3D Marchenko applications: implementation and examples

Joeri Brackenhoff[1,2]*, Jan Thorbecke[1], Giovanni Meles[1,3], Victor Koehne[4],
Diego Barrera[4] and Kees Wapenaar[1]

[1]*Department of Geoscience and Engineering, Delft University of Technology, Stevinweg 1, Delft, 2628CN, The Netherlands,* [2]*Department of Earth Sciences, ETH Zürich, Sonneggstrasse 5, Zurich, 8092, Switzerland,* [3]*Institute of Earth Sciences, University of Lausanne, Lausanne, 1015, Switzerland, and* [4]*Senai Cimatec, Av. Orlando Gomes 2845, Salvador, Bahia 41650-010, Brazil*

ABSTRACT

We implement the 3D Marchenko equations to retrieve responses to virtual sources inside the subsurface. For this, we require reflection data at the surface of the Earth that contain no free-surface multiples and are densely sampled in space. The required 3D reflection data volume is very large and solving the Marchenko equations requires a significant amount of computational cost. To limit the cost, we apply floating point compression to the reflection data to reduce their volume and the loading time from disk. We apply the Marchenko implementation to numerical reflection data to retrieve accurate Green's functions inside the medium and use these reflection data to apply imaging. This requires the simulation of many virtual source points, which we circumvent using virtual plane-wave sources instead of virtual point sources. Through this method, we retrieve the angle-dependent response of a source from a depth level rather than of a point. We use these responses to obtain angle-dependent structural images of the subsurface, free of contamination from wrongly imaged internal multiples. These images have less lateral resolution than those obtained using virtual point sources, but are more efficiently retrieved.

Key words: Signal processing, Seismics, Numerical study.

INTRODUCTION

The Marchenko method, originally derived for the field of quantum physics (Marchenko, 1955), was introduced in the field of geophysics about a decade ago. The method employs reflection data that are recorded at the surface of the Earth to create sources and receivers inside the Earth while internal multiples are handled correctly. Because these sources and receivers are not physically located inside the subsurface but instead are created from the reflection data, they are called virtual sources and receivers. The original Marchenko equation was introduced in the field of quantum physics in the 1950s by Marchenko (1955). More recently, the method was developed in the field of geophysics by Broggini and Snieder (2012)

for 1D media and was further extended for 2D and 3D applications by Wapenaar *et al.* (2013), Broggini *et al.* (2014), Slob *et al.* (2014) and Behura *et al.* (2014). These authors showed that the reflection data and Green's function, that is, the impulse response of a medium, can be related via a so-called focussing function, which is a wavefield that focusses from an open boundary to a location in the subsurface, often referred to as the focal location. This relation can be rewritten into a Marchenko-type equation in 3D and subsequently be solved using only the reflection data and an estimation of the first arrival of a wavefield from the focal location. The first arrival can be easily obtained through the use of a background velocity model.

The Marchenko method has been used for many applications since it was introduced in the field of geophysics. It has been applied to create images of the subsurface that are free of artefacts related to the internal multiples (Broggini *et al.*,

*E-mail: johannes.brackenhoff@erdw.ethz.ch

2014; Meles *et al.*, 2015; Ravasi, 2017; Matias *et al.*, 2018). However, the method has also been used to obtain the homogeneous Green's function in the subsurface (Brackenhoff *et al.*, 2019a, 2019b), remove internal multiples from reflection data (Zhang and Staring, 2018) or to retrieve plane-wave responses in the subsurface (Meles *et al.*, 2018). These applications used acoustic reflection data that contained no free-surface multiples. It has been shown that the Marchenko method can also be applied using elastic reflection data (da Costa Filho *et al.*, 2014; Wapenaar, 2014; Zhou *et al.*, 2019; Reinicke, 2020) and reflection data that contain free-surface multiples (Singh *et al.*, 2015; Slob and Wapenaar, 2017). Although the theory has been fully developed for 3D applications, most publications are applied to 2D datasets. The method has been successfully applied in 2D and 3D using both synthetic and field data (Ravasi *et al.*, 2016; van der Neut and Wapenaar, 2016; Pereira *et al.*, 2019; Staring and Wapenaar, 2019; Lomas and Curtis, 2020; Staring and Wapenaar, 2020a), however, the requirements of the reflection data are highly demanding, as the reflection data need to be well sampled in both time and space. This type of dense sampling is hard to obtain in the field for 3D acquisitions and furthermore, the 3D datasets are considerably larger in storage size than their 2D counterparts. Lomas and Curtis (2020) made a comparison between results obtained with a full 3D acquisition and results obtained with linear seismic acquisition arrays, both recorded over a 3D medium. The authors showed that while the 2D approximation can yield good results, if one wants to take into account the full 3D effects, such as out-of-plane reflections, a 3D version of the Marchenko method is required.

The Marchenko equations can be solved in a variety of ways, the most common way is to use an iterative scheme (Broggini *et al.*, 2014). However, the equations can also be solved using a least-squares inversion (van der Neut *et al.*, 2015; Ravasi, 2017; Ravasi and Vasconcelos, 2021) or the iterative scheme can be combined with adaptive subtraction (da Costa Filho *et al.*, 2018; Staring and Wapenaar, 2020a). The least-squares inversion is computationally feasible for 2D reflection data, however, for 3D data this method becomes computationally expensive. The adaptive subtraction is more robust to imperfections in the reflection data, however, due to its adaptive nature, the subtraction can attenuate physical events that are coinciding with multiples.

Various open-source implementations of the Marchenko method have been released over the years. The first of these implementations was written in the Madagascar environment, employing the 2D iterative scheme (Broggini *et al.*, 2014). Thorbecke *et al.* (2017) published a C implementation of the

Marchenko method that employs an iterative scheme for 2D reflection data that contain no free-surface multiples. More recently, an implementation in Python using the PyLops package was published that solves the method using least-squares inversion in 3D (Ravasi and Vasconcelos, 2021). In this paper, we extend the 2D iterative implementation of Thorbecke *et al.* (2017) to work with 3D reflection data. We discuss the practical challenges of the 3D extension, as the theoretical extension is straightforward, and show applications of the code.

We start by briefly considering the theory of the Marchenko method which forms the basis of the implementation. Next, we discuss the changes made to the 2D implementation to extend the implementation for 3D datasets. We use a floating point algorithm by Lindstrom (2014), called ZFP, to compress the reflection data to limit the required storage space and to reduce the loading time of the data. We also use a 3D Eikonal solver to obtain the first arrivals in the subsurface, to limit the modelling time that would be required if finite-difference modelling was used. We apply our 3D code to a subsection of the complex 3D overthrust model by Aminzadeh *et al.* (1997) for two applications. First, we retrieve Green's functions related to point sources inside the model and compare the results to a directly modelled reference . Next, we use these retrieved wavefields and focussing functions to apply structural imaging, based on the double-focussing approach (Staring *et al.*, 2018) and show how the Marchenko method attenuates artefacts in the image, which are related to internal multiples. As this is a computationally expensive process in 3D, we demonstrate how the retrieval of Green's functions can be performed for an entire depth level, through the use of plane-wave sources, based on the work by Meles *et al.* (2018). This allows us to obtain a structural image of the subsurface much more efficiently, although only certain angles of reflectivity are obtained. We therefore retrieve the reflectivity for multiple angles in order to obtain a more complete structural image of the subsurface.

THEORY

Our 3D implementation of the Marchenko method follows the same approach as the 2D implementation by Thorbecke *et al.* (2017). As the specifics of the implementation are described in great detail in that paper, we will only consider the most important equations, which form the basis of the implementation.

The Marchenko method allows one to retrieve Green's functions $G^{-,\pm}(\mathbf{x}_R, \mathbf{x}_A, t)$ in the subsurface, which are the impulse responses at point $\mathbf{x}_R$ at the surface of the Earth at time

t to a virtual point source at $\mathbf{x}_A$ in the subsurface. Note that we use a Cartesian coordinate system, where $\mathbf{x} = (x_1, x_2, x_3)$ and x_1 and x_2 are the horizontal coordinates and x_3 is the depth coordinate. The first superscript in Green's function indicates whether a wave is propagating downwards or upwards at $\mathbf{x}_R$, using + or -, respectively, and the second superscript follows a similar notation for the radiation direction of the virtual source at $\mathbf{x}_A$. In order to retrieve Green's functions, we require a reflection response $R(\mathbf{x}_R, \mathbf{x}_S, t)$, which is measured at location $\mathbf{x}_R$ at transparent surface $\mathbb{S}_0$, usually the surface of the Earth, as a result of a dipole source $\mathbf{x}_S$ at the same surface. The reflection response contains no free-surface multiples because of the assumption that $\mathbb{S}_0$ is transparent. Field datasets do contain this type of multiples, however, several robust techniques exist to remove these events (Verschuur *et al.*, 1992; Amundsen, 2001; van Groenestijn and Verschuur, 2009). Alternatively, the Marchenko method can also be adjusted to directly work with data containing free-surface multiples (Singh *et al.*, 2015), however, this approach is not considered in this paper.

Green's functions and reflection response can be related to each other via focussing functions $f_1^{\pm}(\mathbf{x}_R, \mathbf{x}_A, t)$, which are propagating downward or upward at $\mathbf{x}_R$, depending on the superscript. These focussing function focus from the surface to a focal point $\mathbf{x}_A$ in the subsurface, while correctly accounting for internal multiples. The relations between Green's functions, focussing functions and reflection response are given as (Wapenaar *et al.*, 2014b, equations (11) and (12))

$$G^{-,+}(\mathbf{x}_R, \mathbf{x}_A, t) + f_1^-(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$= \int_{\mathbb{S}_0} \int_0^{\infty} R(\mathbf{x}_R, \mathbf{x}_S, t') f_1^+(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S, \qquad (1)$$

$$G^{-,-}(\mathbf{x}_R, \mathbf{x}_A, -t) + f_1^+(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$= \int_{\mathbb{S}_0} \int_{-\infty}^0 R(\mathbf{x}, \mathbf{x}_S, -t') f_1^-(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S. \qquad (2)$$

Note that these relations are based on the decomposed wavefields. The focussing functions can be combined together to create a total focussing function:

$$f_2(\mathbf{x}_A, \mathbf{x}_R, t) = f_1^+(\mathbf{x}_R, \mathbf{x}_A, t) - f_1^-(\mathbf{x}_R, \mathbf{x}_A, -t), \qquad (3)$$

where $f_2(\mathbf{x}_A, \mathbf{x}_R, t)$ is a focussing function that focusses from below to a focal point $\mathbf{x}_R$ at the surface. Similarly, following Wapenaar (2020), the decomposed Green's functions can be used to construct the total Green's function, $G(\mathbf{x}_R, \mathbf{x}_A, t)$:

$$G(\mathbf{x}_R, \mathbf{x}_A, t) = G^{+,+}(\mathbf{x}_R, \mathbf{x}_A, t) + G^{+,-}(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$+ G^{-,+}(\mathbf{x}_R, \mathbf{x}_A, t) + G^{-,-}(\mathbf{x}_R, \mathbf{x}_A, t). \qquad (4)$$

Note that if $\mathbf{x}_R$ is located at $\mathbb{S}_0$, the wavefield is purely upgoing because the surface is transparent and the medium above the surface is homogeneous. In other words, $G^{+,+}(\mathbf{x}_R, \mathbf{x}_A, t) = G^{+,-}(\mathbf{x}_R, \mathbf{x}_A, t) = 0$ and total Green's function consists only of $G^{-,+}(\mathbf{x}_R, \mathbf{x}_A, t)$ and $G^{-,-}(\mathbf{x}_R, \mathbf{x}_A, t)$, which are Green's functions in equations (1) and (2).

Green's functions and focussing functions on the left-hand side of equations (1) and (2) can be separated from each other based on causality relations. The direct arrival of Green's function, $G_d(\mathbf{x}_R, \mathbf{x}_A, t)$, and the direct arrival of the downgoing focussing function, $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$, are each other's inverse and therefore $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, -t)$ and $G_d(\mathbf{x}_R, \mathbf{x}_A, t)$ have the same arrival time $t_d(\mathbf{x}_R, \mathbf{x}_A)$, however, their amplitudes are different (van der Neut *et al.*, 2015). Furthermore, aside from the direct arrival, the entire wavefield of $G(\mathbf{x}_R, \mathbf{x}_A, t)$ arrives after $t_d(\mathbf{x}_R, \mathbf{x}_A)$ and the entire wavefield of $f_1^+(\mathbf{x}_A, \mathbf{x}_R, -t)$ and $f_1^-(\mathbf{x}_A, \mathbf{x}_R, t)$ arrives before $t_d(\mathbf{x}_R, \mathbf{x}_A)$, which means that the wavefields are separated in time, except for their direct arrival. Using these relations, we introduce an offset-dependent time-windowing function,

$$\Theta(\mathbf{x}_R, \mathbf{x}_A, t) = \theta(t + (t_d(\mathbf{x}_R, \mathbf{x}_A) - \epsilon))$$
$$- \theta(t - (t_d(\mathbf{x}_R, \mathbf{x}_A) - \epsilon)), \qquad (5)$$

where $\theta(t)$ is the Heaviside step function and ϵ indicates a small constant, which is required to account for the band-limited nature of the wavefields. For simplicity, we will use the substitution $\Theta(\mathbf{x}_R, \mathbf{x}_A, t) = \Theta$. Applying Θ to the Green's function removes all events, however, when the window is applied to the focussing function only the first arrival is removed, while the coda will remain. Because the window is symmetric in time, it does not matter whether the wavefields are time-reversed when the windowing function is applied. The window is applied to equations (1) and (2) in order to remove Green's functions from the equations

$$f_1^-(\mathbf{x}_R, \mathbf{x}_A, t) = \Theta \int_{\mathbb{S}_0} \int_0^{\infty} R(\mathbf{x}_R, \mathbf{x}_S, t') f_1^+(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S, \qquad (6)$$

$$f_1^+(\mathbf{x}_R, \mathbf{x}_A, t) - f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$= \Theta \int_{\mathbb{S}_0} \int_{-\infty}^0 R(\mathbf{x}_R, \mathbf{x}_S, -t') f_1^-(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S. \qquad (7)$$

Equations (6) and (7) are the coupled Marchenko equations. Note that in equation (7), we subtract the direct arrival from the downgoing focussing function to account for the removal of this direct arrival by the windowing function. Assuming that the direct arrival is known, in this system we have

two unknowns and two equations, allowing us to solve it in an iterative manner. The reflection response is the known quantity and will not change, while the upgoing and downgoing focussing function will be updated according to

$$f_{1,k}^-(\mathbf{x}_R, \mathbf{x}_A, t) = \Theta \int_{\mathbb{S}_0} \int_0^\infty R(\mathbf{x}_R, \mathbf{x}_S, t') f_{1,k}^+(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S, \tag{8}$$

$$f_{1,k+1}^+(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$= \Theta \int_{\mathbb{S}_0} \int_{-\infty}^0 R(\mathbf{x}_R, \mathbf{x}_S, -t') f_{1,k}^-(\mathbf{x}_S, \mathbf{x}_A, t - t') \mathrm{d}t' \mathrm{d}\mathbf{x}_S$$
$$+ f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t), \tag{9}$$

where k indicates the iteration number. In equation (9), we added the direct arrival to both sides of the equation, so that the full downgoing focussing function is retrieved on the left-hand side. To start the scheme, a first estimation is required. For the first estimation, we assume that the scattering coda of the downgoing focussing function is equal to zero, so that

$$f_{1,0}^+(\mathbf{x}_R, \mathbf{x}_A, t) = f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t). \tag{10}$$

As it is difficult to estimate $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$, one can instead model $G_d(\mathbf{x}_R, \mathbf{x}_A, t)$, as it can be easily obtained using a background velocity model. $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$ is often approximated using the final arrival of the time-reversed Green's function, $G_d(\mathbf{x}_A, \mathbf{x}_R, -t)$. As mentioned before, $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$ and $G_d(\mathbf{x}_A, \mathbf{x}_R, -t)$ have the same arrival time, however, there is an amplitude difference. Using a time-reversed direct arrival instead of an inverted direct arrival as a first estimation of the focussing function will cause errors in the final result that are proportional to the transmission losses of the medium (Broggini *et al.*, 2014; Brackenhoff, 2016; van der Neut *et al.*, 2018). When $G_d(\mathbf{x}_R, \mathbf{x}_A, t)$ is modelled in a smooth version of the medium, instead of the exact medium, errors will be present in any case, which is why the time reversal is more often used than the inversion of the direct arrival of Green's function. Note that in imaging by deconvolution, discussed in the section 'Marchenko imaging', these amplitude errors should be compensated for.

IMPLEMENTATION

Algorithm

The implementation that is used in this paper uses a single compute kernel. By considering equations (8) and (9), it can be seen that both these equations use the same operation,

namely a convolution of the reflection response or its time reversal with either the downgoing or upgoing focussing function, followed by an application of the time window Θ. It is more efficient to perform the convolution in the frequency domain, while the time window is applied in the time domain. We therefore consider these two operations separately in the kernel. Because the kernel is same for both equations (8) and (9), the equations can be written as a series expansion, similar to equations (B-4) and (B-5) of Wapenaar *et al.* (2014b):

$$-f_{1,k}^-(\mathbf{x}_R, \mathbf{x}_A, -t) = \sum_{i=0}^k N_{2i}(\mathbf{x}_A, \mathbf{x}_R, t), \tag{11}$$

$$f_{1,k+1}^+(\mathbf{x}_R, \mathbf{x}_A, t) = \sum_{i=0}^k N_{2i+1}(\mathbf{x}_A, \mathbf{x}_R, t) + f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$$
$$= \sum_{i=-1}^k N_{2i+1}(\mathbf{x}_A, \mathbf{x}_R, t), \tag{12}$$

where

$$N_i(\mathbf{x}_A, \mathbf{x}_R, -t) = -\Theta R N_{i-1}(\mathbf{x}_A, \mathbf{x}_R, t), \tag{13}$$

$$RN_{i-1}(\mathbf{x}_A, \mathbf{x}_R, \omega)$$
$$= \sum_{j=1}^{n_S} R(\mathbf{x}_R, \mathbf{x}_{S,j}, \omega) N_{i-1}(\mathbf{x}_A, \mathbf{x}_{S,j}, \omega) \frac{\Delta x_1 \Delta x_2 \Delta t}{n_t}, \tag{14}$$

$$N_{-1}(\mathbf{x}_A, \mathbf{x}_R, t) = f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t). \tag{15}$$

Equation (14) is performed in the frequency domain and therefore the functions depend on the angular frequency ω instead of time. The data are transformed to the frequency domain using the Fourier transform and as a result, the convolution between the reflection response and N_i becomes a simple multiplication. To account for the discrete Fourier transform, the data need to be divided by the number of time samples n_t and a scaling with the temporal sampling Δt is required for the numerical time convolution. We have also replaced the integral over the source locations by a numerical approximation, namely a sum over the source positions, where $\mathbf{x}_{S,j}$ indicates the jth source location of a total of n_S source locations. Furthermore, for this approximation the data need to be scaled by the spatial sampling in the horizontal direction x_1 and x_2. Δx_1 and Δx_2 indicate the spatial sampling at $\mathbf{x}_{S,j}$ in horizontal direction x_1 and x_2, respectively, and can be different. For the 2D implementation, Δx_2 can be dropped and the scaling only depends on one horizontal direction.

3D challenges

The basic algorithm that is described by equations (8)–(14) is the same as was used in the 2D implementation by Thorbecke *et al.* (2017) and the implementation is very similar, as the only addition is the integration along the second horizontal direction. However, there are additional complications that come with the 3D extension.

First, $R(\mathbf{x}_R, \mathbf{x}_S, \omega)$ has to be considered. The 3D reflection data must be pre-processed to comply with the assumptions made in the derivation of the Marchenko equations (Wapenaar *et al.*, 2014a). This processing has to at least include (Brackenhoff *et al.*, 2019b)

- removal of free-surface multiples;
- deconvolution with source wavelet;
- crossline interpolation to avoid aliasing.

The removal of free-surface multiples and the deconvolution of the wavelet is beyond the scope of this paper, however, the crossline interpolation is a relevant topic as it relates to one of the most critical requirements of the reflection data, namely dense spatial sampling. Our implementation assumes that this requirement has been fulfilled, however, as a result the size of the reflection data will become very large. Loading the pre-processed reflection data from disk to memory is generally an inexpensive task in 2D, however, loading the full 3D reflection data matrix takes much longer and the storage space that is required for the data is much larger. To mitigate these problems a ZFP-based compression algorithm is used (Lindstrom, 2014). The ZFP compression has been successfully applied for the purpose of reducing the data size of seismic waveforms for waveform tomography with minimal errors (Lindstrom *et al.*, 2016). Before the reflection data are compressed, they are transformed to the frequency domain, as the reflection data are only used in the frequency domain to circumvent the convolution, and only the data in the frequency band of interest are compressed and stored to disk. Typically the lossless ZFP compression reduces the 3D data size by a factor of 4, which decreases the storage space of the data and the read-in time to memory. The program `TWtransform` (explained in Appendix A.2) transforms uncompressed reflection time-data to the frequency domain, applies ZFP compression (based on tolerance) on a selected frequency range and writes the compressed data to disk. The data on disk contain a special compressed header that includes all location information present in the uncompressed Segy/SU headers that are needed in the 3D Marchenko program. The 3D Marchenko program has multiple options to read the reflection data, which can be done in the time domain, frequency domain or compressed frequency domain. In the Appendix section 'marchenko3D', the most important options and parameter settings of the 3D Marchenko program, `marchenko3D`, are explained in more detail.

Aside from the reflection data, the first estimation of the focussing function is also required, as dictated by equation (15). To model this first arrival, a background velocity model is required. Using this model, the first arrival can be calculated using, for example a finite-difference modelling code. In this paper, we use a 3D finite-difference modelling code `fdelmodc3D` that can be found in the OpenSource library by Thorbecke and Brackenhoff (2019). While the calculation of a single first arrival is feasible using a finite-difference code, it is very computationally expensive to calculate it for a large amount of focal points in this way. Furthermore, storing all these wavefields separately would require a large amount of storage space. As an alternative to the finite-difference modelling, an Eikonal, or ray-based solver can be employed. Using this solver, the first arrival of multiple focal locations can be efficiently computed and stored on disk. We employ the Eikonal solver `raytime3D` from the OpenSource library to calculate these first arrivals. This code is based on the theory by Vidale (1990) and aside from the calculation of the arrival time, it also estimates a geometric spreading factor of the wavefield, based on the concepts by Spetzler and Angelov (2005). This amplitude estimation is important, as a uniform amplitude for all arrival times would cause significant artefacts in the result of the Marchenko method. This amplitude estimation is not exact and effects, such as transmission losses, are not taken into account, therefore there will be errors introduced in the amplitude of the estimated Green's function. The first arrival times that are estimated, either through the finite-difference modelling code or the Eikonal solver, are also used to construct the time window Θ.

When the reflection data, time window and the first arrival are available, they can be used to compute the focussing functions through the use of equations (8)–(14). When the iterative scheme has converged to a solution, the final estimations of the upgoing and downgoing focussing functions can be used in equations (1) and (2) to compute the decomposed Green's functions. The `marchenko3D` program has the same functionality as its 2D counterpart `marchenko` to run multiple focal locations simultaneously, which means that the 3D reflection data only have to be read from the disk once. This is especially useful in 3D as loading the 3D reflection data from disk takes up a significant portion of time compared to the computing time of the Marchenko method.
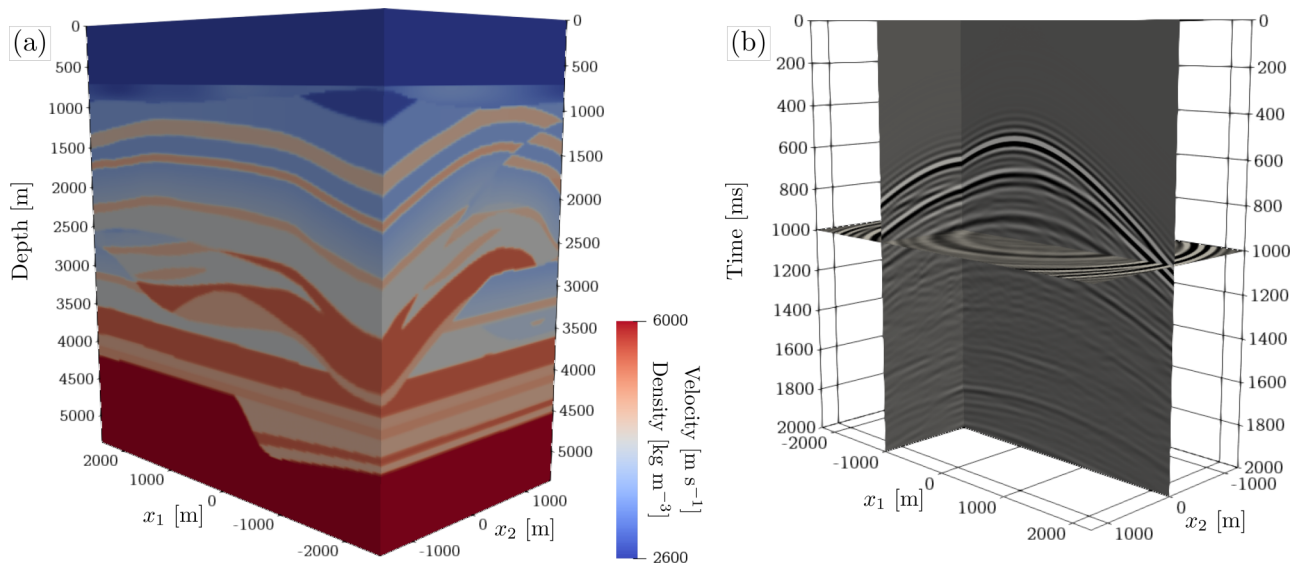
Figure 1 (a) Subsection of the overthrust model used for both velocity and density contrasts. (b) A common-source record, with source position $\mathbf{x}_S = (0, 0, 0)$ and receivers at $\mathbf{x}_R$, containing a wavelet with a flat spectrum between 5 and 25 Hz.

NUMERICAL EXAMPLES

We demonstrate the application of the 3D Marchenko scheme for two different applications, namely Green's function retrieval and imaging. We perform these applications for two types of virtual sources, namely point sources and plane-wave sources.

To take into account complex scattering in three dimensions, we demonstrate the method on numerical data that were modelled in a subsection of the SEG/EAGE overthrust model from Aminzadeh *et al.* (1997), which is publicly available from the SEG Wiki (https://wiki.seg.org/wiki/SEG/ EAGE_Salt_and_Overthrust_Models). We select a subsection of the model, as a recording set-up over the full extent of the model is too large to fit in the memory of our computing nodes. Furthermore, this reduces the modelling time for the reflection response. We insert a layer with constant velocity and density above the model as a water layer, to simulate a marine setting. The velocity and density of the subsection are shown in Figure 1(a). The values of the density model are chosen to be the same as those of the velocity model to ensure strong reflections. We use the `fdelmodc3D` code to model the reflection response. An example of a shot record from a source at the surface in the centre of the model is shown in Figure 1(b). For the full reflection response, we use a fixed spread acquisition, where the source is modelled at every receiver position. In the inline x_1-direction the sources and receivers are distributed from −2250 to 2250 m with a spac-

ing of 25 m and in the crossline x_2-direction the acquisition ranges from −1250 to 1250 m with a spacing of 50 m. Note that the sampling distances in the inline and crossline directions are not equal, as is often the case for acquisition set-ups in the field. For the modelling of the reflection data, we apply a wavelet with a flat frequency spectrum that introduces ringing in the time domain. This approximates the prerequisite of deconvolving for the source wavelet (Thorbecke *et al.*, 2017). The frequency spectrum of this wavelet is flat between 5 and 25 Hz, and tapered to zero with a cosine window from 5 to 0 Hz and from 25 to 30 Hz. The limited range of the frequency spectrum is chosen for modelling runtime purposes. The reflection data are modelled for 4.0 s with a temporal sampling of 4 ms. This temporal sampling complies with the industry standard, but note that a sampling of 16 ms would suffice. Before the reflection data are compressed, the size of the dataset is 387 GB. After the frequency range is selected and the ZFP compression is applied, the size of the dataset is reduced to 46 GB.

Green's function retrieval

First, we consider the most basic application of the Marchenko method, namely, Green's function retrieval for a point source. To generate the necessary first arrival, we model a point source at $\mathbf{x}_A = (0, 0, 1025)$m using an 11 Hz Ricker wavelet and record the response at the surface of the medium. We separate the first arrival from the coda of the wavefield to
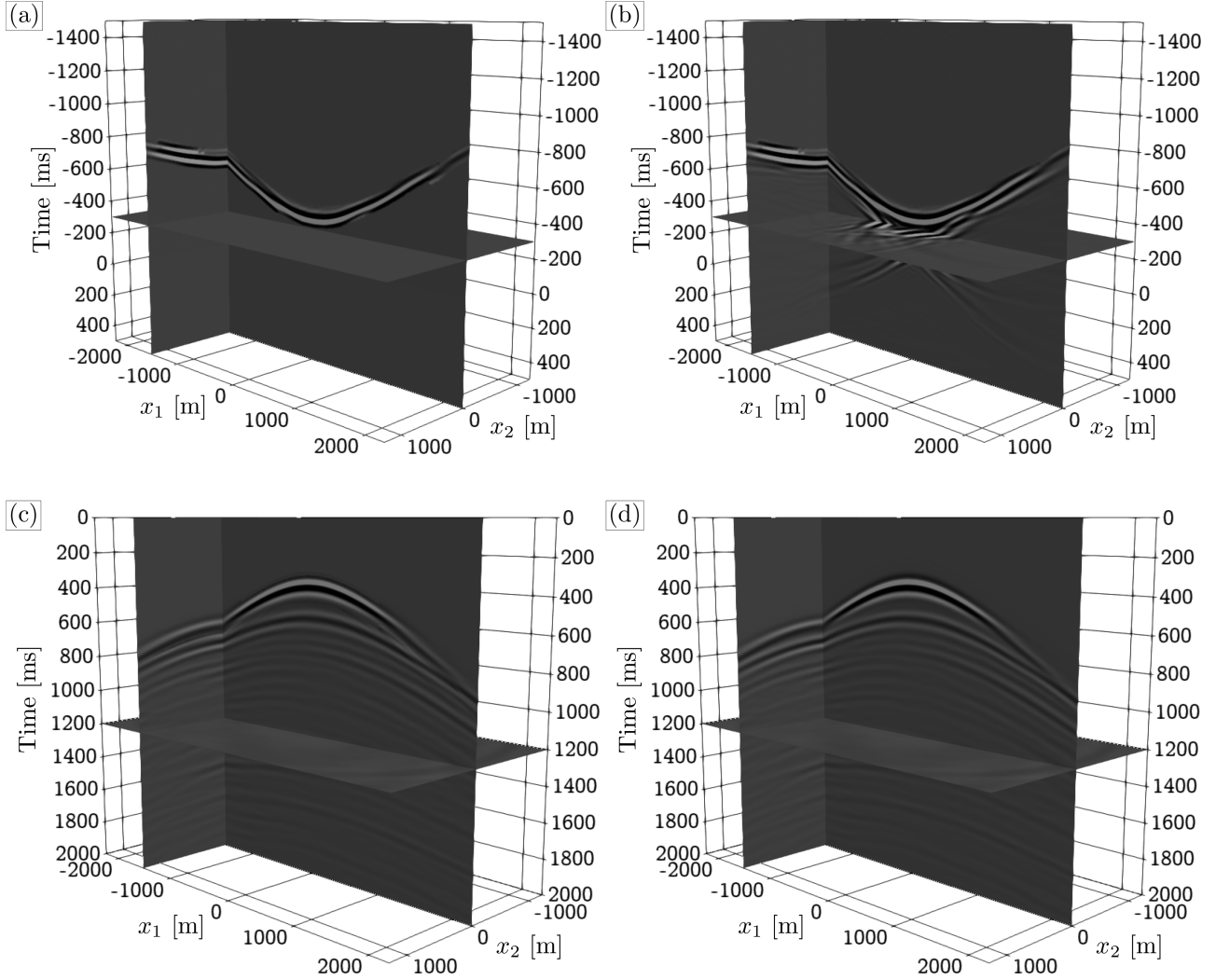
Figure 2 (a) First arrival $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$, modelled in the exact medium, (b) Focussing function $f_2(\mathbf{x}_A, \mathbf{x}_R, t)$ and (c) Green's function $G(\mathbf{x}_R, \mathbf{x}_A, t)$, both obtained through use of the Marchenko method and (d) Reference Green's function $G_{\mathrm{ref}}(\mathbf{x}_R, \mathbf{x}_A, t)$, modelled directly in the exact medium. All wavefields contain an 11 Hz Ricker wavelet, are clipped at the same value and $\mathbf{x}_A = (0, 0, 1025)$m.

obtain $G_d(\mathbf{x}_R, \mathbf{x}_A, t)$. Instead of simply time-reversing this first arrival, we instead invert it to obtain the true $f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_A, t)$, which is shown in Figure 2(a). In the field, it is unlikely that we would be able to achieve this, as one would need to have the exact model available, however, here we wish to demonstrate that the method can theoretically retrieve accurate amplitudes. We use the `marchenko3D` code to retrieve the decomposed focussing functions and Green's functions. We obtain the total focussing function $f_2(\mathbf{x}_A, \mathbf{x}_R, t)$ and the total Green's function $G(\mathbf{x}_R, \mathbf{x}_A, t)$ using equations (3) and (4), respectively, and show these functions in Figure 2(b,c).

The convergence rate of the Marchenko method for a point source at $\mathbf{x}_A = (0, 0, 1025)$m is shown in Figure 3 as the solid red curve. This convergence rate is defined as the rate of the L_1 norm of the total energy in the update N_i from equation (13) to that of the L_1 norm of the total energy in the original update N_0. Note that a significant amount of iterations is required for convergence and that the convergence curve seems to flatten out instead of converging to zero. The low convergence rate has to do with the complexity of the model. Because the model contains a large amount of reflectors and has a great variety of structures, the amount of iterations that are required is high. The maximum convergence is related to the aperture of the reflection response. The convergence to a relatively high energy level indicates that many limited-aperture artefacts are present in the update fields N_i. These limited-aperture
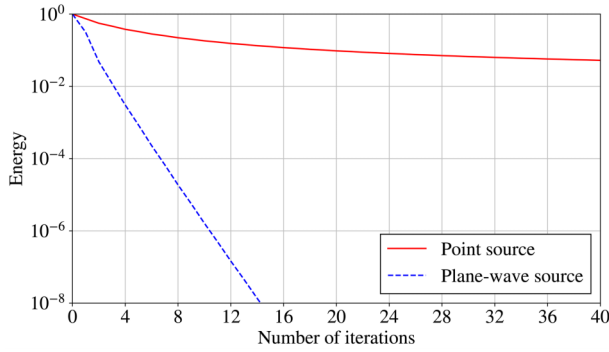
Figure 3 Logarithmic convergence rate $\frac{|N_i|_{L_1}}{|N_0|_{L_1}}$ of Green's function retrieval in the overthrust model for a virtual point source at $\mathbf{x}_A = (0, 0, 1025)$m in solid red and a virtual plane-wave source for $\mathbf{p}_A = (80\,\mu\text{s m}^{-1}, 40\,\mu\text{s m}^{-1}, 1025\,\text{m})$ in dashed blue. The y-axis is clipped at 10^{-8} as we do not expect anything below this energy level to significantly contribute to the final result.

effects are caused by the acquisition footprint, especially in the crossline direction. The complexity of the model causes scattering at large angles, which means that a larger aperture is required to properly capture all the scattered events. This is caused by the assumption of the Marchenko method that $\mathbb{S}_0$ extends infinitely in the horizontal directions, whereas in the field, it will always be limited.

To determine whether the retrieved Green's function is accurate, we compare it to a reference Green's function $G_{\text{ref}}(\mathbf{x}_R, \mathbf{x}_A, t)$. This Green's function was obtained by modelling the source directly at $\mathbf{x}_A = (0, 0, 1025)$m and recording the response at the surface of the medium. It is shown in Figure 2(d). A visual inspection between the retrieved Green's function and the reference Green's function shows that the two wavefields appear to be similar, with a strong correlation between the arrival times and amplitudes of the events. The convergence to the correct solution is further supported by the decrease in energy per update, which can be interpreted from the solid red curve in Figure 3.

To make a more accurate comparison, we show a direct comparison between some of the traces of the two functions in Figure 4. The traces of the retrieved Green's function are shown in black, while the traces of the reference Green's function are shown in dashed red. Note that in this display, we apply a time gain of $t^{1.6}$ to boost the amplitude of the events at later times. When the traces are compared, it can be seen that the match is excellent. The traces in Figure 4(a), which have a fixed offset in the crossline x_2 direction of 500 m, are nearly identical for the two functions. There is a very slight error before the first event, which is caused by the window,

and at high offsets from the source position the quality of the retrieved Green's function decreases slightly. This is because the events that have to be retrieved at this offset are not all captured by the recording aperture. This effect can be seen in Figure 4(b), where the inline x_1 offset is fixed at 1500 m. The errors are slightly larger here because the entire panel is closer to the edge of the aperture. However, overall the match is still very strong and this demonstrates the capability of the Marchenko method to properly retrieve Green's function from surface reflection data.

Marchenko imaging

Green's functions and focussing functions that are retrieved using the Marchenko method can be employed for a variety of schemes. One of the most commonly used applications is imaging, where one aims to obtain the reflectivity of the subsurface in a region of interest. The reflectivity of the medium $\bar{R}_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_A, t)$ at a target depth is related to the decomposed Green's functions as (Amundsen, 2001; Wapenaar *et al.*, 2014b)

$$G^{-,+}(\mathbf{x}_B, \mathbf{x}_R, t)$$
$$= \int_{\mathbb{S}_A} \int_0^{\infty} \bar{R}_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_A, t - t')G^{+,+}(\mathbf{x}_A, \mathbf{x}_R, t')\mathrm{d}t'\mathrm{d}\mathbf{x}_B, \quad (16)$$

where $\mathbb{S}_A$ is a surface inside the medium at the top of the target zone and the reciprocity relations $G^{-,+}(\mathbf{x}_B, \mathbf{x}_R, t) = G^{-,+}(\mathbf{x}_R, \mathbf{x}_B, t)$ and $G^{+,+}(\mathbf{x}_A, \mathbf{x}_R, t) = G^{-,-}(\mathbf{x}_R, \mathbf{x}_A, t)$ have been employed (Wapenaar, 2020). Through the use of a multidimensional deconvolution (MDD) process, the reflectivity can be obtained from the decomposed Green's functions that are obtained using the Marchenko method. By employing this process, the overburden above the target depth is removed and only interactions from the target zone below $\mathbb{S}_A$ are preserved. Furthermore, if there are any errors present in the amplitude of $f^+_{1,d}$, the same errors will be present in both $G^{-,+}$ and $G^{-,-}$, so these errors will be removed for the larger part by the MDD process. However, properly applying the MDD, especially for 3D recording set-ups, is very difficult and prone to errors (Staring *et al.*, 2018). A more stable way of obtaining the reflectivity is through the use of the double-focussing method (van der Neut *et al.*, 2018). Here, no MDD is applied, and the decomposed focussing function is used. The double-focussing method can be expressed as

$$R_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_A, t)$$
$$= \int_{\mathbb{S}_0} \int_0^{\infty} F^+(\mathbf{x}_R, \mathbf{x}_B, t')G^{-,+}(\mathbf{x}_R, \mathbf{x}_A, t - t')\mathrm{d}t'\mathrm{d}\mathbf{x}_R, \quad (17)$$
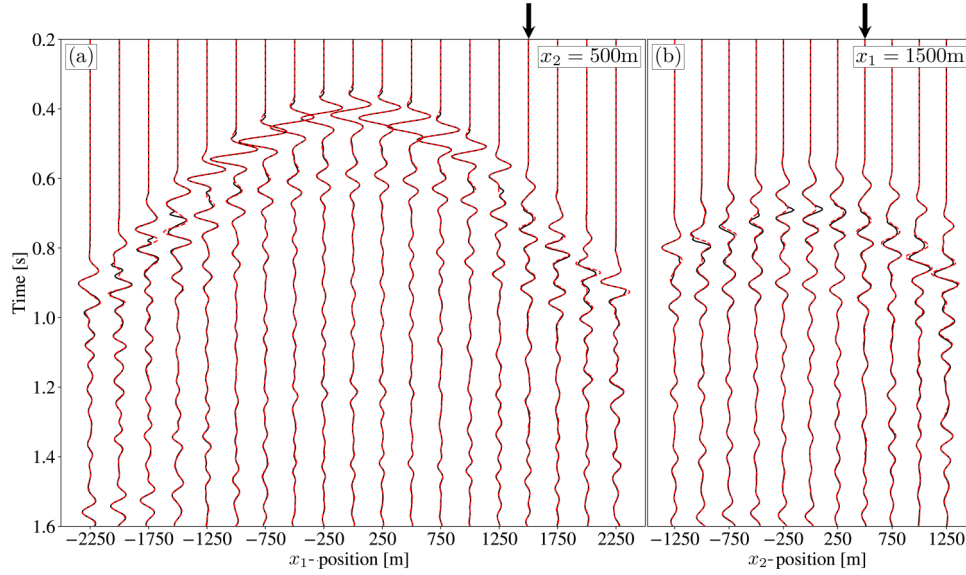
Figure 4 Comparison between the reference Green's function $G_{\text{ref}}(\mathbf{x}_R, \mathbf{x}_A, t)$ in dashed red and Green's function $G(\mathbf{x}_R, \mathbf{x}_A, t)$ obtained through the use of the Marchenko method in solid black for (a) $x_2 = 500$ m and (b) $x_1 = 1500$ m and $\mathbf{x}_A = (0, 0, 1025)$m. All wavefields contain an 11 Hz Ricker wavelet and have a gain applied of $t^{1.6}$ for display purposes. The black arrows on top indicate where the panels in (a) and (b) intersect.

with

$$\partial_t F^+(\mathbf{x}_R, \mathbf{x}_B, t) = -\frac{2}{\rho}\partial_3 f_1^+(\mathbf{x}_R, \mathbf{x}_B, t). \tag{18}$$

In (17), $R_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_A, t) = G^{-,+}(\mathbf{x}_B, \mathbf{x}_A, t)$. Unlike the MDD approach, the double-focussing approach does not remove the overburden, but only redatums the sources and receivers from the surface of the medium to the target depth. Additionally, because there is no longer a deconvolution employed, the amplitude errors associated with $f_{1,d}^+$ are not removed. This means that we will not obtain the true amplitudes of the subsurface reflectors, however, we obtain the location of the reflectors, which means that we can retrieve an accurate structural image of the subsurface. The entire process is stable however, because no deconvolution is employed. To obtain an image in a location inside the medium, without any influence of internal multiples, one can select the zero time, zero offset of the reflectivity in that location. This is defined as

$$r_{\text{im}}(\mathbf{x}_B) = R_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_B, t = 0). \tag{19}$$

In this equation, the location of the virtual source and the virtual receiver coincides. For the imaging, only the first-time sample is of interest. Hence, equation (17) can be adjusted because the coda of the focussing function is meant to eliminate artefacts in the coda of the redatumed reflection response. Therefore, equation (17) can be applied without the coda of

the focussing function,

$$r_{\text{im}}(\mathbf{x}_B) = R_{\text{tar}}(\mathbf{x}_B, \mathbf{x}_B, t = 0)$$
$$= \int_{\mathbb{S}_0}\int_0^\infty F_d^+(\mathbf{x}_R, \mathbf{x}_B, t')G^{-,+}(\mathbf{x}_R, \mathbf{x}_B, -t')\mathrm{d}t'\mathrm{d}\mathbf{x}_R, \tag{20}$$

with

$$\partial_t F_d^+(\mathbf{x}_R, \mathbf{x}_B, t) = -\frac{2}{\rho}\partial_3 f_{1,d}^+(\mathbf{x}_R, \mathbf{x}_B, t). \tag{21}$$

In equation (20), only the zero-time sample is retrieved, however, as stated before, the errors will be present in the coda of the redatumed reflection response and the zero-time sample will be unaffected. The result can therefore also be used for the purpose of imaging without artefacts caused by internal multiples. Note that, similar to equation (17), any errors in $f_{1,d}^+$ are not accounted for. The double-focussing method has been successfully employed for the purpose of imaging using 3D acquisitions in the field, for an example, see Staring and Wapenaar (2020a).

To test the imaging potential of our `marchenko3D` program, we employ the double-focussing method. The required Green's functions and focussing functions are obtained by the Marchenko method, however, we make three alterations to the approach that was used for Green's function retrieval. Firstly, instead of using the exact medium for estimating our first arrival, we apply a smoothing algorithm to create a background velocity model. Secondly, the first arrival is now modelled
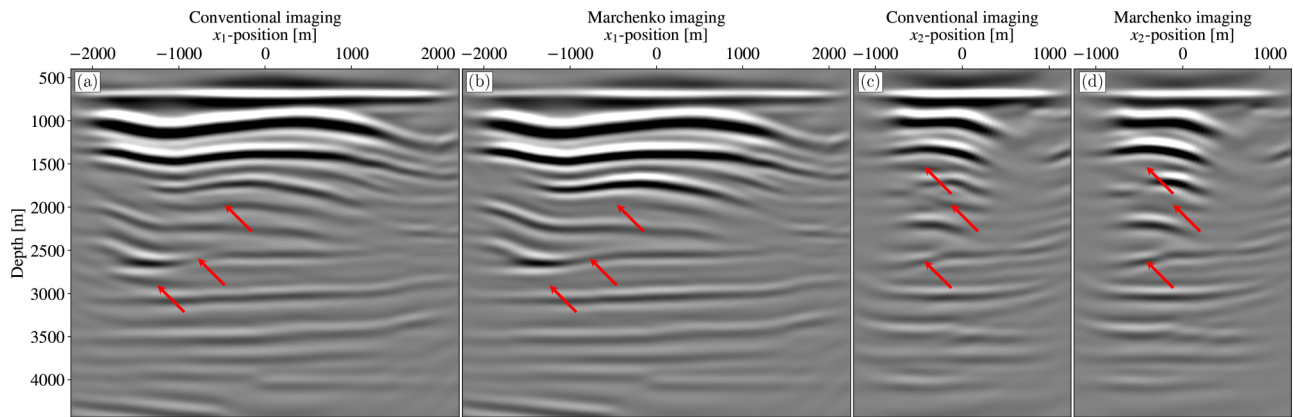
Figure 5 Image of the overthrust model along a fixed x_2 value of 0 m using (a) conventional imaging and (b) Marchenko imaging after 30 iterations, and image of the overthrust model along a fixed x_1 value of 0 m using (c) conventional imaging and (d) Marchenko imaging after 30 iterations. The locations of artefacts that are attenuated by the Marchenko imaging are indicated by the red arrows.

using a 3D Eikonal solver instead of using a finite-difference method. Finally, we no longer invert the first arrival, but simply take the time reversal, as the changes to the modelling process of the first arrival will produce amplitude errors regardless. The changes in our approach are made to simulate field conditions and to keep the modelling time feasible. The Marchenko method needs to be performed for every location in the subsurface where an image is desired and therefore, the computational costs for producing an image are high.

Note that while the use of an Eikonal solver decreases the modelling time of the first arrival significantly, the use of finite-difference modelling would be a more accurate process, as an Eikonal solver does not represent the full physics of the wave equation. In media with strong velocity variations, the first arriving event may not be a purely propagating wavefield and may contain a refracted wave. The image obtained using an Eikonal solver will therefore be an approximation, however, it is computationally more feasible to apply the method like this in practice for a large amount of focal locations.

We obtain the image of two cross sections of the overthrust model, one inline cross section for a fixed x_2-offset of 0 m and one crossline cross section for a fixed x_1-offset of 0 m. The two cross sections intersect each other in their respective centres. The focal locations are placed along a depth range of 400–4400 m with a sampling of 25 m. The range of the focal points in the inline direction is from −2250 to 2250 m, with a sampling of 25 m and in the crossline direction the range is set from −1250 to 1250 m with a sampling distance of 50 m. Similar to Green's function retrieval, we perform 40 iterations for each focal point. After the focussing function and

Green's function for each focal location are obtained using the Marchenko method and equations (1) and (2), they are used in equation (17) to obtain the local reflectivity. The zero-time sample is then extracted to obtain the reflectivity at that exact location, following equation (19). For comparison, we also performed this approach using equation (20) instead of (17), however, for the imaging result, there were no differences, as is expected from the theory. Figure 5(b) shows the reflectivity for each focal location for the inline direction and 5(d) shows the same for the crossline direction. Conventional images for the inline and crossline directions are shown in Figure 5(a,c), respectively. As can be seen from the figure, the subsurface is complex and hard to resolve. Due to the small frequency bandwidth, the resolution of the images is limited, however, there are still artefacts present caused by the internal multiples, as indicated by the red arrows. The Marchenko imaging attenuates these artefacts, which shows that even on complex 3D models, our `marchenko3D` code can produce good results. The downside is that the computational costs remain high, even if compressed reflection data are used. Imaging all the points in the inline direction for a single depth level takes around 19,800 seconds or five and a half hours, running in parallel on 40 Intel E5-2560 cores with a clock speed of 2.3 GHz on a node using 256 GB of 2133 MHz RAM. Imaging the entire 3D medium using this approach may therefore prove too demanding unless a very powerful computing machine is available. A potential solution could be the use of graphics processing units (GPUs) for the implementation of the Marchenko method, which has shown promise for 2D Marchenko applications, for example in Koehne *et al.* (2021).

Plane-wave Green's function retrieval

Due to the high computational costs of creating 3D images in the subsurface, it is hard to efficiently apply imaging in practice using virtual point sources. As an alternative to the standard Marchenko imaging, one can image an entire depth level at once instead of just a single point. This can be achieved through the use of virtual plane-wave sources instead of virtual point sources. The idea of combining plane waves with the Marchenko method was first proposed by Meles *et al.* (2018), who applied the method with success in 2D settings. Here, we wish to implement the approach for 3D settings. The concept of the plane-wave method is that the focussing function no longer focusses to a single focal location in the subsurface, but rather to a focal depth. This gives a limited amount of information because only certain angles of the wavefield are considered. The plane-wave focussing functions $\tilde{f}_1^{\pm}(\mathbf{x}, \mathbf{p}_A, t)$ are related to the focussing functions as

$$\tilde{f}_1^{\pm}(\mathbf{x}, \mathbf{p}_A, t) = \int_{\mathbb{S}_A} f_1^{\pm}(\mathbf{x}, \mathbf{x}_A, t - \mathbf{p} \cdot \mathbf{x}_{H,A}) d\mathbf{x}_A, \tag{22}$$

where $\mathbf{x}_{H,A} = (x_{1,A} - x_{1,c}, x_{2,A} - x_{2,c})$, $\mathbf{p} = (p_1, p_2)$ and $\mathbf{p}_A = (\mathbf{p}, x_{3,A})$. Here, $p_1 = \sin\alpha\cos\beta/c$ and $p_2 = \sin\alpha\sin\beta/c$ are the horizontal ray parameters, where α is the dip angle, β is the azimuth angle and $c = c(\mathbf{x})$ is the propagation velocity of the medium. $x_{1,c}$ and $x_{2,c}$ are the horizontal coordinates of the central location of the plane-wave source. By including these coordinates in the definition of $\mathbf{x}_{H,A}$, we ensure that the centre of the plane-wave source always has an emission time of $t = 0$. Equation (22) states that the plane-wave focussing function $\tilde{f}_1^{\pm}$ is the integral of the focussing functions $f_1^{\pm}$ over all possible focal points $\mathbf{x}_A$ at $\mathbb{S}_A$ for a certain dip and azimuth given by the ray parameters. As such, $\tilde{f}_1^{\pm}(\mathbf{x}, \mathbf{p}_A, t)$ does not focus to a single location, but instead focusses as a dipping wave, dictated by $\mathbf{p}$, to a single depth level $x_{3,A}$. In case the ray parameters are zero, this wave is horizontal instead of dipping.

We also define the plane-wave Green's functions $\tilde{G}^{-,\pm}(\mathbf{x}_R, \mathbf{p}_A, t)$ which are related to Green's functions

$$\tilde{G}^{-,\pm}(\mathbf{x}_R, \mathbf{p}_A, t) = \int_{\mathbb{S}_A} G^{-,\pm}(\mathbf{x}_R, \mathbf{x}_A, t - \mathbf{p} \cdot \mathbf{x}_{H,A}) d\mathbf{x}_A. \tag{23}$$

The plane-wave Green's functions behave similar to the regular Green's functions, however, instead of being the impulse response of a point source, the plane-wave Green's functions are the response of a medium to a plane-wave source at $x_{3,A}$, which generates a dipping wave as dictated by the ray parameters.

The plane-wave versions of the focussing functions and Green's functions can be combined with the equations (1) and

(2) simply by applying the integration over all possible focal points for a set of ray parameters

$$\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}'_A, t) + \tilde{f}_1^{-}(\mathbf{x}_R, \mathbf{p}'_A, t)$$
$$= \int_{\mathbb{S}_0} \int_0^{\infty} R(\mathbf{x}_R, \mathbf{x}_S, t') \tilde{f}_1^{+}(\mathbf{x}_S, \mathbf{p}'_A, t - t') dt' d\mathbf{x}_S, \tag{24}$$

$$\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A, -t) + \tilde{f}_1^{+}(\mathbf{x}_R, \mathbf{p}'_A, t)$$
$$= \int_{\mathbb{S}_0} \int_{-\infty}^0 R(\mathbf{x}, \mathbf{x}_S, -t') \tilde{f}_1^{-}(\mathbf{x}_S, \mathbf{p}'_A, t - t') dt' d\mathbf{x}_S. \tag{25}$$

In these equations, we define $\mathbf{p}'_A = (-\mathbf{p}, x_{3,A})$, which all plane-wave wavefields contain with the exception of $\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A, -t)$. This is because this particular wavefield is time-reversed in equation (2). The change to plane waves has not affected the reflection response, only the focussing functions and Green's functions, which means that no changes to the reflection response have to be made. Similar to the original equations, the Green's functions in the system of equations (24) and (25) can be suppressed through the use of a windowing function. The change to plane waves does affect the causality relations however, especially because $\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A, t)$ and $\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}'_A, t)$ are dipping at different angles. Hence, the time window needs to be adjusted so the dipping plane waves are properly handled (Meles *et al.*, 2020). These windows are defined as

$$\tilde{\Theta} = \tilde{\Theta}(\mathbf{x}_R, \mathbf{p}_A, t) = \theta(t - t_a) - \theta(t - t_b),$$
$$t_a = -\tilde{t}_d(\mathbf{x}_R, \mathbf{p}_A) + \epsilon, \tag{26}$$
$$t_b = \tilde{t}_d(\mathbf{x}_R, \mathbf{p}'_A) - \epsilon,$$

where $\tilde{t}_d(\mathbf{x}_R, \mathbf{p}'_A)$ is the first possible arriving event of $\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}'_A, t)$ and $-\tilde{t}_d(\mathbf{x}_R, \mathbf{p}_A)$ is the last arriving event of $\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A, -t)$. As a result, the windowing function is no longer symmetrical in time, instead, the window has parallel planes. For a further derivation of this window, see Wapenaar *et al.* (2021). Applying this window to equations (24) and (25) results in the plane wave coupled Marchenko equations

$$\tilde{f}_1^{-}(\mathbf{x}_R, \mathbf{p}'_A, t) = \tilde{\Theta} \int_{\mathbb{S}_0} \int_0^{\infty} R(\mathbf{x}_R, \mathbf{x}_S, t') \tilde{f}_1^{+}(\mathbf{x}_S, \mathbf{p}'_A, t - t') dt' d\mathbf{x}_S, \tag{27}$$

$$\tilde{f}_1^{+}(\mathbf{x}_R, \mathbf{p}'_A, t) - \tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}'_A, t)$$
$$= \tilde{\Theta} \int_{\mathbb{S}_0} \int_{-\infty}^0 R(\mathbf{x}_R, \mathbf{x}_S, -t') \tilde{f}_1^{-}(\mathbf{x}_S, \mathbf{p}'_A, t - t') dt' d\mathbf{x}_S. \tag{28}$$

Note that (27) and (28) have a very similar structure to equations (6) and (7). They can therefore be solved with the

same Marchenko method as was used for the point-source focussing functions. The only things that have to be adjusted are the input and the windowing function. The first estimation of the focussing function can once again be estimated using the direct arrival

$$\tilde{f}_{1,0}^{+}(\mathbf{x}_R, \mathbf{p}_A', t) = \tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}_A', t). \qquad (29)$$

However, because we cannot compute this first arrival directly, we have to model the direct arrival of Green's function and time-reverse it. The plane wave is modelled by placing point sources along a horizontal plane in the medium of interest and emitting with a time delay, so that the emission time of a location along the horizontal plane is $\delta(t - \mathbf{p} \cdot \mathbf{x}_{H,A})$. By taking the direct arrival of Green's function, we obtain $\tilde{G}_d^{+}(\mathbf{x}_R, \mathbf{p}_A, t)$, which we can then relate to the direct arrival of the focussing function as

$$\tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}_A', t) = \tilde{G}_d(\mathbf{x}_R, \mathbf{p}_A, -t). \qquad (30)$$

Note that according to equation (30), the original ray parameters that were used to model the plane-wave source and obtain $\tilde{G}_d(\mathbf{x}_R, \mathbf{p}_A, t)$ are the opposite of the ray parameters that are included in the first estimation of the focussing function $\tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}_A', t)$. Hence, when we use equation (25) to obtain $\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A', t)$, it is dipping according to the original ray parameters and when we use equation (24) to obtain $\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}_A', t)$ it is dipping according to the opposite of the original ray parameters.

We retrieve Green's function in the subsurface at a depth level of $x_3 = 1025$ m. We use a plane wave that is dipping in both the inline and the crossline direction as determined by ray parameters $\mathbf{p} = (80, 40)\mu$s m^{-1}. Because the velocity of the medium varies at this depth, we cannot convert this to a single angle, however, the majority of this depth level corresponds to a velocity of 4500 m s^{-1}. Using this velocity we calculate a representative dip and azimuth angle, namely $\alpha = 23.6°$ and $\beta = 26.6°$. In order to obtain the full Green's function, we run the Marchenko method twice, once using $\mathbf{p}$ to obtain $\tilde{G}^{-,-}(\mathbf{x}_R, \mathbf{p}_A, t)$ and once using $-\mathbf{p}$ to obtain $\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}_A, t)$, and use these results in equation (4). The initial estimation of the focussing function $\tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}_A', t)$, which we obtained in the exact model for accurate amplitudes, is shown in Figure 6(a), while the full focussing function $\tilde{f}_2(\mathbf{p}_A', \mathbf{x}_R, t)$ is shown in (b). To construct this focussing function, we apply the plane-wave definition to equation (3):

$$\tilde{f}_2(\mathbf{p}_A', \mathbf{x}_R, t) = \tilde{f}_1^{+}(\mathbf{x}_R, \mathbf{p}_A', t) - \tilde{f}_1^{-}(\mathbf{x}_R, \mathbf{p}_A, -t). \qquad (31)$$

The convergence of the Marchenko method using $\mathbf{p}$ is shown as the dashed blue curve in Figure 3. If $-\mathbf{p}$ is used, the graph is

visually identical. Note that the method converges much faster if a plane-wave source is used instead of a point source, this is caused by the fact that only certain angles are covered by the dipping plane wave, hence the amount of propagation directions that needs to be resolved is much lower. The full Green's function $\tilde{G}(\mathbf{x}_R, \mathbf{p}_A, t)$, which is obtained through the use of the two runs of the Marchenko method is shown in Figure 6(c).

We compare the retrieved Green's function to a reference Green's function $\tilde{G}_{ref}(\mathbf{x}_R, \mathbf{p}_A, t)$, that was directly modelled in the medium and is shown in Figure 6(d). Visually, the two Green's functions appear to be very similar. There are some slight differences caused by curving events originating from the edges of the aperture. These are caused by the modelling of the plane-wave source as opposed to a point source. The amplitude along the wavefront is nearly constant for a plane-wave source, while for a point source it decreases near the edges. Therefore, the edge effects are more pronounced for the plane-wave sources. When the first arrival is isolated from the coda, these effects are largely removed because they are present in the coda, hence the difference in the results.

To further investigate the accuracy of the retrieved Green's function, we construct a similar figure as Figure 4, where we display traces of the two Green's function next to each other. The result is shown in Figure 7. The accuracy has decreased in comparison with the standard point-source Green's function retrieval, which is also a result from the plane-wave set-up. The errors caused by placing the source near the edges contribute more to the final result than when point sources are used, hence the decrease in quality. Overall, the match is still very strong, for both the amplitude and the phase of the events.

Plane-wave imaging

We perform plane-wave imaging through the double-focussing method, similar to equation (17). We apply the integral over all possible focal points, following (22) and (23),

$$\tilde{R}_{tar}(\mathbf{x}_B, \mathbf{p}_A, t)$$
$$= \int_{\mathbb{S}_0} \int_0^{\infty} F^{+}(\mathbf{x}_R, \mathbf{x}_B, t')\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}_A, t - t')dt'd\mathbf{x}_R. \qquad (32)$$

In equation (32), $\tilde{G}^{-,+}$ and $\tilde{R}_{tar}$ are both plane-wave responses, while F^{+} still is the focussing function related to a focal point instead of a focal depth. We can use the redatumed receiver response to formulate an image condition for each location in the surface

$$\tilde{r}_{im}(\mathbf{x}_B, \mathbf{p}_A) = \tilde{R}_{tar}(\mathbf{x}_B, \mathbf{p}_A, t = \mathbf{p} \cdot \mathbf{x}_{H,B}). \qquad (33)$$
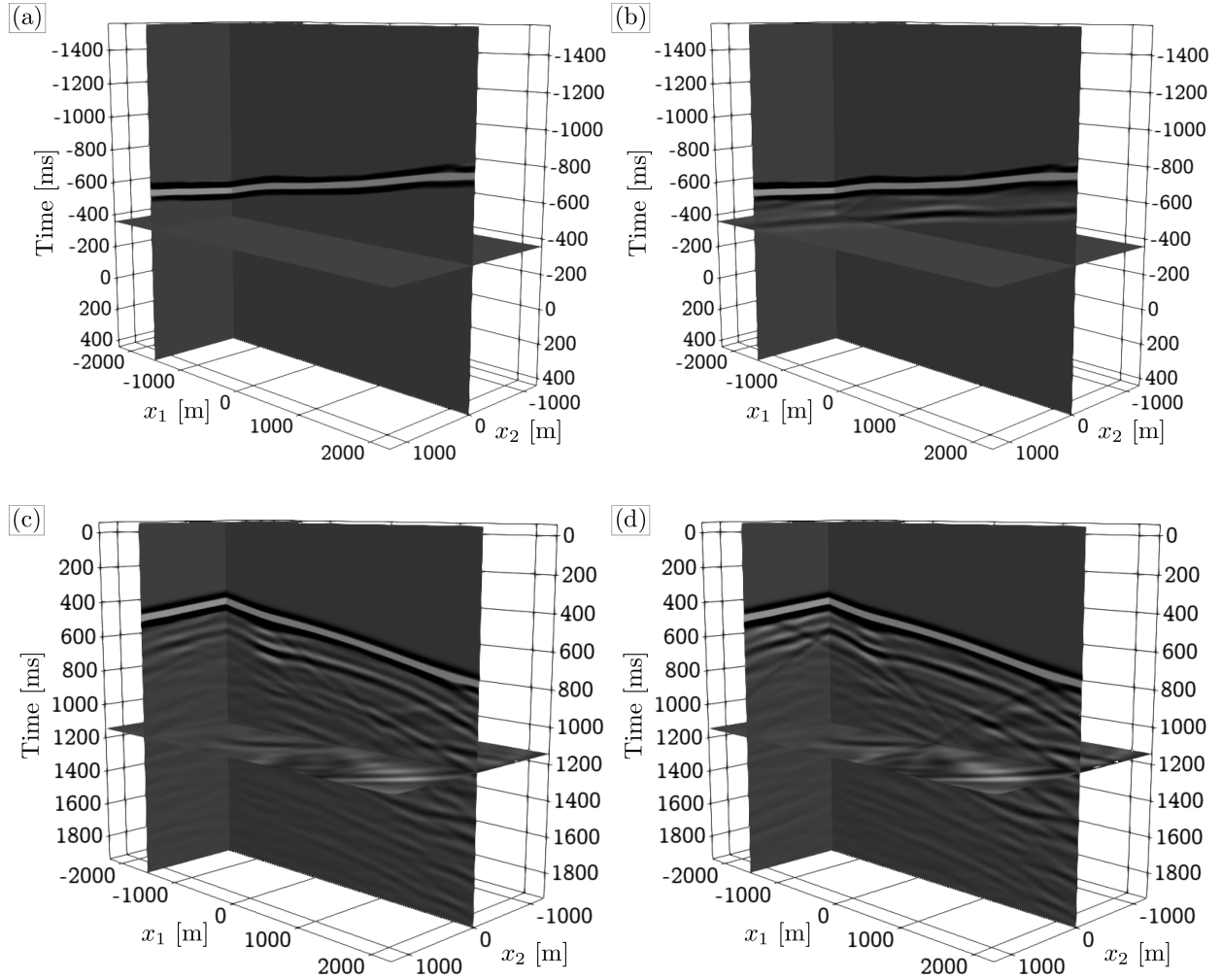
Figure 6 (a) First arrival $\tilde{f}_{1,d}^{+}(\mathbf{x}_R, \mathbf{p}_A', t)$, modelled in the exact medium, (b) focussing function $\tilde{f}_2(\mathbf{p}_A', \mathbf{x}_R, t)$ and (c) Green's function $\tilde{G}(\mathbf{x}_R, \mathbf{p}_A, t)$, both obtained through use of the Marchenko method and (d) reference Green's function $\tilde{G}_{\mathrm{ref}}(\mathbf{x}_R, \mathbf{p}_A, t)$, modelled directly in the exact medium. All wavefields contain an 11 Hz Ricker wavelet, are clipped at the same value and $\mathbf{p}_A = (80\,\mu\mathrm{s}\,\mathrm{m}^{-1}, 40\,\mu\mathrm{s}\,\mathrm{m}^{-1}, 1025\,\mathrm{m})$.

Here, $\tilde{r}_{\mathrm{im}}(\mathbf{x}_B, \mathbf{p}_A)$ is the reflectivity in the subsurface at location $\mathbf{x}_B$ given a certain angle dictated by $\mathbf{p}_A$. The time sample that is selected from the redatumed receiver response is changed, depending on the angle of the incoming wave, which is required due to the time-delay associated with the angle of the wavefront. By choosing different values for $\mathbf{p}_A$, a reflection image can be obtained for various angles.

Note that in order to obtain the correct angle-dependent reflectivity, both $F^+(\mathbf{x}_R, \mathbf{x}_B, t)$ and $\tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}_A, t)$ need to have accurate amplitudes. As we mentioned, this is hard to achieve using the Marchenko method in practice. As a result, we will not retrieve the true angle-dependent reflectivity of the subsurface, however, we can obtain a structural image of

the subsurface related to specific ray parameters. In order to obtain the true angle-dependent reflectivity, equation (32) can be rewritten as a deconvolutional process, similar to equation (16), but this is beyond the scope of this paper, as we are interested in efficiently obtaining a structural image of the subsurface.

The procedure using equation (32) has a notable disadvantage in that not only the plane-wave response needs to be obtained, but also the focussing function for each focal point in the subsurface, which means that the computational effort compared to using equation (17) in fact increases. However, if we instead use an approach similar to equation (20), we can avoid the computation of the full focussing function. We
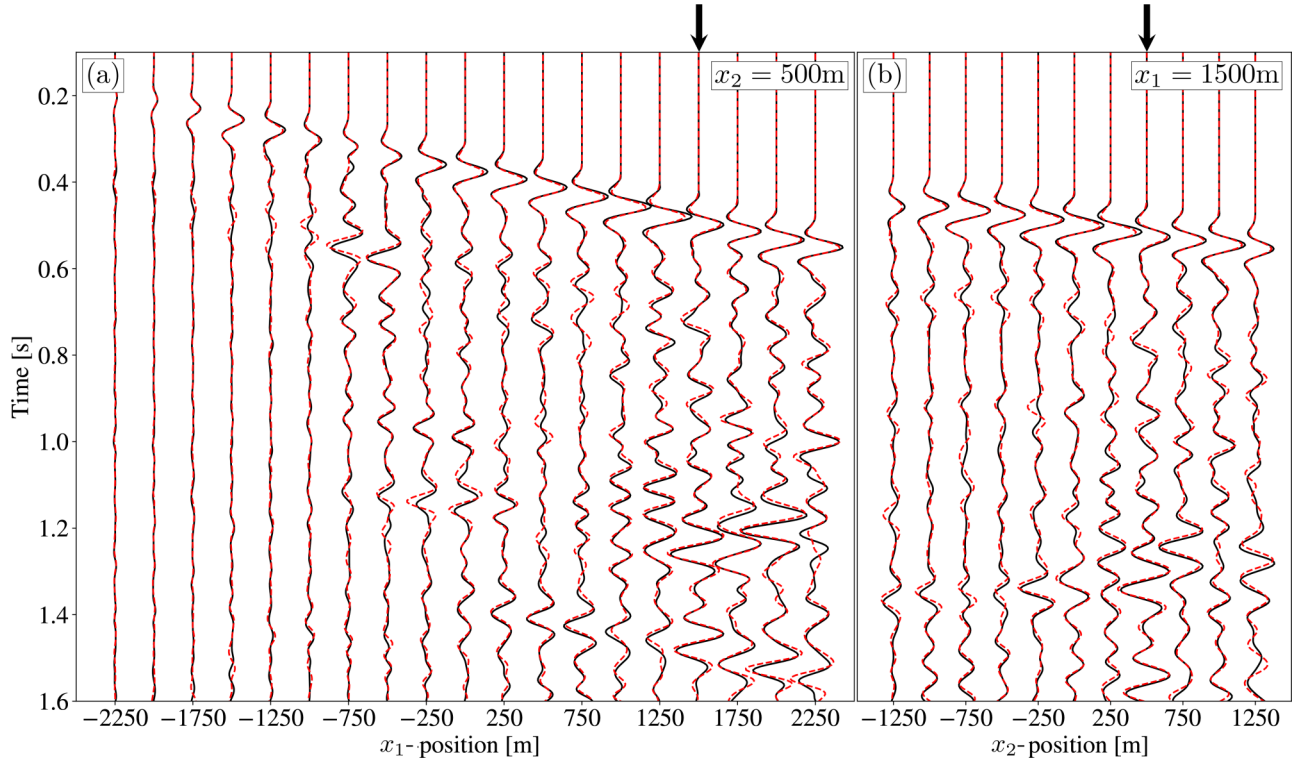
Figure 7 Comparison between the reference Green's function $\tilde{G}_{\text{ref}}(\mathbf{x}_R, \mathbf{p}_A, t)$ in dashed red and Green's function $\tilde{G}(\mathbf{x}_R, \mathbf{p}_A, t)$ obtained through the use of the Marchenko method in solid black for (a) $x_2 = 500$ m and (b) $x_1 = -1500$ m and $\mathbf{p}_A = (80\,\mu s\,m^{-1}, 40\,\mu s\,m^{-1}, 1025m)$. All wavefields contain an 11 Hz Ricker wavelet and have a gain applied of $t^{1.6}$ for display purposes. The black arrows on top indicate where the panels in (a) and (b) intersect.

rewrite equation (32) combined with (33) as

$$\tilde{r}_{\text{im}}(\mathbf{x}_B, \mathbf{p}_A) = \tilde{R}_{\text{tar}}(\mathbf{x}_B, \mathbf{p}_A, t = \mathbf{p} \cdot \mathbf{x}_{H,B})$$

$$= \int_{\mathbb{S}_0} \int_0^{\infty} F_d^+(\mathbf{x}_R, \mathbf{x}_B, t') \tilde{G}^{-,+}(\mathbf{x}_R, \mathbf{p}_A, \mathbf{p} \cdot \mathbf{x}_{H,B}$$

$$-t')\mathrm{d}t'\mathrm{d}\mathbf{x}_R. \tag{34}$$

We model only the first arrival for each point we want to image and do not need to apply the Marchenko method to obtain the full focussing function, saving considerable computational costs.

To demonstrate the results of the method, we apply equation (33) to the overthrust model. We retrieve a horizontal plane wave, that is, $\mathbf{p} = (0, 0)$, and image the same section as we did in Figure 5(b). The first arrival for the plane wave $\tilde{f}_{1,d}^+$ is obtained through the use of finite-difference modelling, while the first arrival $f_{1,d}^+$ that is used in equation (21) and subsequently in equation (34) is obtained using the Eikonal solver. The result is shown in Figure 8(e). For comparison, for the standard Marchenko image we perform the Marchenko method for 161 depths and 181 lateral positions, for a total

of 29,141 different points, while for the plane-wave imaging, we perform the Marchenko method only for the 161 depths. The computation times for all these depths using plane-wave sources is around five and a half hours, which is similar to the computational time of all focal points in the inline direction for a single depth when the standard Marchenko imaging is employed. Because the first arrivals that are used for the imaging step are all obtained using an Eikonal solver, the computational costs of retrieving these events are negligible compared to the computational costs of the Marchenko method. The image that is obtained using the horizontal plane wave contains many similar reflectors as were seen in the standard Marchenko image in Figure 5(b). However, because we are only considering reflectivity that is related to certain ray parameters, the reflectors are not as well resolved as the standard Marchenko image.

We consider additional angles of reflectivity in order to resolve the events in the plane-wave image more clearly (Meles *et al.*, 2018). These images are shown in Figure 8, for varying ray parameters. The ray parameter in the inline direction takes values of -80, 0 and $80\,\mu s\,m^{-1}$ and the ray
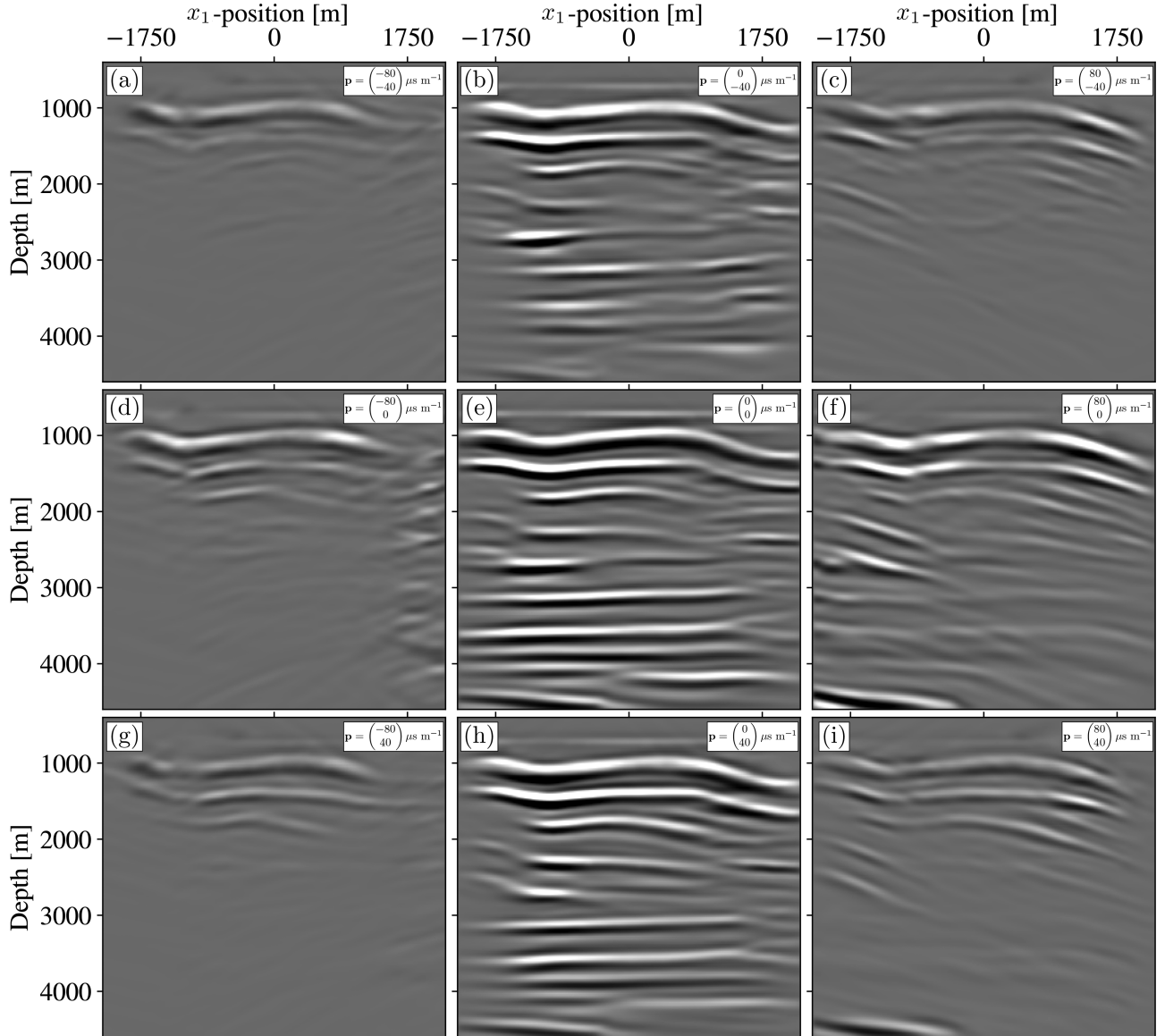
Figure 8 Angle-dependent reflectivity of the subsurface obtained through the use of Marchenko imaging for different ray parameters.

parameter in the crossline direction takes values of −40, 0 and 40 μs m^{-1}. The central column shows the result when the inline ray parameter is zero and we can see that these panels contain the most energy. This indicates that a large amount of the reflectors have a small dipping angle in the inline direction. The deepest part of the image also contains less energy when the crossline ray parameter is negative rather than positive, which indicates that the reflectors at this depth are dipping positively in the crossline direction. The images constructed with a negative ray parameter in the inline direction contain considerably less energy than when the ray parameter is positive. The energy is mostly constrained to the shallow part of

the image and the edges. Especially near the edges, it is hard to determine whether this is caused by the edge of the aperture or by genuine reflector direction. The images that have a dip in the inline direction contain more energy and have reflectors at various depths and at distances away from the edge.

Figure 8 gives insight into the reflector dip of the subsurface. We combine these results together to create a single image, which is shown in Figure 9(a). For comparison, we plot the result of the standard Marchenko image in Figure 9(b). Most of the major reflectors are present at the same locations and with similar reflector strength in both imaging
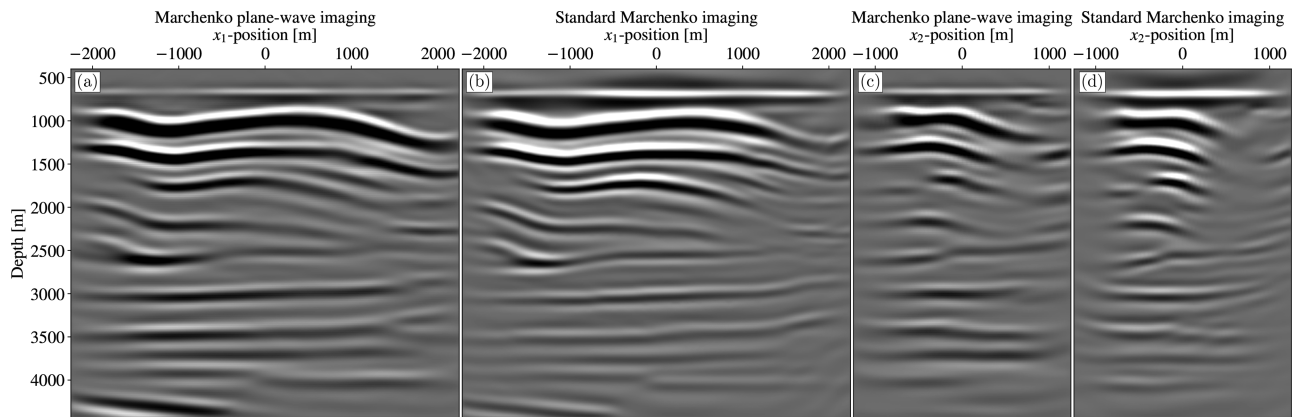
Figure 9 Image of the overthrust model along a fixed x_2 value of 0 m using (a) plane-wave and (b) standard Marchenko imaging after 30 iterations, and image of the overthrust model along a fixed x_1 value of 0 m using (c) plane-wave and (d) standard Marchenko imaging after 30 iterations. The image in (a) is constructed by combining the plane-wave images in Figure 8 and the image in (c) is constructed in a similar way.

results. The result of the standard Marchenko imaging is better resolved because this approach takes into account all angles of reflections within the range that is measured by the aperture. While the result of the plane-wave images is of a lesser quality, the computational costs are significantly lower, even if several images associated with different ray parameters have to be retrieved. We repeat the retrieval for both the standard Marchenko approach and plane-wave Marchenko approach to retrieve the image in the crossline direction. The result for the plane-wave Marchenko image is shown in Figure 9(c) and for the standard Marchenko image in Figure 9(d). From these images, we can draw similar conclusions as we did from the images in the inline direction. The same events are resolved however, the resolution of the standard Marchenko image is higher than that of the plane-wave image. It should also be noted that the plane-wave image does not contain any of the artefacts that are present when conventional imaging is used, as the plane-wave Marchenko method still attenuates all the internal multiple artefacts.

DISCUSSION

Our examples have all been performed using synthetic data. For field data, additional complications would arrive, for example subsampling in space of the reflection data. The crossline spacing of the sources and receivers tends to be much sparser than the inline spacing, so some kind of interpolation needs to be applied (Staring and Wapenaar, 2020b). In case the data have irregular sampling, the standard Marchenko scheme can be adjusted to handle this (van IJsseldijk and Wapenaar, 2020; Haindl *et al.*, 2021). In case the data are well sampled, it has been shown that the size of the data volume can become much larger than has been considered in this paper, to the size of hundreds of terabytes. While our compression-based Marchenko method could assist with using this data volume size, it is advisable to consider only a subset of the data in this case.

As an additional consequence to the large datasize in 3D, the computational costs of retrieving virtual responses using the Marchenko method increase significantly. While our approach can assist with reducing the time reading data from and writing data to disk, the actual computational process is not affected. Possible solutions are the efficient use of distributed computing (Ravasi and Vasconcelos, 2021) or the use of graphics processing units (GPUs) (Koehne *et al.*, 2021). While the use of GPUs for efficient computation has gained much attraction in recent years and is an exciting development, the storage space on GPUs is limited. In the case of large datasets, such as the ones required for the Marchenko method, a substantial number of GPUs would be required. Furthermore, the transfer of data to and from the GPUs can be a significant bottleneck. Further technological developments of GPUs could make this approach more viable in the future. Current state-of-the-art GPUs would require efficient computation distribution to minimize idle times or an effective compression algorithm could be developed that would allow for the application of the Marchenko method without decompression.

One of the other approaches for computational efficiency that we presented in this paper is the use of virtual plane-wave sources instead of virtual point sources. It should be noted that this approach is a most effective alternative for retrieving the response of a large number of virtual source points in the subsurface that are on the same depth level. On the

other hand, when this number is sufficiently small, using the virtual point sources yields a higher accuracy and more angle information. Datasets recorded by dense ocean bottom node arrays could benefit greatly from the plane-wave approach, however, even for more traditional streamer set-ups, the approach can yield benefits, especially for large 3D apertures.

The imaging strategy that we have applied in this paper is based on retrieving the local reflectivity, either at a point or a depth level, and using this to construct the structural image. An alternative to this approach would be to redatum the sources and receivers to a single depth level using the multi-dimensional deconvolution (MDD) approach, so that a reflection response of the target zone without overburden contributions is constructed. A more straightforward imaging technique such as reverse time migration can then be used to image the target zone. This approach works best if the strongest reflectors are located in the overburden and the target zone is well defined. The area below the overburden can still cause internal multiples however, so care needs to be taken when this approach is employed.

Our data are compressed through the use of the ZFP algorithm, which offers accuracy based on tolerance. We used a value of 1×10^{-7}, which means that we allow errors in the compression that are seven orders of magnitude lower than the maximum amplitude. As we are dealing with single-point precision, this gives us the maximum accuracy possible. By setting the tolerance to a higher value, the data can be compressed further, however, this will introduce numerical errors into the compressed data. Depending on the reduction of data space and the desired accuracy, this may be acceptable for certain applications, however, this is beyond the scope of this paper.

CONCLUSIONS

We have discussed the implementation of the 3D Marchenko method while taking into account several practical considerations. The standard Marchenko implementation requires high-quality reflection data and while it is possible to achieve this, the size of the data volume in 3D is considerably large. In order to handle the large data volume and reduce the computational costs of the 3D Marchenko method, we used a ZFP compression algorithm. This reduced the size of the data and limited the loading time. The 3D Marchenko method was then used to retrieve Green's functions in the subsurface. The result closely matched a reference result and as such the retrieved Green's functions could be used for the purpose of imaging. While the result of the imaging shows attenuation

of artefacts related to internal multiples, the computational costs of imaging a large section of the subsurface in 3D was high.

To limit the computational costs of 3D Marchenko imaging further, virtual plane waves were utilized. This approach allows for the imaging of an entire depth level in one go, instead of just a single point in the subsurface. The computational benefit in 2D is already significant, however, in 3D it is even more potent. The retrieval of the plane waves in 3D has the limitation that only a certain angle of reflectivity can be retrieved, while the standard Marchenko imaging using point sources theoretically retrieves all possible angles that are present in the data. By combining the images associated with different angles of reflection together, the quality of the plane-wave imaging was improved, while still functioning at a significantly lower computational cost than the standard Marchenko imaging. From an image constructed using the plane-wave Marchenko approach a target zone can be efficiently determined and the standard Marchenko approach using virtual point sources can then be employed to obtain an accurate image in this target zone.

Given the size of the data volume and availability of computational resources to process that data, the (target oriented) point-source Marchenko method gives the most accurate results. However, if the computing power is limited, the advantage of utilizing efficient approaches to implement the method becomes evident. The use of compressed data and Eikonal solvers can drastically reduce the storage volume that is required for the data, while the plane-wave approach can be used to efficiently determine a target zone, which can then be further explored using standard Marchenko imaging. For a wider application of the Marchenko method, future research should include consideration of how the solutions to the Marchenko method can be efficiently computed, either through practical or theoretical improvements.

DATA AVAILABILITY STATEMENT

The SEG/EAGE overthrust model and additional information can be downloaded from https://wiki.seg.org/wiki/ SEG/EAGE_Salt_and_Overthrust_Models. The model is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of the license, visit https: //creativecommons.org/licenses/by/4.0 or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA. The license permits any user to freely copy and redistribute the material in any medium or format. Users are free to remix, transform and build upon the material for any purpose, including commercially.

The software for the finite-difference modelling, the Marchenko method and the Eikonal solver can all be found at https://github.com/JanThorbecke/OpenSource. Demo scripts for the methods in this paper can be found in the directory `marchenko3D/demo/marchenko3D/oneD/` in the OpenSource directory. The 3D figures in this paper were created using the ParaView software (Ayachit, 2015). The data that are used in this paper are generally stored in the format of Seismic Unix (SU) by Cohen and Stockwell (2001).

ORCID

Joeri Brackenhoff https://orcid.org/0000-0001-7697-1194
Jan Thorbecke https://orcid.org/0000-0002-5219-0868
Giovanni Meles https://orcid.org/0000-0002-2960-9587
Victor Koehne https://orcid.org/0000-0002-7338-1735
Diego Barrera https://orcid.org/0000-0003-1228-7283
Kees Wapenaar https://orcid.org/0000-0002-1620-8282

REFERENCES

Aminzadeh, F., Brac, J. and Kunz, T. (1997) *SEG/EAGE 3-D Salt and Overthrust Models. SEG/EAGE 3-D Modeling Series, No. 1: Distribution CD of Salt and Overthrust Models*. Tulsa, OK: Society of Exploration Geophysicists.

Amundsen, L. (2001) Elimination of free-surface related multiples without need of the source wavelet. *Geophysics*, 66(1), 327–341.

Ayachit, U. (2015) *The Paraview Guide: A Parallel Visualization Application*. Clifton Park, NY: Kitware, Inc.

Behura, J., Wapenaar, K. and Snieder, R. (2014) Autofocus imaging: image reconstruction based on inverse scattering theory. *Geophysics*, 79(3), A19–A26.

Brackenhoff, J. (2016) *Rescaling of incorrect source strength using Marchenko redatuming*. M.Sc. Thesis, Delft University of Technology, Delft. Available at http://resolver.tudelft.nl/uuid: 0f0ce3d0-088f-4306-b884-12054c39d5da Accessed October 5, 2021.

Brackenhoff, J., Thorbecke, J. and Wapenaar, K. (2019a) Monitoring induced distributed double-couple sources using Marchenko-based virtual receivers. *Solid Earth*, 10(4), 1301–1319.

Brackenhoff, J., Thorbecke, J. and Wapenaar, K. (2019b) Virtual sources and receivers in the real Earth: considerations for practical applications. *Journal of Geophysical Research: Solid Earth*, 124(11), 11802–11821.

Broggini, F. and Snieder, R. (2012) Connection of scattering principles: a visual and mathematical tour. *European Journal of Physics*, 33(3), 593.

Broggini, F., Snieder, R. and Wapenaar, K. (2014) Data-driven wavefield focusing and imaging with multidimensional deconvolution: numerical examples for reflection data with internal multiples. *Geophysics*, 79(3), WA107–WA115.

Broggini, F., Wapenaar, K., van der Neut, J. and Snieder, R. (2014) Data-driven Green's function retrieval and application to imaging with multidimensional deconvolution. *Journal of Geophysical Research: Solid Earth*, 119(1), 425–441.

Cohen, J. and Stockwell, J.J. (2001) *CWP/SU: Seismic Unix Release No. 43R4: An Open Source Software Package for Seismic Research and Processing*. Illinois St., Golden, CO: Colorado School of Mines.

da Costa Filho, C.A., Ravasi, M., Curtis, A. and Meles, G.A. (2014) Elastodynamic Green's function retrieval through single-sided Marchenko inverse scattering. *Physical Review E*, 90, 063201.

da Costa Filho, C.A., Tant, K., Curtis, A., Mulholland, A. and Moran, C.M. (2018) Using laboratory experiments to develop and test new Marchenko and imaging methods. In: *89th Annual Meeting SEG*. Tulsa, OK: Society of Exploration Geophysicists. pp. 4352–4356.

Haindl, C., Ravasi, M. and Broggini, F. (2021) Handling gaps in acquisition geometries: improving Marchenko-based imaging using sparsity-promoting inversion and joint inversion of time-lapse data. *Geophysics*, 86(2), S143–S154.

Koehne, V., Santos, M., Santos, R., Barrera, D., Brackenhoff, J. and Thorbecke, J. (2021) "A multi-GPU benchmark for 2D Marchenko imaging," SEG Technical Program Expanded Abstracts, 2939-2943. https://doi.org/10.1190/segam2021-3594688.1

Lindstrom, P. (2014) Fixed-rate compressed floating-point arrays. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2674–2683.

Lindstrom, P., Chen, P. and Lee, E.-J. (2016) Reducing disk storage of full-3D seismic waveform tomography (F3DT) through lossy online compression. *Computers and Geosciences*, 93, 45–54.

Lomas, A. and Curtis, A. (2020) Marchenko methods in a 3-D world. *Geophysical Journal International*, 220(1), 296–307.

Marchenko, V.A. (1955) On reconstruction of the potential energy from phases of the scattered waves. *Doklady Akademii Nauk SSSR*, 104(5), 695–698.

Matias, M.M.A., da Pestana, R.C. and van der Neut, J. (2018) Marchenko imaging by unidimensional deconvolution. *Geophysical Prospecting*, 66(9), 1653–1666.

Meles, G.A., Löer, K., Ravasi, M., Curtis, A. and da Costa Filho, C.A. (2015) Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry. *Geophysics*, 80(1), A7–A11.

Meles, G.A., Wapenaar, K. and Thorbecke, J. (2018) Virtual plane-wave imaging via Marchenko redatuming. *Geophysical Journal International*, 214(1), 508–519.

Meles, G.A., Zhang, L., Thorbecke, J., Wapenaar, K. and Slob, E. (2020) Data-driven retrieval of primary plane-wave responses. *Geophysical Prospecting*, 68(6), 1834–1846.

Pereira, R., Ramzy, M., Griscenco, P., Huard, B., Huang, H., Cypriano, L. and Khalil, A. (2019) Internal multiple attenuation for OBN data with overburden/target separation. In: *90th Annual Meeting SEG*. Tulsa, OK: Society of Exploration Geophysicists. pp. 4520–4524.

Ravasi, M. (2017) Rayleigh–Marchenko redatuming for target-oriented, true-amplitude imaging. *Geophysics*, 82(6), S439–S452.

Ravasi, M. and Vasconcelos, I. (2021) An open-source framework for the implementation of large-scale integral operators with flexible, modern HPC solutions – enabling 3D Marchenko imaging by least squares inversion. *Geophysics*, 86, 1–74.

Ravasi, M., Vasconcelos, I., Kritski, A., Curtis, A., da Costa Filho, C.A. and Meles, G.A. (2016) Target-oriented Marchenko imaging of a North Sea field. *Geophysical Journal International*, 205(1), 99–104.

Reinicke, C. (2020) *Elastodynamic Marchenko inverse scattering*. Ph.D. Thesis, Delft University of Technology, Delft.

Singh, S., Snieder, R., Behura, J., van der Neut, J., Wapenaar, K. and Slob, E. (2015) Marchenko imaging: imaging with primaries, internal multiples, and free-surface multiples. *Geophysics*, 80(5), S165–S174.

Slob, E. and Wapenaar, K. (2017) Theory for Marchenko imaging of marine seismic data with free surface multiple elimination. In: *79th EAGE Conference and Exhibition*. Houten, The Netherlands: European Association of Geoscientists and Engineers.

Slob, E., Wapenaar, K., Broggini, F. and Snieder, R. (2014) Seismic reflector imaging using internal multiples with Marchenko-type equations. *Geophysics*, 79(2), S63–S76.

Spetzler, J. and Angelov, P. (2005) Ray perturbation theory for traveltime and amplitude attributes in 3D heterogeneous media with weak anisotropy. In: *75th Annual Meeting SEG*. Tulsa, OK: Society of Exploration Geophysicists. pp. 2366–2369.

Staring, M., Pereira, R., Douma, H., van der Neut, J. and Wapenaar, K. (2018) Source-receiver Marchenko redatuming on field data using an adaptive double-focusing method. *Geophysics*, 83(6), 1–48.

Staring, M. and Wapenaar, K. (2019) Interbed demultiple using Marchenko redatuming on 3D field data of the Santos Basin. In: *16th International Congress of the Brazilian Geophysical Society*. Rio de Janeiro: Brazilian Geophysical Society, SBGf.

Staring, M. and Wapenaar, K. (2020a) 3D Marchenko internal multiple attenuation on narrow azimuth streamer data of the Santos Basin, Brazil. *Geophysical Prospecting*, 68(6), 1864–1877.

Staring, M. and Wapenaar, K. (2020b) Three-dimensional Marchenko internal multiple attenuation on narrow azimuth streamer data of the Santos basin, Brazil. *Geophysical Prospecting*, 68(6), 1864–1877.

Thorbecke, J. and Brackenhoff, J. (2019) *OpenSource: code for geophysical 3D/2D finite difference modelling, Marchenko algorithms, 2D/3D x-w migration and utilities*. Available at: https://doi.org/10.5281/zenodo.3374728.

Thorbecke, J., Slob, E., Brackenhoff, J., van der Neut, J. and Wapenaar, K. (2017) Implementation of the Marchenko method. *Geophysics*, 82(6), WB29–WB45.

van der Neut, J., Brackenhoff, J., Staring, M., Zhang, L., de Ridder, S., Slob, E. and Wapenaar, K. (2018) Single- and double-sided Marchenko imaging conditions in acoustic media. *IEEE Transactions on Computational Imaging*, 4(1), 160–171.

van der Neut, J., Thorbecke, J., Wapenaar, K. and Slob, E. (2015) Inversion of the multidimensional Marchenko equation. In: *77th EAGE Conference and Exhibition*. Houten, The Netherlands: European Association of Geoscientists and Engineers.

van der Neut, J., Vasconcelos, I. and Wapenaar, K. (2015) On Green's function retrieval by iterative substitution of the coupled Marchenko equations. *Geophysical Journal International*, 203(2), 792–813.

van der Neut, J. and Wapenaar, K. (2016) Adaptive overburden elimination with the multidimensional Marchenko equation. *Geophysics*, 81(5), T265–T284.

van Groenestijn, G. and Verschuur, D. (2009) Estimating primaries by sparse inversion and application to near-offset data reconstruction. *Geophysics*, 74(3), A23–A28.

van IJsseldijk, J. and Wapenaar, K. (2020) Adaptation of the iterative Marchenko scheme for imperfectly sampled data. *Geophysical Journal International*, 224(1), 326–336.

Verschuur, D.J., Berkhout, A. and Wapenaar, C. (1992) Adaptive surface-related multiple elimination. *Geophysics*, 57(9), 1166–1177.

Vidale, J.E. (1990) Finite difference calculation of traveltimes in three dimensions. *Geophysics*, 55(5), 521–526.

Wapenaar, K. (2014) Single-sided Marchenko focusing of compressional and shear waves. *Physical Review E*, 90, 063202.

Wapenaar, K. (2020) Reciprocity and representation theorems for flux-and field-normalised decomposed wave fields. *Advances in Mathematical Physics*, 2020, 9540135–9540150.

Wapenaar, K., Brackenhoff, J., Dukalski, M., Meles, G., Reinicke, C., Slob, E., Staring, M., Thorbecke, J., van der Neut, J. and Zhang, L. (2021) Marchenko redatuming, imaging and multiple elimination, and their mutual relations. *Geophysics*, 0, 1–103.

Wapenaar, K., Broggini, F., Slob, E. and Snieder, R. (2013) Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Green's function retrieval, and their mutual relations. *Physical Review Letters*, 110, 084301.

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E. and Snieder, R. (2014a) Green's function retrieval from reflection data, in absence of a receiver at the virtual source position. *The Journal of the Acoustical Society of America*, 135(5), 2847–2861.

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E. and Snieder, R. (2014b) Marchenko imaging. *Geophysics*, 79(3), WA39–WA57.

Zhang, L. and Staring, M. (2018) Marchenko scheme based internal multiple reflection elimination in acoustic wavefield. *Journal of Applied Geophysics*, 159, 429–433.

Zhou, J., Zhang, W., Lei, T., Wang, D. and Li, J. (2019) Reconstructing the primaries of PP and PS waves by elastic Marchenko method. In: *90th Annual Meeting SEG*. Tulsa, OK: Society of Exploration Geophysicists. pp. 4700–4704.

APPENDIX: INPUT FOR marchenko3D AND AUXILIARY PROGRAMS

A.1 marchenko3D

The marchenko3D program has the following parameters and options:

```
marchenko3D - Iterative Green's function and focussing functions retrieval in 3D
marchenko3D file_tinv= file_shot= [optional parameters]
Required parameters:
First arrival input options:
file_tinv= .............. direct arrival from focal point: G_d
file_ray= ............... direct arrival from raytimes
Shot data input options:
file_shot= .............. Reflection response (time data): R(t)
file_shotw= ............. Reflection response (frequency data): R(w)
file_shotzfp= ........... Reflection response (frequency compressed data): zfp[R(w)]
Optional parameters:
INTEGRATION
ampest=0 ................ Estimate a scalar amplitude correction with depth (=1)
tap=0 ................... lateral taper focussing(1), shot(2) or both(3)
ntap=0 .................. number of taper points at boundaries
fmin=0 .................. minimum frequency in the Fourier transform
fmax=70 ................. maximum frequency in the Fourier transform
MARCHENKO ITERATIONS
niter=10 ................ number of iterations
MUTE-WINDOW
file_amp= ............... amplitudes for the raytime estimation
file_wav= ............... Wavelet applied to the raytime data
above=0 ................. mute above(1), around(0) or below(-1) the travel times of the first arrival
shift=12 ................ number of points above(positive)/below(negative) travel time for mute
hw=8 .................... window in time samples to look for maximum in next trace
smooth=5 ................ number of points to smooth mute with cosine window
MUTE-WINDOW
plane_wave=0 ............ enable plane-wave illumination function
src_anglex=0 ............ angle of the plane wave in the x-direction
src_angley=0 ............ angle of the plane wave in the y-direction
src_velox=0 ............. velocity of the plane wave in the x-direction
src_veloy=0 ............. velocity of the plane wave in the y-direction
REFLECTION RESPONSE CORRECTION
scale=2 ................. scale factor of R for summation of Ni with G_d (only for time shot data)
pad=0 ................... amount of samples to pad the reflection series
HOMOGENEOUS GREEN'S FUNCTION RETRIEVAL OPTIONS
file_homg= .............. output file with homogeneous Green's function
The homogeneous Green's function is computed if a filename is given
file_inp= ............... Input source function for the retrieval
scheme=0 ................ Scheme for the retrieval
........................ scheme=0 Marchenko homogeneous Green's function retrieval with G source
........................ scheme=1 Marchenko homogeneous Green's function retrieval with f2 source
........................ scheme=2 Marchenko Green's function retrieval with source depending on virtual
```

```
receiver location
.......................... scheme=3 Marchenko Green's function retrieval with G source
.......................... scheme=4 Marchenko Green's function retrieval with f2 source
.......................... scheme=5 Classical homogeneous Green's function retrieval
.......................... scheme=6 Marchenko homogeneous Green's function retrieval with multiple G sources
.......................... scheme=7 Marchenko Green's function retrieval with multiple G sources
.......................... scheme=8 f1+ redatuming
.......................... scheme=9 f1- redatuming
.......................... scheme=10 2i IM(f1) redatuming
cp=1000.0 ............... Velocity of upper layer for certain operations
rho=1000.0 ............. Density of upper layer for certain operations
IMAGING
file_imag= .............. output file with image
The image is computed if a filename is given
OUTPUT DEFINITION
file_green= ............. output file with full Green function(s)
file_gplus= ............. output file with G+
file_gmin= .............. output file with G-
file_f1plus= ........... output file with f1+
file_f1min= ............ output file with f1-
file_f2= ............... output file with f2
file_ampscl= ........... output file with estimated amplitudes
file_iter= ............. output file with -Ni(-t) for each iteration
compact=0 .............. Write out homg and imag in compact format
........................ WARNING! This write-out cannot be displayed with SU
zfp=0 .................. Write out the standard output in compressed zfp format
tolerance=1e-7 ......... accuracy of the zfp compression,
verbose=0 .............. silent option; >0 displays info
```

The input of the 3D Marchenko method is similar to the 2D implementation, requiring reflection data and the first arrival time from the focal point. However, due to the large size of 3D data, loading the reflection data from disk using the `file_shot` option can become time-consuming. To mitigate this problem, the 3D implementation gives two alternate options of loading the data from disk. The first is loading the shot data in the frequency domain using `file_shotw`. These data have been pre-transformed to the frequency domain, which avoids the Fourier transform that was required on the shot data in the time domain. Alternatively, the frequency domain data can be compressed using the ZFP algorithm (Lindstrom, 2014) before they are loaded from disk, to reduce the file size. The code requires one of these three data types as input, and the latter two options can be obtained using the `TWtransform` module.

The second required input, the first arrival time, can be passed to the code in two ways. The first is by loading a shot record using the `file_tinv` option. An alternative is using the arrival times that are calculated by a 3D Eikonal solver, an example of which is the `raytime3D` program that is part of the software distribution, based on the work by Vidale (1990). The `raytime3D` program also computes a geometric spreading factor that can be used to estimate the amplitude of the first arrivals. This file can be read in using the `file_amp` option. To approximate seismic broadband data, a wavelet can be read into the code as well using `file_wav`. The file needs to contain a wavelet with no time shift and have the same temporal sampling as the reflection data. For both types of input for the first arrival, multiple focal points can be read in at the same time.

The number of iterations required for convergence depends on the reflection strengths and on the number of events in the model; a complex model will need more iterations. Typically, the number of iterations is chosen between 8 and 20. Setting the `verbose=2` option will compute the convergence of the algorithm by printing the energy of the iteration update $N_i(t)$ relative to the initial value $N_0(t)$. The

energy in the update term $N_i(t)$ should become smaller in each iteration.

The `marchenko3D` program is capable of computing an image for the focal points and outputting them directly through use of the `imaging3D` module of the code. To do this, one can simply set the `file_imag` option and the code computes an image point for each given focal point.

The program also contains an additional module called `homogeneous3D`, which is used for the purpose of retrieving the wavefield between two focal points in the subsurface. By setting the option `file_homg` to a correct path, the wavefield is computed according to the scheme set by the option `scheme`. Most of these schemes are explained in Brackenhoff *et al.* (2019a, 2019b). All the schemes require that the module is given input data by the `file_inp` option. This input file needs to be sampled at the same positions as the data that are computed in the main `marchenko3D` program with the same sam-

ple length and distance. It is recommended to run the module for one focal position first before using it on a large amount of focal positions.

To compensate for the transmission losses, an approximate amplitude correction can be used (van der Neut *et al.*, 2018). This estimation can be added to the results by setting the `ampest` option equal to 1. Using the `file_ampscl` option, the estimated amplitude corrections are written out for each focal point.

The code to produce examples of the Marchenko method on a flat layered model in this paper can be found in the directory `marchenko3D/demo/marchenko3D/oneD`. The `README` file in that directory explains in detail how to run the scripts. The SEG/EAGE overthrust model by Aminzadeh *et al.* (1997) can be found on the SEG wiki at https://wiki.seg.org/wiki/SEG/EAGE_Salt_and_Overthrust_Models.

A.2 TWtransform

```
TWtransform - Transform data from uncompressed time domain to compressed frequency
domain
TWtransform file_in= file_out= [optional parameters]
Required parameters:
file_in= ................ File containing the uncompressed time domain data
file_out= ............... Output for the (compressed) frequency domain data
Optional parameters:
verbose=1 ............... silent option; >0 displays info
fmin=0 .................. minimum frequency in the output
fmax=70 ................. maximum frequency in the output
mode=1 .................. sign of the frequency transform
zfp=0 ................... (=1) compress the transformed data using zfp
tolerance=1e-3 .......... accuracy of the zfp compression, smaller values
give more accuracy to the compressed data but
will decrease the compression rate
weight=2.0 .............. scaling of the reflection data
```

The `TWtransform` program is intended to reduce the file size of the reflection data and to limit the amount of time that is required for loading the reflection data. The program transforms the time data to the frequency domain, which already reduces the file size depending on the frequency range set by

`fmin` and `fmax`. The file size can be further reduced using the `zfp` option, which will compress the frequency data using the ZFP compression by Lindstrom (2014), with an accuracy set by the input parameter `tolerance`. The program also puts a custom header on the data to further reduce the file size.