Simultaneous full waveform inversion for sources and anelastic models

Scott Keating and Kris Innanen

ABSTRACT

Full waveform inversion (FWI) can provide accurate estimates of a variety of mechanical properties in the subsurface. Very often, FWI is used to recover subsurface properties in situations where source properties are well understood. Other seismic inversion strategies exist for the recovery of unknown source properties, and these typically assume a known, simple model of subsurface mechanical properties. In practice, there exist many situations in which a better understanding of both the subsurface model and the seismic sources is desirable. As each of these properties can influence our estimate of the other, it may sometimes be necessary to simultaneously recover both a subsurface model and a characterization of sources. In this report, we propose such a simultaneous inversion, using a FWI approach. We show in a simple synthetic example that this approach can be effective in recovering sources and subsurface models, though some results suggest that sequential inversion approaches may be similarly effective in many settings.

INTRODUCTION

Seismic inversion approaches use measured seismic waves to generate estimates of the conditions which caused those waves. Two very broad categories of seismic inversion are those approaches which estimate the physical properties of the medium the waves travel through, and those which estimate the seismic sources that generate the measured waves. Both of these approaches are widely used on seismic data for a variety of applications. Inversion approaches for seismic media include tomography (e.g. Nolet, 1987), amplitude versus offset (AVO) analysis (e.g. Smith and Gidlow, 1987), migration (e.g. Stolt, 1978), and full waveform inversion (Tarantola, 1984). These types of inversion are very useful in exploration seismology and characterization of the Earth's crust and interior. Inversion approaches for seismic sources generally try to estimate the location, wavelets, and moment tensors of sources. These approaches have applications in microseismic monitoring (e.g. Maxwell et al., 2010), earthquake detection and characterization (e.g. Stich et al., 2010), and in nuclear test monitoring (Alvizuri and Tape, 2018).

Generally, when inverting for one set of properties, the others are held fixed, and assumed to be accurate. When trying to characterize the source mechanisms of earthquakes, for instance, it is common to assume that the medium the seismic waves travel through is well understood. While this simplifies the inversion problem considered, it also introduces a source of uncertainty in the inversion results: the medium estimate will not, in general, be perfectly accurate, and this will introduce errors in the estimated source properties. Because the sources and the medium interact in a nonlinear way to generate the wavefields measured, there will always be an uncertainty associated with inverting for only one of these properties while holding an estimate of the other fixed. In some cases, this uncertainty may be relatively large, suggesting that simultaneous inversion for both medium and sources may be necessary.

One inversion approach that has the capability to recover both source and medium properties is full waveform inversion (FWI). When originally proposed, FWI was suggested as a strategy that could determine both the acoustic properties of the subsurface and the time-dependent source functions at known source locations (Tarantola, 1984). In practice, FWI has developed primarily as an approach for the estimation of complex seismic media (Virieux and Operto, 2009), but source inversion approaches have been suggested as well. In this report, we will develop an FWI approach for simultaneous inversion of anelastic media, point-source locations and moment tensors. This simultaneous strategy may help to reduce uncertainty for applications where neither the source nor medium properties are well understood in advance.

THEORY

In FWI, the inverse problem is generally framed as an attempt to minimize the data misfit, subject to an assumed wave propagation model linking the wavefield and subsurface together. A general version of this problem can be stated as

$$\mathbf{m}_{min} = \underset{\mathbf{m}}{\text{argmin}}\, \phi(\mathbf{u}, \mathbf{m}, \mathbf{d}) \text{ subject to } C(\mathbf{u}, \mathbf{m}) = 0, \tag{1}$$

where $\mathbf{m}$ is a subsurface model, $\mathbf{m}_{min}$ is the inversion result, $\mathbf{u}$ is a simulated wavefield, $\mathbf{d}$ is the measured data, ϕ is an objective function, (which is small when $\mathbf{u}$ is consistent with measured data and $\mathbf{m}$ is consistent with prior information but larger otherwise), and $C = 0$ holds only when the wave equation assumed in the inversion is satisfied. In this report, we assume specific forms for the objective function and wave equation term, giving the optimization problem

$$\mathbf{m}_{min} = \underset{\mathbf{m}}{\text{argmin}}\, \phi = \underset{\mathbf{m}}{\text{argmin}} \sum_{j=1}^{N_F} \sum_{k=1}^{N_S} \frac{1}{2} ||\mathbf{R}\mathbf{u}_{j,k} - \mathbf{d}_{j,k}||^2, \text{ subject to } \mathbf{S}(\mathbf{m})\mathbf{u} = \mathbf{f}, \tag{2}$$

where $\mathbf{R}$ is a sampling matrix representing receiver measurement, S is a finite-difference forward modeling operator, $\mathbf{f}$ is a source term. In our discussion here, we will not explicitly state the sums over sources and frequencies for simplicity.

We can calculate the gradient for this optimization problem by using the adjoint state method. We start by considering the Lagrangian of the problem:

$$L = \frac{1}{2} ||\mathbf{R}\mathbf{u} - \mathbf{d}||_2^2 + < \mathbf{S}(\mathbf{m})\mathbf{u} - (\mathbf{f}_R + i\mathbf{f}_I), \kappa >, \tag{3}$$

where κ is an as-yet unconstrained Lagrange multiplier, $\mathbf{f}$ has been explicitly split into real and imaginary parts, and $<,>$ represents an inner product, such that $< \mathbf{a}, \mathbf{b} >= \sum_{k=1}^{N_S} \mathbf{a}_k^\dagger \mathbf{b}_k$. When $\mathbf{u} = \bar{\mathbf{u}}$ such that $\mathbf{S}\bar{\mathbf{u}} = \mathbf{f}$ (or, equivalently, we consider a wavefield satisfying the wave equation), then $L(\bar{\mathbf{u}}) = \phi$. We can then calculate the derivative of the objective function by considering the derivative of $L(\bar{\mathbf{u}})$. If we wish to find the derivative with respect to an inversion variable $\mathbf{x}$, we observe that

$$\frac{d\phi}{d\mathbf{x}} = \frac{dL(\bar{\mathbf{u}})}{d\mathbf{x}} = \frac{\partial L}{\partial \bar{\mathbf{u}}} \frac{\partial \bar{\mathbf{u}}}{\partial \mathbf{x}} + \frac{\partial L}{\partial \mathbf{x}}. \tag{4}$$

In practice, $\frac{\partial \bar{\mathbf{u}}}{\partial \mathbf{x}}$ is very difficult to calculate. This term can be eliminated by choosing κ such that $\frac{\partial L}{\partial \bar{\mathbf{u}}} = 0$. Assessing this derivative, it is evident that the $\bar{\kappa}$ which satisfies this condition can be calculated from

$$\frac{\partial L}{\partial \bar{\mathbf{u}}} = \mathbf{R}^T(\mathbf{R}\bar{\mathbf{u}} - \mathbf{d}) + \mathbf{S}^\dagger \bar{\kappa} = 0. \tag{5}$$

By choosing the Lagrange multiplier $\bar{\kappa}$, the derivative of the objective function is reduced from equation 4 to

$$\frac{d\phi}{d\mathbf{x}} = \frac{dL(\bar{\mathbf{u}}, \bar{\kappa})}{d\mathbf{x}} = \frac{\partial L(\bar{\mathbf{u}}, \bar{\kappa})}{\partial \mathbf{x}}. \tag{6}$$

It then follows from equation 3 that

$$\frac{d\phi}{d\mathbf{m}_i} = < \frac{\partial \mathbf{S}}{\partial \mathbf{m}_i}\bar{\mathbf{u}}, \bar{\kappa} >, \tag{7}$$

where m_i is the ith element of $\mathbf{m}$,

$$\frac{d\phi}{d\mathbf{f}_{R_i}} = -\Re(\bar{\kappa_i}), \tag{8}$$

and

$$\frac{d\phi}{d\mathbf{f}_{I_i}} = \Im(\bar{\kappa_i}). \tag{9}$$

From these derivatives, it is simple to calculate the derivative of the objective function with respect to any set of variables determining $\mathbf{m}$ and $\mathbf{f}$. A set of variables more restrictive than arbitrary $\mathbf{f}_R$ and $\mathbf{f}_I$ is likely necessary due to the large dimensionality of these variables: each has a number of elements equal to the number of points in the wavefield grid used in forward modeling multiplied by the number of sources considered. Consider an inversion variable $\mathbf{f}_x$, that controls some aspect of the structure of $\mathbf{f}$. The variable $\mathbf{f}_x$ might control the position or moment tensor of a point source, for instance. The derivative of the objective function with respect to such a variable is a simple extension of the derivative with respect to $\mathbf{f}$:

$$\frac{d\phi}{d\mathbf{f}_x} = \sum_i \frac{d\phi}{d\mathbf{f}_{R_i}}\frac{d\mathbf{f}_{R_i}}{d\mathbf{f}_x} + \frac{d\phi}{d\mathbf{f}_{I_i}}\frac{d\mathbf{f}_{I_i}}{d\mathbf{f}_x} = \sum_i -\Re(\bar{\kappa_i})\frac{d\mathbf{f}_{R_i}}{d\mathbf{f}_x} + \Im(\bar{\kappa_i})\frac{d\mathbf{f}_{I_i}}{d\mathbf{f}_x}. \tag{10}$$

Hessian-vector product

A key element in several numerical optimization strategies is the calculation of Hessian-vector products. Here, we will be considering this product while employing the Gauss-Newton approximation of the Hessian, so the residual-dependent part of the Hessian is neglected. We refer to the TN method when using the Gauss-Newton approximation as the truncated Gauss Newton (TGN) method.

The main term of the objective function in equation 2 is of the form $\phi = \frac{1}{2}\gamma^2$. The second derivative of an objective function like this with respect to inversion variable $\mathbf{x}$ is

$$\frac{\partial^2 \phi}{\partial \mathbf{x}^2} = \frac{\partial^2 \gamma}{\partial \mathbf{x}^2}^T \gamma + \frac{\partial \gamma}{\partial \mathbf{x}}^T \frac{\partial \gamma}{\partial \mathbf{x}} \tag{11}$$

Under the Gauss-Newton approximation, we neglect the first term on the right-hand side in equation 11 due to the assumption that γ is small.

For the FWI objective we define in equation 2, the Gauss-Newton Hessian is given by the relation

$$J^{\dagger}\mathbf{R}^{T}\mathbf{R}J, \tag{12}$$

where J is the Jacobian matrix $\frac{d\mathbf{u}}{d\mathbf{x}}$. The matrix J is too costly to directly calculate in FWI, but we can use the adjoint state method to avoid the need for such a calculation. We can begin by noting that the derivative of the function

$$\mathbf{h} =< \mathbf{u}(\mathbf{x}), \mathbf{w} >, \tag{13}$$

where **w** is an arbitrary vector, with respect to **x** is

$$\nabla\mathbf{h} = J^{\dagger}\mathbf{w}. \tag{14}$$

Consequently, if **w** is chosen to be $\mathbf{R}^{T}\mathbf{R}J\mathbf{v}$, then the Hessian-vector product $H_{GN}\mathbf{v}$ is equal to the derivative of **h**. This derivative can be calculated in exactly the same way as the gradient was by considering the Lagrangian

$$L =< \mathbf{u}(\mathbf{x}), \mathbf{w} > + < \mathbf{S}(\mathbf{x})\mathbf{u} - \mathbf{f}, \xi > \tag{15}$$

instead of equation 3. The same procedure follows, but instead of requiring Lagrange multiplier $\bar{\kappa}$ to satisfy equation 5, the removal of $\frac{\partial\bar{\mathbf{u}}}{\partial\mathbf{x}}$ requires that the Lagrange multiplier for this problem, ξ, satisfies

$$\mathbf{S}^{\dagger}\bar{\xi} = -\mathbf{w}. \tag{16}$$

As before, if the Lagrange multiplier satisfies this condition, then

$$\frac{dL}{d\mathbf{m}_i} =< \frac{\partial\mathbf{S}}{\partial\mathbf{m}_i}\bar{\mathbf{u}}, \bar{\kappa} >, \tag{17}$$

$$\frac{dL}{d\mathbf{f}_{R_i}} = -\Re(\bar{\kappa_i}), \tag{18}$$

and

$$\frac{dL}{d\mathbf{f}_{I_i}} = \Im(\bar{\kappa_i}). \tag{19}$$

The calculation of $\bar{\xi}$ does require that **w** is known, however, and we cannot directly calculate J. We can calculate the product of J with the vector **v** through consideration of the derivative of the forward problem with respect to variables x_i multiplied by vector elements v_i. Using the relation $\mathbf{S}\mathbf{u} - \mathbf{f} = 0$, it can be shown that

$$\frac{\partial(\mathbf{S}\mathbf{u} - \mathbf{f})v_{m_i}}{\partial m_i} = \mathbf{S}\left(\frac{\partial\mathbf{u}}{\partial m_i}v_{m_i}\right) + \mathbf{u}\left(\frac{\partial\mathbf{S}}{\partial m_i}v_{m_i}\right) = 0, \tag{20}$$

$$\frac{\partial(\mathbf{S}\mathbf{u} - \mathbf{f})v_{fR_i}}{\partial f_{R_i}} = \mathbf{S}\left(\frac{\partial\mathbf{u}}{\partial f_{R_i}}v_{fR_i}\right) - \Re(v_{fR_i}) = 0, \tag{21}$$

$$\frac{\partial(\mathbf{S}\mathbf{u} - \mathbf{f})v_{fI_i}}{\partial f_{I_i}} = \mathbf{S}\left(\frac{\partial\mathbf{u}}{\partial f_{I_i}}v_{fI_i}\right) - i\Im(v_{fI_i}) = 0, \tag{22}$$

where v_{m_i} is the element of $\mathbf{v}$ corresponding to the ith element of $\mathbf{m}$, v_{fR_i} is the element corresponding to the ith element of $\mathbf{f}_R$ and v_{fI_i} is the element corresponding to the ith element of $\mathbf{f}_I$. A sum over i for the term $S\left(\frac{\partial \mathbf{u}}{\partial x_i} v_i\right)$ is equal to the product $S(J\mathbf{v})$, so equations 20-22 are equivalent to

$$\mathbf{S}(J_{\mathbf{m}\mathbf{v}_m}) = -\mathbf{u} \sum \left(\frac{\partial \mathbf{S}}{\partial m_i} v_{m_i} \right), \tag{23}$$

$$\mathbf{S}(J_{\mathbf{f}_R}\mathbf{v}_{f_R}) = \Re(\mathbf{v}_{fR_i}), \tag{24}$$

and

$$\mathbf{S}(J_{\mathbf{f}_R}\mathbf{v}_{f_I}) = \Im(\mathbf{v}_{fI_i}). \tag{25}$$

These equations can be solved for the product $J\mathbf{v}$. Using this term to define $\mathbf{w}$ as $\mathbf{R}^T\mathbf{R}J\mathbf{v}$, we can solve equation 16 for ξ. With this choice for ξ, the derivatives $\frac{dL}{d\mathbf{x}}$ from equations 17-19 become the Gauss-Newton Hessian vector product $H_{GN}\mathbf{v}$.

Forward modeling

The anelastic finite-difference approach we use for modeling seismic waves here is described in more detail in Keating (2020). The main development necessary for the treatment of source-model simultaneous inversion is the treatment of a more general class of seismic sources. In particular, we are concerned here with recovery of the locations and moment tensor components of point-sources in the two dimensional problem. For a point source, the m, n component of the moment tensor represents a derivative of the displacement in the m direction with respect to the n direction. In the finite-difference approach we use here, the seismic wavefield is defined in terms of vertical and horizontal displacement, so simulating a point-source simply requires that the correct derivatives be introduced in the source term. For accurate representation of a point source, it is essential that a source be represented using only a very small region of the finite-difference model. Accordingly, we will employ a first-order finite-difference approach. In one dimension, the finite difference approximation of $M_{zz} = 1$ at location $z = (n + \frac{1}{2})\Delta z$ (a location midway between two finite-difference grid cell centers) would be $\frac{-1}{\Delta z}$ at $z = n\Delta z$, and $\frac{1}{\Delta z}$ at $z = (n+1)\Delta z$. This is illustrated in figure 1. For a source not located equidistant between finite-difference cell centers, it will no longer be possible to accurately represent the derivative with just two grid cells. More generally, three grid cells will be needed to approximate a derivative in the z direction for a first order, centered finite difference approach at an arbitrary location. The finite-difference weights used for a source location between two finite difference grid lines will be a weighted average of the weights which would be used for a source at either of the bounding grid lines, as shown in figure 2. Extending this concept to two dimensions, we can see that sources will generally be represented by nine cell (three by three) finite difference terms, representing an average between the four cell finite difference approximations at each of the four grid cell intersections bounding the source location. An example of a 2D source term is shown in figure 3. Similarly, in three dimensions, 27 cell finite difference terms (three by three by three) represent averages of the eight cell finite difference approximations at each of the eight grid cell intersections bounding the source location. This finite-difference approximation of 2D point-source moment tensors is well suited to

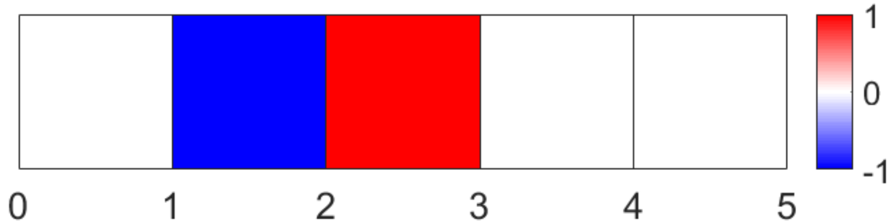

FIG. 1. Scaled weights for approximating a derivative in one dimension, for source location $z = 2$.

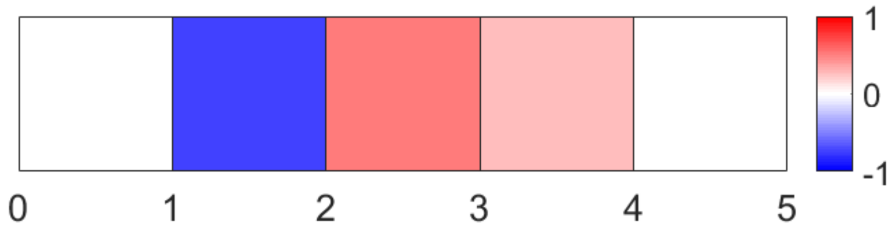

FIG. 2. Scaled weights for approximating a derivative in one dimension, for source location $z = 2.25$.

local optimization, as the gradient of the objective function with respect to source location is well defined in this case: small changes in source location simply correspond to a small re-weighting of the source term values.

Optimization

For inversion procedures designed to recover both source properties and subsurface structure, two broad strategies exist. Firstly, and most conventionally, sequential inversion can be used, wherein either structure or sources are first estimated while holding the other fixed, and then the other property is estimated while holding the first fixed. A sequential approach can often provide good estimates of both source properties and structure, but relies heavily on the assumption that these properties relate to the data approximately independently of one another. In reality, the quality of our estimates of each of these properties strongly affects our ability to estimate the other, so a sequential inversion relies on the presence of sufficiently accurate estimates of at least one of these properties. If accurate estimates are difficult or impossible to obtain, for instance, in cases where source locations are unknown and very complex media are present, it may be more appropriate to adopt the second of these approaches: simultaneous inversion. In simultaneous inversion, both properties are recovered at the same time in the inversion. This allows for the inversion to account for how a source estimate may change when certain features are introduced in the subsurface model, for instance. In this report, we are most interested in the case where simultaneous inversion is appropriate.

Inversion for source terms only can often use powerful global optimization strategies due to the relatively small number of variables needed for source description. Inversions which recover potentially complex subsurface models are not amenable to this type of approach, due to the very large number of variables needed in this problem. Instead, local optimization strategies, which use derivative based approaches to minimize an objective function, are used for this type of problem. While this can allow for a more efficient

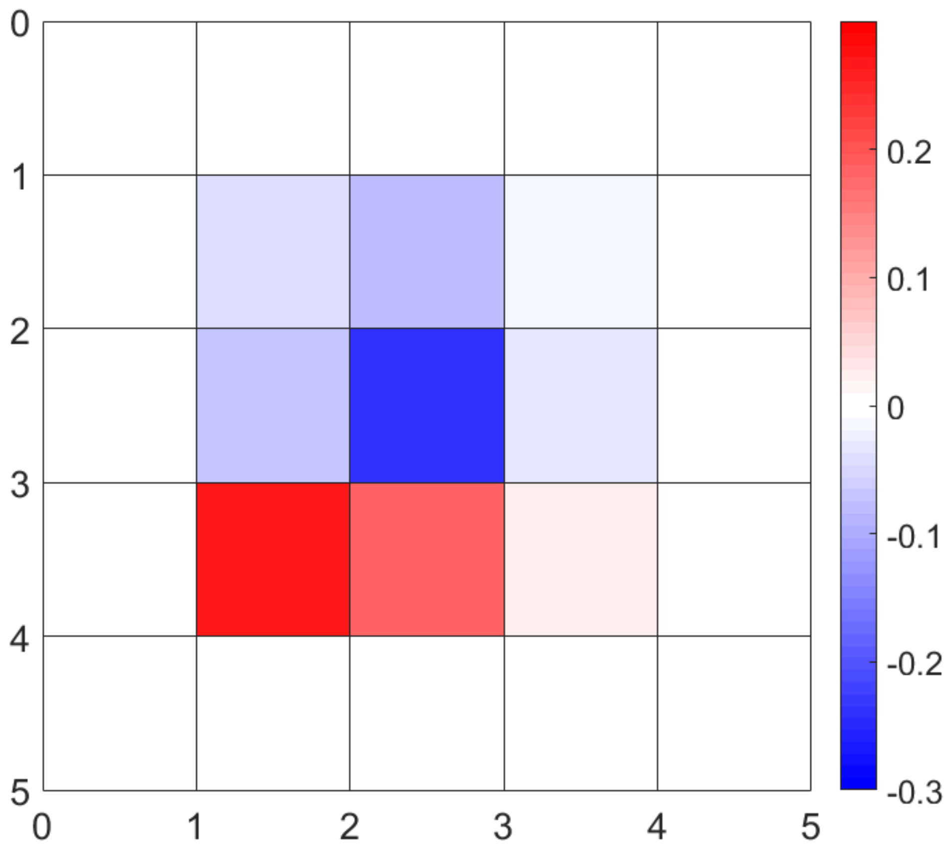

FIG. 3. Scaled weights for approximating an example 2D moment tensor.

inversion, it also introduces susceptibility to inversion results representing local minima, rather than accurate solutions to the inversion problem. A simultaneous inversion for both sources and subsurface models must use variables to represent both, so, as in the case of inversion for subsurface models, local optimization approaches must be used. Here, we will use the truncated Gauss-Newton approach, which uses an approximation of the Newton descent direction. This type of approach is described in detail by Métivier et al. (2013).

Parameterization

The variables $\mathbf{m}_i$ which define the subsurface model will each typically define both 1) a physical property or set of properties affecting wave propagation which they define and 2) a space-dependent function describing the relation of the variable with the spatial values of these properties. In this report, we consider viscoelastic inversion, and so the variables we use will define P- and S- wave velocities v_P and v_S, P- and S- wave quality factors Q_P and Q_S, and density ρ. Specifically, the physical properties defined by $\mathbf{m}_i$ will be v_P^{-2}, v_S^{-2}, Q_P^{-1}, Q_S^{-1}, and ρ. These variables have the advantage of relating to the viscoelastic wave equation in relatively simple ways (Pratt, 1990). Spatially, the most common approach in FWI is to let each variable define a scaled version of one of the physical properties at a single finite-difference grid cell (effectively acting as a Kronecker delta function). This parameterization could be problematic in simultaneous source-model inversion, because it allows for the subsurface model to change on scales smaller than the finite-difference footprint of the source terms, effectively allowing for the model to change the moment tensor or source location. This type of inter-parameter confusion is chiefly an artifact of the finite difference modeling approach, but may seriously harm inversion results (as model updates may be dominated by spurious source modifying features). To avoid this type of issue, we instead define the subsurface model in terms of Gaussian regions centered at each finite difference point. The different spatial parameterizations are illustrated in figure 4. In the Gaussian-type parameterization, the value of the parameter centered at one FD grid cell has a strong influence on the model near that location and a smaller influence on the model at locations some distance away. We define the characteristic width of these Gaussian regions to be larger than the three by three footprint of the source terms in order to remove the ability of model updates to change effective source terms. This should improve the convergence of the inversion, while somewhat reducing the possible resolution of the inversion. This type of approach is discussed in detail in Keating and Innanen (2018).

Numerical examples

In order to test our simultaneous source-model inversion, we consider a synthetic test in this section. The subsurface model and sources to be inverted are shown in figures 5 and 6. The subsurface model is based on a section of the Marmousi model, while the sources are randomly distributed with arbitrary moment tensors. The starting model used for the inversion is based on the assumption that there is little prior information available; the initial model of anelastic properties is constant in each parameter, the initial source location estimates are randomly placed within 70 m of their true location (based on the assumption that such a coarse estimate could be made even based on a very poor velocity

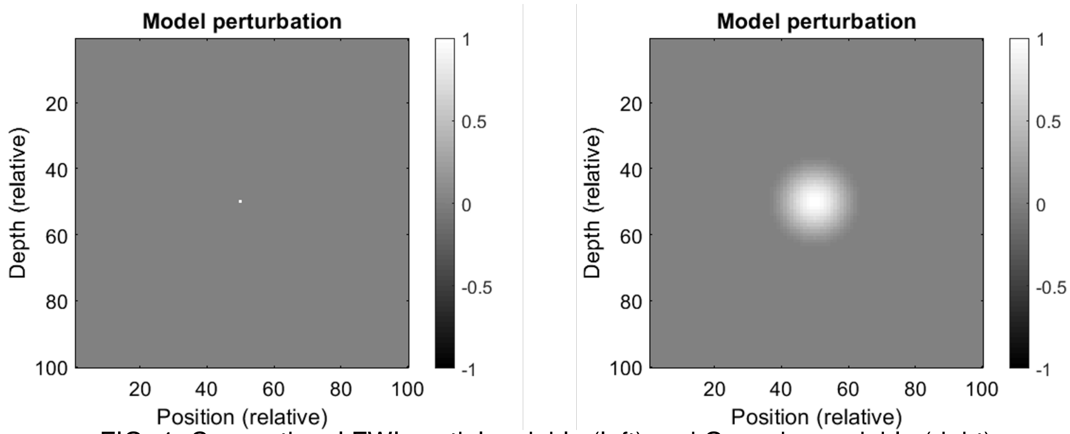

FIG. 4. Conventional FWI spatial variable (left) and Gaussian variable (right).

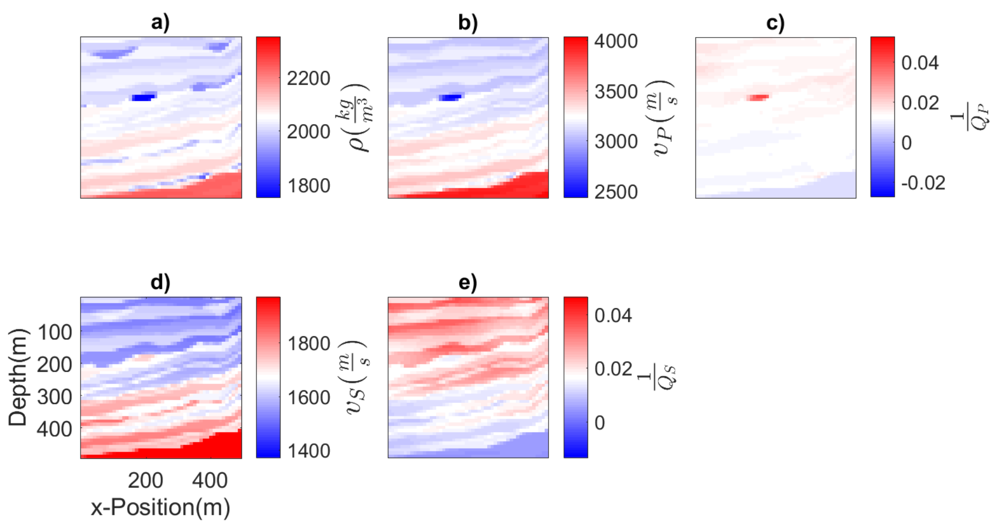

FIG. 5. True physical properties of the synthetic model.

model), and the initial source moment tensor terms are random. The starting model is shown in figure 7, and the initial sources are compared with the true sources in figure 6. We discretize the model into 10m by 10m grid cells for our finite difference modeling.

In our inversion, we use a multiscale approach with 20 total frequency bands each of six frequencies. Each band has data starting at 1 Hz, and includes five other frequencies, linearly spaced to the maximum of each band, starting a 2 Hz in the first band, and linearly increasing to 20 Hz by the last band. At each frequency band, two iterations of truncated Gauss-Newton optimization are used, with twenty inner iterations per FWI iteration. The source estimates after the first five frequency bands are shown in figure 8, while those recovered at the end of the inversion process are shown in figure 9. The corresponding subsurface models are shown in figures 10 and 11. By examining the results after five frequency bands (maximum frequency of about 6 Hz), we can see that there is very good recovery of moment tensors and source locations (figure 8), but only a very crude estimate of the subsurface model (figure 10). This suggests that the early stages of the inversion are dominated by improvements in source estimation rather than model updating. In the final inversion result, there is little improvement in the source estimation (figure 9), as the estimate was already quite good after five frequency bands, but the subsurface model improves

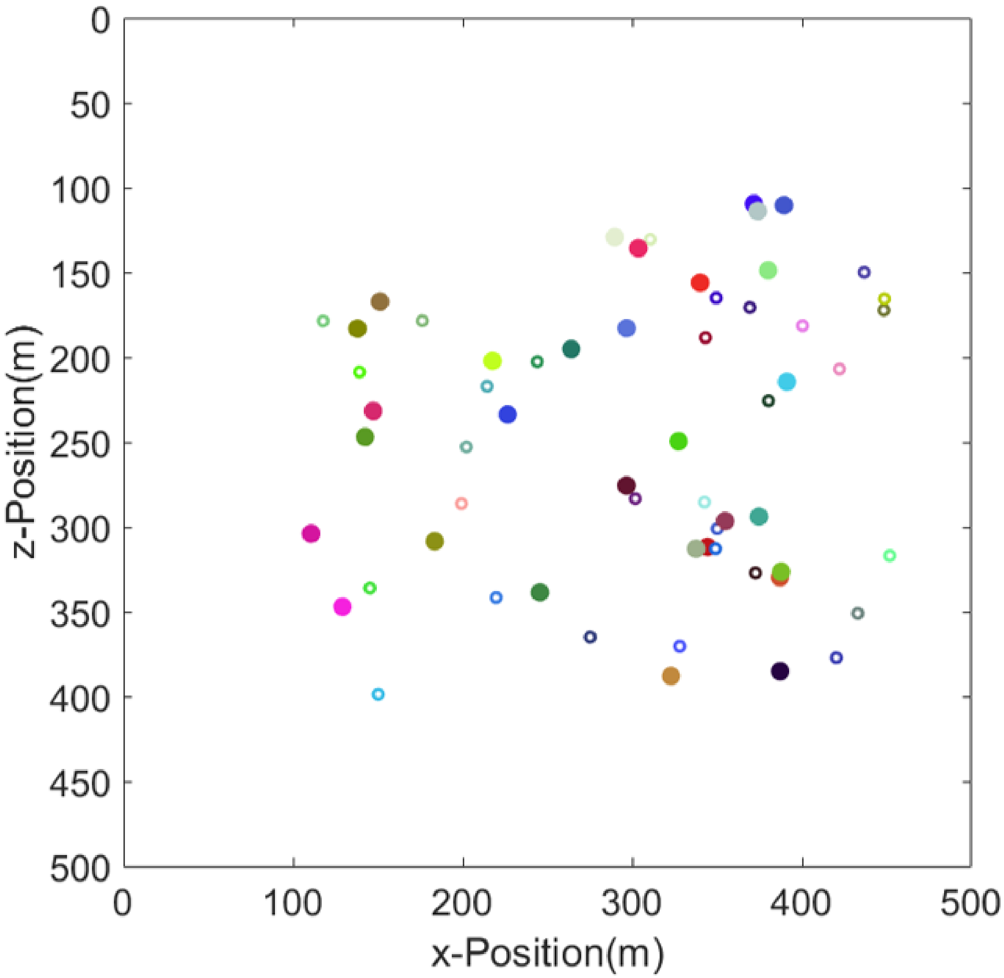

FIG. 6. True sources (dots with white centers) and initial source estimates (solid dots). The colors represent the 2D moment tensor, with the M11 value represented by red (0.5 red is 0 M11), M12 represented by green, and M22 represented by blue.

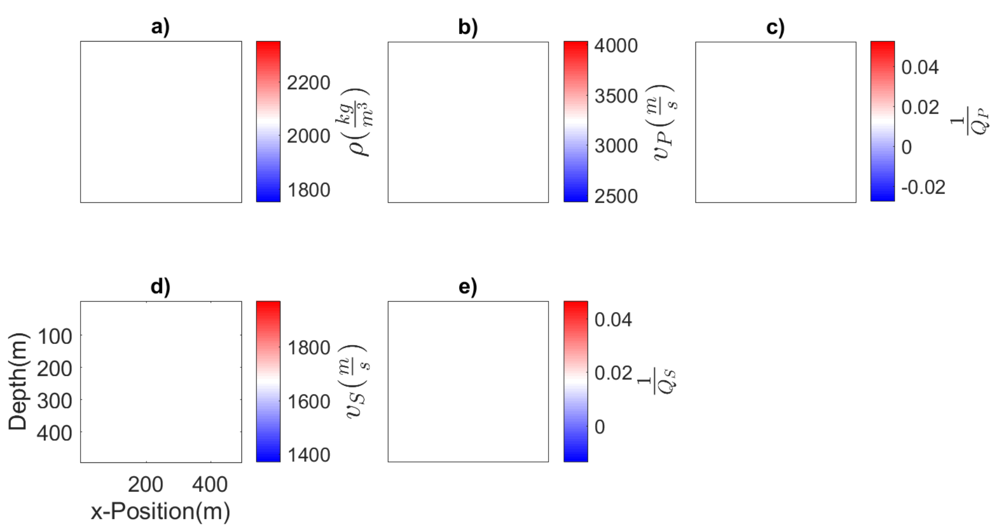

FIG. 7. Initial estimate of physical properties.

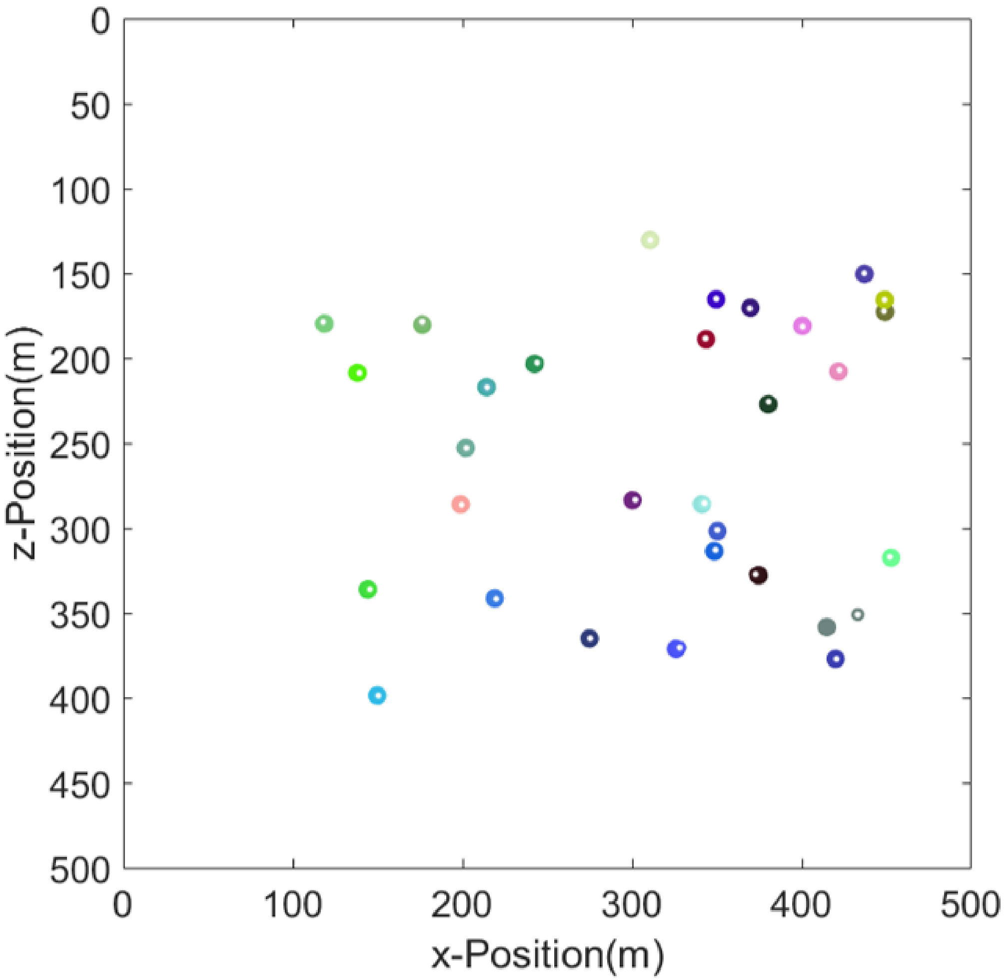

FIG. 8. True sources (dots with white centers) and source estimates after five frequency bands (solid dots). The colors represent the 2D moment tensor, with the M11 value represented by red (0.5 red is 0 M11), M12 represented by green, and M22 represented by blue.

considerably (figure 11). Both sets of properties have reasonably accurate estimates at the end of the inversion procedure.

DISCUSSION

The numerical examples we present here show that, at least in a small, synthetic example, a source-model simultaneous inversion is feasible. Interestingly, however, the behaviour exhibited by the simultaneous approach almost seems to suggest a naturally sequential strategy. In the early iterations of the inversion, the source estimate improves considerably, with moment tensors and locations estimated very accurately. The subsurface model in the same early iterations establishes only very approximately the long-wavelength structure. Conversely, at later iterations, there is little improvement in the source estimate, while the subsurface model improves substantially. Collectively, these results suggest that our inversion, while having the capability to update both sources and model simultaneously, instead begins with a heavy focus on source estimation, and primarily updates the subsurface model afterwards. Moreover, in this case, the source estimate is quite accurate

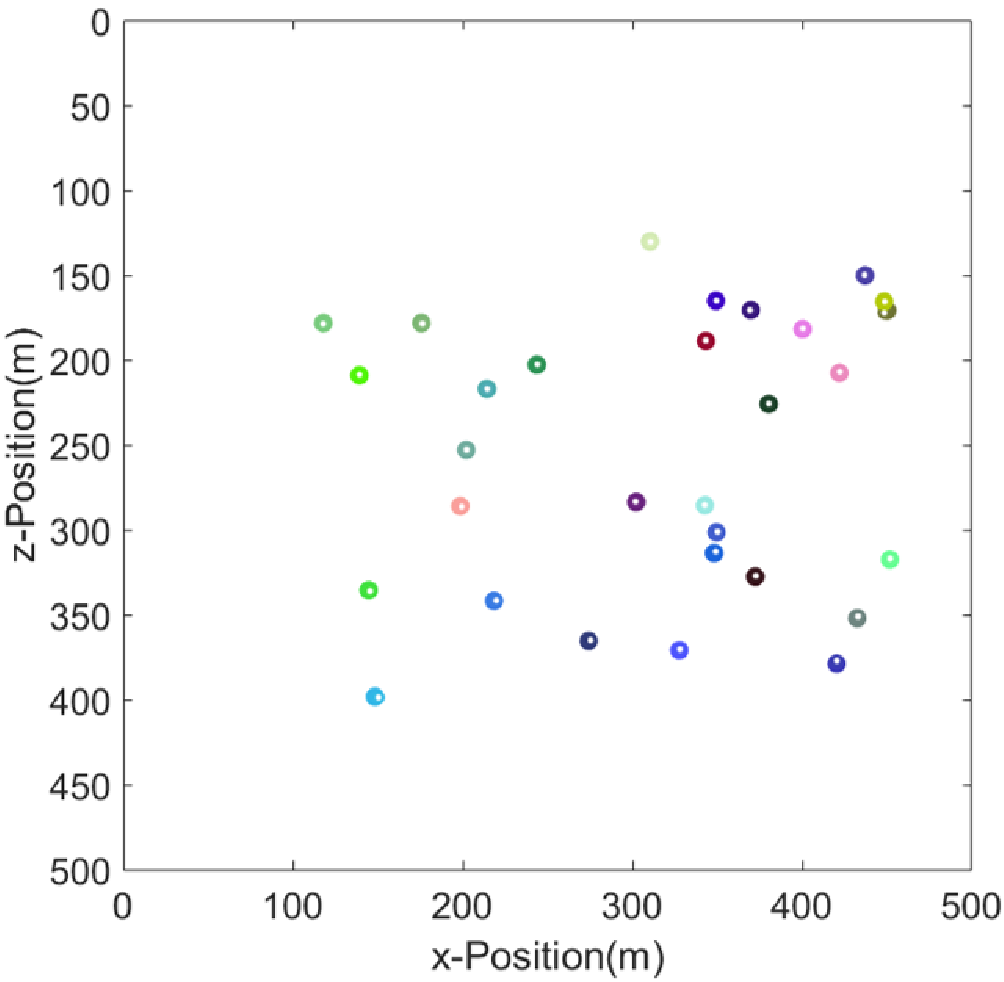

FIG. 9. True sources (dots with white centers) and final source estimates (solid dots). The colors represent the 2D moment tensor, with the M11 value represented by red (0.5 red is 0 M11), M12 represented by green, and M22 represented by blue.

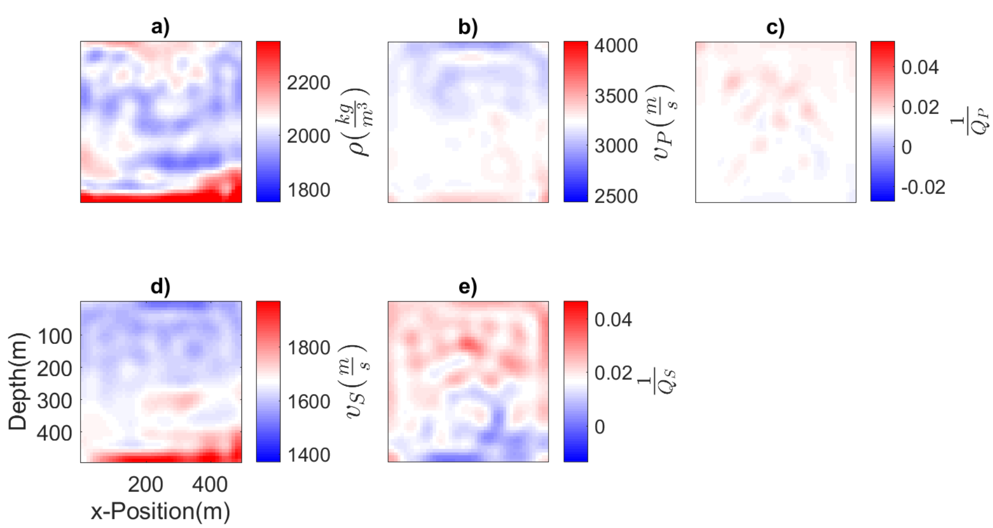

FIG. 10. Estimate of physical properties after five frequency bands.

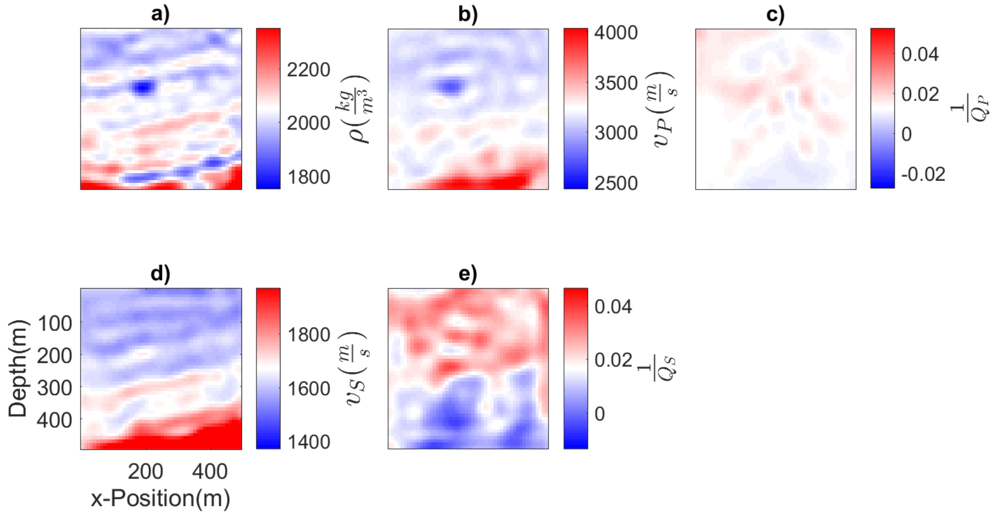

FIG. 11. Final estimate of physical properties.

even when the subsurface model is still quite poor, suggesting that the nearly sequential approach is essentially effective, even for the relatively complex model we consider. This raises the question of how complex a subsurface model or how poor an initial estimate of it must be before simultaneous inversion offers substantial improvement. This is an important question for further study.

In this report, a very simplified treatment of frequencies has been used, where the frequency spectrum of each source is assumed to be constant, known, and to extend into the bandwidth of frequencies typically used in FWI. Inversion for source frequency spectra could add new complexities to the problem, particularly those of defining an appropriate parameterization. A full frequency spectrum for each source requires many more degrees of freedom than can likely be constrained in the inversion, so a compact but effective parameterization would need to be defined. Another challenge facing this approach is the likely frequency spectrum of seismic sources: in microseismic problems, for instance, the dominant frequencies of unknown sources may be significantly higher than the typical FWI frequency band. In this case, a simultaneous approach may not be appropriate, given the very high cost of high frequency FWI.

CONCLUSIONS

In this report, we have investigated the theory for a simultaneous source-model full waveform inversion. Notably, the source and model variable derivatives make use of the same wave propagation solutions, so the main per-iteration computational burden of the sequential inversion approach is not significantly changed in the simultaneous approach. The frequency domain objective function derivative expressions we derive are general, but in a numerical example, we focus on a viscoelastic problem, with point sources in two dimensions.

ACKNOWLEDGEMENTS

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 grants CRDPJ 461179-13 and CRDPJ 543578-19. Scott Keating was also supported by the Earl D. and Reba C. Griffin Memorial Scholarship.

REFERENCES

Alvizuri, C., and Tape, C., 2018, Full moment tensor analysis of nuclear explosions in north korea: Seismological Research Letters, **89**, No. 6, 2139–2151.

Keating, S., 2020, Viscoelastic full-waveform inversion: treating attenuation uncertainty, characterizing cross-talk, and quantifying confidence in inversion results: Ph.D. thesis, University of Calgary.

Keating, S., and Innanen, K. A., 2018, Using multi-resolution truncated newton optimization for cross-talk reduction in fwi: CREWES Annual Report, **30**.

Maxwell, S. C., Rutledge, J., Jones, R., and Fehler, M., 2010, Petroleum reservoir characterization using downhole microseismic monitoring: Geophysics, **75**, No. 5, no. 5, 75A129–75A137.

Métivier, L., Brossier, R., Virieux, J., and Operto, S., 2013, Full waveform inversion and the truncated Newton method: SIAM Journal on Scientific Computing, **35**, No. 2, no. 2, B401–B437.

Nolet, G., 1987, Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics, vol. 5: Springer Science & Business Media.

Pratt, R. G., 1990, Frequency-domain elastic wave modeling by finite differences; a tool for crosshole seismic imaging: Geophysics, **55**, No. 5, 626–632.

Smith, G., and Gidlow, P., 1987, Weighted stacking for rock property estimation and detection of gas: Geophysical prospecting, **35**, No. 9, 993–1014.

Stich, D., Martín, R., and Morales, J., 2010, Moment tensor inversion for iberia–maghreb earthquakes 2005–2008: Tectonophysics, **483**, No. 3-4, 390–398.

Stolt, R. H., 1978, Migration by fourier transform: Geophysics, **43**, No. 1, 23–48.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, No. 6, no. 6, WCC1–WCC26.