Hessian-based multiparameter fractional viscoacoustic full waveform inversion

Guangchi Xing and Tieyuan Zhu, Department of Geosciences, The Pennsylvania State University

SUMMARY

Recent progress on fractional modeling enables incorporating seismic attenuation into wavefield simulation in an accurate and efficient way. But its inverse problem, i.e., the multiparameter viscoacoustic full waveform inversion (FWI), still suffers from various issues, especially the crosstalk between velocity and attenuation. In this study, we integrate the Hessian information via the Newton-CG framework and develop the multiparameter fractional viscoacoustic FWI algorithm. It significantly mitigates the crosstalk problems and sheds light upon simultaneous inversion for both velocity and Q models.

INTRODUCTION

The traditional full waveform inversion (FWI) produces high-resolution velocity model. Beyond seismic velocity, seismic attenuation could provide complementary constraints on subsurface physical properties (e.g., temperature, fluid saturation, and mineral composition). Hence, it would be advantageous to develop the multiparameter viscoacoustic FWI to invert for both velocity and attenuation (Q) models simultaneously. However, one major challenge faced by the multiparameter FWI is the crosstalk (or trade-off) phenomenon, where the data-synthetic residual introduced by one model parameter is mistakenly assigned to another (Hak and Mulder, 2011; Kamei and Pratt, 2013; Alkhalifah and Plessix, 2014). In the viscoacoustic problem, in particular, both kinetic and dynamic information could be attributed to either velocity or Q heterogeneity. Thus, we would expect a serious crosstalk between these two parameters, especially in the classic gradient-based FWI.

One approach to mitigate the crosstalk is to incorporate the information of the 2nd-order Fréchet derivative, i.e., the Hessian. In particular, the multiparameter Hessian has a block structure (Operto et al., 2013) and the off-diagonal blocks measure the coupling effect between different model parameters. Thus, applying the inverse Hessian operator on the gradient, as in the Newton's method (Nocedal and Wright, 2006), is expected to suppress the crosstalk between parameters (Innanen, 2014; Pan et al., 2016). Besides, the Hessian carries the information about the uneven illumination, the finite-frequency effects and the 2nd-order scattering. Hence, the artifacts generated by these effects can be mitigated by the inverse Hessian operator as well (Pratt et al., 1998). However, explicitly calculating, storing and inverting the Hessian matrix requires huge memory space and computational load beyond the current hardware capability, especially for the large-scale 3-D problems (Métivier et al., 2014).

The Newton-CG (also know as truncated Newton) algorithm offers a promising way out of the computational difficulty. At each iteration, it approximates the gradient preconditioned by the inverse Hessian by solving a linear system using a conjugate gradient internal loop (Nocedal and Wright, 2006). This algorithm is computationally feasible as long as the Hessian-vector product can be accessed in an efficient matrix-free fashion. In this respect, previous studies have proposed the 2nd-order adjoint-state method (Fichtner and Trampert, 2011; Métivier et al., 2014; Yang et al., 2018) and the scattering-integral approach (Yang et al., 2016) to compute the Hessian-vector product.

Recently, we proposed a time-domain viscoacoustic wave equation aided by fractional Laplacian operators (Xing and Zhu, 2019b), which guarantees good accuracy and computational efficiency (Zhu and Harris, 2014). Besides, the adjoint of the wave propagator and the Fréchet derivatives associated with this equation have been developed (Xing and Zhu, 2019a). Based on this wave equation, here, we develop a Hessian-based multiparameter fractional viscoacoustic FWI by incorporating the Hessian via the Newton-CG algorithm. Preliminary results indicate that the preconditioned gradient approximated by the internal loop can significantly mitigate the crosstalk issue.

VISCOACOUSTIC FORWARD MODELING

Seismic attenuation of frequency-independent Q media can be described by three parameters (Kjartansson, 1979): the reference angular frequency ω_0, its corresponding phase velocity c_0 and a the strength of attenuation $\gamma = \frac{1}{\pi} arctan(\frac{1}{Q})$. To model the wavefield in such media, we recently propose a fractional viscoacoustic wave equation (Xing and Zhu, 2019b):

$$\mathbf{L}u = (\mathbf{L_0} + \mathbf{L_1} + \mathbf{L_2})u = f, \tag{1}$$

where u is the pressure wavefield, f is the source term. The wave propagator $\mathbf{L}$ has three contributions, i.e., the lossless acoustic wave propagator $\mathbf{L_0}$, the phase dispersion corrector $\mathbf{L_1}$ and the amplitude loss corrector $\mathbf{L_2}$:

$$\mathbf{L_0} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2, \tag{2}$$

$$\mathbf{L_1} = -\gamma\frac{\omega_0}{c}(-\nabla^2)^{\frac{1}{2}} + \gamma\frac{c}{\omega_0}(-\nabla^2)^{\frac{3}{2}}, \tag{3}$$

$$\mathbf{L_2} = (\pi\gamma\frac{1}{c}(-\nabla^2)^{\frac{1}{2}} - \pi\gamma^2\frac{1}{\omega_0}\nabla^2)\frac{\partial}{\partial t}, \tag{4}$$

where $c = c_0 cos(\frac{\pi\gamma}{2})$ is a parameter with the unit of velocity. In this study, we adopt this wave equation to conduct the forward modeling because of its accuracy and computational efficiency (Xing and Zhu, 2019b).

MULTIPARAMETER VISCOACOUSTIC FWI

Our proposed FWI algorithm has a nested loop structure. The internal loop uses the CG iterations to approximately solve the linear system $H\Delta m = -g$, where H is the Hessian matrix, Δm is the model update vector and g is the gradient, i.e., the first order Fréchet derivative. In the external loop, we update both velocity (c) and attenuation (γ) models according to the vector Δm.

Gradient and adjoint propagator

The computation of the gradient g can be formulated using the adjoint-state method (Plessix, 2006). Besides a forward simulation (Equation 1), this process consists of an adjoint simulation and a wavefield interaction:

$$\mathbf{L}^*\lambda = R^*(Ru - d), \tag{5}$$

$$g = -\langle \lambda, \frac{\partial \mathbf{L}}{\partial m} u \rangle_{\mathbb{U}}, \tag{6}$$

where $*$ denotes adjoint; R is the operator that samples the wavefield at the receivers, which makes R^* the operator that pads zeros at non-receiver locations; d is the data; λ denotes the adjoint wavefield; and $\langle \cdot, \cdot \rangle_{\mathbb{U}}$ is the inner product in the state variable (wavefield) space $\mathbb{U}$.

It has been argued that the adjoint viscoacoustic propagator has the form $\mathbf{L}^* = \mathbf{L_0} + \mathbf{L_1} - \mathbf{L_2}$ (Xing and Zhu, 2019a). Physically, it means that the adjoint wavefield λ preserves the velocity dispersion ($\mathbf{L_1}$ term) while compensates ($-\mathbf{L_2}$ term) instead of attenuating the amplitude. Nevertheless, the adjoint wavefield is simulated in a reverse-time order, which makes the reversed adjoint wavefield attenuate the amplitude just like the forward one. The interaction between the two wavefields (Equation 6) requires the operator $\frac{\partial \mathbf{L}}{\partial m}$, the explicit form of which can be derived from Equations (2)-(4).

Hessian-vector product validation

The internal CG loop requires computing the Hessian-vector product in an efficient way. Instead of calculating the Hessian matrix explicitly, the 2nd-order adjoint-state method offers an option to conduct the multiplication in a matrix-free fashion (Métivier et al., 2014; Yang et al., 2018; Fichtner and Trampert, 2011). In particular, for an arbitrary model space vector ν, the product $H\nu$ for each source can be computed at the cost of two more simulations (one forward and one adjoint) and three more interactions:

$$\mathbf{L}\mu_2 = -\langle \nu, \frac{\partial \mathbf{L}}{\partial m} u \rangle_{\mathbb{M}}, \tag{7}$$

$$\mathbf{L}^*\mu_1 = R^*R\mu_2 - \langle \nu, (\frac{\partial \mathbf{L}}{\partial m})^* \lambda \rangle_{\mathbb{M}}, \tag{8}$$

$$H\nu = -\langle \mu_1, \frac{\partial \mathbf{L}}{\partial m} u \rangle_{\mathbb{U}} - \langle \mu_2, \frac{\partial \mathbf{L}^*}{\partial m} \lambda \rangle_{\mathbb{U}} - \langle \nu, \langle \lambda, \frac{\partial^2 \mathbf{L}}{\partial m^2} u \rangle_{\mathbb{U}} \rangle_{\mathbb{M}}, \tag{9}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{M}}$ denotes the inner product in the model space $\mathbb{M}$; μ_2 and μ_1 are the 2nd-order forward and adjoint wavefields, respectively.

To illustrate the physical meaning of the product $H\nu$, we assume that ν represents a scatterer with a perturbation of model parameters (c or γ). As a result, the source of the 2nd-order forward wavefield μ_2 has only non-zero values at the location of the scatterer ν (Equation 7), which suggests that μ_2 propagates outward from the scatterer. The two terms on the right hand side of the Equation (8) indicate that the 2nd-order adjoint wavefield μ_1 has two sources: the first source (Source A) is located at the receiver while the second source (Source B) is at the scatterer. Their corresponding wavefields are denoted by μ_{1A} and μ_{1B}, respectively.

The Hessian-vector product $H\nu$ consists of three interactions (Equation 9) denoted by $(H\nu)_1$, $(H\nu)_2$ and $(H\nu)_3$. In particular, $(H\nu)_1$ has two parts, $(H\nu)_{1A}$ and $(H\nu)_{1B}$, corresponding to μ_{1A} and μ_{1B}, respectively. The interaction between μ_{1A} and the forward wavefield u gives rise to $(H\nu)_{1A}$. Hence, it is influenced by the Fresnel zone between the source and the receiver. Similarly, $(H\nu)_{1B}$ marks the area between the source and the scatterer while $(H\nu)_2$ is sensitive to the area between the receiver and the scatterer. The last interaction $(H\nu)_3$ only has non-zero values at the location of the scatterer and it accounts for the non-linearity introduced by the parameterization (Fichtner and Trampert, 2011).

Internal CG loop

Equipped with the efficient computation of both gradient and the Hessian-vector product, we can construct the internal CG loop following Nocedal and Wright (2006), as shown in Algorithm 1.

Algorithm 1 CG method to solve $H\Delta m = -g$

Input: g, k_{max} (max iter.), J_{tol} (residual tolerance)
Require: Method to compute $H\nu$ for an arbitrary ν
Output: Model parameter update Δm
Initialization: $\Delta m \leftarrow 0$, $r \leftarrow g$, $p \leftarrow -r$, $J_1 \leftarrow r^T r$, $k \leftarrow 0$
while $J_1 > J_{tol}$ & $k < k_{max}$ **do**
  $\xi \leftarrow Hp$
  $\alpha \leftarrow \frac{J}{p^T \xi}$
  $r \leftarrow r + \alpha \xi$s
  $\Delta m \leftarrow \Delta m + \alpha p$
  $J_2 \leftarrow r^T r$
  $\beta \leftarrow \frac{J_2}{J_1}$
  $p \leftarrow -r + \beta p$
  $J_1 \leftarrow J_2$
  $k \leftarrow k + 1$
end while

The early iterations of the CG algorithm account for the largest eigenvalues of the Hessian. Hence, a truncation strategy for the internal loop has an intrinsic regularization effect (Kaltenbacher et al., 2008; Métivier et al., 2014). Thus, we only run the CG internal loop for a few iterations ($k_{max} \sim 5$) in our viscoacoustic FWI problem. After that, we conduct a line search to determine the step lengths for different parameter classes.

Step length computation

The most popular line search methods for FWI base on either parabola fitting (e.g., Köhn et al., 2012) or Born modeling (e.g., Pica et al., 1990). In the multiparameter viscoacoustic FWI, we need to take into account the interaction between different parameter classes in optimizing the step lengths for both velocity (α_c) and attenuation (α_γ). Here, we generalize the Born-based method and propose a Born-based multiparameter step length computation (BMPSLC) algorithm.

We separate the model update Δm into two vectors $\Delta m = \Delta m_c + \Delta m_\gamma$, where Δm_c only has non-zero values in the first half (velocity part) while Δm_γ only has non-zero values in the second half (attenuation part). Conducting the Taylor expansion of the objective function after the model update, followed by zeroing its partial derivatives with respect to α_c and α_γ, yields:

$$\langle \frac{\delta f}{\delta m}\Delta m_c, \frac{\delta f}{\delta m}\Delta m_c\rangle_{\mathbb{D}}\alpha_c + \langle \frac{\delta f}{\delta m}\Delta m_\gamma, \frac{\delta f}{\delta m}\Delta m_c\rangle_{\mathbb{D}}\alpha_\gamma = \langle d - f(m), \frac{\delta f}{\delta m}\Delta m_c\rangle_{\mathbb{D}}, \quad (10)$$

$$\langle \frac{\delta f}{\delta m}\Delta m_c, \frac{\delta f}{\delta m}\Delta m_\gamma\rangle_{\mathbb{D}}\alpha_c + \langle \frac{\delta f}{\delta m}\Delta m_\gamma, \frac{\delta f}{\delta m}\Delta m_\gamma\rangle_{\mathbb{D}}\alpha_\gamma = \langle d - f(m), \frac{\delta f}{\delta m}\Delta m_\gamma\rangle_{\mathbb{D}}, \quad (11)$$

where f is the forward modeling operator, which makes $\frac{\delta f}{\delta m}$ the linear Born modeling operator; $\langle\cdot,\cdot\rangle_{\mathbb{D}}$ denotes the inner product in data space $\mathbb{D}$. Two Born modelings are required to obtain the coefficients of Equations (10)-(11). In our study, we conduct the Born modelings in a finite-difference fashion:

$$\frac{\delta f}{\delta m}\Delta m_i \approx \frac{f(m+\epsilon\Delta m_i) - f(m)}{\epsilon}, \qquad (i = c, \gamma) \quad (12)$$

where ϵ is a small scalar. So, α_c and α_γ can be achieved by solving the linear system of Equations (10)-(11) at the cost of two additional forward modeling per source.

NUMERICAL EXAMPLES

Hessian-vector product

We set up a homogeneous 2-D target model on a 400×200 grid with spacing of 8 m in both x and z directions. The target model has a quality factor $Q = 80$ and a reference phase velocity $c_0 = 3.05$ km/s at 20 Hz. We put a source at (0.24, 0.8) km with a 20 Hz Ricker wavelet and one receiver is located at (2.96, 0.8) km. Using this target model, we simulate the synthetic seismogram, which is considered as the ground truth data. Meanwhile, we build a homogeneous initial model on the same grid with slightly different velocity and attenuation ($c_0 = 3.0$ km/s, $Q = 100$). To demonstrate the physical meaning of the Hessian-vector product, we set up the model space vector ν as a scatterer with 20 m/s velocity perturbation located at (1.36, 0.64) km. In another word, the vector ν has only one non-zero entry, which is located in its first half (velocity part).

We conducted two forward simulations (Equations 1 and 7) and two adjoint simulations (Equations 5 and 8). Taking the interactions between the resultant wavefields (Equation 9) leads to the Hessian-vector product ($H\nu$), which includes both velocity (c) and attenuation (γ) parts. Four constituent parts ($(H\nu)_{1A}$, $(H\nu)_{1B}$, $(H\nu)_2$ and $(H\nu)_3$) are shown in Figure 1(a)-(j) along with their summation. As expected, each of the four constituent parts has the shape as described in the previous section.

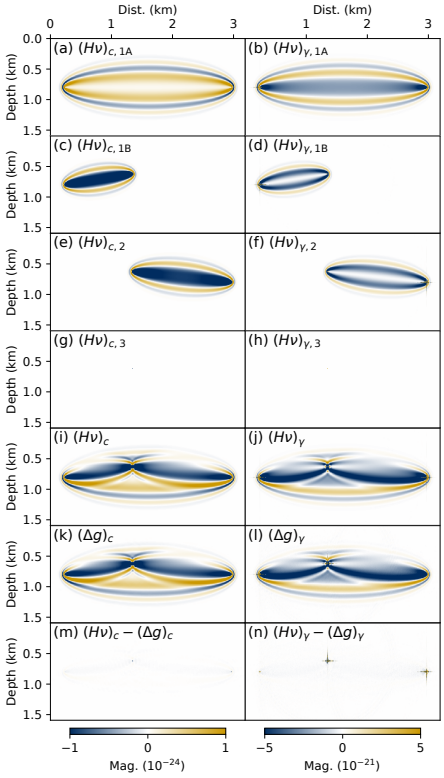

Figure 1: Hessian-vector product. Left: velocity (c) part; right: attenuation (γ) part. (a)-(h): Constituent parts of the product; (i)-(j) 2nd-order adjoint-state method product; (k)-(l) finite-difference method product; (m)-(n) residual between two methods.

To validate the accuracy of the algorithm, we use the finite-difference method to compute the product with the same ν. The definition of the Hessian indicates

$$H\nu \approx g(m+\nu) - g(m) \quad (13)$$

when ν is sufficiently small, where $g(m)$ and $g(m+\nu)$ are the gradients evaluated at the initial model m and the perturbed model $m+\nu$, respectively. We conducted this finite-difference computation and the resultant Hessian-vector product is shown in Figure 1(k)-(l). It turns out that the products computed using both methods are almost the same; and their residuals (Figure 1m-n) are negligible except for some artifacts at source locations (either forward or adjoint, either first-order or second-order), which can be removed by a Gaussian taper applied on the source region wavefields.

2-D Cross-well FWI example

We set up a 2-D model on a 101×151 grid (0.8 m interval) with a cross-well acquisition system: 15 sources on the left edge and 36 receivers on the right edge. The target model (Figure 2) has a homogeneous background ($c_0 = 3.0$ km/s at 150 Hz, $Q = 100$) with a velocity anomaly ($c_0 = 3.05$ km/s at 150 Hz) on the top left and an attenuation anomaly ($Q = 50$) on the bottom right. We start from homogeneous initial models without the anomalies of either velocity or attenuation.

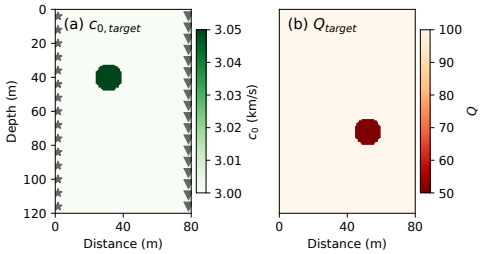

Figure 2: Target models: (a) phase velocity; (b) Q. Stars denote sources and triangles represent receivers.

For the classic gradient-based FWI, the update direction is the negative gradient ($-g$); while for our proposed Hessian-based method, the direction is the (approximate) preconditioned negative gradient (Δm) produced by the internal CG loop (Algorithm 1). At the very first iteration of our example, we show the directions computed using both methods in Figure 3. In the gradient-based FWI, both velocity and attenuation direction can capture the anomalies but the velocity anomaly is mistakenly mapped to the attenuation direction. In particular, the maximum value of the attenuation direction is located at the velocity anomaly (Figure 3b), leading to a significant crosstalk. The Hessian-based direction, on the contrary, is generally sharper than the gradient-based ones. More importantly, the crosstalk artifact in the attenuation direction is substantially reduced (Figure 3d), which confirms the capability of the internal CG loop to mitigate the crosstalk.

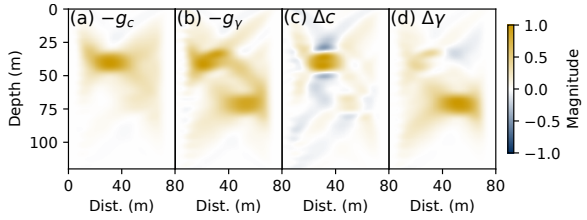

Figure 3: Normalized model update direction: (a)-(b) gradient-based method; (c)-(d) proposed Hessian-based method. Left: velocity (c); right: attenuation (γ).

In Figure 4, we show the final velocity and attenuation models after 20 iterations. Both anomalies are recovered with little crosstalk artifacts. The magnitudes of the anomalies, however, are slightly less than the target model, while the shapes of both anomalies are wider than they are supposed to be. Both imperfections are due to the limitation of the cross-well acquisition system, where the horizontal resolution is better than the vertical one, leading to less constraints on the width of the anomalies and the reduction in the magnitude. If we add 9 sources on the top margin and 23 receivers on the bottom, the resultant final models (after 20 iterations) are shown in Figure 5. With better coverage, the shapes and magnitudes of both anomalies are much improved.

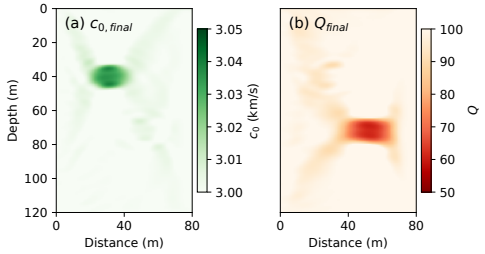

Figure 4: Final models: (a) phase velocity; (b) Q.

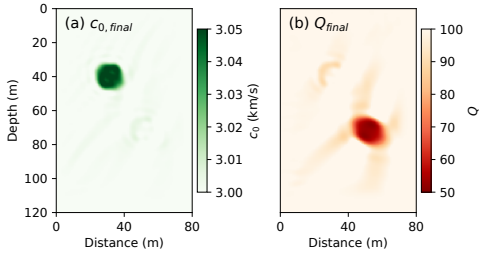

Figure 5: Final models for the FWI with more sources and receivers: (a) phase velocity; (b) Q.

CONCLUSIONS

Based on the recently proposed fractional wave equation, we develop the multiparameter viscoacoustic FWI to invert for both velocity and attenuation simultaneously. To deal with the crosstalk between model parameters, we incorporate the Hessian via the Newton-CG framework. In particular, we use the 2nd-order adjoint-state method to compute the Hessian-vector product in the internal CG loop; and we develop a Born-based multiparameter step length computation algorithm by considering the interaction between different parameter classes. Preliminary results indicate that the crosstalk can be significantly mitigated, and both velocity and attenuation models can be properly recovered.

ACKNOWLEDGMENTS

This work was supported by the U.S. DOE Contract DE-FE0031544 and the NSF Grant EAR 1919650. The authors would like to thank Dr. Chao Huang for very helpful discussion on the FWI algorithm.

REFERENCES

Alkhalifah, T., and R.-É. Plessix, 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study: Geophysics, **79**, R91–R101.

Fichtner, A., and J. Trampert, 2011, Hessian kernels of seismic data functionals based upon adjoint techniques: Geophysical Journal International, **185**, 775–798.

Hak, B., and W. A. Mulder, 2011, Seismic attenuation imaging with causality: Geophysical Journal International, **184**, 439–451.

Innanen, K., 2014, Reconciling seismic avo and precritical reflection fwi-analysis of the inverse hessian, *in* SEG Technical Program Expanded Abstracts 2014: Society of Exploration Geophysicists, 1022–1027.

Kaltenbacher, B., A. Neubauer, and O. Scherzer, 2008, Iterative regularization methods for nonlinear ill-posed problems: Walter de Gruyter, **6**.

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

Kjartansson, E., 1979, Constant Q-wave propagation and attenuation: Journal of Geophysical Research: Solid Earth, **84**, 4737–4748.

Köhn, D., D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: Geophysical Journal International, **191**, 325–345.

Métivier, L., F. Bretaudeau, R. Brossier, S. Operto, and J. Virieux, 2014, Full waveform inversion and the truncated newton method: quantitative imaging of complex subsurface structures: Geophysical Prospecting, **62**, 1353–1375.

Nocedal, J., and S. Wright, 2006, Numerical optimization: Springer Science & Business Media.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

Pan, W., K. A. Innanen, G. F. Margrave, M. C. Fehler, X. Fang, and J. Li, 2016, Estimation of elastic constants for hti media using gauss-newton and full-newton multiparameter full-waveform inversion: Geophysics, **81**, R275–R291.

Pica, A., J. Diet, and A. Tarantola, 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: Geophysics, **55**, 284–292.

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.

Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss–newton and full newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Xing, G., and T. Zhu, 2019a, Fréchet kernels based on a fractional viscoacoustic wave equation, *in* SEG Technical Program Expanded Abstracts 2019: Society of Exploration Geophysicists, 1455–1459.

———, 2019b, Modeling frequency-independent q viscoacoustic wave propagation in heterogeneous media: Journal of Geophysical Research: Solid Earth.

Yang, J., Y. Liu, and L. Dong, 2016, Simultaneous estimation of velocity and density in acoustic multiparameter full-waveform inversion using an improved scattering-integral approach: Geophysics, **81**, R399–R415.

Yang, P., R. Brossier, L. Métivier, J. Virieux, and W. Zhou, 2018, A time-domain preconditioned truncated newton approach to visco-acoustic multiparameter full waveform inversion: SIAM Journal on Scientific Computing, **40**, B1101–B1130.

Zhu, T., and J. M. Harris, 2014, Modeling acoustic wave propagation in heterogeneous attenuating media using decoupled fractional laplacians: Geophysics, **79**, T105–T116.