

# Joint Migration Inversion continuous equations and discretized solution via multiparameter Gauss-Newton method

C. A. M. Assis\* (CEP/UNICAMP), Y. Pan (Karlsruhe Institute of Technology), J. Schleicher (IMECC/UNICAMP & INCT-GP), T. Bohlen (Karlsruhe Institute of Technology), and H. B. Santos (CEP/UNICAMP & INCT-GP)

Copyright 2019, SBGf - Sociedade Brasileira de Geofísica.

This paper was prepared for presentation at the  $16^{\prime h}$  International Congress of the Brazilian Geophysical Society, held in Rio de Janeiro, Brazil, 19-22 August, 2019.

Contents of this paper were reviewed by the Technical Committee of the 16<sup>th</sup> International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

## Abstract

Joint Migration Inversion (JMI) is an independent approach to solve the seismic inverse problem, based on decoupled imaging and inversion operators. Here, we review the background equations in their continuous form. We then proceed to test the JMI methodology using the multiparameter Gauss-Newton method to estimate simultaneously image and slowness updates, and we compare the results to those of the conventionally used steepest-descent method. Our numerical results show that the Gauss-Newton method can provide velocity models with improved resolution, albeit at a higher cost. These results demonstrate that the JMI implementation under the assumptions discussed here can provide a good depth migrated image and a satisfying initial velocity model for a subsequent Full Waveform Inversion.

## Introduction

Joint Migration Inversion (JMI) has been proposed as a new approach to the seismic inverse problem (Berkhout, 2014). The method's distinguishing characteristics are the ability to reproduce intrabed multiples even in a smooth velocity model and a computational cost below finite difference based methods.

JMI is built upon the seismic wavefield decomposition into its downgoing and upgoing components. This approach intrinsically decouples high spatial frequency information, represented by scattering operators, from low spatial frequency information represented by the velocity model (Wapenaar, 1996).

By means of the migration part, the velocity model allows for the positioning of events. In turn, from the resulting image, scattering operators are built to update the velocity model. Since the methodology deals mainly with reflections below the critical angle, it is approximately analogous to Reflection Waveform Inversion (Xu et al., 2012).

Because of the higher cost of wave-based methods, raybased tomography methods are still routinely used in oil and gas exploration (Jones, 2010). In terms of cost and quality, JMI offers an intermediate solution between ray-based methods and Full Waveform Inversion (FWI), accounting for finite frequency effects, but not demanding event picking and with lower computational cost than finite difference solutions of the wave equation.

This work aims to derive the JMI continuous equations and solve the inverse problem using the Gauss-Newton method to estimate simultaneously image and slowness updates. Usually, the methodology is presented in the discrete matrix form, and the parameter updating is done by alternating between imaging and tomography (Masaya and Verschuur, 2018). In the continuous form, the equations may be slightly more cumbersome, but this helps to state the method and its assumptions as clear as possible. The results of our implementation indicate that JMI is a promising methodology for velocity model building, which may assist in reducing the number of necessary iterations in a subsequent FWI application.

## **Forward problem**

## Differential form

Consider the acoustic wave equation decoupled into its downgoing and upgoing components. We present the correspondent equations for a continuous model. According to Ursin et al. (2012), the down/upgoing wavefields  $P^{\pm}(\mathbf{x}, \omega)$  in the frequency-space domain  $(\omega - \mathbf{x})$ for a downgoing volumetric source  $S^+ = S^+(\mathbf{x}_s, \omega)$  must satisfy

$$\frac{\partial P^{+}}{\partial z} = iH_{1}(\mathbf{x}, \boldsymbol{\omega})P^{+}(\mathbf{x}, \boldsymbol{\omega}) + \underbrace{R^{-}(\mathbf{x}, \boldsymbol{\omega})P^{-}(\mathbf{x}, \boldsymbol{\omega})}_{secondary source} + S^{+}, \quad (1)$$

$$\frac{\partial P^{-}}{\partial z} = -iH_{1}(\mathbf{x}, \boldsymbol{\omega})P^{-}(\mathbf{x}, \boldsymbol{\omega}) + \underbrace{R^{+}(\mathbf{x}, \boldsymbol{\omega})P^{+}(\mathbf{x}, \boldsymbol{\omega})}_{secondary source}, \quad (2)$$

where *i* is the imaginary unit;  $\mathbf{x} = (x, z)$  is the coordinate vector, with *x* indicating the horizontal coordinate and *z* the depth coordinate increasing downward;  $\mathbf{x}_s$  is the source position, and  $\omega$  is the angular frequency. Moreover,  $H_1$  is the square-root operator and *R* is the reflection operator, with the superscript + indicating incidence from above at some model position  $\mathbf{x}$  and – denotes incidence from below. Note that as a consequence of the directional decoupling, horizontally propagating waves are not defined (Ursin et al., 2012).

The square-root operator  $H_1$  is defined in such a way that its double application results in the Helmholtz operator  $H_2$ , i.e.,

$$H_2(\mathbf{x},\boldsymbol{\omega}) = H_1(\mathbf{x},\boldsymbol{\omega})H_1(\mathbf{x},\boldsymbol{\omega}), \qquad (3)$$

where the Helmholtz operator, related to the 2D acoustic

wave equation, is defined as

$$H_2(\mathbf{x},\boldsymbol{\omega}) = \boldsymbol{\omega}^2 \boldsymbol{\sigma}^2(\mathbf{x}) + \frac{\partial}{\partial x} \left( \frac{1}{\rho} \frac{\partial}{\partial x} \cdot \right).$$
(4)

Here,  $\sigma$  is the slowness (velocity inverse) and  $\rho$  is density.

In a homogeneous model, the square-root operator in the frequency-wavenumber domain is equivalent to the vertical wavenumber

$$k_z = \omega \sqrt{\sigma^2 - k_x^2 / \omega^2}, \qquad (5)$$

where  $k_x$  is the horizontal wavenumber. The squareroot operator and the scattering operators represent convolutions in the horizontal coordinate, viz.,

$$(H_1 P^{\pm})(\mathbf{x}, \boldsymbol{\omega}) = \int_{-\infty}^{\infty} \mathscr{H}_1(x - x', z, \boldsymbol{\omega}, \sigma(x', z)) P^{\pm}(x', z, \boldsymbol{\omega}) dx',$$
(6)

$$(\mathbf{R}^{+}P^{+})(\mathbf{x},\boldsymbol{\omega}) = \int_{-\infty}^{\infty} \mathscr{R}^{+}(x-x',z,\boldsymbol{\omega})P^{+}(x',z,\boldsymbol{\omega})dx', \qquad (7)$$

where  $\mathscr{H}_1$  and  $\mathscr{R}^+$  are the kernels of the square-root and reflection operator, respectively. A detailed discussion about the square-root operator can be found in Grimbergen et al. (1998). For the purpose of this work, it will not be necessary to specify this operator explicitly.

## Integral form

The downgoing Green's function  $G^+$  for our modeling equations, in a medium without scattering, must satisfy

$$\frac{\partial G^{+}}{\partial z}(\mathbf{x};\mathbf{x}_{s},\boldsymbol{\omega}) = iH_{1}(\mathbf{x},\boldsymbol{\omega})G^{+}(\mathbf{x};\mathbf{x}_{s},\boldsymbol{\omega}) + \delta(\mathbf{x}-\mathbf{x}_{s})\delta(t), \quad (8)$$

The upgoing analogous,  $G^-$ , is obtained with a change of signal of the imaginary unit or equivalently

$$G^{-}(\mathbf{x}_{s};\mathbf{x},\boldsymbol{\omega}) = \left[G^{+}(\mathbf{x};\mathbf{x}_{s},\boldsymbol{\omega})\right]^{\dagger}, \qquad (9)$$

where † denotes the adjoint operator.

Relating equations (1) and (2) with the correspondent Green's function the following integral representation is obtained (Wapenaar, 1996)

$$P^{\pm}(x,z',\boldsymbol{\omega}) = \int_{-\infty}^{\infty} G^{\pm}(x,z';x,z,\boldsymbol{\omega})P^{\pm}(x,z,\boldsymbol{\omega})dx, \qquad (10)$$

or in compact notation, which considers the integral representation and the convolution over the lateral coordinate implicitly,

$$P^{\pm}(\mathbf{x}',\boldsymbol{\omega}) = G^{\pm}(\mathbf{x}';\mathbf{x},\boldsymbol{\omega})P^{\pm}(\mathbf{x},\boldsymbol{\omega}).$$
(11)

In the same notation, and for brevity displaying only the z dependence, the modeling equations can be written as

$$P^{+}(z'-\varepsilon) = G^{+}(z';z)Q^{+}(z+\varepsilon) + G^{+}(z';z_{s})S^{+}(z_{s}+\varepsilon),$$
 (12)

$$P^{-}(z'+\varepsilon) = G^{-}(z';z)Q^{-}(z-\varepsilon), \qquad (13)$$

where a small vertical displacement  $\varepsilon$  was introduced to indicate the relative position of each wavefield with respect to its incidence point. Moreover, symbols

$$Q^{\pm} = R^{\mp}(\mathbf{x}, \boldsymbol{\omega}) P^{\mp}(\mathbf{x}, \boldsymbol{\omega}), \qquad (14)$$

denote the secondary sources for each propagation direction.

#### Recursive modeling and further assumptions

Note that equations (12) and (13) account for full scattering. It is possible to obtain a recursive linearized relation between the wavefields that accounts for higher order scattering using a procedure similar to the Born expansion, commonly applied to the two-way wave equation. This can be done by defining two wavefields, one in a model without scattering and another in a model with the same kinematic characteristics but including scatterers. Following this procedure, we obtain the recursive equations

$$P_{j+1}^{+}(z') = G^{+}(z';z) \underbrace{\left[R^{-}(z)P_{j}^{-}(z) + P_{0}^{+}(z)\right]}_{Q_{j}^{+}}, \quad (15)$$

$$P_{j+1}^{-}(z') = G^{-}(z';z) \underbrace{\left[ R^{+}(z) P_{j+1}^{+}(z) \right]}_{Q_{j+1}^{-}},$$
(16)

where the subscript j indicates the scattering order, with the zero-order terms given by

$$P_0^+(z') = G^+(z';z_s)S^+(z_s), \tag{17}$$

$$P_0^{-}(z') = G(z';z)R^{+}(z)P_0^{+}(z).$$
(18)

In this way,  $P_0^+$  denotes the downgoing direct wave,  $P_0^-$  represents the upgoing waves reflected once,  $P_1^+$  are all twice reflected downgoing waves, and  $P_1^-$  stands for upgoing waves that bounced three times, and so on. This recursive modeling procedure is closely related to the work of Bremmer (1951). Note that the only necessary information to account for higher-order scattering at iteration j+1 is the source wavefield  $S^+$  and the downgoing wavefield from the last iteration  $P_j^-$ . To simplify the presentation, we assumed a 2D acoustic model, but the extension to 3D is straightforward.

Still, for simplicity, we additionally assume the density to be constant, medium parameters to be locally homogeneous, and that the scattering operators/coefficients are angle independent. This last assumption leads to scattering operators that are independent of frequency and allows to substitute the convolution in equation (7) by a direct multiplication (de Bruin et al., 1990). Additionally, the following relations hold (Berkhout, 2014) for an acoustic model:

$$T^+ = I + R^+ \,, \tag{19}$$

$$R^- = -R^+, \qquad (20)$$

$$T^{-} = I - R^{+}$$
 (21)

where I is the identity operator. The transmission effects are introduced during the discretization process. Concerning propagation, the locality assumption allows us to approximately implement the Green's function for laterally varying models using the approach of Thorbecke et al. (2004). Under similar assumptions, any other oneway extrapolation technique can be used.

#### Inverse problem

#### Lagrange multipliers

Here we are interested in estimating the reflection coefficient  $R^+$  and the logarithm,  $\sigma_L$ , of slowness  $\sigma$ , i.e.,

3

 $\sigma_L = \ln(\sigma)$ . After inversion of these parameters, the other scattering coefficients are estimated using equations (19), (20), and (21). Considering the regular least-squares misfit function for a single frequency and a single shot gather, we define the following Lagrangian function

$$L(P^{-}, R^{+}, \sigma_{L}, \mu) = \frac{1}{2} \underbrace{\|d^{-} - \mathscr{S}^{-}P^{-}\|_{2}^{2}}_{+ \Re \left\{ \left\langle \underbrace{\frac{\partial P^{-}}{\partial z} + iH_{1}P^{-} - R^{+}P^{+}}_{F(P^{-}, R^{+}, \sigma_{L})}, \mu \right\rangle \right\}}_{(22)}$$

where  $\mu$  is a Lagrange multiplier or adjoint variable;  $\mathscr{S}^-$  samples the wavefield at the receiver positions;  $d^-$  is the upgoing measured data;  $\|.\|_2^2$  denotes the squared  $L_2$  norm;  $\Re\{.\}$  denotes the real-part operator; and  $\langle .,. \rangle$  denotes dot product.

At this point, it is important to emphasize that it would be more complete to include the downgoing wavefield, the so-called source-side wavefield, and formulate all the scattering operators as a function of  $R^+$ . These extensions are left for future work.

Considering that  $\hat{P}^-$  satisfies equation (2), we conclude that

$$L(\hat{P}^{-}, R^{+}, \sigma_{L}, \mu) = \frac{1}{2} \|d^{-} - \mathscr{S}^{-} \hat{P}^{-}\|_{2}^{2}.$$
 (23)

Fixing  $\sigma_L$  and taking the derivative of equation (22) with respect to  $R^+$  yields

$$\Re\left\{\left[\frac{\partial F(\hat{P}^{-},R^{+},\sigma_{L})}{\partial R^{+}}\right]^{\dagger}\mu\right\} + \frac{\partial L(\hat{P}^{-},R^{+},\sigma_{L},\mu)}{\partial P^{-}}\frac{\partial \hat{P}^{-}}{\partial R^{+}}$$
$$= \frac{\partial f(R^{+},\sigma_{L})}{\partial R^{+}}, \quad (24)$$

where

$$\frac{\partial F(\hat{P}^-, R^+, \sigma_L)}{\partial R^+} = \frac{\partial Q^-}{\partial R^+},$$
$$= P^+.$$
(25)

Defining the adjoint state  $\hat{\mu}$  such that

$$\frac{\partial L(\dot{P}^{-}, R^{+}, \sigma_{L}, \hat{\mu})}{\partial P^{-}} = 0, \qquad (26)$$

we conclude that  $\hat{\mu}$  must satisfy

$$\frac{\partial \hat{\mu}}{\partial z} - iH_1 \hat{\mu} = [\mathscr{S}^-]^{\dagger} (d^- - \mathscr{S}^- \hat{P}^-), \qquad (27)$$

and the imaging gradient is given by

$$\frac{\partial f(R^+, \sigma_L)}{\partial R^+} = \Re\left\{ \underbrace{\left[P^+\right]}_{J_{\varrho}}^{\dagger} \hat{\mu} \right\}.$$
 (28)

In an analogous way, now fixing  $R^+$  and derivating equation (22) with respect to  $\sigma_L$ , we find

$$\Re\left\{\left[\frac{\partial F(\hat{P}^{-},R^{+},\sigma_{L})}{\partial\sigma_{L}}\right]^{\dagger}\mu\right\}+\frac{\partial L(\hat{P}^{-},R^{-},\sigma_{L})}{\partial P^{-}}\frac{\partial \hat{P}^{-}}{\partial\sigma_{L}}=\frac{\partial f}{\partial\sigma_{L}}.$$
(29)

Using the same strategy for the adjoint state  $\hat{\mu}$ , equation 27, and with

$$\frac{\partial F(\hat{P}^-, R^+, \sigma_L)}{\partial \sigma_L} = i \frac{\partial H_1}{\partial \sigma_L} \hat{P}^-, \qquad (30)$$

we find the tomography gradient to be given by

$$\frac{\partial f(R^+, \sigma_L)}{\partial \sigma_L} = \Re \left\{ \underbrace{\left[ i \frac{\partial H_1}{\partial \sigma_L} \hat{P}^- \right]^{\dagger}}_{J_{\sigma_L}} \hat{\mu} \right\}.$$
 (31)

Observe that the forward-modeling partial derivatives with respect to  $R^+$  and  $\sigma_L$  are fully decoupled (see  $J_R$  given by equation 25 and  $J_{\sigma_L}$  given by equation 30). The former depends on the downgoing wavefield, and the latter requires the upgoing wavefield.

## Gauss-Newton method

The Gauss-Newton method is an approximate approach to account for the effect of the inverse Hessian on the gradient. Some benefits of the method are its capability to focus the gradient by reducing band-limitation effects from acquisition and consequently speed up the convergence of the inversion process (Pratt et al., 1998). Adding to that, the Gauss-Newton approach can reduce cross-talk between parameter classes in multiparameter inversion (Pan et al., 2018). Using the LSMR algorithm (Fong and Saunders, 2011), we implemented the multiparameter Gauss-Newton method by solving the least-squares problem

$$\min_{\Delta\sigma_{L_k},\Delta R_k^+} \left\| \mathscr{S}^{-} \begin{bmatrix} J_{\sigma_L} & J_R \end{bmatrix} \begin{bmatrix} \Delta\sigma_{L_k} \\ \Delta R_k^+ \end{bmatrix} - (d^{-} - \mathscr{S}^{-} \hat{P}^{-}) \right\|_2^2, \quad (32)$$

where  $\Delta \sigma_{L_k}$  and  $\Delta R_k^+$  are the update directions at iteration k. The parameters are updated according to  $\sigma_{L_{k+1}} = \sigma_{L_k} + \alpha_k \Delta \sigma_{L_k}$  and  $R_{k+1}^+ = R_k^+ + \beta_k \Delta R_k^+$ . The step-lengths are calculated with the subspace method (Kennett et al., 1988). During the iterative solution of equation (32) using LSMR, at each iteration, one linearized modeling and one adjoint modeling are performed. We compare the LSMR results to those obtained using the steepest-descent method by alternating between imaging and tomography. The steepest-descent directions were obtained from equations 28 and 31, normalized by the corresponding squared wavefield Jacobian stacked over shots and frequencies, which corresponds to the deconvolution imaging condition.

#### Numerical tests

We applied JMI to two synthetic data sets. One model is similar to the one used in Masaya and Verschuur (2018), consisting of a lens and fine layering at the bottom. The second model is the modified Marmousi2 (Pan et al., 2018). Both synthetic data sets were generated with a 20 Hz Ricker wavelet. The same algorithm was used for modeling and inversion with a multiscale approach. In the discussion below, when we indicate the Gauss-Newton computational cost, we consider that one linearized modeling has approximately the same computational cost of one adjoint modeling.



Figure 1: Lens model: Initial velocity model for JMI.

## Lens model

The acquisition geometry for the first test consisted of 40 shots spaced by 50 m and receivers at all surface grid points. The vertical and lateral spatial sampling of the model was 12.5 m. Figure 1 shows the initial velocity model for the first test, a linear vertical gradient model. The initial image was a null array. Figures 2(a) and 2(b) show the exact velocity and image, respectively.

For the inversion, the frequency stages were divided into four equal intervals with fixed minimum frequency. The minimum frequency was set to the first sample after 0 Hz and the starting maximum frequency was 15 Hz. The number of iterations per stage was set to a maximum of 20 and a jump to a new stage was imposed if the relative decrease of the objective function was less than 1%.

Overall, the results were acceptably close to the exact model (see Figure 2). The Gauss-Newton solution, Figure 2(e), exhibits a slightly higher resolution and more precision in reproducing the length of the lens than the steepest-descent result, Figure 2(c). However, the fine layers below the lens show a slightly stronger pull-up in the Gauss-Newton image, Figure 2(f) than in the steepestdescent image, Figure 2(d). Other indicators for the quality of the optimization strategies are the data residuals and model residuals. Figure 3 shows these indicators in plots over iterations. The data and model residuals show that the Gauss-Newton method provided model updates which converged considerably faster towards the exact velocity. The average number of inner iterations in the Gauss-Newton approach was 4.3. Therefore, the computational cost of each iteration was approximately 8.6 times that of the steepest-descent method.

## Marmousi2 model

Figures 4 to 6 show the same sequence of figures for the Marmousi2 model. Figure 4 depicts the initial velocity model. The initial image was again a null array. Figures 5(a) and 5(b) show the exact velocity and image, respectively. In this case, the vertical and lateral spatial sampling was 5 m. The acquisition was made with 22 shots spaced by 150 m and receivers at all surface grid points. The frequency stages were divided into four intervals with an increment of 10 Hz, and the first stage was defined as 0-10 Hz. The maximum number of iterations per stage was set to 20, skipping to a new stage if the relative decrease of the objective function was less than 5%. Both the steepest-descent and Gauss-Newton methods, Figures 5(c) and 5(e), were capable of introducing several details into the initial velocity model, Figure 4. The steepest-descent result makes an overall

somewhat smoother impression than the Gauss-Newton velocity model. Gauss-Newton converged slower in the first stage than steepest descent but subsequently provided a more significant decrease of the data/model residuals, Figures 6(a) and 6(b). The resulting images, Figures 5(d) and 5(f), exhibit approximately the same quality, with the gas lens being visibly better resolved in the Gauss-Newton image. The average number of inner iterations to solve the Gauss-Newton approximation was 6.2. Therefore the computational cost of each iteration was approximately 12.4 times that of steepest descent.

## Conclusion

In this paper, we have studied joint migration/inversion (JMI). Differently from what is usually done in the literature on JMI, we have reviewed the continuous forms of the differential and integral equations that form the background of the JMI methodology. Considering the inverse problem, in general, the JMI gradients are derived from the integral equations after discretization. We obtained the gradients using the original underlying differential equations together with the continuous form of the Lagrange multipliers. Note that transmission effects are not represented explicitly in the adopted equations. These effects are introduced during discretization to guarantee the wavefield continuity between adjacent layers at boundaries. We have also discussed the implementation of the multiparameter Gauss-Newton method to simultaneously estimate updates for the scattering operator and the medium slowness. Our numerical tests on two synthetic models of different degrees of geologic complexity indicate that the computationally more expensive Gauss-Newton method can provide higher resolution in the resulting velocity models than the steepest-descent method. The quality of the estimated images turned out to be somewhat independent of the tested optimization algorithm. In conclusion, JMI is a promising methodology for imaging and velocity model building. However, the assumptions considered here, mainly the angle independence of the scattering operators, should be removed in order to have a method capable of dealing with real data. Overall, the obtained velocity models seem to be of sufficient quality to serve as initial models for a subsequent FWI.

## Acknowledgments

The authors are grateful to Renat Shigapov (Karlsruhe Institute of Technology) for valuable discussions and Prof. Dr. Eric Verschuur (Delft University of Technology) that kindly provided basic JMI algorithms at the beginning of this work. This study was funded by Petrobras, ANP, PRH-PB15, and PRH-PB230. The first author also thanks the exchange program scholarship provided by Capes/CNPq/DAAD process no. {88887.161391/2017-00}. Additional support for the authors was provided by the sponsors of the *Wave Inversion Technology (WIT) Consortium*, Hamburg, Germany.

## References

- Berkhout, A. G., 2014, Review paper: An outlook on the future of seismic imaging, Part III: Joint migration inversion: Geophysical Prospecting, **62**, 950–971.
- Bremmer, H., 1951, The W.K.B. approximation as the first term of a geometric-optical series: Communications on Pure and Applied Mathematics, **4**, 105–115.



Figure 2: Results from the Lens model. (a) Exact velocity model; (b) Exact image. Results from tomography with steepest descent: (c) Velocity; (d) Image. Results from tomography with Gauss-Newton: (e) Velocity; (f) Image.



Figure 3: Results from the Lens model, residual over iterations. (a) Data residual; (b) Model residual.



Figure 4: Marmousi2 model: Initial velocity model for JMI.

- de Bruin, C. G. M., C. P. A. Wapenaar, and A. J. Berkhout, 1990, Angle-dependent reflectivity by means of prestack migration: GEOPHYSICS, **55**, 1223–1234.
- Fong, D. C.-L., and M. Saunders, 2011, LSMR: An iterative algorithm for sparse least-squares problems: SIAM Journal on Scientific Computing, 33, 2950–2971.
- Grimbergen, J. L. T., F. J. Dessing, and K. Wapenaar, 1998, Modal expansion of one-way operators in laterally varying media: GEOPHYSICS, **63**, 995–1005.
- Jones, I., 2010, Tutorial: Velocity estimation via ray-based tomography: First Break, 28.



Figure 5: Results from the Marmousi2 model. (a) Exact velocity model; (b) Exact image. Results from tomography with steepest descent: (c) Velocity; (d) Image. Results from tomography with Gauss-Newton: (e) Velocity; (f) Image.



Figure 6: Results from the Marmousi2 model, residual over iterations. (a) Data residual; (b) Model residual.

- Kennett, B. L. N., M. S. Sambridge, and P. R. Williamson, 1988, Subspace methods for large inverse problems with multiple parameter classes: Geophysical Journal International, **94**, 237–247.
- Masaya, S., and D. J. E. Verschuur, 2018, Iterative reflectivity-constrained velocity estimation for seismic imaging: Geophysical Journal International.
- Pan, W., Y. Geng, and K. A. Innanen, 2018, Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and hussar land data set application: Geophysical Journal International, **213**, 1305–1333.
- Pratt, G., C. Shin, and Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341– 362.

- Thorbecke, J. W., K. Wapenaar, and G. Swinnen, 2004, Design of one-way wavefield extrapolation operators, using smooth functions in WLSQ optimization: GEOPHYSICS, **69**, 1037–1045.
- Ursin, B., Ø. Pedersen, and B. Arntsen, 2012, Fluxnormalized wavefield decomposition and migration of seismic data: GEOPHYSICS, 77, S83–S92.
- Wapenaar, C. P. A., 1996, One-way representations of seismic data: Geophysical Journal International, 127, 178–188.
- Xu, S., D. Wang, F. Chen, Y. Zhang, and G. Lambare, 2012, Full waveform inversion for reflected seismic data: Presented at the 74th EAGE Conference and Exhibition incorporating EUROPEC 2012, EAGE Publications BV.