Received: 6 March 2022 | Accepted: 18 January 2023

DOI: 10.1111/1365-2478.13323

Geophysical Prospecting | EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

ORIGINAL ARTICLE

A matrix-free variant of the distorted Born iterative method for seismic full-waveform inversion

Kui Xiang[1,2] | **Morten Jakobsen**[1] | **Kjersti Solberg Eikrem**[2] | **Geir Nævdal**[2]

[1]Department of Earth Science, University of Bergen, Bergen, Norway

[2]NORCE Norwegian Research Centre AS, Bergen, Norway

Correspondence
Kui Xiang, Department of Earth Science, University of Bergen, Postboks 7803, 5020 Bergen, Norway.
Email: Kui.Xiang@uib.no

Funding information
China Scholarship Council; Research Council of Norway, Grant/Award Number: 267769

[Correction added on 11 September 2023, after first online publication: The copyright line was changed.]

Abstract

The distorted Born iterative method reduces a nonlinear inverse scattering problem to a sequence of (ill-posed) linear inverse scattering problems that can be solved using a regularized least-squares formulation. This method was originally applied to two-dimensional electromagnetic problems but has been implemented to solve acoustic and electromagnetic problems in three dimensions. It has also been applied to seismic problems but only for moderately large two-dimensional models. Previous applications of the distorted Born iterative method to seismic inverse scattering were based on a matrix representation of the relevant integral operators. The matrix-based implementation is simple and transparent but not very suitable for large-scale computations since the memory requirement and computational cost scales like N^2 and N^3, where N is the number of grid blocks. In this paper, we introduce a matrix-free variant of the distorted Born iterative method, which is much more suitable for large-scale problems, since the memory requirements and computational cost have been reduced to N and $N\log(N)$, respectively. Our matrix-free implementation utilizes a fast-Fourier-transform-accelerated iterative method to solve the linear system that results after discretization. In the computation of Fréchet derivatives, we avoid the direct computation of Green's functions for heterogeneous media, by solving an equivalent direct scattering problem. The adjoint of the Fréchet derivative operator is also computed in an efficient matrix-free manner, by exploiting the physical interpretation and reciprocity of the Green's function. We illustrate the potential of the matrix-free variant of the distorted Born iterative method using synthetic waveform data for large two- and three-dimensional models. The main conclusion is that the distorted Born iterative method has been transformed into a more practical tool for seismic as well as electromagnetic and acoustic nonlinear inverse scattering.

KEYWORDS
acoustics, full waveform, inverse problem, mathematical formulation, seismics

INTRODUCTION

In exploration seismology, full waveform inversion (FWI) has emerged as the ultimate method for reconstruction of the subsurface properties using all the kinematic and dynamic information contained in the observed waveform data (Virieux & Operto, 2009). From a mathematical perspective, FWI can be regarded as an optimization problem where the goal is to minimize a data misfit function that measures the difference between the calculated and observed waveform data (Taran-

tola, 1984). It can be implemented in both the time (Tarantola, 1984) and frequency domains (Pratt et al., 1998). This paper only discusses a particular scattering-based implementation in the frequency domain. From a physical perspective, FWI can be regarded as an inverse scattering problem, where the goal is to reconstruct the properties of a scattering domain from observations performed outside this domain. This inverse scattering problem is nonlinear because the relationship between the observed waveforms and the unknown scattering potential is nonlinear. In both pure and applied sciences, various methods exist for solving such nonlinear inverse scattering problems (Pike & Sabatier, 2002). In this study, we focus on the so-called distorted Born iterative (DBI) method.

The DBI method was originally proposed to solve the inverse electromagnetic scattering problem for a two-dimensional (2D) reconstruction of permittivity (Chew & Wang, 1990; Wang & Chew, 1989). It reduces the nonlinear problem to a series of linear problems based on the distorted Born approximation. There are many implementations of the DBI method in microwave imaging and ultrasound tomography (Belkebir et al., 1997; Chandra et al., 2015; Huthwaite & Simonetti, 2011; Hopfer et al., 2017; Souvorov et al., 1998; Semenov et al., 2005; Semenov & Corfield, 2008). In seismic exploration, people are more familiar with the scattering-integral approach (Chen et al., 2007; Djebbi & Alkhalifah, 2020; Tao & Sen, 2013). Essentially, the scattering-integral approach can be viewed as a DBI method since they are both derived from the scattering theory and they both use the Green's function and wavefield in the background medium to form the sensitivity kernel. Recently, the DBI method and/or the scattering-integral approach have been widely investigated in seismic FWI. Tao and Sen (2013) introduced the scattering-integral approach within frequency-domain FWI and compared it with the conventional FWI. Jakobsen and Ursin (2015) used a T-matrix method of the multiple scattering theory to perform nonlinear seismic waveform inversion. Alkhalifah and Wu (2016) proposed a Born series multiple scattering approach and gave insights into the importance of multiple scattering in Alkhalifah et al. (2018). Liu et al. (2015) developed a matrix-free Born kernel FWI scheme by accumulating decomposed vector–scalar products based on the scattering-integral approach. Yang et al. (2016) used the scattering-integral approach to invert velocity and density simultaneously. Malovichko et al. (2017) applied the DBI method with the quasi-analytical approximation to three-dimensional FWI. Eikrem et al. (2019) combined the DBI method with an iterated extended Kalman filter to perform the Bayesian inversion of seismic data. Huang et al. (2019) used the DBI method to achieve a target-oriented inversion for estimating the reservoir changes with the time-lapse seismic data. Huang et al. (2020) estimated the uncertainties of five parameters for Bayesian anisotropic elastic FWI using the DBI method with the iterated extended Kalman filter.

All those developments demonstrate that the DBI method can accurately estimate the underground physical properties; however, until recently, its practical implementations were uncommon due to its substantial computational cost. So there is still a great demand to further improve the efficiency of the DBI method. Therefore, in this paper, we develop an efficient FWI scheme based on a variant of the DBI method with the adjoint method (Hesford & Chew, 2010; Tarantola, 1984; Tao & Sen, 2013) and the fast forward solver (Eikrem et al., 2021; Osnabrugge et al., 2016). Remis and Van den Berg (2000) and Oristaglio and Blok (2012) have demonstrated that the DBI method is consistent with the Gauss–Newton method in optimization theory. Therefore, the approximate Hessian matrix and its inverse are also needed when solving the normal equation at each iteration of traditional DBI methods. To avoid saving and inverting the huge Hessian matrix, we use the conjugate gradient method with the Fréchet operator and its adjoint to replace the Hessian and its inverse. However, the matrix–vector multiplication at each conjugate gradient iteration is still very costly and the background Green's function and wavefield need to be updated and stored for the construction of the Fréchet operator and its adjoint after the update of the background model. Inspired by Hesford and Chew (2006, 2010), we found that the physical meaning of the Fréchet operator and its adjoint represent forms of wavefields. Based on that, we use the adjoint-state method to find the formulation of the adjoint operator. Further, we rewrite the Fréchet and adjoint operators in new expressions that only include wavefield operations. The new wavefields can be obtained via the fast forward solver with fastFourier transform. In the new variant of the DBI method, only the Green's function for a homogeneous background medium needs to be saved and it is saved in a compact format due to its Toeplitz structure.

This paper is organized as follows: We first review the traditional DBI method; then, we introduce the Fréchet operator, adjoint operator and normal equation; further, we present a matrix-free implementation of the Fréchet and adjoint operator; finally, we provide numerical examples to test the validity of the proposed method.

METHODOLOGY

The distorted Born iterative method

In the frequency domain, the propagation of seismic waves in the acoustic approximation is governed by the Helmholtz equation (Červený, 2001; Jakobsen & Ursin, 2015; Morse & Feshbach, 1954)

$$\left[\nabla^2 + \frac{\omega^2}{v^2(\mathbf{x})}\right] p(\mathbf{x},\omega) = -S(\mathbf{x},\omega), \tag{1}$$

where $p(\mathbf{x},\omega)$ is the wavefield (pressure in acoustic medium) and ω is the angular frequency. $v(\mathbf{x})$ and $S(\mathbf{x},\omega)$ are the velocity and the source, respectively.

Considering a background medium $v_b(\mathbf{x})$, the wavefield within it, $p_b(\mathbf{x},\omega)$, satisfy

$$\left[\nabla^2+\frac{\omega^2}{v_b^2(\mathbf{x})}\right]p_b(\mathbf{x},\omega)=-S(\mathbf{x},\omega), \tag{2}$$

By adding $[\omega^2/v_b^2(\mathbf{x})]p(\mathbf{x},\omega)$ on both sides of (1) and then subtracting (2), we have

$$\left[\nabla^2+\frac{\omega^2}{v_b^2(\mathbf{x})}\right]\left[p(\mathbf{x},\omega)-p_b(\mathbf{x},\omega)\right]=-\left[\frac{\omega^2}{v^2(\mathbf{x})}-\frac{\omega^2}{v_b^2(\mathbf{x})}\right]p(\mathbf{x},\omega). \tag{3}$$

Equation (3) can be expressed in a volume-integral form as

$$\delta p(\mathbf{x},\omega)=\omega^2\int_\Omega d\mathbf{x}' G_b(\mathbf{x},\mathbf{x}',\omega)\delta m(\mathbf{x}')p(\mathbf{x}',\omega), \tag{4}$$

where $\delta p(\mathbf{x},\omega)=p(\mathbf{x},\omega)-p_b(\mathbf{x},\omega)$ is the wavefield perturbation, Ω is the scattering domain where the scattering potential is non-zero and $\delta m(\mathbf{x})$ is the model-parameter perturbation that is defined as

$$\delta m(\mathbf{x})=\frac{1}{v^2(\mathbf{x})}-\frac{1}{v_b^2(\mathbf{x})}. \tag{5}$$

The Green's function $G_b(\mathbf{x},\mathbf{x}',\omega)$ is the solution to

$$\left[\nabla^2+\frac{\omega^2}{v_b^2(\mathbf{x})}\right]G_b(\mathbf{x},\mathbf{x}',\omega)=-\delta\left(\mathbf{x}-\mathbf{x}'\right), \tag{6}$$

where $\delta(\mathbf{x}-\mathbf{x}')$ is the Dirac delta function that represents a unit impulse. By comparing (2) and (6), we can find that the physical meaning of Green's function is the wavefield generated by the unit impulse. So the total wavefield is the superposition of the wavefields generated by all unit impulses. If the background medium is smooth, the Green's function can be constructed by ray theory (Červenỳ, 2001). In an arbitrary background medium, the Green's function can be numerically solved by using the finite difference method (Kirchner & Shapiro, 2001) or the integral equation method (Jakobsen & Ursin, 2015). More details about the Green's function can be found in Arfken and Weber (1999). For the simplicity of notation, we suppress the dependency on ω for the wavefield p, the source term S and the Green's function G in the following formulas.

In practical implementations, the wavefield from the real model is only recorded at the receiver positions. By assuming the real model is close to the background model, we can approximate the wavefield within the real model by the wavefield within the background model. Based on the above analysis, we rewrite (4) as

$$\delta p(\mathbf{r})\approx\omega^2\int_\Omega d\mathbf{x}' G_b(\mathbf{r},\mathbf{x}')\delta m(\mathbf{x}')p_b(\mathbf{x}'),\quad \mathbf{r}\in D, \tag{7}$$

where $\delta p(\mathbf{r})=p(\mathbf{r})-p_b(\mathbf{r})$ is the wavefield perturbation at receiver position and D is the observed data domain. Equation (7) gives a linear relation between the wavefield perturbation and model perturbation to solve δm from δp. In the distorted Born iterative (DBI) method, we first assume a known background model and calculate the background Green's function G_b and background wavefield p_b. Then we can get δp and solve δm from δp. After we have δm, we can update the background model and use the updated model as the new background model to calculate the new δm until the background model is close enough to the real model.

The normal equation

Equation (7) can be expressed in an operator form as

$$\mathcal{F}\delta\mathbf{m}=\delta\mathbf{p}, \tag{8}$$

where $\mathcal{F}$ is the Fréchet operator, which is represented by

$$[\mathcal{F}X](\mathbf{r})=\omega^2\int_\Omega d\mathbf{x}' G_b(\mathbf{r},\mathbf{x}')X(\mathbf{x}')p_b(\mathbf{x}'). \tag{9}$$

In most cases of seismic exploration, determining $\delta\mathbf{m}$ from (8) is ill-posed. To solve this problem, we adopt the generalized Tikhonov method (Menke, 2012), which yields a closed-form solution to (8) by minimizing the objective function

$$\mathcal{E}(\delta\mathbf{m})=\|\delta\mathbf{p}-\mathcal{F}\delta\mathbf{m}\|_2^2+\lambda^2\|\delta\mathbf{m}\|_2^2, \tag{10}$$

where $\|\cdot\|_2^2$ is the L_2-norm and λ is the regularization parameter that balances the influence of the fidelity term and the regularization term. There are various strategies for selecting the regularization parameter (Constable et al., 1987; Farquharson & Oldenburg, 2004; Hansen, 1998). It can also be determined by a self-adaptive cooling scheme, in which an initial regularization parameter is established at the start of the iteration process and subsequently decreased as the number of iterations increases (Jakobsen & Ursin, 2015). Minimizing (10) is equivalent to solving the regularized normal equation (Virieux & Operto, 2009)

$$(\mathcal{F}^\dagger\mathcal{F}+\lambda^2)\delta\mathbf{m}=\mathcal{F}^\dagger\delta\mathbf{p}, \tag{11}$$

where $\mathcal{F}^\dagger$ is the adjoint of the Fréchet operator, which can be obtained through the inner product rule of the operator and its

adjoint as

$$[\mathcal{F}^{\dagger}Y](\mathbf{x}) = \left\{ \omega^2 p_b(\mathbf{x}) \left[\int_D d\mathbf{r} G_b(\mathbf{x},\mathbf{r}) Y^*(\mathbf{r}) \right] \right\}^*, \quad (12)$$

where $\{\cdot\}^*$ denotes the complex conjugation. More details about the derivation of the adjoint operator can be found in Appendix A.

Matrix-free implementation of the Fréchet and adjoint operators

In the traditional DBI method, Equation (11) is solved by matrix inversion (Huang et al., 2020; Jakobsen & Ursin, 2015). However, the direct solver requires a lot of computational time and memory. Therefore, we solve for the model update $\delta\mathbf{m}$ via the conjugate gradient method in conjunction with the matrix-free implementation of the Fréchet and adjoint operators. By setting $J(\mathbf{x}') = X(\mathbf{x}')p_b(\mathbf{x}')$, the Fréchet operator can be rewritten as

$$[\mathcal{F}X](\mathbf{r}) = \omega^2 \int_\Omega d\mathbf{x}' G_b(\mathbf{r},\mathbf{x}') J(\mathbf{x}'). \quad (13)$$

The integral part of Equation (13) can be interpreted as the observed wavefield from a heterogeneous background medium due to the source J (Hesford and Chew, 2010), so that Equation (13) can be rewritten as

$$[\mathcal{F}X](\mathbf{r}) = \omega^2 p_J(\mathbf{r}). \quad (14)$$

In Equation (14), $p_J(\mathbf{r})$ can be solved from the data equation (Jakobsen & Ursin, 2015)

$$p_J(\mathbf{r}) = p_J^{(0)}(\mathbf{r}) + \omega^2 \int_\Omega d\mathbf{x} G^{(0)}(\mathbf{r},\mathbf{x}') \chi_b(\mathbf{x}') p_J(\mathbf{x}'). \quad (15)$$

Here, $p_J(\mathbf{x}')$ is the wavefield within the background medium, which can be solved from

$$p_J(\mathbf{x}) = p_J^{(0)}(\mathbf{x}) + \omega^2 \int_\Omega d\mathbf{x} G^{(0)}(\mathbf{x},\mathbf{x}') \chi_b(\mathbf{x}') p_J(\mathbf{x}'), \quad (16)$$

where

$$p_J^{(0)}(\mathbf{x}) = \int_\Omega d\mathbf{x}' G^{(0)}(\mathbf{x},\mathbf{x}') J(\mathbf{x}') \quad (17)$$

is the incident wavefield caused by the source J,

$$\chi_b(\mathbf{x}') = \frac{1}{v_b^2(\mathbf{x}')} - \frac{1}{v_0^2} \quad (18)$$

is the contrast between the heterogeneous background medium v_b and the homogeneous medium v_0, and $G^{(0)}(\mathbf{x},\mathbf{x}')$ is the Green's function in the homogeneous medium. The analytic formulas of the homogeneous Green's function are found in Červený (2001) as

$$G^{(0)}(\mathbf{x},\mathbf{x}') = \begin{cases} \frac{1}{4} i H_0^{(1)}(k_0 r), & \text{in } 2D, \\ \frac{1}{4\pi r} \exp(i k_0 r), & \text{in } 3D, \end{cases} \quad (19)$$

where $r = \|\mathbf{x} - \mathbf{x}'\|$ and $H_0^{(1)}$ is the Hankel function of the first kind and zeroth order.

In the context of the adjoint operator, the part in square brackets of Equation (12) can be interpreted as the wavefield back-propagating from the receiver to the scattering domain, as per the physical interpretation of the Green's function $G_b(\mathbf{x},\mathbf{r})$. So we rewrite (12) as

$$[\mathcal{F}^{\dagger}Y](\mathbf{x}) = \left[\omega^2 p_b(\mathbf{x}) p_Y(\mathbf{x})\right]^*, \quad (20)$$

where

$$p_Y(\mathbf{x}) = \int_D d\mathbf{r} G_b(\mathbf{x},\mathbf{r}) Y^*(\mathbf{r}) \quad (21)$$

is the back-propagating wavefield. Here, $p_Y(\mathbf{x})$ can be solved from

$$p_Y(\mathbf{x}) = p_Y^{(0)}(\mathbf{x}) + \omega^2 \int_\Omega d\mathbf{x}' G^{(0)}(\mathbf{x},\mathbf{x}') \chi_b(\mathbf{x}') p_Y(\mathbf{x}'), \quad (22)$$

where

$$p_Y^{(0)}(\mathbf{x}) = \int_D d\mathbf{r} G^{(0)}(\mathbf{x},\mathbf{r}) Y^*(\mathbf{r}). \quad (23)$$

In the new expressions of the Fréchet derivative (14) and its adjoint operator (20), only three vectors p_J, p_b and p_Y remain. Their related equations, (15), (16) and (22), are equations of the Lippmann–Schwinger type, which can be easily solved by iterative solvers, such as the preconditioned Born series (Eikrem et al., 2021; Jakobsen et al., 2020), the conjugate gradient method (Hestenes & Stiefel, 1952) and so on. Because the homogeneous Green's function $G^{(0)}$ depends on $\mathbf{x} - \mathbf{x}'$, the integrals that include it can be calculated quickly and accurately by utilizing the fast Fourier transform. This is crucial for speeding up the overall inversion method.

In the conventional DBI method, Equation (11) is transformed into matrix form after discretization and is then solved using matrix inversion. The computational complexity and storage requirement of the matrix inversion method scale like N^3 and N^2, respectively, where N is the number of grid blocks to divide the computational domain. In our method, the scales of the computational cost and storage requirement are

ALGORITHM 1 Pseudo code for matrix-free distorted Born iterative (DBI) method

FDM: Forward modeling solver
CGM: Conjugate Gradient Method
Setup tolerance η, maximum iteration k, initial regularization λ_{ini} and decrease factor a
for $\omega = \omega_{\min}$ to $\omega_{\max}$ **do**
  $i = 0, \epsilon_d = 1, \lambda = \lambda_{ini}$
  $\mathbf{p} = \mathrm{FDM}(\mathbf{m}_{ini})$
  $\delta\mathbf{p} = \mathbf{p} - \mathbf{p}_{obs}$
  $\mathbf{g} = \mathcal{F}^{\dagger}\delta\mathbf{p}$
  while $\epsilon_d > \eta$ & $i < k$ **do**
    $i = i + 1$
    $\delta\mathbf{m} = \mathrm{CGM}(\mathcal{F}^{\dagger}, \mathcal{F}, \mathbf{g}, \lambda)$
    $\mathbf{m} = \mathbf{m}_{ini} + \delta\mathbf{m}$
    $\mathbf{p}' = \mathbf{p}, \epsilon_d' = \epsilon_d$
    $\mathbf{p} = \mathrm{FDM}(\mathbf{m})$
    $\epsilon_d = \|\mathbf{p} - \mathbf{p}_{obs}\| / \|\mathbf{p}_{obs}\|$
    if $\epsilon_d < \epsilon_d'$ **then**
      $\mathbf{m}_{ini} = \mathbf{m}, \lambda = a\lambda$
      $\delta\mathbf{p} = \mathbf{p} - \mathbf{p}_{obs}$
      $\mathbf{g} = \mathcal{F}^{\dagger}\delta\mathbf{p}$
    else
      $\mathbf{m} = \mathbf{m}_{ini}, \epsilon_d = \epsilon_d', \mathbf{p} = \mathbf{p}', \lambda = \lambda/a$
    end if
  end while
end for

$N\log(N)$ and N, respectively. Details about the matrix-free DBI method can be found in Algorithm 1.

NUMERICAL EXAMPLES

Our code is written in the MATLAB language, and all the numerical examples in this paper were run on a computer with an Intel i7-7700 CPU and 64 GB of RAM.

Conceptual reservoir model

In the first numerical example, we create a conceptual reservoir model to compare the performance of the matrix-free and conventional matrix-based distorted Born iterative (DBI) methods. Figure 1 shows the real and initial velocity models. The model has been discretized into 249 × 71 grids, each of area 20 m×20 m. The measurements are produced by 71 sources and 249 receivers that are evenly distributed on the surface of the model. For each source, we use a Ricker wavelet of 7.5 Hz central frequency and sample it with a sampling interval $\Delta t = 0.004$ s and a total record length $T = 3$ s. We employ the fast Fourier transform to transform the Ricker wavelet into the frequency domain to generate frequency component data. In this example, six frequency components have been used for inversion: 3, 7.5, 10, 12, 15 and 18 Hz. The preconditioned Born series (Jakobsen et al., 2020; Osnabrugge et al., 2016), with the hierarchical preconditioner (Eikrem et al., 2021), is our forward modelling solver. The rank and level for construction of the hierarchical preconditioner (Eikrem et al., 2021) at each frequency are 4, 4, 4, 4, 4, 4 and 20, 40, 50, 60, 70, 80, respectively. More details about the hierarchical preconditioner, the rank and level and the preconditioned Born series can be found in Appendix B.

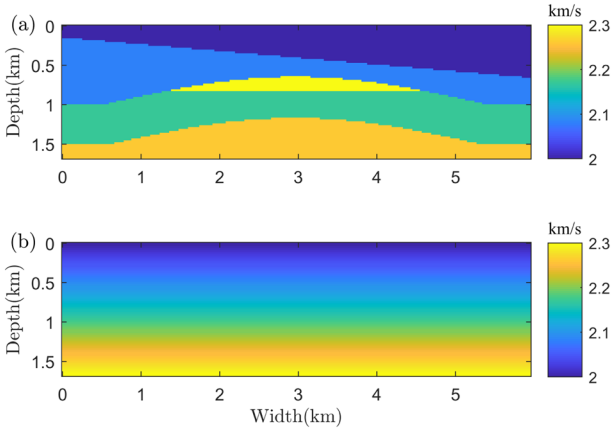

FIGURE 1 Real (a) and initial (b) velocity model used in the first numerical example.

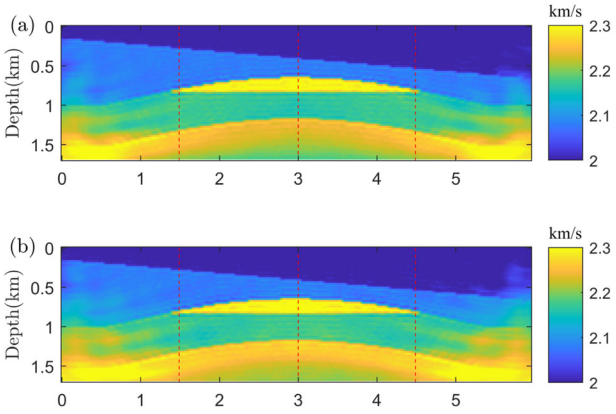

FIGURE 2 Inverted velocity using (a) matrix-based DBI and (b) matrix-free DBI.

We use the anti-reflection boundary layer (Osnabrugge et al., 2021) to reduce the artificial boundary reflection from the computational boundaries. For both the matrix-free DBI and matrix-based DBI, the tolerance η of normalized data error is set to 0.01 and the maximum number of iterations k of model updates is set to 10. The inner conjugate-gradient-method loop is stopped when the relative change in the residual falls below a tolerance of 0.1 or the number of iterations reaches 30. At the beginning of each frequency, we set the initial value of the regularization parameter and the decrease factor as $\lambda_{\mathrm{ini}} = 0.0005$ and $a = 0.1$, respectively.

The inverted velocity of the matrix-based DBI and matrix-free DBI is shown in Figure 2. The results of these two DBI methods are similar and acceptable. In order to better observe the details of the inverted model, we display three profiles of the inverted velocities as shown in Figure 3. In Figure 3, we can see that the differences between the profiles of the matrix-free DBI method and the matrix-based DBI method are minimal. In the upper and middle parts, the inverted profiles of these two methods are both close to the real model.

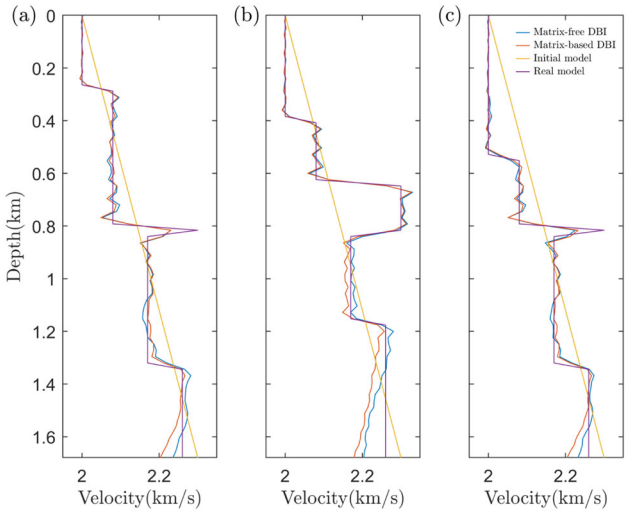

FIGURE 3 Profile comparison of different velocity models at (a) $x = 1.488$ km, (b) $x = 3$ km and (c) $x = 4.488$ km. The yellow line denotes the initial velocity model. The purple line denotes the real velocity model. The red line denotes the inverted velocity model using conventional matrix-based DBI. The blue line denotes the inverted velocity model using matrix-free DBI. These three locations are highlighted by the dashed red lines in Figure 2.

The lower parts are not well inverted. One explanation is that the bottoms lack illumination due to the absorbing boundary layer. To quantify the difference between the updated model m and the real model m_{true}, we compute the normalized model error ϵ_m defined as

$$\epsilon_m = \|\mathbf{m} - \mathbf{m}_{\text{true}}\| / \|\mathbf{m}_{\text{true}}\|. \tag{24}$$

Figure 4 shows how the normalized data errors and model errors change with the number of iterations for matrix-based and matrix-free DBI at different frequencies, in which it is visible that both methods work well and ϵ_m of the matrix-based DBI method converges to a smaller value than the matrix-free DBI method. That is because the solution of the normal equation obtained via the conjugate gradient method in matrix-free DBI is not as accurate as the solution obtained by the direct matrix inversion method in matrix-based DBI. However, we recognize that their differences are small. The computational time of the matrix-based DBI is 2.90×10^4 s, while the matrix-free DBI is 2.77×10^3 s.

Marmousi model

To further investigate the performance of the proposed matrix-free DBI method in a more realistic case, we use a resampled Marmousi2 model (Martin et al., 2006), as shown in Figure 5a, for our second numerical example. This model has a size of 7.6 km in the horizontal direction and 2.08 km

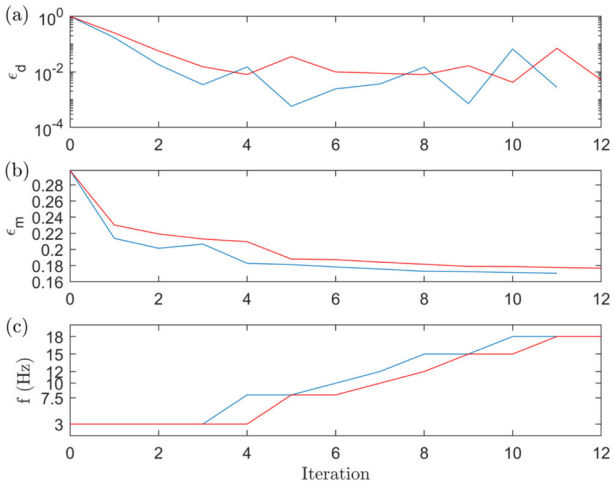

FIGURE 4 Convergence diagrams of the conventional matrix-based DBI and the matrix-free DBI at all frequencies: (a) normalized data error ϵ_d versus the number of iterations, (b) normalized model error ϵ_m versus the number of iterations and (c) the frequency corresponding to each iteration. The blue line denotes using matrix-based DBI. The red line denotes using matrix-free DBI.

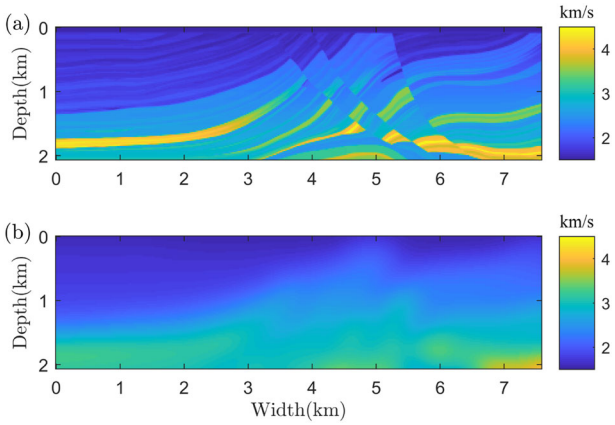

FIGURE 5 Marmousi model used in the second numerical example: (a) real velocity model and (b) initial velocity model.

in the vertical direction. We discretized it into 380×104 grid points. The size of each grid point is 20 m $\times$ 20 m. The initial model (Figure 5b) we used in this example is a smoothed version of the real model, which is generated by filtering the real model with a Gaussian smoothing filter with a standard deviation of 10. There are 104 sources and 380 receivers uniformly distributed along the surface to simulate the synthetic data. We use the same Ricker wavelet as in the first numerical example as the source. We choose six frequencies (3, 7.5, 10, 12, 15 and 18 Hz) to carry out this numerical example.

Because the memory demand of the matrix-based DBI method exceeds our computer's memory, only the matrix-free DBI method has been implemented here. The forward modelling solver is still the preconditioned Born series. The rank

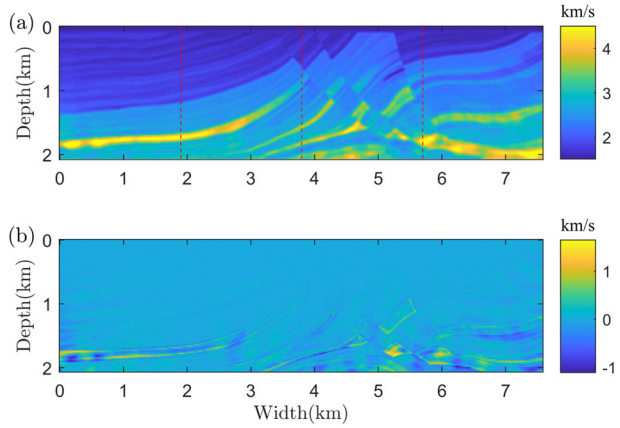

FIGURE 6 Marmousi model: (a) inverted velocity model. (b) The difference between the real model and the inverted model.

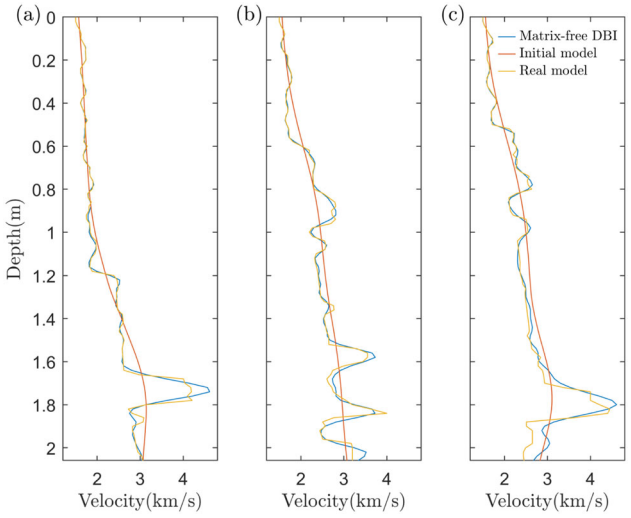

FIGURE 7 Profile comparison of different Marmousi velocity models at (a) $x = 1.9$ km, (b) $x = 3.8$ km and (c) $x = 5.7$ km. The red line denotes the initial velocity model. The yellow line denotes the real velocity model. The blue line denotes the inverted velocity model using matrix-free DBI. These three locations are highlighted by the dashed red lines in Figure 6a.

and level used to construct the preconditioner for different frequencies (from low to high) are 5, 5, 5, 5, 5, 4 and 20, 50, 60, 70, 80, 140, respectively. The initial regularization parameter and the decrease factor are set as $\lambda_{\text{ini}} = 0.001$ and $a = 0.1$, respectively. The maximum number of iterations and the residual tolerance of the conjugate gradient method are set as 20 and 0.1, respectively. The maximum number of iterations and the tolerance for the model update are set as $k = 10$ and $\eta = 0.01$, respectively. The inverted velocity in Figure 6 shows that most of the details of the Marmousi model have been well reconstructed. Figure 7 shows a profile comparison of different velocities at different locations. We can see that the model inverted by the matrix-free method approximates the real model closely. The convergence diagrams of the normalized data error and model error with iteration number at all frequencies are plotted in Figure 8, which shows that the matrix-free DBI method also works well for the complicated model. The computational time of this example is 3.57×10^4 s.

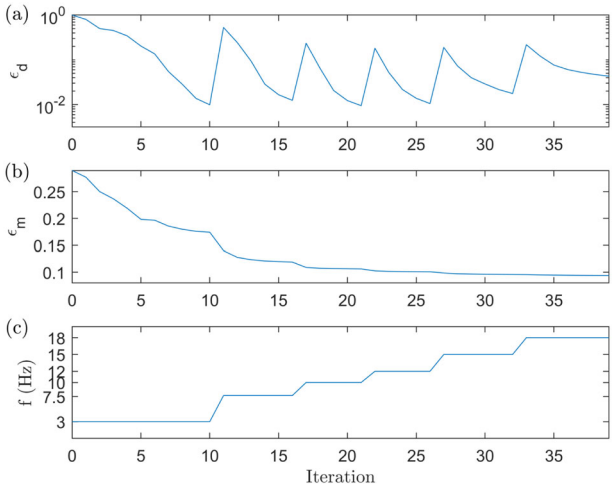

FIGURE 8 Convergence diagrams of the matrix-free DBI in the Marmousi example: (a) normalized data error ϵ_d versus the number of iterations at all frequencies, (b) normalized model error ϵ_m versus the number of iterations at all frequencies and (c) the frequency corresponding to each iteration.

We also test our method in the presence of noise. The noise has been added by using this formula (Jakobsen & Ursin, 2015)

$$\mathbf{d}^{\text{noisy}} = \mathbf{d} + \frac{|\mathbf{d}|}{\sqrt{\text{SNR}}} \cdot \frac{\sigma}{|\sigma|}, \tag{25}$$

where σ is a vector of independent random numbers obtained from a Gaussian distribution and SNR is the signal-to-noise ratio used to measure the noise level. We have tried three noise levels: 5%, 10% and 20% with different SNRs: 26, 20 and 14 dB. The inverted velocities are shown in Figure 9. We can see that a higher noise level causes a deterioration in the inverted results. However, even when the noise level is as high as 20%, the structure of the Marmousi model can still be reconstructed.

Three-dimensional model

To demonstrate the feasibility of our method in a three-dimensional (3D) inversion context, we consider a 3D velocity model (shown in Figure 10), which is resampled from part of the 3D overthrust model (Aminzadeh et al., 1997). The resampled model consists of $128 \times 64 \times 32$ grid blocks. The size of each grid block is 20 m $\times$ 20 m $\times$ 20 m. To better visualize

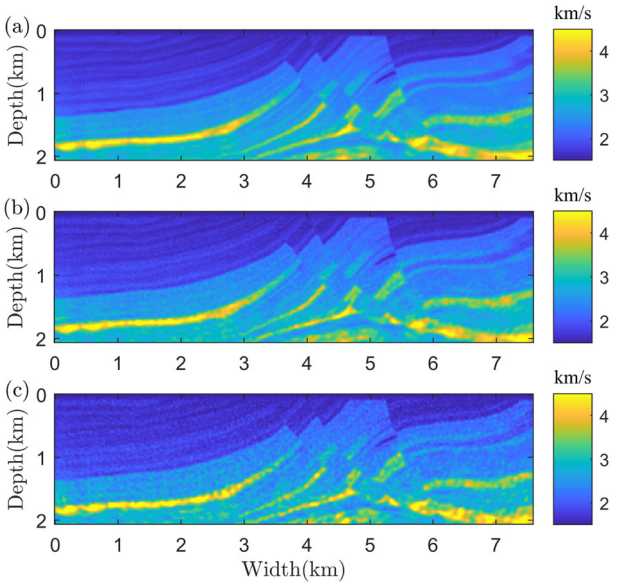

FIGURE 9 Noise test for matrix-free DBI in the Marmousi example. The inverted models with different noise levels: (a) 5%, (b) 10% and (c) 20%.

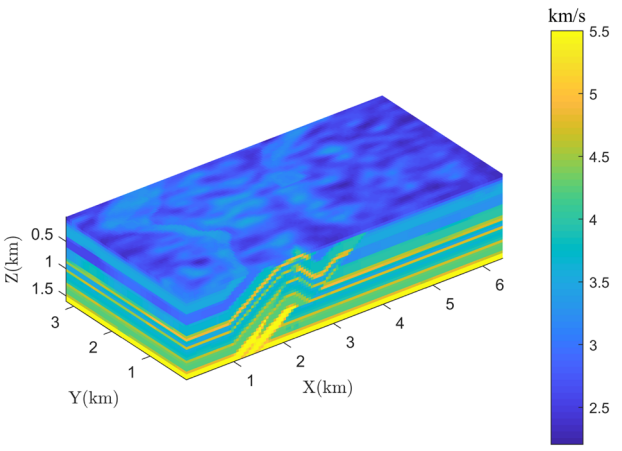

FIGURE 10 Part of the 3D overthrust model, which was used in the third numerical example.

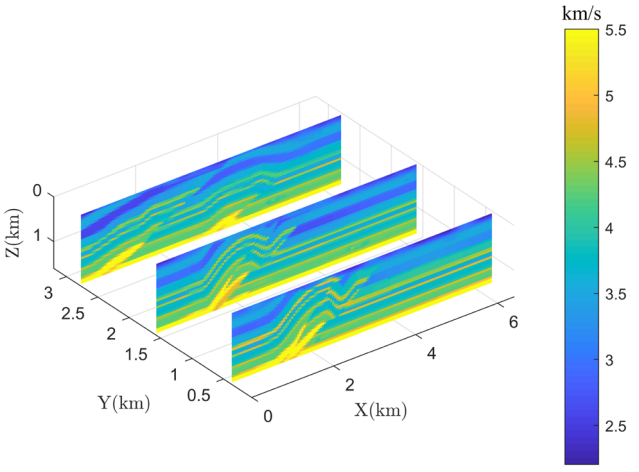

FIGURE 11 Three slices to show the inner structure of the 3D real model.

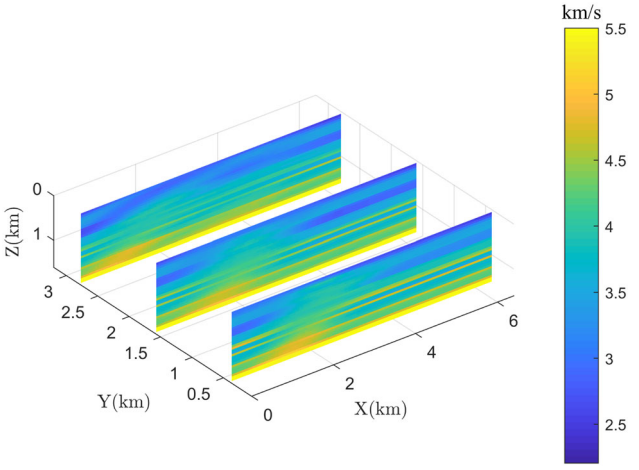

FIGURE 12 Three slices to show the inner structure of the 3D initial model used for the 3D inversion test. The 3D initial model is a smooth version of the 3D real model shown in Figure 10.

the model's inner structure, we only show three slices of the model in Figure 11, the same for the initial and inverted models. The initial model (Figure 12) we used for this example is generated by filtering the real model with a Gaussian smoothing filter with a standard deviation of 8. We set 64 sources and 256 receivers uniformly distributed on the top of the model to get the synthetic data. The same Ricker wavelet as used in the previous numerical examples has been used to simulate the source. To show that other forward modelling solvers can also be used in our method, we adopt the generalized minimal residual method (Saad & Schultz, 1986) to solve the forward problem in this example. Due to the limitation of our computer's memory, we do not include the absorbing boundary here but one can easily include it when the computer's memory is large enough.

Six frequencies (3, 5, 7.5, 10, 12 and 15 Hz) were used for this example. Due to the memory problem of the matrix-based DBI method, only the matrix-free DBI method has been implemented here. The tolerance, maximum number of iterations, initial regularization parameter and decrease factor used in Algorithm 1 are set as $\eta = 0.001$, $k = 10$, $\lambda_{\text{ini}} = 0.001$ and $a = 0.1$, respectively. The maximum number of iterations and residual tolerance for the conjugate gradient method are set to 20 and 0.1, respectively. The inverted model is shown in Figure 13, where we can see that the matrix-free DBI method achieves a satisfactory 3D inversion result. The convergence diagrams in Figure 14 show that the normalized data error and model error both converge well at each frequency. The computational time of this example is 9.88×10^4 s.

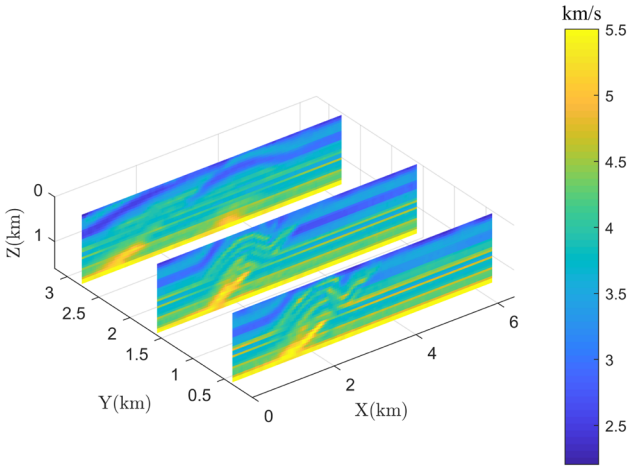

FIGURE 13 Three slices to show the inner structure of the 3D inverted model.

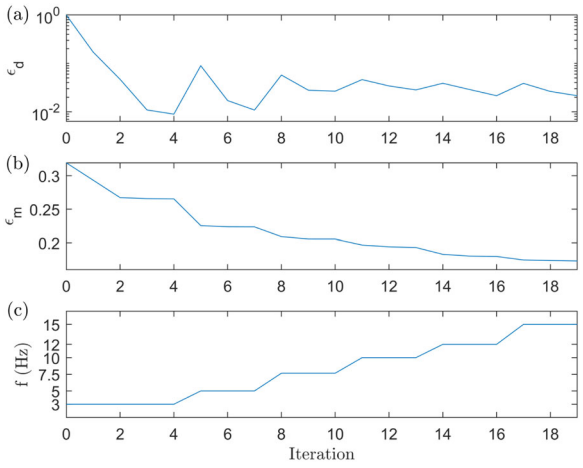

FIGURE 14 Convergence diagrams of the matrix-free DBI in the 3D example: (a) normalized data error ϵ_d versus the number of iterations at all frequencies, (b) normalized model error ϵ_m versus the number of iterations at all frequencies and (c) the frequency corresponding to each iteration.

CONCLUDING REMARKS

We have improved the classic distorted Born iterative (DBI) method based on the integral equation for the mid- and large-size seismic inversion problems. The need to invert the huge approximate Hessian matrix at each iteration was eliminated via the implicit implementation of the Fréchet derivative and its adjoint operators from their physical interpretations. The efficiency of the overall method was improved via the use of the fast-Fourier-transform-accelerated forward solvers. The preconditioned Born series and the Krylov-space method used in this paper provide the requisite for efficiency, reducing the computational cost and storage demand of the Fréchet operator, the adjoint operator, the approximate Hessian and the inverse of the approximate Hessian. The improved computational cost and storage demand make the DBI method more efficient and practical for seismic applications. Numerical examples show the accuracy and effectiveness of this matrix-free DBI method in two- and three-dimensional cases as well as the effect of noise. We compared the computational cost of the conventional DBI method and the matrix-free DBI method for a two-dimensional conceptual reservoir model. The results show that the computational time of the matrix-free DBI method is reduced by approximately 90% compared to the conventional DBI method under that configuration. There are two loops for each frequency in the matrix-free DBI method: the outer loop of the model update and the inner loop of the conjugate gradient method. The stopping criteria for these two loops have a significant impact on the computational cost of the matrix-free DBI method. For example, the computational time can be further reduced if we set a higher tolerance of the conjugate gradient method while the inverted result will be worse. One could explore adopting different termination criteria for these two loops to see if cost-benefit trade-offs are worth it for actual implementations. In addition, the matrix-free DBI method does not require a lot of memory, which makes it a potential choice to be implemented for moderate- or large-scale inverse problems on small computers. Moreover, this matrix-free DBI method provides a connection between the adjoint state method and the integral equation method. There is great potential to design more efficient full waveform inversion algorithms from existing integral equation methods that are commonly used for electromagnetic and acoustic applications. In future work, extending the proposed matrix-free DBI method to more general cases, such as elastic media, anisotropic media and so on, will be investigated and implemented.

ACKNOWLEDGEMENTS

The authors acknowledge the China Scholarship Council for the financial support for Kui Xiang's study in Norway. Eikrem, Nævdal and Jakobsen would like to acknowledge the Research Council of Norway (RCN) for the Petromaks II project 267769 (Bayesian inversion of 4D seismic waveform data for quantitative integration with production data) for financial support.

DATA AVAILABILITY STATEMENT

All numerical experiments in this study are based on synthetic data, no real data involved.

ORCID

Kui Xiang https://orcid.org/0000-0002-4683-919X
Morten Jakobsen https://orcid.org/0000-0001-8861-1938

REFERENCES

Aki, K. & Richards, P.G. (1980) *Quantitative seismology: Theory and methods*. W. R. Freeman and Company.

Alkhalifah, T., Sun, B.B. & Wu, Z. (2018) Full model wavenumber inversion: Identifying sources of information for the elusive middle model wavenumbers. *Geophysics*, 83(6), R597–R610.

Alkhalifah, T. & Wu, Z. (2016) Multiscattering inversion for low-model wavenumbers. *Geophysics*, 81(6), R417–R428.

Aminzadeh, F., Brac, J. & Kunz, T. (1997) *3-D Salt and Overthrust models*. SEG/EAGE 3-D Modeling Series No. 1. SEG Houston, TX: Society of Exploration Geophysicists.

Arfken, G.B. & Weber, H.J. (1999) *Mathematical methods for physicists*. College Park, MD: American Association of Physics Teachers.

Belkebir, K., Kleinman, R.E. & Pichot, C. (1997) Microwave imaging location and shape reconstruction from multifrequency scattering data. *IEEE Transactions on Microwave Theory and Techniques*, 45(4), 469–476.

Červenỳ, V. (2001) *Seismic ray theory*. Cambridge, UK: Cambridge University Press.

Chandra, R., Zhou, H., Balasingham, I. & Narayanan, R.M. (2015) On the opportunities and challenges in microwave medical sensing and imaging. *IEEE Transactions on Biomedical Engineering*, 62(7), 1667–1682.

Chen, P., Jordan, T.H. & Zhao, L. (2007) Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods. *Geophysical Journal International*, 170(1), 175–181.

Chew, W.C. & Wang, Y.M. (1990) Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method. *IEEE Transactions on Medical Imaging*, 9(2), 218–225.

Constable, S.C., Parker, R.L. & Constable, C.G. (1987) Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52(3), 289–300.

Djebbi, R. & Alkhalifah, T. (2020) Hybrid frequency-domain full-waveform inversion using ray+Born sensitivity kernels. *Geophysics*, 85(4), R339–R347.

Eikrem, K.S., Nævdal, G. & Jakobsen, M. (2019) Iterated extended Kalman filter method for time-lapse seismic full-waveform inversion. *Geophysical Prospecting*, 67(2), 379–394.

Eikrem, K.S., Nævdal, G. & Jakobsen, M. (2021) Iterative solution of the Lippmann–Schwinger equation in strongly scattering acoustic media by randomized construction of preconditioners. *Geophysical Journal International*, 224(3), 2121–2130.

Farquharson, C.G. & Oldenburg, D.W. (2004) A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. *Geophysical Journal International*, 156(3), 411–425.

Hackbusch, W. (1999) A sparse matrix arithmetic based on *H*-matrices. part I: introduction to *H*-matrices. *Computing*, 62(2), 89–108.

Halko, N., Martinsson, P.G. & Tropp, J.A. (2011) Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2), 217–288.

Hansen, P.C. (1998) *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Philadelphia, PA: SIAM.

Hesford, A.J. & Chew, W.C. (2006) A frequency-domain formulation of the Frechet derivative to exploit the inherent parallelism of the distorted Born iterative method. *Waves in Random and Complex Media*, 16(4), 495–508.

Hesford, A.J. & Chew, W.C. (2010) Fast inverse scattering solutions using the distorted Born iterative method and the multilevel fast multipole algorithm. *The Journal of the Acoustical Society of America*, 128(2), 679–690.

Hestenes, M.R. & Stiefel, E. (1952) Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6), 409.

Hopfer, M., Planas, R., Hamidipour, A., Henriksson, T. & Semenov, S. (2017) Electromagnetic tomography for detection, differentiation, and monitoring of brain stroke: a virtual data and human head phantom study. *IEEE Antennas and Propagation Magazine*, 59(5), 86–97.

Huang, X., Eikrem, K.S., Jakobsen, M. & Nævdal, G. (2020) Bayesian full-waveform inversion in anisotropic elastic media using the iterated extended Kalman filter. *Geophysics*, 85(4), C125–C139.

Huang, X., Jakobsen, M., Nævdal, G. & Eikrem, K.S. (2019) Target-oriented inversion of time-lapse seismic waveform data. *Communications in Computational Physics*, 27, 1–27.

Huthwaite, P. & Simonetti, F. (2011) High-resolution imaging without iteration: a fast and robust method for breast ultrasound tomography. *The Journal of the Acoustical Society of America*, 130(3), 1721–1734.

Jakobsen, M., Huang, X. & Wu, R.S. (2020) Homotopy analysis of the Lippmann–Schwinger equation for seismic wavefield modelling in strongly scattering media. *Geophysical Journal International*, 222(2), 743–753.

Jakobsen, M. & Ursin, B. (2015) Full waveform inversion in the frequency domain using direct iterative T-matrix methods. *Journal of Geophysics and Engineering*, 12(3), 400–418.

Kirchner, A. & Shapiro, S.A. (2001) Fast repeat-modelling of time-lapse seismograms. *Geophysical Prospecting*, 49(5), 557–569.

Liu, Y., Yang, J., Chi, B. & Dong, L. (2015) An improved scattering-integral approach for frequency-domain full waveform inversion. *Geophysical Journal International*, 202(3), 1827–1842.

Malovichko, M., Khokhlov, N., Yavich, N. & Zhdanov, M. (2017) Approximate solutions of acoustic 3D integral equation and their application to seismic modeling and full-waveform inversion. *Journal of Computational Physics*, 346, 318–339.

Martin, G.S., Wiley, R. & Marfurt, K.J. (2006) Marmousi2: an elastic upgrade for Marmousi. *The Leading Edge*, 25(2), 156–166.

Menke, W. (2012) *Geophysical data analysis: discrete inverse theory: MATLAB edition*. New York: Academic Press.

Morse, P.M. & Feshbach, H. (1954) *Methods of theoretical physics*. New York: McGraw-Hill.

Oristaglio, M. & Blok, H. (2012) *Wavefield imaging and inversion in electromagnetics and acoustics*. Cambridge, UK: Cambridge University Press.

Osnabrugge, G., Benedictus, M. & Vellekoop, I.M. (2021) Ultra-thin boundary layer for high-accuracy simulations of light propagation. *Optics Express*, 29(2), 1649–1658.

Osnabrugge, G., Leedumrongwatthanakun, S. & Vellekoop, I.M. (2016) A convergent Born series for solving the inhomogeneous Helmholtz equation in arbitrarily large media. *Journal of Computational Physics*, 322, 113–124.

Pike, E.R. & Sabatier, P.C. (2002) *Scattering: Scattering and inverse scattering in pure and applied science*. Academic Press.

Pratt, R.G., Shin, C. & Hick, G. (1998) Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133(2), 341–362.

Remis, R.F. & Van den Berg, P. (2000) On the equivalence of the Newton-Kantorovich and distorted Born methods. *Inverse Problems*, 16(1), L1.

Saad, Y. & Schultz, M.H. (1986) GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3), 856–869.

Semenov, S.Y., Bulyshev, A.E., Abubakar, A., Posukh, V.G., Sizov, Y.E., Souvorov, A.E., Van Den Berg, P.M. & Williams, T.C. (2005) Microwave-tomographic imaging of the high dielectric-contrast objects using different image-reconstruction approaches. *IEEE Transactions on Microwave Theory and Techniques*, 53(7), 2284–2294.

Semenov, S.Y. & Corfield, D.R. (2008) Microwave tomography for brain imaging: feasibility assessment for stroke detection. *International Journal of Antennas and Propagation*, 2008.

Souvorov, A.E., Bulyshev, A.E., Semenov, S.Y., Svenson, R.H., Nazarov, A.G., Sizov, Y.E. & Tatsis, G.P. (1998) Microwave tomography: a two-dimensional Newton iterative scheme. *IEEE Transactions on Microwave Theory and Techniques*, 46(11), 1654–1659.

Tao, Y. & Sen, M.K. (2013) Frequency-domain full waveform inversion with a scattering-integral approach and its sensitivity analysis. *Journal of Geophysics and Engineering*, 10(6), 065008.

Tarantola, A. (1984) Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8), 1259–1266.

Tarantola, A. (2005) *Inverse problem theory and methods for model parameter estimation*. Philadelphia, PA: SIAM.

Virieux, J. & Operto, S. (2009) An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6), WCC1–WCC26.

Wang, Y. & Chew, W.C. (1989) An iterative solution of the two-dimensional electromagnetic inverse scattering problem. *International Journal of Imaging Systems and Technology*, 1(1), 100–108.

Yang, J., Liu, Y. & Dong, L. (2016) Simultaneous estimation of velocity and density in acoustic multiparameter full-waveform inversion using an improved scattering-integral approach. *Geophysics*, 81(6), R399–R415.

How to cite this article: Xiang, K., Jakobsen, M., Eikrem, K.S. & Nævdal, G. (2023) A matrix-free variant of the distorted Born iterative method for seismic full-waveform inversion. *Geophysical Prospecting*, 71, 431–442. https://doi.org/10.1111/1365-2478.13323

APPENDIX A: DERIVATION OF THE ADJOINT OPERATOR

The adjoint operator $\mathcal{F}^{\dagger}$ must satisfy (Tarantola, 2005)

$$\langle [\mathcal{F}X](\mathbf{r}), Y(\mathbf{r})\rangle_D = \left\langle X(\mathbf{x}), [\mathcal{F}^{\dagger}Y](\mathbf{x})\right\rangle_{\Omega}, \tag{A.1}$$

where $<\cdot>$ denotes an inner product of two vectors in the Hilbert space. Using the definition of the inner product to expand (A.1) yields

$$\int_D d\mathbf{r}[\mathcal{F}X](\mathbf{r})Y^*(\mathbf{r}) = \int_{\Omega} d\mathbf{x}X(\mathbf{x})\left\{[\mathcal{F}^{\dagger}Y](\mathbf{x})\right\}^*, \tag{A.2}$$

where $\{\cdot\}^*$ represents complex conjugation. Substituting the Fréchet operator (9) into the left side of (A.2) and changing the integral orders, we have

$$\int_D d\mathbf{r}\left[\omega^2\int_{\Omega} d\mathbf{x}G_b(\mathbf{r},\mathbf{x})X(\mathbf{x})p_b(\mathbf{x})\right]Y^*(\mathbf{r}) = \int_{\Omega} d\mathbf{x}X(\mathbf{x})\omega^2 p_b(\mathbf{x})\left[\int_D d\mathbf{r}G_b(\mathbf{r},\mathbf{x})Y^*(\mathbf{r})\right]. \tag{A.3}$$

Comparing the right-hand side of (A.3) and (A.2) obtains

$$\left\{[\mathcal{F}^{\dagger}Y](\mathbf{x})\right\}^* = \omega^2 p_b(\mathbf{x})\left[\int_D d\mathbf{r}G_b(\mathbf{r},\mathbf{x})Y^*(\mathbf{r})\right]. \tag{A.4}$$

Taking the complex conjugate on both sides of (A.4) and using the reciprocity property of the Green's function (Aki & Richards, 1980), we finally yield

$$[\mathcal{F}^{\dagger}Y](\mathbf{x}) = \left\{\omega^2 p_b(\mathbf{x})\left[\int_D d\mathbf{r}G_b(\mathbf{x},\mathbf{r})Y^*(\mathbf{r})\right]\right\}^*. \tag{A.5}$$

APPENDIX B: THE PRECONDITIONED BORN SERIES

The details of the preconditioned Born series are presented here for completeness. Considering the background medium v_b in (4) is a homogeneous medium v_0, Equation (4) can be rewritten as

$$p(\mathbf{x},\omega) = p^{(0)}(\mathbf{x},\omega) + \omega^2\int_{\Omega} d\mathbf{x}'G^{(0)}(\mathbf{x},\mathbf{x}',\omega)\chi(\mathbf{x}')p(\mathbf{x}',\omega), \tag{B.1}$$

where

$$p^{(0)}(\mathbf{x},\omega) = \int d\mathbf{x}'G^{(0)}(\mathbf{x},\mathbf{x}',\omega)S(\mathbf{x}',\omega) \tag{B.2}$$

is the wavefield within the homogeneous medium and

$$\chi(\mathbf{x}') = \frac{1}{v^2(\mathbf{x}')} - \frac{1}{v_0^2} \tag{B.3}$$

is the contrast between the real and the homogeneous medium. $G^{(0)}(\mathbf{x},\mathbf{x}',\omega)$ is the Green's function for the homogeneous background medium which can be found in Equation (19).

For numerical implementations, we discretized the scattering domain Ω into N uniform grid blocks, each of area Δs. The wavefield and the parameter within each discretized grid block are assumed to be constant. After discretization and rearrangement, Equation (B.1) can be written compactly in matrix form as

$$\mathbf{p} = \mathbf{p}^{(0)} + \mathbf{G}^{(0)}\mathbf{V}\mathbf{p}, \tag{B.4}$$

where $\mathbf{V} = \mathrm{Diag}(\boldsymbol{\chi})$ is the scattering potential matrix with the scattering potential at all scattering points on the diagonal. In Equation (B.4), the wavefields $\mathbf{p}$, $\mathbf{p}^{(0)}$ and the scattering potential vector $\boldsymbol{\chi}$ are $N\times 1$ vectors, the Green's function $\mathbf{G}^{(0)}$ and the scattering potential matrix $\mathbf{V}$ are $N\times N$ matrices.

Equation (B.4) can be solved directly by matrix inversion, but matrix inversion includes an inversion of a full large matrix, which is very costly. It can also be solved by the Born series (Morse & Feshbach, 1954), but this series only converges if the scattering potential is weak enough. In order to solve Equation (B.4) efficiently for a strongly scattering medium, we introduce the preconditioned Born series (Eikrem et al., 2021; Osnabrugge et al., 2016)

$$\mathbf{p} = \sum_{k=0}^{\infty} \mathbf{M}^k \mathbf{H} \mathbf{p}^{(0)}, \tag{B.5}$$

where $\mathbf{H}$ is the preconditioner we will discuss in detail later and

$$\mathbf{M} = \mathbf{I} - \mathbf{H} + \mathbf{H}\mathbf{G}^{(0)}\mathbf{V}. \tag{B.6}$$

The preconditioned Born series (B.5) is guaranteed to converge if the spectral radius σ of $\mathbf{M}$, denoted $\sigma(\mathbf{M})$, is less than unity. In Equation (B.6), $\mathbf{H}$ can be selected to make sure $\sigma(\mathbf{M}) < 1$. Jakobsen et al. (2020) gave examples of a few simple $\mathbf{H}$, all of which are diagonal matrices. For general implementation, Eikrem et al. (2021) proposed a non-diagonal $\mathbf{H}$ based on the randomized low-rank approximation (Halko et al., 2011) and the hierarchical matrix (Hackbusch, 1999). We adopt the hierarchical preconditioner from Eikrem et al. (2021) in this paper.

If $\mathbf{M} = 0$, $\sigma(\mathbf{M}) = 0$ so that $\sigma(\mathbf{M}) < 1$. But when we set $\mathbf{M} = 0$ in Equation (B.6), $\mathbf{H} = (\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V})^{-1}$, which is the same as matrix inversion. The key idea of Eikrem et al. (2021) is to find an $\mathbf{H}$ that is not equal to $(\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V})^{-1}$ but approximate to $(\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V})^{-1}$. Then $\sigma(\mathbf{M}) \neq 0$ but $\sigma(\mathbf{M}) \approx 0$, which still satisfies $\sigma(\mathbf{M}) < 1$. In Eikrem et al. (2021), $\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V}$ is approximated by a hierarchical matrix (Hackbusch, 1999). To construct the hierarchical matrix, first we need to divide the square matrix $\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V}$ into four equal-sized square blocks. Then, we keep the blocks on the off-diagonal unchanged and continue to divide the diagonal blocks into four equal parts. Next, we keep repeating the previous step to divide the diagonal blocks further. The number of divisions is the level. Finally, we will divide the full matrix $\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V}$ as shown in Figure B.1, which is an example of the four-level division. In Figure B.1, the grey blocks are kept as full matrices and the white blocks are approximated by the multiplication of two low-rank matrices

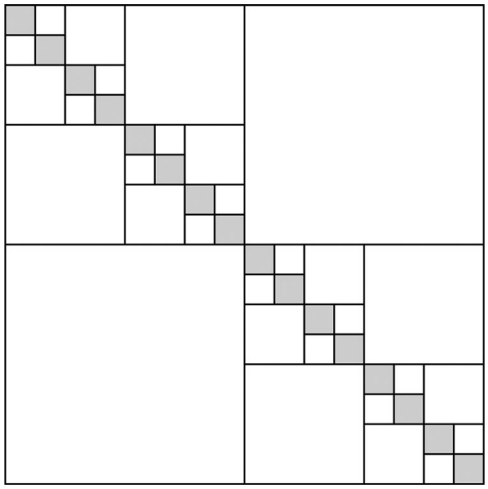

FIGURE B.1 Structure of the hierarchical matrix that approximates $\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V}$: the grey blocks are full matrices and the white are approximated by low-rank matrices.

$$\mathbf{W} = \mathbf{A}\mathbf{B}^T, \tag{B.7}$$

where $\mathbf{W}$ is an $n \times n$ matrix representing any white block, $\mathbf{A}$ and $\mathbf{B}$ are both $n \times k$ matrices and k is the rank whose value is much smaller than n. After approximating all the white blocks with low-rank matrices, we get the hierarchical matrix. Next, we use the rule of the 2×2 block matrix inversion recursively to find the inverse of the hierarchical matrix

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{pmatrix}. \tag{B.8}$$

This inverse is the approximation of $(\mathbf{I} - \mathbf{G}^{(0)}\mathbf{V})^{-1}$, which can be used as the preconditioner $\mathbf{H}$ in the preconditioned Born series. Although all the above steps are formulated with matrices, in practical applications, we can use them with randomized methods (Halko et al., 2011) and fastFourier transform without needing to form the full matrix, which significantly reduces memory consumption. A more detailed description of the hierarchical preconditioner can be found in Eikrem et al. (2021).