# Viscoelastic FWI: solving for $\mathrm{Q_P}$, $\mathrm{Q_S}$, $\mathrm{V_P}$, $\mathrm{V_S}$, and density

Scott Keating and Kris Innanen


## ABSTRACT

Useful information about seismic amplitudes is often neglected or underused in full waveform inversion due to the common neglect of elastic or attenuative physics considered in the inversion. To make better use of measured data, we present a frequency domain, viscoelastic full waveform inversion. Inter-parameter cross-talk is demonstrated to be a major concern in this problem, not easily prevented through comprehensive acquisition geometry or relatively intensive numerical optimization. We propose a strategy for cross-talk mitigation based on prioritizing the transmissive effects of the $Q$ variables.


## INTRODUCTION

In concept, full waveform inversion (FWI) attempts to find the subsurface model best describing the full information content of a seismic experiment (Tarantola, 1984). While this ambitious goal is far from being realized, FWI approaches have proved to be effective tools for the recovery of intermediate scale P-wave velocity models. The FWI strategy in this context, while highly successful, relies heavily on a constant-density, acoustic model of wave propagation. This means that inversion results are mostly restricted to explaining the phase information of P-wave arrivals contained in measured data, falling well short of the conceptual promise of the technique. Significant efforts are being made to make more complete formulations of FWI more practical.

Important sources of information in seismic data are neglected in acoustic FWI approaches, significantly amplitudes. Amplitude information has been employed with great success in exploration geophysics in the analysis of amplitude variation with offset (AVO) to better understand the elastic properties of the subsurface. In order to reproduce accurate AVO effects it is necessary to consider an elastic model of the subsurface. While elastic formulations of FWI were proposed shortly after the original acoustic algorithm (Tarantola, 1986), most progress which has been made in FWI has centered on the constant density acoustic case. This is partly due to the considerable computational cost of modeling elastic wave propagation, but also to the difficulty of correctly treating amplitude information in FWI Virieux and Operto (2009). Seismic waves are considerably attenuated in the subsurface, and if this is not considered in the inversion, FWI may still struggle to correctly interpret amplitude information.

Attenuation plays a major role in seismic wave propagation, but is often neglected in elastic FWI formulations. While the treatment of attenuation as characterized by $Q$ in FWI has been studied, most research has been done in a viscoacoustic setting (e.g., Hicks and Pratt, 2001; Malinowski et al., 2011; Kamei and Pratt, 2013; Métivier et al., 2015; Plessix et al., 2016; Keating and Innanen, 2016b, 2017b). These approaches have shown success, but suffer from the same neglect of elastic information as the constant density acoustic algorithm. The viscoelastic FWI problem is investigated less frequently, but has a greater potential to accurately reproduce seismic wave propagation.

There are many challenges associated with moving from constant density acoustic FWI to viscoelastic FWI. One major concern is the very large number of variables in the inversion. Recovering five physical properties instead of one means that the number of variables in a conventional viscoelastic FWI is five times greater. The seismic inverse problem is already ill-posed and this expansion requires that limited data resolve an even larger number of variables. The much slower speeds of S-waves mean that much finer finite difference or finite element grids need to be considered, increasing the cost of wavefield modeling. Inter-parameter cross-talk, where different physical properties are confused with one another in the inversion, is another serious concern. In this report, we discuss viso-elastic full waveform inversion, with a focus on inter-parameter cross-talk.

## THEORY

### Forward modeling

Numerical simulations of wave propagation are key in full waveform inversion. A frequency domain finite difference solution to the visco-elastic wave equation is calculated here by solving the system described by (Pratt, 1990):

$$\omega^2 \rho u_x + \frac{\partial}{\partial x}\left[\tilde{\lambda}\left(\frac{\partial u_x}{\partial x}+\frac{\partial u_z}{\partial z}\right)+2\tilde{\mu}\frac{\partial u_x}{\partial x}\right]+\frac{\partial}{\partial z}\left[\tilde{\mu}\left(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z}\right)\right]+f=0 \tag{1}$$

and

$$\omega^2 \rho u_z + \frac{\partial}{\partial z}\left[\tilde{\lambda}\left(\frac{\partial u_x}{\partial x}+\frac{\partial u_z}{\partial z}\right)+2\tilde{\mu}\frac{\partial u_z}{\partial z}\right]+\frac{\partial}{\partial x}\left[\tilde{\mu}\left(\frac{\partial u_z}{\partial x}+\frac{\partial u_x}{\partial z}\right)\right]+g=0, \tag{2}$$

where $\omega$ is the angular frequency, $\rho$ is density, $u_x$ and $u_z$ are, respectively, the horizontal and vertical displacements, $f$ and $g$ are the corresponding source terms, and $\tilde{\lambda}$ and $\tilde{\mu}$ are the complex, frequency dependent Lamé parameters. Assuming a Kolsky-Futterman model of attenuation (Kolsky, 1956; Futterman, 1962), these are defined in terms of $\rho$, the P and S wave speeds, $v_P$ and $v_S$, and Q values $Q_P$ and $Q_S$ by

$$\tilde{\mu}=\left\{v_S\left[1+\frac{1}{Q_S}\left(\frac{1}{\pi}\log\frac{\omega}{\omega_0}+\frac{i}{2}\right)\right]\right\}^2\rho \tag{3}$$

and

$$\tilde{\lambda}=\left\{v_P\left[1+\frac{1}{Q_P}\left(\frac{1}{\pi}\log\frac{\omega}{\omega_0}+\frac{i}{2}\right)\right]\right\}^2\rho-2\tilde{\mu}, \tag{4}$$

where $\omega_0$ is a reference frequency. This system was solved using the finite difference equations described in Pratt (1990), which takes the form

$$Su=h, \tag{5}$$

where $u$ is a vector containing $u_x$ and $u_z$, $h$ is a vector containing $f$ and $g$, and $S$ is a Helmholtz matrix containing the finite difference coefficients. The matrix $S$ applies a finite difference star to $u$ approximating the wave equation in equations 1 and 2. Perfectly matched layers (Berenger, 1994) were used to prevent reflections from the boundaries of the model.

## Optimization

The objective function for the FWI problem discussed here is given by

$$\phi = \sum_{x_s,\omega} \frac{1}{2}||D - Ru(m)||_2^2, \tag{6}$$

where $D$ represents the measured data, $R$ is a matrix representing the receiver sampling of the wavefield, $x_s$ represents the source locations, $\omega$ is angular frequency, and $m$ is the estimated model of the subsurface. The gradient of this function with respect to $m$ was calculated by Tarantola (1984). It is expressed in a simple, general form by Metivier et al. (2013) as

$$\frac{\partial \phi}{\partial m} = \langle \frac{\partial S}{\partial m} u, \lambda \rangle, \tag{7}$$

where $\langle , \rangle$ represents an inner product, and

$$S^\dagger \lambda = R^T (Ru - D). \tag{8}$$

Because the matrix $\frac{\partial S}{\partial m}$ is very sparse, this derivative can be inexpensively calculated for a large number of variables, the chief cost being the calculation of $u$ and $\lambda$. The model variables $m$ considered in this report are $\frac{1}{v_P^2}$, $\frac{1}{Q_P}$, $\frac{1}{v_S^2}$, $\frac{1}{Q_S}$, and $\rho$. These properties assume very different values, and this causes the amplitudes of their respective gradients to differ widely. The gradients with respect to each model variable were scaled according to the amplitude of the initial model here.

The optimization strategy used here is the truncated Gauss Newton (TGN) method. This approach approximates the Gauss Newton update, given by

$$H_{GN} \Delta m = -g \quad , \tag{9}$$

where $\Delta m$ is the update, $g$ is the gradient of the objective function with respect to the model parameters, and $H_{GN}$ is the residual independent part of the Hessian, or second derivative of the objective function with respect to the model parameters. The TGN method approximates the Gauss-Newton step by attempting to solve equation 9 through an iterative numerical optimization. The TGN approach as used here is described in detail in Keating and Innanen (2016a), and is based on the work of Metivier et al. (2013). The approximation is calculated iteratively, in an 'inner loop' of the problem, performed at each FWI iteration. The accuracy and cost of the approximation increase with the number of inner loop iterations performed. The inner loop is ended when a maximum number of iterations are reached, or the inequality

$$|H\Delta m + g|_2 \leq \eta |g|_2 \tag{10}$$

is satisfied, where $\eta$ is a chosen forcing term.

## Cross-talk

Cross-talk is the phenomenon in which the data signatures of different physical properties are confused in FWI. The challenges associated with cross-talk are a major obstacle

to the effective implementation of multiparameter FWI. Cross-talk in FWI is generally introduced by either insufficient data in the inversion, or by an inadequate numerical optimization approach. Seismic surveys generally collect sufficient information to enable discrimination between different elastic properties in most locations. If sufficiently broad ranges of frequencies are considered in the inversion, it becomes possible to distinguish between $Q$ terms and other viscoelastic properties as well.

With sufficient data, global optimization strategies should, in theory, be able to recover a model with little or no cross-talk. Unfortunately, such strategies are unviable on the scales that FWI is typically applied. Instead, linearized, local optimization techniques are typically used. These attempt to reduce the objective function by searching in directions based on the derivatives of the objective function with respect to the model variables at the current model location. This often leads to cross-talk in cases where changes in different physical properties can all partially reduce the same data residual. The second-derivative information considered in Newton optimization helps to account for this confusion, and can substantially reduce cross-talk (e.g. Innanen, 2014). Unfortunately, this approach is also typically too expensive for use in FWI. To reduce computational costs, TGN optimization is used here instead. Unfortunately, the solution of the system in equation 9 is extremely difficult to approximate in the viscoelastic case due to the very large size of $H_{GN}$. As a conventional FWI approach will have 5 times as many variables in a viscoelastic inversion as in a single parameter inversion, the matrix $H$ will contain 25 times the number of elements. This makes it very computationally expensive to approximate the solution to equation 9. Due to this cost, elastic implementations of FWI often use other strategies in conjunction with second derivative information to reduce cross-talk.

The model parameterization used in the inversion is an important consideration when trying to minimize cross-talk. Different physical properties are sensitive to different types of data in an FWI scheme. The response of seismic waves to changes in these physical properties can be studied through the use of radiation patterns. These are measures of how small perturbations in a given parameter change the wavefield. Study of radiation patterns can reveal which variables are sensitive to which data. Figure 1 shows the radiation patterns of point perturbations of different visco-elastic parameters for frequency domain data at 15Hz. The source position is marked with a red star, and the model is perturbed at the green dot. Different variables have different radiation patterns, so while a change in P-wave velocity will alter the wavefield at all scattering angles, a change in density mostly alters the back-scattered wavefield. These differences have been used in multiparameter FWI to prevent cross-talk by choosing model parameterizations in which each variable has a highly distinct radiation pattern for the data considered. For instance, because the radiation patterns of P-wave velocity and density both have large amplitudes for most back-scattering angles, small angle reflection data may introduce severe cross-talk unless expensive optimization techniques are used. This parameterization can is undesirable for reflection-type data. If transmission type data are considered instead, little cross-talk between $v_P$ and $\rho$ will occur even with gradient based optimization approaches due to the limited impact of density on the data considered. Appropriate choice of model parameterization has been used to effectively suppress cross-talk in several multiparameter problems (e.g. Operto et al., 2013).

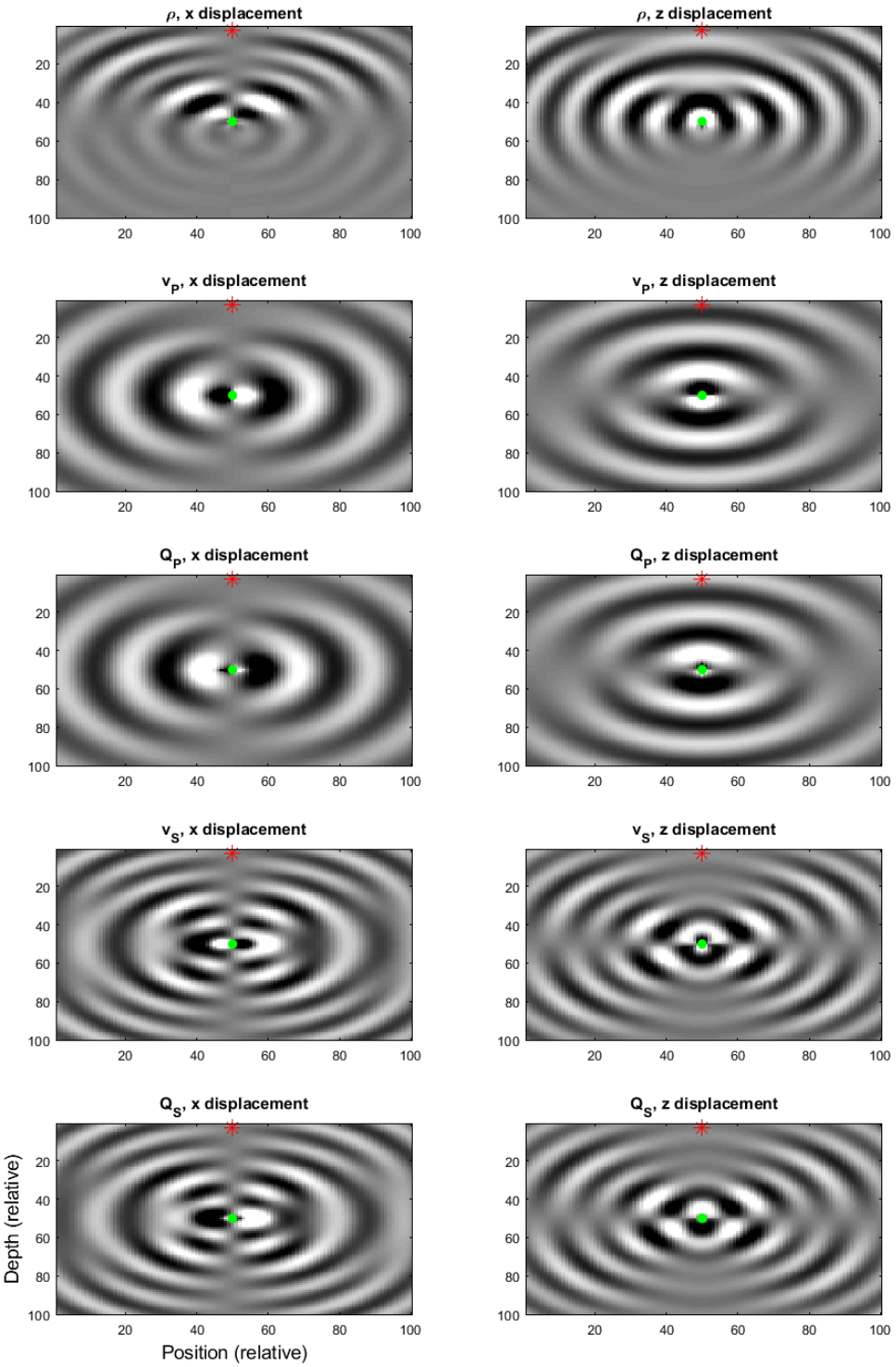


FIG. 1. X displacement (left) and Z displacement (right) of radiation patterns. Red stars indicate source locations, and green dots are the location of the perturbed variables. The variables perturbed, from top to bottom, are $\rho$, $\frac{1}{v_P^2}$, $\frac{1}{Q_P}$, $\frac{1}{v_S^2}$, and $\frac{1}{Q_S}$

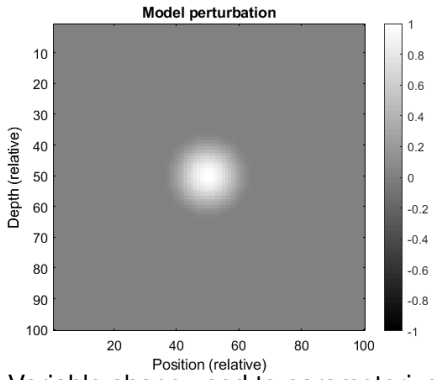


FIG. 2. Variable shape used to parameterize $Q$ model.

An unfortunate feature of the viscoelastic problem is the difficulty of using a data selection approach to prevent cross-talk between $Q_P$ and $v_P$, or $Q_S$ and $v_S$. The radiation patterns of these variables, shown in figure 1, have strong amplitudes at the same scattering angles. These radiation patterns have different frequency dependence, but at no frequency is one negligible relative to the other. In consequence, some effective discrimination techniques that work for closely related problems are not effective when trying to prevent cross-talk with $Q$.

## Strategies

The optimization strategy investigated in this report is an attempt to reduce cross-talk in the visoelastic problem. The goals of the approach are to frame an inversion problem of fewer variables, to mitigate cross-talk with $Q$ variables arising from reflection events, and to limit extreme variations in the $Q$ model. To achieve these goals, the definition of the $Q$ model was changed. In Keating and Innanen (2017a), the concept of defining variables with spatial extent in the FWI problem was explored. Here we investigate the use of long-wavelength $Q$ variables to reduce cross-talk in the inversion.

While reflections can be introduced by contrasts in $Q$ (Lines and Vasheghani, 2008), the attenuative and dispersive effects of $Q$ are more often significant in seismic exploration. Cross-talk can be introduced when reflections are mistakenly attributed to $Q$ contrasts instead of changes in elastic properties. If it is assumed that reflections from $Q$ contrasts are negligible in the measured data, it is possible choose a set of variables parameterizing the $Q$ model which do not suffer from this mode of cross talk. An example of this type of variable shown in figure 2, as are the corresponding radiation patterns in figure 3. As these variables describe a long wavelength change in the model, their capacity to introduce reflections is very small. This should help to reduce cross-talk from reflection-type features in the data. The large spatial size of these variables means that relatively few of them should be sufficient to describe the model. This means that the dimensionality of the optimization problem can be reduced, allowing for second derivative information to be calculated at lower cost.

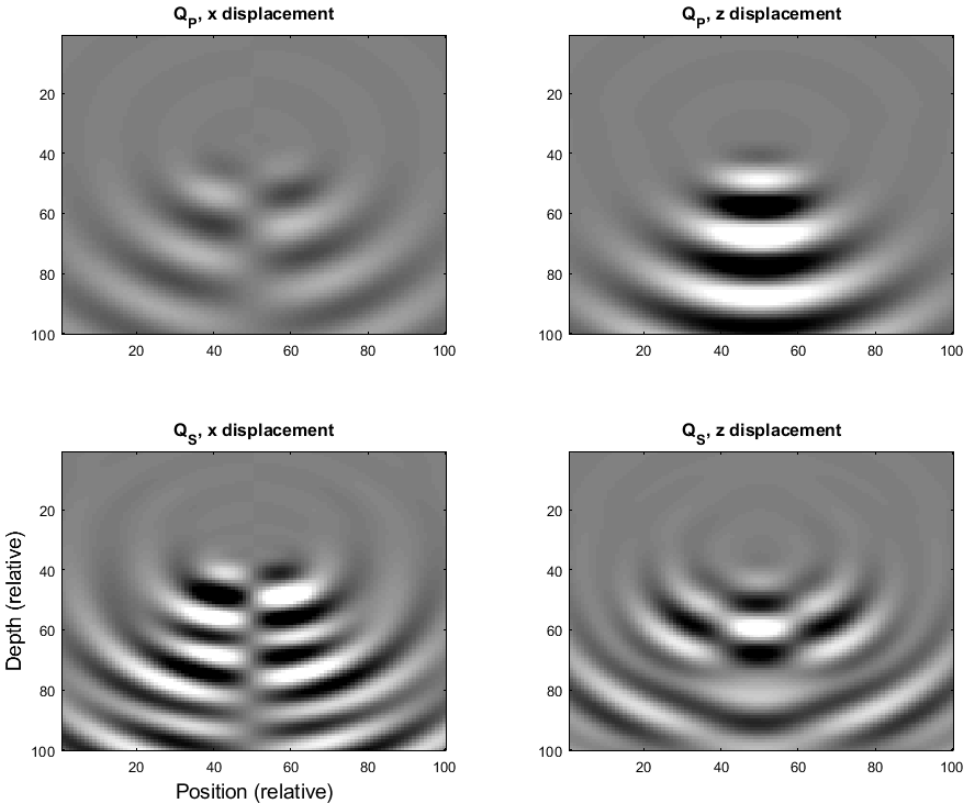


FIG. 3. X displacement (left) and Z displacement (right) of radiation patterns. The variables perturbed, from top to bottom, are $\frac{1}{Q_P}$, and $\frac{1}{Q_S}$, over a region of the form shown in figure 2

| | |
|---|---|
| FWI iterations per frequency band | 1 |
| Minimum frequency | 1 Hz |
| Max frequency for band n | (1 + n) Hz |
| Frequency bands used | 20 |
| Frequencies per band | 10 |
| Forcing term | 0.001 |

Table 1. Optimization parameters used in the numerical examples.

## NUMERICAL EXAMPLES

To investigate the extent of inter-parameter cross-talk in the anelastic FWI problem, numerical tests were performed on the model in figure 4. In these tests, the starting estimates for the inversion were constant values for each parameter. Very low frequencies (1 Hz) were simulated in the 'measured data', preventing cycle-skipping and local minima problems. Different acquisition geometries were investigated in order to assess their impact on the relative degree of cross-talk observed. Explosive sources were simulated in all tests.

In the first test, the geometry considered was meant to resemble surface seismic acquistion. 98 receivers were placed at 5 m depth, and 48 explosive sources were placed at 10 m depth. The frequencies and forcing term used in the optimization are listed in table 1. A maximum of five inner loop iterations were allowed in the truncated Newton approach, and a conventional FWI parameterization (scattering patterns in figure 1) was used. The inversion result from this test is shown in figure 5. Cross-talk can be identified in this simple model by spatially misplaced features. Considerable cross-talk exists between the density and other parameters, evident due to the considerable changes in the density model above the true density contrast. Cross-talk into the $v_S$ model is limited, while the ball-shaped anomalies in the recovered $Q$ models indicate extreme cross-talk from the corresponding velocity models.

Figure 6 shows the results of an inversion using the same model and data, but a more powerful optimization strategy. Twenty inner loop iterations were allowed in this test. The recovered model is substantially improved by the more costly optimization. In particular, the extent of cross talk from $v_S$ into $v_P$, and from all variables into density is substantially decreased. Other cross-talk remains significant. Strong changes in the velocity models below 400 m suggest that cross-talk from density into these variables is considerable. This is supported by the radiation patterns in figure 1, which are highly similar for velocity and density at reflected scattering angles. The very strong signature of the velocity anomalies in the recovered $Q$ models persists in this case, indicating that the cross talk between these parameters is also difficult to prevent. This is again supported by the very similar radiation patterns in figure 1. These results suggest that when considering reflection type information in a surface acquisition geometry, cross-talk between density and other parameters is difficult to prevent, as is the cross-talk into $Q$ parameters.

In a second set of numerical tests, the acquisition geometry is changed to include transmission information. 98 receivers at 495 m depth were added to the previous acquisition. The additional scattering angles considered in this test should make it easier to prevent cross-talk between elastic properties. The inversion result using the inexpensive optimiza-

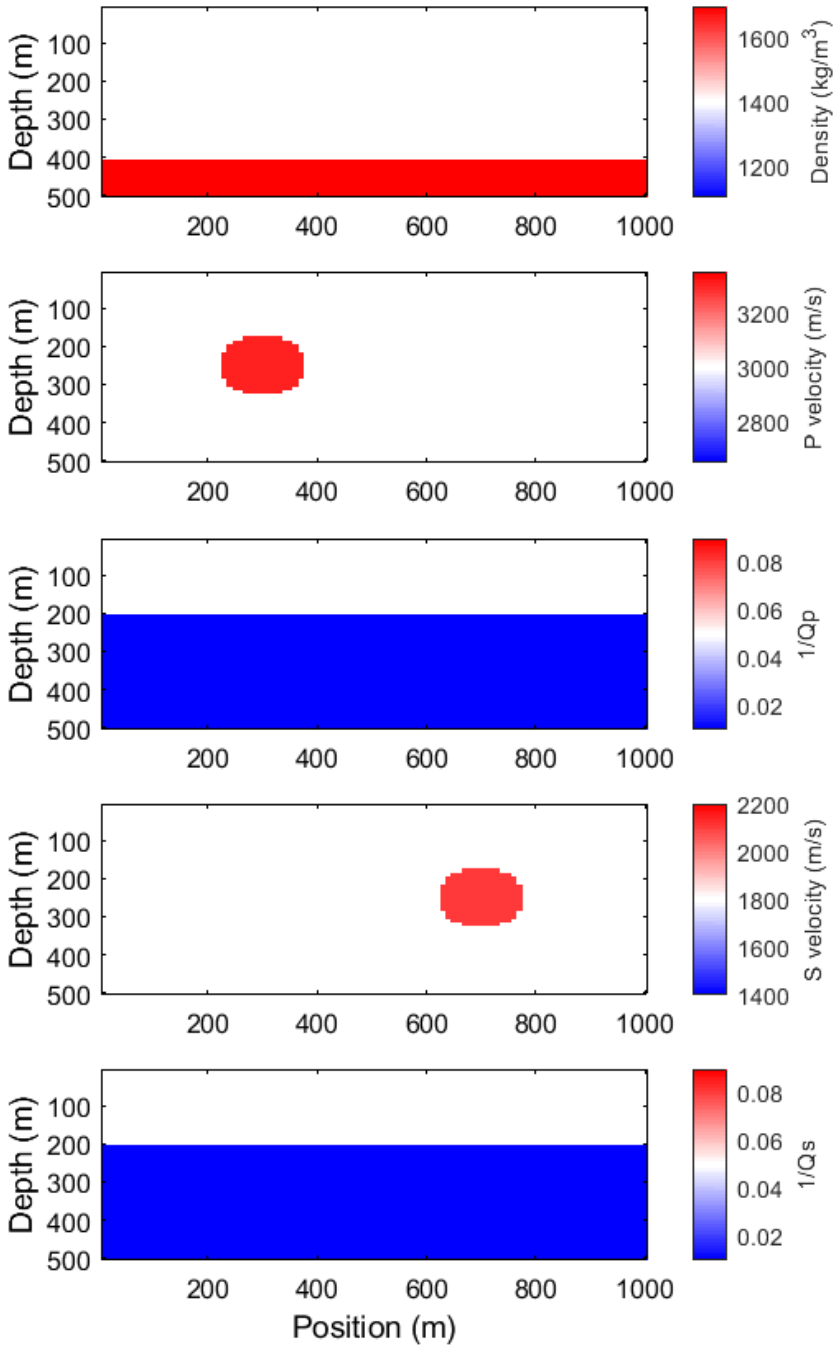


FIG. 4. True model used for the numerical examples.

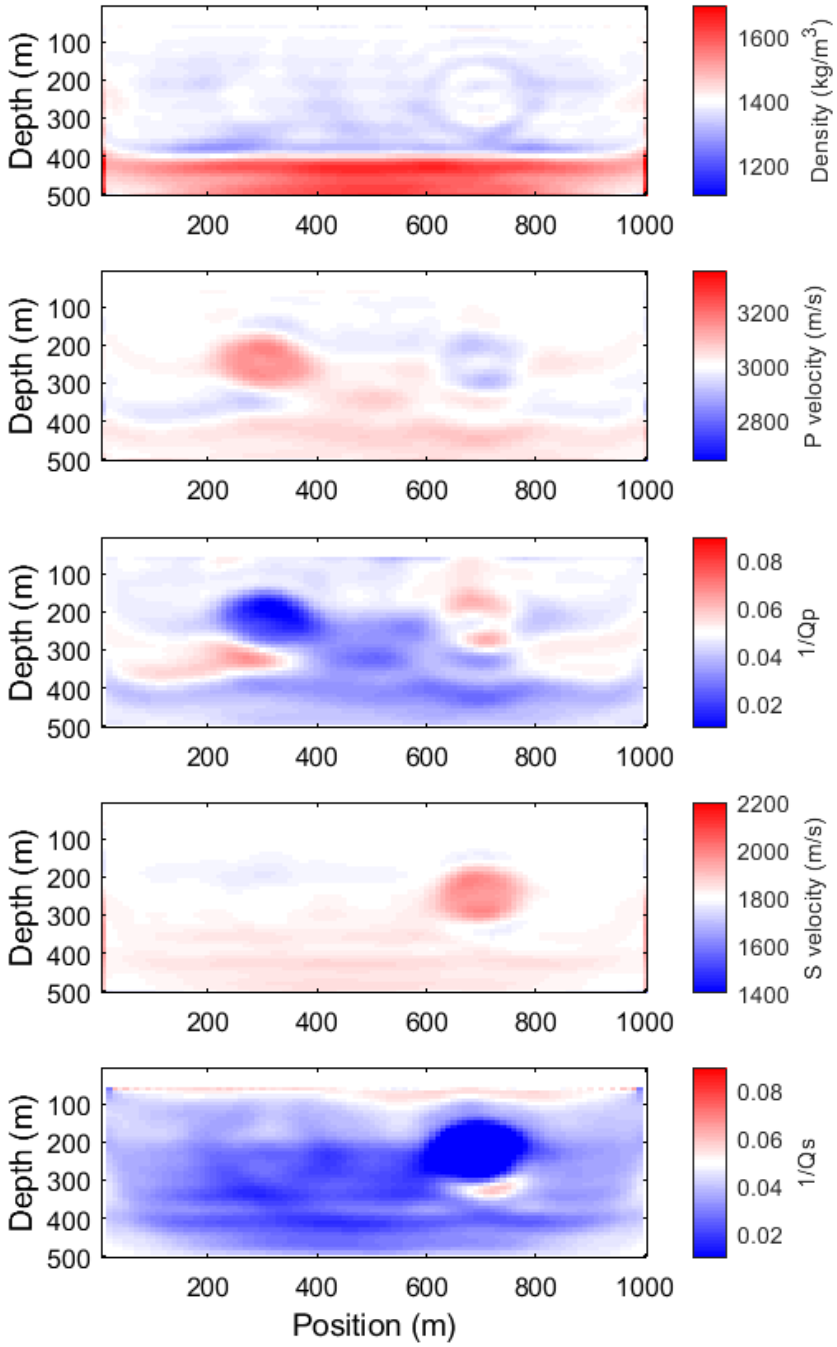


FIG. 5. Recovered elastic model with surface acquisition, and five maximum inner iterations in the TGN inversion.

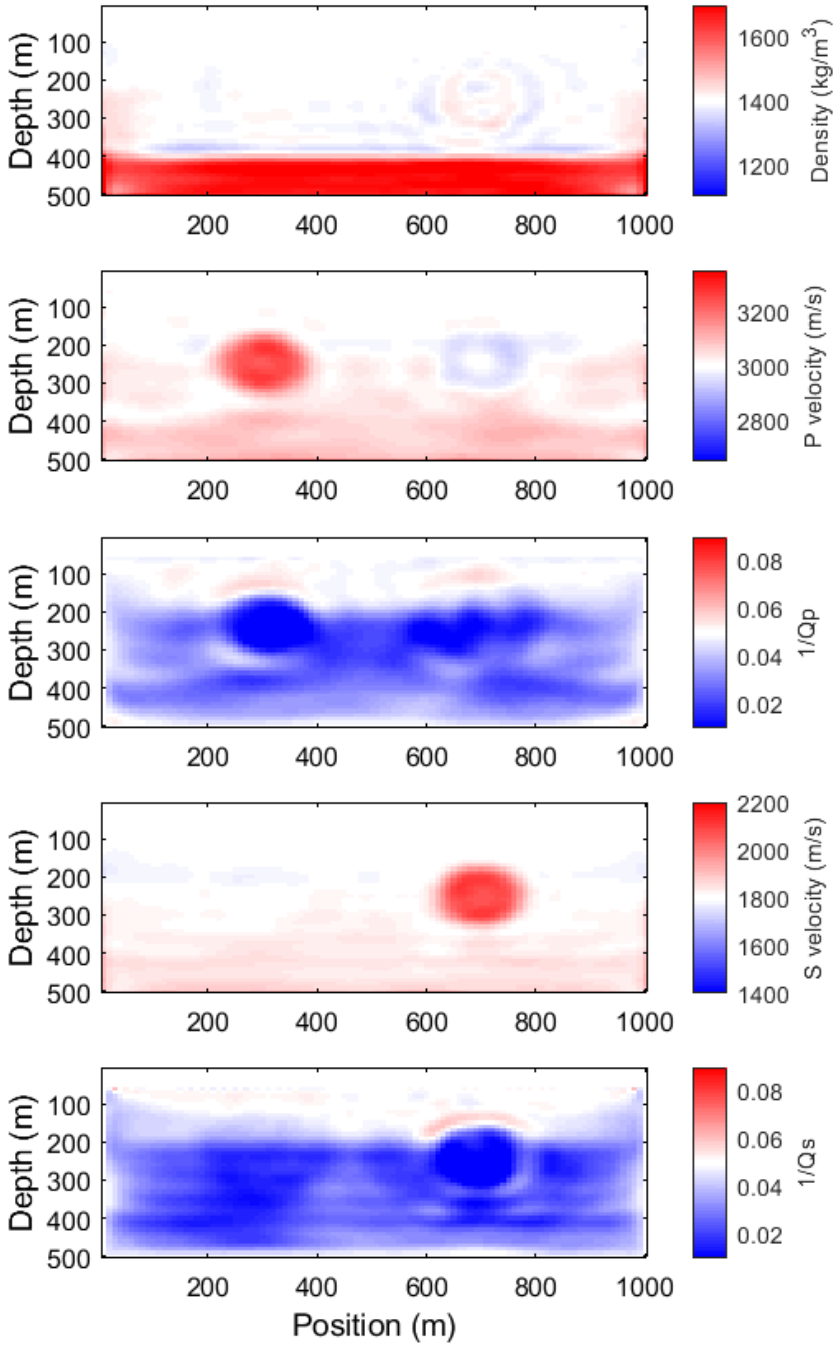


FIG. 6. Recovered anelastic model with surface acquisition, and twenty maximum inner iterations in the TGN inversion.

tion approach (maximum five inner loop iterations) is shown in figure 7. As expected, cross-talk between the velocity variables and density is relatively small in this example. Strong cross-talk from velocity terms into $Q$ remains present, despite the improved acquisition. Interestingly, there is considerable cross-talk from $Q_P$ or $Q_S$ into density. Unlike the other cross-talk behaviour, this occurs despite the very different radiation patterns of these variables at transmissive scattering angles. The $Q_P$ model recovered here does not correctly identify the upper boundary of the high $Q$ region. This could be a result of the transmission-type data dominating the inversion, giving limited ability to identify this contrast. The $Q_S$ model is still dominated by reflection-type effects, as evidenced by the high-to-low contrast at the location of the $v_S$ anomaly. Cross-talk from $Q_P$ into $v_P$ is evident, lower velocities are obtained above the $Q_P$ anomaly, and higher values are found below.

The more comprehensive acquisition was also tested with the more expensive inversion approach (maximum twenty inner loop iterations). This approach reduced the cross-talk from $v_P$ into $Q_P$ as shown in figure 8. Once again, however, substantial cross-talk remains in the model. The cross-talk from $Q_P$ into $v_P$ and $\rho$ remains prominent, and that between $v_S$ and $Q_S$ remains severe. These modes of cross-talk are different from the persistent modes in the surface-only acquistion inversion.

In the final test, the model parameterization was changed, using the transmission-type $Q$ variables strategy described in the previous section. Using the inexpensive optimization approach, the result in figure 9 was obtained. In comparison to the equivalent test with traditional parameterization (figure 7) a number of improvements are evident. Most model estimates improve, with the possible exception of the $Q_S$ model, which underestimates $Q$ in the lower region. More improvement is needed however, as there is still substantial $Q$ related cross-talk, especially into density.

## DISCUSSION

The simple numerical tests here present an inversion problem with comprehensive acquisition, no noise, and an inversion which considers wave propagation in exactly the same way as the forward modeling used to generate the data. For the tests on transmission-type acquisition geometry, receivers are placed in unrealistic locations, which would not be accessible in a real seismic survey. In these and other respects, this numerical problem is substantially easier than application of FWI to real seismic data. Despite these considerable advantages the inversion struggles to distinguish different physical properties. This suggests that more sophisticated inversion strategies need to be employed if cross-talk is to be successfully mitigated.

The cross-talk observed between $Q$ and $\rho$ is interesting because of the limited similarity of the corresponding radiation patterns. This may arise due to the similarity of the back-scattered $\rho$ radiation pattern and the forward scattered $Q$ radiation patterns. In effect, the $\rho$ model may introduce an additional contrast in order to increase reflection amplitudes to offset underestimated $Q$ between the receivers and the reflector. This cross-talk needs to be investigated further.

In these examples the $Q_P$ model suffered less cross-talk with the more complete ac-

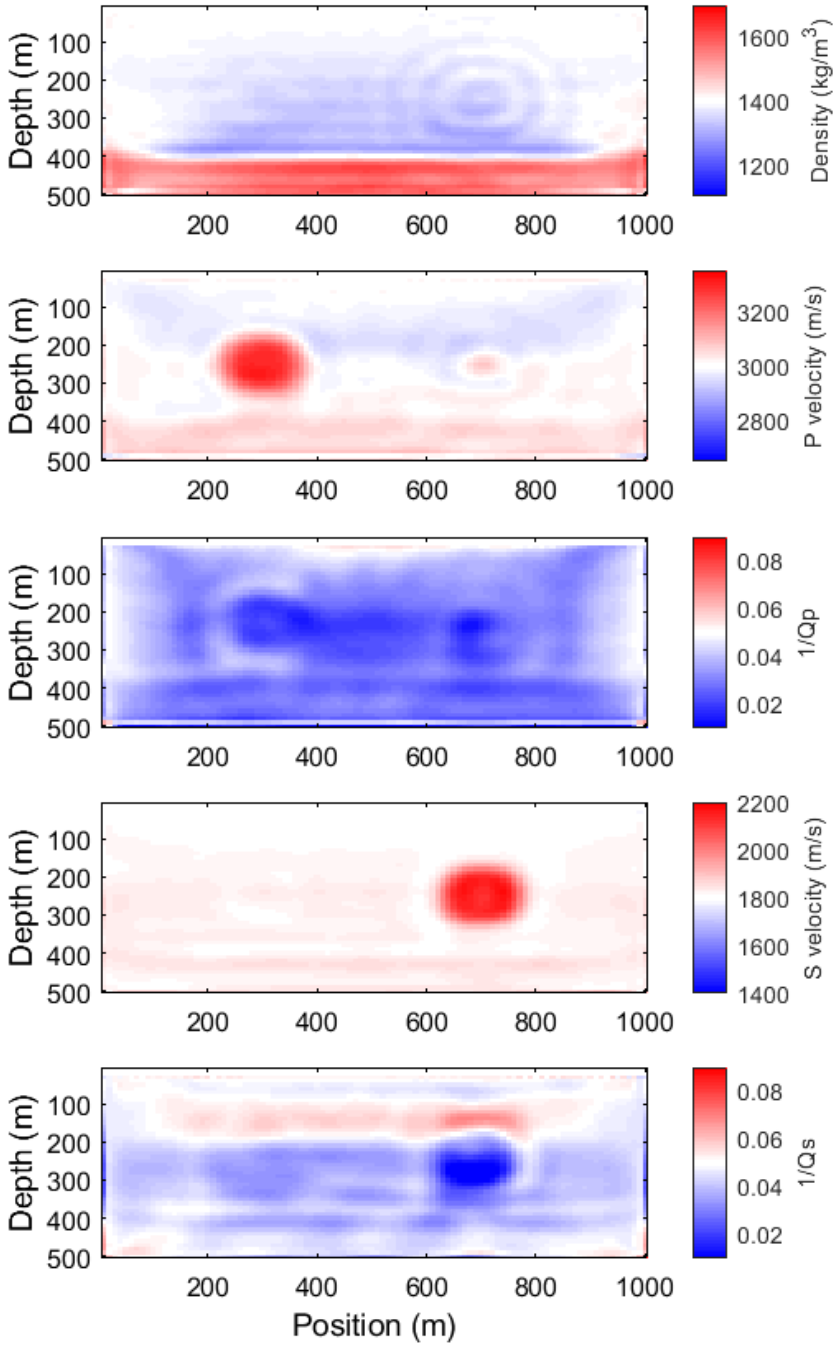


FIG. 7. Recovered anelastic model with comprehensive acquisition, and five maximum inner iterations in the TGN inversion.

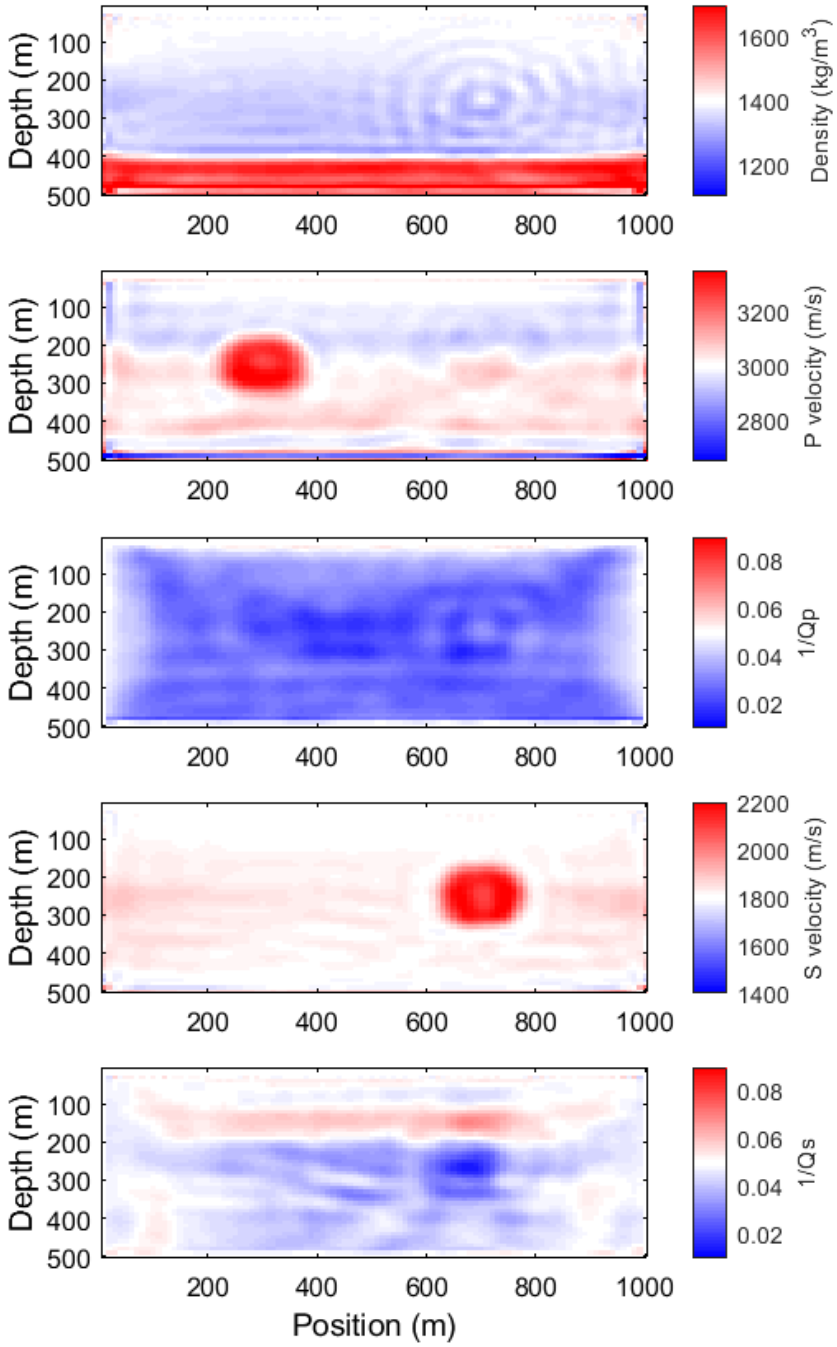


FIG. 8. Recovered anelastic model with comprehensive acquisition, and twenty maximum inner iterations in the TGN inversion.

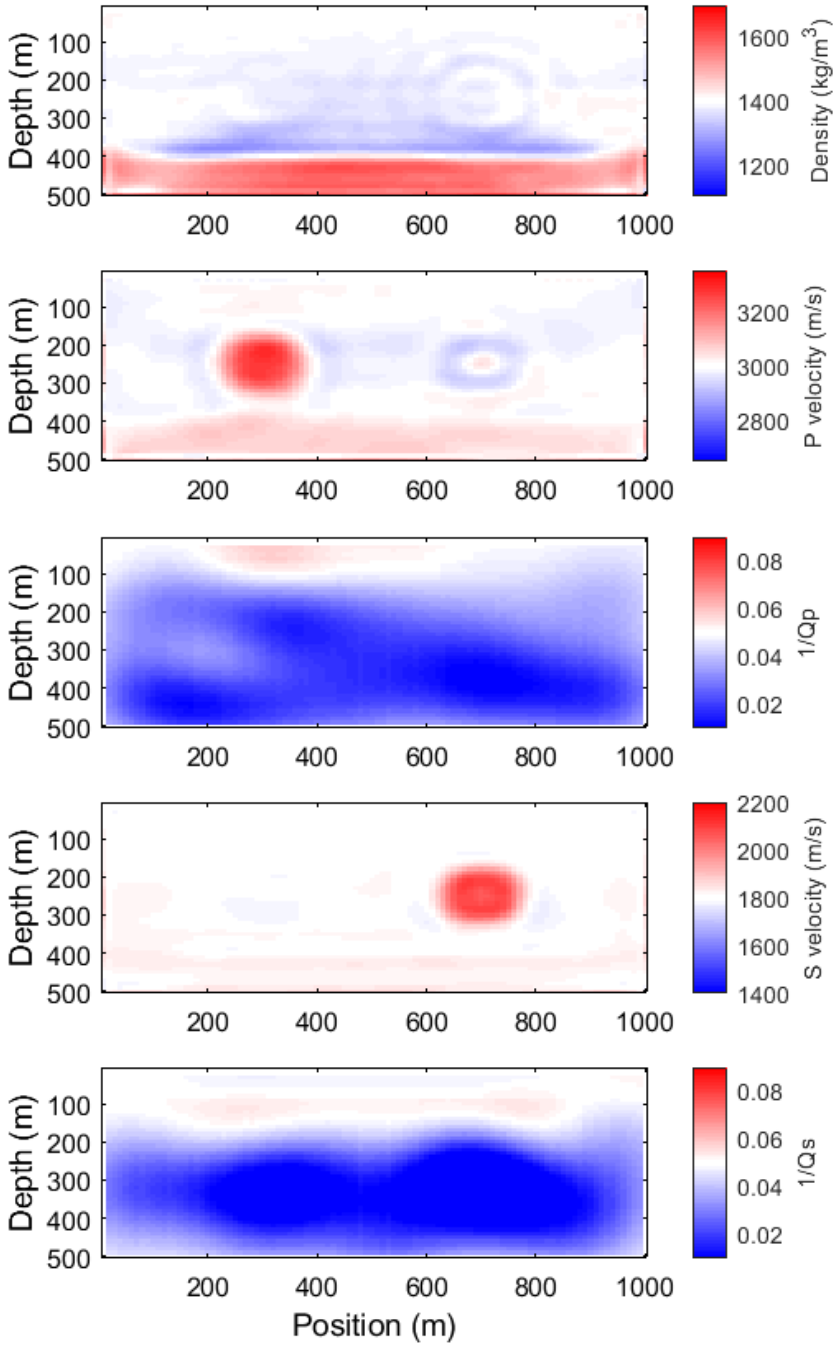


FIG. 9. Recovered anelastic model with comprehensive acquisition, alternate $Q$ variables, and five maximum inner iterations in the TGN inversion.

quisition, but the $Q_S$ model did not. This may have to do with the explosive sources used, which generated only P-waves. The S-waves interacting with the $Q_S$ anomaly were generally upgoing converted waves, so the bottom recievers may have had only limited impact. This is also a topic for further investigation.

## CONCLUSIONS

Visco-elastic full waveform inversion introduces several complications to the FWI problem, cross-talk significantly among them. Considerable cross-talk was observed between the different parameters, making a hierarchical, sequential inversion difficult. The strategy of using long wavelength variables for reducing cross-talk into $Q$ was motivated by the reduced cross-talk expected from reflection-type events, and the reduced number of variables needed to parameterize the problem. While these variables were able to parameterize the model more efficiently and had promising radiation patterns, there was no substantial reduction in cross-talk using these parameters.

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

## REFERENCES

Berenger, J., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of computational physics, **114**, 185–200.

Futterman, W. I., 1962, Dispersive body waves: Journal of Geophysical Research, **67**, 5279–5291.

Hicks, G., and Pratt, R., 2001, Reflection waveform inversion using local descent methods: Estimating attenuation and velocity over a gas-sand deposit: Geophysics, **66**, 598–612.

Innanen, K., 2014, Seismic avo and the inverse hessian in precritical reflection full waveform inversion: Geophys. J. Int., **199**, 717–734.

Kamei, R., and Pratt, R., 2013, Inversion strategies for visco-acoustic waveform inversion: Geophysical Journal International, **194**, 859–884.

Keating, S., and Innanen, K., 2016a, Applying the truncated newton method in anacoustic full waveform inversion: CREWES Annual Report, **28**.

Keating, S., and Innanen, K., 2016b, Challenges in anacoustic full waveform inversion: CREWES Annual Report, **28**.

Keating, S., and Innanen, K., 2017a, Multi-resolution newton optimization in full waveform inversion: CREWES Annual Report, **29**.

Keating, S., and Innanen, K. A., 2017b, Characterizing and mitigating fwi modeling errors due to uncertainty in attenuation physics: SEG Expanded Abstracts.

Kolsky, H., 1956, The propagation of stress pulses in viscoelastic solids: Philosophical Magazine, **1**, 693–710.

Lines, L., and Vasheghani, F., 2008, Reflections on q: CREWES Annual Report, **20**.

Malinowski, M., Operto, S., and Ribodetti, A., 2011, High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full-waveform inversion: Geophysical Journal International, **186**, 1179–1204.

Métivier, L., Brossier, R., Operto, S., and Virieux, J., 2015, Acoustic multi-parameter FWI for the reconstruction of P-wave velocity, density and attenuation: preconditioned truncated Newton approach: SEG Expanded Abstracts, 1198–1203.

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.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metivier, L., and Virieux, J., 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, 1040–1054.

Plessix, R. E., Stopin, A., Kuehl, H., Goh, V., and Overgaag, K., 2016, Visco-acoustic full waveform inversion: EAGE Expanded Abstracts.

Pratt, R., 1990, Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging: Geophys. J. Int., **133**, 341–362.

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

Tarantola, A., 1986, A strategy for nonlinear inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

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