

We E108 04

## A Low-rank Approximation to the Hessian for 3D CSEM Gauss-Newton Inversion

M. Amaya\* (Norwegian University of Science and Technology), L. Boman (Electromagnetic Geoservices ASA) & J.P. Morten (Electromagnetic Geoservices ASA)

### SUMMARY

---

Use of controlled-source electromagnetics in increasingly challenging exploration applications has led to the requirement for more powerful 3D inversion approaches. For 3D cases, application of Gauss-Newton algorithms is limited by the computational cost required to compute the Hessian matrix and solve for the model update. We consider a low-rank approximation to the Hessian matrix, which has the potential to reduce the numerical complexity drastically. The scheme is based on phase encoding groups of sources instead of incorporating sources individually. We demonstrate the feasibility of the approach by numerical examples and present an analysis of the errors introduced by the approximation.

## Introduction

The application of 3D controlled-source electromagnetics (CSEM) to image and characterize targets in increasingly challenging environments has motivated the development of more powerful inversion methods. The target response often represents a small perturbation of the measured signal response from complex background resistivity variations. When the geological understanding is limited, we require the 3D CSEM inversion to reconstruct from the data not only the potential hydrocarbon reservoir target, but also an accurate representation of the background resistivity variation and the structural framework.

The Gauss–Newton optimization algorithm is known to work well for inversion of CSEM data when assumptions of lower spatial dimensionality can be applied (Abubakar et al., 2006; Mittet et al., 2007). When a 3D model description is required, and when the input from state-of-the-art 3D acquisition is to be used, the numerical complexity of the Gauss–Newton algorithm can be very large (Abubakar et al., 2009; Sasaki, 2011). The large size of the Jacobian and Hessian matrices, as well as the number of 3D forward simulations can be a severe limitation. This has been addressed by several authors considering schemes to reduce the numerical cost by e.g. model reparameterization (Lin et al., 2013), and input data decimation (Schwarzach and Haber, 2011). Gradient-based approaches to 3D CSEM inversion, like conjugate-gradient and quasi-Newton (Mackie et al., 2007; Støren et al., 2008), are less computationally demanding, and are now commonly used. However, these approaches are most accurate when a good background model has been built. The construction of the background model can be a demanding task if the geology is complex and if little other geophysical data is available.

In this paper we present a Hessian approximation based on the superposition of phase-encoded sources. This approach leads to a low-rank representation of the Hessian matrix, and alleviates the computational cost of constructing and storing this matrix as well as the solution of the Gauss–Newton equation. We show by numerical examples how the approximation is able to capture important features of the Hessian, at a numerical cost that is up to two orders of magnitude smaller than the exact calculation.

## Gauss–Newton optimization and Hessian approximation

The inversion of CSEM data is formulated as an optimization problem  $\sigma = \arg \min_{\sigma \in \mathcal{M}} \varepsilon(\sigma)$ , where  $\sigma$  is a 3D conductivity model in the set  $\mathcal{M}$  of models compatible with a priori information, and  $\varepsilon$  is the cost function. The cost function includes both regularization terms and a data misfit term,

$$\varepsilon_{\text{Data}}(\sigma) = \sum_{F,i,f,\mathbf{r}_{\text{rx}},\mathbf{r}_{\text{tx}}} |W_i^F(\mathbf{r}_{\text{rx}}|\mathbf{r}_{\text{tx}},f)\Delta F_i(\mathbf{r}_{\text{rx}}|\mathbf{r}_{\text{tx}},f;\sigma)|^2. \quad (1)$$

Here  $\Delta F = F^{\text{Obs}} - F^{\text{Synth}}$  represents the difference between observed and synthetic fields ( $F = E$  for electric and  $F = H$  for magnetic),  $W$  is a datum weight (typically inverse measurement uncertainty),  $i$  are the spatial components ( $x, y$ ) of the field recordings,  $f$  are the frequencies,  $\mathbf{r}_{\text{rx}}$  is a receiver position, and  $\mathbf{r}_{\text{tx}}$  is a source position. The shorthand notation  $\kappa = (F, i, f, \mathbf{r}_{\text{rx}}, \mathbf{r}_{\text{tx}})$  will uniquely label a measurement.

The non-linear optimization problem is solved by iteratively updating the conductivity model. The Gauss–Newton equation for model updates  $\Delta\sigma$  is  $\mathbf{H}\Delta\sigma = -\mathbf{g}$  where  $\mathbf{H} = \mathbf{J}^\dagger\mathbf{J} + \text{c.c.}$  is the Hessian matrix constructed from the Jacobian matrix  $\mathbf{J}$ , and  $\mathbf{g} = \sum_{\kappa} W_{\kappa}\Delta F_{\kappa}^*(\mathbf{J})_{\kappa} + \text{c.c.}$  is the model parameter gradient of the cost function. The c.c. denotes complex conjugated term. The Jacobian is a complex  $N \times M$  matrix where  $N$  is the number of data samples, and  $M$  is the number of model parameters. The Jacobian can be constructed from Green functions,

$$(\mathbf{J})_{\kappa,\mathbf{r}} = W_i^F(\mathbf{r}_{\text{rx}}|\mathbf{r}_{\text{tx}},f)\sum_m G_{i,m}^{F,J}(\mathbf{r}_{\text{rx}}|\mathbf{r},f)\sum_n G_{m,n}^{E,J}(\mathbf{r}|\mathbf{r}_{\text{tx}},f)J_n(\mathbf{r}_{\text{tx}},f), \quad (2)$$

where  $\mathbf{r}$  is the position in the model,  $G_{m,n}^{F,J}$  denotes the Green function for field  $F$ , component  $m$ , from a unit electric current source in direction  $n$ , and  $J_n$  is a component of the source ( $m, n = x, y, z$ ). It

is straightforward to generalize the expression in (2) to the anisotropic and discrete case. From this expression, we see that explicit construction of the Jacobian requires the Green function associated with every source position  $\mathbf{r}_{\text{tx}}$  to be calculated. The Hessian matrix is a real  $M \times M$  matrix, with rank given by the number of rows in the Jacobian, i.e.  $\text{rank}(\mathbf{H}) = 2 N_F N_f N_i N_{\text{rx}} N_{\text{tx}}$  (Grayver et al., 2013). Here, and throughout, the notation  $N_a$  denotes the number of unique elements of index  $a$ . For a state-of-the-art 3D CSEM survey and with a realistic model representation, the numerical complexity involved with the construction of  $\mathbf{H}$  and the solution for the model update can be very large. The number of forward solutions required can be of order  $10^4$ , and the dense linear system for  $\Delta\sigma$  can be of size  $10^6 \times 10^6$ .

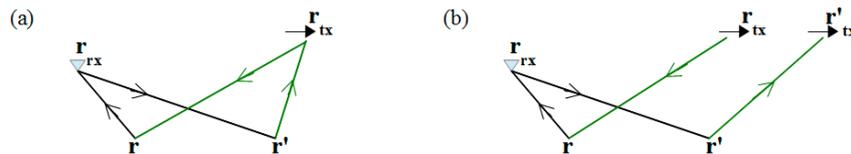
In this paper, we consider a low-rank approximation where sources in (2) are combined after encoding with a random phase factor, i.e. we construct  $\sum_{\kappa \in g} e^{i\phi_\kappa} (\mathbf{J})_\kappa$  for a group of sources  $g$  associated with a receiver channel, and where  $\phi_\kappa$  are uniformly distributed random numbers in the interval  $[0, 2\pi)$ . The number of source groups  $N_g$  and the grouping scheme will be discussed below. Following this approach, the factors of  $\mathbf{J}$  associated with the sources in a source group and a specific receiver channel can be calculated from a single simultaneous-source (super-shot) forward solution of the Maxwell equations. We denote the output of such simulation,

$$\tilde{G}_{i,m,\mathbf{r}_{\text{tx}},g}^F(\mathbf{r}, f) = \sum_{n, \mathbf{r}_{\text{tx}} \in g} W_i^F(\mathbf{r}_{\text{tx}}|\mathbf{r}_{\text{tx}}, f) J_n(\mathbf{r}_{\text{tx}}, f) e^{i\phi_{E,i,\mathbf{r}_{\text{tx}},f}} G_{m,n}^{E,J}(\mathbf{r}|\mathbf{r}_{\text{tx}}, f). \quad (3)$$

The approximate Hessian matrix  $\tilde{\mathbf{H}}$  following from the Jacobian constructed in this approach becomes,

$$\tilde{\mathbf{H}}(\mathbf{r}, \mathbf{r}') = \sum_{F,i,f,\mathbf{r}_{\text{tx}},g} \left[ \sum_m G_{i,m}^{F,J}(\mathbf{r}_{\text{tx}}|\mathbf{r}, f) \tilde{G}_{i,m,\mathbf{r}_{\text{tx}},g}^F(\mathbf{r}, f) \right] \times \left[ \sum_n G_{i,n}^{F,J}(\mathbf{r}_{\text{tx}}|\mathbf{r}', f) \tilde{G}_{i,n,\mathbf{r}_{\text{tx}},g}^F(\mathbf{r}', f) \right]^* + \text{c.c.} \quad (4)$$

The rank of the approximation is given by the number of terms in the outer sum, i.e.  $\text{rank}(\tilde{\mathbf{H}}) = 2 N_F N_i N_f N_{\text{rx}} N_g$ . Note that both the rank and the storage requirement to construct  $\tilde{\mathbf{H}}$  scale by  $N_g$  instead of  $N_{\text{tx}}$  as for  $\mathbf{H}$ . The reduction in numerical complexity from the approximation is described below, but it is realistic that the ratio  $N_{\text{tx}}/N_g$  can be of order 10 – 100 meaning a dramatic decrease in complexity. Consider now the errors introduced by the approximation to the Hessian in (4). Due to the summation over source positions in  $\tilde{\mathbf{H}}$ , the approximation will include terms involving two different source positions. Such terms are not present in  $\mathbf{H}$ , see Figure 1.



**Figure 1** (a) Diagram representing terms included in the Hessian  $\mathbf{H}$ . Arrows correspond to Green functions, with reverse directions indicating complex conjugation. Current factors  $W_i^F(\mathbf{r}_{\text{tx}}|\mathbf{r}_{\text{tx}}, f) J_n(\mathbf{r}_{\text{tx}}, f) e^{i\phi_{E,i,\mathbf{r}_{\text{tx}},f}}$  are associated with source positions. (b) Diagram representing the additional cross-talk terms introduced into  $\tilde{\mathbf{H}}$ , where two separate source positions give a contribution.

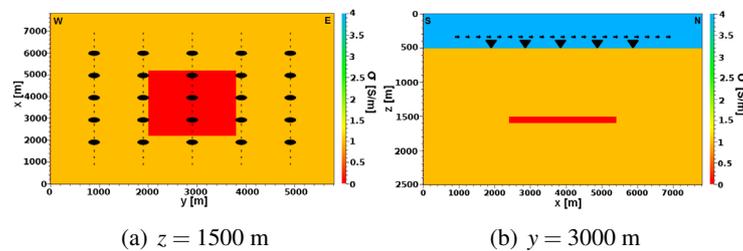
We refer to these errors as “cross-talk” and denote their contribution  $\eta$ , such that  $\tilde{\mathbf{H}} = \mathbf{H} + \eta$ . The source-diagonal terms, involving only a single  $\mathbf{r}_{\text{tx}}$ , are the terms that make up the exact Gauss–Newton Hessian  $\mathbf{H}$ . The random phase-factors  $e^{i\phi_\kappa}$  introduced in (3) will cancel in the source-diagonal terms since they enter as an absolute value. However, for the cross-talk terms, where two different source points are involved, the phase factors remain and act to suppress the cross-talk in the outer summation in (4). This is similar to applications of phase encoding in seismic modeling, see e.g. Bansal et al. (2013). The number of source-diagonal terms contributing to  $\mathbf{H}$  is proportional to  $N_{\text{tx}}$ . The number of cross-talk terms contributing to  $\eta$  will scale with the number of sources as  $N_{\text{tx}}^2$  (assuming  $N_g = 1$ ). However, the magnitude  $|\eta|$  should scale linearly in  $N_{\text{tx}}$  due to the random phase of the cross-talk terms from the  $e^{i\phi_\kappa}$  factors, and by analogy to a Gaussian random walk. Further, the number of significant

terms contributing to  $\eta$  should be less than  $N_{\text{tx}}^2$  due to the exponential decay of the Green functions for cross-talk terms where  $|\mathbf{r}_{\text{tx}} - \mathbf{r}'_{\text{tx}}|$  is large. Thus the magnitude  $|\eta|$  scales by the number of sources in a source group as  $(N_{\text{tx}})^\alpha$  with  $\alpha < 1$ . In summary, the asymptotic behavior of the approximation is feasible since  $\lim_{N_{\text{tx}} \rightarrow \infty} |\eta|/|\mathbf{H}| = 0$ . We can reduce the error of the approximation by increasing the number of source groups  $N_g$ . In fact, using the maximum  $N_g = N_{\text{tx}}$  makes  $\tilde{\mathbf{H}}$  identical to  $\mathbf{H}$ , but in this case there is no reduction in computational cost. We can optimize the approximation by constructing the source groups with maximum separation between the spatial locations of sources. In this case, each cross-talk contribution shown in Figure 1 (b) will be smaller compared to a source-diagonal contribution in Figure 1 (a) by the decay of the Green functions over distance  $|\mathbf{r}_{\text{tx}} - \mathbf{r}'_{\text{tx}}|$ .

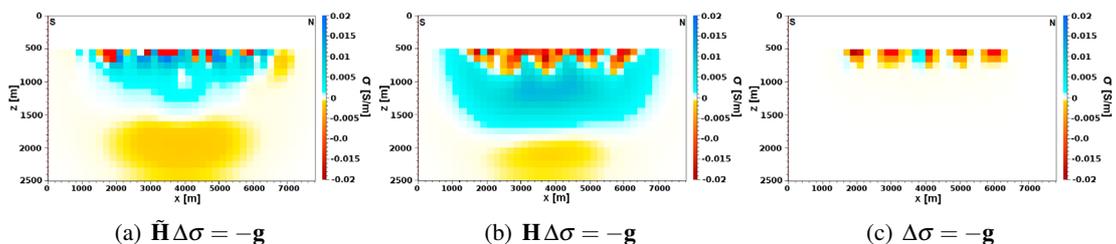
The approximation  $\tilde{\mathbf{H}}$  to the Hessian will be better for some parts of the matrix than others. A good implementation into an inverse scheme can use the approximation where the errors are below tolerance, and use the Gauss–Newton Hessian when accuracy is critical. In a standard Gauss–Newton implementation, the number of forward computations,  $N_{\text{sim}}$ , is mainly driven by the number of source Green functions needed, that is  $N_{\text{sim}} \sim N_{\text{tx}}$ . Likewise, the memory needed to store the Jacobian matrix scales linearly with  $N_{\text{tx}}$ . When the Hessian matrix is built using source groups and following the approach described above, both the number of simulations and the memory requirements for the Jacobian scale as  $N_g$  instead of  $N_{\text{tx}}$ . The low-rank approximation allows a Hessian representation using considerably less memory when  $N_g$  is small by storing the quantities in square brackets in (4). In addition to the savings on number of forward solutions and memory requirements, we expect that solving the Gauss–Newton equation can be done very efficiently by exploiting the low-rank property of the Hessian matrix in (4).

## Results

In this section we show model updates computed for an example CSEM survey, using both the Gauss–Newton update and the approximation scheme described above. In the example, the inversion parameterization is a regular grid with cell size  $200 \text{ m} \times 200 \text{ m} \times 100 \text{ m}$  and the total number of cells is 28275. The survey layout is detailed in Figure 2 with a total of 10500 data samples.



**Figure 2** True model and survey layout for the example CSEM survey. The water conductivity is 4 S/m, and the water depth is 500 m. A resistor is located at 1.5 km depth, with dimensions  $3 \text{ km} \times 2 \text{ km} \times 0.1 \text{ km}$ , and conductivity 0.02 S/m. The formation conductivity is 1 S/m. There are 5 towlines and 25 receivers recording  $E_x$  and  $E_y$  at 0.25 and 1.0 Hz. The source distance is 300 m along towlines.



**Figure 3** Model updates  $\Delta\sigma$  at  $y = 3000 \text{ m}$ .

The Gauss–Newton equation was solved using a conjugate gradient method with a small stabilizer and an initial model with the correct background conductivity. Figure 3 shows the solution for three cases using (a) the approximate Hessian matrix  $\tilde{\mathbf{H}}$ , (b) the exact Gauss–Newton Hessian matrix  $\mathbf{H}$ , and (c)

steepest descent ( $\mathbf{H} \rightarrow \mathbf{1}$ ). The approximate Hessian was obtained using one source group per receiver channel ( $N_g = 1$ ). Comparing (a) with (b) and (c), we see that the solution to the Gauss–Newton equation with the approximate Hessian matrix is qualitatively more similar to a solution with the exact Hessian matrix than a steepest descent solution. In particular, much of the sensitivity information in  $\mathbf{H}$  remains in  $\tilde{\mathbf{H}}$  as seen at depth. Table 1 shows key characteristics for the computational cost of inversion for the example survey as well as for a larger, more realistically sized 3D CSEM survey. As is shown in the table, the number of forward simulations is reduced with a factor 2.6, however in a larger survey the reduction in computational cost can be much larger while keeping the error at the same level.

Case	Numerical cost	Standard GN	Approximate $\tilde{\mathbf{H}}$	Ratio
Example	Forward simulations	260	100	2.6
Example	Jacobian memory	4.4 GB	43.1 MB	105
Realistic	Forward simulations	11800	1000	11.8
Realistic	Jacobian memory	127.4 TB	68.7 GB	1900

**Table 1** Computational cost comparison for exact Gauss–Newton and the approximation scheme, for the simple example case shown in Figures 2 and 3, as well as a large-scale realistic survey.

The data for the realistic survey in Table 1 were obtained from a survey area of 30 km×20 km×4 km, with 10 towlines at 2 km line separation and a source distance of 100 m along the towlines. In total we obtain 5700 source positions recorded at 100 receiver sites, measuring the horizontal components of electric and magnetic fields at 4 frequencies. The same discretization as in the smaller example is assumed, for two anisotropy components. The survey has a total of 14.59 million data samples. Using three source groups ( $N_g = 3$ ) we keep the simultaneous source separation at 300 m such as in the smaller example survey discussed above. The size of the Hessian will be 2.6 TB, but the Jacobian representation in the approximation scheme offers a 40-fold reduction in size. The approximation error,  $\eta$ , could be reduced by increasing the number of source groups, but the computational cost would increase.

## Conclusions and acknowledgement

We have described a low-rank approximation to the Hessian for Gauss–Newton 3D inversion of CSEM data. The scheme is based on superposition of phase-encoded sources, and we have demonstrated the potential to significantly reduce both the number of forward simulations and memory requirements for inversion. We thank Research Council of Norway (PETROMAKS project 217223) and EMGS ASA for supporting this work.

## References

- Abubakar, A., Habashy, T., Druskin, V., Alumbaugh, D., Zerelli, A. and Knizhnerman, L. [2006] Two-and-half-dimensional forward and inverse modeling for marine CSEM. *SEG Expanded Abstracts 2006*, **150**, 750.
- Abubakar, A., Habashy, T.M., Li, M. and Liu, J. [2009] Inversion algorithms for large-scale geophysical electromagnetic measurements. *Inverse Problems*, **25**, 123012.
- Bansal, R. et al. [2013] Simultaneous-source full-wavefield inversion. *The Leading Edge*, **32**, 1100.
- Grayver, A.V., Streich, R. and Ritter, O. [2013] Three-dimensional parallel distributed inversion of CSEM data using a direct forward solver. *Geophysical Journal International*, **193**, 1432.
- Lin, Y., Li, M., Abubakar, A. and Habashy, T. [2013] A wavelet-based model compression method for three-dimensional electromagnetic data inversion. *SEG Expanded Abstracts 2013*, **138**, 707.
- Mackie, R., Watts, M. and Rodi, W. [2007] Joint 3D inversion of marine CSEM and MT data. *SEG Expanded Abstracts 2007*, **116**, 574.
- Mittet, R., Maulana, H., Brauti, K. and Wicklund, T.A. [2007] CMP inversion of marine CSEM data. *EGM 2007 International Workshop*.
- Sasaki, Y. [2011] Gauss-newton-based 3D joint inversion of marine CSEM and MT data. *73rd EAGE Conference & Exhibition, 2011, C027*.
- Schwarzbach, C. and Haber, E. [2011] Finite element based inversion for electromagnetic problems using stochastic optimization. *SEG Expanded Abstracts 2011*, **110**, 567.
- Støren, T., Zach, J. and Maaø, F. [2008] Gradient calculations for 3D inversion of CSEM data using a fast finite-difference time-domain modelling code. *70th EAGE Conference & Exhibition, 2008, P194*.