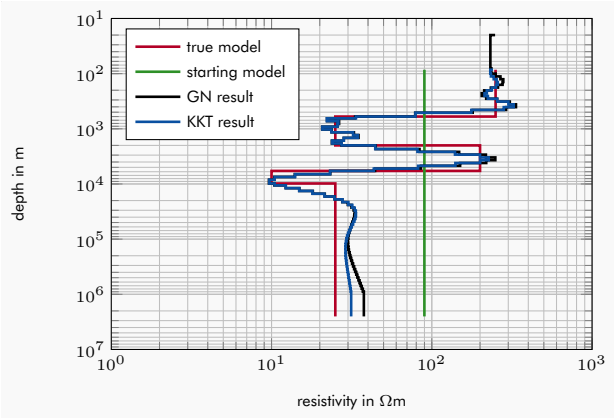

Figure 1. Inversion results for 1D magnetotellurics.

In the following, we describe the current state of our investigations. In the first section, we therefore explain the governing equations of the forward problem. The second part deals with the formulation of the inverse problem using the all-at-once approach. There we show how to build up the KKT matrix, essentially, by incorporating the system matrix of the forward problem and its derivatives.

FORWARD PROBLEM

For magnetotellurics the behaviour of the electric field $\mathbf{E}$ in a 3D earth is described by the homogeneous curl-curl equation

$$\nabla \times \nabla \times \mathbf{E} + i\omega\mu_0\sigma\mathbf{E} = 0 \tag{1}$$

with ω as the angular frequency and $\omega = 2\pi f$ where f is the frequency, i as the imaginary unit, σ as the electrical conductivity and μ_0 as the induction constant. Using the secondary field approach this equation changes to

$$\nabla \times \nabla \times \mathbf{E}_s + i\omega\mu_0\sigma\mathbf{E}_s = -i\omega\mu_0(\sigma - \sigma_0)\mathbf{E}_p, \tag{2}$$

where σ_0 is a background conductivity model, usually a layered or a homogeneous halfspace having an analytical solution. The total electric field $\mathbf{E}$ is the sum of both, primary and secondary electric fields $\mathbf{E}_p$ and $\mathbf{E}_s$. The primary electric field $\mathbf{E}_p$ is a function of depth z and the secondary electric field $\mathbf{E}_s$ vanishes on the boundary Γ of the model. The right-hand side of eq. 2 represents the sources of the secondary field. Together with appropriate Dirichlet

boundary conditions the curl-curl equation is discretized on a finite-difference grid providing the forward operator $\mathbf{A}(\boldsymbol{m})$. For 3D magnetotelluric measurements two polarisation directions are needed. Then the forward problem can be written as

$$\mathbf{A}(\boldsymbol{m})[\boldsymbol{u}_1\boldsymbol{u}_2] = [\boldsymbol{b}_1(\boldsymbol{m})\boldsymbol{b}_2(\boldsymbol{m})] \tag{3}$$

with $\boldsymbol{m}$ as a vector containing logarithmic model parameters $m_j = log\sigma_j, j = 1, ..., M$, and M as the total number of model cells. The two right-hand side vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ contain secondary source terms and boundary values. The vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ denote the complex secondary electric field $\mathbf{E}_s^{1,2} = [\mathbf{E}_s^x, \mathbf{E}_s^y, \mathbf{E}_s^z]^\top$ for either polarisation 1 or 2 within the whole discretized region.

The 3D forward operator was validated using published data of the 3D-2 COMMEMI model (Zhdanov & Weaver, 1997). In Figure 2 and 3 our results agree very well with those of Xiong (1992).

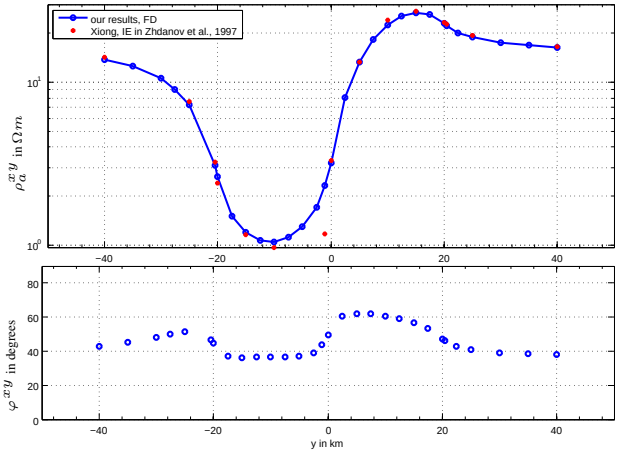

Figure 2. Forward modelling results for 3D magnetotellurics at a frequency of $f = 0.01\,\mathrm{Hz}$: xy-polarisation.

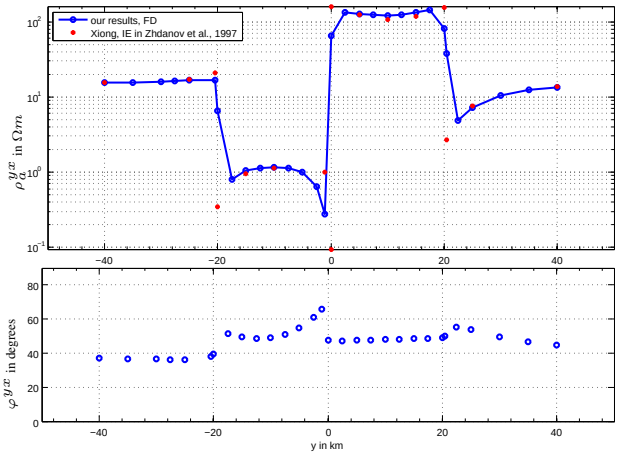

Figure 3. Forward modelling results for 3D magnetotellurics at a frequency of $f = 0.01\,\mathrm{Hz}$: yx-polarisation.

INVERSE PROBLEM

The all-at-once approach is based on Newton's method. As described by Haber, Ascher, and Oldenburg (2000) for the 1D magnetotelluric problem the Lagrangian $\mathcal{L}$ is formed by setting up the optimization problem in a constrained form. In doing so, two forward problems (one for each polarisation direction) is incorporated using Lagrange multipliers $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ in the objective function:

$$\begin{aligned}\mathcal{L}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{m}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) &= f(\boldsymbol{u}_1, \boldsymbol{u}_2) \\ &+ \tfrac{\beta}{2}||\mathbf{W}(\boldsymbol{m} - \boldsymbol{m}_{\text{ref}})||^2 \\ &+ \boldsymbol{\lambda}_1^T[\mathbf{A}(\boldsymbol{m})\boldsymbol{u}_1 - \boldsymbol{b}_1(\boldsymbol{m})] \\ &+ \boldsymbol{\lambda}_2^T[\mathbf{A}(\boldsymbol{m})\boldsymbol{u}_2 - \boldsymbol{b}_2(\boldsymbol{m})],\end{aligned} \tag{4}$$

where the first term on the right-hand side is the data residual. This is challenging to calculate because simulated electric and magnetic field values do not correspond with measured data consisting of impedances. We overcome this problem in manipulating measured impedances with simulated magnetic fields to compute data residuals in terms of electric fields:

$$\begin{aligned}f(\boldsymbol{u}_1, \boldsymbol{u}_2) &= \tfrac{1}{2}\left\|\begin{pmatrix}E_x^1 & E_x^2\\ E_y^1 & E_y^2\end{pmatrix} - \mathbf{Z}\begin{pmatrix}H_x^1 & H_x^2\\ H_y^1 & H_y^2\end{pmatrix}\right\|^2 \\ &= \tfrac{1}{2}\left\|\left[\begin{pmatrix}Q_{Ex}\\ Q_{Ey}\end{pmatrix} - \mathbf{Z}\begin{pmatrix}Q_{Hx}\\ Q_{Hy}\end{pmatrix}\right]\begin{pmatrix}\boldsymbol{u}_1\\ \boldsymbol{u}_2\end{pmatrix}^\top\right\|^2 \\ &= \tfrac{1}{2}\left\|(\mathbf{Q_E} - \mathbf{Z}\mathbf{Q_H})\begin{pmatrix}\boldsymbol{u}_1\\ \boldsymbol{u}_2\end{pmatrix}^\top\right\|^2\end{aligned} \tag{5}$$

with the impedance matrix $\mathbf{Z} = \begin{pmatrix}Z_{xx} & Z_{xy}\\ Z_{yx} & Z_{yy}\end{pmatrix}$ consisting of measured data.

Matrices $\mathbf{Q_E}$ and $\mathbf{Q_H}$ are observation operators that extract the corresponding data (electric and magnetic fields at the measurement locations) from the model response $\begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2\end{pmatrix}$.

The second term of eq. (4) is a regularisation term. For minimizing the Lagrangian via Newton's method, the gradient $\nabla\mathcal{L}$ and the Hessian $\nabla^2\mathcal{L}$ need to be calculated. The gradient is a vector containing the partial derivatives of $\mathcal{L}$ with respect to all five parameters $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{m}$, $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$:

$$\nabla\mathcal{L} = \begin{pmatrix}\mathcal{L}_{\boldsymbol{u}_1}\\ \mathcal{L}_{\boldsymbol{u}_2}\\ \mathcal{L}_{\boldsymbol{m}}\\ \mathcal{L}_{\boldsymbol{\lambda}_1}\\ \mathcal{L}_{\boldsymbol{\lambda}_2}\end{pmatrix} \tag{6}$$

$$\mathcal{L}_{\boldsymbol{u}_1} = [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\boldsymbol{u}_1\\\boldsymbol{u}_2\end{pmatrix}^\top + \mathbf{A}^\top\boldsymbol{\lambda}_1 \quad (7)$$

$$\mathcal{L}_{\boldsymbol{u}_2} = [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{0}\\\mathbf{1}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\boldsymbol{u}_1\\\boldsymbol{u}_2\end{pmatrix}^\top + \mathbf{A}^\top\boldsymbol{\lambda}_2 \quad (8)$$

$$\mathcal{L}_{\boldsymbol{m}} = \beta\mathbf{W}^\top\mathbf{W}(\boldsymbol{m} - \boldsymbol{m}_{\text{ref}}) + (\mathbf{G}_1 - \mathbf{B}_1)^\top\boldsymbol{\lambda}_1 + (\mathbf{G}_2 - \mathbf{B}_2)^\top\boldsymbol{\lambda}_2 \quad (9)$$

$$\mathcal{L}_{\boldsymbol{\lambda}_1} = \mathbf{A}\boldsymbol{u}_1 - \boldsymbol{b}_1 \quad (10)$$

$$\mathcal{L}_{\boldsymbol{\lambda}_2} = \mathbf{A}\boldsymbol{u}_2 - \boldsymbol{b}_2 \quad (11)$$

with the following matrices containing first derivatives $\mathbf{G}_1 = \partial_m\mathbf{A}\cdot\boldsymbol{u}_1$, $\mathbf{G}_2 = \partial_m\mathbf{A}\cdot\boldsymbol{u}_2$, $\mathbf{B}_1 = \partial_m\boldsymbol{b}_1$, and $\mathbf{B}_2 = \partial_m\boldsymbol{b}_2$.

The Hessian $\nabla^2\mathcal{L}$ (or KKT matrix $\mathcal{H}_{kkt}$) contains the second derivatives of the Lagrange function with respect to all five parameters:

$$\mathcal{H}_{kkt} = \begin{pmatrix}
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{u}_1} & \mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{u}_2} & \mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{m}} & \mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{\lambda}_1} & \mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{\lambda}_2}\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{u}_1} & \mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{u}_2} & \mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{m}} & \mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{\lambda}_1} & \mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{\lambda}_2}\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{u}_1} & \mathcal{L}_{\boldsymbol{m},\boldsymbol{u}_2} & \mathcal{L}_{\boldsymbol{m},\boldsymbol{m}} & \mathcal{L}_{\boldsymbol{m},\boldsymbol{\lambda}_1} & \mathcal{L}_{\boldsymbol{m},\boldsymbol{\lambda}_2}\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{u}_1} & \mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{u}_2} & \mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{m}} & \mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_1} & \mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2}\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{u}_1} & \mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{u}_2} & \mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{m}} & \mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_1} & \mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_2}
\end{pmatrix}$$

where

$$\begin{aligned}
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{u}_1} &= [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}^\top\\
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{u}_2} &= [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{0}\\\mathbf{1}\end{pmatrix}^\top\\
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{m}} &= \mathbf{K}_1\\
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{\lambda}_1} &= \mathbf{A}^\top\\
\mathcal{L}_{\boldsymbol{u}_1,\boldsymbol{\lambda}_2} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{u}_1} &= [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{0}\\\mathbf{1}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{1}\\\mathbf{0}\end{pmatrix}^\top\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{u}_2} &= [(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{0}\\\mathbf{1}\end{pmatrix}^\top]^\top(\mathbf{Q_E} - \mathbf{ZQ_H})\begin{pmatrix}\mathbf{0}\\\mathbf{1}\end{pmatrix}^\top\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{m}} &= \mathbf{K}_2\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{\lambda}_1} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{u}_2,\boldsymbol{\lambda}_2} &= \mathbf{A}^\top\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{u}_1} &= \mathbf{K}_1^\top\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{u}_2} &= \mathbf{K}_2^\top\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{m}} &= \beta\mathbf{W}^\top\mathbf{W} + \mathbf{R}_1 - \mathbf{D}_1 + \mathbf{R}_2 - \mathbf{D}_2\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{\lambda}_1} &= (\mathbf{G}_1 - \mathbf{B}_1)^\top\\
\mathcal{L}_{\boldsymbol{m},\boldsymbol{\lambda}_2} &= (\mathbf{G}_2 - \mathbf{B}_2)^\top\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{u}_1} &= \mathbf{A}\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{u}_2} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{m}} &= \mathbf{G}_1 - \mathbf{B}_1\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_1} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{u}_1} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{u}_2} &= \mathbf{A}\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{m}} &= \mathbf{G}_2 - \mathbf{B}_2\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_1} &= \mathbf{0}\\
\mathcal{L}_{\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_2} &= \mathbf{0}
\end{aligned}$$

with the following matrices containing first and second derivatives with respect to model parameters $\boldsymbol{m}$:
$\mathbf{K}_1 = \partial_m\mathbf{A}^\top\cdot\boldsymbol{\lambda}_1$, and $\mathbf{K}_2 = \partial_m\mathbf{A}^\top\cdot\boldsymbol{\lambda}_2$,
$\mathbf{R}_1 = \partial_m\mathbf{G}_1^\top\cdot\boldsymbol{\lambda}_1$, and $\mathbf{R}_2 = \partial_m\mathbf{G}_2^\top\cdot\boldsymbol{\lambda}_2$,
$\mathbf{D}_1 = \partial_m\mathbf{B}_1^\top\cdot\boldsymbol{\lambda}_1$, and $\mathbf{D}_2 = \partial_m\mathbf{B}_2^\top\cdot\boldsymbol{\lambda}_2$.

The Newton method then reads

$$\mathcal{H}_{kkt}\cdot\begin{pmatrix}\boldsymbol{\delta u}_1\\\boldsymbol{\delta u}_2\\\boldsymbol{\delta m}\\\boldsymbol{\delta\lambda}_1\\\boldsymbol{\delta\lambda}_2\end{pmatrix} = -\begin{pmatrix}\mathcal{L}_{\boldsymbol{u}_1}\\\mathcal{L}_{\boldsymbol{u}_2}\\\mathcal{L}_{\boldsymbol{m}}\\\mathcal{L}_{\boldsymbol{\lambda}_1}\\\mathcal{L}_{\boldsymbol{\lambda}_2}\end{pmatrix}. \quad (12)$$

with $\delta\mathbf{u_1}$, $\delta\mathbf{u_2}$, $\delta\mathbf{m}$, $\delta\boldsymbol{\lambda_1}$ and $\delta\boldsymbol{\lambda_2}$ as parameter updates. After each Newton iteration the inversion produces updates for all five different types of parameters:

$$\begin{pmatrix}\boldsymbol{u}_1^{n+1}\\\boldsymbol{u}_2^{n+1}\\\boldsymbol{m}^{n+1}\\\boldsymbol{\lambda}_1^{n+1}\\\boldsymbol{\lambda}_2^{n+1}\end{pmatrix} = \begin{pmatrix}\boldsymbol{u}_1^{n}\\\boldsymbol{u}_2^{n}\\\boldsymbol{m}^{n}\\\boldsymbol{\lambda}_1^{n}\\\boldsymbol{\lambda}_2^{n}\end{pmatrix} + \begin{pmatrix}\boldsymbol{\delta u}_1\\\boldsymbol{\delta u}_2\\\boldsymbol{\delta m}\\\boldsymbol{\delta\lambda}_1\\\boldsymbol{\delta\lambda}_2\end{pmatrix}. \quad (13)$$

Thus, explicit forward modelling after each iteration is not necessary.

OUTLOOK

$\mathcal{H}_{kkt}$ is a blockwise singular matrix where only cross-diagonal blocks are invertible. As a part of a large but sparse linear system of equations, it can be transformed into a positive definite matrix by permutation of block rows and columns.

The reduced Hessian method as described by Haber and Ascher (2001) is going to be the approach we will use to solve the permuted KKT system. The first try is the Gauß-Newton approximation where second-order information (like $\mathbf{R}_1$ and $\mathbf{R}_2$) is ignored.

ACKNOWLEDGMENTS

W.W. likes to thank Eldad Haber for his support via personal communication.

References

Haber, E., & Ascher, U. M. (2001). Preconditioned all-at-once methods for large, sparse parameter estimation problems. *Inverse Problems*, *17*, 1847 - 1864.

Haber, E., Ascher, U. M., & Oldenburg, D. (2000). On optimization techniques for solving nonlinear inverse problems. *Inverse Problems*, *16*, 1263 - 1280.

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

Zhdanov, M. S., & Weaver, J. T. (1997). Methods for modelling electromagnetic fields. *Journal of Applied Geophysics*, *37*, 133 - 271.