

Quantify parameter resolution for 3D elastic and anisotropic full-waveform inversion via multi-parameter Hessian probing

Wenyong Pan, Kris Innanen, Junxiao Li

ABSTRACT

Estimating multiple physical parameters in subsurface using full-waveform inversion (FWI) methods suffers from parameter crosstalk challenge. The inherent ambiguities among different physical parameters make the inverse problems much more non-linear. The parameter resolution is highly influenced by the parameterization selected for describing the subsurface elastic and anisotropic medium. Quantify the parameter resolution for determining the optimal parameterization in multi-parameter FWI becomes essential for reducing the parameter crosstalk difficulty. In recent years, researchers devote significant effort for evaluating the resolving abilities of different parameterizations for elastic and anisotropic full-waveform inversion based on the analytic solutions of the Fréchet derivative wavefields (so called "scattering" or "radiation" patterns). However, these studies may not be able to evaluate the resolving abilities of different parameter classes completely because of the inherent defects of the scattering patterns. The goal of this research is to develop new strategies for quantifying the parameter resolution for multi-parameter elastic and anisotropic full-waveform inversion. We find that the multi-parameter Hessian, the second-order derivative of the misfit function, provides direct and complete measurements of the inter-parameter trade-off. The investigations based on scattering patterns can be interpreted as an asymptotic approximation of the multi-parameter Gauss-Newton Hessian. With the block-diagonal approximation of the multi-parameter Gauss-Newton Hessian, we are able to assess the parameter resolution by taking the geometrical spreading and complex model into consideration. Furthermore, with the adjoint-state technique, we are able to calculate one column of the multi-parameter Hessian (multi-parameter point spread function), defined as parameter resolution kernel in this research, with which we can evaluate the inter-parameter mapping of different parameters at adjacent positions locally by considering finite-frequency effects. With the help of random probing technique, we can infer the characteristics of the multi-parameter Hessian within the whole model at affordable computation cost for large-scale inverse problems. Thus, with the multi-parameter Hessian, the parameter resolution can be assessed more completely.

INTRODUCTION

In recent decades, researchers have devoted significant effort to the development of full-waveform inversion (FWI) techniques, which promise to provide high-resolution estimates of subsurface elastic and anisotropic properties (Lailly, 1983; Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009). In mono-parameter FWI, researches focus on overcoming cycle-skipping difficulty for recovering P-wave velocity due to the lack of low frequencies and inaccurate initial model (Ma and Hale, 2012; Warner and Guasch, 2014; Wu et al., 2014; Li and Demanet, 2016; Zhu and Fomel, 2016; Pan et al., 2016b). If anisotropy and elasticity are ignored, the model estimations will be contaminated by the mappings from elastic and anisotropic effects in the data. It is also necessary to reconstruct the subsurface elastic and anisotropic parameters for reservoir characterization. However, simultaneously

reconstructing multiple physical parameters is much more challenging. The parameter crosstalk artifacts involved by the inherent ambiguities (or coupling effects) among different physical parameters significantly increase the non-linearity of the inverse problems (Operto et al., 2013; Innanen, 2014a).

For mitigating the parameter crosstalk difficulty in multi-parameter FWI, a lot of studies have been carried out for the parameter resolution analysis based on the analytic solutions of the Frechét derivative wavefields ("scattering" or "radiation" patterns) for different parameter classes (Tarantola, 1986; Gholami et al., 2013b,a; Alkhalifa and Plessix, 2014; Kamath and Tsvankin, 2014; Podgornova et al., 2015). Coupling effects appear between two different physical parameters, if the scattered wavefields due to the perturbations of the two physical parameters overlap at the range of scattering (or azimuthal) angle. These analyses are essential for determining the optimal parameterization, acquisition geometry and inversion strategies for multi-parameter FWI. Gholami et al. (2013a) investigated the scattering patterns of parameters resulting from various parameterizations of the acoustic FWI. Similar types of analyses have been used to identify an optimal parameterization for acoustic VTI FWI. (Alkhalifa and Plessix, 2014) highlights the power of the horizontal P-wave velocity in reducing the number of parameters for VTI FWI. Podgornova et al. (2015) analyzed the resolution limits of multi-parameter inversion based on the singular value decomposition of far-field linearized inversion operator in the wavenumber domain. The analysis of amplitude variations of the scattering patterns help to understand the parameter crosstalk (or trade-off) between parameters. The parameter resolution in multi-parameter FWI is highly influenced by the parameter class in the inversion process. It is important to determine an optimal parameterization for elastic and anisotropic FWI for avoiding the parameter crosstalk. So the studies on parameter resolution analysis lies at the center of elastic and anisotropic FWI. Köhn et al. (2015) studied the resolution ability of the 21 elastic constants in triclinic anisotropic media taking into consideration various acquisition systems and configurations. For elastic orthorhombic media, they conclude that it will be very hard to recover all of the parameters needed to describe such anisotropy. Studies focused on the ability of full-waveform inversion in recovering anisotropic properties. However, the multi-parameter FWI for anisotropic media is prone to entrapment in local minimum because we invert for a large number of parameters from given seismic data, which results in possible trade-off between parameters (Bergslid et al., 2015). For transverse isotropy with a horizontal symmetry axis (HTI media), the simplest azimuthal anisotropic model for describing vertical cracks, reflection seismic signatures can be described by five independent elastic constants (Rüger, 1997; Tsvankin, 1997b,a). The HTI media can also be described by P-wave velocity, S-wave velocity, density and Thomsen parameters ϵ , δ and γ . Inverting azimuthally anisotropic parameters are essential important for fractured reservoir characterization. In this research, we focus on resolving abilities of azimuthally anisotropic parameters in full-waveform inversion.

We find that these parameter resolution studies based on scattering patterns suffer from a number of inherent drawbacks: (1) the background model is assumed to be isotropic and homogeneous; (2) incident plane-wave with far-field approximation is employed; (3) finite-frequency effects are ignored; (4) the parameter crosstalk due to different physical parameters at different locations, which is especially important when low frequencies are deficient, are not considered; (5) the parameter crosstalk due to higher-order scattering

effects are not considered; (6) the travel time information is not considered. One goal of this research is to develop new techniques for parameter resolution analysis in elastic and anisotropic media.

The Hessian operator describes the geometry of the objective function in terms of curvature or convexity (Fichtner and van Leeuwen, 2015). Hessian provides the most direct measurements of resolution and trade-offs as it describes the change of the misfit when \mathbf{m} is slightly perturbed to $\mathbf{m} + \Delta\mathbf{m}$ (Fichtner and van Leeuwen, 2015). The multi-parameter Hessian has a block structure. The off-diagonal element in the diagonal block $H_{\theta\theta}(\mathbf{x}, \mathbf{x}')$ defines the spatial correlation of the same model parameter \mathbf{m}_θ at position \mathbf{x} and \mathbf{x}' . The diagonal element $H_{\theta\theta}(\mathbf{x}, \mathbf{x})$ of the off-diagonal block measures the coupling of model parameter \mathbf{m}_θ and model parameter \mathbf{m}_θ at the same position \mathbf{x} . The off-diagonal element in the off-diagonal block measure the coupling of the model parameter \mathbf{m}_θ at \mathbf{x} and the model parameter \mathbf{m}_θ at position \mathbf{x}' . It has been proved that the multi-parameter Hessian can be employed to suppress the parameter trade-off for multi-parameter FWI (Innanen, 2014b; Pan et al., 2016a). We also notice that the resolving abilities of different parameter classes can be quantified with multi-parameter Hessian. A further study reveals that the parameter resolution studies based on scattering patterns is actually equivalent to an asymptotic approximation of the multi-parameter Gauss-Newton Hessian with a perfect acquisition survey. Overlapping the scattering patterns due to different physical parameters can be considered as the diagonal element in the off-diagonal blocks of the multi-parameter Gauss-Newton Hessian. In this research, we propose to analyze the parameter resolution for multi-parameter FWI using the multi-parameter Hessian.

However, the Hessian is a large and dense matrix. It is impractical to construct multi-parameter Hessian explicitly for large-scale inverse problems. We can extract partial information from Hessian for analysis. The diagonal elements of the off-diagonal blocks in multi-parameter Gauss-Newton Hessian, which can be constructed by adjoint-state technique explicitly, measure the geometrical spreading and the finite-frequency effects for parameter resolution analysis at the same location. To quantify the coupling effects of different physical parameters at different locations, we need one column (or row) of the multi-parameter Hessian. In this paper, we will show that one column of multi-parameter Hessian is proportional to multi-parameter point spread function, which can be constructed by spike probing technique with adjoint-state method. What's more, the multi-parameter full Hessian also considers the parameter crosstalk due to second-order scattering effects. The multi-parameter point spread functions can only quantify the parameter resolution locally. To measure the parameter trade-off in the whole model, a random probing technique can be employed (An, 2012; Trampert et al., 2013; Fichtner and van Leeuwen, 2015). Fichtner and van Leeuwen (2015) demonstrated that autocorrelations of random-model applications to the Hessian yields various resolution measures, including direction- and position-dependent resolution lengths and the strength of the inter-parameter mappings.

In this paper, we first review the basic principle of full-waveform inversion. Then, we revisit the parameter resolution issue in multi-parameter FWI and explain the disadvantages the scattering patterns for parameter resolution studies. We then explain how to quantify the parameter resolution by probing the multi-parameter Hessian. In this numerical section, we construct the sensitivity kernels for anisotropic parameters using adjoint-state method.

We also show the diagonal elements in block matrices of the multi-parameter Hessian in orthorhombic media. We give numerical examples to show that the column of the multi-parameter Hessian can be extracted for quantifying the parameter resolution for 3D elastic and anisotropic full-waveform inversion.

REVIEW OF LEAST-SQUARES WAVE-EQUATION INVERSION

Full-waveform inversion (FWI) seeks to estimate the subsurface (an) elastic and anisotropic properties by iteratively minimizing the differences between recorded data $\mathbf{d}_{obs}(\mathbf{x}_s, \mathbf{x}_r, t)$ and synthetic data $\mathbf{d}_{syn}(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{m})$ simulated from an estimated model \mathbf{m} . The misfit function is formulated as a l_2 norm:

$$\Phi(\mathbf{m}) = \sum_{s=1}^S \sum_{r=1}^R \int_0^T \|\Delta \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t)\|^2 dt, \quad (1)$$

where $\Delta \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t) = \mathbf{d}_{obs}(\mathbf{x}_s, \mathbf{x}_r, t) - \mathbf{d}_{syn}(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{m})$ is the data residual, \mathbf{x}_s ($s = 1, \dots, S$) and \mathbf{x}_r ($r = 1, \dots, R$) indicate source and receiver locations, S and R are the maximum source and receiver indexes, and T represents maximum recording time. In order to solve the inverse problem, the next step is to find the model which minimizes the adopted cost function. Within the Newton optimization framework, Fréchet derivative of the misfit function can be written as:

$$\nabla_{\mathbf{m}} \Phi(\mathbf{m}) = \sum_{s=1}^S \sum_{r=1}^R \int_0^T \int_{\Omega(\mathbf{x})} \nabla_{\mathbf{m}(\mathbf{x})} \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{m}) [\mathbf{d}_{obs}(\mathbf{x}_s, \mathbf{x}_r, t) - \mathbf{d}_{syn}(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{m})] d^3 \mathbf{x} dt, \quad (2)$$

where $\nabla_{\mathbf{m}(\mathbf{x})} \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t; \mathbf{m})$ indicates the Fréchet derivative wavefield (or Jacobian matrix) recorded at the receiver locations due to model perturbation at position \mathbf{x} . Generally, for the large-scale inverse problems in seismic tomography, constructing the Jacobian matrix explicitly is considered to be unaffordable. The adjoint-state method is always employed to construct the gradients by cross-correlating the forward modelling wavefield and backward propagated data residual wavefield. Considering the general anisotropic model, characterized by ρ and c_{ijkl} , where ρ and c_{ijkl} ($i, j, k, l = 1, 2, 3$) denote density and elastic constant tensor with Einstein summation convention, the perturbed n th displacement field due to model perturbation $\Delta \rho$ and Δc_{ijkl} at position \mathbf{x} is expressed as:

$$\begin{aligned} \nabla_{\mathbf{m}(\mathbf{x})} d_n(\mathbf{x}_s, \mathbf{x}_r, t') = & - \int_0^T \int_{\Omega(\mathbf{x})} [\Delta \rho(\mathbf{x}) G_{ni}(\mathbf{x}, \mathbf{x}_r, T - t') \partial_{t'}^2 u_i(\mathbf{x}, \mathbf{x}_s, t') \\ & + \Delta c_{ijkl}(\mathbf{x}) \partial'_j G_{ni}(\mathbf{x}, \mathbf{x}_r, T - t') \partial'_k u_l(\mathbf{x}, \mathbf{x}_s, t')] d^3 \mathbf{x} dt', \end{aligned} \quad (3)$$

where $G_{ni}(\mathbf{x}, \mathbf{x}_r, T - t')$ is the Green's tensor. Substituting equation (3) into equation (7) gives the gradients of the model parameters:

$$\begin{aligned} \nabla_{\mathbf{m}} \Phi(\mathbf{m}) = & - \sum_{s=1}^S \sum_{r=1}^R \int_0^T [d_i^{obs}(\mathbf{x}_s, \mathbf{x}_r, t) - d_i^{syn}(\mathbf{x}_s, \mathbf{x}_r, t)] \\ & \times \int_0^t \int_{\Omega(\mathbf{x})} [\delta \rho(\mathbf{x}) G_{ni}(\mathbf{x}, \mathbf{x}_r, T - t') \partial_{t'}^2 u_i(\mathbf{x}_s, \mathbf{x}, t') \\ & + \Delta c_{ijkl}(\mathbf{x}) \partial'_j G_{ni}(\mathbf{x}, \mathbf{x}_r, t - t') \partial'_k u_l(\mathbf{x}_s, \mathbf{x}, t')] d^3 \mathbf{x} dt' dt. \end{aligned} \quad (4)$$

Considering the reciprocity of the Green's tensor, we can define the adjoint wavefield as:

$$u_k^\dagger(\mathbf{x}', t') = \int_0^{t'} \int_{\Omega} G_{ki}(\mathbf{x}', \mathbf{x}; t' - t) f_i^\dagger(\mathbf{x}, t) d^3\mathbf{x} dt, \quad (5)$$

where $f_i^\dagger(\mathbf{x}, t)$ is the adjoint source:

$$f_i^\dagger(\mathbf{x}, t) = \sum_{s=1}^S \sum_{r=1}^R [d_i^{obs}(\mathbf{x}_s, \mathbf{x}_r, T - t) - d_i^{sym}(\mathbf{x}_s, \mathbf{x}_r, T - t)] \delta(\mathbf{x} - \mathbf{x}_r). \quad (6)$$

The Fréchet derivative of the misfit function can be re-written as:

$$\nabla_{\mathbf{m}} \Phi(\mathbf{m}) = \int_{\Omega(\mathbf{x})} [K_\rho(\mathbf{x}) \Delta\rho(\mathbf{x}) + K_{c_{ijkl}} \Delta c_{ijkl}(\mathbf{x})] d^3\mathbf{x}, \quad (7)$$

The 3-D waveform misfit kernels K_ρ and $K_{c_{ijkl}}$ represent Fréchet derivatives with respect to density and the elastic parameters are defined by

$$K_\rho(\mathbf{x}) = - \int_0^T \Delta\rho(\mathbf{x}) \mathbf{u}^\dagger(\mathbf{x}, T - t) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t) dt, \quad (8)$$

$$K_{c_{ijkl}}(\mathbf{x}) = - \sum_{s=1}^S \sum_{r=1}^R \int_0^T \int \partial_j G_{ni}(\mathbf{x}_r, \mathbf{x}, T - t) \Delta c_{ijkl}(\mathbf{x}) \partial_k u_l(\mathbf{x}, \mathbf{x}_s, t) dt, \quad (9)$$

In general anisotropic media, the gradients for the 21 elastic constants are listed in Appendix A. In 3D HTI media, 5 elastic constants (c_{11} , c_{13} , c_{33} , c_{44} and c_{55}) are needed for

describing the media. Gradients for the 5 elastic constants can be expressed as:

$$\begin{aligned}
K_{c_{11}}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int \partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \Delta c_{11}(\mathbf{x}) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t) dt, \\
K_{c_{13}}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int [\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) (\partial_y u_y(\mathbf{x}, \mathbf{x}_s, t) + \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t)) \\
&\quad + (\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t)) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)] \Delta c_{13}(\mathbf{x}) dt, \\
K_{c_{33}}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int (\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t)) \\
&\quad \times (\partial_y u_y(\mathbf{x}, \mathbf{x}_s, t) + \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t)) \Delta c_{33}(\mathbf{x}) dt, \\
K_{c_{44}}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int [2\partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t) \\
&\quad + 2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t) - (\partial_y u_z(\mathbf{x}, \mathbf{x}_s, t) + \partial_z u_y(\mathbf{x}, \mathbf{x}_s, t)) \\
&\quad \times (\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_z G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t))] \Delta c_{44}(\mathbf{x}) dt, \\
K_{c_{55}}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int (\partial_z G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t)) \\
&\quad \times (\partial_z u_x(\mathbf{x}, \mathbf{x}_s, t) + \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t)) (\partial_y G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t)) \\
&\quad \times (\partial_y u_x(\mathbf{x}, \mathbf{x}_s, t) + \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t)) \Delta c_{55}(\mathbf{x}) dt, \\
K_{\rho}(\mathbf{x}) &= - \sum_{s=1}^S \sum_{r=1}^R \int (\partial_t^2 u_x(\mathbf{x}, \mathbf{x}_s, t) G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \\
&\quad + \partial_t^2 u_y(\mathbf{x}, \mathbf{x}_s, t) G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) + \partial_t^2 u_z(\mathbf{x}, \mathbf{x}_s, t) G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t)) \Delta \rho(\mathbf{x}) dt.
\end{aligned} \tag{10}$$

We can also describe the 3D HTI media using the velocity parameterization (α , β , ρ , ε , δ and γ). The gradients for these parameters can be obtained with the chain rule:

$$\begin{aligned}
 K_\alpha(\mathbf{x}) &= 2\rho\alpha K_{c_{33}}(\mathbf{x}) \\
 &\quad + 2\rho\alpha(2\varepsilon + 1)K_{c_{11}}(\mathbf{x}) + 2\rho \frac{2\delta\alpha^3 - \delta\alpha\beta^2 + 2\rho\alpha(\alpha^2 - \beta^2)}{\sqrt{2\delta\alpha^2(\alpha^2 - \beta^2)^2 + (\alpha^2 - \beta^2)^2}} K_{c_{13}}(\mathbf{x}), \\
 K_\beta(\mathbf{x}) &= 2\rho\beta K_{c_{55}}(\mathbf{x}) + \frac{2\rho\beta}{(2\gamma + 1)} K_{c_{44}}(\mathbf{x}) \\
 &\quad + \frac{-2\delta\rho\alpha^2\beta - 2\rho\beta(\alpha^2 - \beta^2)}{\sqrt{2\delta\alpha^2(\alpha^2 - \beta^2) + (\alpha^2 - \beta^2)^2}} K_{c_{13}}(\mathbf{x}) - 2\rho\beta K_{c_{13}}(\mathbf{x}), \\
 K_\varepsilon(\mathbf{x}) &= 2\rho\alpha^2 K_{c_{11}}(\mathbf{x}), \\
 K_\delta(\mathbf{x}) &= \frac{\rho\alpha^2(\alpha^2 - \beta^2)}{\sqrt{2\alpha^2\delta(\alpha^2 - \beta^2) + (\alpha^2 - \beta^2)^2}} K_{c_{13}}(\mathbf{x}), \\
 K_\gamma(\mathbf{x}) &= -\frac{2\rho\beta^2}{(1 + 2\gamma)^2} K_{c_{44}}(\mathbf{x}), \\
 K'_\rho(\mathbf{x}) &= K_\rho(\mathbf{x}) + \alpha^2 K_{c_{33}}(\mathbf{x}) + \beta^2 K_{c_{55}}(\mathbf{x}) + \frac{\beta^2}{1 + 2\gamma} K_{c_{55}}(\mathbf{x}) \\
 &\quad + \left(\sqrt{2\alpha^2\delta(\alpha^2 - \beta^2) + (\alpha^2 - \beta^2)^2} - \beta^2 \right) K_{c_{13}}(\mathbf{x}).
 \end{aligned} \tag{11}$$

Within the Newton framework, the search direction $\Delta\mathbf{m}$ can be obtained by solving the following Newton linear system:

$$\mathbf{H}\Delta\mathbf{m} = -\nabla\Phi, \tag{12}$$

where \mathbf{H} is the Hessian matrix, the second derivative of the misfit with respect to the model parameters:

$$\begin{aligned}
 \mathbf{H} &= \int_{\Omega(\mathbf{x})} \int_{\Omega(\mathbf{x}')} \Delta\rho(\mathbf{x}) H_{\rho\rho}(\mathbf{x}, \mathbf{x}') \Delta\rho(\mathbf{x}') + \Delta\rho(\mathbf{x}') H_{\rho c_{ijkl}}(\mathbf{x}, \mathbf{x}') :: \Delta c_{ijkl}(\mathbf{x}') \\
 &\quad + \Delta c_{ijkl}(\mathbf{x}) H_{c_{ijkl}c_{i'j'k'l'}}(\mathbf{x}, \mathbf{x}') :: \Delta c_{i'j'k'l'}(\mathbf{x}') d^3\mathbf{x} d^3\mathbf{x}'.
 \end{aligned} \tag{13}$$

The Hessian matrix measures the amount of inter-parameter "blurring" produced by the adjoint operator. The Hessian matrix works like a blurring filter to the model perturbation, which form the blurred updates (gradient vectors) for the model parameters. The inverse Hessian matrix works as de-blurring filter to the gradient and remove the finite-frequency effects (Pratt et al., 1998; Pan et al., 2014b,a, 2015). In multi-parameter FWI, the multi-parameter Hessian has a block structure. The multi-parameter Gauss-Newton Hessian $\tilde{\mathbf{H}}$ has a similar structure with multi-parameter full Hessian but only considers the first-order scattering effects and ignores the second-order scattering effects. For the general anisotropic media, the multi-parameter Gauss-Newton Hessian can be expressed as:

$$\begin{aligned}
 \tilde{H}_{\rho\rho}(\mathbf{x}, \mathbf{x}') &= \sum_{s=1}^S \sum_{r=1}^R \int \int \int G_{ni}(\mathbf{x}_r, \mathbf{x}, t - t') \partial_{t'}^2 u_i(\mathbf{x}, \mathbf{x}_s, t') \\
 &\quad \times G_{ni'}(\mathbf{x}_r, \mathbf{x}', t - t'') \partial_{t''}^2 u_{i'}(\mathbf{x}', \mathbf{x}_s, t'') dt' dt'' dt,
 \end{aligned} \tag{14}$$

$$\tilde{H}_{\rho c}(\mathbf{x}, \mathbf{x}') = \sum_{s=1}^S \sum_{r=1}^R \int \int \int G_{ni}(\mathbf{x}_r, \mathbf{x}, t - t') \partial_{t'}^2 u_i(\mathbf{x}, \mathbf{x}_s, t') \partial_j G_{ni'}(\mathbf{x}_r, \mathbf{x}', t - t'') \partial_k u_l(\mathbf{x}', \mathbf{x}_s, t'') dt' dt'' dt, \quad (15)$$

$$\tilde{H}_{c\rho}(\mathbf{x}, \mathbf{x}') = \sum_{s=1}^S \sum_{r=1}^R \int \int \int \partial_j G_{ni}(\mathbf{x}_r, \mathbf{x}, t - t') \partial_k u_l(\mathbf{x}, \mathbf{x}_s, t') G_{ni'}(\mathbf{x}_r, \mathbf{x}', t - t'') \partial_{t''}^2 u_{i'}(\mathbf{x}', \mathbf{x}_s, t'') dt' dt'' dt, \quad (16)$$

$$\tilde{H}_{cc}(\mathbf{x}, \mathbf{x}') = \sum_{s=1}^S \sum_{r=1}^R \int \int \int \partial_j G_{ni}(\mathbf{x}_r, \mathbf{x}, t - t') \partial_k u_l(\mathbf{x}, \mathbf{x}_s, t') \partial_{j'} G_{ni'}(\mathbf{x}_r, \mathbf{x}', t - t'') \partial_{k'} u_{l'}(\mathbf{x}', \mathbf{x}_s, t'') dt' dt'' dt, \quad (17)$$

where when $\mathbf{x} = \mathbf{x}'$, we can obtain the off-diagonal elements in the block matrices and when $\mathbf{x} \neq \mathbf{x}'$, we can obtain the diagonal elements of the block matrices.

Considering a 3D subsurface model with $M = N_x \times N_y \times N_z$ nodes and N_P physical parameters ($\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_P$) are assigned to describe the properties of each node. The multi-parameter Hessian is a $M \times N_P$ by $M \times N_P$ square and symmetric matrix with N_P diagonal blocks and $N_P(N_P - 1)$ off-diagonal blocks. Each block is a M by M square matrix. Figure 1 describes the structure the multi-parameter Gauss-Newton Hessian.

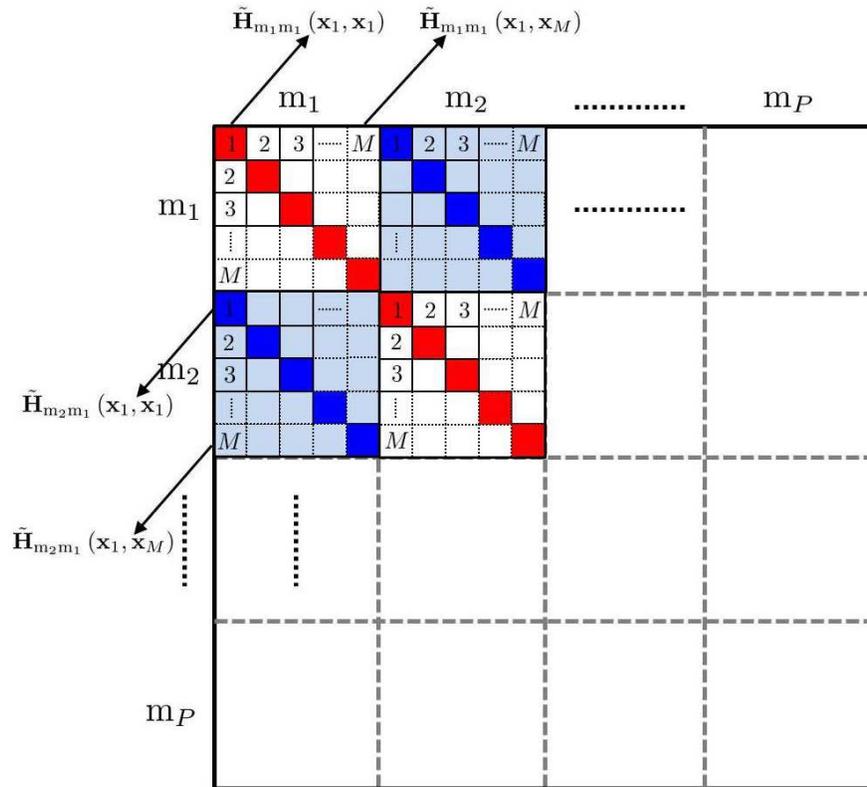


FIG. 1. The schematic diagram of multi-parameter Gauss-Newton Hessian.

The elements in the multi-parameter Gauss-Newton Hessian matrix can be classified into 4 types. Each element in the multi-parameter Gauss-Newton Hessian can be denoted as $\tilde{H}_{\mathbf{m}_p \mathbf{m}_q}(\mathbf{x}, \mathbf{x}')$, where \mathbf{x} and \mathbf{x}' represent the row and column indexes of the element. When $\mathbf{m}_p = \mathbf{m}_q$ and $\mathbf{x} = \mathbf{x}'$, we can obtain the diagonal elements in the diagonal blocks, which measure the geometrical spreading, as indicated in the red grids of Figure 12 (e.g., $\tilde{H}_{\mathbf{m}_1 \mathbf{m}_1}(\mathbf{x}_1, \mathbf{x}_1)$). When $\mathbf{m}_p = \mathbf{m}_q$ and $\mathbf{x} \neq \mathbf{x}'$, we can obtain the off-diagonal elements in the diagonal blocks, which measure the space trade-off, as indicated by the white grids of Figure 12 (e.g., $\tilde{H}_{\mathbf{m}_1 \mathbf{m}_1}(\mathbf{x}_1, \mathbf{x}_M)$). When $\mathbf{m}_p \neq \mathbf{m}_q$ and $\mathbf{x} = \mathbf{x}'$, we obtain the diagonal elements in the off-diagonal blocks, which measure the parameter trade-off, as indicated by the blue grids in Figure 12 (e.g., $\tilde{H}_{\mathbf{m}_1 \mathbf{m}_2}(\mathbf{x}_1, \mathbf{x}_1)$). When $\mathbf{m}_p \neq \mathbf{m}_q$ and $\mathbf{x} \neq \mathbf{x}'$, we obtain the off-diagonal elements in the off-diagonal blocks, which measure the space and parameter trade-off, as indicated by the light blue grids in Figure 12 (e.g., $\tilde{H}_{\mathbf{m}_1 \mathbf{m}_2}(\mathbf{x}_1, \mathbf{x}_M)$). One row/column in multi-parameter Gauss-Newton Hessian indicates the correlation of Fréchet derivative wavefield due to one physical parameter at a specific position with the Fréchet derivative wavefields due to other physical parameters at other locations. For example, the third column in the multi-parameter Gauss-Newton Hessian shown in Figure 1 can be written as:

$$\tilde{H}_{\mathbf{m}_3}(\mathbf{x}_3) = \sum_{r=1}^R \sum_{s=1}^S \sum_{p=1}^P \int_{\Omega(\mathbf{x})} \nabla_{\mathbf{m}_p(\mathbf{x})} \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t) \nabla_{\mathbf{m}(\mathbf{x}_3)} \mathbf{d}(\mathbf{x}_s, \mathbf{x}_r, t) d^3 \mathbf{x} dt \quad (18)$$

The following figure show the physical interpretation of the elements in multi-parameter Gauss-Newton Hessian.

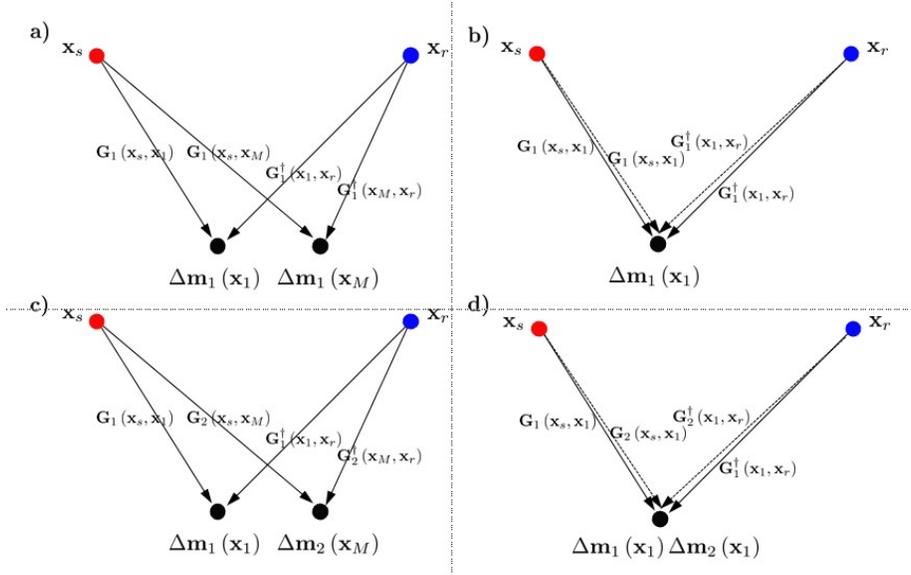


FIG. 2. Physical interpretation of the multi-parameter Gauss-Newton Hessian. (a) and (b) show the off-diagonal and diagonal elements in the diagonal blocks of the multi-parameter Gauss-Newton Hessian. (c) and (d) show the off-diagonal and diagonal elements in the off-diagonal blocks of the multi-parameter Gauss-Newton Hessian.

PARAMETER CROSSTALK DIFFICULTY IN MULTI-PARAMETER FWI

Reconstructing multiple physical parameters using FWI suffers from parameter crosstalk difficulty. Perturbation of one physical parameter will map into the updates for other phys-

ical parameters resulting from the coupling effects between these different physical parameters, which increase the non-linearity of the inverse problem significantly. For understanding the parameter crosstalk artifacts in multi-parameter FWI, we can first review the Newton equation system for multi-parameter inverse problem. Considering that two different physical parameters \mathbf{m}_θ and \mathbf{m}_ϑ are inverted simultaneously, the search directions $\Delta\mathbf{m}_\theta$ and $\Delta\mathbf{m}_\vartheta$ can be obtained by solving the following equation (19):

$$\begin{pmatrix} \mathbf{H}_{\theta\theta} & \mathbf{H}_{\theta\vartheta} \\ \mathbf{H}_{\vartheta\theta} & \mathbf{H}_{\vartheta\vartheta} \end{pmatrix} \begin{pmatrix} \Delta\mathbf{m}_\theta \\ \Delta\mathbf{m}_\vartheta \end{pmatrix} = - \begin{pmatrix} \nabla_\theta\Phi \\ \nabla_\vartheta\Phi \end{pmatrix}, \quad (19)$$

where $\nabla_\theta\Phi$ and $\nabla_\vartheta\Phi$ represent the corresponding gradient vectors. As we can see, the multi-parameter Hessian consists of four block matrices. The two diagonal blocks $\mathbf{H}_{\theta\theta}$ and $\mathbf{H}_{\vartheta\vartheta}$ measure the spatial correlations of the same physical parameter. The off-diagonal blocks $\mathbf{H}_{\theta\vartheta}$ and $\mathbf{H}_{\vartheta\theta}$ measure the trade-offs between model parameters θ and ϑ . Equation (19) can be re-written in an integral formulation:

$$\nabla_\theta\Phi = - \int_{\Omega(\mathbf{x})} \int_{\Omega(\mathbf{x}')} H_{\theta\theta}(\mathbf{x}, \mathbf{x}') \Delta m_\theta(\mathbf{x}') d\mathbf{x}' d\mathbf{x} - \int_{\Omega(\mathbf{x})} \int_{\Omega(\mathbf{x}')} H_{\theta\vartheta}(\mathbf{x}, \mathbf{x}') \Delta m_\vartheta(\mathbf{x}') d\mathbf{x}' d\mathbf{x}, \quad (20)$$

$$\nabla_\vartheta\Phi = - \int_{\Omega(\mathbf{x})} \int_{\Omega(\mathbf{x}')} H_{\vartheta\theta}(\mathbf{x}, \mathbf{x}') \Delta m_\theta(\mathbf{x}') d\mathbf{x}' d\mathbf{x} - \int_{\Omega(\mathbf{x})} \int_{\Omega(\mathbf{x}')} H_{\vartheta\vartheta}(\mathbf{x}, \mathbf{x}') \Delta m_\vartheta(\mathbf{x}') d\mathbf{x}' d\mathbf{x}. \quad (21)$$

We observe that the model perturbation $\Delta\mathbf{m}_\vartheta$ maps into the gradient update $\nabla_\theta\Phi$ and the model perturbation $\Delta\mathbf{m}_\theta$ maps into the gradient update $\nabla_\vartheta\Phi$, as indicated by the second and first terms in equations (20) and (21) respectively. As we can see, the gradient updates suffer from parameter crosstalk artifacts, which are formed by blurring the model perturbation with the off-diagonal blocks of the multi-parameter Hessian. Features that seem well reconstructed may be accidental artifacts due to other parameters.

Different parameterizations have different resolving abilities. It is necessary and essential to quantify the resolving ability of a parameter class for avoiding parameter crosstalk. The parameterization with weakest ambiguities should be selected for reconstructing the multiple physical parameters.

Parameter resolution analysis with scattering patterns

In recent years, researchers devote intensive efforts for understanding the resolution abilities of different parameterizations for multi-parameter FWI. These studies are carried out based on the so called "scattering pattern". The scattering pattern evaluates the amplitude variation of the Fréchet derivative wavefield due to the perturbation of one type of physical parameter with varying the scattering or azimuthal angles. Identical or nearly identical variations are one of the key mechanisms of parameter crosstalk in multi-parameter FWI (Operto et al., 2013). The overlapping the scattering patterns due to different physical parameters are used to evaluate the ambiguity between these two types of parameters.

However, these parameter resolution studies based on scattering patterns suffer from several drawbacks because of its inherent defects. First, the background model is assumed to be isotropic and homogeneous (as shown in Figure 3), which means that the complex geological structures and complex physics are not considered. Second, it assumes that the an incident plane-wave with far-field approximation is scattered due to a local point heterogeneity or a horizontal reflector. This means that it does not consider the influences of geometrical spreading and complex model perturbations. Third, only the amplitudes of the analytic Fréchet derivative wavefields due to different physical parameters at the same location are considered. Thus, the finite-frequency effects and travel time information are ignored. Furthermore, these studies ignore the parameter crosstalk due to different physical parameters at adjacent locations, which actually are especially important when low wavenumber components of the model parameters are deficient. Further studies reveal that the parameter resolution studies based on scattering patterns can be interpreted as constructing an asymptotic approximation of the multi-parameter Gauss-Newton Hessian with a perfect survey. Overlapping the scattering patterns due to different physical parameters at position \mathbf{x} is equivalent to correlating two Fréchet derivative wavefields at position \mathbf{x} , which is equivalent to one element in the diagonal elements of the off-diagonal blocks in multi-parameter Gauss-Newton Hessian, as shown in Figure 1. However, we also observe that the scattering patterns analysis only consider the coupling effects in the wavefield level, which are measured by the off-diagonal blocks of the multi-parameter Gauss-Newton Hessian. It is more appropriate to measure the coupling effects in the objective function level, which are measured by the multi-parameter full Hessian. In this condition, the parameter crosstalk due to multi-parameter second-order scattering effects should also be evaluated.

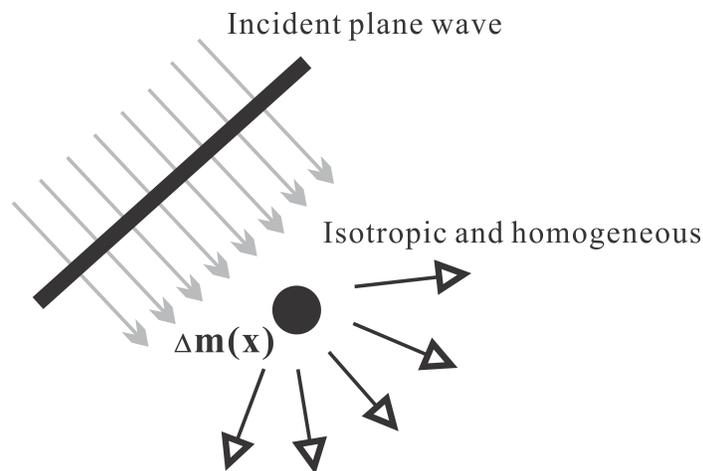


FIG. 3. The schematic diagram of scattering pattern.

To evaluate the resolution abilities of different parameter classes accurately and fully, in this research, we propose to quantify the parameter resolution with the multi-parameter Hessian, as discussed in the following section.

Quantify the parameter resolution with multi-parameter Hessian

As discussed in the previous sections, the off-diagonal blocks of the multi-parameter full Hessian measure the parameter trade-off for multi-parameter FWI. To quantify the

space and parameter resolution completely, a multi-parameter full Hessian matrix is needed. However, constructing the multi-parameter full Hessian explicitly is considered to be unaffordable for large-scale inverse problems especially for 3D model. In this section, we will show how to extract partial Hessian information for quantifying the parameter resolution.

Block diagonal approximation of the multi-parameter Gauss-Newton Hessian

Employing the adjoint-state technique, we can construct the diagonal elements of the off-diagonal blocks in the multi-parameter Gauss-Newton Hessian explicitly for a specific model, which can be defined as the block diagonal approximation of a multi-parameter Gauss-Newton Hessian $\tilde{\mathbf{H}}_D$. The elements of this type of multi-parameter Hessian approximation measures correlations of the Fréchet derivative wavefields due to different physical parameters at one location in the interested model. Compared to the analysis based on the scattering patterns, the block diagonal approximation $\tilde{\mathbf{H}}_D$ take the geometrical spreading, finite-frequency, and complex model into consideration.

Furthermore, it can also be used as an effective preconditioner, which can help to compensate the geometrical spreading and mitigate the parameter crosstalk artifacts in the inversion process (Innanen, 2014b). Obtaining the block diagonal approximation $\tilde{\mathbf{H}}_D$ needs constructing the receiver-side Green's functions from all of the receivers explicitly, which is also an expensive task. A better choice is to ignore the receiver-side Green's functions and the source-side Green's functions can be calculated in the forward modelling process without additional cost.

Multi-parameter point spread function

Even though, we are able to construct the diagonal elements of the block matrices in a multi-parameter Gauss-Newton Hessian directly. It is still not able to quantify the parameter crosstalk due the perturbations of different physical parameters at adjacent locations. Considering an arbitrary element $H_{\theta\vartheta}(\mathbf{x}, \mathbf{x}')$ in multi-parameter Hessian:

$$H_{\theta\vartheta}(\mathbf{x}, \mathbf{x}') = \nabla_{\theta(\mathbf{x})} \nabla_{\vartheta(\mathbf{x}')} \Phi, \quad (22)$$

which measures the coupling effects of model parameter \mathbf{m}_θ at position \mathbf{x} and model parameter \mathbf{m}_ϑ at position \mathbf{x}' in the misfit function level. The column in the multi-parameter Hessian corresponding to model parameter \mathbf{m}_θ at position \mathbf{x} can be written as:

$$\mathbf{H}_{\theta\vartheta}(\mathbf{x}) = \int_{\Omega(\mathbf{x}')} H_{\theta\vartheta}(\mathbf{x}, \mathbf{x}') d^3\mathbf{x}', \quad (23)$$

which measures the coupling effects of model parameter \mathbf{m}_θ at \mathbf{x} with the model parameter \mathbf{m}_ϑ at all other positions in the interested volume Ω . Considering equations (20) and (21), suppose that one point perturbation of model parameter \mathbf{m}_θ is located at position $\tilde{\mathbf{x}}$ ($\Delta\mathbf{m}_\theta = \delta(\mathbf{x} - \tilde{\mathbf{x}})$) and there is no perturbation for model parameter \mathbf{m}_ϑ ($\Delta\mathbf{m}_\vartheta = 0$). The above

equations become:

$$\begin{aligned}\nabla_{\theta}\Phi &= -\int_{\Omega(\mathbf{x})}\int_{\Omega(\mathbf{x}')}H_{\theta\theta}(\mathbf{x},\mathbf{x}')\delta(\mathbf{x}'-\tilde{\mathbf{x}})d^3\mathbf{x}'d^3\mathbf{x}=-\int_{\Omega(\mathbf{x})}H_{\theta\theta}(\mathbf{x},\tilde{\mathbf{x}})d^3\mathbf{x}=-\mathbf{H}_{\theta\theta}(\tilde{\mathbf{x}}), \\ \nabla_{\vartheta}\Phi &= -\int_{\Omega(\mathbf{x})}\int_{\Omega(\mathbf{x}')}H_{\vartheta\theta}(\mathbf{x},\mathbf{x}')\delta(\mathbf{x}'-\tilde{\mathbf{x}})d^3\mathbf{x}'d^3\mathbf{x}=-\int_{\Omega(\mathbf{x})}H_{\vartheta\theta}(\mathbf{x},\tilde{\mathbf{x}})d^3\mathbf{x}=-\mathbf{H}_{\vartheta\theta}(\tilde{\mathbf{x}}).\end{aligned}\tag{24}$$

The gradient vectors are blurred representations of the model perturbations with limited wavenumber content and resolution. As we see, gradient update $\nabla_{\theta}\Phi$ due to the point-localized perturbation $\Delta\mathbf{m}_{\theta}$ at position $\tilde{\mathbf{x}}$ is a conservative estimate of spike model perturbation blurred by the diagonal block $\mathbf{H}_{\theta\theta}$ of the multi-parameter Hessian. Thus, it only suffers from spatial trade-off. However, for the gradient update $\nabla_{\vartheta}\Phi$, it is a blurred estimation by off-diagonal block $\mathbf{H}_{\vartheta\theta}$ for model parameter \mathbf{m}_{ϑ} . Because there is no model perturbation in model parameter \mathbf{m}_{ϑ} , the point-localized model perturbation in \mathbf{m}_{θ} maps into the gradient update for \mathbf{m}_{ϑ} , which is defined as parameter crosstalk artifacts. The off-diagonal blocks in multi-parameter Hessian introduce unwanted updates of the model parameters that have initially not been perturbed.

Furthermore, we notice that the gradient vectors $\nabla_{\theta}\Phi$ and $\nabla_{\vartheta}\Phi$ due to a point-localized model perturbation is proportional to one column in the multi-parameter Hessian specified by the position $\tilde{\mathbf{x}}$, which is known as multi-parameter point spread function (MPSF) in seismic tomography. Here, in this research, we employ the multi-parameter point spread function, defined as parameter resolution kernel, to quantify the resolution abilities of different parameterizations. Point spread functions corresponding to different physical parameters at different locations will be quite different. These phenomenons will not be described by the block diagonal Hessian approximation, which does not consider the energy away from the center of the point scatter.

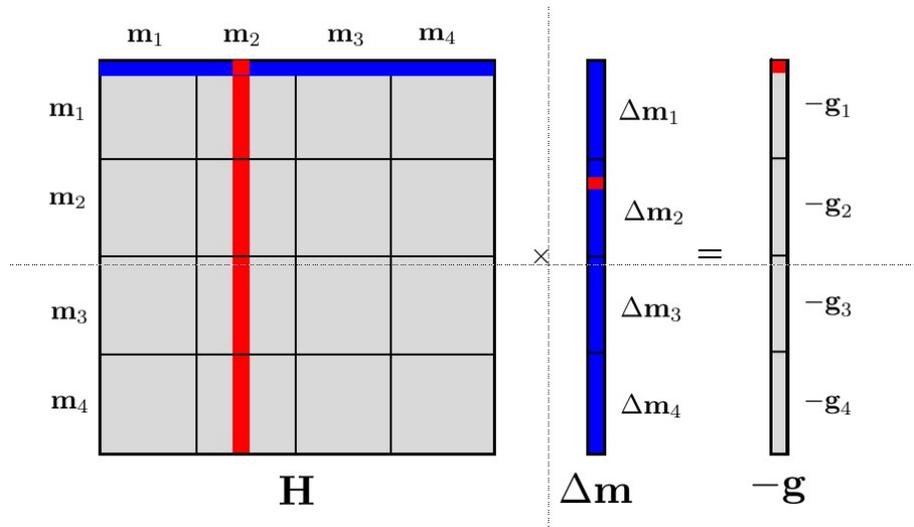


FIG. 4. The schematic diagram of spike probing.

One problem of the multi-parameter PSF is its local signature, which means that we can only quantify the parameter resolution at a local position. If the full blurring effects are to be

understood, then a spike test for each model parameter would need to be run independently, and the results collated. A single spike test of this sort would retrieve the information equivalent of only one column or row of the Hessian matrix \mathbf{H} . Ideally, for a specific model, the PSFs should be computed for every model parameter at every position within the volume of interest, which will be prohibitively expensive. Reviewing the following equation:

$$\nabla\Phi = - \int H(\mathbf{x}, \mathbf{x}') \Delta m(\mathbf{x}') d\mathbf{x}' d\mathbf{x}. \quad (25)$$

The volume of the inter-parameter point spread functions describe the strength of apparent heterogeneities in one parameter that result from trade-offs from other parameter, and vice versa. We notice that the above integral is equivalent to sampling the Hessian with the model perturbation vector. The Hessian can be sampled by a point-localized heterogeneity or a series of discrete model perturbations. Thus, applying the Hessian to the whole model perturbation vector can be approximated by applying the Hessian to a small number of model perturbations. By sampling Hessian with a suitable set of model perturbations, we can gather as much second-derivative information as needed for our purposes, though at the expense of potentially prohibitive computational requirements. It is therefore the purpose of this paper to develop a sampling strategy of the Hessian that operates with as few model perturbations as possible while leading to an approximation of Hessian that is physically meaningful and interpretable (Fichtner and van Leeuwen, 2015; Zhu and Fomel, 2016).

The multi-parameter Hessian has the ability to quantify both spatial and inter-parameter resolution. The power of the point spread functions lies in its applicability to infinite-dimensional model spaces that we typically encounter in tomographic problems where we seek quantities that are continuously distributed in space.

NUMERICAL EXPERIMENTS

In this numerical section, we first calculate the sensitivity kernels for the elastic and anisotropic parameters. Then, we construct the diagonal elements in the block matrices of the multi-parameter Gauss-Newton Hessian. We also give the 3D multi-parameter point spread functions of anisotropic parameters.

Sensitivity kernels of the model parameters in elastic and anisotropic media

In this example, we calculate the 3D sensitivity kernels for the anisotropic and elastic parameters in 3D HTI media based on adjoint-state method. We assume the anisotropic model parameters are: P-wave velocity $\alpha = 3500\text{m/s}$, S-wave velocity $\beta = 2000\text{m/s}$, density $\rho = 2000\text{kg/m}^3$, $\varepsilon = -0.14$, $\delta = -0.18$ and $\gamma = -0.15$. The initial model is isotropic: P-wave velocity $\alpha = 3200\text{m/s}$, S-wave velocity $\beta = 1800\text{m/s}$, density $\rho = 1800\text{kg/m}^3$. Figure 5 shows the 3D view of acquisition geometry. One source is location at (400m, 400m, 40m). One receiver is located at (400m, 400m, 760m). We use finite-difference method for solving the anisotropic and elastic wave equation with a standard stagger grid method. A CPML boundary condition is applied on all of the boundaries of the model. Note that in numerical experiments presented in this research, we used the pressure source for forward modelling and different types of sources should lead to different results.

Figures 6a-6c show the observed data (z component displacement field), synthetic data and adjoint source. Figures 7a-7f show the sensitivity kernels K_α , K_β , K_ρ , K_ε , K_δ and K_γ for P-wave velocity, S-wave velocity, density, ε , δ and γ respectively.

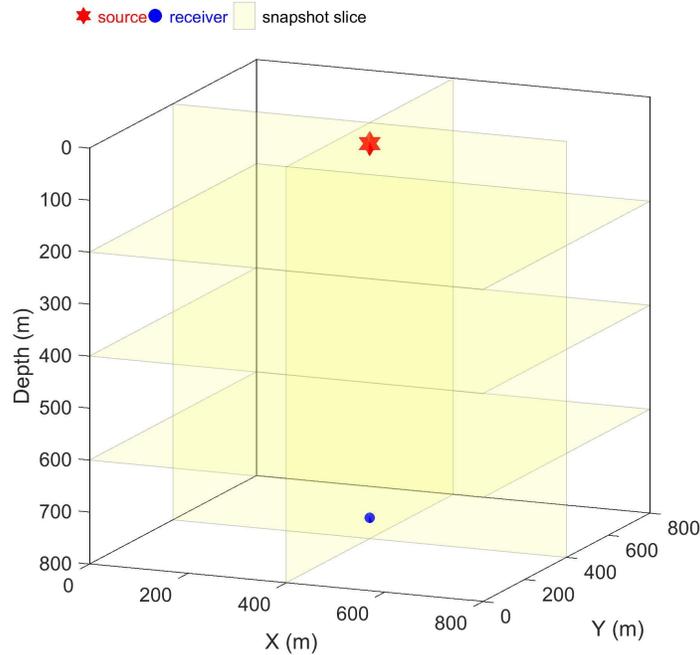


FIG. 5. The 3D acquisition geometry. The red star and blue circle represent the positions of source and receiver.

Diagonal elements of the block matrices in multi-parameter Gauss-Newton Hessian

As discussed in previous section, the diagonal elements in the diagonal blocks and off-diagonal blocks of the multi-parameter Gauss-Newton Hessian can be constructed explicitly using adjoint-state technique. Here, we give numerical examples to show the diagonal elements of the block matrices in multi-parameter Gauss-Newton Hessian for 3D anisotropic FWI. The first numerical example is gradients and diagonal Hessian for the elastic constants in 3D orthorhombic media. The 3D orthorhombic media is homogeneous with the elastic constants of $c_{11} = 25.34$ GPA, $c_{12} = 12.26$ GPA, $c_{13} = 13.26$ GPA, $c_{22} = 17.4$ GPA, $c_{23} = 11.58$ GPA, $c_{33} = 24.5$ GPA, $c_{44} = 5.68$ GPA, $c_{55} = 6.78$ GPA, $c_{66} = 5.56$ GPA. Figures 8a-8i show the diagonal elements of the diagonal blocks $\tilde{\mathbf{H}}_{c_{11}c_{11}}$, $\tilde{\mathbf{H}}_{c_{22}c_{22}}$, $\tilde{\mathbf{H}}_{c_{33}c_{33}}$, $\tilde{\mathbf{H}}_{c_{12}c_{12}}$, $\tilde{\mathbf{H}}_{c_{13}c_{13}}$, $\tilde{\mathbf{H}}_{c_{23}c_{23}}$, $\tilde{\mathbf{H}}_{c_{44}c_{44}}$, $\tilde{\mathbf{H}}_{c_{55}c_{55}}$, and $\tilde{\mathbf{H}}_{c_{66}c_{66}}$ for multi-parameter Hessian in orthorhombic media. Figures 9a-9i show the diagonal elements of the off-diagonal blocks $\tilde{\mathbf{H}}_{c_{11}c_{33}}$, $\tilde{\mathbf{H}}_{c_{22}c_{33}}$, $\tilde{\mathbf{H}}_{c_{12}c_{33}}$, $\tilde{\mathbf{H}}_{c_{13}c_{33}}$, $\tilde{\mathbf{H}}_{c_{23}c_{33}}$, $\tilde{\mathbf{H}}_{c_{44}c_{33}}$, $\tilde{\mathbf{H}}_{c_{55}c_{33}}$, and $\tilde{\mathbf{H}}_{c_{66}c_{33}}$, which measure the parameter trade-off of c_{33} with other elastic constants. Note that here we ignore the receiver-side Green functions and the amplitudes of the diagonal Hessian elements are normalized at the same depth slice.

From Figures 8 and 9, we can see that the energy distribution for different elastic constants in the diagonal elements of the multi-parameter Hessian are quite different. The energy distributions should match the scattering patterns of these elastic constants (Pan et al., 2016a). The diagonal elements of the multi-parameter Gauss-Newton Hessian are able to

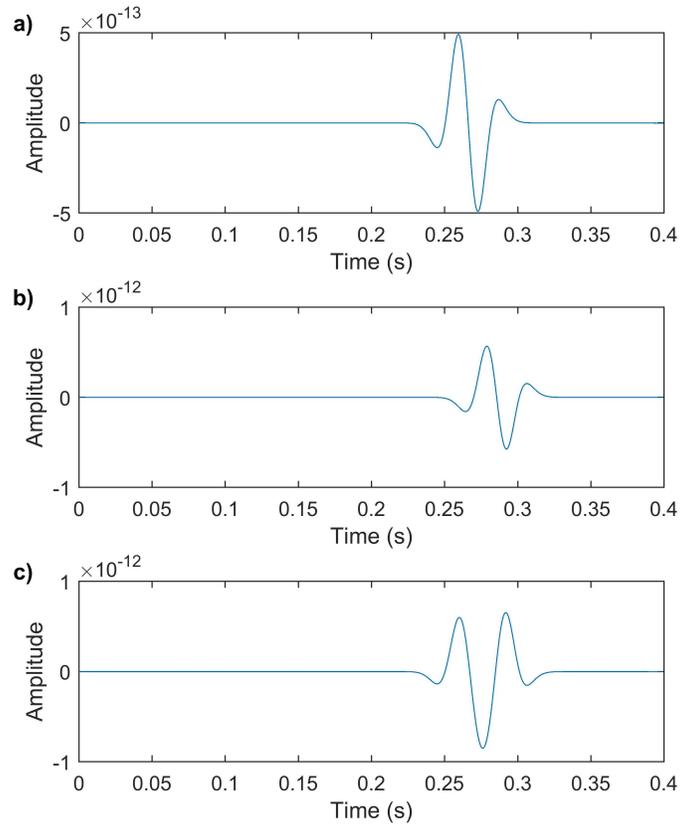


FIG. 6. (a), (b), and (c) show the observed data (z component displacement field), synthetic data and adjoint source.

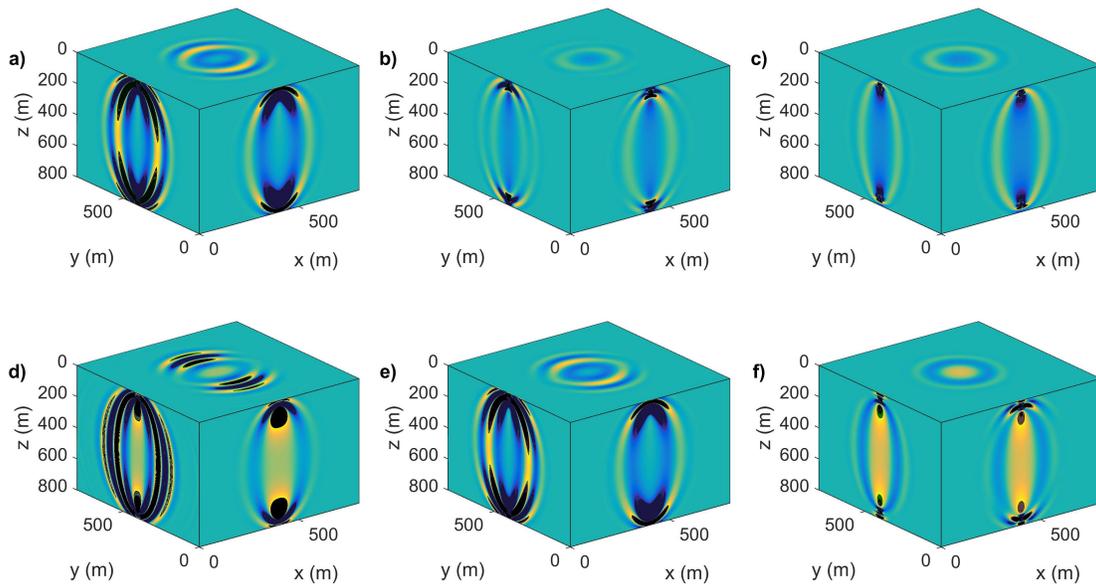


FIG. 7. (a), (b), (c), (d), (e) and (f) show the sensitivity kernels of P-wave velocity, S-wave velocity, density, ϵ , δ and γ .

quantify the parameter resolution by taking geometrical spreading and finite-frequency effects into consideration. They are not able to evaluate the parameter trade-off due different

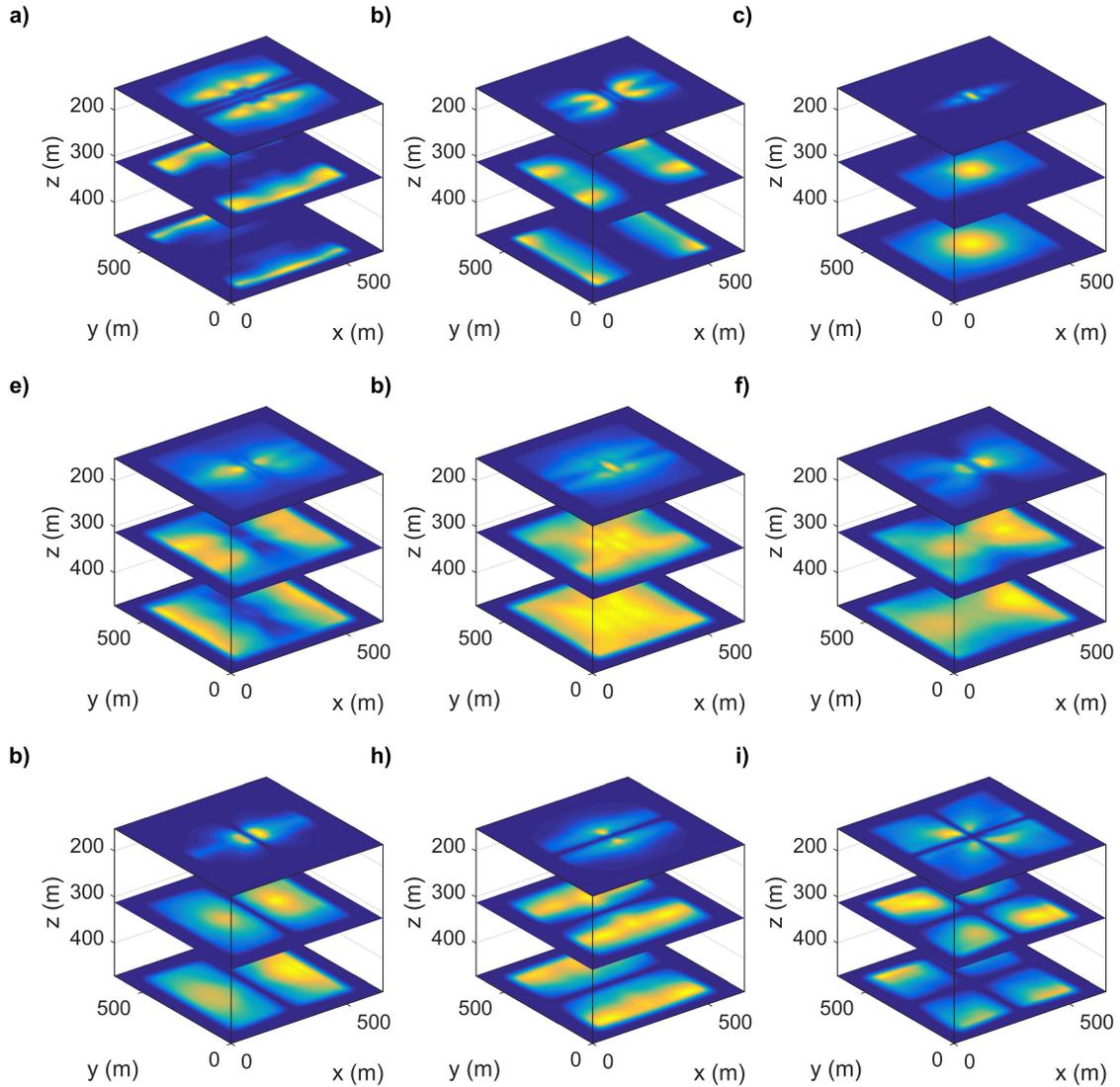


FIG. 8. (a)-(i) show the diagonal elements of the diagonal blocks $\tilde{\mathbf{H}}_{c_{11}c_{11}}$, $\tilde{\mathbf{H}}_{c_{22}c_{22}}$, $\tilde{\mathbf{H}}_{c_{33}c_{33}}$, $\tilde{\mathbf{H}}_{c_{12}c_{12}}$, $\tilde{\mathbf{H}}_{c_{13}c_{13}}$, $\tilde{\mathbf{H}}_{c_{23}c_{23}}$, $\tilde{\mathbf{H}}_{c_{44}c_{44}}$, $\tilde{\mathbf{H}}_{c_{55}c_{55}}$, and $\tilde{\mathbf{H}}_{c_{66}c_{66}}$.

physical parameters at different locations, which can be achieved by the columns or rows of the multi-parameter Hessian, as shown in the following examples.

Multi-parameter point spread function

In this section, we will first illustrate the parameter resolution kernels for elastic and isotropic parameters. Figure 10 shows the acquisition geometry. The model is isotropic and homogeneous with P-wave velocity $\alpha=2000$ m/s, $\beta=1400$ m/s and $\rho=1200$ kg/m³. A total of 64 sources are distributed regularly along the 4 boundaries of the model with a spacing of 60 m. A total of 192 receivers are arranged along 4 boundaries of the model regularly with a spacing of 60 m.

To calculate the columns of the multi-parameter Hessian, we first apply a +10% point

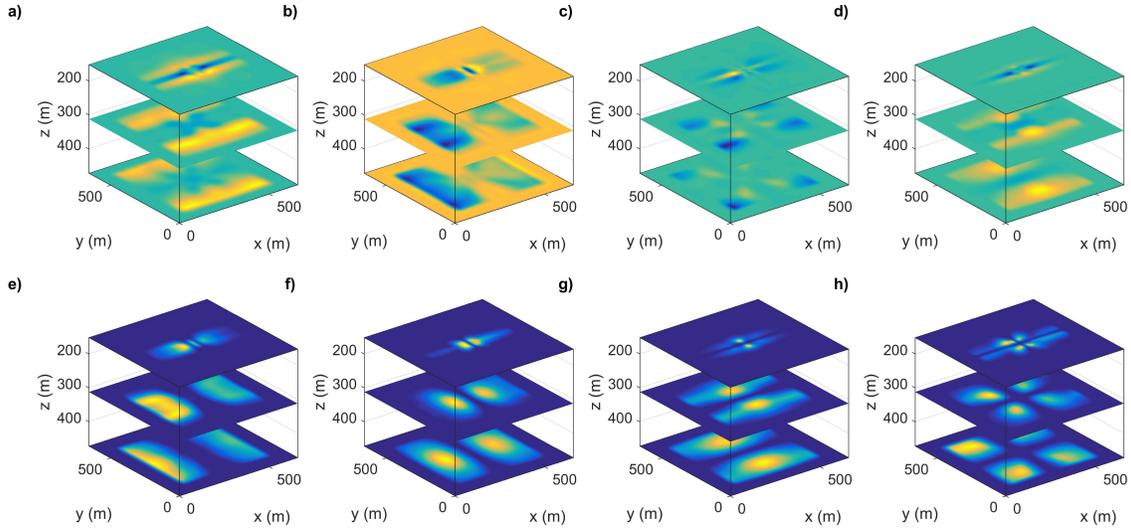


FIG. 9. (a)-(h) show the diagonal elements of the off-diagonal blocks $\tilde{\mathbf{H}}_{c_{11}c_{33}}$, $\tilde{\mathbf{H}}_{c_{22}c_{33}}$, $\tilde{\mathbf{H}}_{c_{12}c_{33}}$, $\tilde{\mathbf{H}}_{c_{13}c_{33}}$, $\tilde{\mathbf{H}}_{c_{23}c_{33}}$, $\tilde{\mathbf{H}}_{c_{44}c_{33}}$, $\tilde{\mathbf{H}}_{c_{55}c_{33}}$, and $\tilde{\mathbf{H}}_{c_{66}c_{33}}$.

perturbation of P-wave velocity at $\mathbf{x}=(0.5 \text{ km}, 0.5 \text{ km})$ and calculate the gradients for all of the three parameters, which forms the multi-parameter point spread functions $\mathbf{H}_{\alpha\alpha}(\mathbf{x})$, $\mathbf{H}_{\alpha\beta}(\mathbf{x})$ and $\mathbf{H}_{\alpha\rho}(\mathbf{x})$. We then apply a +10% point perturbation of S-wave velocity at $\mathbf{x}=(0.5 \text{ km}, 0.5 \text{ km})$ and calculate the gradients, which correspond to $\mathbf{H}_{\beta\alpha}(\mathbf{x})$, $\mathbf{H}_{\beta\beta}(\mathbf{x})$ and $\mathbf{H}_{\beta\rho}(\mathbf{x})$. Finally, we apply point perturbation of density and calculate the gradients, which are $\mathbf{H}_{\rho\alpha}(\mathbf{x})$, $\mathbf{H}_{\rho\beta}(\mathbf{x})$ and $\mathbf{H}_{\rho\rho}(\mathbf{x})$. We rearrange the multi-parameter point spread functions, as shown in Figures 11a and 11b. Figure 11c show the point spread functions after normalization, from which we are able to evaluate the inter-parameter mapping for elastic FWI. We conclude that the mapping from P-wave velocity to S-wave velocity and density are weak. Positive P-wave perturbation produces negative artifacts in S-wave velocity and density. Positive S-wave velocity perturbation results in strong and negative mappings in P-wave velocity and density models. Positive density perturbation produces positive artifacts in P-wave velocity and S-wave velocity. These predictions are verified by the numerical examples in Pan and Innanen (2016).

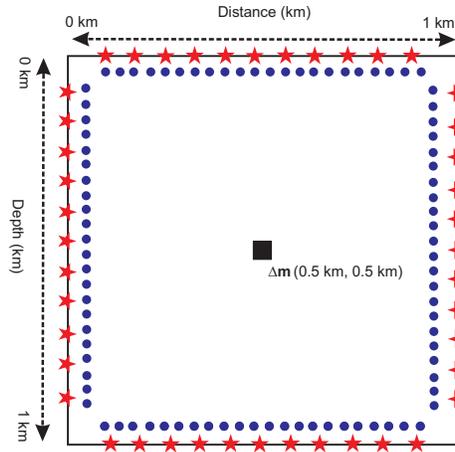


FIG. 10. The acquisition geometry for constructing the PSFs. Red stars and blue circles represent the locations of the sources and receivers.

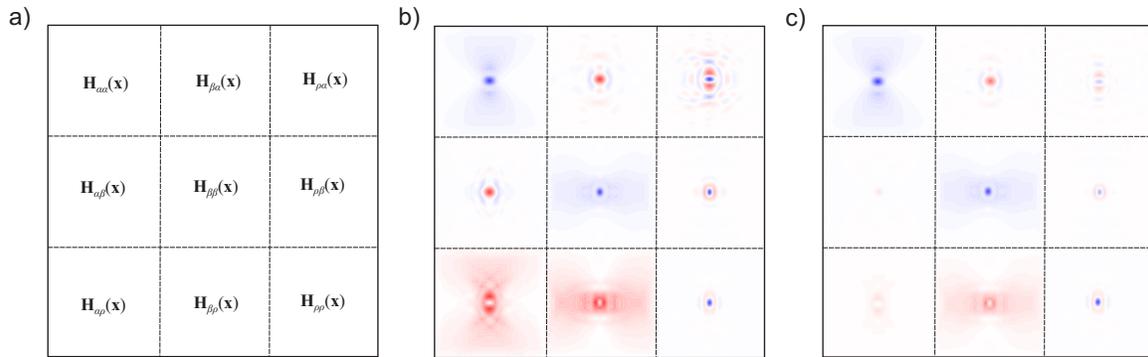


FIG. 11. (a), (b) and (c) show the multi-parameter point spread functions in elastic and isotropic media.

In the next example, we design a 3D HTI model for examining the multi-parameter point spread functions. The model parameters are: P-wave velocity $\alpha = 3200\text{m/s}$, S-wave velocity $\beta = 1800\text{m/s}$, density $\rho = 1800\text{kg/m}^3$, $\varepsilon = -0.14$, $\delta = -0.18$ and $\gamma = -0.15$. Figure 12 shows the acquisition survey with a plan view. The positions of sources and receivers are denoted by red stars and blue circles. Nine sources at a depth of 32m are excited simultaneously. A total of 1681 receivers are arranged regularly with a depth of 32m.

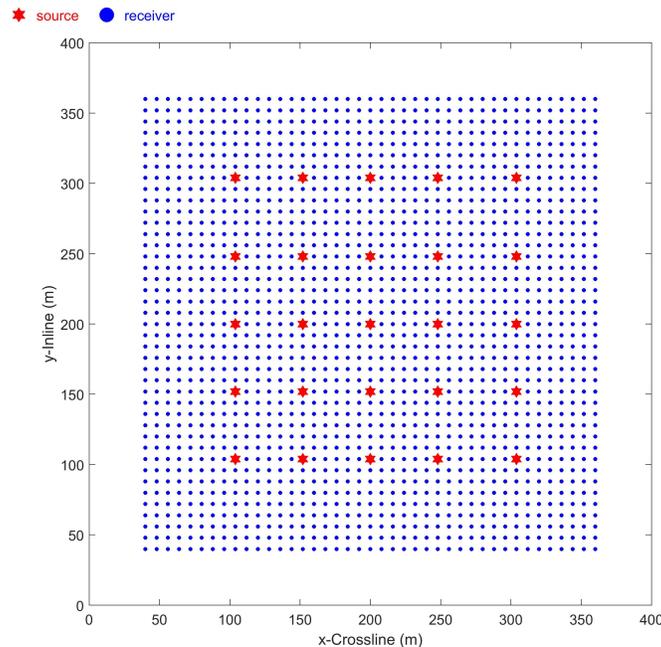


FIG. 12. The 3D acquisition geometry. The red stars and blue circles represent the positions of sources and receivers.

We next carry out numerical experiments for constructing the multi-parameter point spread function. Considering that the model perturbation of P-wave velocity is point located at position $\tilde{\mathbf{x}}=(200\text{ m}, 200\text{ m}, 200\text{ m})$, the gradient of the point scatter leads to the determination of one column of the multi-parameter Hessian. The z , x , and y components of the data residuals due the point perturbation of P-wave velocity are illustrated in Figures 13a, 13b, and 13c respectively. Note that because we only applied +5% perturbation of

P-wave velocity, the data residuals are contaminated by artifacts.

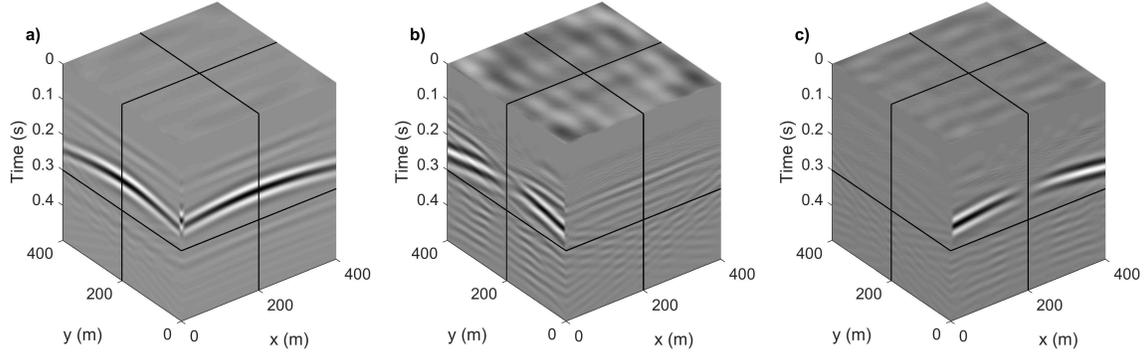


FIG. 13. Data residuals due to perturbation of P-wave velocity. (a) z -component; (b) x -component; (c) y -component.

A P-wave velocity perturbation located at $\tilde{\mathbf{x}}$ will map into a blurred S-wave velocity heterogeneity with the shape of $\mathbf{H}_{\alpha\beta}(\tilde{\mathbf{x}})$, a blurred density heterogeneity with the shape of $\mathbf{H}_{\alpha\rho}(\tilde{\mathbf{x}})$, a blurred ε heterogeneity with the shape of $\mathbf{H}_{\alpha\varepsilon}(\tilde{\mathbf{x}})$, a blurred δ heterogeneity with the shape of $\mathbf{H}_{\alpha\delta}(\tilde{\mathbf{x}})$ and a blurred δ heterogeneity with the shape of $\mathbf{H}_{\alpha\gamma}(\tilde{\mathbf{x}})$.

Figure 14a shows the point spread function $\mathbf{H}_{\alpha\alpha}(\tilde{\mathbf{x}})$. Figures 14b-14f show the multi-parameter point spread functions $\mathbf{H}_{\alpha\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\delta}(\tilde{\mathbf{x}})$ and $\mathbf{H}_{\alpha\gamma}(\tilde{\mathbf{x}})$, which indicate the inter-parameter mappings due to model perturbation $\Delta\mathbf{m}(\tilde{\mathbf{x}})$ to model parameters β , ρ , ε , δ and γ respectively. There are several aspects for understanding the inter-parameter trade-offs with the multi-parameter point spread functions: the strength of the point spread function, the polarity of the point spread function and the spreading range of the point spread function.

As we can see from Figure 14, the perturbation of model parameter $\Delta\mathbf{m}_\alpha(\tilde{\mathbf{x}})$ results in parameter crosstalk artifacts in other parameters. However, we observe that strength, polarity, spreading width of the parameter crosstalk artifacts for different parameters are quite different. Positive P-wave velocity perturbation results in negative and positive parameter crosstalk artifacts in S-wave velocity and density. However, the strength of artifacts are very weak, which means that the parameter cross-talk artifacts due to P-wave velocity may not be significant for S-wave velocity and density. However, parameter crosstalk artifacts in the anisotropic parameters ε , δ and γ are very strong, which will make it quite difficult to invert anisotropic parameters.

We then calculate the multi-parameter point spread functions due to the model perturbation $\Delta\mathbf{m}_\beta(\tilde{\mathbf{x}})$ for investigating the influences of S-wave velocity on the inversion of other model parameters. Figures 15a-15c show the z component, x component and y component of the data residuals due to model perturbation of S-wave velocity. Figures 16a-16f show the multi-parameter point spread functions $\mathbf{H}_{\beta\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\gamma}(\tilde{\mathbf{x}})$ due to the model perturbation $\Delta\mathbf{m}_\beta(\tilde{\mathbf{x}})$. We notice that the parameter crosstalk artifacts introduced by S-wave velocity perturbation are very strong, which means that the S-wave velocity perturbations make it difficult to invert P-wave velocity, density and the anisotropic parameters.

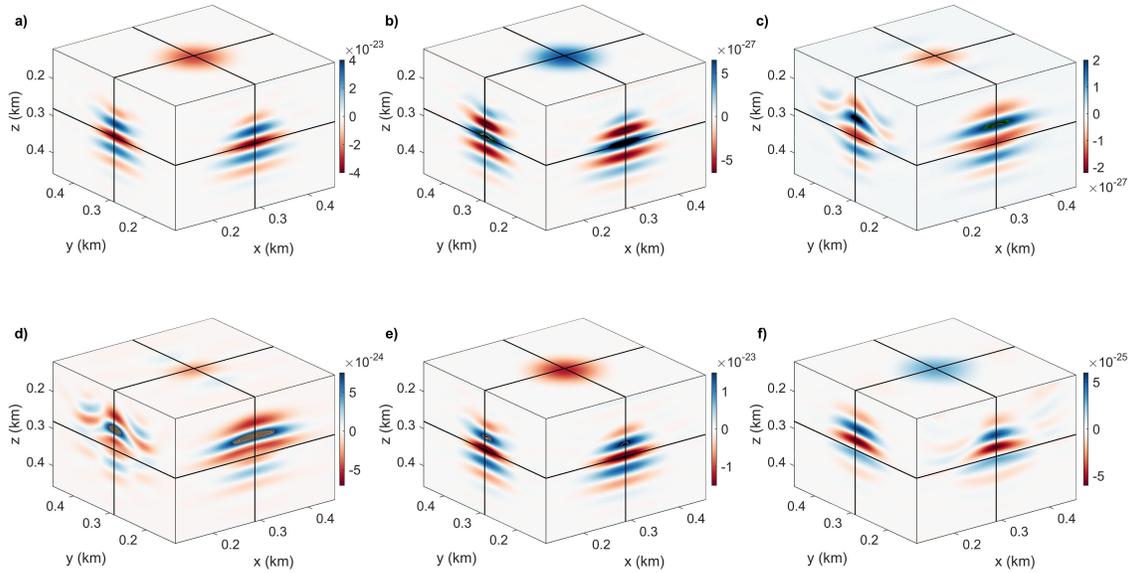


FIG. 14. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\alpha\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\epsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\alpha\gamma}(\tilde{\mathbf{x}})$.

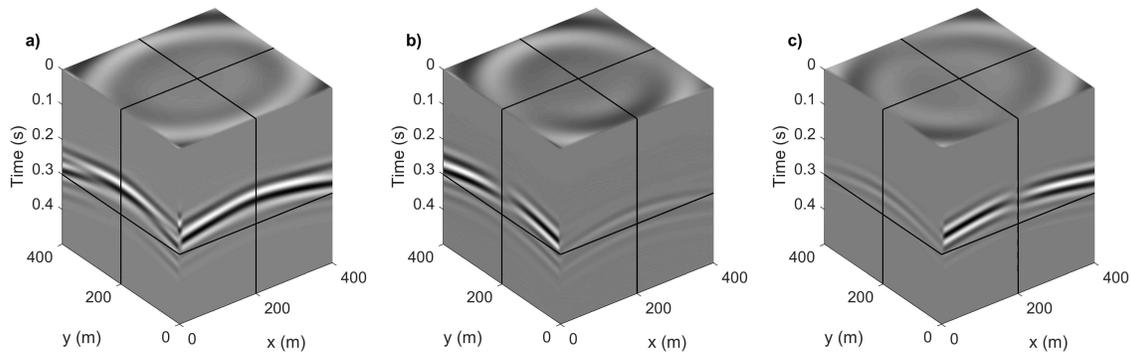


FIG. 15. Data residuals due to perturbation of S-wave velocity. (a) z -component; (b) x -component; (c) y -component.

Figures 17, 19, 21 and 23 show the data residuals due to model perturbations of density, anisotropic parameters ϵ , δ and γ respectively. Using similar approaches, we can calculate the multi-parameter point spread functions due to density, anisotropic parameters ϵ , δ and γ , as shown in Figures 18, 20, 22, and 24. From our observations, we conclude that Thomsen parameters ϵ , δ and γ will not produce significant parameter crosstalk artifacts for P-wave velocity, S-wave velocity and density. However, P-wave velocity, S-wave velocity and density perturbations will make these anisotropic parameter difficult to be inverted.

DISCUSSION

It is known that the multi-parameter Hessian can help to reduce the parameter crosstalk artifacts. However, it is extremely extensive to construct the Hessian matrix explicitly for large-scale inverse problems. The Hessian-free methods provide effective strategies of applying inverse Hessian to precondition the gradient (Pan et al., 2017). In these types of methods, only Hessian-vector products are required and the search direction is obtained by

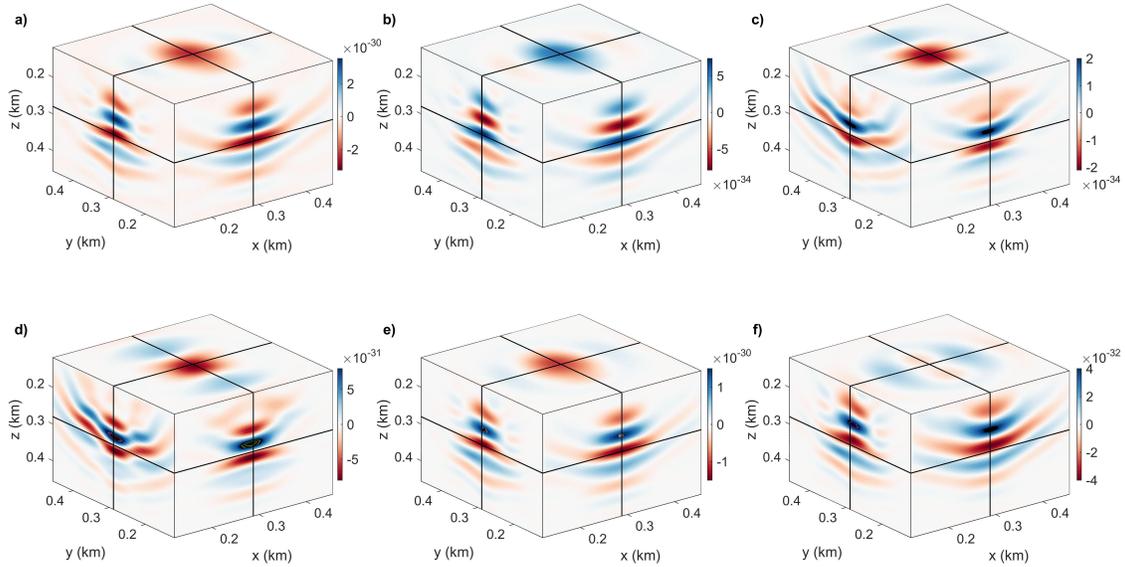


FIG. 16. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\beta\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\beta\gamma}(\tilde{\mathbf{x}})$.

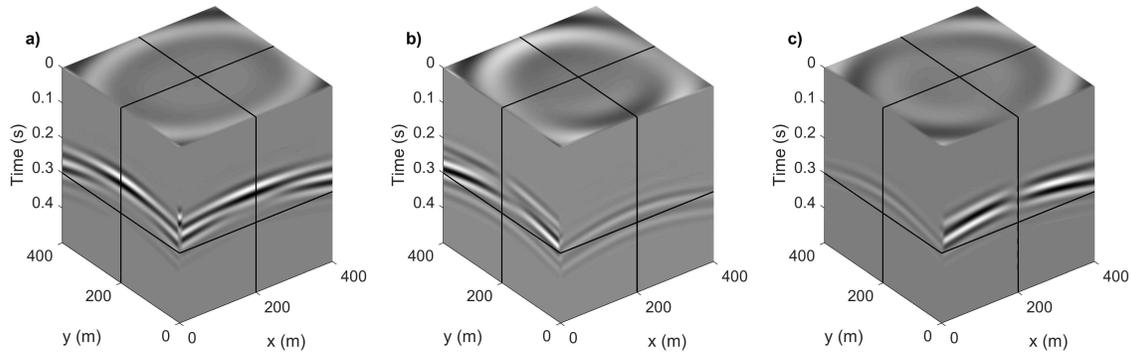


FIG. 17. Data residuals due to perturbation of density. (a) z -component; (b) x -component; (c) y -component.

solving the Newton equation iteratively with an conjugate-gradient algorithm (Pan et al., 2017). Most of Gauss-Newton FWI applications are implemented in frequency domain. From my personal experience, for frequency domain Hessian-free Gauss-Newton FWI, only a limited number of inner iterations are affordable. Time domain Gauss-Newton FWI should be more stable. In this research, we propose to quantify the parameter resolution for multi-parameter FWI in elastic and anisotropic media via probing the multi-parameter Hessian. Furthermore, we notice that the probing technique can be used to construct the Hessian-vector products in time domain, which can be used to implement Gauss-Newton FWI in time domain.

CONCLUSION

In this report, we analyze the disadvantages of scattering patterns for parameter resolution analysis in multi-parameter FWI. To quantify the parameter resolution more completely, we propose to employ Hessian information. With the adjoint-state technique, we

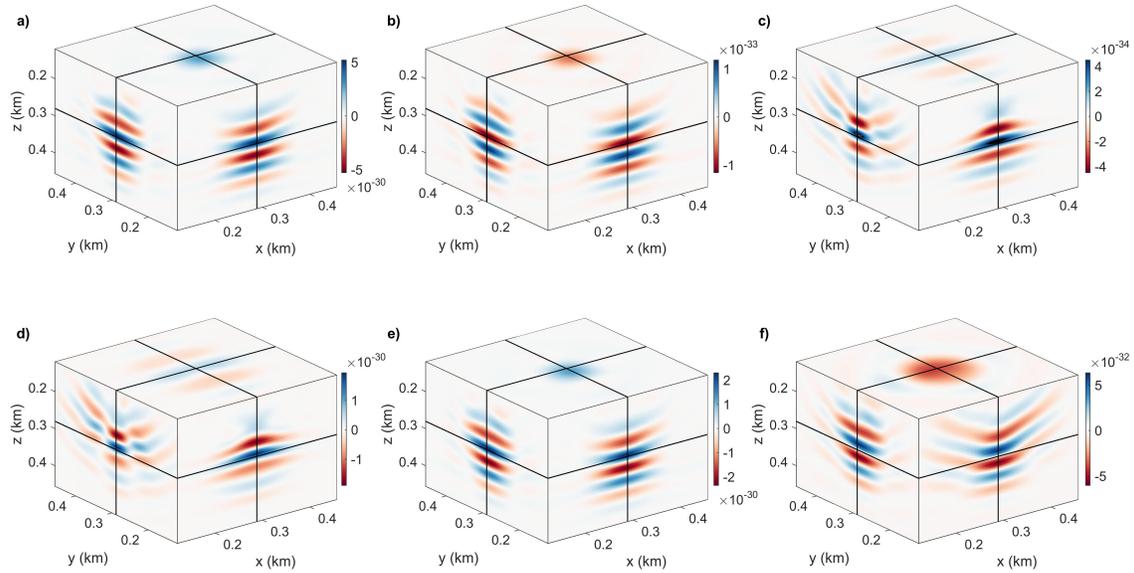


FIG. 18. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\rho\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\rho\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\rho\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\rho\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\rho\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\rho\gamma}(\tilde{\mathbf{x}})$.

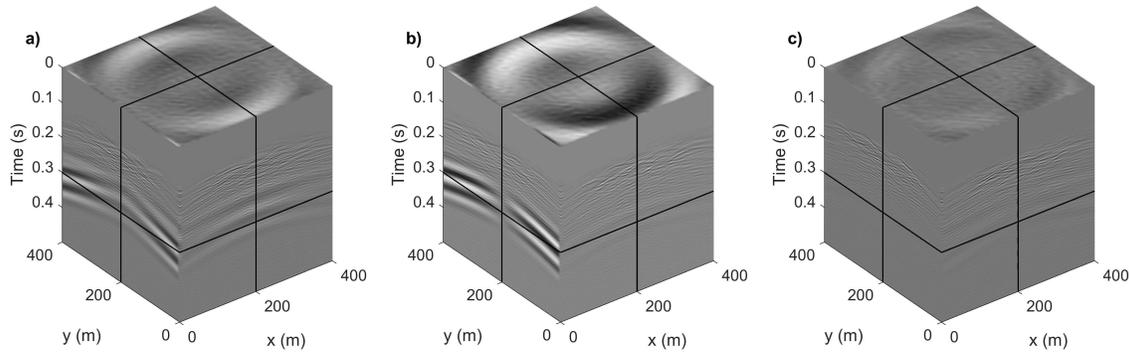


FIG. 19. Data residuals due to perturbation of ε . (a) z -component; (b) x -component; (c) y -component.

are able to calculate the diagonal elements of the block matrices in multi-parameter Hessian explicitly. We can also extract one column from multi-parameter Hessian, which is named as the multi-parameter point spread function, with the spike probing technique. Thus, it is possible to evaluate the parameter crosstalk of different parameters at different locations. In the numerical modelling section, we give numerical examples to verify our ideas.

Gradients of the elastic constants in general anisotropic media In this appendix, we give the gradient expressions for the 21 elastic constants in general anisotropic media. Equation (26) illustrates the gradient expression

$$K_{c_{ijkl}}(\mathbf{x}) = - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \partial_j G_{ni}(\mathbf{x}_r, \mathbf{x}, T-t) \delta c_{ijkl}(\mathbf{x}) \partial_k u_l(\mathbf{x}, \mathbf{x}_s, t) dt \quad (26)$$

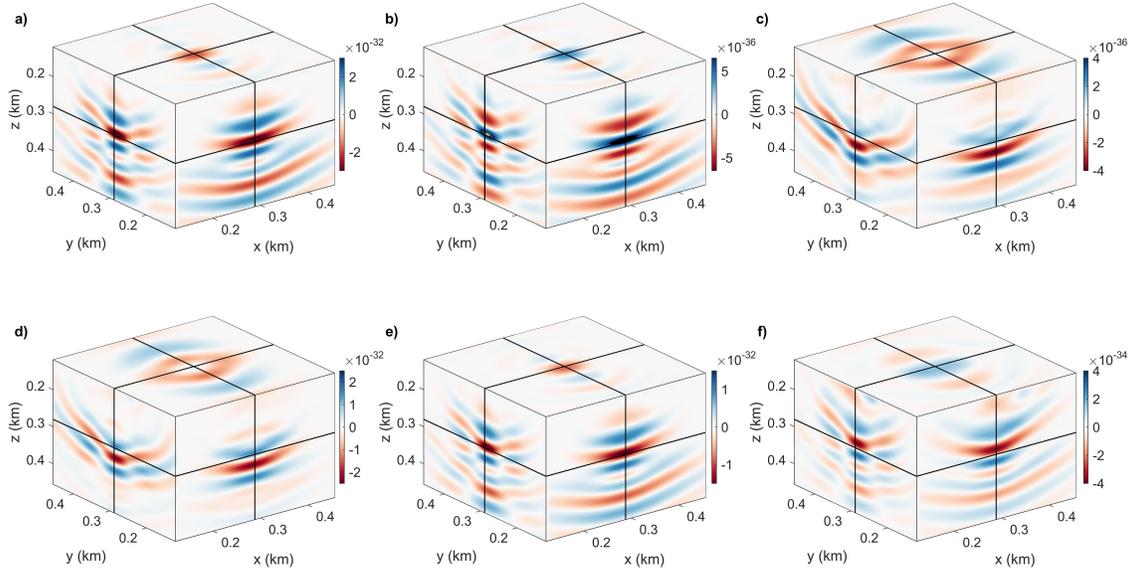


FIG. 20. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\epsilon\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\epsilon\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\epsilon\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\epsilon\epsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\epsilon\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\epsilon\gamma}(\tilde{\mathbf{x}})$.

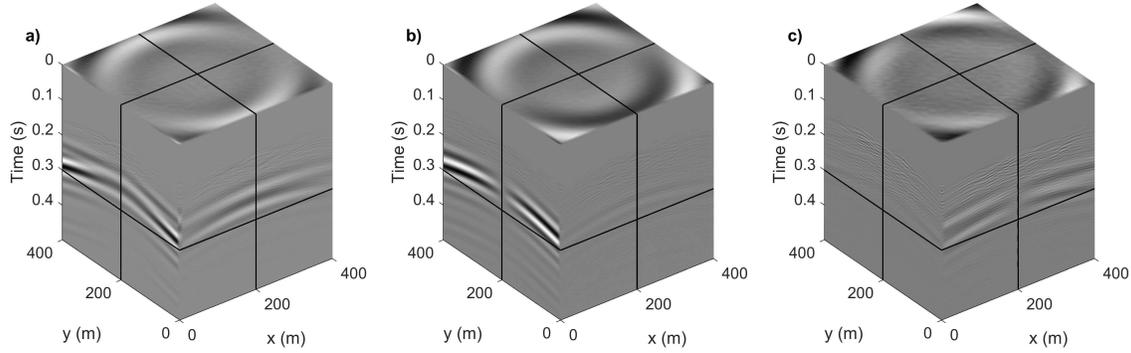


FIG. 21. Data residuals due to perturbation of δ . (a) z -component; (b) x -component; (c) y -component.

The gradients for the 21 elastic constants can be written explicitly as:

$$\begin{aligned}
 K_{c_{11}}(\mathbf{x}) &= -\sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{11}(\mathbf{x}) \partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t) dt, \\
 K_{c_{12}}(\mathbf{x}) &= -\sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{12}(\mathbf{x}) (\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + \partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{13}}(\mathbf{x}) &= -\sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{13}(\mathbf{x}) (\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + \partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{14}}(\mathbf{x}) &= -\sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{14}(\mathbf{x}) (2\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)) dt,
 \end{aligned} \tag{27}$$

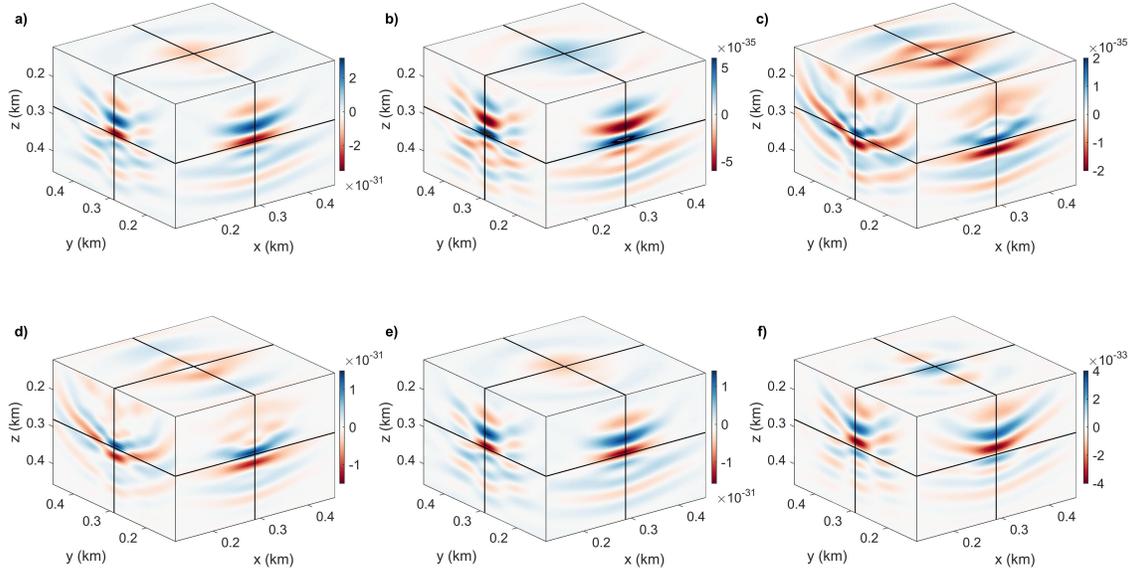


FIG. 22. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\delta\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\delta\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\delta\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\delta\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\delta\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\delta\gamma}(\tilde{\mathbf{x}})$.

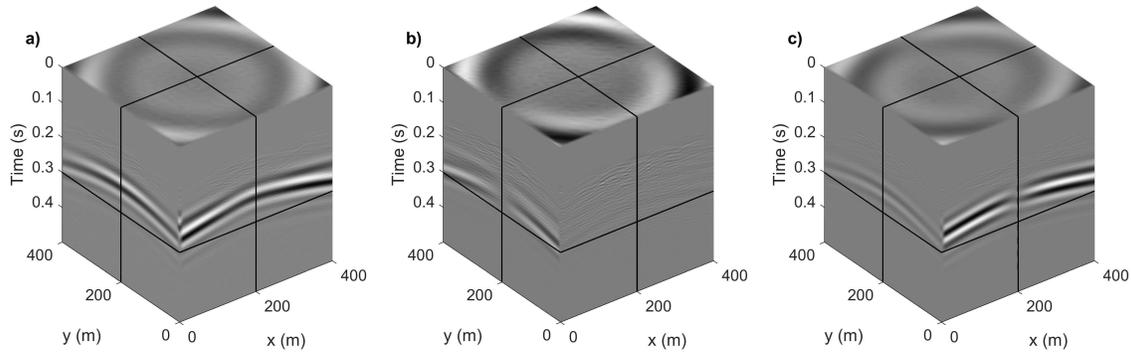


FIG. 23. Data residuals due to perturbation of γ . (a) z -component; (b) x -component; (c) y -component.

$$\begin{aligned}
 K_{c_{15}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{15}(\mathbf{x}) (2\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{16}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{16}(\mathbf{x}) (2\partial_x G_{nx}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_x(\mathbf{x}, \mathbf{x}_s, t)) dt, \quad (28) \\
 K_{c_{22}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{22}(\mathbf{x}) \partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t) dt, \\
 K_{c_{23}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{23}(\mathbf{x}) (2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t)) dt,
 \end{aligned}$$

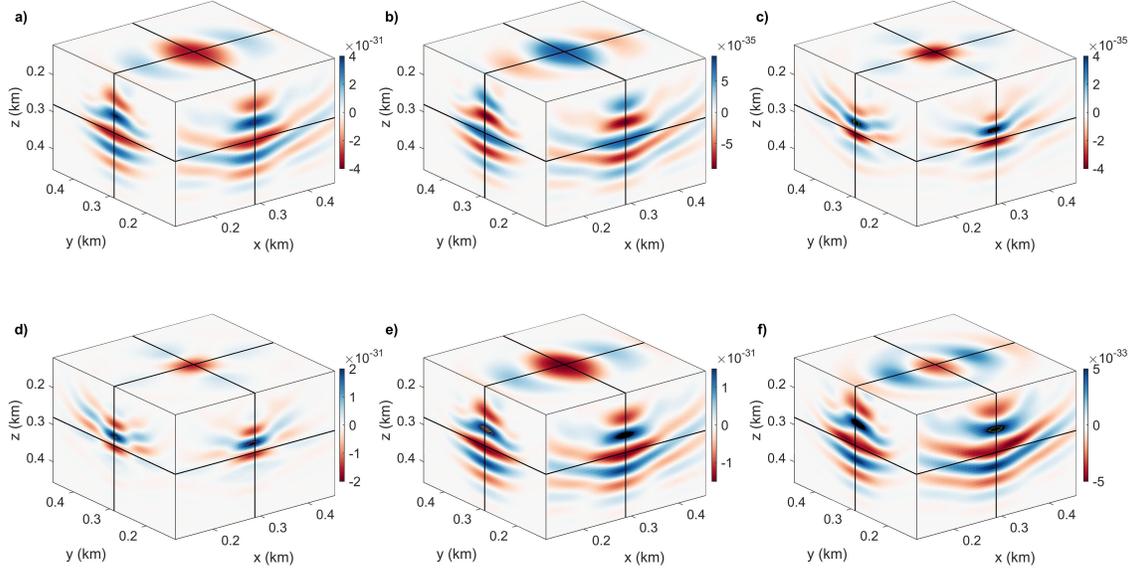


FIG. 24. (a), (b), (c), (d), (e) and (f) show the multi-parameter point spread functions $\mathbf{H}_{\gamma\alpha}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\gamma\beta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\gamma\rho}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\gamma\varepsilon}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\gamma\delta}(\tilde{\mathbf{x}})$, $\mathbf{H}_{\gamma\gamma}(\tilde{\mathbf{x}})$.

$$\begin{aligned}
 K_{c_{24}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{24}(\mathbf{x}) (2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{25}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{25}(\mathbf{x}) (2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{26}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{26}(\mathbf{x}) (2\partial_y G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_y(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{33}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{33}(\mathbf{x}) \partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t) dt, \\
 K_{c_{34}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{34}(\mathbf{x}) (2\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_y z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t)) dt,
 \end{aligned} \tag{29}$$

$$\begin{aligned}
 K_{c_{35}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{35}(\mathbf{x}) (2\partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{36}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{36}(\mathbf{x}) (2\partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_z u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 2\partial_z G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t)) dt, \quad (30) \\
 K_{c_{44}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{44}(\mathbf{x}) 4\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t) dt, \\
 K_{c_{45}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{45}(\mathbf{x}) (4\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 4\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{46}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{46}(\mathbf{x}) (4\partial_y G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 4\partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_y u_z(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{55}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{55}(\mathbf{x}) 4\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_z(\mathbf{x}, \mathbf{x}_s, t) dt, \\
 K_{c_{56}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{56}(\mathbf{x}) (4\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t) \\
 &\quad + 4\partial_x G_{nz}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t)) dt, \\
 K_{c_{66}}(\mathbf{x}) &= - \sum_{r=1}^R \sum_{s=1}^S \int_0^T \Delta c_{66}(\mathbf{x}) 4\partial_x G_{ny}(\mathbf{x}_r, \mathbf{x}, T-t) \partial_x u_y(\mathbf{x}, \mathbf{x}_s, t) dt.
 \end{aligned}
 \tag{31}$$

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. Author 1 was also supported by a SEG/Chevron scholarship and eyes high international doctoral scholarship. A special thanks to Doug S. Phillips of Information Technologies, University of Calgary, for his kind help of running parallel codes on Lattice cluster. Thanks also to Compute Canada for providing high-performance computing facilities.

REFERENCES

- Alkhalifa, T., and Plessix, R., 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytic parameter resolution study: *Geophysics*, **79**, R91–R101.
- An, M., 2012, A simple method for determining the spatial resolution of a general inverse problem: *Geophysical Journal International*, **191**, 849–864.

- Bergslid, T. S., Raknes, E. B., and Arntsen, B., 2015, The influence of anisotropy on elastic full-waveform inversion: SEG Expanded Abstracts, 1425–1429.
- Fichtner, A., and van Leeuwen, T., 2015, Resolution analysis by random probing: Journal of Geophysical Research: Solid Earth, **120**, 5549–5573.
- Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A., and Virieux, J., 2013a, Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? part 2: Synthetic and real data case studies from valhall: Geophysics, **78**, R107–R124.
- Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., and Virieux, J., 2013b, Which parametrization for acoustic VTI full waveform inversion? - part 1: sensitivity and trade-off analysis: Geophysics, **78**, R81–R105.
- Innanen, K. A., 2014a, Reconciling seismic AVO and precritical reflection FWI-analysis of the inverse Hessian: SEG Technical Program Expanded Abstracts, 1022–1027.
- Innanen, K. A., 2014b, Seismic AVO and the inverse Hessian in precritical reflection full waveform inversion: Geophysical Journal International, **199**, 717–734.
- Kamath, N., and Tsvankin, I., 2014, Sensitivity analysis for elastic full-waveform inversion in vti media: SEG Expanded Abstracts, 1162–1166.
- Köhn, D., Hellwig, O., Nil, D. D., and Rabbel, W., 2015, Waveform inversion in triclinic anisotropic media—a resolution study: Geophysical Journal International, **201**, 1642–1656.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration: Conference on Inverse Scattering, Theory and Applications, SIAM, Expanded Abstracts, 206–220.
- Li, Y., and Demanet, L., 2016, Full-waveform inversion with extrapolated low-frequency data: Geophysics, **81**, R339–R348.
- Ma, Y., and Hale, D., 2012, Quasi-Newton full-waveform inversion with a projected Hessian matrix: Geophysics, **77**, R207–R216.
- 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, **32**, 1040–1054.
- Pan, W., and Innanen, K. A., 2016, Elastic full-waveform inversion: density effects, cycle-skipping, and inter-parameter mapping: CREWES Annual Report, 1–16.
- Pan, W., Innanen, K. A., and Liao, W., 2017, Accelerating Hessian-free Gauss-Newton full-waveform inversion via *l*-BFGS preconditioned conjugate-gradient algorithm: Geophysics, 1–10.
- Pan, W., Innanen, K. A., and Margrave, G. F., 2014a, A comparison of different scaling methods for least-squares migration/inversion: EAGE Expanded Abstracts, We G103 14.
- Pan, W., Innanen, K. A., Margrave, G. F., and Cao, D., 2015, Efficient pseudo-Gauss-Newton full-waveform inversion in the τ - p domain: Geophysics, **80**, R225–R14.
- Pan, W., Innanen, K. A., Margrave, G. F., Fhler, M. C., Fang, X., and Li, J., 2016a, Estimation of elastic constants for HTI media using Gauss-Newton and full-Newton multiparameter full-waveform inversion: Geophysics, **81**, R275–R291.
- Pan, W., Innanen, K. A., Margrave, G. F., and Keating, S., 2016b, Mitigate cycle-skipping for full-waveform inversion by band-limited impedance inversion and POCS: GeoConvention, 1–5.
- Pan, W., Margrave, G. F., and Innanen, K. A., 2014b, Iterative modeling migration and inversion (IMMI): Combining full waveform inversion with standard inversion methodology: SEG Technical Program Expanded Abstracts, 938–943.

- Podgornova, O., Leaney, S., and Liang, L., 2015, Analysis of resolution limits of VTI anisotropy with full waveform inversion: SEG Expanded Abstracts, 1188–1192.
- Pratt, R. G., Shin, C., and Hicks, G. J., 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Rüger, A., 1997, P-wave reflection coefficients for transversely isotropic models with vertical and horizontal axis of symmetry: *Geophysics*, **62**, 713–722.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: *Geophysics*, **51**, 1893–1903.
- Trampert, J., Fichtner, A., and Ritsema, J., 2013, Resolution tests revisited: the power of random numbers: *Geophysical Journal International*, **192**, 676–680.
- Tsvankin, I., 1997a, Anisotropic parameters and p-wave velocity for orthorhombic media: *Geophysics*, **62**, 1292–1309.
- Tsvankin, I., 1997b, Reflection moveout and parameter estimation for horizontal transverse isotropy: *Geophysics*, **62**, 614–629.
- Virieux, A., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.
- Warner, M., and Guasch, L., 2014, Adaptive waveform inversion: theory: SEG Expanded Abstracts, 1089–1093.
- Wu, W., Luo, J., and Wu, B., 2014, Seismic envelope inversion and modulation signal model: *Geophysics*, **79**, WA13–WA24.
- Zhu, H., and Fomel, S., 2016, Building good starting models for full-waveform inversion using adaptive matching filtering misfit: *Geophysics*, **81**, U61–U72.