Efficient Preconditioned Least-Squares Wave-Equation Migration

Abolhassani, Siamak; Verschuur, Dirk Jacob

DOI
10.1190/geo2023-0048.1

Publication date
2024

Document Version
Final published version

Published in
Geophysics

GEOPHYSICS, VOL. 89, NO. 3 (MAY-JUNE 2024); P. S275–S288, 17 FIGS.
10.1190/GEO2023-0048.1

DOI:10.1190/geo2023-0048.1

Efficient preconditioned least-squares wave-equation migration

Siamak Abolhassani[1] and Dirk Jacob Verschuur[1]

ABSTRACT

Since the appearance of wave-equation migration (WEM), many have tried to improve the resolution and effectiveness of this technology. Least-squares wave-equation migration is one of those attempts that try to fill the gap between migration assumptions and reality in an iterative manner. However, these iterations do not come cheap. A proven solution to limit the number of least-squares iterations is to correct the gradient direction within each iteration via the action of a preconditioner that approximates the inverse Hessian. However, the Hessian computation, or even the Hessian approximation computation, in large-scale seismic imaging problems involves an expensive computational bottleneck, making it unfeasible. Therefore, we develop an efficient computation of the Hessian approximation operator in the context of one-way WEM in the space-frequency domain. We build the Hessian approximation operator depth by depth, considerably reducing the operator size each time it is calculated. We prove the validity of our method with two numerical examples. We then extend our proposal to the framework of full-wavefield migration, which is based on WEM principles but includes interbed multiples. Finally, this efficient preconditioned least-squares full-wavefield migration is successfully applied to a data set with strong interbed multiple scattering.

INTRODUCTION

Seismic migration, also called seismic imaging, has been an imperative tool in characterizing the earth's subsurface geologic structures in the search for subsurface resources over the past 50 years. Numerous research efforts have been attempted to contribute to the theoretical developments of modern seismic migration, being one- or two-way wave-equation migration (WEM), notable among which are Claerbout (1971) and Claerbout and Doherty (1972) for introducing finite-difference migration, Stolt (1978) and Gazdag (1978) for inventing and developing migration in the wavenumber-frequency domain, Berkhout and Wulfften (1979) for introducing migration as a spatial deconvolution in the space-frequency domain, and also Baysal et al. (1983), Whitmore (1983), and McMechan (1983) for pioneering the use of reverse-time migration (RTM).

Most WEM algorithms share the same imaging mechanism. They first propagate the source wavefield forward into the medium while also propagating the receiver wavefield backward into the medium. They then construct the subsurface image by applying an imaging condition (Jones, 2014) to the forward and backward propagated wavefields at every trial image point. Note that WEM can also be interpreted and implemented as a generalized diffraction stack migration (Schuster, 2002; Zhan et al., 2014).

Despite the similarities, WEM algorithms differ in a couple of aspects, among which are the dimension — either time or depth — along which they propagate the wavefields and the scalar Helmholtz equation — either one-way or two-way — they solve to propagate the wavefields. Generally speaking, the most popular WEM algorithms can be split into two main kinds: one-way WEM and RTM (Etgen et al., 2009; Jones, 2014). Although RTM propagates wavefields along the time axis through the numerical solution of the two-way Helmholtz equation via direct methods such as finite-difference, spectral-element, and finite-element modeling, WEM propagates wavefields along the depth axis through the numerical solution of the one-way Helmholtz equation. WEM algorithms, to avoid the well-known computational overhead of direct solutions to the Helmholtz equation, are mostly built on the solution of an approximation to the square root of the Helmholtz operator, which is cheap, specifically for 3D cases (Mulder and Plessix, 2004). Although RTM outperforms WEM in addressing large propagation angles (greater than $\pm 90°$), therefore delivering superior images in complex subsurface regions, it is a more computationally intensive and memory-demanding operation. Hence, WEM is still a frequently used migration technique in the industry for high-frequency imaging or large-scale data sets (Mehta et al., 2017) if steep reflecting structures are not expected.

Manuscript received by the Editor 30 March 2023; revised manuscript received 17 January 2024; published ahead of production 19 February 2024; published online 25 April 2024.
[1]Delft University of Technology, Department of Imaging Physics (ImPhys), Laboratory of Computational Imaging, Faculty of Applied Sciences, Delft, The Netherlands. E-mail: s.abolhassani@tudelft.nl (corresponding author); d.j.verschuur@tudelft.nl.

DOI:10.1190/geo2023-0048.1

In addition, WEM principles may be used in so-called full-wavefield modeling (FWMod) (Berkhout, 2014b) to generate multiple scattering to apply full-wavefield migration (Berkhout, 2014a). From another perspective, however, it may be argued that interpreting such comparative statements between WEM and RTM as the outright dismissal of one over the other is a misconception. In fact, the preference for one over the other could be significantly affected by various factors, particularly the geologic complexity of the medium, the current stage of the project — whether it is in the early stage or full production mode — and, more importantly, efficient project resource allocation.

The WEM and RTM algorithms have proven effective on their own, provided that a few fundamental migration assumptions are satisfied: (a) their input data set is regularly sampled, (b) their input data set is free of multiples if the first-order Born approximation is followed, (c) there is no amplitude problem in their input data set due to the source energy variation/dissipation, (d) there is an accurate migration velocity model, (e) there is a broadband source function, and (f) there is an accurate migration operator. Otherwise, their output would not be an ideal representation of the subsurface reflectivity model, as they suffer from migration artifacts (Jones, 2018).

To resolve better images and close the gap between the migration assumptions and reality, the iterative least-squares migration (LSM) concept was introduced and leveled up any standard migration algorithm into a local minimization problem (e.g., Cole and Karrenbach, 1992; Chavent and Plessix, 1999; Duquet and Marfurt, 1999; Nemeth et al., 1999). While the number of iterations increases, any LSM technique can effectively suppress the part of the migration artifacts that appear due to irregular acquisition geometry, a band-limited source function, and geometric spreading (Huang et al., 2014). This iterative process, however, comes with high computational costs, so keeping the number of iterations to a minimum is crucial.

To reduce the number of least-squares iterations, one remedy is to precondition the gradient vector in each iteration with an approximation to the reciprocal of the Hessian matrix-operator (Pratt et al., 1998). As a result, the model perturbation vector $\Delta\mathbf{m}$ reads

$$\mathbf{H}^{a}\Delta\mathbf{m} = -\mathbf{g}, \tag{1}$$

in which $\mathbf{g}$ and $\mathbf{H}^{a}$ denote the gradient and Hessian approximation, respectively, and are given by

$$\mathbf{g} = \mathbf{J}^{\dagger}\Delta\mathbf{d}, \tag{2}$$

$$\mathbf{H}^{a} = \mathbf{J}^{\dagger}\mathbf{J}, \tag{3}$$

where $\Delta\mathbf{d}$ is the data error vector, † denotes the adjoint, and $\mathbf{J}$ is the forward (Born) modeling operator, or Jacobian matrix-operator, each column of which represents the wavefield scattered by a small perturbation of the model parameter m_i (i denotes the location of the model parameter here), while all the other model parameters are kept fixed. It mathematically follows that although the diagonal Hessian elements carry the scattered wavefield autocorrelations, the off-diagonal Hessian elements carry the scattered wavefield crosscorrelations of the neighboring model parameters in the medium (Operto et al., 2013). Although it is expected that the scattered wavefields exhibit only autocorrelations (i.e., only the on-diagonal coefficients in $\mathbf{H}^{a}$ have values), this is not the case. Indeed, due to the limited bandwidth of seismic sources and the proximity of model parameters in space, they are also partially cross-correlated (i.e., both the on-diagonal and off-diagonal coefficients in $\mathbf{H}^{a}$ have values). It follows that the Hessian approximation matrix has a diagonally dominant structure rather than a pure diagonal one, as noted by Pratt et al. (1998).

As is clear from equation 1, each element of the gradient vector is a weighted sum of all the model perturbations. This clarifies how blurred a migrated image will be using a forward-modeling operator with nonunitary columns (i.e., $\mathbf{H}^{a}$ is not an identity matrix) and, at the same time, explains why LSM, after the action of the inverse Hessian approximation on the gradient, converges faster (Aoki and Schuster, 2009). However, for large-scale seismic imaging problems, even such an approximation to the Hessian operator has a considerable computational burden — if it is rebuilt/updated in each iteration — as it requires each iteration to construct, invert, and multiply a matrix of size: number of model parameters × number of model parameters.

To make the computation of the Hessian operator feasible, some have replaced it with a diagonal approximation (e.g., Beydoun and Mendes, 1989; Chavent and Plessix, 1999; Shin et al., 2001), and others have approximated the entire set of coefficients with a manageable computational burden. Unlike the reciprocal of the diagonal coefficients, which only corrects the migration image for amplitude-related artifacts, the reciprocal of the entire set of coefficients can correct the migration image for bandwidth- and amplitude-related artifacts. Therefore, to gain a better image resolution, there have been many proposals to approximate the entire Hessian matrix, either in the data domain or image domain, among which the most remarkable are as follows. Hu et al. (2001) approximate the inverse Hessian matrix by a migration deconvolution filter constructed based on a single reference acoustic velocity profile $v(z)$, leaving laterally invariant deconvolution filters. For a layered medium with lateral velocity variations, Yu et al. (2006) construct a set of different deconvolution filters based on multiple acoustic velocity profiles to apply to different parts of the migrated image. Guitton (2004) approximates the inverse Hessian matrix by designing a matching convolution filter for each point in the image space to locally match a reference image to its corresponding migrated image in a least-squares sense. For a target-oriented imaging problem, Valenciano et al. (2006) compute a sparse-structure Hessian matrix explicitly through the crosscorrelation of the source and receiver Green's functions within the target area. Lecomte (2008) and Fletcher et al. (2016) calculate the Hessian approximation matrix using point spread functions (PSFs), in which each PSF measures the scattered wavefield for every point in the image space. Metivier et al. (2014) estimate the product of the Hessian approximation matrix and any vector that lies within the subsurface model space through the second-order adjoint-state method and then solve equation 1 using conjugate-gradient iterations. Assis and Schleicher (2021) use the same technique as Metivier et al. (2014) but in the context of one-way reflection waveform inversion (RWI). Lu et al. (2018), in the context of viscoacoustic anisotropic WEM, compute $\mathbf{J}$ and $\mathbf{J}^{\dagger}$ implicitly and then solve equation 1 using an iterative sparse solver of the least-squares QR factorization family. In their recent study, Yang et al. (2021) approximate the inverse Hessian by comparing the S-transform spectra of two images of the subsurface: one obtained through Born migration and the other obtained by remodeling the Born image and then remigrating it.

Over the past few years, there have also been significant studies in transforming the migration operator from an adjoint to a

pseudoinverse under the high-frequency assumption, commonly known as a "true-amplitude migration operator," mostly taking into account the geometric spreading loss. In such a context, Zhang et al. (2007) achieve a true-amplitude WEM by introducing the high-frequency approximation of the true-amplitude downgoing (forward) and upgoing (backward) wavefields at every trial image point. Kiyashchenko et al. (2007) propose a true-amplitude correlation-based imaging condition applicable to a finite-difference solution of the one-way wave equation, mathematically equivalent to the high-frequency approximation of the LSM solution (Bleistein et al., 2001). There exist alternative true-amplitude migration schemes that explicitly take into account the transmission effect as well (e.g., Deng and McMechan, 2007). Ten Kroode (2012) derives a pseudoinverse to the subsurface extended scale-separated Kirchhoff modeling operator, and Hou and Symes (2015) obtain a pseudoinverse to the subsurface extended Born modeling operator. In the same spirit as Hou and Symes (2015), Chauris and Cocher (2017) construct a pseudoinverse to the subsurface extended Born modeling operator after linearizing the phase of the Born operator, thereby promoting shorter offsets, smaller dips, and vertical rays in practice. A pseudoinverse operator, more importantly, can be advantageous in a least-squares framework also, as it can be applied to the residual data to speed up the convergence rate of the least-squares inversion scheme that relies merely on the steepest descent update direction (e.g., Hou and Symes, 2016).

This paper aims to present a cost-friendly Hessian approximation operator in the context of the space-frequency domain least-squares one-way wave-equation migration (LS-WEM) relying on angle-independent reflection/transmission coefficients (Berkhout, 1982, 2014a). To this purpose, we build the Hessian approximation operator depth by depth, significantly reducing the operator size each time it is calculated (Abolhassani and Verschuur, 2022). We also rebuild/update the Hessian approximation operator within each iteration, taking into account the updated transmission effects. This paper is organized as follows: First, we present our preconditioned LS-WEM theory (forward and inverse problems), in which the reciprocal of our proposed depth-based Hessian approximation operator preconditions the gradient direction. This is followed by two numerical examples for two synthetic data sets, including only primary reflections, one from a velocity model with a lens-shaped inclusion and the other from the SEG/EAGE overthrust velocity model. In addition, we evaluate the effectiveness of our proposed approach within the context of full-wavefield migration (Berkhout, 2014a) with a numerical example for a synthetic data set including strong interbed multiples. We finish with a final discussion and conclusion.

PRECONDITIONED LEAST-SQUARES ONE-WAY WAVE-EQUATION MIGRATION

The theoretical and algorithmic aspects of the preconditioned least-squares one-way wave-equation migration (PLS-WEM) are presented here.

Forward problem

The Kirchhoff integral for homogeneous fluids describes how to model the pressure wavefield at an arbitrary point inside a closed surface S when the pressure wavefield and the normal component of the particle velocity everywhere on S are known. The Kirchhoff integral can be simplified by choosing a plane surface for S and absorbing boundary conditions for the one-way Green's function (see, e.g., Berkhout and Wapenaar, 1989). The resultant integral is then represented as the Rayleigh II integral, which is the basis here for the one-way forward and adjoint wavefield extrapolations in a 2D inhomogeneous acoustic medium.

With an extrapolation step bounded by origin and destination depth levels, the 2D forward extrapolated wavefield, in which the wavefields existing at all lateral positions located at the origin depth level (z_ℓ) are extrapolated in the $+z$ direction toward a lateral position located at the destination depth level ($z_{\ell+1}$) (Figure 1), is defined as (known as the Rayleigh II integral)

$$p^+(x_j, z_{\ell+1}, \omega_f) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{p}^+(k_x, z_\ell, \omega_f) e^{-i\left(\sqrt{\left(\frac{\omega_f}{v_j}\right)^2 - k_x^2}\right)|\Delta z|} e^{-ik_x x_j} dk_x, \tag{4}$$

where v_j is the acoustic velocity at the position $(x_j, z_{\ell+1})$, ω_f denotes a given angular frequency component, k_x is the horizontal wavenumber, Δz denotes the laterally constant extrapolation step (the vertical distance between the origin and destination depth levels) and is assumed small enough to be considered homogeneous vertically, and $p^+(x_j, z_{\ell+1}, \omega_f)$ and $\tilde{p}^+(k_x, z_\ell, \omega_f)$ are the monochromatic downgoing acoustic wavefields that read the following forward and inverse Fourier conventions:

$$\tilde{p}(k_{x_j}) = \int_{-\infty}^{+\infty} p(x) e^{ik_{x_j} x} dx, \tag{5}$$

$$p(x_j) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{p}(k_x) e^{-ik_x x_j} dk_x. \tag{6}$$

Inserting equation 5 into equation 4 gives a convolution integral along the x-axis:

$$p^+(x_j, z_{\ell+1}, \omega_f) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} p^+(x, z_\ell, \omega_f) \overbrace{\left(\int_{-\infty}^{+\infty} e^{-i\left(\sqrt{\left(\frac{\omega_f}{v_j}\right)^2 - k_x^2}\right)|\Delta z|} e^{-ik_x(x_j - x)} dk_x\right)}^{w^+ : \text{extrapolation kernel}} dx$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} p^+(x, z_\ell, \omega_f) w^+_{z_{\ell+1}; z_\ell}(x_j - x, \omega_f) dx, \tag{7}$$

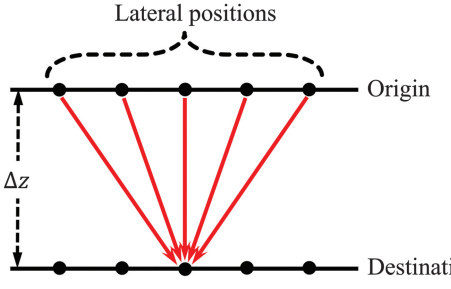

Figure 1. Extrapolation from all existing lateral positions located at the origin depth level toward a lateral position located at the destination depth level. Note that Δz denotes the extrapolation step.

in which the downward wavefield extrapolation is marked by w^+, and x_j denotes the convolution lag. Rewriting equation 7 into the vector-matrix form gives the following equations (Berkhout, 1982):

$$\mathbf{p}^+(z_{\ell+1},\omega_f) = \mathbf{W}^+_{z_{\ell+1};z_\ell}\mathbf{p}^+(z_\ell,\omega_f), \tag{8}$$

$$\mathbf{p}^-(z_\ell,\omega_f) = \mathbf{W}^-_{z_\ell;z_{\ell+1}}\mathbf{p}^-(z_{\ell+1},\omega_f), \tag{9}$$

where $\mathbf{p}^-(z_\ell,\omega_f)$ is the monochromatic upgoing acoustic wavefield at z_ℓ, $\mathbf{p}^+(z_\ell,\omega_f)$ is the monochromatic downgoing acoustic wavefield at z_ℓ, $\mathbf{p}^+(z_{\ell+1},\omega_f)$ is the monochromatic downgoing acoustic wavefield at $z_{\ell+1}$, $\mathbf{p}^-(z_{\ell+1},\omega_f)$ is the monochromatic upgoing acoustic wavefield at $z_{\ell+1}$, $\mathbf{W}^+_{z_{\ell+1};z_\ell}$ is the downward ($z_\ell \rightarrow z_{\ell+1}$) wavefield extrapolation matrix-operator, and $\mathbf{W}^-_{z_\ell;z_{\ell+1}}$ is the upward ($z_{\ell+1} \rightarrow z_\ell$) wavefield extrapolation matrix-operator.

With an extrapolation step without lateral velocity variations ($v_j = v$ everywhere at the destination depth level), $\mathbf{W}$ becomes a Toeplitz matrix-operator, i.e., each row vector-operator of $\mathbf{W}$ comes with a finite length of n_x and contains the same elements as others but is moved by one element to the right compared with its preceding row. However, with an extrapolation step including lateral velocity variations (a different v_j for every x_j at the destination depth level), $\mathbf{W}$ turns into a space-variant convolution matrix-operator, i.e., each row vector-operator of $\mathbf{W}$ is defined based on a locally averaged velocity, so it comes with a short spatial extent of L such that $L \ll n_x$ (Berkhout, 1982; Thorbecke et al., 2004).

To model the monochromatic angle-independent primary reflection data at the earth's surface (traditional surface seismic data excluding horizontally propagating waves) for a given shot location, the following equations, extrapolating the seismic source signature downward from z_0 (the earth's surface) to z_N and then upward from z_N to z_0, are solved in the space-frequency domain:

$$\mathbf{p}^-_{\text{mod}}(z_0,\omega_f) = \sum_{m=N}^{1}\mathbf{U}^-_{z_0;z_m}(\mathbf{r}^\cup(z_m) \circ \mathbf{p}^+_{\text{mod}}(z_m,\omega_f)), \tag{10}$$

$$\mathbf{p}^+_{\text{mod}}(z_m,\omega_f) = \mathbf{U}^+_{z_m;z_0}\mathbf{s}^+(z_0,\omega_f), \tag{11}$$

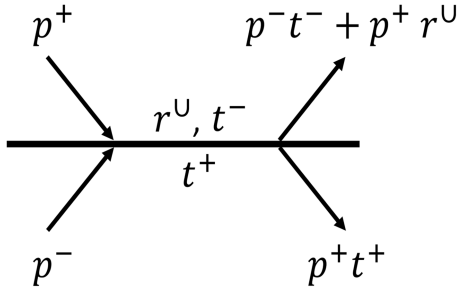

Figure 2. Wavefield updating in PWMod with angle-independent reflections at a given depth level, where p^+ is the downgoing wavefield, p^- is the upgoing wavefield, $r^\cup$ represents the upward reflectivity, and t^+ and t^- denote the downward and upward transmission.

$$\mathbf{U}^+_{z_m;z_0} = \left[\prod_{n=m-1}^{1}\mathbf{W}^+_{z_{n+1};z_n}\mathbf{T}^+(z_n)\right]\mathbf{W}^+_{z_1;z_0}, \tag{12}$$

$$\mathbf{U}^-_{z_0;z_m} = \left[\prod_{n=1}^{m-1}\mathbf{W}^-_{z_{n-1};z_n}\mathbf{T}^-(z_n)\right]\mathbf{W}^-_{z_{m-1};z_m}, \tag{13}$$

in which $\mathbf{p}^-_{\text{mod}}(z_0,\omega_f)$ denotes the monochromatic upgoing wavefield received at the depth level z_0, $\mathbf{p}^+_{\text{mod}}(z_m,\omega_f)$ indicates the monochromatic downgoing wavefield received at the depth level z_m, $\mathbf{r}^\cup(z_m)$ represents the angle-independent upward reflectivity vector-operator at z_m, $\mathbf{T}^+(z_n)$ is the downward transmission diagonal operator and reads $\mathbf{T}^+ = \mathbf{I} + \text{diag}(\mathbf{r}^\cup)$, $\mathbf{T}^-$ is the upward transmission diagonal operator and reads $\mathbf{T}^- = \mathbf{I} - \text{diag}(\mathbf{r}^\cup)$, $\mathbf{U}^+_{z_m;z_0}$ is called the total downward extrapolator and contains all the downward extrapolation operators ($\mathbf{W}^+$) coupled with the downward transmission operators required to reach from z_0 to z_m, $\mathbf{U}^-_{z_0;z_m}$ is called the total upward extrapolator and contains all the upward extrapolation operators ($\mathbf{W}^-$) coupled with the upward transmission operators required to reach from z_m to z_0, $\mathbf{s}^+(z_0,\omega_f)$ indicates the monochromatic downgoing physical source at z_0, N is the total number of the depth levels; and finally the symbol $\circ$ means the Hadamard product (Figure 2). Henceforward, we refer to the previously described primary wavefield modeling approach as PWMod.

As is clear from equations 10–13, PWMod calculates the earth's primary two-way response via a recursive summation (equation 10) in depth by including the reflection and transmission effects. The recursive summation, indeed, includes the multiplication of the total upward extrapolator and the upward scattered wavefield at each depth level.

Inverse problem

LS-WEM is a seismic depth migration technology based on a data-fitting process that seeks the earth's reflectivity model. LS-WEM iteratively minimizes a data error functional in a least-squares sense (Schuster, 2017). To minimize the data error functional, LS-WEM follows a migration-demigration cycle, as depicted in Figure 3. The LS-WEM cycle starts with an accurate migration velocity model fixed in the cycle and a zero-reflectivity model variable in the cycle. The modeled data are then generated via PWMod, and the data error is calculated and back-projected into the reflectivity model to build/update the reflectors via a steepest descent algorithm. The cycle iterates until the data error almost vanishes.

To introduce preconditioning in the LS-WEM algorithm, we seek the least-squares solution of the linear system (Lines and Treitel, 1984; Tarantola, 1984)

$$\Delta\mathbf{d}^- = \mathbf{J}\Delta\mathbf{r}^\cup, \tag{14}$$

where $\Delta\mathbf{d}^-$ represents the error between the observed and modeled data at z_0, $\Delta\mathbf{r}^\cup$ denotes the total upward reflectivity perturbation, and $\mathbf{J}$ is given by

$$\mathbf{J} = \frac{\partial\mathbf{p}^-_{\text{mod}}(z_0)}{\partial\mathbf{r}^\cup}. \tag{15}$$

Therefore, to solve equation 14, we minimize the following error functional (Shin et al., 2001; Jang et al., 2009; Oh and Min, 2013) as

a quadratic function of $\Delta\mathbf{r}^{\cup}(\omega_f)$:

$$\mathcal{C} = \frac{1}{2}\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\|\Delta\mathbf{d}_s^-(\omega_f) - \mathbf{J}_s(\omega_f)\Delta\mathbf{r}^{\cup}(\omega_f)\|_2^2, \tag{16}$$

in which

$$\Delta\mathbf{d}_s^-(\omega_f) = \mathbf{p}_{\text{obs},s}^-(z_0,\omega_f) - \mathbf{p}_{\text{mod},s}^-(z_0,\omega_f,\mathbf{r}^{\cup}), \tag{17}$$

and $\mathbf{p}_{\text{obs},s}^-(z_0,\omega_f)$ is the monochromatic multiple-free observed data recorded at the earth's surface for shot s, $\mathbf{p}_{\text{mod},s}^-(z_0,\omega_f,\mathbf{r}^{\cup})$ represents the monochromatic modeled primary data at the earth's surface for shot s, $\Delta\mathbf{d}_s^-(\omega_f)$ is the monochromatic residual data for shot s, ω_f represents an angular frequency component, N_s is the total number of shot locations, N_f indicates the total number of frequency components, and $\mathbf{r}^{\cup}$ is ordered as

$$\mathbf{r}^{\cup} = \begin{bmatrix}\mathbf{r}^{\cup}(z_0)\\ \mathbf{r}^{\cup}(z_1)\\ \vdots\\ \mathbf{r}^{\cup}(z_N)\end{bmatrix}. \tag{18}$$

Minimizing equation 16 with respect to $\Delta\mathbf{r}^{\cup}(\omega_f)$ gives the descent direction

$$\Delta\mathbf{r}^{\cup} = -\sum_{f=1}^{N_f}\Delta\mathbf{r}^{\cup}(\omega_f), \tag{19}$$

in which

$$\Delta\mathbf{r}^{\cup}(\omega_f) = \Re\left\{\left(\sum_{s=1}^{N_s}\overbrace{\mathbf{J}_s^{\dagger}(\omega_f)\mathbf{J}_s(\omega_f)}^{\mathbf{H}_s^{a}(\omega_f)}\right)\right\}^{-1}\Re\left\{\sum_{s=1}^{N_s}\overbrace{\mathbf{J}_s^{\dagger}(\omega_f)\Delta\mathbf{d}_s^-(\omega_f)}^{\mathbf{g}_s(\omega_f)}\right\}, \tag{20}$$

and, to link the linearized and nonlinear inverse problems, $\mathbf{g}$ and $\mathbf{H}^{a}$ are referred to as the gradient and Hessian approximation (also known as the Gauss-Newton Hessian approximation). The introduced frequency-dependent preconditioning operation here exhibits certain parallels with the deconvolution imaging condition introduced by Valenciano and Biondi (2003) that accounts for the source deconvolution.

Using equation 20, the reflectivity model can be updated iteratively via

$$\mathbf{r}_{k+1}^{\cup} = \mathbf{r}_k^{\cup} + \alpha_k\Delta\mathbf{r}_k^{\cup}, \tag{21}$$

where α denotes the minimization step length, and k denotes the current iteration. Because with PWMod, the acoustic wavefield is accessible depth by depth, the gradient vector in equation 20 can be rewritten as

$$\mathbf{g}_s(\omega_f) = \begin{bmatrix}\mathbf{g}_s(z_0,\omega_f)\\ \vdots\\ \mathbf{g}_s(z_N,\omega_f)\end{bmatrix} = \overbrace{\begin{bmatrix}\frac{\partial\mathbf{p}_{\text{mod},s}^-(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_0)} & \cdots & \frac{\partial\mathbf{p}_{\text{mod},s}^-(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_N)}\end{bmatrix}^{\dagger}}^{\text{total Jacobian}}\Delta\mathbf{d}_s^-(\omega_f), \tag{22}$$

in which each element of $\mathbf{g}$ is a vector representing the gradient associated with the model parameters located at a given depth level, and each column of the total Jacobian denotes the partial derivative of the upgoing modeled wavefield at the receiver locations with respect to the model parameters located at a given depth level. What is important to note here is that equation 22 allows us to construct not only the gradient vector depth by depth but also the Jacobian matrix, paving the way for constructing the Hessian approximation matrix with a reduced number of elements at each depth level.

Each column of the total Jacobian matrix, after calculating the corresponding partial derivatives, is simplified to

$$\frac{\partial\mathbf{p}_{\text{mod},s}^-(z_0,\omega_f)}{\partial\mathbf{r}^{\cup}(z_m)} = \begin{bmatrix}\mathbf{U}_{z_0;z_m}^-\begin{pmatrix}p_{1_{\text{mod},s}}^+(z_m,\omega_f)\\ 0\\ \vdots\\ 0\end{pmatrix}\mathbf{U}_{z_0;z_m}^-\begin{pmatrix}0\\ p_{2_{\text{mod},s}}^+(z_m,\omega_f)\\ \vdots\\ 0\end{pmatrix}\\ \cdots\mathbf{U}_{z_0;z_m}^-\begin{pmatrix}0\\ 0\\ \vdots\\ p_{n_{x_{\text{mod},s}}}^+(z_m,\omega_f)\end{pmatrix}\end{bmatrix}, \tag{23}$$

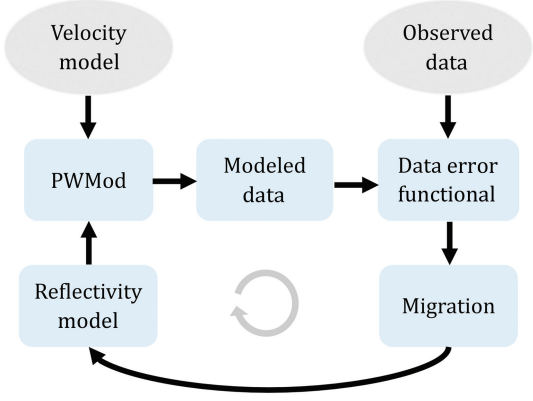

Figure 3. The LS-WEM cycle: LS-WEM starts with an accurate migration velocity model (fixed in the cycle) and a zero-reflectivity model (variable in the cycle). Then, using PWMod, the modeled data are generated and contrasted with the multiple-free observed data using a data error functional. The data error is then back-projected into the reflectivity model to build/update the reflectors through a scaled steepest descent algorithm. The cycle repeats itself until the data error almost disappears.

where n_x denotes the number of grid points at z_m (the inversion and modeling grids are similar), and $p^+_{j_{\mathrm{mod},s}}(z_m,\omega_f)$ is a complex number representing the downgoing modeled wavefield at the jth grid point of z_m. As the vectors multiplied by $\mathbf{U}^-_{z_0;z_m}$ in equation 23 each have only one active element, the multiplications can be simplified to

$$\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_m,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)} = [p^+_{1_{\mathrm{mod},s}}(z_m,\omega_f)\mathbf{U}^-_{*,1_{z_0;z_m}} \quad p^+_{2_{\mathrm{mod},s}}(z_m,\omega_f)\mathbf{U}^-_{*,2_{z_0;z_m}} \cdots \quad p^+_{n_{x_{\mathrm{mod},s}}}(z_m,\omega_f)\mathbf{U}^-_{*,n_{x_{z_0;z_m}}}], \tag{24}$$

in which $\mathbf{U}^-_{*,j_{z_0;z_m}}$ denotes the jth column of $\mathbf{U}^-_{z_0;z_m}$.

According to equation 3, the Hessian approximation matrix for each pair of source and frequency is expressed as

$$\mathbf{H}^{\mathrm{a}}_s(\omega_f) = \left[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}}\right]^{\dagger}\left[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}}\right], \tag{25}$$

and as can be seen, it shows a square and symmetric structure with the massive dimension of $n_m \times n_m$, where n_m represents the total number of model parameters in the whole medium. As an alternative to constructing one massive, computationally unfeasible Hessian approximation operator all at once, where it is required to perform the crosscorrelation between the partial derivative wavefields associated with the model parameters located at all depth levels (the whole medium), we split up the massive Hessian approximation operator into several smaller operators, each calculated for the model parameters located at a different depth level (n_x model parameters are located at each depth level). To do so, it suffices to write

$$\mathbf{H}^{\mathrm{a}}_s(z_m,\omega_f) = \left[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)}\right]^{\dagger}\left[\frac{\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)}{\partial \mathbf{r}^{\cup}(z_m)}\right], \tag{26}$$

where the reciprocal of $\mathbf{H}^{\mathrm{a}}_s(z_m,\omega_f)$ compensates for the geometric spreading and spatial correlations of the neighboring model parameters while also conducting source deconvolution.

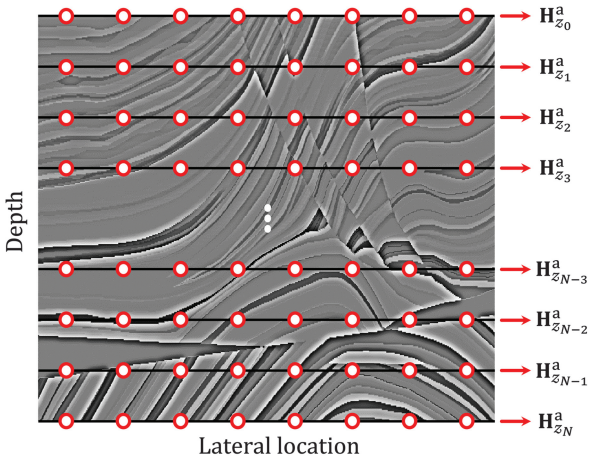

Figure 4. Decomposing the massive Hessian approximation operator into several small operators, each related to a depth level.

Therefore, to precondition every depth-level gradient, we calculate an individual tiny Hessian operator with the dimension of $n_x \times n_x$, where $n_x = n_m/N$ (Figure 4), reducing the number of model parameters by a factor of N whenever the Hessian operator is calculated. This is feasible because equation 24 enables us to construct $\partial \mathbf{p}^-_{\mathrm{mod},s}(z_0,\omega_f)/\partial \mathbf{r}^{\cup}$ for each depth level in the medium. This obviously allows us to decompose the Hessian approximation operator into several small operators, each of which only carries the correlation of the partial derivative wavefields associated with a single depth level. These minimal operators are computationally cheap to build, store, and invert in each iteration, enabling us to precondition the gradient vector efficiently. Figure 5 represents the structure of such an approximate Hessian.

NUMERICAL EXAMPLES OF PLS-WEM

To investigate the effect of the suggested approximate Hessian while solving equation 19 in each iteration, we contrast PLS-WEM with a scaled version of LS-WEM. In this scaled version, the LS-WEM gradient vector is scaled by the diagonal components of an inverse approximate Hessian (see Plessix and Mulder, 2004, equation 27), where the Hessian operator itself is given by

$$\mathbf{H} = \frac{\partial^2 \mathcal{C}'}{\partial \mathbf{r}^{\cup 2}} = \frac{\partial^2}{\partial \mathbf{r}^{\cup 2}}\left(\frac{1}{2}\sum_{f=1}^{N_f}\sum_{s=1}^{N_s}\|\Delta \mathbf{d}^-_s(\omega_f)\|_2^2\right), \tag{27}$$

and the approximate Hessian operator is derived by neglecting the terms in the aforementioned derivative that depend on the residual data, as discussed in Cova and Innanen (2013).

Both the PLS-WEM and LS-WEM methods are tested here on two synthetic data sets associated with a lens-shaped inclusion model and the SEG/EAGE overthrust model.

Lens-shaped inclusion model

In this example, our true model is a homogeneous model with a lens-shaped anomaly in the middle (Figure 6). The true model is

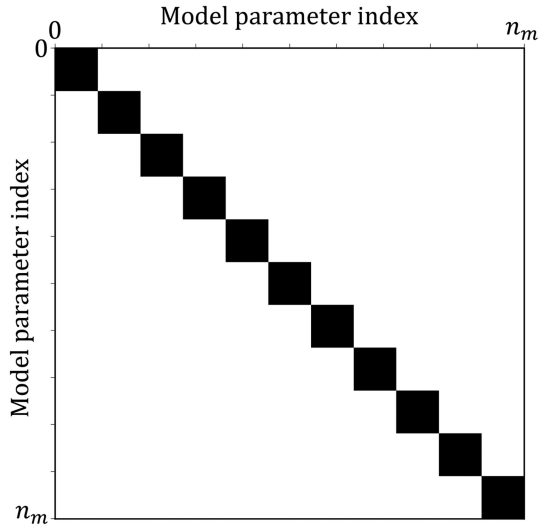

Figure 5. Proposed small Hessian approximation operators assembled into one big operator. Each block is a matrix of $n_x \times n_x$ related to a depth level, where $n_x = n_m/N$ and $N = 11$.

represented by 201 × 173 grid points in the horizontal and vertical directions. The horizontal and vertical grid sizes are 15 and 7 m, respectively. Forty-one shotpoints with 75 m spacing are laid out on top of the model, and 201 receivers with 15 m spacing are used per shot to record the reflection data (fixed-spread acquisition). The source function used in this example is a Ricker wavelet with a dominant frequency of 10 Hz. Our forward-modeling tool for generating the observed data set (only primaries) is PWMod, which is applied to the true model, the so-called inverse crime. The recording time is 1.4 s, and the largest offset used in the migration process is 3 km. Using an accurate migration velocity model with sharp reflectors, we apply LS-WEM and PLS-WEM in such a setup to evaluate the action of the reciprocal of the suggested depth-based Hessian approximation operator. Figure 7 depicts the results. Figure 7a, 7c, and 7e shows the LS-WEM results after 1, 3, and 5 iterations, respectively. Figure 7b, 7d, and 7f shows the PLS-WEM results after 1, 3, and 5 iterations, respectively.

The results show that the reciprocal of the suggested depth-based Hessian approximation operator significantly impacts the gradient vector from very early iterations. The image deconvolution and preserved amplitudes are clearly the most noticeable impacts of using the suggested preconditioner. Therefore, the PLS-WEM images demonstrate superior focusing compared with the LS-WEM images. As seen in Figure 7e and 7f, even after 5 least-squares iterations, the LS-WEM image quality does not reach the image quality of PLS-WEM after its first iteration.

To compare the convergence characteristics of LS-WEM and PLS-WEM, their normalized data error values in each iteration in the log scale are plotted in Figure 8a. As is evident, the PLS-WEM method shows a faster convergence ratio and also reaches smaller data errors, which means its estimated image fits the observed data better.

SEG/EAGE overthrust model

In the next example, our true model is one vertical slice of the 3D SEG/EAGE overthrust model (Figure 9) (Aminzadeh et al., 1994). The selected model contains 501 × 151 grid points in the horizontal and vertical directions, respectively, with a grid interval of 24 m. One hundred and one shotpoints with an interval of 120 m are used on top of the model, and 501 receivers per shot with an interval of 24 m collect the reflection data (fixed-spread acquisition). PWMod is used to generate the observed reflection data (only primaries) using a Ricker wavelet with a dominant frequency of 10 Hz as the source function, the so-called inverse crime. The trace length is 2.6 s, and the largest offset used in the migration process is 4 km. With the aid of an accurate migration velocity model including sharp reflectors, both LS-WEM and PLS-WEM are tested in such a setup to examine the action of the reciprocal of the suggested depth-based Hessian approximation operator on the gradient. The results are shown in Figure 10. Figure 10a, 10c, and 10e shows the LS-WEM results after 1, 3, and 5 iterations, respectively. Figure 10b, 10d, and 10f shows the PLS-WEM results after 1, 3, and 5 iterations, respectively.

The results confirm the effectiveness of the action of the suggested preconditioner on the gradient. From the first iteration, balanced-amplitude reflectors and image deconvolution are again the most visible outcomes of applying the depth-based Hessian approximation inverse on the gradient vector. Although LS-WEM, in its early iterations, leaves us with an image affected by the limited bandwidth of the seismic data, PLS-WEM mitigates such an unfavorable effect and estimates a high-resolution image from early iterations. As demonstrated in Figure 10e and 10f, even after 5 iterations, the unfocused reflection energies are visible in the LS-WEM image, but the PLS-WEM image appears to be focused clearly.

In Figure 11, a specific region of the estimated images after 5 iterations, specified by the dashed yellow rectangles in Figure 10e and 10f, is presented to obtain the magnified sections for deeper investigation. Although in Figure 11b (PLS-WEM), the faults and curvatures are accurately imaged together with focused, stronger, and accurate reflectors, in Figure 11a (LS-WEM), the faults and curvatures are mapped inaccurately together with unfocused weaker reflectors. Figure 12 also compares the 1D vertical reflectivity profiles estimated by LS-WEM and PLS-WEM with the true profile at the lateral location of 6 km.

The convergence properties of LS-WEM and PLS-WEM are compared in Figure 8b by plotting their data error values for each

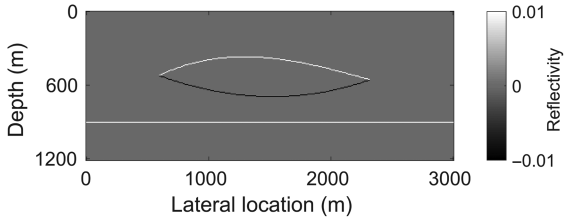

Figure 6. True lens-shaped inclusion reflectivity model.

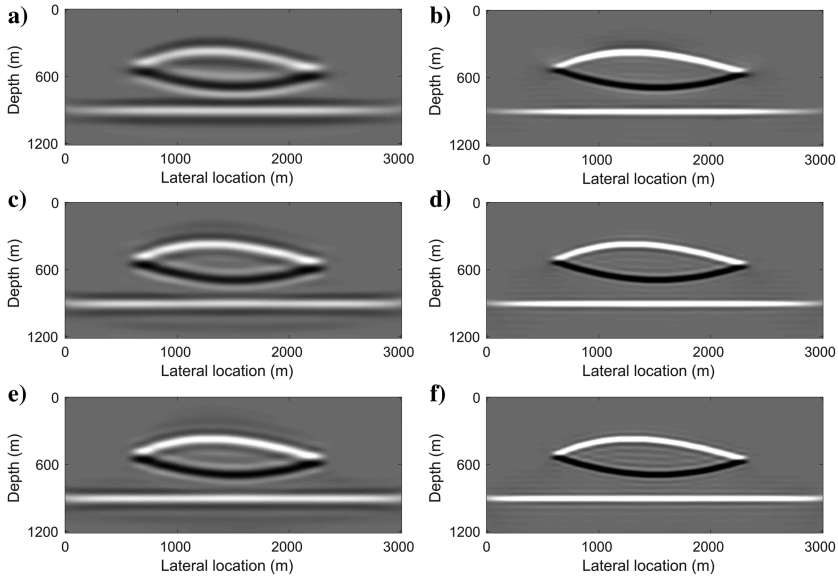

Figure 7. Estimated reflectivity models using the LS-WEM and PLS-WEM methods after 1, 3, and 5 iterations associated with the true lens-shaped inclusion model shown in Figure 6. (a, c, and e) Estimated reflectivity models by LS-WEM after 1, 3, and 5 iterations, respectively. (b, d, and f) Estimated reflectivity models by PLS-WEM after 1, 3, and 5 iterations, respectively.

iteration. As expected, the PLS-WEM method shows a higher convergence ratio and achieves smaller data errors, which means that its estimated image is more in line with the observed data. Figure 13 represents the initial and final residual data, confirming the convergence properties.

APPLICATION TO LEAST-SQUARES FULL-WAVEFIELD MIGRATION

In the preceding examples, we ignored multiple scattering while modeling the seismic data. As a result, to avoid matching data events with different scattering orders, known as cross-talk noise or multiple imprints, the observed data events caused by multiple scattering were required to be detected and eliminated prior to the migration process using a multiple elimination algorithm (Berkhout and Verschuur, 1997, 2005; Weglein et al., 1997, 2003; van Groenestijn and Verschuur, 2009; Ypma and Verschuur, 2013; Slob et al., 2014; Siahkoohi et al., 2019; Zhang and Slob, 2020; Thorbecke et al., 2021). Although eliminating multiple scattering offers benefits, it can also be a challenging task or exceed the budgetary limits in some seismic applications. Therefore, from this perspective, it would be ideal if a migration algorithm could effectively manage multiple scattering together with first-order scattering waves.

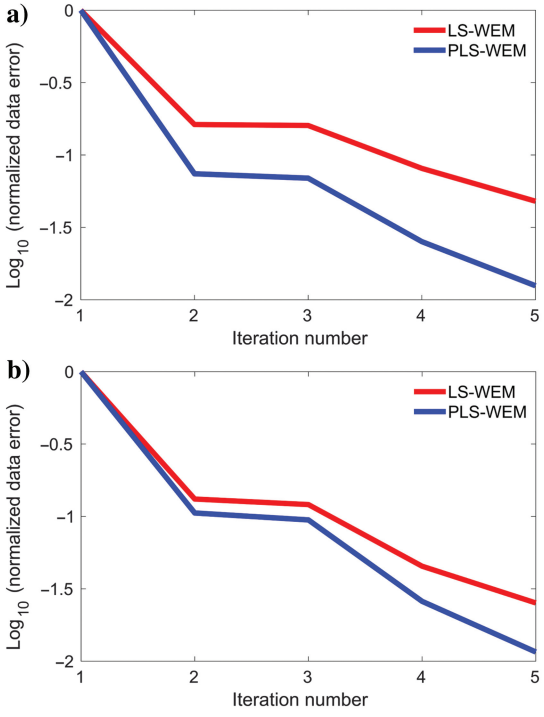

Figure 8. Convergence curves: (a) the lens-shaped inclusion model and (b) the SEG/EAGE overthrust model.

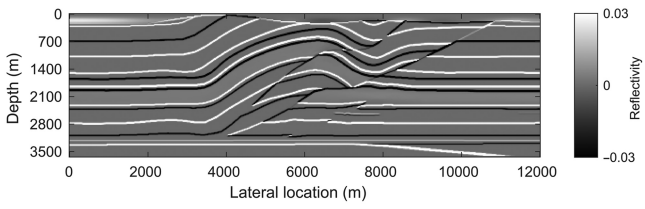

Figure 9. True SEG/EAGE overthrust reflectivity model.

Some studies have shown that when a migration algorithm effectively manages multiple scattering, it can enhance illumination, especially in areas where primary imaging fails to provide adequate illumination due to missing/masked-by-multiple primary reflections in the observed data (Berkhout and Verschuur, 2016; Davydenko and Verschuur, 2017; Lu et al., 2018; Slob et al., 2021). Moreover, when reflectors are better focused from early iterations through an effective gradient preconditioning, multiple scattering modeling will be more accurate, which should result in a better reduction of the cross-talk noise on the final stacked image. This is indeed expected as the outcome of a clearer distinction between the primary and multiple reflections during data matching, especially when they overlap.

Taking into account interbed multiples in modeling, the following example first examines the successful application (enhanced resolution and stronger reflectivities) and subsequently investigates the expected improvement (better cross-talk noise reduction), if any, of the suggested depth-based preconditioner within the context of the least-squares full-wavefield migration (LS-FWM) algorithm, which was first introduced by Berkhout (2014a).

LS-FWM is a nonlinear least-squares one-way wave-equation migration technique that uses the full (primary and multiples) two-way earth's response by incorporating multiple scattering (via iterative modeling) into the forward-modeling theory already presented in the forward problem section. LS-FWM and LS-WEM share the same inversion cycle (Figure 3), with the difference being that LS-FWM uses FWMod (Berkhout, 1982, 2014b) rather than PWMod. The LS-FWM gradient vector, similar to the LS-WEM gradient vector, in each iteration is scaled by the diagonal components of the inverse approximate Hessian derived by the linear part of equation 27.

Angle-independent FWMod models the angle-independent primary reflections at the earth's surface by extrapolating the seismic source signature first downward from z_0 to z_N and then upward from z_N to z_0 by accounting for reflection and transmission effects, which is called one roundtrip/iteration. Subsequent FWMod roundtrips contribute to new orders of multiple scattering while still explaining the reflection and transmission effects. Compared with PWMod, FWMod takes into account the multiple scattering by replacing equation 11 with

$$\mathbf{p}^{+}_{\mathrm{mod}}(z_m,\omega_f)=\sum_{m=1}^{N}\mathbf{U}^{+}_{z_m;z_0}(\mathbf{s}^{+}(z_0,\omega_f)+\mathbf{r}^{\cap}(z_m)\circ\mathbf{p}^{-}_{\mathrm{mod}}(z_m,\omega_f)), \tag{28}$$

where $\mathbf{r}^{\cap}(z_m)$ represents the angle-independent downward reflectivity vector-operator at z_m, and $\mathbf{r}^{\cap}=-\mathbf{r}^{\cup}$ (acoustic medium assumption) (Figure 14).

As is clear, FWMod, in its first roundtrip, calculates the earth's two-way response (primary reflections) via a recursive summation in depth by including the reflection and transmission effects. FWMod has control over multiple scattering, and each roundtrip (other than the first roundtrip) adds an order of scattering to the primaries. As a result, with FWMod, modeling the primary reflections, surface-related multiples, interbed multiples, and total reflection wavefields are easily accessible.

PLS-FWM minimizes the same error functional as PLS-WEM, i.e., equation 16, with the difference that $\mathbf{p}^-_{\text{obs}}$ and $\mathbf{p}^-_{\text{mod}}$ in the equation are now the full (primaries and multiples) observed and modeled data at the earth's surface. Other details of the method can be found in Berkhout (2014a). As a result, for deriving the suggested preconditioned model update in the PLS-FWM framework, nothing mathematically changes from what is already derived for PLS-WEM, as FWMod can be seen as an iterative PWMod, where the upgoing/downgoing wavefields stored at different depth levels in former roundtrips are used as the input wavefields for the subsequent roundtrips. Therefore, the same theory is used here to build the preconditioned model update for PLS-FWM in each iteration.

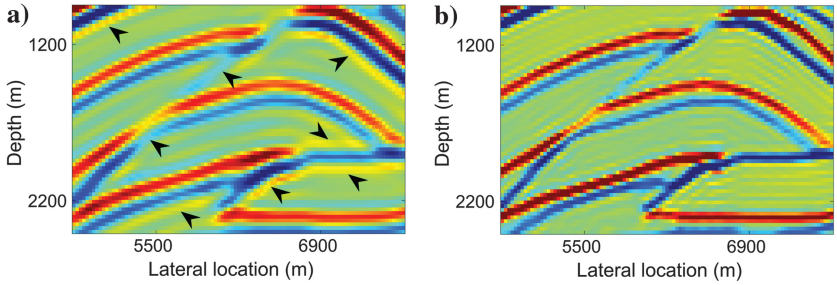

Figure 11. Magnified sections of the estimated SEG/EAGE overthrust reflectivity models by LS-WEM and PLS-WEM after 5 iterations, shown in Figure 10e and 10f. (a) Magnified section of the estimated reflectivity model by LS-WEM and (b) magnified section of the estimated model by PLS-WEM.

NUMERICAL EXAMPLE OF PLS-FWM

To evaluate the application and impact of the suggested depth-based preconditioner within the framework of least-squares full-wavefield migration, we use a multireflector wedge model placed between two horizontal reflectors, in which each reflector is only due to the density contrast, as we use a homogeneous velocity of 2000 m/s to build the reflectivity model. This extreme example yields alternating local reflection coefficients of $r = 1/5$ and $r = -1/5$ in the model, causing medium-strength interbed multiples in the data compared with the reflection coefficients used in Slob et al. (2021). Figure 15 shows the true resonant wedge density model in kg/m^3. To add the resonant feature to the model, the

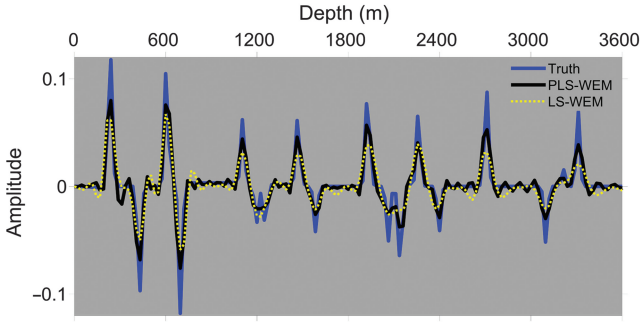

Figure 12. Estimated 1D reflectivity profiles using the LS-FWM and PLS-FWM methods after 5 iterations associated with the SEG/EAGE overthrust model. The profiles are taken from the estimated reflectivity models shown in Figure 10e and 10f and represent the lateral location of 6 km.

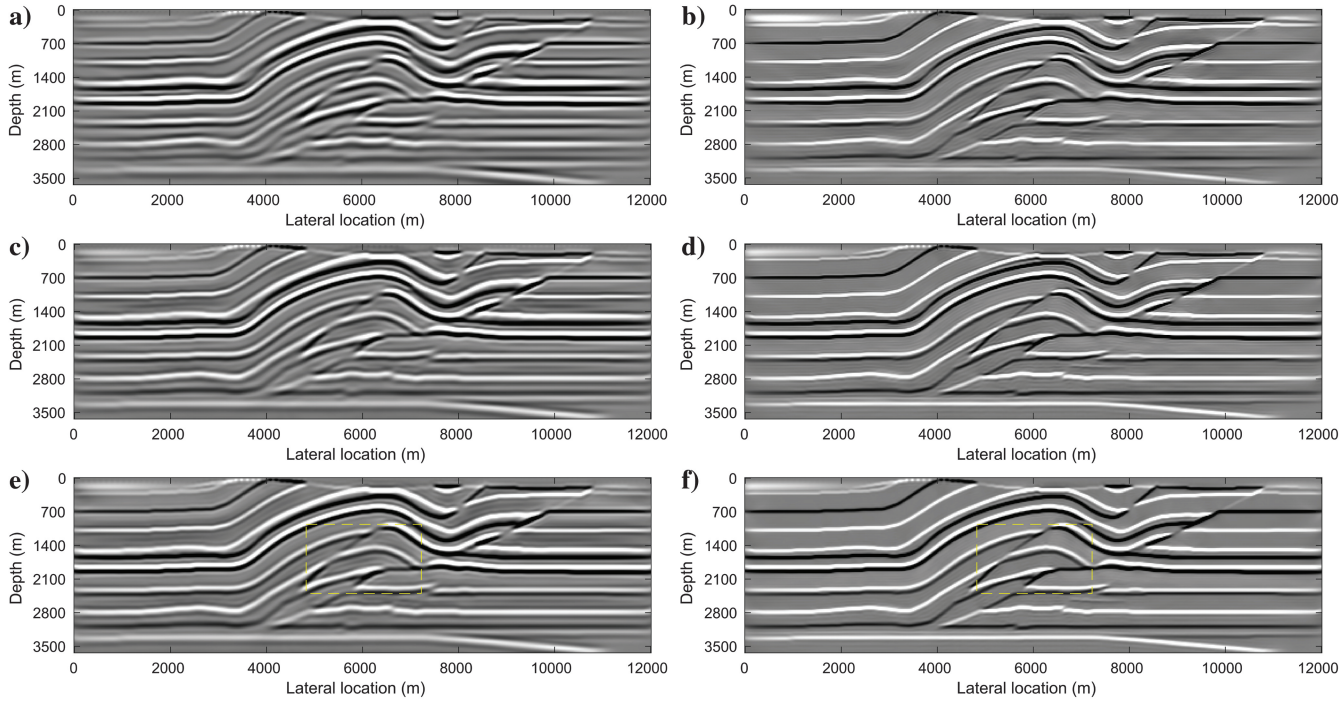

Figure 10. Estimated reflectivity models using the LS-WEM and PLS-WEM methods after 1, 3, and 5 iterations associated with the true SEG/EAGE overthrust reflectivity model shown in Figure 9. (a, c, and e) Estimated reflectivity models by LS-WEM after 1, 3, and 5 iterations, respectively. (b, d, and f) Estimated reflectivity models by PLS-WEM after 1, 3, and 5 iterations, respectively.

dipping layers in the wedge part all have the same thickness. Besides, whereas each layer on the left side has a thickness of 100 m, the thickness approaches zero on the right. An acoustic finite-difference modeling scheme with a grid size of 1 m generates the observed reflection data. A 15 Hz Ricker wavelet is used as a source wavelet. Thirty-one shots with a 120 m shot spacing are deployed on top of the model, such that the first shot is placed at 200 m and the last shot at 3800 m. Each observed shot gather includes 401 traces with 10 m trace spacing (fixed-spread acquisition), and the length of the traces is 2.92 s. With the same source-receiver setup and the true homogeneous velocity model, FWMod, with a horizontal grid size of 10 m and a vertical grid size of 4 m, generates the modeled reflection data in every iteration. The maximum frequency used in the migration is 40 Hz, and the largest offset used in the migration is 4 km. Note that in Slob et al. (2021) it is already shown that such a model can be viewed as an extreme scenario when interbed multiples mask primaries. With this experimental setup, LS-FWM and PLS-FWM are used here to assess the extent of the achievement of the reciprocal of the suggested depth-based Hessian approximation operator in preconditioning the gradient direction within the context of full-wavefield migration. Figure 16 shows the results.

Figure 16a and 16b shows the LS-FWM and PLS-FWM output images after 5 iterations, respectively. For the purpose of analysis, the output images can be laterally split into three main regions: the first 1000 m of the lateral distance of the image (the left-edge region), from 1000 to 3000 m of the lateral distance of the image (the middle region), and the last 1000 m of the lateral distance of the image (the right-edge region). For the left-edge region as deep as the dipping layers, LS-FWM delivers an alternating sequence of positive and negative amplitudes. Consequently, it is challenging to determine which ones represent the real dipping reflectors. In contrast, by removing the source signature and balancing the reflector amplitudes, PLS-FWM delivers a clean view of the real dipping reflectors there. For the middle region as deep as the dipping layers, we observe that PLS-FWM, compared with LS-FWM, effectively performs better and maps the reflectors with less blurriness. For the right-edge region as deep as the dipping layers, where the resolution limit is reached, we notice that for recovering the fifth dipping reflector (the yellow arrow), whereas PLS-FWM does not stably function, LS-FWM functions slightly more stably.

We can also see that after 5 iterations, both images exhibit a common characteristic, that is, the cross-talk noise, or equally ghost reflectors, generated in the image owing to the presence of interbed multiples in the data. As a result, the last real dipping reflector is not retrieved as strongly as the others, as denoted by the green arrow in Figure 16b, and several ghost reflectors are also built, as indicated by the blue arrows in Figure 16b.

Figure 16c and 16d shows the LS-FWM and PLS-FWM results after 15 iterations, respectively. Comparing the images reveals that the ghost reflectors existing in the fifth-iteration images are now attenuated after 15 iterations, as shown by the blue arrows in Figure 16d. In addition, we observe that the weak real dipping reflector in the fifth-iteration images is now recovered much stronger after addressing both primaries and interbed multiples (see also

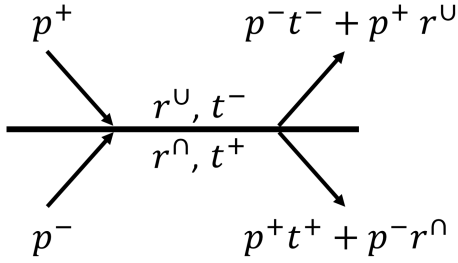

Figure 14. Wavefield updating in FWMod with angle-independent reflections at a given depth level, where p^+ is the downgoing wavefield, p^- is the upgoing wavefield, $r^\cup$ represents the upward reflectivity, $r^\cap$ shows the downward reflectivity, and t^+ and t^- denote the downward and upward transmission, respectively.

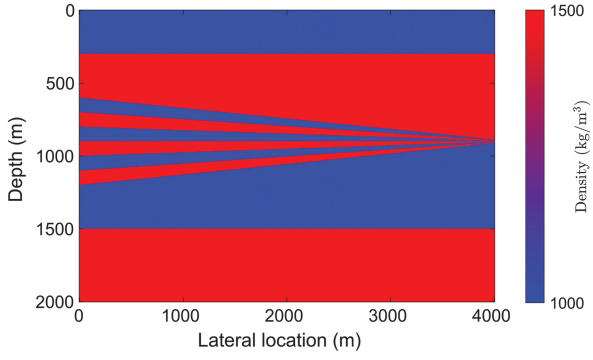

Figure 15. Variations in the true resonant wedge density model in kg/m^3. The change in density between the layers is 50%, and the dipping layers all have identical thicknesses.

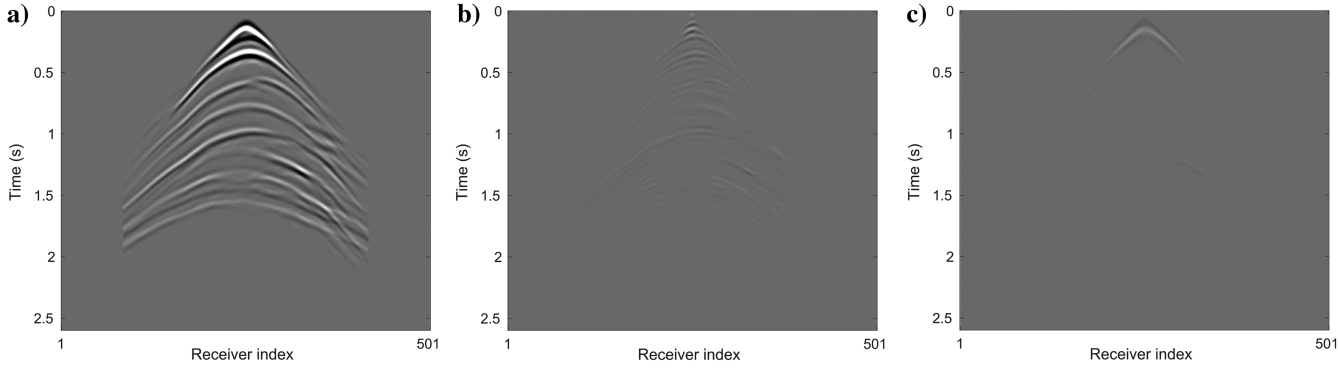

Figure 13. Residual data associated with the SEG/EAGE overthrust model. (a) Initial residual data, (b) final residual data with LS-WEM (iteration 5), and (c) final residual data with PLS-WEM (iteration 5).

Slob et al., 2021), as denoted by the green arrows in Figure 16d, but still not as strong as the other real dipping reflectors. We also notice that the ghost reflectors are marginally better reduced by PLS-WEM, as denoted by the red arrows in Figure 16c (see also the "Discussion" section).

Figure 17 compares the 1D reflectivity profiles corresponding to the LS-FWM and PLS-FWM images in iteration 15, depicted in Figure 16c and 16d, at the lateral location of 2150 m. It clearly proves the accuracy of the PLS-FWM image in comparison to the LS-FWM image in terms of the resolved resolution, reflectivity magnitudes, and the reduction of multiple cross talk (the red arrows). Therefore, we can state that PLS-FWM outperforms LS-FWM in terms of image quality.

DISCUSSION

Despite the limitations outlined in the "Introduction" section, the industry has continued to embrace the practice of mapping primary energies back into the image space using LS-WEM due to its notable advantage of lower computational cost when compared with LS-RTM. Although LS-WEM has proven itself quite effective in recovering decent images, we observed that our newly developed PLS-WEM algorithm outperforms LS-WEM. It shows faster convergence and superior resolution. Moreover, it is efficient in computation when compared with an alternative solution that computes a huge Hessian approximation operator for the entire medium. Although feasible, it still remains expensive in terms of time compared with LS-WEM, with the current implementation and contemporary computing resources. However, based on our internal

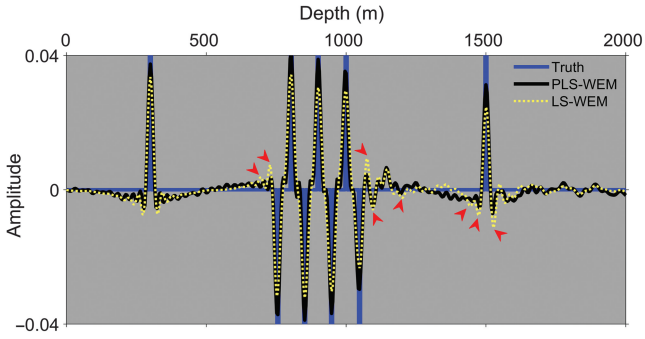

Figure 17. Estimated 1D reflectivity profiles using the LS-FWM and PLS-FWM methods after 15 iterations associated with the resonant wedge model. The profiles are taken from the estimated reflectivity models shown in Figure 10e and 10f and represent the lateral location of 2150 m.

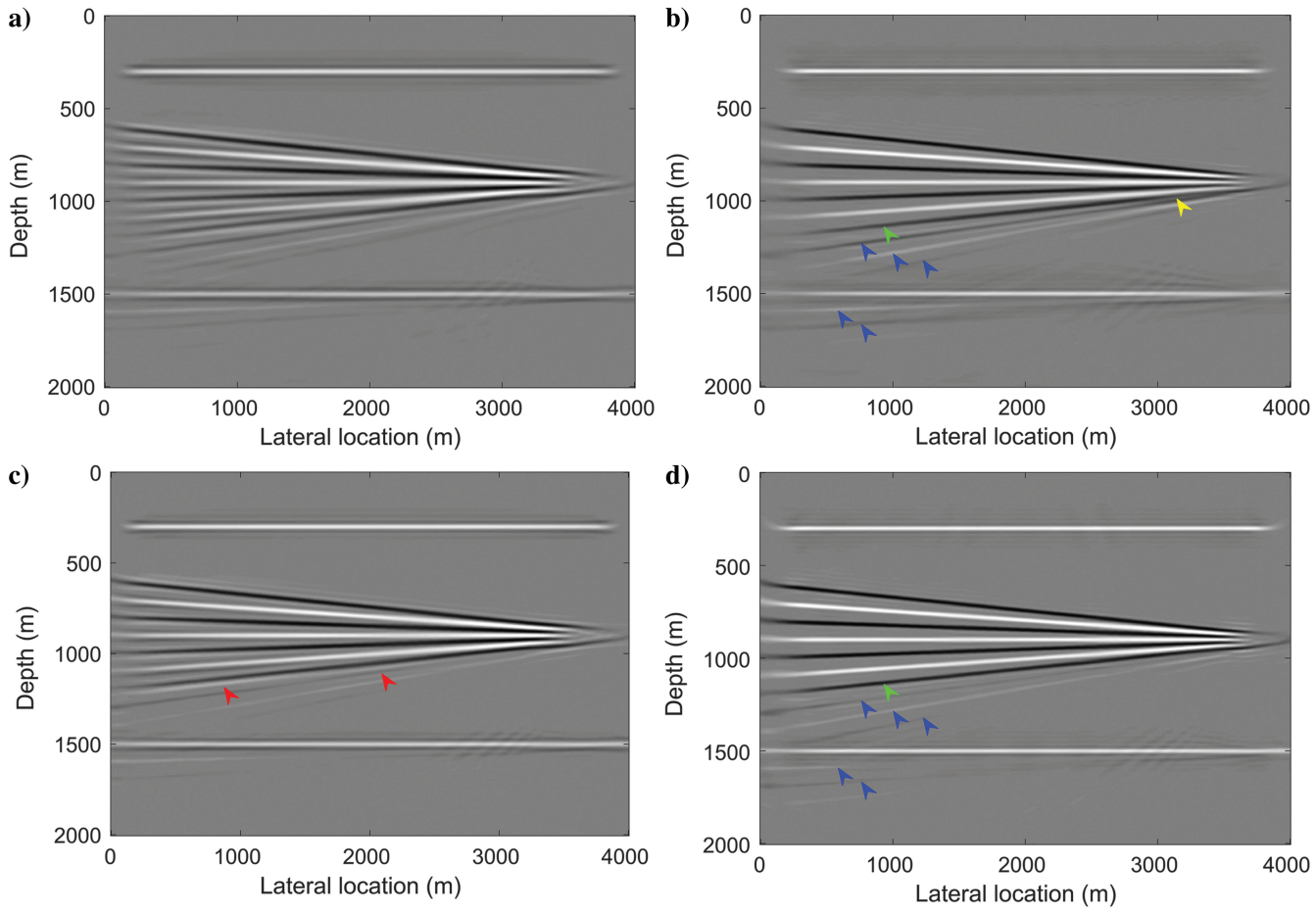

Figure 16. Estimated reflectivity models using the LS-FWM and PLS-FWM methods after 5 and 15 iterations associated with the true resonant wedge density model shown in Figure 15. (a and c) Estimated reflectivity models by LS-WEM after 5 and 15 iterations, respectively. (b and d) Estimated reflectivity models by PLS-WEM after 5 and 15 iterations, respectively.

investigations, PLS-WEM with source subsampling in the computation of the Hessian approximation operators (Matharu and Sacchi, 2019) — one source out of three — can generate similar high-resolution images, comparable to this study in half the iterations of the LS-WEM, although its wall time for each iteration is observed to be double to triple of LS-WEM.

It is important to acknowledge that both the LS-WEM and PLS-WEM algorithms presented in this paper may underperform due to flaws in the multiple elimination stage, poor wavelet estimation, the angle-independent reflection coefficient assumption, steep reflectors, and an inaccurate migration velocity model.

In the case of the full-wavefield migration example, although we expected better resolution and stronger reflection coefficients from PLS-FWM compared with LS-FWM, we also hoped for more reduction of cross-talk noise. According to the resonant wedge example results, even though PLS-FWM achieves greater resolution and stronger reflectivities, the cross-talk noise is not better reduced. As previously stated, better-focused reflectors are expected to end in more accurate modeling of multiple scattering and, hence, better attenuation of multiple imprints. However, this doesn't seem to work as well as we thought it would, although it gave slight improvements in reducing the cross-talk noise (see the red arrows in Figures 16c and 17). This capacity may be partially underused due to the double-edged nature of preconditioning; the preconditioner does its function on both real and ghost reflectors in each iteration, giving high-resolution ghost and real reflectors for modeling, leading to a conflict between the real and fake reflections in the corresponding data minimization problem, still ending up in a local minimum. From another perspective, the underperformance could also be attributed to the highly nonlinear nature of the resonant wedge experiment. Lastly, the underperformance might be linked to the FWMod convergence issue with spatially inhomogeneous media (McMaken, 1986). To ascertain the main cause, further investigation must be conducted.

Typically, the convergence of LSM exhibits a faster rate in the initial iterations and a slower rate in the subsequent iterations. In Figure 12a and 12b, it can be observed that the data error between the second and third iterations remains almost identical, indicating a minimal change in the image update. The observed phenomenon can be attributed to the nonlinear dynamics introduced into the modeling process by the implicit accounting for the transmission effects in the modeled data, as shown by the factor $(1 + r^{\cup})(1 - r^{\cup}) = 1 - (r^{\cup})^2$. In the first iteration of LSM, there is no information about the reflection coefficients. This lack of information may lead to an overcorrection of the transmission coefficients with the nonlinear factor, resulting in modeled data that are not amplitude-consistent with the observed data. To correct this, the next iteration of LSM performs data amplitude balancing, resulting in a relatively similar data error when compared with the previous iteration. With each subsequent iteration, this issue gradually irons itself out, bringing the modeled data into close alignment with the observed data.

Superior imaging can also help build high-resolution and strong tomographic updates in RWI. RWI is described as a method for constructing a low-wavenumber velocity model via confining the full-waveform inversion sensitivity kernel along the transmission-after-reflection wavepaths. RWI sequentially solves a multiparameter primary reflection data-driven inverse problem with velocity and reflectivity as the model parameters. In RWI, reflectors are first mapped using a migration technology, and then the tomographic update is built based on the mapped reflectors. As a result, in such a flow, a higher-resolution image with stronger amplitudes may lead to a more consistent and stronger tomographic update, something that our proposed PLS-WEM can facilitate.

CONCLUSION

We mathematically showed how the Hessian approximation operator and its reciprocal as the gradient preconditioner could be efficiently built for LS-WEM in a depth marching regime. We showed that our proposed cost-friendly PLS-WEM algorithm, thanks to PWMod, comes with a minimal computational effort in each iteration in comparison to its alternatives that compute a massive Hessian approximation for the entire medium. Using two numerical examples, we verified how the proposed preconditioner effectively cuts the migration artifacts generated by the band-limited nature of seismic data and preserves the reflectivity amplitudes. We also confirmed that the improvement in migration images results in a faster convergence ratio and a better data fit for PLS-WEM. We also used the proposed preconditioner in the context of full-wavefield migration. With an extreme scenario generating strong interbed multiples, we showed that PLS-FWM outperforms LS-FWM in the same two ways that PLS-WEM outperforms LS-WEM. That is, PLS-FWM removes the source signature from the image quickly and recovers stronger and more accurate reflectivities than LS-FWM. We also observed that PLS-FWM only provided limited improvement in minimizing the cross-talk noise, contrary to our expectations.

ACKNOWLEDGMENTS

The authors express their gratitude to the sponsors of the Delphi consortium for their stimulating discussions, valuable feedback, and financial support.

DATA AND MATERIALS AVAILABILITY

No data have been required for this paper.

REFERENCES

Abolhassani, S., and E. Verschuur, 2022, Fast Gauss-Newton full-wavefield migration: Second International Meeting for Applied Geoscience & Energy, SEG, Expanded Abstracts, 2709–2713, doi: 10.1190/image2022-3751008.1.

Aminzadeh, F., N. Burkhard, L. Nicoletis, F. Rocca, and K. Wyatt, 1994, SEG/EAGE 3-D modeling project: 2nd update: The Leading Edge, **13**, 949–952, doi: 10.1190/1.1437054.

Aoki, N., and G. T. Schuster, 2009, Fast least-squares migration with a deblurring filter: Geophysics, **74**, no. 6, WCA83–WCA93, doi: 10.1190/1.3155162.

Assis, C. A., and J. Schleicher, 2021, Introduction of the Hessian in joint migration inversion and improved recovery of structural information using image-based regularization: Geophysics, **86**, no. 6, R777–R793, doi: 10.1190/geo2020-0730.1.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, doi: 10.1190/1.1441434.

Berkhout, A. J., 1982, Seismic migration: Imaging of acoustic energy by wave field extrapolation: Elsevier.

Berkhout, A. J., 2014a, Review paper: An outlook on the future of seismic imaging, Part II: Full-wavefield migration: Geophysical Prospecting, **62**, 931–949, doi: 10.1111/1365-2478.12154.

Berkhout, A. J., 2014b, Review paper: An outlook on the future seismic imaging, Part I: Forward and reverse modelling: Geophysical Prospecting, **62**, 911–930, doi: 10.1111/1365-2478.12161.

Berkhout, A. J., and D. J. Verschuur, 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations: Geophysics, **62**, 1586–1595, doi: 10.1190/1.1444261.

Berkhout, A. J., and D. J. Verschuur, 2005, Removal of internal multiples with the common-focus-point (CFP) approach: Part 1 — Explanation of the theory: Geophysics, **70**, no. 3, V45–V60, doi: 10.1190/1.1925753.

Berkhout, A. J., and D. J. Verschuur, 2016, Enriched seismic imaging by using multiple scattering: The Leading Edge, **35**, 128–133, doi: 10.1190/tle35020128.1.

Berkhout, A. J., and C. P. A. Wapenaar, 1989, One-way versions of the Kirchhoff integral: Geophysics, **54**, 460–467, doi: 10.1190/1.1442672.

Berkhout, A. J., and P. V. Wulfften, 1979, Migration in terms of spatial deconvolution: Geophysical Prospecting, **27**, 261–291, doi: 10.1111/j.1365-2478.1979.tb00970.x.

Beydoun, W. B., and M. Mendes, 1989, Elastic ray-Born L2-migration/inversion: Geophysical Journal International, **97**, 151–160, doi: 10.1111/j.1365-246X.1989.tb00490.x.

Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of multidimensional seismic imaging and inversion: Springer.

Chauris, H., and E. Cocher, 2017, From migration to inversion velocity analysis: Geophysics, **82**, no. 3, S207–S223, doi: 10.1190/geo2016-0359.1.

Chavent, G., and R. E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: Geophysics, **64**, 508–515, doi: 10.1190/1.1444557.

Claerbout, J. F., 1971, Towards a unified theory of reflector mapping: Geophysics, **36**, 467–481, doi: 10.1190/1.1440185.

Claerbout, J. F., and S. Doherty, 1972, Downward continuation of moveout-corrected seismogram: Geophysics, **37**, 741–768, doi: 10.1190/1.1440298.

Cole, S., and M. Karrenbach, 1992, Least-squares Kirchhoff migration: Stanford Exploration Project (SEP) Report.

Cova, R., and K. A. Innanen, 2013, Approximate- vs. full-Hessian in FWI: 1D analytical and numerical experiments: CREWES Research Report.

Davydenko, M., and D. J. Verschuur, 2017, Full-wavefield migration: Using surface and internal multiples in imaging: Geophysical Prospecting, **65**, 7–21, doi: 10.1111/1365-2478.12360.

Deng, F., and G. A. McMechan, 2007, True-amplitude prestack depth migration: Geophysics, **72**, no. 3, S155–S166, doi: 10.1190/1.2714334.

Duquet, B., and K. J. Marfurt, 1999, Filtering coherent noise during prestack depth migration: Geophysics, **64**, 1054–1066, doi: 10.1190/1.1444613.

Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, no. 6, WCA5–WCA17, doi: 10.1190/1.3223188.

Fletcher, R. P., D. Nichols, R. Bloor, and R. T. Coates, 2016, Least-squares migration — Data domain versus image domain using point spread functions: The Leading Edge, **35**, 157–162, doi: 10.1190/tle35020157.1.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351, doi: 10.1190/1.1440899.

Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024, doi: 10.1190/1.1778244.

Hou, J., and W. Symes, 2016, Approximate Gauss-Newton iteration for full-waveform inversion: 86th Annual International Meeting, SEG, Expanded Abstracts, 1163–1168, doi: 10.1190/segam2016-13878327.1.

Hou, J., and W. W. Symes, 2015, An approximate inverse to the extended Born modeling operator: Geophysics, **80**, no. 6, R331–R349, doi: 10.1190/geo2014-0592.1.

Hu, J., G. T. Schuster, and P. Valasek, 2001, Poststack migration deconvolution: Geophysics, **66**, 939–952, doi: 10.1190/1.1444984.

Huang, Y., G. Dutta, W. Dai, X. Wang, G. T. Schuster, and J. Yu, 2014, Making the most out of least-squares migration: The Leading Edge, **33**, 954–960, doi: 10.1190/tle33090954.1.

Jang, U., D.-J. Min, and C. Shin, 2009, Comparison of scaling methods for waveform inversion: Geophysical Prospecting, **57**, 49–59, doi: 10.1111/j.1365-2478.2008.00739.x.

Jones, I. F., 2014, Tutorial: Migration imaging conditions: First Break, **32**, 45–55, doi: 10.3997/1365-2397.2014017.

Jones, I. F., 2018, Velocities, imaging, and waveform inversion: EAGE.

Kiyashchenko, D., R.-E. Plessix, B. Kashtan, and V. Troyan, 2007, A modified imaging principle for true-amplitude wave-equation migration: Geophysical Journal International, **168**, 1093–1104, doi: 10.1111/j.1365-246X.2006.03187.x.

Lecomte, I., 2008, Resolution and illumination analyses in PSDM: A ray-based approach: The Leading Edge, **27**, 650–663, doi: 10.1190/1.2919584.

Lines, L., and S. Treitel, 1984, A review of least-squares inversion and its application to geophysical problems: Geophysical Prospecting, **32**, 159–186, doi: 10.1111/j.1365-2478.1984.tb00726.x.

Lu, S., F. Liu, N. Chemingui, A. Valenciano, and A. Long, 2018, Least-squares fullwavefield migration: The Leading Edge, **37**, 46–51, doi: 10.1190/tle37010046.1.

Matharu, G., and M. Sacchi, 2019, A subsampled truncated-Newton method for multiparameter full-waveform inversion: Geophysics, **84**, no. 3, R333–R340, doi: 10.1190/geo2018-0624.1.

McMaken, H., 1986, On the convergence of the Bremmer series for the Helmholtz equation in 2-D: Wave Motion, **8**, 277–283, doi: 10.1016/S0165-2125(86)80050-6.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420, doi: 10.1111/j.1365-2478.1983.tb01060.x.

Mehta, K., M. Hugues, O. Hernandez, D. E. Bernholdt, and H. Calandra, 2017, One-way wave equation migration at scale on GPUs using directive based programming: IEEE International Parallel and Distributed Processing Symposium, 224–233.

Metivier, L., F. Bretaudeau, R. Brossier, S. Operto, and J. Virieux, 2014, Full waveform inversion and the truncated Newton method: Quantitative imaging of complex subsurface structures: Geophysical Prospecting, **62**, 1353–1375, doi: 10.1111/1365-2478.12136.

Mulder, W. A., and R. E. Plessix, 2004, A comparison between one-way and two-way wave-equation migration: Geophysics, **69**, 1491–1504, doi: 10.1190/1.1836822.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221, doi: 10.1190/1.1444517.

Oh, J.-W., and D.-J. Min, 2013, Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion: Geophysical Journal International, **194**, 322–347, doi: 10.1093/gji/ggt107.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054, doi: 10.1190/tle32091040.1.

Plessix, R. E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitudepreserving migration: Geophysical Journal International, **157**, 975–987, doi: 10.1111/j.1365-246X.2004.02282.x.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, doi: 10.1046/j.1365-246X.1998.00498.x.

Schuster, G. T., 2002, Reverse-time migration = generalized diffraction stack migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1280–1283, doi: 10.1190/1.1816888.

Schuster, G. T., 2017, Seismic inversion: SEG.

Shin, C., C. S. Jang, and D.-J. Min, 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592–606, doi: 10.1046/j.1365-2478.2001.00279.x.

Siahkoohi, A., D. J. Verschuur, and F. J. Herrmann, 2019, Surface-related multiple eliminations with deep learning: 89th Annual International Meeting, SEG, Expanded Abstracts, 4629–4634, doi: 10.1190/segam2019-3216723.1.

Slob, E., K. Wapenaar, F. Broggini, and R. Snieder, 2014, Seismic reflector imaging using internal multiples with Marchenko-type equations: Geophysics, **79**, no. 2, S63–S76, doi: 10.1190/geo2013-0095.1.

Slob, E., L. Zhang, and E. Verschuur, 2021, Marchenko multiple elimination and full-wavefield migration in a resonant pinch-out model: Geophysics, **86**, no. 5, WC1–WC9, doi: 10.1190/geo2020-0867.1.

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, 23–48, doi: 10.1190/1.1440826.

Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015, doi: 10.1111/j.1365-2478.1984.tb00751.x.

ten Kroode, F., 2012, A wave-equation-based Kirchhoff operator: Inverse Problems, **28**, 115013, doi: 10.1088/0266-5611/28/11/115013.

Thorbecke, J., L. Zhang, K. Wapenaar, and E. Slob, 2021, Implementation of the Marchenko multiple elimination algorithm: Geophysics, **86**, no. 2, F9–F23, doi: 10.1190/geo2020-0196.1.

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, doi: 10.1190/1.1778246.

Valenciano, A. A., and B. Biondi, 2003, 2-D deconvolution imaging condition for shot-profile migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1059–1062, doi: 10.1190/1.1817454.

Valenciano, A. A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, no. 4, A35–A38, doi: 10.1190/1.2213359.

van Groenestijn, G. J., and D. J. Verschuur, 2009, Estimating primaries by sparse inversion and application to near-offset data reconstruction: Geophysics, **74**, no. 3, A23–A28, doi: 10.1190/1.3111115.

Weglein, A. B., F. V. Araujo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse Problems, **19**, R27–R83, doi: 10.1088/0266-5611/19/6/R01.

Weglein, A. B., F. A. Gasparotto, P. M. Carvalho, and R. H. Stolt, 1997, An inverse-scattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, 1975–1989, doi: 10.1190/1.1444298.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385, doi: 10.1190/1.1893867.

Yang, J., J. Huang, Z. Li, H. Zhu, G. A. McMechan, and X. Luo, 2021, Approximating the Gauss-Newton Hessian using a space-wavenumber filter and its applications in least-squares seismic imaging: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–13, doi: 10.1109/TGRS.2021.3064852.

Ypma, F. H. C., and D. J. Verschuur, 2013, Estimating primaries by sparse inversion, a generalized approach: Geophysical Prospecting, **61**, 94–108, doi: 10.1111/j.1365-2478.2012.01095.x.

Yu, J., J. Hu, G. T. Schuster, and R. Estill, 2006, Prestack migration deconvolution: Geophysics, **71**, no. 2, S53–S62, doi: 10.1190/1.2187783.

Zhan, G., W. Dai, M. Zhou, Y. Luo, and G. T. Schuster, 2014, Generalized diffraction-stack migration and filtering of coherent noise: Geophysical Prospecting, **62**, 427–442, doi: 10.1111/1365-2478.12086.

Zhang, L., and E. Slob, 2020, A field data example of Marchenko multiple elimination: Geophysics, **85**, no. 2, S65–S70, doi: 10.1190/geo2019-0327.1.

Zhang, Y., S. Xu, N. Bleistein, and G. Zhang, 2007, True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations: Geophysics, **72**, no. 1, S49–S58, doi: 10.1190/1.2399371.

Biographies and photographs of the authors are not available.