Brazilian Journal of Geophysics, **42** (2024), 35, 1–14
Brazilian Geophysical Society
ISSN 2764-8044
DOI 10.22564/brjg.v42i3.2330

PRECONDITIONING OF LEAST-SQUARES REVERSE-TIME MIGRATION WITH THE INVERSE-SCATTERING IMAGING CONDITION

Natiê A. Albano [1*], Jessé C. Costa [2], Jörg Schleicher [3], and Amélia Novais [3]

[1]Universidade Estadual de Campinas – Unicamp, CEPETRO / INCT-GP, Campinas, SP, Brazil
[2]Universidade Federal do Pará – UFPA / INCT-GP, Belém, PA, Brazil
[3]Universidade Estadual de Campinas – Unicamp, DMA/IMECC / CEPETRO / INCT-GP, Campinas, SP, Brazil

*Corresponding author: natieaalbano@gmail.com

ABSTRACT. The Inverse-Scattering Imaging Condition (ISIC) is an imaging condition for Reverse-Time Migration (RTM) that attempts to recover the medium reflectivity. It is theoretically based on the asymptotic inverse to the Born approximation and can be represented in several approximately equivalent forms. Its application leads to more reliable reflectivity estimates and strongly reduces backscattering artifacts. In this work, we demonstrate that an ISIC formulation involving a Laplacian filter can be used as an effective preconditioning for Least-Squares RTM (LSRTM). The Laplacian-filter ISIC does not increase the computational cost over conventional imaging conditions. Our numerical experiments using synthetic seismic data from the Marmousi II model demonstrate that this preconditioning leads to faster convergence and superior final images of LSRTM, both in the image domain (ID-LSRTM) and data domain (DD-LSRTM), in this way actually reducing computational cost and turnaround time.

Keywords: illumination compensation in LSRTM; reflectivity recovery; ISIC; seismic-migration preconditioning; backscattering-noise reduction.

INTRODUCTION

Seismic migration is a process that aims at generating detailed images of reflecting subsurface structures. This can be realized by means of various methods, one prominent example being reverse-time migration (RTM) (Schultz and Sherwood, 1980; Baysal et al., 1983; McMechan, 1983). Unlike other methods, RTM is based on the full wave equation, granting it the capability to produce images in models with strong velocity contrasts. Moreover, this approach offers the particular advantage of allowing to accurately image reflectors with arbitrary dip, making use of prismatic and diving waves (Leveille et al., 2011).

The RTM methodology relies on extrapolating the (forward) source and (backward) receiver wavefield, culminating in the construction of the migrated image via the application of the so-called imaging condition. The conventional imaging condition employed in reverse time migration (RTM) is a zero-lag crosscorrelation of the source and receiver wavefields, which finds its origins in the work of Claerbout (1971).

In spite of the many advantages of RTM, there are a few drawbacks. Depending on the geological complexity of the medium, the extrapolated wavefields can encompass waves traveling in both ascending and descending directions, such as diving waves, head waves and prismatic waves (Fletcher et al., 2005). In this situation, the correlation between forward and backward propagated wavefields introduces low-wavenumber noise in the image, commonly referred to as backscattering noise.

Another factor that frequently degrades the amplitude distribution in the final migrated image is uneven illumination of the subsurface target area. This may be caused by inadequate positioning of the sources and receivers and by the focusing and defocusing of the extrapolated wavefields, possibly including geometrical-spreading effects.

To mitigate these problems, a common practice involves post-processing the image resulting from zero-lag crosscorrelation as introduced byClaerbout (1971). Two frequently employed corrections are the application of a Laplacian filter to reduce backscattering noise, and the correction of the image amplitudes by the subsurface illumination of the source wavefield to account for focusing and defocusing effects (Biondi, 2006). This leads to the improved crosscorrelation-with-Laplacian-filter-and-illumination-compensation imaging condition (CLIIC).

Although the CLIIC produces a good estimate of the structures of subsurface reflectors, their amplitudes do not correspond to the true reflection coefficients of the subsurface (Kiyashchenko et al., 2007). Also, the image quality is frequently degraded by additional problems, such as the limited aperture of the acquisition geometry, illumination variations and band-limited frequency content of the source wavelet. Moreover, mathematically speaking, migration is the adjoint rather than the inverse operation to forward modeling, which causes a blur of the image points (Claerbout, 1992).

To further improve the migrated seismic images and make the amplitudes of the reflectors proportional to reflection coefficients, Kiyashchenko et al. (2007) and Op't Root et al. (2012) tried to formulate approximate inverse operators to forward modeling. Based on inverse-scattering theory, the latter authors carried out a microlocal analysis to derive the so-called inverse-scattering imaging condition (ISIC) that makes RTM the asymptotic inverse operation to acoustic forward modeling in the Born approximation.

The original ISIC formulation of Op't Root et al. (2012) and its first time-domain implementations consist of the sum of two separately generated images, one of which is formed by the correlation of the time derivatives of the downward extrapolated source and receiver wavefields and the other by the correlation of the corresponding spatial derivatives (Whitmore and Crawley, 2012; Pestana et al., 2014; Fang et al., 2017). The first implementations demonstrate that the ISIC can help to produce seismic images in which backscattering noise is greatly reduced and amplitudes become correlated with reflection coefficients. Because this original form of the ISIC requires the generation and sum of two images, its application is computationally more expensive than a conventional crosscorrelation imaging condition, even in the form of the CLIIC, where Laplacian filtering and illumination correction are included.

Recently, Albano et al. (2023) have compared several implementational forms of the ISIC, which are theoretically equivalent, except for a slightly different approximation of the inherent illumination correction. In two of these ISIC forms, the two images are mathematically combined into a single one. This is achieved by using the wave equation to substitute the temporal wavefield derivatives by spatial ones. In this way, the combination of the derivatives gives rise to the Laplacian operator, making the computation cost of these ISIC versions the same as that of the CLIIC.

However, the formulation of an adequately corrected migration operator as an asymptotic inverse to forward modeling is not able to solve all image deficiencies. Though it significantly reduces backscattering and image blurring, it does not consider the acquisition footprint or the band-limited nature of the data (Albano et al., 2023). These effects can be taken into account when migration is formulated as an inverse problem in the sense of Tarantola (1984), i.e., minimizing the discrepancy between the observed and modeled data. In its linearized form, with the data discrepancy measured by the squared data residual, this inverse problem is called least-squares migration. It aims at iteratively converting the adjoint operator into the best possible estimate of the true inverse operator, given a (fixed) background velocity model, in this way determining the model perturbations (also called reflectivity) with the best possible resolution (Nemeth et al., 1999; Dai and Schuster, 2009; Trad, 2015). The least-squares formulation can be applied to any seismic migration method. Here, we specifically consider least-squares reverse-time migration (LSRTM, see, e.g., Dai et al., 2012; Zhang et al., 2015; Feng and Schuster, 2017).

LSRTM can be implemented in two distinct domains, the data domain (DD-LSRTM) and the image domain (ID-LSRTM). In the data domain, it requires the iterative solution of a linear system. At each iteration, the scattered wavefield is modeled using the Born approximation. Subsequently, the residual between the modeled and the observed reflected data is migrated. The update of the reflectivity model is carried out in the direction of the steepest descent of the quadratic data residual (Duan et al., 2017; Yang et al., 2019). More advanced and costlier methods attempt to accelerate convergence by preconditioning the gradient direction with (an approximation of) the inverse Hessian.

In the image domain, the residual is defined as the difference between the migrated images of the modeled and the observed data. In this domain, the combined effect of all practical limitations, whether arising from the observed data, acquisition geometry, or employed migration method, can be simulated by determining the impulse responses to point sources distributed throughout the model, the so-called point-spread functions. The least-squares solution to the original inverse problem can then be expressed as a deconvolution of the initial migrated image with the inverses of these point-spread functions, eliminating the need for an iterative solution (Aoki and Schuster, 2009; Fletcher et al., 2016; Schuster, 2017). As a consequence, ID-LSRTM yields migrated images with

enhanced resolution at a reduced computational cost as compared to DD-LSRTM.

In this work, we investigated if the ISIC can help to precondition LSRTM so as to improve the convergence and/or increase the quality of the resulting migrated image without increasing the cost over the CLIIC. For this purpose, we implemented ISIC preconditioning for both ID-LSRTM and DD-LSRTM. By means of numerical experiments using the Marmousi II synthetic dataset, we show that, in both approaches, the preconditioned algorithm reaches a lower residual and converges faster to the solution, in this way reducing the computational cost.

LSRTM

Any kind of wave equation represents the wavefield to be determined as a nonlinear function of the medium parameters. The Born approximation linearizes the forward problem, calculating a wavefield perturbation as a function of a perturbation of the (fixed) background velocity distribution. In this linearized formulation of the forward problem, the modeled data $\mathbf{d}$ are obtained as a result of applying a linear modeling operator $\mathbf{L}$ (which depends only on the velocity background) to the model reflectivity $\mathbf{m}$ (Snieder and Trampert, 1999). In matrix form, this can be written as

$$\mathbf{d} = \mathbf{L}\mathbf{m}. \tag{1}$$

The inverse problem can then be formulated as a least-squares problem, where the objective is to find that particular model $\mathbf{m}^*$ that minimizes the square of the mismatch between the modeled data $\mathbf{d}$ and the observed reflection data $\mathbf{d}_{obs}$. In other words, the problem becomes to minimize the objective function

$$J_{DD}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{d}_{obs} - \mathbf{d}\right\|^2, \tag{2}$$

where the sums are carried out over all sources and receivers. This is achieved in an iterative form, where each iteration uses a migration of the data residual $\mathbf{d} - \mathbf{d}_{obs}$ to estimate a model update $\Delta\mathbf{m}$ that further reduces the value of the objective function in the data domain, equation 2. In this way, DD-LSRTM aims at constructing the best possible migrated image $\mathbf{m}^*$ for a given background velocity distribution (Nemeth et al., 1999; Prucha and Biondi, 2002; Dai et al., 2012; Duveneck et al., 2021).

According to Menke (1989), the least-squares solution to the minimization problem 2, given the forward modeling operation 1, can be expressed as

$$\mathbf{L}^T\mathbf{L}\mathbf{m}^* = \mathbf{L}^T\mathbf{d}_{obs}, \tag{3}$$

where $\mathbf{L}^T$ denotes the adjoint to the modeling operator $\mathbf{L}$.

Recognizing this operator as a migration, to be carried out as an RTM here, we can identify the right-hand side of equation 3 as the migrated data

$$\mathbf{m}_{mig} = \mathbf{L}^T\mathbf{d}_{obs}. \tag{4}$$

Applying the migration operator $\mathbf{L}^T$ to both data terms in equation 2, we can reformulate the least-squares problem as the one of minimizing the objective function in the image domain

$$J_{ID}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{L}^T\mathbf{L}\mathbf{m} - \mathbf{m}_{mig}\right\|^2, \tag{5}$$

with the same theoretical solution given in equation 3. We can then proceed iteratively in a similar way as previously to find the solution that minimizes equation 5, where at each iteration, the data $\mathbf{d}$ need to be remodeled and then migrated.

However, equation 3 allows for a different interpretation. It demonstrates that the best model $\mathbf{m}^*$ in a least-squares sense is obtained when the migrated data $\mathbf{m}_{mig}$ can be represented by multiplying $\mathbf{m}^*$ with a blurring operator, $\mathbf{\Psi} = \mathbf{L}^T\mathbf{L}$. If $\mathbf{\Psi}$ can be, at least approximately, inverted, we can represent the desired solution $\mathbf{m}^*$ as

$$\mathbf{m}^* = \mathbf{\Psi}^{-1}\mathbf{m}_{mig}. \tag{6}$$

This gives rise to an alternative way of approximately determining $\mathbf{m}^*$ in the image domain, directly from $\mathbf{m}_{mig}$, using an estimate of $\mathbf{\Psi}$ or $\mathbf{\Psi}^{-1}$, without the need for an iterative minimization of the objective function 5.

Because it is generally easier to estimate $\mathbf{\Psi}$ rather than $\mathbf{\Psi}^{-1}$, the solution 6 is usually constructed by (iteratively) solving the linear problem $\mathbf{\Psi}\mathbf{m}^* = \mathbf{m}_{mig}$, where the migrated image $\mathbf{m}_{mig}$ and an estimate of $\mathbf{\Psi}$ are known (Guitton, 2004; Valenciano, 2008; Fletcher et al., 2016; Guo and Wang, 2020; Osorio et al., 2021).

DD-LSRTM

In the data domain, LSRTM aims at iteratively estimating the solution $\mathbf{m}^*$ defined in equation 3 that minimizes the objective function 2. The iterative procedure is obtained by noting that, at $\mathbf{m}^*$, we expect the gradient of the objective function in equation 2 to be zero, i.e., $\nabla J(\mathbf{m}^*) = 0$. Under the assumption that we are at a point $\mathbf{m}$ close to $\mathbf{m}^*$, i.e.,

$$\mathbf{m}^* = \mathbf{m} + \Delta\mathbf{m}\,, \qquad \Delta\mathbf{m} \ll \mathbf{m}\,, \tag{7}$$

we can write, up to first order in $\Delta\mathbf{m}$,

$$\nabla J(\mathbf{m}^*) = \nabla J(\mathbf{m}) + \mathbf{H}(\mathbf{m})\Delta\mathbf{m} = 0\,, \tag{8}$$

where $\mathbf{H}(\mathbf{m})$ denotes the Hessian of the objective function. As long as $\mathbf{H}(\mathbf{m})$ is invertible, we can conclude that the difference between the present model $\mathbf{m}$ and the desired model $\mathbf{m}^*$ is

$$\Delta\mathbf{m} = -\mathbf{H}^{-1}(\mathbf{m})\nabla J(\mathbf{m})\,. \tag{9}$$

Since equation 8 is a linear approximation of the gradient at $\mathbf{m}^*$, equation 9 is also an approximation. It gives rise to an iterative model updating technique. In its most basic version, the inverse of the Hessian is replaced by a multiple of the identity matrix, resulting in (see, e.g., Schuster and Liu, 2019)

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha \nabla J^k(\mathbf{m}), \tag{10}$$

where α denotes the step size. In this work, we adopt the iterative model-updating procedure described by equation 10, known as the gradient method.

Also, we consider the forward problem to be described by the Born approximation to the acoustic wave equation for constant density (see Appendix A), given by

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = \frac{m(\mathbf{x})}{c_0^2(\mathbf{x})}\frac{\partial^2 p_0(t,\mathbf{x})}{\partial t^2}, \tag{11}$$

where $\Delta p(t,\mathbf{x})$ is the scattered wavefield, and $c_0(\mathbf{x})$ and $p_0(t,\mathbf{x})$ denote the wave-propagation velocity and the acoustic pressure wavefield in the background medium. Note that equation 11 depends linearly on the medium reflectivity, $m(\mathbf{x})$, which represents the model parameter to be recovered by LSRTM.

In this situation, the gradient of the objective function can be expressed as

$$\nabla J(\mathbf{m}) = \sum_{\mathbf{x}_s} \int_{\Omega(\mathbf{x})} d\Omega \frac{1}{c_0^2(\mathbf{x})} \int_0^T dt \left[p_r(t,\mathbf{x}) \frac{\partial^2 p_s(t,\mathbf{x})}{\partial t^2} \right], \tag{12}$$

where $p_s(t,\mathbf{x})$ denotes the source-side wavefield at an image point $\mathbf{x}$, modeled directly by means of equation 11, and $p_r(t,\mathbf{x})$ is the receiver-side wavefield at the same image point, obtained by backward propagation of the data residual $\mathbf{d} - \mathbf{d}_{obs}$, i.e., $p_r(t,\mathbf{x})$ is the solution of the adjoint wave equation

$$\frac{1}{c^2}\frac{\partial^2 p_r(t,\mathbf{x})}{\partial t^2} - \nabla^2 p_r(t,\mathbf{x}) = \sum_{\mathbf{x}_r} \left[p_{obs}(t,\mathbf{x}) - \Delta p(t,\mathbf{x}) \right] \delta(\mathbf{x} - \mathbf{x}_r). \tag{13}$$

Thus, the gradient of the objective function is the result of the zero-lag crosscorrelation between the second temporal derivative of the forward modeled source wavefield $p_s(t,\mathbf{x})$ and the backward modeled receiver wavefield $p_r(t,\mathbf{x})$, weighted by the inverse square of the velocity. In other words, it is determined by an RTM of the data residual.

ID-LSRTM

In the image domain, LSRTM aims at minimizing the objective function 5. However, as mentioned earlier, the iterative minimization in the image domain is rather expensive, because it requires a modeling plus migration at every iteration. An alternative approach to ID-LSRTM makes use of equation 3, which is the least-squares solution to the minimization problem in both formulations of equations 2 and 5. It can be reformulated in the image domain as

$$\mathbf{\Psi}\mathbf{m}^* = \mathbf{m}_{mig} \ , \tag{14}$$

where $\mathbf{\Psi} = \mathbf{L}^T\mathbf{L}$ is the blurring operator that contains the combined effect of the modeling and migration operators. Equation 14 expresses the fact that the migrated image $\mathbf{m}_{mig}$ is a blurred version of the reflectivity that results from the convolutional relationship between the subsurface reflectivity and the subsequent application of the modeling and migration operators (Valenciano, 2008). Thus, instead of the computationally expensive procedure of iteratively minimizing the objective function 5, another way of estimating the desired model $\mathbf{m}^*$ is to approximately invert the blurring operator $\mathbf{\Psi}$.

Because of its high computational cost, an exact implementation of the blurring operator is highly impractical. For this reason, several approaches to approximately computing this function have been proposed in the literature. Guitton (2004) directly approximates $\mathbf{\Psi}^{-1}$ by employing a set of non-stationary matching filters. Valenciano et al. (2006) introduce a target-oriented, sparse version of the blurring operator, significantly reducing its dimensions. In a concurring approach, Jiang and Zhang (2019) propose a strategy involving a localized approach that amounts to breaking down the blurring operator into a series of smaller matrices, which are computationally manageable.

A particularly attractive procedure to estimate the blurring operator $\mathbf{\Psi}$ is to study its impulse responses for specific isolated points in the image. In other words, one replaces the model $\mathbf{m}^*$ in equation 14 by a $\boldsymbol{\delta}(\mathbf{x})$ matrix representing a point source at image point $\mathbf{x}$ and carries out a modeling followed by a migration. The result is

$$\mathbf{\Psi}\boldsymbol{\delta}(\mathbf{x}) = \psi(\mathbf{x}), \tag{15}$$

known as the point-spread function (PSF) for image point $\mathbf{x}$ (Lecomte, 2008). Generally, a PSF is spatially limited around the image point for which it is calculated. Thus, after determining PSFs for a sufficiently dense grid of image points, they can be used to locally deconvolve the migrated image, in this way reconstructing an approximation to $\mathbf{m}^*$ at an acceptable cost (Valenciano, 2008; Fletcher et al., 2016; Osorio et al., 2021).

In this work, we estimate the PSFs as follows. Initially, we generate linearized data using a model consisting of a grid of point scatterers added to the given background model. The grid points must be adequately spaced to strike a balance: they should be spread far enough apart to prevent interference of the PSFs while being also positioned closely enough for effective interpolation, ensuring comprehensive coverage of the entire model. Then, the wavefield re-

sulting from Born modeling in this model is injected at the receiver positions and backward propagated in time. Finally, the PSFs are obtained by applying an imaging condition using the forward and backward propagated wavefields at each of the considered grid points.

When migrating the scattered data originating from the diffraction points, the imaging condition will exhibit constructive interference only at image points in the vicinity of actual scatterer positions, in this way generating the PSFs associated with these image points, as illustrated in Figure 1.

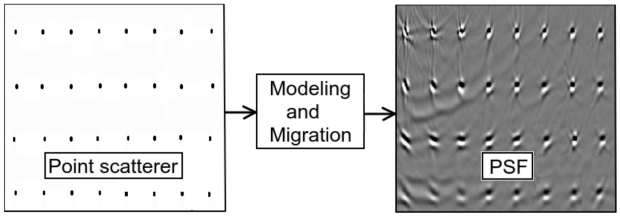

Figure 1: Schematic explanation of the generation of PSFs. Scattered data are generated with the model on the left-hand side and then migrated to obtain the image on the right-hand side. The blurred image of each diffractor is the PSF associated with the corresponding location in the model.

Each PSF $\psi(\mathbf{x})$ forms a column in the overall blurring matrix $\mathbf{\Psi}$ in equation 14. A window is used to extract a single PSF from the image in Figure 1. This windowing technique helps ensure that the solution remains free from unwanted influence, in this way assuring the accuracy of the estimated PSFs. At image points where no PSF can be estimated because of the required spacing between the point scatterers, it is necessary to calculate them using interpolation. Frequently, it is necessary to perform additional PSF estimations with displaced grids of point scatterers, because the modeling-plus-migration procedure can result in rather broad or irregular PSFs.

Specifically, for acoustic modeling as adopted in this work, the forward wavefield is modeled using the Born approximation 11 and the backpropagation carried out solving equation 13, where the adjoint sources are substituted with the scattered wavefield at the receivers. The imaging condition commonly applied in the literature is the crosscorrelation of the forward and backward propagated wavefields (Claerbout, 1971), enhanced with a Laplacian filter and illumination compensation (CLIIC), given by

Each PSF $\psi(\mathbf{x})$ forms a column in the overall blurring matrix $\mathbf{\Psi}$ in equation 14. A window is used to extract a single PSF from the image in Figure 1. This windowing technique helps ensure that the solution remains free from unwanted influence, in this way assuring the accuracy of the estimated PSFs. At image points where no PSF can be estimated because of the required spacing between the point scatterers, it is necessary to calculate them using interpolation. Frequently, it is necessary to perform additional PSF estimations with displaced grids of point scatterers, because the modeling-plus-migration procedure can result in rather broad or irregular PSFs.

Specifically, for acoustic modeling as adopted in this work, the forward wavefield is modeled using the Born approximation 11 and the backpropagation carried out solving equation 13, where the adjoint sources are substituted with the scattered wavefield at the receivers. The imaging condition commonly applied in the literature is the crosscorrelation of the forward and backward propagated wavefields (Claerbout, 1971), enhanced with a Laplacian filter and illumination compensation (CLIIC), given by

$$I_{ID}(\mathbf{x}) = \sum_s \frac{1}{P(\mathbf{x},\mathbf{x}_s)} \nabla^2 \int_t dt \left[p_s(t,\mathbf{x}) p_r(t,\mathbf{x})\right], \tag{16}$$

where $p_s(t,\mathbf{x})$ and $p_r(t,\mathbf{x})$ are the source and receiver wavefields, and $P(\mathbf{x},\mathbf{x}_s)$ is the illumination-compensation factor, calculated as the autocorrelation of the source wavefield, i.e.,

$$P(\mathbf{x},\mathbf{x}_s) = \int_t \left[p_s(t,\mathbf{x})\right]^2 dt. \tag{17}$$

INVERSE-SCATTERING IMAGING CONDITION

The inverse-scattering imaging condition (ISIC) as derived by Op't Root et al. (2012) is the asymptotic inverse operator to Born modeling. Its primary purpose is to correctly treat the image amplitudes, ensuring that they are directly proportional to the subsurface reflectivity. The ISIC is typically implemented in the time domain. It forms the image as the sum of two contributions. One of these contributions results from the multiplication of the temporal derivatives of the source and receiver wavefields, weighted with the sloth, i.e., the square of the slowness, at the image point. The second contribution arises from the multiplication of the spatial derivatives of these same wavefields. Finally, after the sum, an illumination-compensation factor guarantees amplitude balancing among different parts of the image (see, e.g., Whitmore and Crawley, 2012; Pestana et al., 2014). Thus, in its most basic form, the ISIC in the time domain reads

$$I(\mathbf{x}) = \sum_s \frac{1}{P(\mathbf{x},\mathbf{x}_s)} \int_t dt \left[\frac{1}{c^2(\mathbf{x})} \frac{\partial p_s(t,\mathbf{x})}{\partial t} \frac{\partial p_r(t,\mathbf{x})}{\partial t} + \frac{\partial p_s(t,\mathbf{x})}{\partial \mathbf{x}} \frac{\partial p_r(t,\mathbf{x})}{\partial \mathbf{x}}\right]. \tag{18}$$

We will refer to this form of the ISIC as the basic ISIC (B-ISIC).

As discussed in Whitmore (2013), equation 18 is intended to mitigate backscatter noise in the image. However, factors like the structural complexity of the medium, the presence of caustics, variable den-

sity, elasticity and anisotropy can lead to incomplete elimination of low-wavenumber artifacts from migration when the two contributions are equally weighted. To address this challenge, the authors propose the use of customized weights for each of the contributions. They demonstrate that a careful choice of these weights can lead to a resulting image that is almost completely free from backscatter noise, in this way offering a better representation of the subsurface. Unfortunately, Whitmore (2013) did not specify the explicit form of the illumination-compensation factor $P(\mathbf{x}, \mathbf{x}_s)$ he used in his experiments. We suppose here that he use the form given in equation 17.

Using microlocal analysis in the frequency domain, Op't Root et al. (2012) derived a theoretically more precise version of the ISIC as an asymptotic inverse of Born scattering. In a recent study, Albano et al. (2023) demonstrated that their form of the ISIC can be realized in the time domain through various expressions that are theoretically approximately equivalent. Three of these ISIC formulations closely resemble equation 18, i.e., they involve the sum of products of the temporal and spatial derivatives of the source and receiver wavefields. The key difference lies in how time derivatives or time integrals of the involved wavefields are used in the expressions. One favorable form reads

$$I_r(\mathbf{x}) = \sum_s \frac{1}{P'(\mathbf{x}, \mathbf{x}_s)} \int_t dt \left[\frac{\partial p_s(t, \mathbf{x})}{\partial t} \frac{\partial q_r(t, \mathbf{x})}{\partial t} + c^2(\mathbf{x}) \frac{\partial p_s(t, \mathbf{x})}{\partial \mathbf{x}} \frac{\partial q_r(t, \mathbf{x})}{\partial \mathbf{x}} \right], \tag{19}$$

where the illumination-compensation factor $P'(\mathbf{x}, \mathbf{x}_s)$ is calculated using the time-differentiated source wavefield, i.e., it is given by

$$P'(\mathbf{x}, \mathbf{x}_s) = \int_t \left[\frac{\partial p_s(t, \mathbf{x})}{\partial t} \right]^2 dt. \tag{20}$$

The main difference of equation 19 with respect to equation 18 is the use of the time-integrated receiver wavefield

$$q_r(t, \mathbf{x}_s) = \int_0^t p_r(t', \mathbf{x}) dt' \tag{21}$$

instead of $p_r(t, \mathbf{x})$ in both contributions, which is compensated by the different illumination compensation. We will refer to version 19 of the ISIC as the integrated ISIC (I-ISIC). Another noteworthy difference is the scale factor of c^2 multiplying both contributions and canceling the $1/c^2$ weight of the time-derivative contribution in equation 18. As shown by Albano et al. (2023), in the form of equation 19, the ISIC does not require amplitude balancing between the two contributions to achieve almost complete elimination of backscattering noise.

Note that because of the necessary time derivatives, both ISIC versions in equations 18 and 19 come at a somewhat higher computational cost than the basic CLIIC of equation 16. Making use of the wave equation, Albano et al. (2023) demonstrated that the ISIC can be reformulated only with spatial derivatives. The resulting ISIC versions can be expressed in terms of the Laplacian operator applied to a single crosscorrelation term, in a very similar way to the CLIIC of equation 16. One of these Laplacian-based ISIC (L-ISIC) versions reads

$$I_r(\mathbf{x}) = \sum_s \frac{c^2(\mathbf{x})}{2R(\mathbf{x}; \mathbf{x}_s)} \nabla^2 \int_t dt \, [r_s(t, \mathbf{x}; \mathbf{x}_s) q_r(t, \mathbf{x})] \, , \tag{22}$$

where $r_s(t, \mathbf{x})$ is twice time-differentiated receiver wavefield. Moreover, in this L-ISIC version, the illumination-compensation factor needs to be modified to

$$R(\mathbf{x}, \mathbf{x}_s) = \int_t [r_s(t, \mathbf{x})]^2 \, dt. \tag{23}$$

When implemented in this form, the L-ISIC has a nearly identical computational cost as the CLIIC. However, it retains the same benefits as the two-term ISIC expression of equation 19.

LSRTM preconditioning based on the ISIC

Taking into account the advantages of the ISIC discussed in the previous section, we study its application as a preconditioner for LSRTM. To avoid a significant increase in computational cost as compared to the conventional approach, we adopt the formulation based on the Laplacian operator with only a single crosscorrelation term, as expressed in equation 22. For the purpose of comparison, we also include the results obtained with the basic form of the ISIC (B-ISIC), equation 18 and with the two-term form using the integrated receiver wavefield (I-ISIC), equation 19.

For DD-LSRTM, we replace the imaging condition 12 in the calculation of the objective-function gradient with the different versions of the ISIC from equations 18, 19, and 22, to proceed with the iterative model-updating procedure of equation 10. For application in ID-LSRTM, we construct the migrated image of the scatterer grid to generate the PSFs using the four tested imaging conditions. Then, we construct matrix $\mathbf{\Psi}$ from the individual PSFs to proceed with the solution of the linear system 14.

NUMERICAL RESULTS

We evaluate the results of LSRTM preconditioned with the L-ISIC, both in the data and in the image domains, by means of experiments with the Marmousi II velocity model (Figure 2). We used the true P-wave velocity model, depicted in Figure 2, to model the observed data, necessary to construct the migrated image, with an implementation of the acoustic wave equation with constant density, equation A1.

The acquisition geometry for this study involved a total of 90 sources, evenly spaced at 100 m inter-

vals along the length of the model. The first source was positioned laterally at 15 m. Additionally, 175 receivers were placed along the surface, also starting at 15 m and regularly spaced at 50 m intervals. The depths of the sources and receivers were 50 m and 125 m, respectively. For reference, the remaining parameters utilized in data modeling are summarized in Table 1.

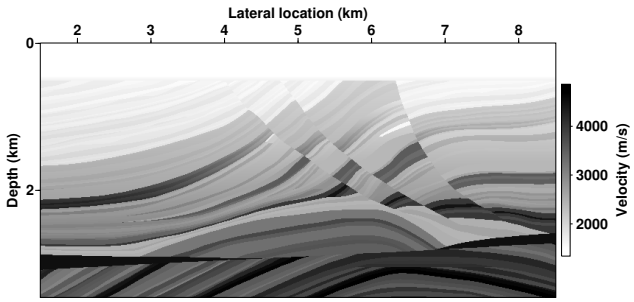

Figure 2: True P-wave velocity distribution of the Marmousi II velocity model.

Table 1: Modeling parameters used for the generation of the observed data.

Source pulse	**Ricker wavelet**
Peak frequency (Hertz)	15.0
Acquisition time (seconds)	5.0
Sampling rate (seconds)	0.004
Boundary conditions	Absorbing boundary
Boundary-layer size (points)	50

Prior to migration, we preprocessed the simulated data to eliminate direct waves. Subsequently, we applied an RTM using a smoothed version of the Marmousi II velocity model (Figure 3a) to generate the initial migrated data. The final objective of LSRTM is to recover the reflectivity (Figure 3b), i.e., the normalized difference between the true model and the smooth background model.

Figure 4 compares the effect of employing the ISIC versions (B-ISIC, I-ISIC and L-ISIC, equations 18, 19, and 22, respectively) with the conventional CLIIC, equation 16, in the initial RTM. We note that all four images present no backscattering artifacts and are similar with respect to resolution. When comparing the initial RTM image obtained with the CLIIC (Figure 4a) with the corresponding images using the ISIC versions (Figures 4b, 4c and 4d), it becomes apparent that the CLIIC image exhibits inferior amplitudes, diminishing from top to bottom, than the images obtained with all three versions of the ISIC. Figure 4b shows that the basic ISIC version of equation 18 slightly improves amplitude balancing, but the image is still very similar to that obtained with the CLIIC of equation 16 (Figure 4a). It is important to stress that our version of the B-ISIC did not need any additional amplitude balancing between the two contributions.

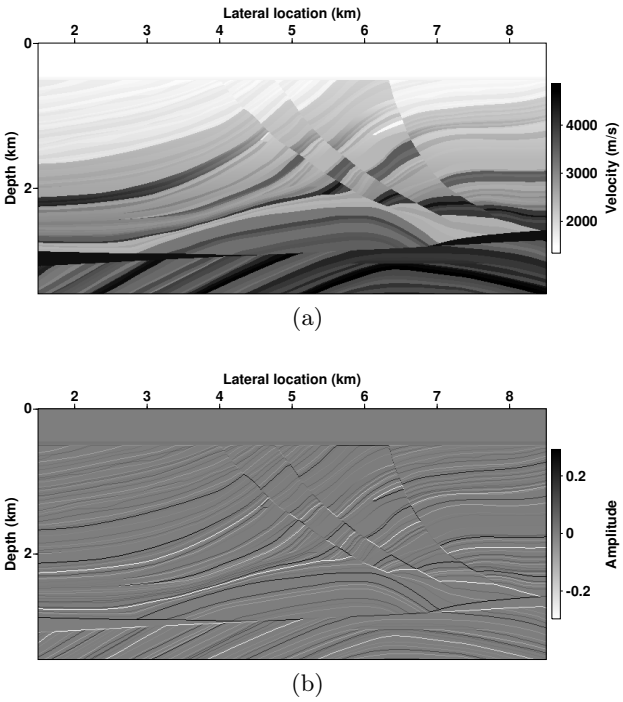

Figure 3: (a) Smoothed Marmousi II velocity model used as the background model for migration. (b) Reflectivity, i.e., normalized difference between (a) and the true model in Figure 2.

The two ISIC versions that theoretically honor the derivation of Op't Root et al. (2012) produce very similar images (Figures 4c and 4d) with amplitude distributions that closely resemble that of the true reflectivity (Figure 3b). The image obtained with the L-ISIC (equation 22, Figure 4d) presents a slightly better resolution than the one with the I-ISIC (equation 19, Figure 4c). It is to be observed that the L-ISIC image (Figure 4d) comes at a computational cost that is practically the same as that of the CLIIC (Figure 4a), while the two-term images of Figures 4b and 4c require more computational effort.

DD-LSRTM

For the implementation in the data domain, the update of the reflectivity model was carried out using ADAM optimization (Kingma and Ba, 2014; Kim et al., 2019), because the available Limited-memory-Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization algorithm (Liu and Nocedal, 1989) did not work equally well with all tested imaging conditions.

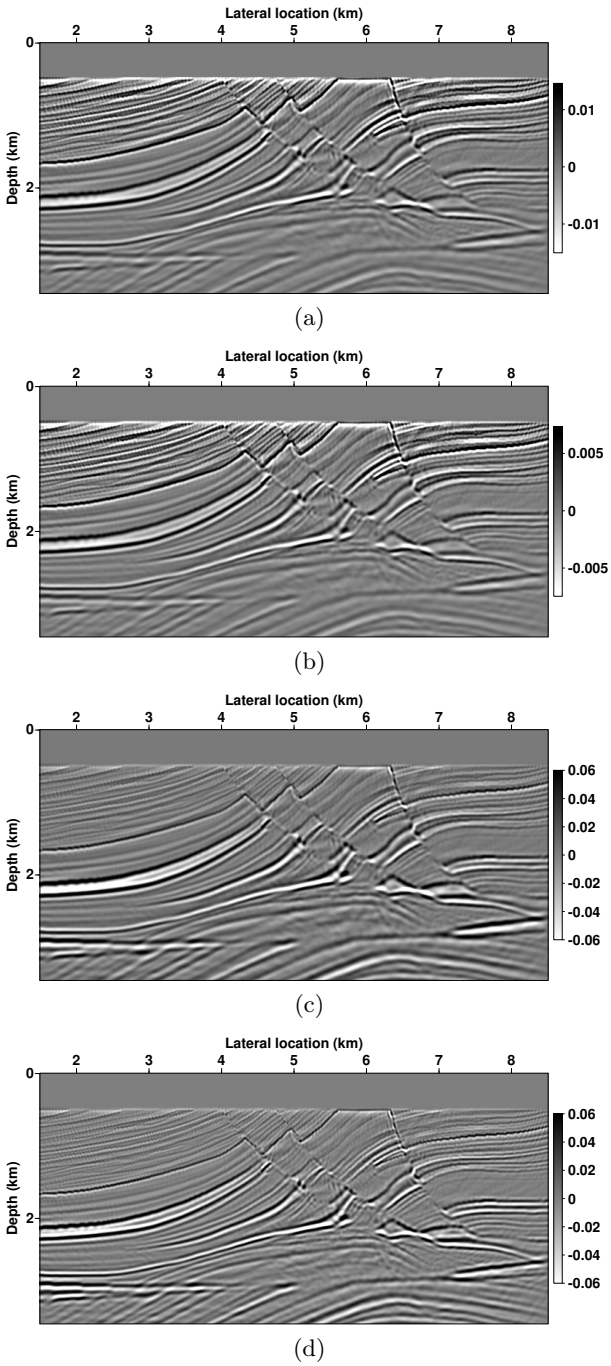

Figure 4: RTM images using different imaging conditions. (a) CLIIC, equation 16, (b) B-ISIC, equation 18, (c) I-ISIC, equation 19, (d) L-ISIC, equation 22.

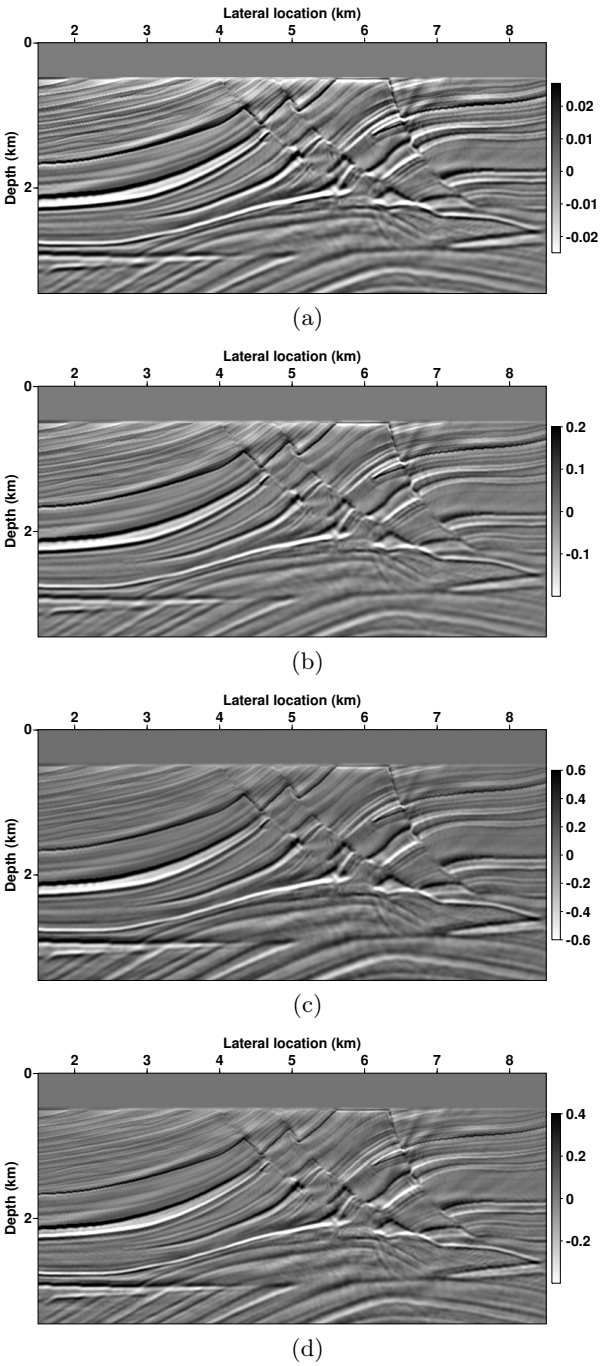

Figure 5: DD-LSRTM images of the Marmousi II model after 30 iterations using different imaging conditions. (a) CLIIC, equation 16, (b) B-ISIC, equation 18, (c) I-ISIC, equation 19, (d) L-ISIC, equation 22.

The LSRTM results in the data domain after 30 iterations with the four imaging conditions discussed here are presented in Figure 5. Figures 5a and 5b present the results of DD-LSRTM with the CLIIC and B-ISIC, and Figures 5c and 5d show the results using the I-ISIC and L-ISIC, respectively. For comparison with the true reflectivity model, please refer to Figure 3b, and to compare with the original RTM images, look at Figure 4.

The differences among the DD-LSRTM images in Figure 5 are much more subtle than among the corresponding original RTM results of Figure 4, indicating that the least-squares process is able to significantly improve the image quality even when starting from a worse initial image. Nevertheless, the overall impression is that the ISIC images (Figures 5b, 5c and 5d) appear cleaner than the CLIIC image (Figure 5a). The amplitudes of the theoretically equivalent I-ISIC

and L-ISIC images (Figures 5c and 5d) are a better match to the true reflectivity (Figure 3b) than the CLIIC and B-ISIC images (Figures 5a and 5b). Again, similarly to the pure RTM images of Figure 4, the overall impression is that the best image in terms of artifacts and resolution is the L-ISIC image (Figure 5d).

ID-LSRTM

In the image domain approach, the first step involves estimating the PSFs in the initial model. For that purpose, we define four grids of point scatterers, each shifted by half the grid spacing of the first grid in the vertical, horizontal and diagonal directions. In each section, the scatterers were spaced at 0.5 km in both directions. This arrangement ensures that the scatterers are placed at sufficient distance from each other to prevent interference between neighboring PSFs. For the unsampled points on the dense velocity grid between the scatterers, we computed the PSFs using bilinear interpolation so as to construct the full blurring matrix $\mathbf{\Psi}$. We then solved the linear system 14 using the Conjugate Gradient optimization algorithm to recover an approximation to $\mathbf{m}^*$.

Figure 6 shows the sets of PSFs estimated using the different imaging conditions: CLIIC (Figure 6a), B-ISIC (Figure 6b), I-ISIC (Figure 6c) and L-ISIC (Figure 6d). We note that the set of PSFs estimated with the I-ISIC (Figure 6c) and the L-ISIC (Figure 6d) exhibit more uniform illumination than the PSFs obtained with the CLIIC (Figure 6a) and the B-ISIC (Figure 6b), particularly in the deeper areas of the model. This is in agreement with the previous observation (see also Albano et al., 2023) that the ISIC provides a better amplitude balance in the image. When comparing the PSFs of the I-ISIC (Figure 6c) and of the L-ISIC (Figure 6d), we note that they are very similar, but the latter are a little more compact than the former. This is consistent with the observation that the migrated image using the L-ISIC (Figure 5d) presents a slightly better resolution than the one using the I-ISIC (Figure 5c). Though the PSFs are designed to remove these effects by deconvolving each image with the associated blurring operator $\mathbf{\Psi}$, it can be expected that any lack of information in the PSFs will lead to less quality in the final least-squares image.

The resulting ID-LSRTM images are shown in Figure 7. At first glance, the four images look very similar, indicating that the deconvolution with the PSFs has done a good job in removing the differences among the images of Figure 4. This could be interpreted as a sign that the employed imaging condition is of little importance in ID-LSRTM. However, closer inspection reveals a few differences like, e.g., the noise in the center of the image at 2 km depth and 5.5 km lateral position, which is rather strong in the CLIIC image (Figure 7a), a bit weaker in the B-ISIC image (Figure 7b), and almost completely absent in the I-ISIC and L-ISIC images (Figures 7c and 7d). Also, there are a few differences in resolution, e.g., at the salt wedges on both sides of the model at about 2.5 km depth, which are visibly best resolved in the L-ISIC image (Figure 7d).

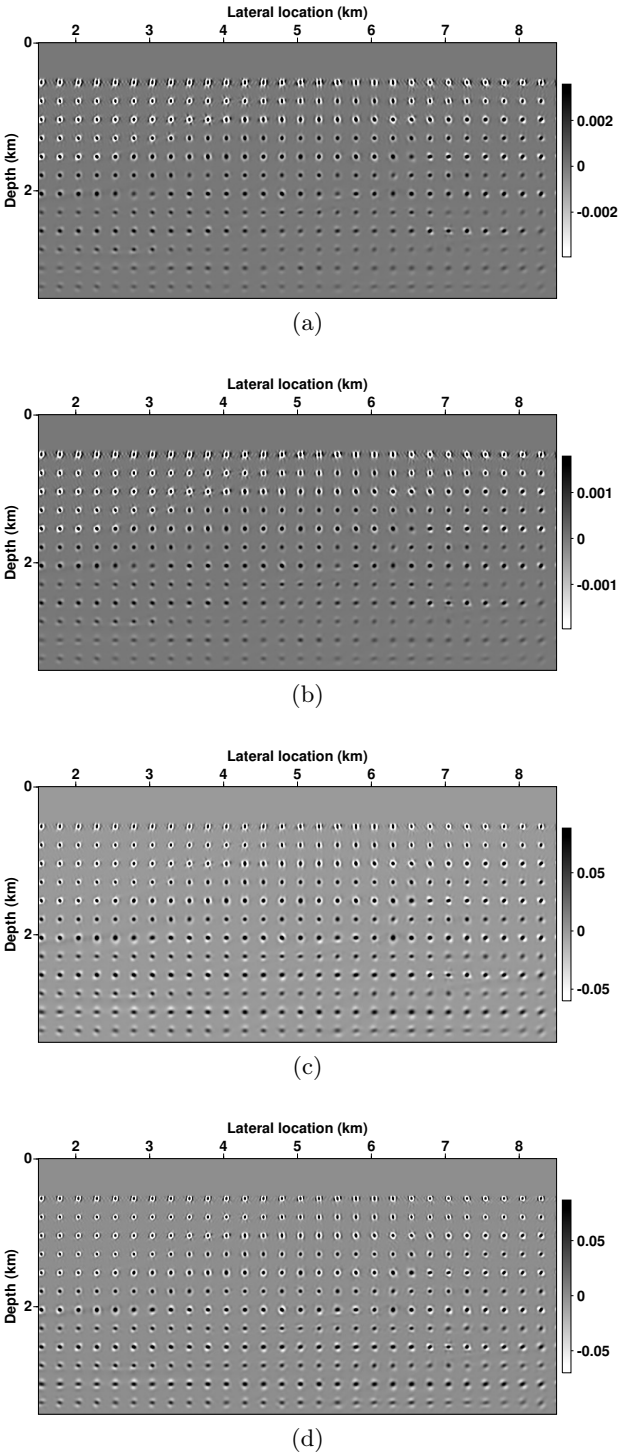

Figure 6: Point-spread functions for a grid of point scatterers as obtained from RTM using different imaging conditions. (a) CLIIC, equation 16, (b) B-ISIC, equation 18, (c) I-ISIC, equation 19, (d) L-ISIC, equation 22.

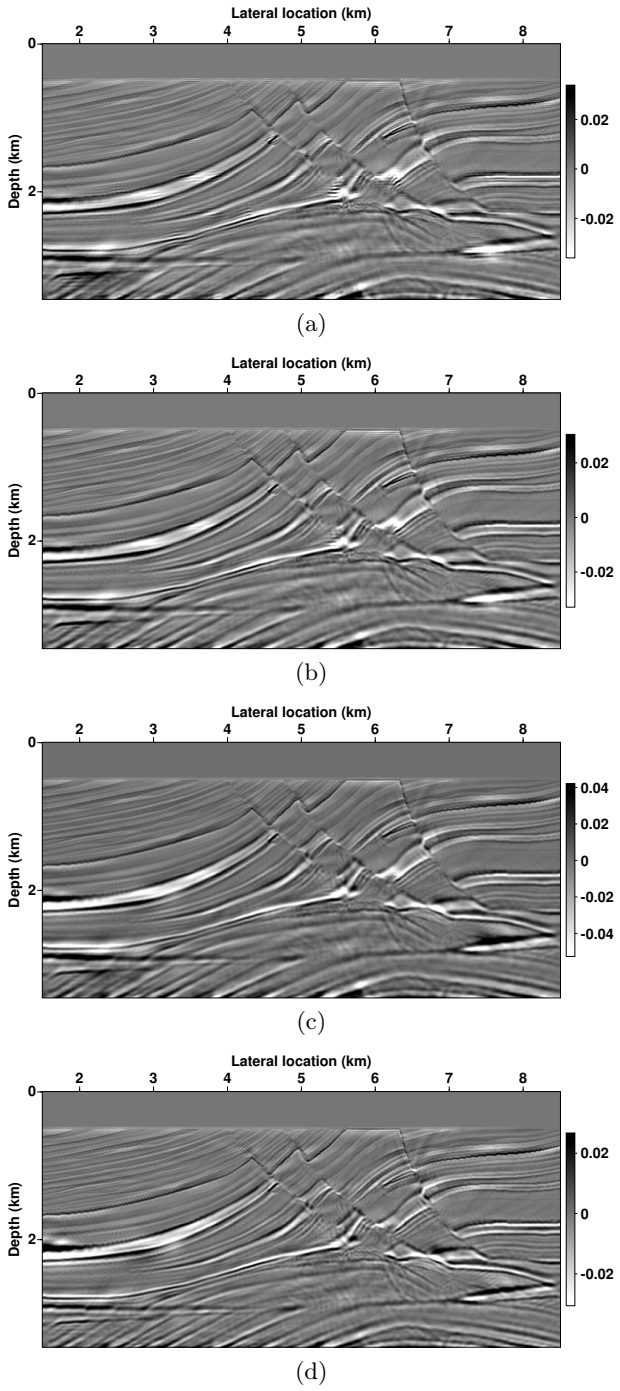

Figure 7: ID-LSRTM image of the reflectivity of the Marmousi II model with different imaging conditions. (a) CLIIC, equation 16, (b) B-ISIC, equation 18, (c) I-ISIC, equation 19, (d) L-ISIC, equation 22.

Comparison

When comparing the images in the data domain (Figure 5) with those in the image domain (Figure 7), we note that the latter are superior in resolution, but appear generally noisier. The PSF deconvolution in the image domain has done a better job of removing the differences caused by the different imaging conditions than the data-domain optimization. It must be noted, though, that at several locations, coherent noise that was very weak in the original RTM images (Figure 4) has been enhanced in the DD-LSRTM images (Figure 5), and even more so in the ID-LSRTM versions (Figure 7).

Convergence rate

There is one significant difference of preconditioning LSRTM with the four different imaging conditions, which is the convergence rate. As illustrated in Figure 8, the convergence rate varies in both the data (Figure 8a) and image (Figure 8b) domains in dependence on the imaging condition. In both domains, convergence with all forms of the ISIC is much faster than when using the CLIIC. The two theoretically correct forms (I-ISIC and L-ISIC) achieve the best convergence, while B-ISIC remains intermediate.

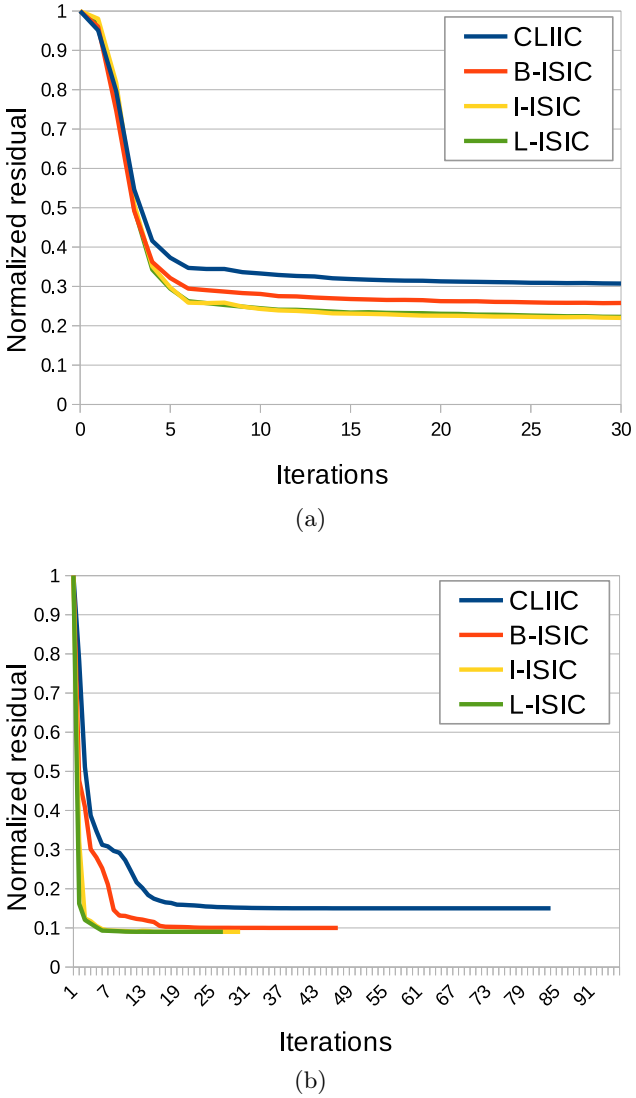

Figure 8: Convergence rate of LSRTM preconditioned with different imaging conditions. (a) Data domain; (b) image domain.

In the data domain, we stopped the inversion after 30 iterations. We see in Figure 8a that I-ISIC or L-ISIC preconditioned LSRTM reaches a residual value at a little more than two thirds of the value obtained with the CLIIC. Even after just 5 iterations, the residual obtained with I-ISIC or L-ISIC preconditioning is already smaller than that using the CLIIC after 30 iterations. In the image domain, the advantage in convergence of the ISIC versions over the CLIIC is even greater. Both I-ISIC and L-ISIC reach the convergence criterion at a residual value of 0.95 after 31 and 28 iterations, respectively, dropping down to below 0.1 in only 7 iterations. B-ISIC takes 48 iterations to converge to 0.1,reaching this level after about 20 iterations, and CLIIC converges only after 85 iterations at a residual of 0.15, reaching this level at about 35 iterations.

When we recall that the CLIIC and L-ISIC are one-term imaging conditions, thus being computationally less expensive than the other two, it becomes clear that L-ISIC preconditioning is the most favorable choice in both domains.

DISCUSSION

We have studied the effect of preconditioning least-squares reverse-time migration (LSRTM) by means of a computationally efficient implementation of the inverse-scattering imaging condition (ISIC) that uses the Laplacian operator (L-ISIC) as recently derived by Albano et al. (2023). The ISIC was derived by Op't Root et al. (2012) using microlocal analysis as the asymptotic inverse to Born scattering. In its original form, it consists of two contributions that are summed to form the final image. It aims at providing amplitudes that are proportional to the medium reflectivity. The implementational form of Albano et al. (2023) is computationally superior to the two-term version (Op't Root et al., 2012), both in terms of quality and cost. When implemented in its basic form (which is slightly different from the correct theoretical expression), additional weights may be needed to balance the two contributions (see, e.g., Whitmore and Crawley, 2012).

The computational cost of the L-ISIC is the same as that of the conventional crosscorrelation imaging condition (Claerbout, 1971) with illumination compensation and Laplacian filtering (CLIIC), but lower than that of the basic form (B-ISIC) and of another two-term version that is theoretically correct (I-ISIC). When applying these imaging conditions in a single RTM, L-ISIC provides clearly superior images (Albano et al., 2023, see also Figure 4). Thus, the natural question arises whether it can be beneficial as a preconditioner for LSRTM. In this work, we have demonstrated that this is indeed the case.

While the differences in image quality after least-squares optimization are strongly reduced, the L-ISIC LSRTM images are still of slightly superior quality than those obtained with the other tested imaging conditions.

CONCLUSIONS

In this work, we have shown that LSRTM can benefit from L-ISIC preconditioning. We tested L-ISIC preconditioning in both the data and image domains. Our numerical tests demonstrated that the final images after L-ISIC preconditioned DD-LSRTM are of comparable quality or even slightly superior than those obtained with standard CLIIC preconditioning or with two other versions of the ISIC. In the image domain, the quality gains are more prominent, with L-ISIC preconditioned ID-LSRTM presenting higher resolution and less migration artifacts. However, the principal advantage of using L-ISIC preconditioning is the notably accelerated convergence as compared to CLIIC preconditioning.

In the data domain, ISIC preconditioned DD-LSRTM achieved convergence after much less iterations than required with the CLIIC. In this domain, each iteration of the optimization algorithm depends on at least two modeling steps (direct and reverse) to calculate the objective-function gradient. Therefore, the L-ISIC preconditioning considerably reduced the overall computational cost of LSRTM over using the conventional CLIIC. While I-ISIC achieves the same convergence rates as L-ISIC, it requires the computation of two contributions. Therefore, each iteration comes at a somewhat higher computational cost than in L-ISIC and CLIIC.

In the implementation in the image domain, the number of iterations of the inversion algorithm of ID-LSRTM with L-ISIC is even stronger reduced over the CLIIC than in the data domain, requiring only about one third of the iterations. Although the solution of the linear system is not the computationally heaviest part of ID-LRSTM, it still results in a considerable reduction of the computational cost.

When comparing the final images obtained from DD-LSRTM and ID-LSRTM with each other, we conclude that the image-domain approach yields the better results. The reflectors are clearly better defined, particularly in regions with low illumination, as for example in the deeper parts of the images.

ACKNOWLEDGMENTS

The authors are grateful for the support of Petrobras as well as the Brazilian agencies CNPq (supporting the INCT-GP) and CAPES (student scholarships).

APPENDIX A

LINEARIZATION OF THE ACOUSTIC WAVE EQUATION

In this paper, we consider the wave propagation to be described by the acoustic wave equation for constant density,

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2} - \nabla^2 p(t,\mathbf{x}) = s(t,\mathbf{x}), \tag{A1}$$

where $c(\mathbf{x})$ is the wave-propagation velocity, $p(t,\mathbf{x})$ is the acoustic pressure wavefield and $s(t,\mathbf{x})$ represents an impulsive point source. As mentioned in the main text, because of its nonlinear relation between the wavefield and the medium velocity, equation A1 cannot be used directly in LSRTM.

For the purpose of linearization, we assume that the true velocity distribution $c(\mathbf{x})$ can be represented as a (small) perturbation of a background velocity distribution $c_0(\mathbf{x})$ (see, e.g., Snieder, 2004), i.e.,

$$c(\mathbf{x}) = c_0(\mathbf{x}) + \delta c(\mathbf{x}), \qquad \delta c(\mathbf{x}) \ll c_0(\mathbf{x})\ , \tag{A2}$$

and that the wavefield $p_0(t,\mathbf{x})$, generated by the same source $s(t,\mathbf{x})$ in a medium with the background velocity, satisfies the corresponding wave equation

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 p_0(t,\mathbf{x})}{\partial t^2} - \nabla^2 p_0(t,\mathbf{x}) = s(t,\mathbf{x}). \tag{A3}$$

The difference of equations A1 and A3 yields

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2} - \frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 p_0(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = 0, \tag{A4}$$

where Δp denotes the difference between the two wavefields, i.e., the solution of equation A1 can be written as

$$p(t,\mathbf{x}) = p_0(t,\mathbf{x}) + \Delta p(t,\mathbf{x}). \tag{A5}$$

Adding and subtracting the term $\frac{1}{c_0^2}\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2}$ in equation A4, it can be recast into the form

$$\left[\frac{1}{c^2(\mathbf{x})} - \frac{1}{c_0^2(\mathbf{x})}\right]\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2} + \frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = 0, \tag{A6}$$

or

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = -\left[\frac{1}{c^2(\mathbf{x})} - \frac{1}{c_0^2(\mathbf{x})}\right]\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2}\ , \tag{A7}$$

which demonstrates that the wavefield difference Δp can be represented as the solution of the wave equation in the background medium, equation A1, with the source term $s(\mathbf{x},t)$ replaced by the secondary sources $s_s(\mathbf{x},t) = -\left[\frac{1}{c^2(\mathbf{x})} - \frac{1}{c_0^2(\mathbf{x})}\right]\frac{\partial^2 p(t,\mathbf{x})}{\partial t^2}$.

Note that equation A7 is exact, but its solution is not actually possible, because p on the right-hand side of equation A7 depends on Δp. The Born approximation assumes small velocity and wavefield perturbations, i.e., $c(\mathbf{x}) - c_0(\mathbf{x}) = \delta c(\mathbf{x}) \ll c_0(\mathbf{x})$ and $\Delta p(t,\mathbf{x}) \ll p_0(t,\mathbf{x})$. In a Taylor expansion up to first order in δc, we can write

$$\frac{1}{c^2(\mathbf{x})} = \frac{1}{[c_0(\mathbf{x}) + \delta c(\mathbf{x})]^2} = \frac{1}{c_0^2(\mathbf{x})} - \frac{2\delta c(\mathbf{x})}{c_0^3(\mathbf{x})} + O(\delta c^2(x)). \tag{A8}$$

Using this, equation A7 becomes

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = -\left[-\frac{2\delta c(\mathbf{x})}{c_0^3(\mathbf{x})} + O(\delta c^2(x))\right]\left(\frac{\partial^2 p_0(t,\mathbf{x})}{\partial t^2} + \frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2}\right)\ . \tag{A9}$$

Under the specific assumption that the second time derivative of the wavefield difference is $O(\delta c(\mathbf{x}))$, the linear approximation of equation A9, i.e., up to first order in $\delta c(\mathbf{x})$, can be represented as

$$\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2 \Delta p(t,\mathbf{x})}{\partial t^2} - \nabla^2 \Delta p(t,\mathbf{x}) = \frac{m(\mathbf{x})}{c_0^2(\mathbf{x})}\frac{\partial^2 p_0(t,\mathbf{x})}{\partial t^2}, \tag{A10}$$

where

$$m(\mathbf{x}) = \frac{2\delta c(\mathbf{x})}{c_0(\mathbf{x})} \tag{A11}$$

denotes the medium reflectivity. Note that this approximation considers only secondary sources that are excited by the background wavefield and neglects those excited by the wavefield difference, which represents the scattered wavefield. In other words, the Born approximation is a single-scattering approximation.

In LSRTM, one aims at improving an estimate for $m(\mathbf{x})$ without changing the background velocity model $c_0(\mathbf{x})$. Therefore, equation A3 for p_0 needs to be solved only once.

REFERENCES

Albano, N. A., J. C. Costa, J. Schleicher, and A. Novais, 2023, On the estimation of reflectivity in reverse time migration: Implementational forms of the inverse-scattering imaging condition: Geophysics, **88**, S1–S16, doi: 10.1190/geo2021-0698.1.

Aoki, N., and G. T. Schuster, 2009, Fast least-squares migration with a deblurring filter: Geophysics, **74**, WCA83–WCA93, doi: 10.1190/1.3155162.

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

Biondi, B. L., 2006, 3D seismic imaging. Investigations in Geophysics Series: SEG. (247 pp, doi: 10.1190/1.9781560801689).

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

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion, **6**: Blackwell Scientific Publications, Boston, USA. (Available at: https://sepwww.stanford.edu/sep/prof/pvi.pdf).

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695, doi: 10.1111/j.1365-2478.2012.01092.x.

Dai, W., and J. Schuster, 2009, Least-squares migration of simultaneous sources data with a deblurring filter, *in* SEG Technical Program Expanded Abstracts 2009: Society of Exploration Geophysicists, 2990–2994. doi: 10.1190/1.3255474.

Duan, Y., A. Guitton, and P. Sava, 2017, Elastic least-squares reverse time migration: Geophysics, **82**, S315–S325, doi: 10.1190/geo2016-0564.1.

Duveneck, E., M. Kiehn, A. Chandran, and T. Kühnel, 2021, Reflection angle/azimuth-dependent least-squares reverse time migration: Geophysics, **86**, S325–S338, doi: 10.1190/geo2020-0701.1.

Fang, X., D. Wu, and F. Niu, 2017, LSRTM enhanced with inverse scattering imaging condition: SEG 2017 Workshop: Full-waveform Inversion and Beyond, Beijing, China, Society of Exploration Geophysicists, 107–110. doi: 10.1190/FWI2017-028.

Feng, Z., and G. T. Schuster, 2017, Elastic least-squares reverse time migration: Geophysics, **82**, S143–S157, doi: 10.1190/geo2016-0254.1.

Fletcher, R. F., P. Fowler, P. Kitchenside, and U. Albertin, 2005, Suppressing artifacts in prestack reverse time migration, *in* SEG Technical Program Expanded Abstracts 2005: Society of Exploration Geophysicists, 2049–2051. doi: 10.1190/1.2148113.

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.

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

Guo, S., and H. Wang, 2020, Image domain least-squares migration with a Hessian matrix estimated by non-stationary matching filters: Journal of Geophysics and Engineering, **17**, 148–159, doi: 10.1093/jge/gxz098.

Jiang, B., and J. Zhang, 2019, Least-squares migration with a blockwise Hessian matrix: A prestack time-migration approach: Geophysics, **84**, R625–R640, doi: 10.1190/geo2018-0533.1.

Kim, S., W. Chung, and S. Shin, 2019, Acoustic full-waveform inversion using Adam optimizer: Geophysics and Geophysical Exploration, **22**, 202–209, doi: 10.7582/GGE.2019.22.4.202.

Kingma, D. P., and J. Ba, 2014, Adam: A method for stochastic optimization: arXiv:1412.6980, doi: 10.48550/arXiv.1412.6980.

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.

Leveille, J. P., I. F. Jones, Z.-Z. Zhou, B. Wang, and F. Liu, 2011, Subsalt imaging for exploration, production, and development: A review: Geophysics, **76**, WB3–WB20, doi: 10.1190/geo2011-0156.1.

Liu, D. C., and J. Nocedal, 1989, On the limited memory BFGS method for large scale optimization: Mathematical Programming, **45**, 503–528, doi: 10.1007/BF01589116.

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.

Menke, W., 1989, Geophysical data analysis: Discrete inverse theory: Academic Press.

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.

Op't Root, T. J. P. M., C. C. Stolk, and M. V. de Hoop, 2012, Linearized inverse scattering based on seismic reverse time migration: Journal de Mathématiques Pures et Appliquées, **98**, 211–238, doi: 10.1016/j.matpur.2012.02.009.

Osorio, L. N., B. Pereira-Dias, A. Bulcão, and L. Landau, 2021, Migration deconvolution using domain decomposition: Geophysics, **86**, S247–S256, doi: 10.1190/geo2020-0352.1.

Pestana, R. C., A. W. G. dos Santos, and E. S. Araujo, 2014, RTM imaging condition using impedance sensitivity kernel combined with Poynting vector: SEG Technical Program Expanded Abstracts, SEG, 3763–3768. doi: 10.1190/segam2014-0374.1.

Prucha, M. L., and B. L. Biondi, 2002, Subsalt event regularization with steering filters, *in* SEG Technical Program Expanded Abstracts 2002: Society of Exploration Geophysicists, 1176–1179. doi: 10.1190/1.1816859.

Schultz, P. S., and J. W. C. Sherwood, 1980, Depth migration before stack: Geophysics, **45**, 376–393, doi: 10.1190/1.1441088.

Schuster, G. T., 2017, Seismic inversion. Investigations in Geophysics Series, No. 20: Society of Exploration Geophysicists. (376 pp, doi: 10.1190/1.9781560803423).

Schuster, G. T., and Z. Liu, 2019, Least squares migration: Current and future directions: 81st EAGE Conference and Exhibition 2019, European Association of Geoscientists & Engineers, 1–5. doi: 10.3997/2214-4609.201901270.

Snieder, R., 2004, A Guided Tour of Mathematical Methods for the Physical Sciences, 2nd ed.: Cambridge University Press. (doi: 10.1017/CBO9781139013543).

Snieder, R., and J. Trampert, 1999, Inverse problems in Geophysics, *in* Wirgin, A. (ed.), Wavefield Inversion. International Centre for Mechanical Sciences, v. 398: Springer, 119–190. doi: 10.1007/978-3-7091-2486-4_3.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Trad, D., 2015, Least squares Kirchhoff depth migration: Implementation, challenges, and opportunities, *in* SEG Technical Program Expanded Abstracts 2015, Proceedings: Society of Exploration Geophysicists, 4238–4242. doi: 10.1190/segam2015-5833655.1.

Valenciano, A. A., 2008, Imaging by wave-equation inversion: PhD thesis, Stanford University. (118 pp, Available at: https://sep.sites.stanford.edu/publications/theses/imaging-wave-equation-inversion-sep-133-2008).

Valenciano, A. A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35–A38, doi: 10.1190/1.2213359.

Whitmore, N. D., 2013, A generalization of RTM inverse scattering imaging conditions: 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013, European Association of Geoscientists & Engineers, cp–348–01020. doi: 10.3997/2214-4609.20130477.

Whitmore, N. D., and S. Crawley, 2012, Applications of RTM inverse scattering imaging conditions, *in* SEG Technical Program Expanded Abstracts 2012: Society of Exploration Geophysicists, 1–6. doi: 10.1190/segam2012-0779.1.

Yang, J., Y. Elita Li, A. Cheng, Y. Liu, and L. Dong, 2019, Least-squares reverse time migration in the presence of velocity errors: Geophysics, **84**, S567–S580, doi: 10.1190/geo2019-0005.1.

Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: Geophysics, **80**, V23–V31, doi: 10.1190/geo2013-0461.1.

Albano, N.A.: code development, theoretical discussions, writing – original draft preparation; **Costa, J.C.**: code development, theoretical discussions, writing – original draft preparation; **Schleicher, J.**: code suggestions, theoretical discussions, writing – original draft preparation; **Novais, A.**: code suggestions, theoretical discussions, writing – original draft preparation.

Received on September 5, 2023 / Accepted on January 10, 2025