Using the parallel MUMPS solver for frequency domain full waveform inversion

Scott Keating and Kris Innanen

ABSTRACT

Frequency domain finite difference (FDFD) wave propagation is often used in full waveform inversion (FWI) research. The major cost in FDFD is the factorization of a large, sparse matrix. This means that solving the problem for an additional source term is inexpensive, but also raises challenging problems for parallelization. Parallelization over source terms offers negligible benefits as no re-factorization of the matrix is required. Parallelization over different frequencies is possible, but limited by the number of frequencies considered, which may be small. This work focuses on using the MUMPS solver, which allows for massive parallelization of the factorization itself.

INTRODUCTION

Much research in full waveform inversion (FWI) is done using frequency domain finite difference wave propagation (e.g. Pratt et al., 1998; Metivier et al., 2013). This approach is appealing for a number of reasons, for small models it is very fast approach, it naturally allows for consideration of specific frequencies in a multi-scale approach, and it easily allows for treatment of attenuation and dispersion, which require the use of convolutional operators in the time domain (Casula and Carcione, 1992). Although frequency domain finite difference (FDFD) wave propagation is impractical for very large models, research in FWI using FDFD on small scales has provided useful insights on FWI (e.g. Pratt et al., 1998; Metivier et al., 2013).

Parallelization is often used in FWI in order to bring more computational power to bear on the problem and decrease the time required to obtain solutions. In the time domain, parallelization can be easily implemented by forward propagating each source on a different CPU simultaneously. Given the large number of sources typically considered in seismic experiments, this allows for a large degree of parallelism. If even greater parallelism is desired, this can be achieved by decomposing the model into sub-domains, each of which can be solved largely independently.

FDFD wave propagation is more difficult to parallelize. Parallelizing over sources provides little benefit due to the relatively small cost of solving for another source once the Helmholtz matrix is factorized. Parallelizing over frequencies is effective, but in the FWI problem only a few frequencies are typically considered at any given iteration, severely restricting the level of parallelism that can be achieved in this way. A more effective way of employing parallel computation is to parallelize the factorization of the Helmholtz matrix. This process is quite complicated, and can be challenging to implement. One existing tool for parallelizing factorization of large, sparse matrices is the MUltifrontal Massively Parallel sparse direct Solver (MUMPS) (Amestoy et al., 2000). MUMPS has been used successfully in FWI research to decrease compute times required (e.g. Metivier et al., 2013). This report outlines the motivation behind using the MUMPS solver.

THEORY

The goal in FWI is to recover a subsurface model which reproduces the observed data. This is typically done by minimizing a measure of the discrepancy between synthetic and measured data, called the objective function. This minimization procedure can require calculation of the objective function, its gradient and second derivative terms. Due to the presence of the synthetic data in the objective function, all of these values require wavefield modeling, typically done using finite differences. This wavefield modeling dominates the cost of the FWI procedure.

Computational costs associated with FWI are very large, causing the procedure to be very time consuming. If sufficient computational resources are available, computation times can be improved by making use of parallel computing, where different computations are performed simultaneously on different machines. Parallel computing is effective if computationally demanding steps can be broken into independent sub-problems. The cost of FWI largely arises from the minimization procedures used, but these are iterative, so each step requires information produced at the previous step. Parallelism in FWI, then, must employ parallel computing within each step of the procedure. Given the large cost of wavefield modeling, the greatest benefits are gained by parallelizing this portion of the algorithm. This report focuses on the details of this parallelization when using frequency domain finite difference wave propagation.

Frequency domain finite difference wave propagation

Frequency domain modeling is useful in FWI research for several reasons. Firstly, it is usually necessary to employ a multiscale approach in FWI (Bunks et al. (1995)), where early iterations consider only low frequency information and high frequency information is gradually introduced at later iterations. The frequency domain lends itself to this approach, as it allows for efficient methods which only model the frequencies that are used at each iteration. Secondly, constant Q attenuation of the form often considered is difficult to model in the time domain, requiring the use of convolutional operators, whereas in the frequency domain it is relatively simple to introduce these attenuation terms. In this report, we use as an example anacoustic wave propagation of the form

$$\left[\omega^2 s(\mathrm{r}) + \nabla^2\right] u(\mathrm{r},\omega) = f(\mathrm{r},\omega) \quad , \tag{1}$$

where the model parameter s is given by

$$s(\mathrm{r},\omega) = \frac{1}{c^2(\mathrm{r})}\left\{1 + \frac{1}{Q(\mathrm{r})}\left[i - \frac{2}{\pi}\log\left(\frac{\omega}{\omega_0}\right)\right]\right\} \quad , \tag{2}$$

c is the acoustic wave velocity, Q is the quality factor, ω_0 is a reference frequency, u is the pressure field, and f is a source term. This equation can be solved for u by the frequency domain finite difference (FDFD) method. In FDFD, the x second derivative operators are discretized as

$$\frac{\partial^2 \mathbf{u}}{\partial x^2} = \frac{u_{h-1,j} - 2u_{h,j} + u_{h+1,j}}{\Delta x^2} \quad , \tag{3}$$

where Δx is the x spacing of the model, and $u_{h,j}$ denotes the pressure field at the hth x position and jth z position (Franklin (2005)). The expression for the discretization of the z

second derivative operators is similar. With the discretized derivative operators, equation 1 can be restated as

$$\left[\omega^2 s_{h,j} u_{h,j} + \frac{u_{h-1,j} - 2u_{h,j} + u_{h+1,j}}{\Delta x^2} + \frac{u_{h,j-1} - 2u_{h,j} + u_{h,j+1}}{\Delta z^2}\right] = f_{h,j} \quad . \tag{4}$$

If positions are mapped to a single 1D index k via $k = (h-1) * N + j$, we can follow Franklin (2005) and restate eq. 4 as

$$\mathbf{Mu} = \mathbf{f} \quad , \tag{5}$$

where **u** and **f** are vectors, and **M** is a sparse matrix. Equation 5 can then be solved for **u**.

Parallelizing over sources

Solving equation 5 typically consists of factorizing **M**, and performing back-substitution to solve for **u**. For a model of N elements, the factorization procedure requires approximately $\mathcal{O}(N^3)$ operations, while the back-substitution requires $\mathcal{O}(N^2)$ operations. Solving the system for a different f does not require a new factorization, only a new back-substitution. Consequently, parallelizing this computation over sources can only reduce the time associated with the $\mathcal{O}(N_S N^2)$ computation associated with one back-substitution for each of N_S sources. In exploration seismology many sources are used, but necessarily there are many more finite difference grid points than there are sources, so the cost of this back-substitution will not dominate the total cost.

Parallelizing over frequencies

FDFD in FWI offers the opportunity to consider only the frequencies desired at each iteration. The number of frequencies considered is usually kept as small as possible to increase efficiency. Each frequency considered requires that the matrix M in equation 5 be changed, and re-factorized. This is the major cost in FDFD wave propagation, and so doing this factorization in parallel for each frequency can offer significant improvements to speed. The small number of frequencies considered, however, severely limits the speed increases which can be achieved with this approach.

MUMPS

The MUMPS solver allows for parallel computation to be used in factorizing the matrix **M**. Like parallelization over frequencies, this reduces the time needed for the most costly part of the FWI procedure. Unlike parallelization over frequencies, however, a large number of processors can be used in parallel for solving this problem. By employing MUMPS, we expect to be able to accelerate FWI computations, allowing for larger models to be considered in future research.

CONCLUSIONS

Frequency domain FWI is a challenging problem to efficiently parallelize to a large degree. Parallel computation over different sources provides little decrease in compute

time, while parallelization over different frequencies can only offer limited parallelism. MUMPS provides a means of implementing frequency domain FWI with a large degree of parallelism, which should decrease compute times, and make research on larger synthetic models feasible.

ACKNOWLEDGMENTS

The authors thank the sponsors of CREWES for continued support. This work was funded by CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 461179-13. Scott Keating was also supported by the Earl D. and Reba C. Griffin Memorial Scholarship..

REFERENCES

Amestoy, D., Duff, I., and L'Excellent, J. Y., 2000, Multifrontal parallel distributed symmetric and unsymmetric solvers: Comput. Methods Appl. Mech. Engrg., **184**, 501–520.

Bunks, C., Salek, F., Zaleski, S., and Chavent, G., 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

Casula, G., and Carcione, J., 1992, Generalized mechanical model analogies of viscoelastic behaviour: Bolletina di Geofisica Teorica ed Applicata, **34**, 235–256.

Franklin, J., 2005, Frequency domain modeling techniques for the scalar wave equation, an introduction: MIT Earth Resources Laboratory Industry Consortia Annual Report, 1–22.

Metivier, L., Brossier, R., Virieux, J., and Operto, S., 2013, Full waveform inversion and the truncated newton method: Siam J. Sci. Comput., **35**, B401–B437.

Pratt, R., C[]Shin, and G[]J[]Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophys. J. Int., **133**, 341–362.