Introduction

Elastic full-waveform inversion (EFWI) is a powerful technique that can be used to infer elastic parameters (such as P- and S-wave velocities) of the medium from seismic data observed at a set of receivers (Mora, 1987; Virieux and Operto, 2009). EFWI is in practice a local optimization problem. Starting from an initial estimate, the medium parameters can be updated iteratively by incrementing them with the derivatives of the misfit functional (which measures the error between the observed and simulated data) with respect to each of the medium parameters. The efficient implementation of the derivative computations require enormous amount of computer memory storage. The large requirements in terms of storage is one the main barriers for the application of this method to large scale 3D problems.

In this paper, we propose and test a strategy based on reverse-time wavefield reconstruction using the Kirchhoff integral that effectively reduces the storage requirements, at the cost of a mere factor of two increase in the computational runtime. Different from methods based on checkpointing (Griewank and Walther, 2000), the method does not produce an exact reproduction of the derivatives. But the artifacts introduced are uncorrelated from source to source, such that, in problems with large number of sources (i.e. 3D problems) the noise will tend to average out. To verify the effect of the imperfect derivatives on the convergence of EFWI, we test the strategy in a moderate size 3D synthetic dataset.

Method and Theory

Elastic full-waveform inversion

Elastic full-waveform inversion is a classical non-linear inverse problem (Mora, 1987). Elastic waves generated by a source propagate in a heterogeneous elastic medium and the displacements are measured at a finite number of receivers. By minimizing the misfit between the observed and the simulated seismic data at the receivers the values of the elastic parameters of the medium can be estimated. There are different measures of the misfit between observed and simulated data, the most common being the least squares difference. In its most simple form, the least-squares norm can be written as

$$S = \frac{1}{2}\|u_i - u_i^o\|^2 = \sum_s \int dt \int dx_r (u_i(x_r,t,s) - u_i^o(x_r,t,s))^2 \qquad (1)$$

where, u_i is the simulated data, and u_i^o is the observed data, x_r are receiver coordinates, t is time, and s is a source index.

The evaluation of the misfit functional (equation 1) is computationally expensive. In addition, the solution space can be very large (generally in the order of millions of unknowns). For these reasons, global optimization solutions to the minimization problem do not work well, and therefore one has to rely on gradient descent based methods. The gradient descent methods require the evaluation of the misfit functional and its gradient with respect to the model parameters at each iteration. To compute the gradient of equation 1 with respect to the parameters, it is customary to employ the adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Plessix, 2006). This method gives the following equations for the gradients with respect to the model parameters m

$$\frac{\partial S}{\partial m}(\mathbf{x}) = \int ds \int dt \, \frac{\partial c_{ijkl}}{\partial m}(\mathbf{x}) \frac{\partial u_l}{\partial x_k}(\mathbf{x},t,s) \frac{\partial \tilde{u}_i}{\partial x_j}(\mathbf{x},T-t,s) \qquad (2)$$

where c_{ijkl} is the elasticity tensor, $\mathbf{x}$ are Cartesian coordinates, u_i are the forward modeled displacement wavefields, $\tilde{u}_i$ are reverse-time modeled residual displacement wavefields. The equations used to compute these wavefields can be found in Mora (1987).

Equation 2 shows that $\partial S/\partial m$ can be computed from the temporal accumulation of a weighted crosscorrelation of the forward modeled source wavefields and the reverse-time modeled adjoint wavefields. In order to perform the crosscorrelation, both forward and reverse-time modeled wavefields must be made accessible at the same time instances. This problem has classically been solved by first modeling and storing all the time instances of the forward wavefields in memory (checkpointing) and then, during the reverse-time modeling, accessing them as necessary. In the case of 3D EFWI, this strategy requires the availability of several hundreds of terabytes of memory for any realistically sized problem. To overcome this challenge, other strategies have been suggested, all of them involving some combination of disk storage and recomputation of the forward wavefields. In the following, we describe a few of these strategies and the their requirements.

Consider an elastic medium $\Omega \subset \mathbf{R}^3$ surrounded by the boundary $\partial\Omega$. For any $(\mathbf{x},t) \in \Omega \times T$, where T is some time interval, we let $u_i(\mathbf{x},t)$ be the particle displacement in direction i, $f_i(\mathbf{x},t)$ is the body force in direction i, and c_{ijkl} is the elasticity tensor. An elastic wave propagating in Ω can be explained by the elastodynamic wave equation for the displacement field,

$$\rho(\mathbf{x})\frac{\partial^2 u_i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x_j}\left[c_{ijkl}(\mathbf{x})\frac{\partial u_l}{\partial x_k}(\mathbf{x},t)\right] = f_i(x_s,t,s). \tag{3}$$

The above equation (equation 3) is solved numerically by rewriting it into a velocity-stress hyperbolic system of first-order partial differential equations, and solve the system using a staggered-grid finite difference method with high order spatial differential operators (Virieux, 1986; Holberg, 1987). In 3D, this system consists of a total of 9 equations with 9 field variables (6 stresses and 3 particle velocities). In addition to the system of 9 equations, perfectly matched layers (PML) are used to attenuate reflections from the boundaries of the modeling aperture. The PML equations are solved only at the boundaries and require some additional auxiliary variables.

Optimal checkpointing method

The only strategy that produce an exact reconstruction of the forward wavefields is the optimal checkpointing method (Griewank and Walther, 2000). It consists in a first instance of computing the forward modeling of the source wavefields and storing a small number of snapshots of these wavefields. Then during the reverse-time modeling of the adjoint wavefields, the forward wavefields necessary for the crosscorrelation are recomputed forward in time starting from the nearest stored time instance. The method requires the storage of snapshots of all fields variables and absorbing boundary auxiliary variables in order to accurately reconstruct the forward wavefields. There is a tradeoff between the number of stored snapshots and the amount of recomputation necessary. The optimal number of snapshots to store is then chosen such as to balance storage and recomputation. A detailed description of the method can be found in Griewank and Walther (2000).

Dirichlet boundary condition method

The cheapest strategies both in terms of storage and recomputation involve the storage of the forward wavefields at the boundaries of the computational grid together with the full snapshots of the last time. Given the lossless (self-adjoint) nature of the elastic wave equation, the wavefields can then be reconstructed in reverse-time from the boundaries. The reconstruction can be performed in two ways. One way uses Dirichlet boundary conditions to reconstruct the wavefields in reverse-time. The other uses the Kirchhoff integral (Mittet, 1994). Here we describe the Dirichlet boundary condition method. The second method is described next. The Dirichlet boundary method requires the storage of all time instances of forward wavefields of the 9 field variables at $\partial\Omega$. The auxiliary PML variables do not need to be stored. The forward wavefields are then reconstructed inside Ω in reverse-time from the boundaries by setting the stored variables as boundary conditions. Since both the forward and adjoint wavefields are now modeled in reverse-time, both wavefields needed for crosscorrelation are available at all time instances. The recomputation ratio of this method is two, due to the fact that the forward wavefields need to be computed twice, once forward in time and once in reverse-time. Figure 1c) and 1d) show an example of the reconstruction of the forward wavefields of the vertical particle velocity ($\dot{u}_z$) using this strategy. The black stippled lines show the boundaries where the fields are stored. Inside the bound-

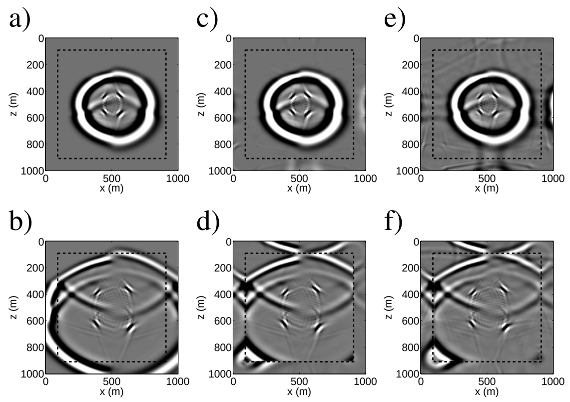

Figure 1 Vertical profiles of snapshots of vertical particle velocity at two different times (0.2 and 0.3 s). Figures a) and b) are exact reproductions; c) and d) are reconstructed with Dirichlet boundary method; e) and f) are reconstructed with the Kirchhoff integral method. The black stippled lines mark the boundaries where the fields are stored, and inside which the wavefields are reverse-time reconstructed.

aries, the wavefields are reconstructed accurately, although with small artifacts in phase and amplitude. Outside the boundaries, there are spurious outgoing waves. These outgoing waves need to be attenuated with PML absorbing boundary strips.

Kirchhoff integral method

According to the Kirchhoff integral, 6 fields (3 particle velocities and 3 normal tractions) need to be known at the boundaries in order to accurately determine the direction, as well as the mode of wave propagation (P or S). Knowing these 6 variables at the boundaries, the wavefields inside Ω can be properly reconstructed in reverse-time from the boundaries in the direction they came from and in the mode they propagated (Mittet, 1994). However, since we are measuring and storing the wavefields at the outer boundaries, we know that the direction of the propagation is outwards. This means that we only need to store 3 wavefields (the three particle velocities, in this case) at the boundaries in order to reconstruct the wavefields inside Ω. The consequence is that the wavefields will be reconstructed in a mirrored fashion to both sides of the boundaries. Similarly to the Dirichlet boundary method, the artificial ghost wavefields can be attenuated with a PML absorbing boundary condition during the reverse-time reconstruction of the wavefields. In practice, as can be seen in Figure 1e) and 1f), the incomplete Kirchhoff reconstruction generate stronger artifacts than the Dirichlet boundary method. However, despite the spurious artifacts, the forward wavefields inside the boundaries are properly reconstructed in both amplitude and phase. This is the most cost-effective strategy since it requires only 3 fields to be stored at the boundaries, and reconstruct the forward wavefields at the cost of only one additional reverse-time modeling. The artifacts inside Ω, in form of spurious waves coming from the injection boundaries, are incoherent from source to source and will tend to stack out in the sum over sources. This is particularly true for problems with a large number of sources (i.e. large 3D problems).

Results

Figure 2a and 2b show vertical profiles through the P- and S-wave velocities (V_p and V_s) of a 3D synthetic elastic model. The model is adapted from the SEG/EAGE 3D Overthrust model. To test if the artifacts of the Kirchhoff method can have an effect in the convergence of EFWI, we compare the results of inversion for the P-wave velocities (V_p), and the S-wave velocities (V_s) using both the Dirichlet boundary condition method and the Kirchhoff integral method. Ideally, the comparison should be done between the exact reconstruction and the Kirchhoff integral method. However, even for this moderate size synthetic model, the optimal checkpointing method require either more storage than we have available, or unreasonable runtimes. For this reason, we are forced to compare two approximate methods. Figures 2c and 2d show the initial V_p and V_s models used to start the inversion. Figures 2e and 2f show results of inversion using the Dirichlet boundary method to compute the gradient of equation 1. Figures 2g and 2h show results of inversion using the Kichhoff integral method to compute the gradient of equation 1. Both the Dirichlet boundary method and Kirchhoff integral method show very similar inversion results, despite having different levels of artifacts in their wavefield reconstructions. This shows that the effect of the artifacts on the inversion results are negligible.

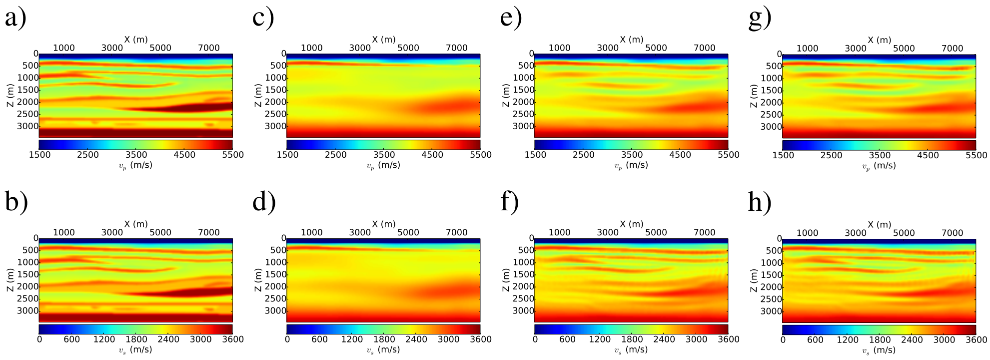

Figure 2 *Vertical profiles of a) true V_p; b) true V_s; c) initial V_p; d) initial V_s; e) inverted V_p using Dirchlet boundary method; f) inverted V_s using Dirichlet boundary method; g) inverted V_p using Kirchhoff integral method; h) inverted V_s using Kirchhoff integral method.*

Conclusions

Although none of the wavefield reconstruction strategies tested here are exact, they allow the inversion of larger scale models than would otherwise possible using the exact checkpointed wavefields. Further work is necessary to test the effect of the imperfect reconstruction against exact checkpointing over large scale 3D models.

Acknowledgements

We acknowledge the sponsors of the ROSE consortium and the BIGCCS centre and its partners for financial support of this work.

References

Chavent, G. and Lemonnier, P. [1974] Identification de la non-linéarité d'une équation parabolique quasilinéaire. *Applied Mathematics & Optimization*, **1**, 121–162.

Griewank, A. and Walther, A. [2000] Algorithm 799: Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Trans. Math. Softw.*, **26**(1), 19–45, ISSN 0098-3500, doi:10.1145/347837.347846.

Holberg, O. [1987] Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophysical Prospecting*, **35**(6), 629–655, ISSN 1365-2478, doi:10.1111/j.1365-2478.1987.tb00841.x.

Lions, J.L. and Magenes, E. [1972] *Nonhomogeneous boundary value problems and applications*. Springer Verlag, Berlin.

Mittet, R. [1994] Implementation of the Kirchhoff integral for elastic waves in staggered-grid modeling schemes. *Geophysics*, **59**(12), 1894–1901, doi:10.1190/1.1443576.

Mora, P. [1987] Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, **52**(9), 1211–1228, doi:10.1190/1.1442384.

Plessix, R.E. [2006] A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, **167**, 495–503.

Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC1–WCC26, doi:10.1190/1.3238367.

Virieux, J. [1986] P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method. *Geophysics*, **51**(4),