



Delft University of Technology

One-way reflection waveform inversion with depth-dependent gradient pre-conditioning

Abolhassani, Siamak; Hoogerbrugge, Leo; Verschuur, Dirk Jacob

DOI

[10.1093/gji/ggae397](https://doi.org/10.1093/gji/ggae397)

Publication date

2025

Document Version

Final published version

Published in

Geophysical Journal International

Citation (APA)

Abolhassani, S., Hoogerbrugge, L., & Verschuur, D. J. (2025). One-way reflection waveform inversion with depth-dependent gradient pre-conditioning. *Geophysical Journal International*, 240(1), 652-672.
<https://doi.org/10.1093/gji/ggae397>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

One-way reflection waveform inversion with depth-dependent gradient pre-conditioning

Siamak Abolhassani , Leo Hoogerbrugge and Dirk Jacob Verschuur

Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, 2628 C.J Delft, The Netherlands.

E-mail: s.abolhassani@tudelft.nl

Accepted 2024 October 25. Received 2024 October 25; in original form 2024 January 30

SUMMARY

Reflection waveform inversion (RWI) is a technique that uses pure reflection data to estimate subsurface background velocity, relying on evolving seismic images. Conventional RWI operates in a cyclic workflow, with two key components in each cycle—migration and reflection tomography. Conventional RWI may result in suboptimal background velocity estimation, partly due to limited or unresolved resolution within each component in each cycle. While gradient pre-conditioning with the reciprocal of Hessian information helps resolve this issue in both components of RWI, it becomes impractical for a large number of model parameters. One-way reflection waveform inversion (ORWI) is a reflection waveform inversion technique in which the forward modelling scheme operates in one direction (downward and then upward) via virtual parallel depth levels within the medium. Leveraging the ORWI framework, we decompose and reduce the linear Hessian operator (also known as the approximate Hessian or Gauss–Newton Hessian) into multiple smaller suboperators. In particular, the diagonal blocks of the monofrequency approximate Hessian operators, each corresponding to a single depth level within the medium, are extracted and inverted to pre-condition the corresponding monofrequency gradients in both the migration and reflection tomography components of ORWI. This depth-dependent gradient pre-conditioning transforms standard ORWI into a high-resolution, yet computationally feasible version aimed at addressing suboptimal velocity estimation, referred to as high-resolution ORWI. The effectiveness of the proposed approach is demonstrated through successful applications to synthetic data examples.

Key words: Acoustic properties; Seismic tomography; Waveform inversion; Wave propagation.

1 INTRODUCTION

Tarantola (1984) introduced the conventional full waveform inversion (FWI) in acoustic approximation, a well-known non-linear data-fitting technique to estimate high-resolution subsurface velocity models. Conventional FWI has become widely adopted for evaluating shallow subsurface targets. Indeed, incorporating refracted and diving waves, it offers a detailed subsurface model and has proven to be highly effective for this purpose (e.g. Vigh *et al.* 2011; Zou *et al.* 2014). However, when attempting to map the targets that lie beyond the reach of refracted and diving waves (“deep targets”), its performance may falter (e.g. Plessix *et al.* 2013; Brittan & Jones 2019). This limitation arises because, while the sampling of high-reflective model wavenumbers is nonlinearly dependent on the sampling of low-propagative model wavenumbers in FWI (Audebert & Ortiz-Rubio 2018), conventional FWI fails to actively sample low-propagative model wavenumbers within the deep targets because of weak transmission-after-reflection wave paths within the FWI gradient, commonly known as “rabbit-ear” wave paths. Scaled by reflection coefficients, the rabbit-ear wave paths exhibit weaker amplitudes compared to other wave paths within the FWI gradient, that is the transmission wave path and migration isochrone (Fig. 1). This failure to sample low-propagative model wavenumbers significantly limits the capture of high-reflective model wavenumbers by conventional FWI, highlighting the significance of model scale separation for reflection tomography (Mora 1989).

Over the years, several variants of migration-based velocity analysis (MVA), each built upon the principle of model scale separation based on Born modelling, have been developed. Among them, the ones incorporating wave-equation forward modelling have stood out due to the improved handling of wave propagation in complex media. Wave-equation MVA (WEMVA) algorithms (Sava & Biondi 2004) typically optimize an image-domain error function measured on the common image gathers (CIGs) either via a classical semblance or a differential

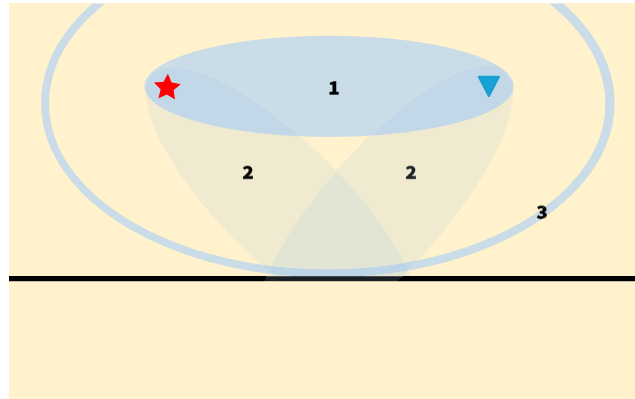


Figure 1. The wave paths constructing the FWI gradient: the transmission wave path, labelled as number 1, the pair of transmission-after-reflection wave paths (rabbit-ear wave paths), labelled as number 2, and the migration isochrone, labelled as number 3. As can be observed, the pair of transmission-after-reflection wave paths appears in a lighter shade, illustrating its much weaker amplitude compared to the other existing wave paths. This occurs as a consequence of scaling by the reflection coefficient of the reflector within the medium. Hence, it can be concluded that the FWI gradient is dominated by the transmission wave path and migration isochrone, while the pair of transmission-after-reflection wave paths remains relatively inactive.

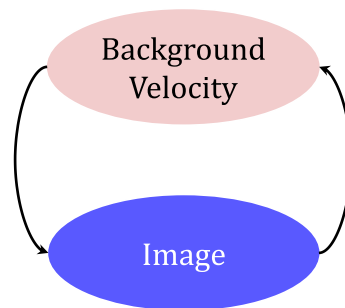


Figure 2. Conventional RWI cycle in which background velocity estimation and image reconstruction alternate.

semblance (DS) function, measuring the flatness (coherency property) in CIGs. While with the classical semblance function, the flatness measure of CIGs is given by the stack power of image amplitudes over surface offsets (Chauris & Noble 2001), the flatness measure of CIGs with the DS function is given by squaring the L2 norm of the derivative of image amplitudes with respect to surface offsets (Shen *et al.* 2003). WEMVA may also be formulated in the subsurface offset domain, where, at the correct velocity, the image is expected to be focused at zero offset. In this domain, measuring the focusing of subsurface-offset CIGs mirrors the use of the classical semblance and DS function with surface CIGs (Shen 2005; Shen & Symes 2008; Lameloise *et al.* 2015). Alternatively, a scattering angle can serve as the extension parameter, with flatness measured on angle-domain CIGs (Biondi & Symes 2004).

Expanding on the foundations of extended-domain WEMVA, Symes (2008) formulated a theory that laid the groundwork for integrating extended-domain WEMVA and FWI, leveraging their synergy. This theory was subsequently validated by Almomin & Biondi (2012). In recent initiatives to enhance MVA, Chauris & Cocher (2017) introduced the notion of inversion velocity analysis (IVA) by substituting standard migration with iterative migration to map cleaner CIGs, reducing migration artifacts and multiple crosstalk. Moreover, they adopted a pseudo-inverse of the Born modelling operator, rather than an adjoint, to reduce the migration iterations. In the same vein, Assis *et al.* (2022) introduced tomographic Hessian-based IVA.

Employing migrated images for velocity model building also triggered another trend of development known as reflection waveform inversion (RWI). RWI represents a wave-equation reflection tomography tool with a data-domain error function (Xu *et al.* 2012; Zhou *et al.* 2012; Berkhout 2012; Wang *et al.* 2013). Conventional RWI in acoustic approximation alternately solves a two-parameter minimization problem for the background velocity and “reflectivity” of the subsurface. Although RWI and the migration-based traveltimes waveform inversion (MBTT; Clément *et al.* 2001) share some similarities, it is important to note that the migration step in MBTT operates on reflectivities in time rather than depth—images in time are largely invariant to velocity, unlike images in depth, but this assumes a 1-D velocity model.

Unlike FWI, RWI is specifically tailored to sample deep targets tomographically harnessing exclusively reflection data through the pair of transmission-after-reflection wave paths. RWI involves estimating the background velocity of the subsurface in a cyclic process as displayed in Fig. 2, relying on an evolving stacked image that fits near-offset observed and modelled “phase information” within each cycle. Therefore, it is expected that RWI naturally suppresses any ambiguities initiated by the phase mismatch in the near-offset reflection data (Provenzano *et al.* 2023), particularly within the deep areas or the areas with challenging data acquisition conditions where FWI suffers severely. To improve the RWI tomogram, specifically within reach of refractions and diving waves, some studies have also investigated the

impact of integrating early-arrival FWI and RWI in order to steer the search direction toward recovering a full (horizontal and vertical) low local-model-wavenumber spectrum for the background velocity model (Zhou *et al.* 2015; Wu & Alkhalifah 2015). For a full review of the other aspects of RWI in recent years, the reader can refer to Yao *et al.* (2020).

Although RWI is built upon the notion of model scale separation, making this separation has been a struggle. To address the issue, several approaches have been put forth. Among others, we reference Alkhalifah (2014) for introducing the scattering-angle filtering solution, Wang *et al.* (2013) for adopting the wavefield decomposition solution, and Xu *et al.* (2012) for leveraging Born modelling as the solution. Additionally, Berkhout (2012, 2014) introduced the joint migration inversion (JMI) technique that makes use of a one-way forward modelling scheme called full-wavefield modelling (FWMod), parametrized based on reflectivity and acoustic background velocity, to fulfill the model scale separation by wavefield decomposition into upgoing and downgoing wavefields. FWMod models full (both primaries and multiples) downgoing and upgoing reflection wavefields within the JMI framework, in which zero-lag cross-correlations of same-direction source and receiver wavefields (either both upwards or both downwards) reconstruct the reflectivity model update (small-scale medium variations), and correlations of opposite-traveling source and receiver wavefields (one upgoing, one downgoing) estimate the background velocity model update (large-scale medium variations).

Adopting the concept of JMI for reflection waveform inversion and constraining FWMod to primary wavefield modelling (PWMod), assuming internal multiples are weak or resolved via pre-processing, this paper represents and investigates upon a one-way reflection waveform inversion technique, hereafter referred to as standard ORWI. The appropriate parameterization in PWMod leads to a natural scale separation—separation of migration isochrones and transmission-after-reflection wavepaths—in standard ORWI. This separation allows for the independent calculation of migration and tomography gradients within each ORWI cycle, thereby freeing tomograms from high-reflective model wavenumbers. Algorithm 1 shows the cyclic workflow of standard ORWI, which involves an alternating sequence between the image and background velocity reconstructions. Standard ORWI cyclically minimizes the error function

$$E = \frac{1}{2} \int d\omega \sum_s \Delta \mathbf{d}_{s,\omega}^- \Delta \mathbf{d}_{s,\omega}^{-\dagger}, \quad (1)$$

where $\Delta \mathbf{d}^-$ is the residual data vector (upgoing) between the observed and modelled reflection data recorded at the Earth's surface, the subscript s indicates a source location, ω shows an angular frequency component, and the superscript \dagger represents the complex conjugate transpose operation.

Algorithm 1 Cyclic workflow of ORWI

Provide a smooth initial background velocity model and set the initial reflectivity model to zero when cycle number $k = 1$.

while convergence not reached **do**

for $i \leftarrow 1$ to m **do** $\triangleright m$ is the total iteration number for updating reflectivity in cycle k .

UPDATE the reflectivity model using one-way wave-equation

 while keeping the background velocity model fixed (migration).

end for

for $j \leftarrow 1$ to n **do** $\triangleright n$ is the total iteration number for updating background velocity in cycle k .

UPDATE the background velocity model using one-way wave-equation

 while keeping the reflectivity model fixed (reflection tomography).

end for

$k = k + 1$ \triangleright The reflectivity model can be reset to zero at this point.

end while

Suboptimal background velocity estimation in RWI can be attributed, among others, to the use of seismic images with limited or unresolved resolution as well as suboptimal preservation of true amplitudes in each cycle (e.g. Hou & Symes 2015; Chauris & Cocher 2017; Gomes & Yang 2017). With the same line of reasoning, one can also link suboptimal background velocity estimation in RWI to the use of tomograms with limited or unresolved resolution. Pratt *et al.* (1998) showed that the Newton optimization method in waveform inversion estimates accurate velocity models with higher resolution than the gradient decent method since it uses the reciprocal of second-order derivatives of the error function (inverse Hessian) to pre-condition the gradient. However, the calculation demand of the Hessian matrix-operator for large-scale problems in seismology renders the Newton method impossibly expensive in computing terms. Several attempts, either in the data or image domain, have been made thus far to approximate an effective yet cost-effective pre-conditioner: migration deconvolution filters (Hu *et al.* 2001; Yu *et al.* 2006), matching filters (Guitton 2004; Aoki & Schuster 2009; Guitton 2017; Guo & Wang 2020;

Yang *et al.* 2021), point spread functions (Lecomte 2008; Fletcher *et al.* 2016; Yang *et al.* 2022), target-oriented solutions (Valenciano *et al.* 2006; Tang 2009), asymptotic pseudo inverses to the Born modelling operator (ten Kroode 2014; Hou & Symes 2015; Chauris & Cocher 2017), linear Hessian (also known as the Gauss–Newton Hessian or approximate Hessian) (Pratt *et al.* 1998), pseudo-Hessian (Choi *et al.* 2008; Jun *et al.* 2015), diagonal approximate Hessian (Chavent *et al.* 1999; Shin *et al.* 2001; Plessix & Mulder 2004), and approximating the application of the Hessian inverse operator on the search direction (Brossier *et al.* 2009; Asnaashari *et al.* 2013; Métivier *et al.* 2013; Assis & Schleicher 2021). Within the context of iterative least-squares one-way wave-equation migration (LS-WEM), while Lu *et al.* (2018) obtained an approximate Hessian by implicitly formulating the modelling operator and its adjoint, Abolhassani & Verschuur (2024) introduced a computationally affordable, depth-dependent approximate Hessian, referred to henceforth as pre-conditioned LS-WEM (PLS-WEM).

This paper aims to increase the resolution power of standard ORWI tomograms, expanding upon the findings of Abolhassani & Verschuur (2024) on depth-dependent pre-conditioning for image reconstruction. While employing depth-dependent pre-conditioners for high-resolution image reconstruction with standard ORWI, we concurrently develop the required mathematical groundwork to integrate the depth-dependent pre-conditioning concept for background velocity estimation. This two-fold effort facilitates more accurate velocity estimation with ORWI. We refer to it as high-resolution ORWI (HR-ORWI).

In the following sections, we will first review the theory of HR-ORWI. Then, we will present a brief overview of the depth-dependent gradient pre-conditioning concept for migration (PLS-WEM). Following that, we will mathematically demonstrate its relevance for tomography within the context of ORWI. Next, we will provide three numerical examples, where the first two compare standard ORWI and HR-ORWI, and the third validates that our cost-friendly high-resolution updates in both migration and tomography loops of HR-ORWI result in a reliable background velocity estimation for the Marmousi2 model. Lastly, we end with discussions and conclusions.

2 HR-ORWI THEORY

The theoretical aspects of HR-ORWI are presented here. We rely on Berkhout (1981, 1982) for the forward problem theory.

2.1 Forward problem

The two-way acoustic wave equation for a homogeneous medium in the space-frequency domain can be written as

$$\frac{\partial^2 u}{\partial z^2} = -\frac{\omega^2}{c^2}u - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}, \quad (2)$$

in which u represents the monochromatic pressure wavefield, and c is the velocity.

Replacing the differentiations in eq. (2) with spatial convolutions (Berkhout 1982) gives

$$\frac{\partial^2 u}{\partial z^2} = -\left(\frac{\omega^2}{c^2}\delta(x)\delta(y) + d_2(x) + d_2(y)\right) * u, \quad (3)$$

where $*$ denotes the spatial convolution, δ means the Dirac delta function, and the operators $d_2(x)$ and $d_2(y)$ represent spatial differential operators of order 2 with respect to x and y , respectively. When the initial condition $u(z_0)$ is available, eq. (3) reads

$$\left.\frac{\partial^2 u(x, y, z)}{\partial z^2}\right|_{z=z_0} = h_2(x, y) * u(x, y, z_0). \quad (4)$$

After splitting the total pressure wavefield into up- and downgoing wavefields ($u = u^+ + u^-$), eq. (4), following its transformation into the wavenumber–frequency domain within the $k_x - k_y$ plane, can be broken down into

$$\begin{aligned} \left.\frac{\partial \tilde{u}^+}{\partial z}\right|_{z=z_0} &= -i\tilde{h}_1^+ \tilde{u}^+(z_0), \\ \left.\frac{\partial \tilde{u}^-}{\partial z}\right|_{z=z_0} &= +i\tilde{h}_1^- \tilde{u}^-(z_0), \end{aligned} \quad (5)$$

where z is pointing downward, \tilde{u}^+ represents the monochromatic downgoing wavefield in the wavenumber domain, and \tilde{u}^- is the monochromatic upgoing wavefield in the wavenumber domain. The partial derivatives $\frac{\partial \tilde{u}^+}{\partial z}$ and $\frac{\partial \tilde{u}^-}{\partial z}$ evaluated at z_0 , are denoted as $\left.\frac{\partial \tilde{u}^+}{\partial z}\right|_{z_0}$ and $\left.\frac{\partial \tilde{u}^-}{\partial z}\right|_{z_0}$, respectively. Here, $i = \sqrt{-1}$, and $\tilde{h}_1^+ = \tilde{h}_1^- = \sqrt{\tilde{h}_2} = \sqrt{k^2 - k_x^2 - k_y^2}$, with $k = \frac{\omega}{c}$.

Inserting the derived partial derivative wavefields into the Taylor expansion for $\tilde{u}^\pm(z_0 \pm |\Delta z|)$ when $k^2 \geq k_x^2 + k_y^2$ yields

$$\begin{aligned} \tilde{u}^+(z_0 + |\Delta z|) &= e^{-i\sqrt{k^2 - k_x^2 - k_y^2}|\Delta z|} \tilde{u}^+(z_0), \\ \tilde{u}^-(z_0 - |\Delta z|) &= e^{-i\sqrt{k^2 - k_x^2 - k_y^2}|\Delta z|} \tilde{u}^-(z_0), \end{aligned} \quad (6)$$

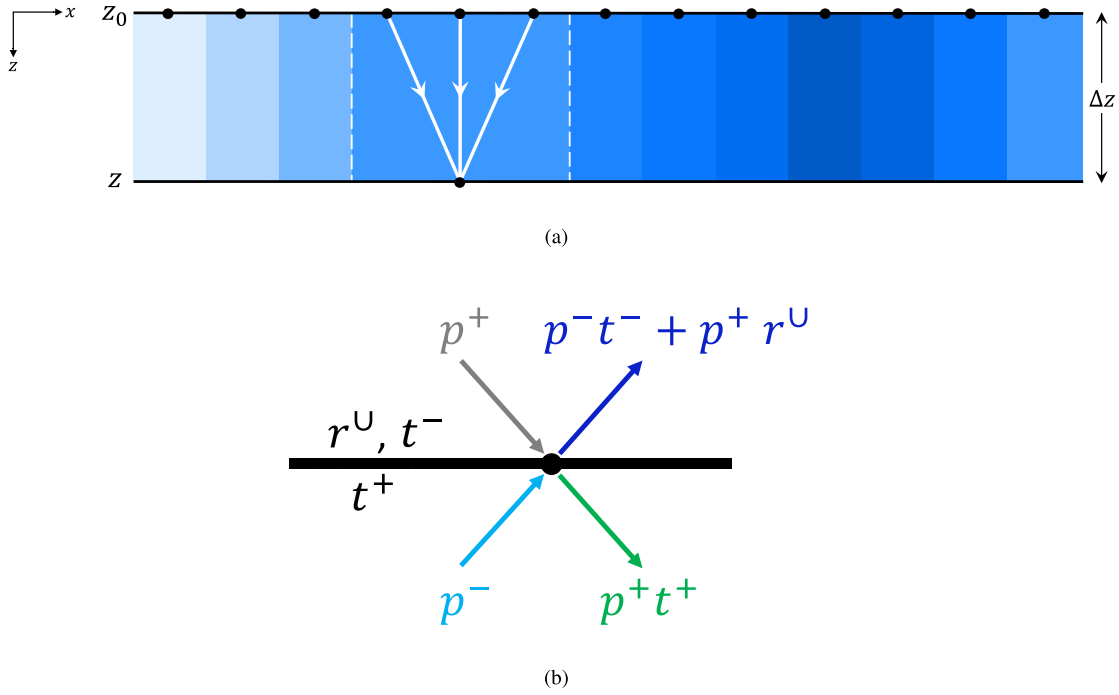


Figure 3. (a) Schematic representation of downward wavefield extrapolation between two virtual depth levels, considering lateral velocity variation. This graphic represents one single row of $\mathbf{W}_{z_0+|\Delta z|;z_0}^+$, which propagates the wavefields to a lateral position at z using a local average velocity across an operator size of three velocity cells. Note that the background velocity cells are defined between the virtual depth levels. (b) Schematic representation of how the incident wavefield interacts (reflection and transmission) at a particular virtual depth level in PWMod, where r^U means an upward reflectivity scalar, t^+ denotes a downward transmission scalar, and t^+ denotes an upward transmission scalar of a given grid point at a specific lateral location.

where Δz denotes the extrapolation distance, $\tilde{w}_{z_0+|\Delta z|;z_0}^+$ is the downward propagator from z_0 to $z_0 + \Delta z$, and $\tilde{w}_{z_0-|\Delta z|;z_0}^-$ is the upward propagator from z_0 to $z_0 - \Delta z$. Reverting to the space-frequency domain, eq. (6) can be expressed as follows

$$u^\pm(x, y, z_0 \pm |\Delta z|) = w_{z_0 \pm |\Delta z|;z_0}^\pm(x, y, \omega) * u^\pm(x, y, z_0). \quad (7)$$

When dealing with extrapolation distances with lateral velocity changes, a matrix formulation is adopted to represent eq. (7) as

$$\mathbf{u}_\omega^+(z_0 + |\Delta z|) = \mathbf{W}_{z_0+|\Delta z|;z_0}^+ \mathbf{u}_\omega^+(z_0), \quad (8)$$

$$\mathbf{u}_\omega^-(z_0 - |\Delta z|) = \mathbf{W}_{z_0-|\Delta z|;z_0}^- \mathbf{u}_\omega^-(z_0), \quad (9)$$

where \mathbf{W} (extrapolation/propagation matrix) is a space-variant convolutional matrix rather than a Toeplitz matrix. Fig. 3 schematically illustrates one row of $\mathbf{W}_{z_0+|\Delta z|;z_0}^+$ propagating the wavefields downwards from z_0 to a lateral position at z using a local average velocity across an operator size of three—selected for illustrative purposes here—velocity cells. Note that the velocity cells are defined between the virtual depth levels. For more details on the space-variant extrapolation matrix and operator size, see Berkhout (1982) and Thorbecke *et al.* (2004). One should note that the space-variant extrapolation matrix used for this research remains an approximation valid under the condition of smooth lateral velocity changes. Interested readers are referred to Sun & Verschuur (2020) and Li *et al.* (2024) for insights on an accurate extrapolation matrix for strong lateral velocity changes.

By discretizing the subsurface into N virtual depth levels separated by Δz (Δz is supposed to be sufficiently small to be considered vertically homogeneous), the fundamental equation governing the modelling of monochromatic primary reflection data recorded at the Earth's surface (z_0) with one monochromatic physical source positioned at z_0 is given by

$$\mathbf{u}^-(z_0) = \sum_{m=1}^N \mathbf{W}_{z_0;z_m}^- \mathbf{r}^U(z_m) \circ \mathbf{W}_{z_m;z_0}^+ \mathbf{s}^+(z_0), \quad (10)$$

with

$$\mathbf{W}_{z_m;z_0}^+ = \mathbf{W}_{z_m;z_{m-1}}^+ \mathbf{T}^+(z_{m-1}) \cdots \mathbf{T}^+(z_1) \mathbf{W}_{z_1;z_0}^+, \quad (11)$$

$$\mathbf{W}_{z_0;z_m}^- = \mathbf{W}_{z_0;z_1}^- \mathbf{T}^-(z_1) \cdots \mathbf{T}^-(z_{m-1}) \mathbf{W}_{z_{m-1};z_m}^-, \quad (12)$$

where $\mathbf{u}^-(z_0)$ is the modelled monochromatic reflection data recorded at the Earth's surface (z_0), $\mathbf{s}^+(z_0)$ is the physical monochromatic downgoing source at z_0 , \mathbf{r}^U is the angle-independent upward reflectivity vector-operator at a layer boundary, \mathbf{T}^\pm is the angle-independent

downward/upward transmission matrix-operator at a layer boundary and explicitly defined by $\mathbf{T}^\pm = \mathbf{I} \pm \text{diag}(\mathbf{r}^\cup)$, and \circ means the Hadamard product (Fig. 3b). Note that despite expressing the forward modelling equations in three dimensions (3-D), for the remainder of this paper, we shall consider only the 2-D case.

2.2 Optimization problem

HR-ORWI minimizes a two-parameter error function by alternating updates of the two parameters: angle-independent upward reflectivity and background velocity (see Algorithm 1). The error function through a linearized forward problem approximation is defined as a quadratic function of frequency-dependent model perturbation vectors

$$E = \frac{1}{2} \int d\omega S(\delta \mathbf{m}_\omega), \quad (13)$$

with

$$S(\delta \mathbf{m}_\omega) = \sum_{s=1}^{N_s} \left(\Delta \mathbf{d}_{s,\omega}^- - \frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}} \delta \mathbf{m}_\omega \right)^\dagger \left(\Delta \mathbf{d}_{s,\omega}^- - \frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}} \delta \mathbf{m}_\omega \right), \quad (14)$$

and

$$\delta \mathbf{m}_\omega = \begin{pmatrix} \delta \mathbf{r}_\omega^\cup \\ \delta \mathbf{c}_\omega \end{pmatrix}, \quad (15)$$

where $\delta \mathbf{m}_\omega$ is the monofrequency total model perturbation vector, $\delta \mathbf{r}_\omega^\cup$ represents the upward reflectivity perturbation vector (the total upward reflectivity perturbation vector $\delta \mathbf{r}^\cup$ is obtained by integrating over frequency), $\delta \mathbf{c}_\omega$ denotes the background velocity perturbation vector (the total background velocity perturbation vector $\delta \mathbf{c}$ is obtained by integrating over frequency), $\Delta \mathbf{d}_{s,\omega}^-$ represents the monochromatic residual data vector between the observed and modelled primary reflection data at z_0 for a given source location, $\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}}$ represents the monochromatic sensitivity matrix—the first-order partial derivative wavefield with respect to the model parameter change—for a given source location, s counts the source locations, and N_s shows the numbers of source locations.

Setting the derivative of eq. (13) with respect to $\delta \mathbf{m}_\omega$ to zero (gradient descent optimization method) gives

$$\frac{\partial E}{\partial \delta \mathbf{m}_\omega} = \frac{\partial S}{\partial \delta \mathbf{m}_\omega} = 0, \quad (16)$$

which leads to

$$\delta \mathbf{m}_\omega = \text{Re} \left\{ \left[\sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}} \right]^\dagger \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}} \right] \right]^{-1} \text{Re} \left\{ \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{m}} \right]^\dagger \Delta \mathbf{d}_{s,\omega}^- \right\} \right\}, \quad (17)$$

with

$$\mathbf{H}_\omega^a = \begin{bmatrix} \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{r}^\cup} \right]^\dagger \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{r}^\cup} \right] & \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{r}^\cup} \right]^\dagger \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}} \right] \\ \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}} \right]^\dagger \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{r}^\cup} \right] & \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}} \right]^\dagger \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}} \right] \end{bmatrix}, \quad (18)$$

and

$$\mathbf{g}_\omega = \begin{pmatrix} \mathbf{g}_\omega^r \\ \mathbf{g}_\omega^c \end{pmatrix}, \quad (19)$$

where \mathbf{H}_ω^a is the monofrequency approximate Hessian matrix for all shots (equivalent to the monofrequency Gauss–Newton Hessian matrix), \mathbf{g}_ω shows the monofrequency gradient vector, \mathbf{g}_ω^r represents the monofrequency gradient vector for reflectivity, and \mathbf{g}_ω^c denotes the monofrequency gradient vector for background velocity—note that with Gauss–Newton optimization in waveform inversion, the local convexity in inverse modeling, despite a nonlinear forward model, resembles dealing with a linear forward model in each iteration.

Following Jang *et al.* (2009) and Oh & Min (2013), the total model perturbation vector, after accounting for the negative sign indicating the descent direction, can be expressed as the linear combination of all the monofrequency model perturbation vectors

$$\delta \mathbf{m} = - \int d\omega \delta \mathbf{m}_\omega, \quad (20)$$

where $\delta \mathbf{m}$ represents the total model perturbation vector.

In the n^{th} cycle, the total model perturbation vector satisfies

$$\mathbf{m}^{(n+1)} = \mathbf{m}^{(n)} + \alpha^{(n)} \delta \mathbf{m}^{(n)}, \quad (21)$$

where α denotes the optimization step length and reads

$$\alpha^{(n)} = \frac{\Delta \mathbf{d}^{-\dagger}(\mathbf{m}^{(n)} + \delta \mathbf{m}^{(n)}) \Delta \mathbf{d}^{-}(\mathbf{m}^{(n)})}{\Delta \mathbf{d}^{-\dagger}(\mathbf{m}^{(n)} + \delta \mathbf{m}^{(n)}) \Delta \mathbf{d}^{-}(\mathbf{m}^{(n)} + \delta \mathbf{m}^{(n)})}, \quad (22)$$

resulting from minimizing the residual data with respect to the optimization step length.

To reduce the computational load of solving the multiparameter inverse problem, we first discard the off-diagonal blocks of the monofrequency approximate Hessian, which represent the coupling between parameter classes. This is achieved by setting the off-diagonal elements of $\mathbf{H}_{\omega}^{\mathbf{a}}$ to zero. As a result, eq. (18) transforms into

$$\mathbf{H}_{\omega}^{\mathbf{a}} = \begin{bmatrix} \mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}} & \\ & \mathbf{H}_{\omega}^{\mathbf{a},\mathbf{c}} \end{bmatrix}, \quad (23)$$

where $\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}}$ shows the monofrequency approximate Hessian matrix to pre-condition $\mathbf{g}_{\omega}^{\mathbf{r}}$, $\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{c}}$ shows the monofrequency approximate Hessian matrix to pre-condition $\mathbf{g}_{\omega}^{\mathbf{c}}$, and zeroes are omitted for brevity. With this, eq. (21) turns into a system of equations with two equations with reduced and simpler terms. We then, instead of solving the two equations simultaneously, opt to solve the equations “independently” (i.e. minimal interdependency assumption between reflectivity and background velocity to further ease the computational load). This process involves an alternating scheme where we first solve the first equation to update \mathbf{r}^{\cup} while keeping \mathbf{c} fixed (migration), and then solve the second equation to update \mathbf{c} while keeping \mathbf{r}^{\cup} fixed (reflection tomography). This alternating scheme continues until convergence is achieved. This enables HR-ORWI to cyclically update both \mathbf{r}^{\cup} and \mathbf{c} , with an alternating scheme offering partial mitigation of the neglected interdependency between \mathbf{r}^{\cup} and \mathbf{c} (see Algorithm 1).

In the following section, we present how to calculate $\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}}$ and $\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{c}}$ in a depth-marching regime. The gradient calculations are explained in detail in Sun *et al.* (2019).

2.3 Depth-dependent gradient pre-conditioning

2.3.1 Reflectivity

Since PWMod allows access to the acoustic wavefield at different virtual depth levels, \mathbf{g}_{ω} for the reflectivity class of parameters can be expressed as

$$\mathbf{g}_{\omega}^{\mathbf{r}} = \begin{pmatrix} \mathbf{g}_{\omega}^{\mathbf{r}}(z_0) \\ \vdots \\ \mathbf{g}_{\omega}^{\mathbf{r}}(z_N) \end{pmatrix} = \begin{pmatrix} \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^{-}(z_0)}{\partial \mathbf{r}^{\cup}(z_0)} \right]^{\dagger} \Delta \mathbf{d}_{s,\omega}^{-} \\ \vdots \\ \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^{-}(z_N)}{\partial \mathbf{r}^{\cup}(z_N)} \right]^{\dagger} \Delta \mathbf{d}_{s,\omega}^{-} \end{pmatrix}, \quad (24)$$

and

$$\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}} = \begin{bmatrix} \mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}}(z_0) & & \\ & \ddots & \\ & & \mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}}(z_N) \end{bmatrix}, \quad (25)$$

where $\mathbf{H}_{\omega}^{\mathbf{a},\mathbf{r}}$ shows a block-diagonal structure, and off-block-diagonal zeroes are here omitted for brevity.

Within the LS-WEM context, Abolhassani & Verschuur (2024) showed that the sensitivity matrix with respect to the upward reflectivity model at a given virtual depth level (z_{ℓ}) for a pair of source and frequency component reads

$$\frac{\partial \mathbf{u}_{s,\omega}^{-}(z_0)}{\partial \mathbf{r}^{\cup}(z_{\ell})} = \begin{bmatrix} \mathbf{W}_{z_0;z_{\ell}}^{-} \begin{pmatrix} u_{1s,\omega}^{+}(z_{\ell}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mathbf{W}_{z_0;z_{\ell}}^{-} \begin{pmatrix} 0 \\ u_{2s,\omega}^{+}(z_{\ell}) \\ \vdots \\ 0 \end{pmatrix} \\ \vdots \\ \mathbf{W}_{z_0;z_{\ell}}^{-} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ u_{n_{xs},\omega}^{+}(z_{\ell}) \end{pmatrix} \end{bmatrix}, \quad (26)$$

in which n_x denotes the number of gridpoints at z_m , and $u_{j,s,\omega}^+(z_\ell)$ is a complex number representing the downgoing modelled wavefield at the j th grid point of z_ℓ .

Integrating eqs (24), (25), and (26) into eq. (17) and then into eq. (20) yield the total angle-independent upward reflectivity perturbation vector $\delta \mathbf{r}^\cup$. Finally, eq. (21) is used to update the total angle-independent upward reflectivity vector \mathbf{r}^\cup in each cycle.

2.3.2 Background velocity

By similarity, we will here derive the sensitivity matrix for the background velocity model at a given virtual depth level (z_ℓ) for a pair of source and frequency component.

We rewrite the monochromatic primary reflection data recorded at the Earth's surface (eq. 10) into the superposition of two upgoing wavefields, one travelling from $[z_{\ell+1}, z_N]$ to z_0 and the other from $[z_1, z_\ell]$ to z_0 ,

$$\mathbf{u}_{s,\omega}^-(z_0) = \sum_{m=\ell+1}^N \underline{\mathbf{W}}_{z_0;z_m}^- \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}_{z_m;z_0}^+ \mathbf{s}_\omega^+(z_0) + \sum_{m=1}^{\ell} \underline{\mathbf{W}}_{z_0;z_m}^- \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}_{z_m;z_0}^+ \mathbf{s}_\omega^+(z_0). \quad (27)$$

The data sensitivity with respect to a single background velocity model parameter located at z_ℓ (i.e. a single column of the respective sensitivity matrix) is found by taking the derivative of eq. (27) with respect to $c_j(z_\ell)$

$$\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial c_j(z_\ell)} = \left\{ \sum_{m=\ell+1}^N \frac{\partial \underline{\mathbf{W}}_{z_0;z_m}^-}{\partial c_j(z_\ell)} \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}_{z_m;z_0}^+ \mathbf{s}_\omega^+(z_0) \right\} + \left\{ \sum_{m=1}^{\ell} \underline{\mathbf{W}}_{z_0;z_m}^- \mathbf{r}^\cup(z_m) \circ \frac{\partial \underline{\mathbf{W}}_{z_m;z_0}^+}{\partial c_j(z_\ell)} \mathbf{s}_\omega^+(z_0) \right\}. \quad (28)$$

After reintroducing $\underline{\mathbf{W}}_{z_0;z_m}^-$ and $\underline{\mathbf{W}}_{z_m;z_0}^+$ as

$$\underline{\mathbf{W}}_{z_0;z_m}^- = \underline{\mathbf{W}}_{z_0;z_\ell}^- \mathbf{T}^-(z_\ell) \underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^- \mathbf{T}^-(z_{\ell+1}) \underline{\mathbf{W}}_{z_{\ell+1};z_m}^-, \quad (29)$$

$$\underline{\mathbf{W}}_{z_m;z_0}^+ = \underline{\mathbf{W}}_{z_m;z_{\ell+1}}^+ \mathbf{T}^+(z_{\ell+1}) \underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+ \mathbf{T}^+(z_\ell) \underline{\mathbf{W}}_{z_\ell;z_0}^+, \quad (30)$$

and plugging them into eq. (28), we get

$$\mathbf{d}_a = \underline{\mathbf{W}}_{z_0;z_\ell}^- \mathbf{T}^-(z_\ell) \frac{\partial \underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^-}{\partial c_j(z_\ell)} \mathbf{T}^-(z_{\ell+1}) \sum_{m=\ell+2}^N \underline{\mathbf{W}}_{z_{\ell+1};z_m}^- \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}_{z_m;z_0}^+ \mathbf{s}_\omega^+(z_0) \quad (31)$$

$$+ \underline{\mathbf{W}}_{z_0;z_\ell}^- \mathbf{T}^-(z_\ell) \frac{\partial \underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^-}{\partial c_j(z_\ell)} \mathbf{r}^\cup(z_{\ell+1}) \circ \underline{\mathbf{W}}_{z_{\ell+1};z_0}^+ \mathbf{s}_\omega^+(z_0),$$

$$\mathbf{d}_b = \sum_{m=\ell+2}^N \underline{\mathbf{W}}_{z_0;z_m}^- \mathbf{r}^\cup(z_m) \circ \underline{\mathbf{W}}_{z_m;z_{\ell+1}}^+ \mathbf{T}^+(z_{\ell+1}) \frac{\partial \underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+}{\partial c_j(z_\ell)} \mathbf{T}^+(z_\ell) \underline{\mathbf{W}}_{z_\ell;z_0}^+ \mathbf{s}_\omega^+(z_0) \quad (32)$$

$$+ \underline{\mathbf{W}}_{z_0;z_{\ell+1}}^- \mathbf{r}^\cup(z_{\ell+1}) \circ \frac{\partial \underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+}{\partial c_j(z_\ell)} \mathbf{T}^+(z_\ell) \underline{\mathbf{W}}_{z_\ell;z_0}^+ \mathbf{s}_\omega^+(z_0),$$

in which given the assumption that $\underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^-$ and $\underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+$ are approximate to each other, the first partial derivatives of the extrapolation operator are defined as

$$\frac{\partial \underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^-}{\partial c_j(z_\ell)} = \frac{\partial \underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+}{\partial c_j(z_\ell)} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \frac{\partial \mathbf{w}_{j,\star}}{\partial c_j} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}_{n_x \times n_x}, \quad (33)$$

and $\mathbf{w}_{j,\star}$ means the j th row of $\underline{\mathbf{W}}_{z_\ell;z_{\ell+1}}^-$ or $\underline{\mathbf{W}}_{z_{\ell+1};z_\ell}^+$, with each $\mathbf{0}$ (zero in boldface) denoting a row vector of dimension $1 \times n_x$.

Fig. 4 provides a schematic representation of the physical interpretation of eqs (31) and (32) for a medium consisting of six virtual depth levels, including z_0 . This illustration focuses on the scenario where c_j is positioned “at the virtual depth level z_2 ” in notation (between z_2 and z_3 in practice).

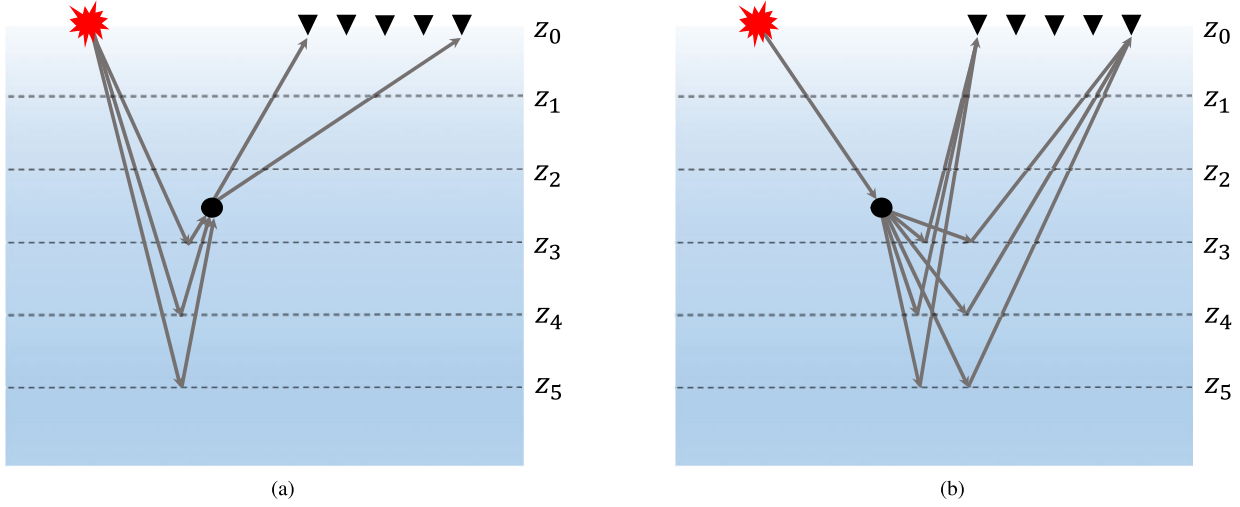


Figure 4. Schematic representation of \mathbf{d}_a and \mathbf{d}_b for one shot, eqs (31) and (32), in a medium with six virtual depth levels, including z_0 . (a) Receiver-side background velocity data sensitivity for one shot (i.e. \mathbf{d}_a). (b) Source-side background velocity data sensitivity for one shot (i.e. \mathbf{d}_b).

Now, after eq. (24), we can write,

$$\mathbf{g}_\omega^c = \begin{pmatrix} \mathbf{g}_\omega^c(z_0) \\ \vdots \\ \mathbf{g}_\omega^c(z_N) \end{pmatrix} = \begin{pmatrix} \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}(z_0)} \right]^\dagger \Delta \mathbf{d}_{s,\omega}^- \\ \vdots \\ \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_N)}{\partial \mathbf{c}(z_N)} \right]^\dagger \Delta \mathbf{d}_{s,\omega}^- \end{pmatrix}, \quad (34)$$

and

$$\mathbf{H}_\omega^{a,c} = \begin{bmatrix} \mathbf{H}_\omega^{a,c}(z_0) & & \\ & \ddots & \\ & & \mathbf{H}_\omega^{a,c}(z_N) \end{bmatrix}, \quad (35)$$

where $\mathbf{H}_\omega^{a,c}$ again shows a block-diagonal structure, and off-block-diagonal zeroes are here omitted for brevity.

Incorporating eqs (34), (35), and (28) into eq. (17) and subsequently into eq. (20) give the total background velocity perturbation vector $\delta \mathbf{c}$. Afterward, eq. (21) is employed to update the total background velocity vector \mathbf{c} in each cycle.

According to eqs (25) and (35), the depth-marching nature of HR-ORWI framework facilitates constructing monofrequency, depth-dependent approximate Hessian operators. Each monofrequency depth-dependent approximate Hessian operator is calculated for model parameters located on a single virtual depth level (with n_x model parameters at each virtual depth level). This calculation involves both autocorrelation and cross-correlation between the partial derivative wavefields associated with a specific depth level, while disregarding any cross-correlation between the partial derivative wavefields of that depth level and those of other virtual depth levels. This choice, facilitated by the frequency-domain, depth-marching nature of one-way propagators, reduces the entire monofrequency approximate Hessian, with a dimension of $n_m \times n_m$ ($n_m = N \times n_x$, representing the total number of model parameters within the entire medium), into multiple suboperators, each tailored to a single virtual depth level (i.e. N operators with a reduced dimension of $n_x \times n_x$). Reducing the entire monofrequency approximate Hessian into N smaller suboperators lowers the memory and processing demands for constructing, storing, and inverting in large-scale problems, while still retaining essential information for pre-conditioning at each depth level. The inverse of $\mathbf{H}_\omega^{a,r}(z_\ell)$ and $\mathbf{H}_\omega^{a,c}(z_\ell)$ accounts for geometric spreading and the spatial correlations of neighbouring model parameters at z_ℓ , while also performing frequency-wise source deconvolution. Compared to the monofrequency Gauss–Newton Hessian (Pratt *et al.* 1998), this approach approximates a block-diagonal representation of the monofrequency Gauss–Newton Hessian by assembling all monofrequency, depth-dependent approximate Hessian matrices into a single large matrix representing the entire medium (Abolhassani & Verschuur 2024).

2.4 HR-ORWI versus standard ORWI

Standard ORWI leverages the same forward problem theory as HR-ORWI and similarly updates reflectivity and background velocity by minimizing the errors in primary reflection waveforms for each class of parameters alternately, meaning that both standard ORWI and HR-ORWI follow the same inversion workflow, i.e. Algorithm 1.

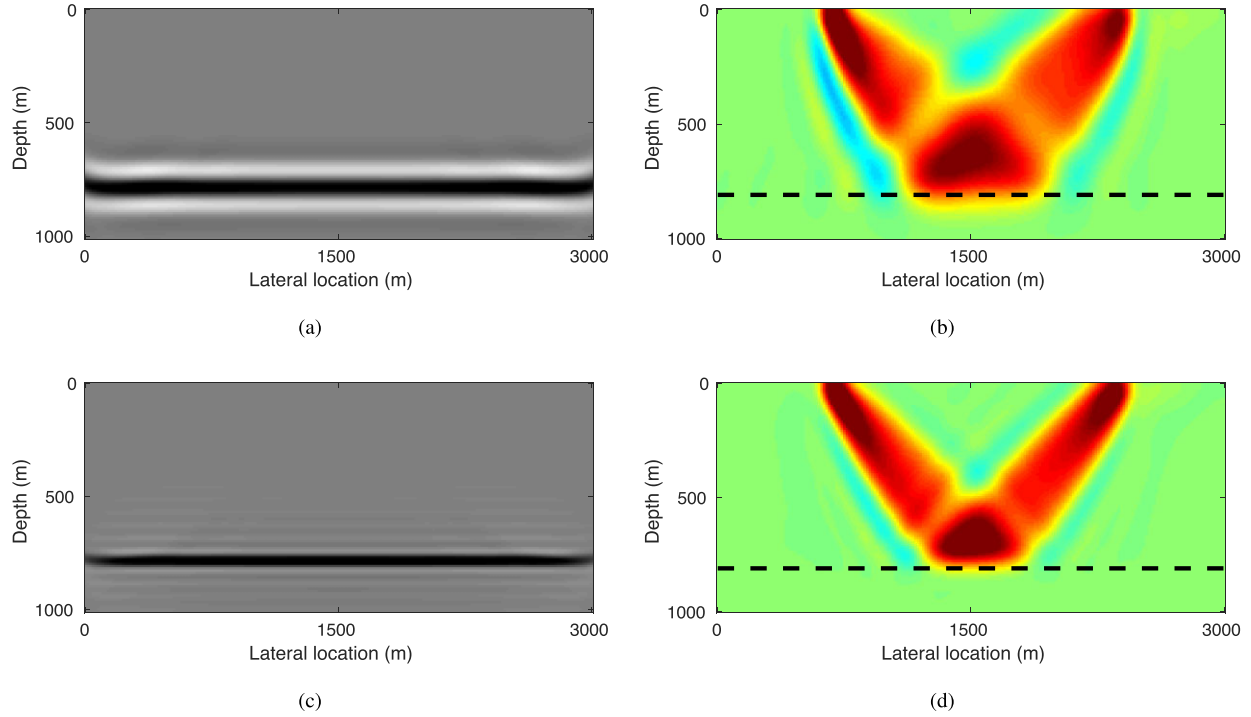


Figure 5. Stacked reflectivity perturbation and one-source–receiver-pair background velocity perturbation in a two-layer model after one cycle ($1 \times$ LS migration and $1 \times$ LS tomography). (a, b) Standard ORWI: (a) Stacked reflectivity perturbation and (b) One-source–receiver-pair background velocity perturbation. (c, d) HR-ORWI: (c) Stacked reflectivity perturbation and (d) One-source–receiver-pair background velocity perturbation.

Unlike HR-ORWI, standard ORWI minimizes eq. (13) as a quadratic function of $\delta \mathbf{m}$ rather than $\delta \mathbf{m}_\omega$. Setting the derivative of eq. (13) with respect to $\delta \mathbf{m}$ to zero (gradient descent optimization method) gives

$$\delta \mathbf{r}^U = -\text{Re} \left\{ \int d\omega \mathbf{H}_\omega^{a,r}(\mathbf{x}, \mathbf{x}) \right\}^{-1} \text{Re} \left\{ \int d\omega \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{r}^U} \right]^\dagger \Delta \mathbf{d}_{s,\omega}^- \right\}, \quad (36)$$

$$\delta \mathbf{c} = -\text{Re} \left\{ \int d\omega \mathbf{H}_\omega^{a,c}(\mathbf{x}, \mathbf{x}) \right\}^{-1} \text{Re} \left\{ \int d\omega \sum_{s=1}^{N_s} \left[\frac{\partial \mathbf{u}_{s,\omega}^-(z_0)}{\partial \mathbf{c}} \right]^\dagger \Delta \mathbf{d}_{s,\omega}^- \right\}, \quad (37)$$

where $\mathbf{H}_\omega^{a,r}(\mathbf{x}, \mathbf{x})$ indicates that only the diagonal elements of the monofrequency approximate Hessian possess values (see Staal 2015). This type of pre-conditioning only accounts for geometric spreading. Lastly, eq. (21) is employed in order to update the total angle-independent upward reflectivity vector \mathbf{r}^U and the total background velocity vector \mathbf{c} in each cycle.

We now use a two-layer model to evaluate the fidelity of the reflectivity and background velocity gradients in HR-ORWI versus Standard ORWI. The upper layer has a homogeneous background velocity of 3500 m s^{-1} , while the bottom layer has a homogeneous background velocity of 3000 m s^{-1} . The interface is at a depth of 800 m . 41 shots are triggered every 75 m on the surface. A 10 Hz Ricker wavelet is used as the source function with a fixed-spread acquisition geometry (the maximum available offset is 3000 m). We fix the migration offset—data offsets used for seismic migration—at 400 m (near-offset imaging) and use an initial background velocity model of 3400 m s^{-1} (i.e. ≈ 3 per cent error). Fig. 5 displays the stacked reflectivity perturbation ($\delta \mathbf{r}^U$) and one-source–receiver-pair background velocity perturbation ($\delta \mathbf{c}$), estimated after one cycle ($1 \times$ LS migration and $1 \times$ LS tomography) of standard ORWI and HR-ORWI.

Comparing Figs 5(a) and (b) (standard ORWI) with Figs 5(c) and (d) (HR-ORWI) reveals that HR-ORWI provides significantly better resolution in both the image and tomogram than standard ORWI. In Figs 5(a) and (b) (standard ORWI), the signature of the source function is visible in both the image and tomogram. The pair of transmission-after-reflection wave paths in Fig. 5(b) (standard ORWI) are unfocused and band-limited, as manifested by the scattering of reflected energies over a larger area compared to Fig. 5(d) (HR-ORWI), which makes information extraction from the data more challenging. In contrast, Figs 5(c) and (d) (HR-ORWI) show superior resolution in both the image and tomogram. This improvement is attributed to the source signature deconvolution effect embedded in the proposed depth-dependent approximate Hessian operators. With better-focused transmission-after-reflection wave paths, HR-ORWI is better positioned to handle the complex interference of the wave paths in challenging geological settings, supporting more reliable search directions than standard ORWI.

With acquisition configurations identical to those used in the experiment associated with Fig. 5, we use the same two-layer model to compare the stacked background velocity perturbations estimated through HR-ORWI and standard ORWI. In the experiment, we use an initial background velocity model of 3000 m s^{-1} (i.e. ≈ 15 per cent error). Additionally, we adopt the offset-selection strategy presented in Appendix A. That is, we fix the migration offset at 800 m (maximum effective migration offset: MEMO), use the mid-to-far-offset residual data ($|\text{offsets}| > 500 \text{ m}$) for tomography, and exclude the cycle-skipped far-offset data for tomography. Fig. 6 compares the respective

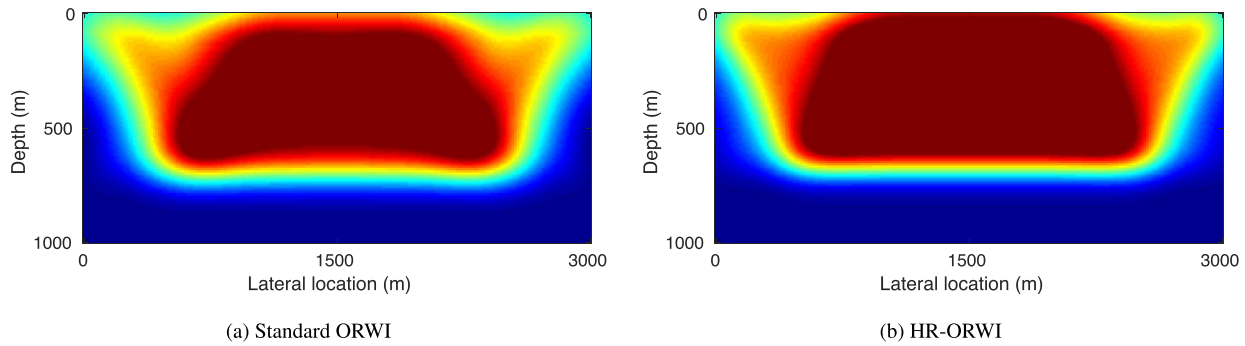


Figure 6. Background velocity perturbation corresponding to the two-layer model, estimated after one cycle ($1 \times$ LS migration and $1 \times$ LS tomography) for (a) standard ORWI and (b) HR-ORWI.

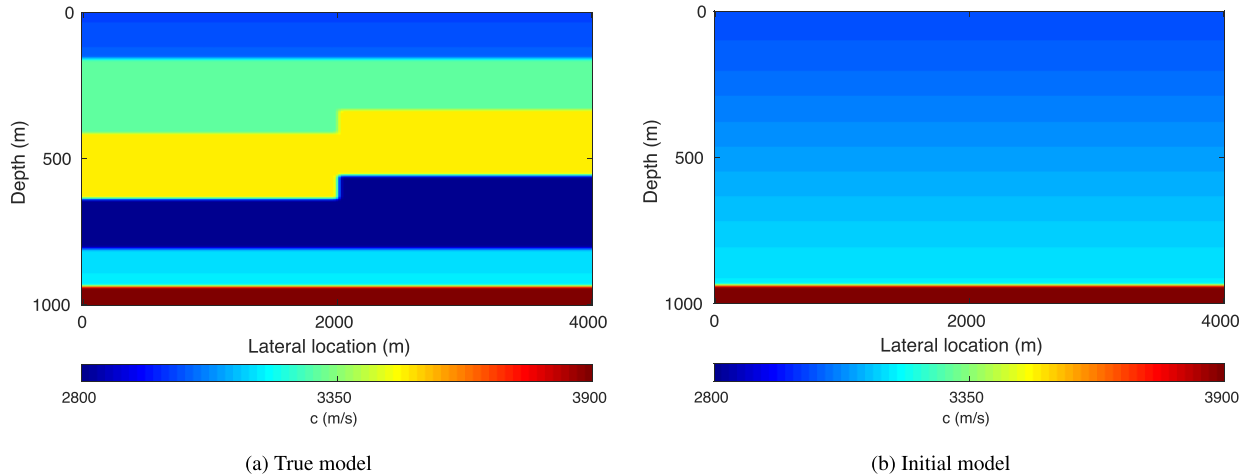


Figure 7. Fault model example. (a) True background velocity model. (b) Initial background velocity model.

background velocity perturbations (δc) after one cycle ($1 \times$ LS migration and $1 \times$ LS tomography) of HR-ORWI and standard ORWI. Upon comparison, it is observed that the superposition of tomographic wavepaths estimated through HR-ORWI exhibits a higher level of geometric consistency and homogeneity with the true layer. This observation highlights the accuracy of HR-ORWI in capturing subsurface features, resulting in a more faithful representation of the truth.

3 NUMERICAL EXAMPLES

First, we validate the effectiveness of HR-ORWI compared to standard ORWI using two synthetic examples. Second, we evaluate the performance of HR-ORWI with synthetic data on the marine-environment Marmousi model (Martin *et al.* 2006). In all the examples, we adopt the offset-selection strategy presented in Appendix A.

3.1 Fault model

In the fault model example (Fig. 7a), the observed data is generated via PWMod (only primary reflections), employing a 10 Hz Ricker wavelet. The fault model is discretized with 201 gridpoints in the horizontal direction (20 m interval) and 201 gridpoints in the vertical direction (5 m interval). 41 shots with 100 m spacing are used, and each shot is recorded by 201 receivers with a fixed-spread acquisition geometry at the surface of the Earth. The maximum available offset in the data is limited to 3000 m. The data is recorded for a duration of 1.6 s. The initial background velocity model is a 1-D linearly increasing gradient model (Fig. 7b), and the initial reflectivity model is zero. We fix the migration offset at 1000 m, use the mid-to-far-offset residual data ($|\text{offsets}| > 500$ m) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle.

The 0–30 Hz frequency band data is inverted at once, that is the inversion process does not involve a multiscaling strategy (Bunks *et al.* 1995). Each cycle of ORWI includes $1 \times$ LS migration and $1 \times$ LS tomography. While Fig. 8 provides a comparison of the estimated background velocity perturbations (δc) via standard ORWI and HR-ORWI after one cycle, Fig. 9 compares the estimated tomograms using standard ORWI and HR-ORWI after 25 and 50 cycles.

Comparing Figs 8(a) and (b) reveals a clear optimization direction with HR-ORWI shortly after the first cycle, where the fault line and balanced-amplitude background velocity perturbations are visible from shallow to deep. Emphasizing the reliability of HR-ORWI, this comparison also highlights the potential confusion that might be introduced with standard ORWI.

By comparing Figs 9(a) and (c) to Figs 9(b) and (d), respectively, it becomes evident that the estimated tomogram via HR-ORWI after 25 cycles closely aligns with, or even excel, the estimated tomogram via standard ORWI after 50 cycles. Further comparison also confirms the

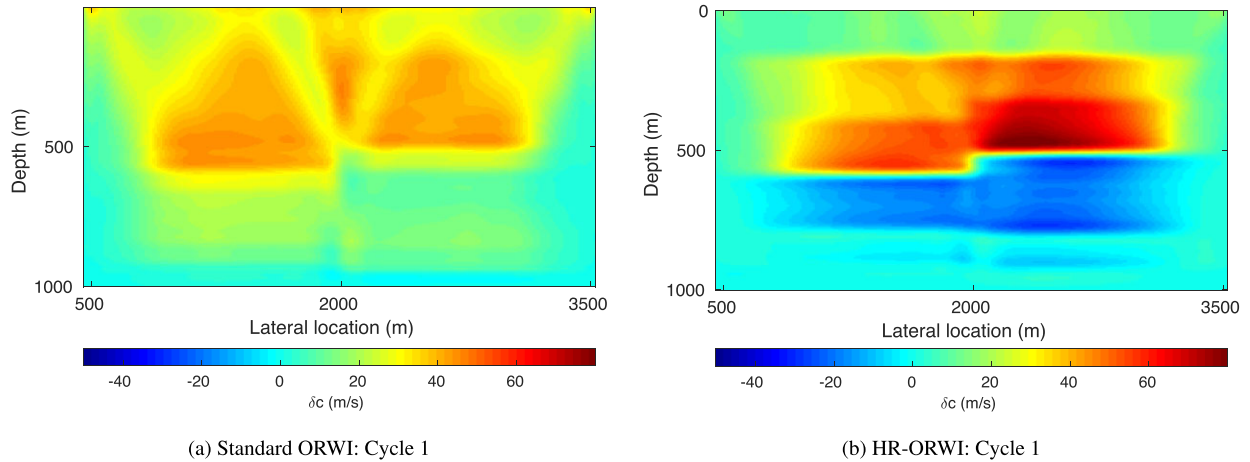


Figure 8. Background velocity perturbations associated with the fault model example estimated after one cycle ($1 \times$ LS migration and $1 \times$ LS tomography) of ORWI. (a) Background velocity perturbation estimated via standard ORWI. (b) Background velocity perturbation estimated via HR-ORWI.

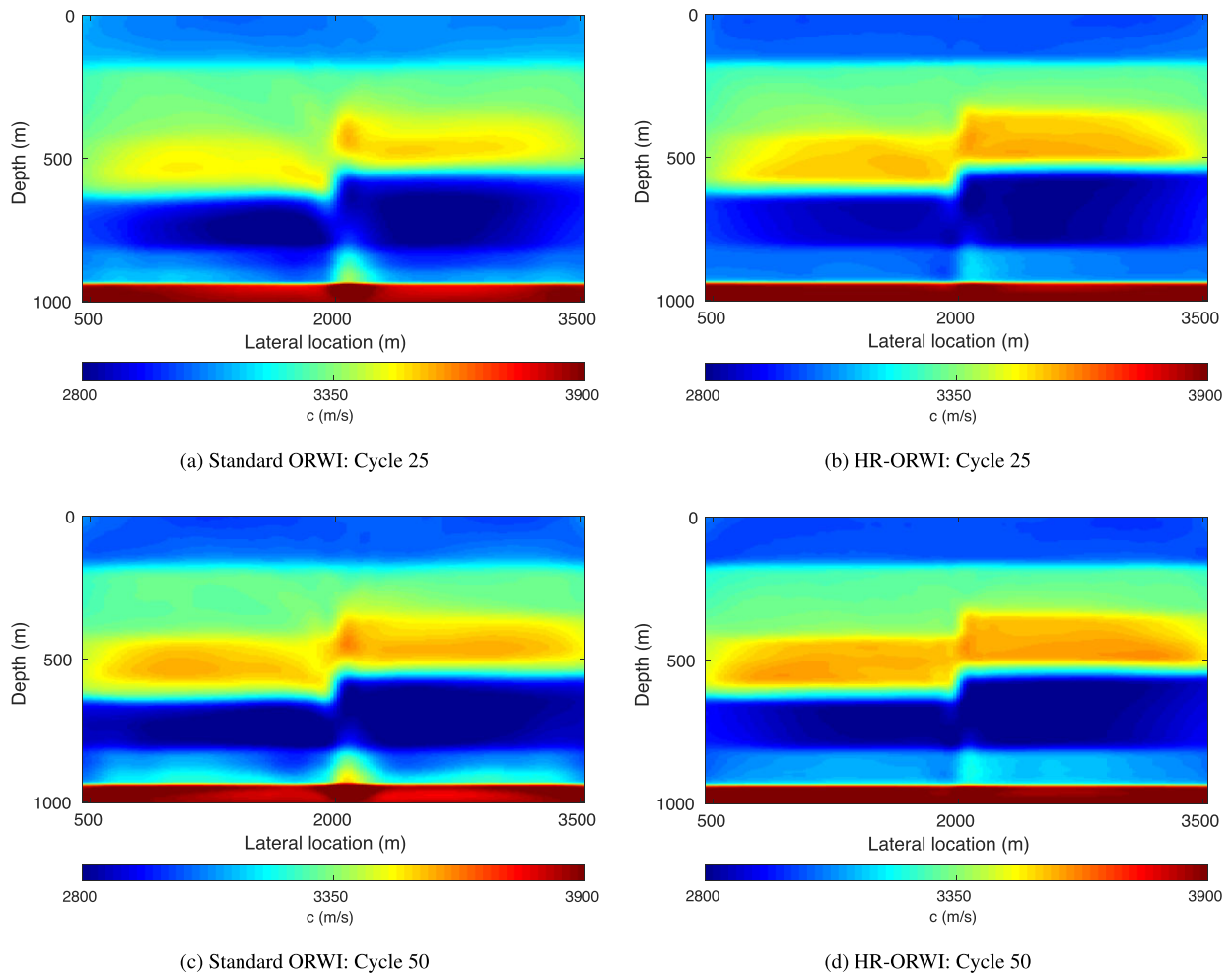


Figure 9. Estimated tomograms associated with the fault model example. (a) and (c) Estimated tomograms after 25 and 50 cycles (each cycle involves $1 \times$ LS migration and $1 \times$ LS tomography) of standard ORWI. (b) and (d) Estimated tomograms after 25 and 50 cycles (each cycle involves $1 \times$ LS migration and $1 \times$ LS tomography) of HR-ORWI.

existence of persistent and non-physical artifacts in tomograms estimated via standard ORWI from the first to the fiftieth cycle. The accuracy of the tomogram estimated via HR-ORWI after 50 cycles (Fig. 9d) is self-evident.

Fig. 10 (convergence history) again validates that HR-ORWI exhibits faster convergence and is able to achieve a better data fit in fewer cycles compared to standard ORWI.

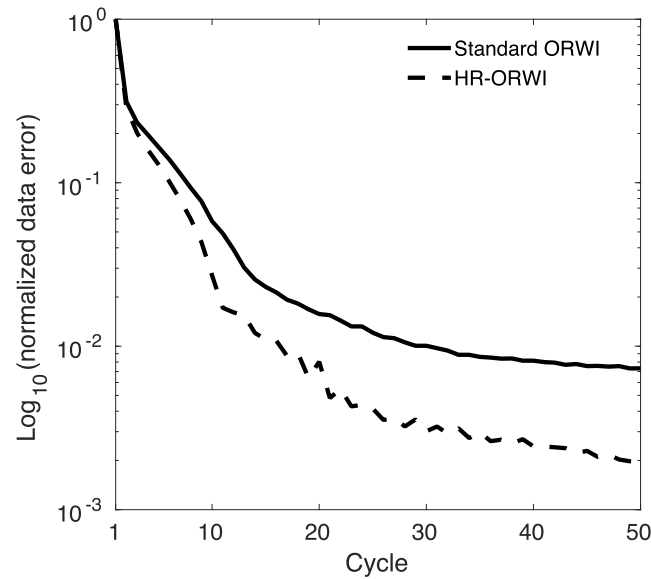


Figure 10. Convergence history (data error curve) associated with the fault model example.

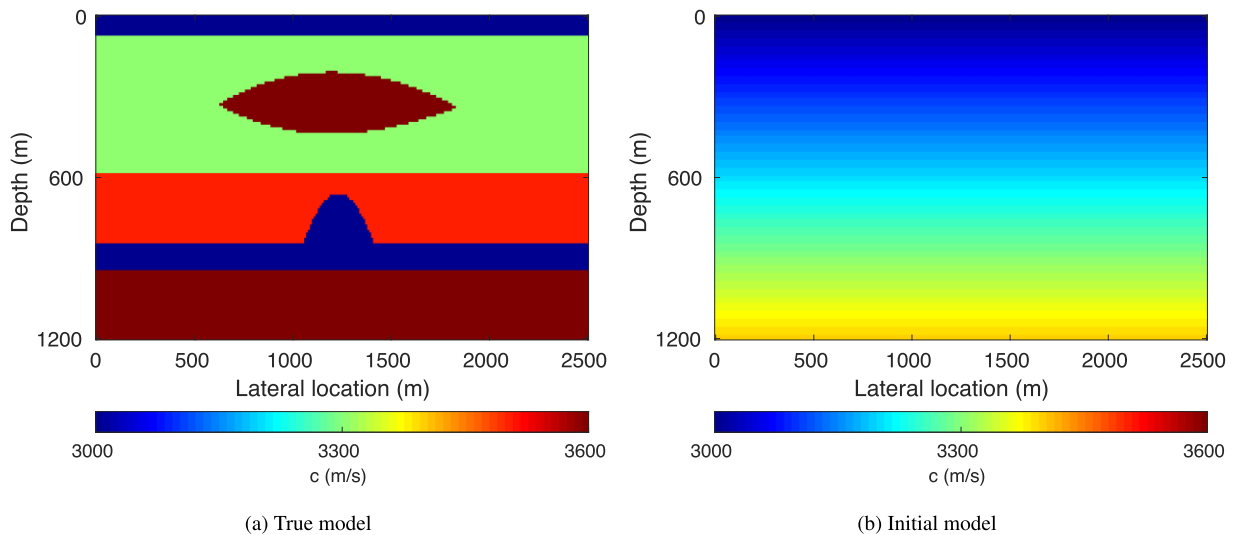


Figure 11. Reservoir model example. (a) True background velocity model. (b) Initial background velocity model.

3.2 Reservoir model

In the reservoir model example, a buried low velocity reservoir lies beneath a lens-shaped geological formation with high velocity (Fig. 11a), discretized into a grid of 251 points horizontally and 121 points vertically, both spaced at 10 m intervals. The observed data set is acquired with a fixed-spread acquisition geometry, including 26 shots with 100 m spacing and 251 receivers, both distributed at the surface of the Earth. The data set contains 1.4 s records, generated with a 15 Hz Ricker wavelet as the source function using PWMod, including only primary reflections. The maximum available offset in the data set is limited to 2000 m. The initial background velocity model is a 1-D linearly increasing gradient model, illustrated in Fig. 11(b), and the initial reflectivity model is zero. We fix the migration offset at 1000 m (MEMO), use the mid-to-far-offset residual data ($|\text{offsets}| > 500$ m) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle. The 0–25 Hz frequency band data is inverted at once (no multiscaling), and each cycle of ORWI includes $1 \times$ LS migration and $1 \times$ LS tomography. Fig. 12 compares the estimated tomograms using standard ORWI and HR-ORWI after 25 and 50 cycles.

Comparing Fig. 12(a) with Fig. 12(b) reveals that standard ORWI, after 25 cycles, still struggles to distinguish between the lens body and the underlying flat layer, whereas HR-ORWI reliably resolves this. Furthermore, while standard ORWI completely misestimates the background velocity of the reservoir body, HR-ORWI shows an accurate retrieval.

Comparing Fig. 12(c) with Fig. 12(b) shows that the tomogram estimated with HR-ORWI after 25 cycles not only aligns with the tomogram estimated with standard ORWI after 50 cycles but also surpasses in accuracy, especially at the reservoir level; this is observed while both share nearly identical data errors according to Fig. 13. The accuracy of the tomogram estimated via HR-ORWI after 50 cycles (Fig. 12d) is self-evident.

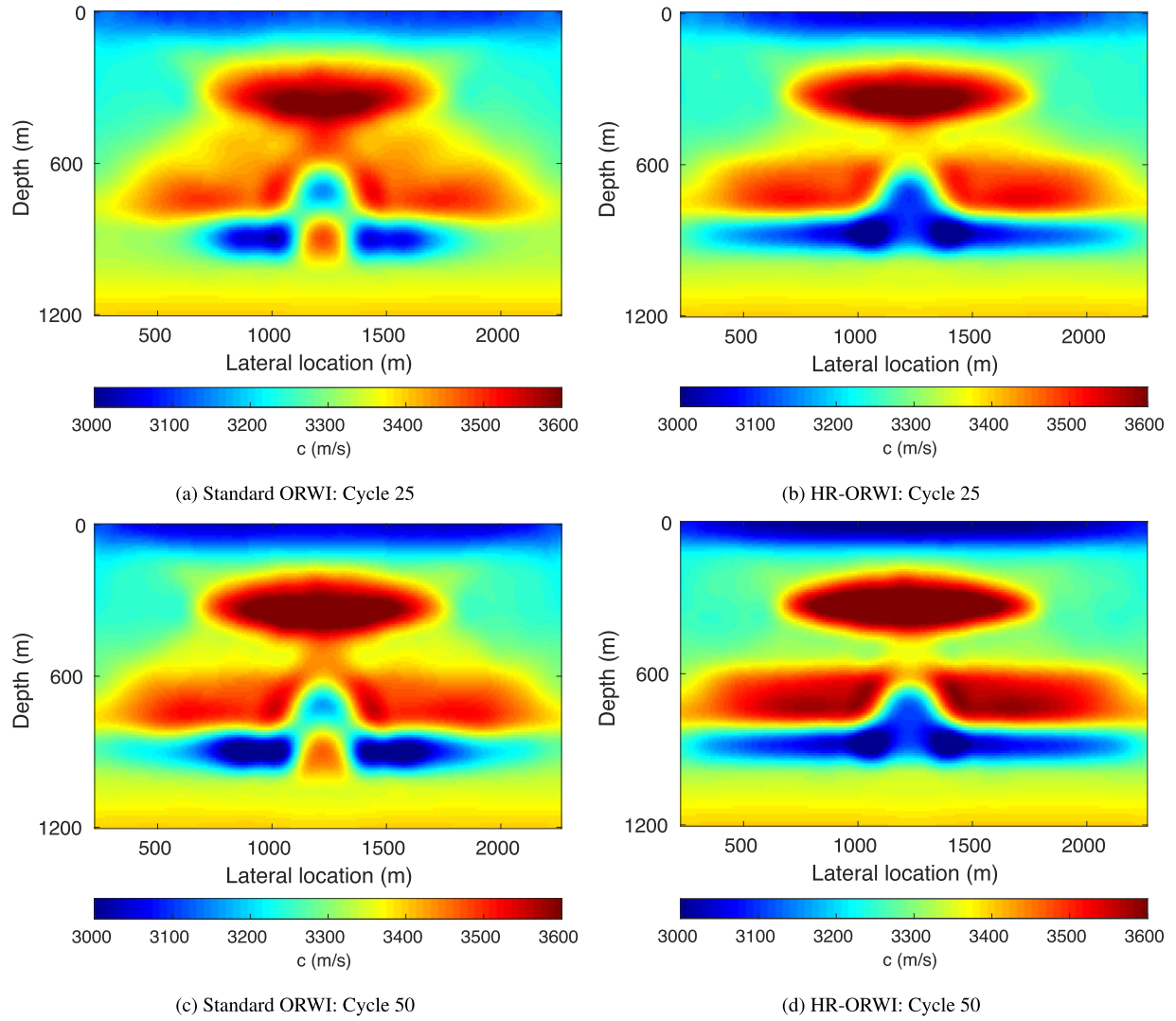


Figure 12. Estimated tomograms associated with the reservoir model example. (a) and (c) Estimated tomograms after 25 and 50 cycles (each cycle involves $1 \times$ LS migration and $1 \times$ LS tomography) of standard ORWI. (b) and (d) Estimated tomograms after 25 and 50 cycles (each cycle involves $1 \times$ LS migration and $1 \times$ LS tomography) of HR-ORWI.

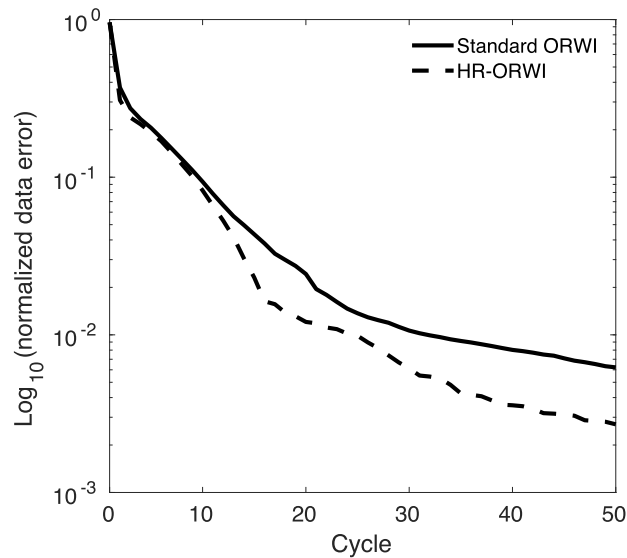


Figure 13. Convergence history (data error curve) associated with the reservoir model example.

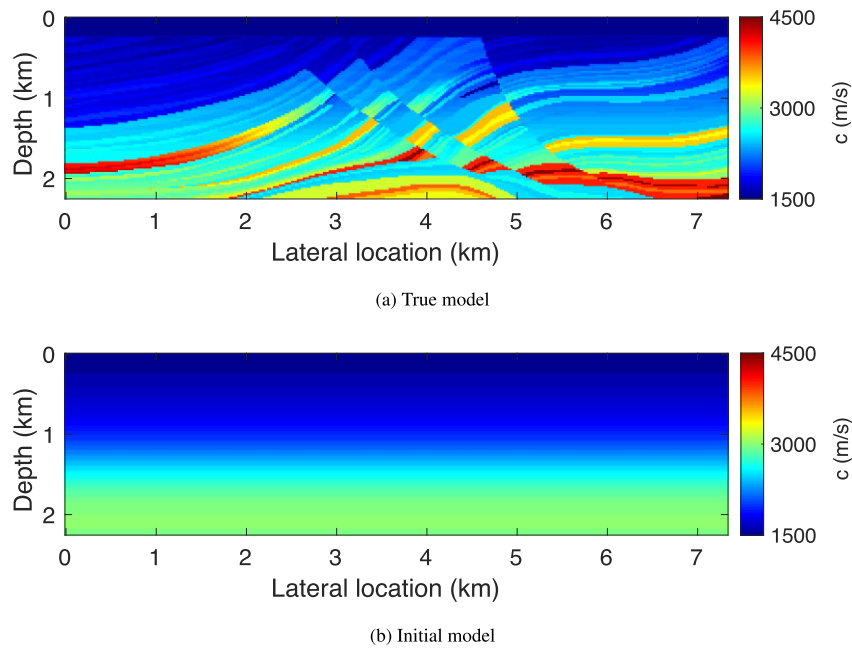


Figure 14. Marmousi2 model example. (a) True background velocity model. (b) Initial background velocity model.

Fig. 13 (convergence history) again verifies that HR-ORWI comes with faster convergence and can reach a better data fit in fewer cycles compared to standard ORWI.

3.3 Marmousi2 model

The Marmousi2 model example (Fig. 14) contains 334×103 gridpoints with intervals of 22 m. The observed data is generated via PWMod (only primary reflections) with a 10 Hz Ricker wavelet as the source function. With a shooting interval of nearly 200 m, 41 shots are triggered at the surface of the Earth, and each shot is recorded by 334 receivers planted on the surface of the Earth (fixed-spread acquisition). The data is recorded for 4.096 s, and the maximum available offset is limited to 4 km. To build the initial background velocity model, the true background velocity model is smoothed out by applying a square Gaussian kernel with a standard deviation of 1100 m. Subsequently, a single vertical profile is chosen to construct the initial 1D model with a water layer on top, as shown in Fig. 14(b). The initial reflectivity model is set to zero. We fix the migration offset at 2 km (MEMO), use the mid-to-far-offset residual data ($|\text{offsets}| > 0.5$ km) for tomography, and exclude the cycle-skipped far-offset data for tomography in each cycle. As we advance through cycles and update the background velocity, a growing number of offsets contribute to tomography as they are no longer cycle-skipped. This process continues until all offsets up to 4 km have been taken into account.

To invert the data within the 0–18 Hz frequency band, a multiscaling strategy is employed, involving three frequency bands: 0–5, 0–11, and 0–18 Hz. For both 0–50 and 0–11 Hz, the inversion process involves 20 cycles, each with 1 least-squares LS migration iteration and 1 LS tomography iteration. The image does not reset to zero after each cycle. For the 0–18 Hz band, the inversion process involves a total of 60 cycles, each with $1 \times$ least-squares LS migration iteration and $1 \times$ LS tomography iteration, and no image reset after each cycle. Note that the reflectivity model is reset back to zero after each frequency scale. The estimated background velocity perturbations (cycles 1 and 100) and tomogram (cycle 100) through HR-ORWI are displayed in Figs 15 and 16, respectively.

The estimated background velocity perturbation after the first cycle (Fig. 15a) illustrates that the pre-conditioned tomographic gradient reliably points in the right direction even from the outset. The main reflections have been correctly interpreted. The tomographic update focuses primarily on reflected events within the depth range of 0–1.5 km. Fig. 15(b) shows the estimated background velocity perturbation after the 100th cycle. It is evident that the tomographic gradient of HR-ORWI successfully samples the low-to-intermediate model wavenumbers of the true background velocity properly in 100 cycles from shallow to deep, though with a depth-dependent decreasing resolution (Williamson 1991).

The estimated tomogram after 100 cycles is shown in Fig. 16(a). In addition, a smoothed version of the true background velocity model with the aid of a square Gaussian kernel with a standard deviation of 75 m and a water layer on top is generated and shown in Fig. 16(a). A comparison of the two models presented in Fig. 16 indicates that the estimated tomogram has a resolution approximately similar to that of the smoothed true model.

To check the quality of the estimated tomogram, different vertical profiles located at the lateral locations of 3000, 3750, and 4600 m are extracted and plotted out in Fig. 17. It can be seen that the estimated background velocity in all extracted lateral locations is smooth but significantly close to the true velocity.

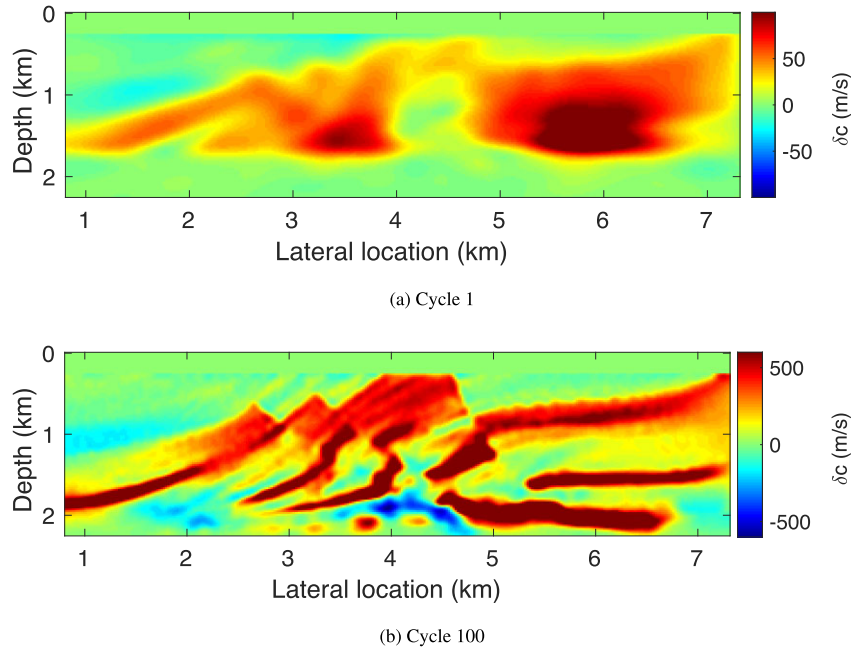


Figure 15. Background velocity perturbations associated with the Marmousi2 model example. (a) Background velocity perturbation estimated after the first cycle of HR-ORWI. (b) Background velocity perturbation estimated after the 100th cycle of HR-ORWI.

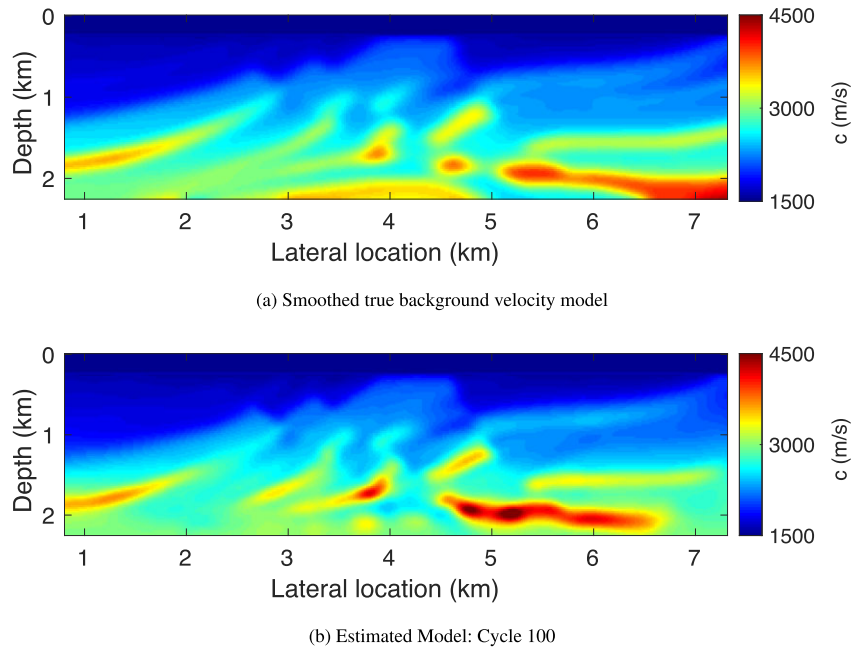


Figure 16. (a) Smoothed true background velocity model of Marmousi2 generated via a square Gaussian kernel with a standard deviation of 75 m and a water layer on top. (b) Estimated tomogram associated with the Marmousi2 model example after 100 cycles of HR-ORWI.

To assess the extent of uplift in the background velocity from cycle 1 to cycle 100, various images using different background velocity models, including true, initial, smoothed and estimated, are presented in Fig. 18. Upon comparing the images mapped with the initial and estimated background velocity models (Figs 18b and d), a significant kinematic uplift can be observed in Fig. 18(d), attributed to the enhanced focus of the reflectors. After comparing the images mapped with the true and estimated background velocity models (Figs 18a and d), a significant kinematic uplift is again confirmed in Fig. 18(d) based on the reasonably accurate positioning of the reflectors. Comparing the images mapped with the smoothed and estimated background velocity models (Figs 18c and d) also suggests that a “maximum” kinematic uplift within the achievable resolution of RWI has occurred in Fig. 18(d).

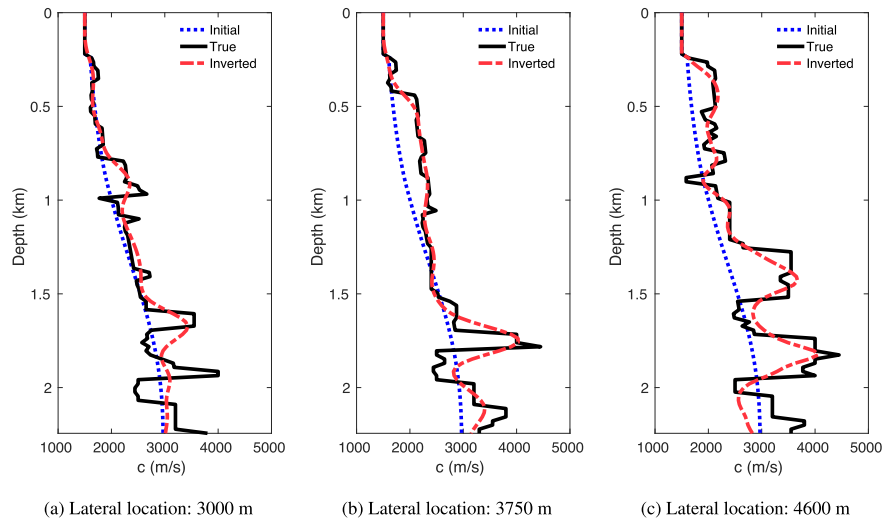


Figure 17. Vertical background velocity profiles associated with the Marmousi2 model example. A vertical profile of the estimated tomogram at the lateral location of (a) 3000 m, (b) 3750 m, and (c) 4600 m.

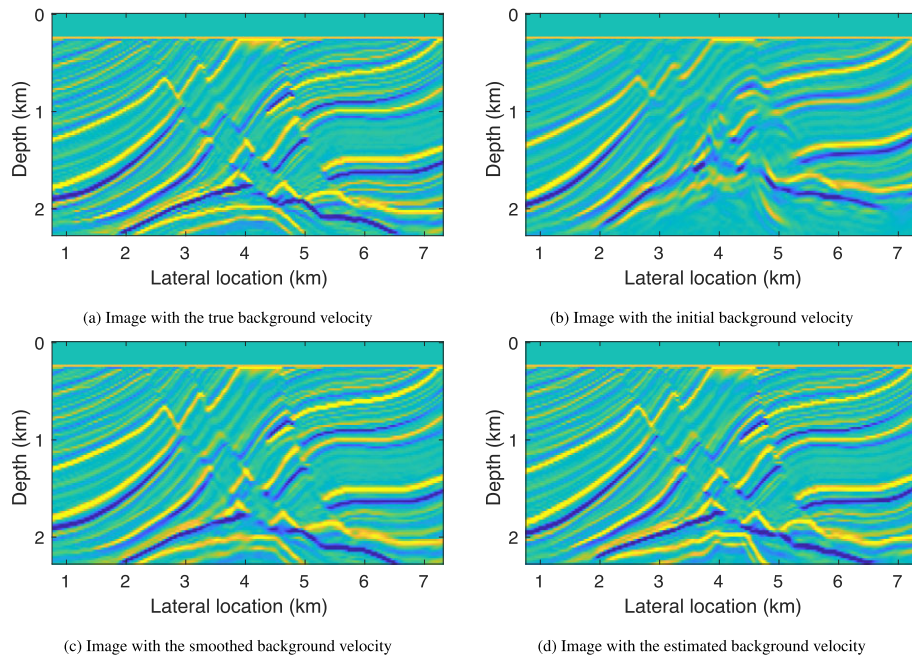


Figure 18. Images using different background velocity models associated with the Marmousi2 model example. The mapped images correspond to the (a) true, (b) initial, (c) smoothed and (d) estimated background velocity models.

4 DISCUSSION

The one-way (up and down) wave-equation extrapolation technique has persisted in the seismic imaging industry despite its weaknesses (for details on weaknesses, see Mulder & Plessix 2004), as it offers computational efficiency when contrasted with direct solutions to the Helmholtz equation. In this research, we demonstrated the potential use of one-way wave-equation forward modelling, that is PWMod, in efficient gradient pre-conditioning to enhance suboptimal images and tomograms resulting from RWI. In terms of constructing, inverting or storing, depth-dependent approximate Hessians computed based on PWMod are more feasible than doing so for pre-conditioners computed at once for all model parameters within the medium. Despite being feasible, it remains time-intensive compared to standard ORWI (i.e. HR-ORWI is six to nine times more expensive but requires half the iterations of standard ORWI, effectively making it three to five times more expensive overall), given the current implementation and the capabilities of the current computing resources. However, there are ways to reduce the time intensity. Among others, source subsampling in the computation of \mathbf{H}^a (Matharu & Sacchi 2019) sounds promising. According to our in-house inquiries, HR-ORWI with source subsampling in constructing \mathbf{H}_ω^a —specifically, one source out of three—could retrieve high-resolution tomograms consistent with this study. Even though each iteration would still take two to three times longer, the retrieval is reached in half the iterations of standard ORWI, making it almost equally expensive overall.

Selecting one technology over another often involves evaluating the cost-versus-quality trade-off. While cost-effectiveness must remain a key priority, compromising accuracy due to budget constraints in exploration seismology can raise significant concerns. In the reservoir model example presented in this paper, we clearly observed how choosing a budget-conscious technology (i.e. standard ORWI) can lead to incorrect subsurface illumination, resulting in misinterpretations of subsurface properties and potentially far-reaching consequences. Using cheaper technology can indeed be a double-edged sword; while it saves money upfront, the outcome could be irreversible. This portrays the complexity of decision-making dynamics in exploration seismology.

Although we did not use any model-based regularization in this study, integrating approaches like structure-oriented regularization guided by the reflectivity model may further enhance the fidelity of tomographic updates in HR-ORWI (Masaya & Verschuur 2018; Yao *et al.* 2019; Provenzano *et al.* 2023).

Real data applications require careful consideration of several additional factors. The negative potential impact of source wavelet estimation errors and the absence of low-frequency data on the stability of HR-ORWI should be carefully explored. Furthermore, addressing amplitude versus offset (AVO) effects in the ORWI process is crucial. Qu & Verschuur (2021) recommend minimizing AVO effects before applying ORWI. In this study, through inverse-crime numerical examples, we assumed reduced AVO effects in the data, facilitating a thorough examination of ORWI convergence. Challenges may also arise in accommodating anisotropy effects and integrating the Q-effect into ORWI, as highlighted in other studies (Alshuhail & Verschuur 2019; Safari & Verschuur 2023).

5 CONCLUSIONS

Limited or unresolved resolution in both migration and reflection tomography components of RWI can lead to suboptimal velocity estimation. To address this within the context of ORWI, we introduced a pre-conditioning strategy that enhanced both the migration and tomography loops, resulting in high-resolution updates for both reflectivity and background velocity. We named this approach HR-ORWI, where diagonal blocks of the monofrequency Gauss–Newton Hessian operators, each corresponding to a specific virtual depth level, are extracted and inverted to pre-condition their corresponding monofrequency gradients. This approach performs a partial deconvolution of the gradient while keeping memory and computational requirements manageable. Through three synthetic examples, we demonstrated that HR-ORWI generates superior background velocity estimations compared to standard ORWI.

ACKNOWLEDGMENTS

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

DATA AVAILABILITY

The code of this research is confidential. All data generated in this research can be shared upon a reasonable request directed to the corresponding author.

REFERENCES

- Abolhassani, S. & Verschuur, D.J., 2024. Efficient preconditioned least-squares wave-equation migration, *Geophysics*, **89**(3), S275–S288.
- Abolhassani, S. & Verschuur, E., 2022. Towards a more robust joint migration inversion, in *83rd EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.
- Alkhalifah, T., 2014. Scattering-angle based filtering of the waveform inversion gradients, *Geophys. J. Int.*, **200**(1), 363–373.
- Almomin, A. & Biondi, B., 2012. Tomographic full waveform inversion: practical and computationally feasible approach, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1–5.
- Alshuhail, A.A. & Verschuur, D.J., 2019. Robust estimation of vertical symmetry axis models via joint migration inversion: Including multiples in anisotropic parameter estimation, *Geophysics*, **84**(1), C57–C74.
- Aoki, N. & Schuster, G., 2009. Fast least-squares migration with a deblurring filter, *Geophysics*, **74**(6), WCA83–WCA93.
- Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P. & Virieux, J., 2013. Regularized seismic full waveform inversion with prior model information, *Geophysics*, **78**(2), R25–R36.
- Assis, C.A. & Schleicher, J., 2021. Introduction of the Hessian in joint migration inversion and improved recovery of structural information using image-based regularization, *Geophysics*, **86**, R777–R793.
- Assis, C.A.M., Chauris, H., Audebert, F. & Williamson, P., 2022. Tomographic hessian-based inversion velocity analysis: a first evaluation, in *83rd EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.
- Audebert, F. & Ortiz-Rubio, D., 2018. Separation of Scales in FWI and RFWI. filling the gap? or exploiting it?, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Berkhout, A., 1981. Wave field extrapolation techniques in seismic migration, a tutorial, *Geophysics*, **46**(12), 1638–1656.
- Berkhout, A.J., 1982. *Seismic Migration, Imaging of Acoustic Energy by Wave Field Extrapolation: A. Theoretical Aspects*, Elsevier, 1–351.
- Berkhout, A.J., 2012. Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging, *Geophysics*, **77**(2), S43–S50.
- Berkhout, A.J., 2014. Review paper: An outlook on the future seismic imaging, part III: Joint migration inversion, *Geophys. Prospect.*, **62**(5), 950–971.
- Biondi, B. & Symes, W.W., 2004. Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging, *Geophysics*, **69**(5), 1283–1298.
- Brittan, J. & Jones, I., 2019. FWI evolution—from a monolith to a toolkit, *Leading Edge*, **38**(3), 179–184.

- Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex on-shore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**(6), WCC105–WCC118.
- Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.
- Chauris, H. & Cocher, E., 2017. From migration to inversion velocity analysis, *Geophysics*, **82**(3), S207–S223.
- Chauris, H. & Noble, M., 2001. Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets, *Geophys. J. Int.*, **144**(1), 14–26.
- Chavent, G. & Plessix, R.E., 1999. An optimal true-amplitude least-squares prestack depth-migration operator, *Geophysics*, **64**(2), 508–515.
- Choi, Y., Min, D.-J. & Shin, C., 2008. Frequency-domain elastic full waveform inversion using the new pseudo-hessian matrix: Experience of elastic Marmousi-2 synthetic data, *Bull. seism. Soc. Am.*, **98**(5), 2402–2415.
- Clément, F., Chavent, G. & Gómez, S., 2001. Migration-based travel-time waveform inversion of 2-D simple structures: A synthetic example, *Geophysics*, **66**(3), 845–860.
- Fletcher, R.P., Nichols, D., Bloor, R. & Coates, R.T., 2016. Least-squares migration—Data domain versus image domain using point spread functions, *Leading Edge*, **35**(2), 157–162.
- Gomes, A. & Yang, Z., 2017. Improving reflection FWI reflectivity using LSRTM in the curvelet domain, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Guitton, A., 2004. Amplitude and kinematic corrections of migrated images for nonunitary imaging operators, *Geophysics*, **69**(4), 1017–1024.
- Guitton, A., 2017. Fast 3D least-squares RTM by preconditioning with nonstationary matching filters, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Guo, S. & Wang, H., 2020. Image domain least-squares migration with a hessian matrix estimated by non-stationary matching filters, *J. Geophys. Eng.*, **17**(1), 148–159.
- Hou, J. & Symes, W.W., 2015. An approximate inverse to the extended born modeling operator, *Geophysics*, **80**(6), R331–R349.
- Hu, J., Schuster, G.T. & Valasek, P., 2001. Poststack migration deconvolution, *Geophysics*, **66**(3), 939–952.
- Jang, U., Min, D.-J. & Shin, C., 2009. Comparison of scaling methods for waveform inversion, *Geophys. Prospect.*, **57**(1), 49–59.
- Jun, H., Park, E. & Shin, C., 2015. Weighted pseudo-hessian for frequency-domain elastic full waveform inversion, *J. Appl. Geophys.*, **123**, 1–17.
- Lameloise, C.-A., Chauris, H. & Noble, M., 2015. Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis, *Geophys. Prospect.*, **63**(2), 391–404.
- Lecomte, I., 2008. A ray-based approach: Resolution and illumination analyses in psdm, *Leading Edge*, **27**(5), 650–663.
- Li, A., Verschuur, D.J., Liu, X. & Abolhassani, S., 2024. An accurate propagator for heterogeneous media in full wavefield migration, *IEEE Trans. Geosci. Remote Sens.*, **62**, 1–10.
- Liang, H., Zhang, H. & Liu, H., 2022. Analysis of the impact of demigration on traveltimes-based reflection full-waveform inversion, in *SEG/AAPG Second International Meeting for Applied Geoscience & Energy, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1002–1006.
- Lu, S., Liu, F., Chemingui, N., Valenciano, A. & Long, A., 2018. Least-squares full-wavefield migration, *Leading Edge*, **37**(1), 46–51.
- Martin, G.S., Wiley, R. & Marfurt, K.J., 2006. Marmousi2: An elastic upgrade for Marmousi, *Leading Edge*, **25**(2), 156–166.
- Masaya, S. & Verschuur, D., 2018. Iterative reflectivity-constrained velocity estimation for seismic imaging, *Geophys. J. Int.*, **214**(1), 1–13.
- Matharu, G. & Sacchi, M., 2019. A subsampled truncated-newton method for multiparameter full-waveform inversion, *Geophysics*, **84**(3), R333–R340.
- Métivier, L., Brossier, R., Virieux, J. & Operto, S., 2013. Full waveform inversion and the truncated newton method, *SIAM J. Sci. Comput.*, **35**(2), B401–B437.
- Mora, P., 1989. Inversion = migration + tomography, *Geophysics*, **54**(12), 1575–1586.
- Mulder, W.A. & Plessix, R.-E., 2004. A comparison between one-way and two-way wave-equation migration, *Geophysics*, **69**(6), 1491–1504.
- Oh, J.-W. & Min, D.-J., 2013. Weighting technique using backpropagated wavefields incited by deconvolved residuals for frequency-domain elastic full waveform inversion, *Geophys. J. Int.*, **194**(1), 322–347.
- Plessix, R.E. & Mulder, W.A., 2004. Frequency-domain finite-difference amplitude-preserving migration, *Geophys. J. Int.*, **157**(3), 975–987.
- Plessix, R.E., Milcik, P., Rynja, H., Stopin, A. & Matson, K., 2013. Multi-parameter full waveform inversion: marine and land examples, *Leading Edge*, **32**(9), 1030–1038.
- Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss–Newton and full newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, **133**(2), 341–362.
- Provenzano, G., Brossier, R. & Métivier, L., 2023. Robust and efficient waveform-based velocity-model building by optimal transport in the pseudotime domain: Methodology, *Geophysics*, **88**(2), U49–U70.
- Qu, S. & Verschuur, D., 2021. An effective scheme of pseudo-time joint migration inversion with an AVO mitigating workflow, in *82th EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.
- Safari, M. & Verschuur, J., 2023. Joint migration inversion including Q effects: towards Q estimation, in *84th EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.
- Sava, P. & Biondi, B., 2004. Wave-equation migration velocity analysis. i. theory, *Geophys. Prospect.*, **52**(6), 593–606.
- Shen, P. & Symes, W.W., 2008. Automatic velocity analysis via shot profile migration, *Geophysics*, **73**(5), VE49–VE59.
- Shen, P., 2005. *Wave equation migration velocity analysis by differential semblance optimization*, Ph.D. dissertation, Rice University, pp. 1–194.
- Shen, P., Symes, W.W. & Stolk, C.C., 2003. Differential semblance velocity analysis by wave-equation migration, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Shin, C., Jang, C.S. & Min, D.-J., 2001. Improved amplitude preservation for prestack depth migration by inverse scattering theory, *Geophys. Prospect.*, **49**(5), 592–606.
- Staal, X.R., 2015. *Combined imaging and velocity estimation by Joint Migration Inversion*, Ph.D. dissertation, Delft University of Technology, pp. 1–121.
- Sun, Y. & Verschuur, E., 2020. Full wavefield modeling using rigorous one-way propagation, reflection and transmission operators, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 2694–2698.
- Sun, Y., Verschuur, E. & Qu, S., 2019. Research note: Derivations of gradients in angle-independent joint migration inversion, *Geophys. Prospect.*, **67**(3), 572–579.
- Symes, W.W., 2008. Migration velocity analysis and waveform inversion, *Geophys. Prospect.*, **56**(6), 765–790.
- Tang, Y., 2009. Target-oriented wave-equation least-squares migration/inversion with phase-encoded hessian, *Geophysics*, **74**(6), WCA95–WCA107.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.
- ten Kroode, F., 2014. A Lie group associated to seismic velocity estimation, in *Inverse Problems from Theory to Applications Conference*, Institute of Physics Publishing, pp. 142–146.
- Thorbecke, J.W., Wapenaar, K. & Swinnen, G., 2004. Design of one-way wavefield extrapolation operators, using smooth functions in wlsq optimization, *Geophysics*, **69**(4), 1037–1045.
- Valenciano, A.A., Biondi, B. & Guitton, A., 2006. Target-oriented wave-equation inversion, *Geophysics*, **71**(4), A35–A38.

- Valensi, R. & Baina, R., 2021. A time consistent waveform inversion (TWIN) method, in *82nd EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.
- van Leeuwen, T. & Mulder, W.A., 2010. A correlation-based misfit criterion for wave-equation traveltime tomography, *Geophys. J. Int.*, **182**(3), 1383–1394.
- Vigh, D., Kapoor, J. & Li, H., 2011. Full-waveform inversion application in different geological settings, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Wang, S., Chen, F., Zhang, H. & Shen, Y., 2013. Reflection-based full waveform inversion (RFWI) in the frequency domain, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1324–1328.
- Williamson, P., 1991. A guide to the limits of resolution imposed by scattering in ray tomography, *Geophysics*, **56**(2), 202–207.
- Wu, Z. & Alkhalifah, T., 2015. Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion, *Geophysics*, **80**(6), R317–R329.
- Xu, S., Wang, D., Chen, F., Lambaré, G. & Zhang, Y., 2012. Inversion on reflected seismic wave, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1–7.
- Yang, J., Huang, J., Li, Z., Zhu, H., McMechan, G.A. & Luo, X., 2021. Approximating the Gauss-Newton Hessian using a space-wavenumber filter and its applications in least-squares seismic imaging, *IEEE Trans. Geosci. Remote Sens.*, **60**, 1–13.
- Yang, J., Huang, J., Zhu, H., McMechan, G. & Li, Z., 2022. An efficient and stable high-resolution seismic imaging method: Point-spread function deconvolution, *J. Geophys. Res.: Solid Earth*, **127**(7), e2021JB023281.
- Yao, G., da Silva, N.V. & Wu, D., 2019. Reflection-waveform inversion regularized with structure-oriented smoothing shaping, *Pure Appl. Geophys.*, **176**(12), 5315–5335.
- Yao, G., Wu, D. & Wang, S., 2020. A review on reflection-waveform inversion, *Petrol. Sci.*, **17**, 334–351.
- Yu, J., Hu, J., Schuster, G.T. & Estill, R., 2006. Prestack migration deconvolution, *Geophysics*, **71**(2), S53–S62.
- Zhou, H., Amundsen, L. & Zhang, G., 2012. Fundamental issues in full waveform inversion, in *SEG Annual International Meeting, Expanded Abstracts*, Society of Exploration Geophysicists, pp. 1–5.
- Zhou, W., Brossier, R., Operto, S. & Virieux, J., 2015. Full waveform inversion of diving waves for velocity model building with impedance inversion based on scale separation, *Geophys. J. Int.*, **202**(3), 1535–1554.
- Zou, Z., Ramos-Martínez, J., Kelly, S., Ronholt, G., Langlo, L., Mavilio, A.V., Chemingui, N. & Lie, J., 2014. Refraction full-waveform inversion in a shallow water environment, in *76th EAGE Annual Conference & Exhibition, Extended Abstracts*, European Association of Geoscientists & Engineers, pp. 1–5.

APPENDIX A: OFFSET-SELECTION STRATEGY

Despite its conceptual appeal and aside from the resolution challenges, ORWI could face other limitations due to other factors: (1) Full-wave inconsistency in the short-offset residual data for tomography due to inconsistent reflectivity and velocity models (Valensi & Baina 2021; Liang *et al.* 2022; Provenzano *et al.* 2023). (2) The standard practice in the RWI community for mitigating the adverse imprint of inconsistent reflectivity and velocity models often involves using short-offset or near-zero offset data for migration (Provenzano *et al.* 2023). However, this can degrade the illumination and signal-to-noise ratio of the reflectivity model and result in an increased number of least-squares migration iterations. (3) Adverse contribution of cycle-skipped long-offset reflection data in tomography (e.g. Abolhassani & Verschuur 2022; Provenzano *et al.* 2023).

A1 Tomography offset: short offsets

To minimize the impact of inconsistent reflectivity and velocity models on the tomographic gradient of ORWI, we propose muting their relevant erroneous tomographic wave paths in the residual data gathers, which are typically established at short offsets.

A2 Migration offset

Building upon the muting of short-offset residual waveforms for tomography, we extend the migration offset to improve both the illumination and the signal-to-noise ratio of the reflectors. We follow Algorithm 2, aiming for an extended offset for migration (rather than short offset) where a significant reduction in the misfit function is impracticable. The misfit indeed remains unchanged as the amplitude fit fails to better itself while mapping more out-of-phase reflection data into the model space. This denotes the MEMO, lying averagely between the short offset and the maximum uncycle-skipped offset in the data. Algorithm 2 is applied as a quality control measure prior to ORWI. This algorithm calculates the misfit function value for various offset ranges up to the maximum offset within the acquisition. Following this, on the misfit-offset graph, we pick an offset right before the curve starts to become nearly horizontal on a logarithmic scale.

Algorithm 2 Migration offset analysis pseudo-code

```

for  $h \leftarrow 0$  to  $\max(\text{offset})$  do
  image  $\leftarrow$  1x LS migration with  $h$ 
   $\mathbf{d}_{\text{mod}}^- \leftarrow$  forward modeling with image and background velocity model
   $\Delta \mathbf{d}^- \leftarrow \mathbf{d}_{\text{obs}} - \mathbf{d}_{\text{mod}}^-$   $\triangleright \Delta \mathbf{d}^-$  is the full-offset residual data
  store the misfit function value associated with  $h$ 
end for

```

A3 Tomography offset: long offsets

A given pair of modelled and observed seismic traces in the time domain are called cycle-skipped traces if their time distance is larger than half a cycle time shift (Virieux & Operto 2009). Inspired by van Leeuwen & Mulder (2010), we defined Algorithm 3 as a time-domain two-step data-selection algorithm excluding the contribution of the cycle-skipped long-offset residual data in the tomography loop of ORWI to obtain a tomographic update free of the damaging effect of cycle-skipping in each iteration. In Algorithm 3, we identify a given pair of time-domain modelled and observed seismic traces as cycle-skipped if their maximum correlation lag exceeds a reference lag determined by the dominant period of the observed trace. Based on this criterion, our time-domain data-selection algorithm relies on the cross-correlation of these traces within a sliding window (referred to as local cross-correlation), safeguarded by a preliminary global cross-correlation of the given traces. This global cross-correlation functions as a protection, protecting subsequent local cross-correlations from irregularities in the modelled waveform. Without this safeguard, the outputs of the local cross-correlations would not meet the required accuracy.

Algorithm 3 Cycle-skipping check pseudo-code

```

▷  $\mathbf{D}_{\text{obs}}$  and  $\mathbf{D}_{\text{mod}}$  are 3D matrices representing the observed and modeled data (full-shot).
▷  $n_t \times n_g \times n_s$  is the size of  $\mathbf{D}_{\text{obs}}$ .
▷  $n_t$  is the total number of time samples.
▷  $n_g$  and  $n_s$  are the total number of receivers and the total number of shots.
▷  $\ell_w$  and  $\ell_r$  are the correlation-window length and reference lag.
▷  $\omega_c$  is the current maximum frequency in  $\mathbf{D}$ .
▷  $\omega_m$  is the peak frequency of the seismic wavelet.
▷  $\text{win}_c$  is the current correlation window.

function CYCLESKIPCHECK( $\mathbf{D}_{\text{obs}}, \mathbf{D}_{\text{mod}}, n_g, n_s, n_t, \ell_w, \ell_r, \omega_c, \omega_m$ )
  TIMETABLE  $\leftarrow$  a matrix of ones of size:  $n_g \times n_s$ 
  for  $i$ shot  $\leftarrow 1$  to  $n_s$  do
     $\varphi \leftarrow$  determine the anchor offset index in the current shot gather
     $\iota \leftarrow \varphi$ 
     $\kappa \leftarrow \varphi - 1$ 
    for  $m \leftarrow [1, 2]$  do
      flag  $\leftarrow$  true
      while  $\left( m = 1 \text{ and flag} = \text{true and } \iota \leq n_g \right) \text{ or } \left( m = 2 \text{ and flag} = \text{true and } \kappa \leq 1 \right)$  do
         $\mathbf{d}_{\text{obs}}, n_{\text{obs}} \leftarrow$  current observed trace and the associated first-break sample
         $\mathbf{d}_{\text{mod}} \leftarrow$  current modeled trace
        if  $\omega_c > \omega_m$  then
           $\mathbf{d}_{\text{obs}}, \mathbf{d}_{\text{mod}} \leftarrow \mathbf{d}_{\text{obs}}[n_{\text{obs}} : n_t], \mathbf{d}_{\text{mod}}[n_{\text{obs}} : n_t]$ 
        else
           $\mathbf{d}_{\text{obs}}, \mathbf{d}_{\text{mod}} \leftarrow \mathbf{d}_{\text{obs}}[1 : n_t], \mathbf{d}_{\text{mod}}[1 : n_t]$ 
        end if
        calculate cross-correlation between  $\mathbf{d}_{\text{obs}}$  and  $\mathbf{d}_{\text{mod}}$  ▷ global cross-correlation
        if max correlation lag  $< \ell_r$  then
           $\text{win} \leftarrow [1 : \ell_w : n_t]$ 
           $b \leftarrow \text{length}(\text{win})$ 
          for  $\tau \leftarrow 1$  to  $b$  do
             $\text{win}_c \leftarrow [1 : \text{win}(\tau + 1)]$ 
            calculate cross-correlation between  $\mathbf{d}_{\text{obs}}$  and  $\mathbf{d}_{\text{mod}}$  in  $\text{win}_c$  ▷ local cross-correlation
            if max correlation lag  $< \ell_r$  then
              if  $\tau = b$  then
                tmp  $\leftarrow n_t$ 
              else
                tmp  $\leftarrow n_{\text{obs}} + \text{win}_c(\text{end})$ 
              end if
              if  $m = 1$  then
                TIMETABLE $[\iota, i\text{shot}] \leftarrow$  tmp
              else
                TIMETABLE $[\kappa, i\text{shot}] \leftarrow$  tmp
              end if
            end if
          end for
          flag  $\leftarrow$  false
          break
        end if
      end for
    else
      flag  $\leftarrow$  false
    end if
    if  $m = 1$  then
       $\iota \leftarrow \iota + 1$ 
    else
       $\kappa \leftarrow \kappa - 1$ 
    end if
  end while
  end for
  end for
end function

```
