



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

An introduction to Marchenko methods for imaging

Citation for published version:

Lomas, A & Curtis, A 2019, 'An introduction to Marchenko methods for imaging', *Geophysics*, vol. 84, no. 2, pp. F35-F45. <https://doi.org/10.1190/geo2018-0068.1>

Digital Object Identifier (DOI):

[10.1190/geo2018-0068.1](https://doi.org/10.1190/geo2018-0068.1)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Geophysics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



An Introduction to Marchenko Methods for Imaging

Angus Lomas¹ and Andrew Curtis^{1,2}

¹*School of GeoSciences, University of Edinburgh,*

Grant Institute, James Hutton Road,

King's Buildings, Edinburgh, EH9 3FE, UK

²*Institute of Geophysics, ETH Zurich,*

8092 Zurich, Switzerland

(March 22, 2020)

Running head: **An Introduction to Marchenko Methods**

ABSTRACT

Geoscientists often have little information about Earth's subsurface heterogeneities prior to mapping them using seismic or other geophysical data. Marchenko methods are a set of novel, data-driven techniques that help us to project surface seismic data to points in the subsurface, to create seismograms as though they had been measured at each point. In so doing Marchenko methods account for many of the complex, multiply-reflected seismic wave interactions that take place in the real Earth's subsurface. The resulting seismograms are the information required to create subsurface images that are more accurate than standard methods. Our aim is to introduce, these concepts with the minimum amount of mathematics required to understand how the Marchenko method can iterate to a solution, and to provide a well-commented, easily editable MATLAB code package for demonstration and training purposes. Green's function estimation using the Marchenko method is first illustrated for a constant velocity, variable density, one-dimensional medium, with results that show a near perfect match when compared to true, synthetically modelled solutions. Similar quality results are shown for variable velocity, two-dimensional Green's function estimation. Finally, we show how these estimates can be used to create images of the subsurface, which, when compared to standard methods contain reduced contamination due to multiple-related artifacts. The accompanying code package includes the two-dimensional dataset required to reconstruct the relevant figures presented herein, and allows readers to experiment with the implementation of the Marchenko method and the application of Marchenko imaging.

INTRODUCTION

The aim of seismic imaging is to map unknown heterogeneities in the Earth's subsurface, given a wavefield measured on or close to the Earth's surface. An approximate seismic wave speed model, usually called a velocity model, is required in order to map the subsurface accurately. This model provides a basic level of understanding about how seismic waves propagate through the subsurface and allows seismic information measured at the surface to be mapped to approximately correct subsurface locations. Much effort goes into estimating the velocity model using migration velocity analysis (Yilmaz, 2001; Sava and Biondi, 2004), travel time tomography (Stork, 1992; Jones, 2010) and full waveform inversion (Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009) but it is always imperfect. In particular it is usually far more smooth than the true Earth and therefore is not kinematically accurate; in other words it does not map waves that reflect from abrupt interfaces to their true subsurface positions. Even when these errors are sufficiently small that the image produced is correctly positioned, the inaccuracies usually cause other additional artifacts to be superimposed on the final image. Artifacts that are often most troublesome are those created by recorded seismic waves that reflect more than once in the subsurface, called multiples. This tutorial explains a set of methods that account for such waves so that these artifacts do not occur.

Marchenko methods (Rose, 2001; Brogгинi et al., 2012) are data-driven methods that use measured surface seismic data and an approximate velocity model to calculate the signal that would have been recorded at the surface if an impulsive, frequency band-limited source had fired at each chosen subsurface image point – including multiples. The estimated signals are called (frequency band-limited) Green's functions, and are exactly the information needed for accurate subsurface imaging (Behura et al., 2014; Wapenaar et al., 2014),

seismic redatuming (Wapenaar et al., 2014) or identifying and removing multiples (Meles et al., 2014, 2016).

The name Marchenko comes from the author of the original work on inverse scattering (Marchenko, 1955) who devised methods to estimate Green's functions in the field of quantum mechanics in one dimension (Snieder (2015) provides more information about this application). More recently a solution to the so-called Marchenko equations was formulated for geophysical applications that allow two and three-dimensional media to be imaged under certain approximations (Wapenaar et al., 2013; Lomas and Curtis, 2017).

This article presents an intuitive introduction to the Marchenko method and its applications. The aim is not to introduce fundamentally new concepts but to provide an easily accessible guide to some of the key concepts and methods that already exist. Additionally, a Marchenko MATLAB code and a relevant dataset accompany this article: the code is well commented, easily editable and adaptable for two-dimensional seismic problems. It is constructed so as to give readers further insight into the workflow used to calculate Green's functions using Marchenko methods, and to allow them to experiment and gain comfort with the methods - rather than being geared towards computationally efficient construction of large seismic images. Nevertheless, all two-dimensional examples presented herein were constructed using this code, hence it is perfectly sufficient to be used to teach and learn about Marchenko methods, and to process small datasets. Other codes exist in the public domain (e.g. Thorbecke et al. (2017)) but our code is designed specifically for user experimentation and so is written in a more intuitively accessible (higher-level) programming language. It is therefore also ideal as an aid to teaching about Marchenko methods in Masters or professional development courses.

Marchenko methods are simplest, most intuitive, and most accurate for one-dimensional problems, so the first section of this article introduces Green’s functions estimation for a simple one-dimensional medium in which full wavefields can be displayed and understood. We then introduce the reader to two-dimensional examples, the accompanying MATLAB code and the application of Marchenko imaging. Throughout this article multiple datasets are used: all are constructed in acoustic media and exclude free surface multiples as such data allow the simplest and most studied form of Marchenko methods to be applied. However, theory exists for Marchenko methods using elastic data (da Costa Filho et al., 2014, 2015; Wapenaar, 2014; Wapenaar and Slob, 2014) and data containing surface-related multiples (Singh et al., 2015, 2016). There are also limited examples of applications to real data (Ravasi et al., 2016; Jia et al., 2018; Wapenaar et al., 2018). For an entirely non-mathematical introduction to Marchenko methods we refer readers to van der Neut et al. (2015c), or for an introduction to one-dimensional Marchenko methods see Cui et al. (2018); for a thorough introduction to the more sophisticated mathematical aspects of the Marchenko methods see Slob et al. (2014b), Wapenaar et al. (2014) or van der Neut et al. (2015b) . Our tutorial fills the niche between these studies by introducing the concepts, the mathematics, and the computational machinery in an accessible way, with a code designed to facilitate experimentation and education.

THE MARCHENKO METHOD

There are multiple applications of Marchenko methods (imaging, redatuming, constructing primaries and multiple removal) but they all have the same foundation, namely Green’s function estimation. Green’s functions are the waves that arrive at a receiver location due to the firing of a spatio-temporally impulsive source. We represent these Green’s functions

as $G(x_0, x_i, t)$ where x_0 is the location of a receiver on the recording surface, x_i is a source point in the subsurface and t represents the time domain. In this syntax each term is a signal with two locations (x): the second always denotes the source location and the first is the receiver location. The Marchenko method estimates Green's functions between an arbitrarily chosen image point (or an artificial or virtual source) within the subsurface, and any point within the surface acquisition array (Figure 1).

[Figure 1 about here.]

The most basic form of Green's function estimation is to assume or estimate an initial approximate velocity model and estimate Green's functions $G(x_0, x_i, t)$ using either ray propagation or wavefield calculation through that model. This is standard practise in reverse time migration (RTM) for example (Baysal et al., 1983). Marchenko methods provide a workflow to estimate Green's functions but decomposed into two constituent parts: the first part consists of all waves that are upgoing ($-$) at the image point in the Earth's subsurface while the second part consists of the downgoing ($+$) waves. This includes components of the wavefield that have undergone multiple reflections, so-called multiples. In other words two Green's functions can be constructed from each subsurface image point which are recorded at the surface: the first G^- contains signals that start at the image point as a source wavefield propagating upward, the second G^+ contains signals that initially propagate downward from the image point and are reflected back up to the surface.

For simplicity let us begin with the one-dimensional Marchenko method. Two pieces of information are needed to calculate the decomposed Green's functions G^+ and G^- . The first is the reflectivity from a point source at the surface measured by a point receiver at the surface, denoted by $R(x_0, x_0, t)$; in the real world this is an idealised version of a

one-dimensional surface seismic reflection data after surface-related multiple removal. The second is an estimate of the direct (non-reflected) wave arrival $T_d(x_i, x_0, t)$ between the surface source and an image point. The decomposed Green's functions $G^{+/-}$ between x_0 and x_i are related to the reflectivity R through additional terms f^+ and f^- which are called focusing functions and are the subject of the next section:

$$G^+(x_0, x_i, t) = f^+(x_0, x_i, t) \otimes R(x_0, x_0, t) - f^-(x_0, x_i, t) \quad (1)$$

$$G^-(x_0, x_i, t) = f^+(x_0, x_i, -t) - R(x_0, x_0, t) \otimes f^-(x_0, x_i, -t) \quad (2)$$

Equations 1 and 2 are both defined in the time domain. Symbol $-t$ (and also later in this article, superscript $*$) denotes time reversal of the signal that precedes it. This is accomplished if we simply flip the positive and negative time axis of the initial signal, an operation that corresponds to complex conjugation in the frequency domain. The symbol \otimes represents a time domain convolution, which is equivalent to multiplication in the frequency domain.

It is worth noting that equations 1 and 2 differ from those given in most of the existing literature on Marchenko methods. To aid intuition we have created a virtual source at the image point inside the subsurface rather than a virtual receiver (the latter is more common). Comparing these cases, the direction of wave propagation is reversed: $G^+(x_0, x_i, t) = G^-(x_i, x_0, t)$. However the property of source-receiver reciprocity states that these are equivalent (identical signals). We will continue to use the virtual source syntax for the remainder of this article.

Focusing Functions

Focusing functions are key for understanding Marchenko methods. Imagine throwing a stone into a still pond on a windless day: ripples diverge from the location of impact, propagating as waves across the water surface. Let us imagine that these ripples are recorded on some closed boundary of receivers that surrounds the impact point. If we waited until all of the energy had settled, we could then use the receivers as sources to inject the recorded wavefield back into the pond. If we do this in time-reversed order (inject the last wave recorded at each receiver first), the original ripples will be recreated, but this time they would converge inwards rather than propagating outwards (Cassereau and Fink, 1992). They will all eventually re-focus at the impact point, then diverge outwards again, creating another wavefield that can be recorded at the receiver boundary. In this thought experiment, the (time reversed) wavefield injected on the boundary, is called a focusing function: it defines exactly which waves we should inject in order to focus the in-going energy at the impact point.

Focusing functions used in Marchenko methods are intuitively similar to those above. The only conceptual differences are that the source point in this case is the subsurface source in Figure 1, and that the receiver boundary is at the Earth's surface and so is only on one side of the source point. Downgoing focusing functions are related to wavefields that if injected at the Earth's surface, would focus (collapse all of their energy to a point) at a specific location in the subsurface (here, the location of any chosen virtual source or image point). However, in the case of focusing in the subsurface, this only occurs in an idealised (truncated) model of the Earth's subsurface structure which is homogeneous below the depth of that point, but which has the true Earth's structure above that depth.

Focusing means that there is a time at which the waves at a certain depth only exist at one specific image point – everywhere else at that depth the wavefield is zero. The function f^+ is the wavefield that we would have to inject at the surface (at point x_0) in order for the wavefield to focus at the image point, and hence this wavefield is downgoing at the surface. Function f^- is the wavefield that we would record at the surface as we inject f^+ in the truncated model, hence f^- is upgoing at the surface. Both wavefields are shifted along the time axis such that the focus occurs at time zero.

Figure 2 includes a standard representation of a focusing function (Slob et al., 2014b) for a simple one-dimensional subsurface model that consists of layers with varying density and a constant velocity. First, Figure 2a shows the wavefield that develops in space and time when a simple impulsive source (convolved with a Ricker wavelet) is injected at the surface at time $t = 0$ (note that time is on the horizontal axis). This consists of a direct wave (the first continuous, linear wave on the left) and a set of (singly and multiply) reflecting waves. At a particular image point in depth (for example $1400m$ - indicated by an arrow in Figure 2) multiple waves arrive and hence there is not a focus of energy. In order to create such a focus additional energy must be injected to cancel all but the direct wave at that point.

[Figure 2 about here.]

The focusing function is the signal at the surface ($depth = 0m$) in Figure 2b, shown in Figure 2c. The downgoing component f^+ is the signal injected at the surface in order to create the focus at $depth = 1400m$, shown by the circle in Figure 2b. The upgoing component f^- is the reflected response observed at the surface from this injected signal. It can be seen that three pulses of energy (α , β and γ) are injected at x_0 in addition to the initial pulse δ to cancel out the reflected components of the wavefield observed in Figure

2a. These three signals together with δ make up the complete downgoing focusing function f^+ and all of the up-coming waves at $depth = 0m$ comprise f^- .

Iterative Solution

The Marchenko method works by first calculating the focusing functions and then using equations 1 and 2 to estimate the Green's functions. While the relationships between focusing functions and Green's functions in those equations is a relatively simple one, they do not explain how one can calculate focusing functions. Several methods have been proposed to do this (Broggini et al., 2014; van der Neut et al., 2015a,b), and here we present the method of Wapenaar et al. (2014) as it can be understood most intuitively.

In a one-dimensional system we assume that we know the reflectivity at the surface $R(x_0, x_0, t)$ as well as the direct arrival between the surface and the image point $T_d(x_i, x_0, t)$ – which identifies the chosen image point x_i . The first step in estimating the focusing functions is to set:

$$f_0^+(x_0, x_i, t) = T_d(x_i, x_0, t)^{-1} \quad (3)$$

Equation 3 inverts the direct arrival, commonly approximated as simply performing a time reversal (switching the time axis of T_d and setting the signal at positive times to zero: $T_d(x_i, x_0, t)^{-1} \approx T_d(x_i, x_0, -t)$). The result is used as a first approximation for f^+ denoted f_0^+ (Wapenaar et al., 2014) and forms the component δ from Figure 2b. This makes intuitive sense: if we time reverse the direct wave between the image point and the surface, it will propagate back to its source point (the image point) and create a pulse of energy there at zero time, just as in the example of ripples on the pond. Unfortunately

though, as it propagates back into the subsurface some of its energy will scatter or reflect from heterogeneities in the Earth, creating a more complex part of the wavefield that will disrupt the focus. Marchenko methods design energy to inject in order to destructively interfere with these scattered waves, reducing them to zero amplitude.

The estimate for f_0^+ can then be used to estimate f_0^- :

$$f_0^-(x_0, x_i) = \theta(x_0, x_i, t) [R(x_0, x_0, t) \otimes f_0^+(x_0, x_i, t)] \quad (4)$$

Within the square brackets equation 4 convolves the initial estimate of f_0^+ with the reflectivity, which is equivalent to injecting f_0^+ into the Earth and recording the result at the surface. Again this is equivalent to injecting the time-reversed wavefield in the pond, and recording the reflecting waves on the source boundary. The additional term in this equation, θ , is a focusing-location dependent window which removes all energy that arrives at times greater than or equal to the direct arrival and is symmetric in time. It may appear counter intuitive to apply a window that removes all energy at these times as this is the data that we are ultimately trying to estimate in the Green's functions. However, this stage of the Marchenko method estimates the focusing functions, and these functions only exist at times *before* the direct arrival and *after* the time reversed direct arrival. Outside this window is where the Green's function exists but its accuracy is dependent on the accuracy of the focusing functions (by equations 1 and 2). Furthermore, this in itself is an approximation as we are assuming that the Green's function and focusing functions can be separated by a windowing operator in the space-time domain, which is not always the case (e.g. when the focusing location is on or near to a subsurface interface). Nevertheless, we work with these approximations and now iterate to a solution as follows.

In Figure 2b three wave packets are labelled (α , β and γ): these are injected at the surface in addition to the initial impulsive source δ that is used to obtain the reflectivity in Figure 2a. These additional wave packets make up M_k^+ which is the coda (later part) of f_k^+ , where k is the number of iterations:

$$f_k^+(x_0, x_i, t) = f_0^+(x_0, x_i, t) + M_k^+(x_0, x_i, t) \quad (5)$$

As a demonstration of how the focusing functions are estimated using equations 3 and 4 (and equations 6 and 7 below), Figures 3 and 4 show a series of ray path diagrams that explain their various travel time relationships. These figures use a similar display format to that presented by van der Neut et al. (2015b). In Figure 3 the three primary reflections from the reflectivity are depicted individually (middle column) and convolved with the inverted direct arrival f_0^+ (left column) from equation 3. The main point of Figure 3 is to show that the results of this convolution are a series of *non*-physical signals, each of which is contained within the pass-window of θ and make up f_0^- . They are not physical because each signal on the right is made up of combinations of energy that has positive (solid) and negative (dashed) travel times which survive the windowing operation in equation 3. Despite being non-physical at this stage, they can be used to construct the focusing functions by progressing them to the next iteration:

$$M_k^+(x_0, x_i, -t) = \theta(x_0, x_i, t) [R(x_0, x_0, t) \otimes f_{k-1}^-(x_0, x_i, -t)] \quad (6)$$

[Figure 3 about here.]

To estimate M_k^+ using equation 6 we start with the estimate f_{k-1}^- produced by the

previous iteration (or the initial iteration f_0^-), time reverse it, and then convolve it with the reflectivity. The same windowing operator (θ) as above is then applied to isolate the focusing functions of interest.

A schematic of the first iteration ($k = 1$) to calculate M_1^+ is shown in Figure 4. The columns on the left of Figure 4 show a subset of the (time reversed) columns on the right of Figure 3, a subset was selected (rows 2 and 3 from Figure 3 only) as these are the only components that contribute to M_1^+ . This column is convolved across rows with primary reflections from R ; again we have only included a subset of these reflections as these are the components that contribute to M_1^+ . The solutions to this convolution step are shown in the third column and the time reversal of this result is given in the final column.

The final column represents M_1^+ and is composed of three signals each of which is made up of the time reversed direct arrival (the left column in Figure 3) plus an additional time lag. This additional time lag is equal to the two-way travel time through one or more subsurface layers. This travel time information is what the Marchenko method requires to accurately account for internal multiples. To demonstrate this we have included the travel times $\alpha = -0.24s$, $\beta = -0.16s$ and $\gamma = 0.16s$ from Figure 2b and 2c in the column depicting M_1^+ in Figure 4. This shows that at $depth = 0m$ we have calculated the travel times of the signals we need to inject into the subsurface to destructively interfere with the multiply-scattered components of the reflectivity so as to cancel them out above the focus point. We have therefore demonstrated that the additional components of the focusing functions that are used in Figure 2b to remove multiples from the seismic reflection data can be formed by convolving the data with itself (and with the direct wave estimate). While the travel times of α , β and γ are correct, the amplitudes of the energy constructed in iteration 1 does not cause perfect cancellation of the internal multiples; amplitudes are corrected in subsequent

iterations.

The waves in M_k^+ can then be injected into the subsurface and what would return to the surface can be calculated (the convolution with R); the results are windowed with θ and whatever remains is added to the estimate of f_0^- :

$$f_k^-(x_0, x_i, t) = f_0^-(x_0, x_i, t) + \theta(x_0, x_i, t) [R(x_0, x_0, t) \otimes M_k^+(x_0, x_i, t)] \quad (7)$$

Equations 6 and 7 are iterated until the solutions for M_k^+ and f_k^- in consecutive iterations have converged to stable values. When the solutions have converged, the total downgoing focusing function f_k^+ can be constructed by summing the inverted direct arrival f_0^+ and M_k^+ using equation 5.

[Figure 4 about here.]

It is worth noting that within equations 4-7 the quantity from the previous iteration is convolved with the reflectivity (R), which in practise always contains a source term (no matter whether real or modelled data are used). To avoid iteratively convolving multiple source terms together, the source wavelet was initially deconvolved from the reflectivity. If the reflectivity was not deconvolved, the effective source wavelet would change with each iteration so consecutive iterations would be inconsistent. An alternative approach that avoids deconvolution is illustrated in the two-dimensional example below.

Green's Function Estimation

Once calculated, the focusing functions can be used to estimate directionally decomposed Green's functions ($G^{+/-}$) using equations 1 and 2; summing those two signals gives the

total Green's function $G(x_0, x_i, t) = G^-(x_0, x_i, t) + G^+(x_0, x_i, t)$. This Green's function represents the signal that would have been measured at the surface if there had been a source at the image point (or vice versa, by source-receiver reciprocity).

[Figure 5 about here.]

As this experiment is synthetic we can test the accuracy of the Green's function estimates by comparing them with the true Green's function – one that is obtained by modelling an actual source at the image point, as shown in Figure 5. The final panels of Figure 5a and 5b show the estimated and true Green's functions for two different subsurface image points. For both examples they match in time, and amplitudes are correct for all arrivals. In the first two panels of both Figure 5a and 5b the Green's functions are shown in decomposed form as obtained directly from equations 1 and 2: we observe well separated events in the upgoing and downgoing components. It can be seen that there are no measured downgoing arrivals in Figure 5a which is to be expected given that downgoing waves from a virtual source at the image point would have to be reflected back upward in order to be recorded at the surface; there are no interfaces below the image point in Figure 5a (as shown by the reflector locations in Figure 2) so no such reflection can occur. By contrast the image point in Figure 5b lies between reflectors and therefore both the downgoing and upgoing signals contain arrivals.

Marchenko methods are only able to construct events that follow ray paths for which the energy was recorded in the original reflectivity (R). Therefore in the examples presented in Figure 5 we have only plotted the estimated Green's function to a maximum time of 1.4s. This equates to times preceding the recording time of the reflectivity (R) minus the travel time of the direct arrival (T_d). This shows that it is important to have sufficiently long

recording times for Marchenko methods to be effective.

So far we have discussed a simple, one-dimensional example which demonstrates the methodology clearly and in which the solutions are essentially perfect. The next section extends the examples to higher dimensions, where the results are more prone to errors.

MARCHENKO METHODS IN HIGHER DIMENSIONS

The example above illustrated for one-dimensional problems that all of the information required to determine Green's functions between a subsurface image point and the surface is contained within just two signals, $R(x_0, x_0, t)$ and $T_d(x_i, x_0, t)$. In two dimensions this is not the case as the reflectivity from one surface source (\mathbf{x}'_0) is measured by multiple receivers (\mathbf{x}_0). It is worth noting that our notation must now change to account for the extra spatial coordinate where $\mathbf{x} = (x_1, x_2)$. Nevertheless, in two or three dimensions concepts similar to those in Figures 3 and 4 hold. Indeed while for one-dimensional problems those diagrams have axes of depth and time, they apply with similar geometries (but incorrect angles of transmission and reflection) to two-dimensional problems if the horizontal time axis is replaced with the horizontal space axis. Each arrival in the desired Green's function at any particular angle is constructed by the interference of other specific arrivals at other particular receiver and source combinations in the reflectivity.

Rather than requiring that we selected specific arrivals to convolve in order to construct each arrival in the Green's function, Marchenko methods sum (integrate) over all possible sources along the acquisition array (boundary) and rely on destructive interference to cancel out unwanted energy. A similar cancellation occurs in reverse time migration (Kaelin and Guitton, 2006) and in seismic interferometry (van Manen et al., 2005, 2006; Wapenaar and

Fokkema, 2006). This only works accurately in two dimensions if the reflectivity is of the correct form, and in practise this means that a vertical spatial derivative (often called a dipole) source or receiver needs to be created or measured. In the two-dimensional examples presented in this article the reflectivity is from a pressure (monopole) source measured by a vertical particle velocity (dipole) receiver.

In the previous section we introduced a set of formulae to estimate Green's functions using Marchenko methods. These formulae are extended to two dimensions by changing the one-dimensional convolutions to multi-dimensional convolutions and integrating across all sources on the acquisition boundary (we denote this boundary as $\partial\mathbb{D}_0$). For example the two-dimensional versions of equations 1 and 2 for source redatuming are:

$$G^+(\mathbf{x}_0, \mathbf{x}_i, t) = \int_{\partial\mathbb{D}_0} f^+(\mathbf{x}'_0, \mathbf{x}_i, t) \otimes R(\mathbf{x}_0, \mathbf{x}'_0, t) d\mathbf{x}'_0 - f^-(\mathbf{x}_0, \mathbf{x}_i, t) \quad (8)$$

$$G^-(\mathbf{x}_0, \mathbf{x}_i, t) = f^+(\mathbf{x}_0, \mathbf{x}_i, -t) - \int_{\partial\mathbb{D}_0} R(\mathbf{x}_0, \mathbf{x}'_0, t) \otimes f^-(\mathbf{x}'_0, \mathbf{x}_i, -t) d\mathbf{x}'_0 \quad (9)$$

To implement equations 8 and 9 the focusing functions need to be available between the focusing location and both the surface sources and receivers (e.g. $f^+(\mathbf{x}_0, \mathbf{x}_i, t)$ and $f^+(\mathbf{x}'_0, \mathbf{x}_i, t)$). In practice the two-dimensional formulation of these functions requires interchangeability between the two. We therefore impose the condition that the source array and receiver array are co-located $\mathbf{x}'_0 = \mathbf{x}_0$.

A further consideration for implementation of the two-dimensional Marchenko method is the direct arrival estimate (T_d) and windowing function (θ) which now need to be estimated in two dimensions as they were above in one dimension. These functions are now calculated between a single focusing location and multiple surface sources/receivers. This increases the

complexity of these functions and the potential for errors in estimating them; nevertheless if this is done following the same workflow as one-dimensional Marchenko methods the relationships discussed in the previous sections still hold.

Green’s Function Estimation in Two Dimensions

In the following two-dimensional example, the source wavelet in the reflectivity is not the same as that in the focusing functions. We use the $20Hz$ Ricker wavelet shown in Figure 6b for the focusing function, whereas we use a flat spectrum wavelet shown in Figure 6a as the source wavelet for our reflectivity. The flat spectrum wavelet is defined in the frequency domain so as to have an amplitude of 1 over the range of frequencies of interest (the frequencies contained within the Ricker wavelet) as demonstrated in Figure 6c. Using this formulation removes the need for deconvolution as it ensures that the frequency content of the updated focusing functions does not change between iterations of the Marchenko method (in each iteration they are simply multiplied by a source wavelet that does not change the shape of the current wavelet within the frequency band of interest (Thorbecke et al., 2017)).

[Figure 6 about here.]

Figure 7 shows a subsurface model that is used to demonstrate two-dimensional Marchenko methods. This subsurface model has both variable density and velocity. There are 188 symmetrically spread sources and receivers co-located at 16 meter intervals along the surface of the model ($depth = 0$). A point is also marked in the subsurface, $\mathbf{x}_i = (1000m, 800m)$ which identifies a chosen image point for Marchenko Green’s function calculation.

[Figure 7 about here.]

Two models have been used to create the two input datasets required for the Marchenko method. The reflectivity from surface sources measured at the surface receivers and excluding free surface multiples has been created using the true model given in Figure 7a. This represents surface seismic reflection data after free-surface multiples have been removed. An estimate of the direct arrival between the surface and the image point has been created using the smooth model in Figure 7b. In practice we do not have access to the true model, hence we have used a smoothed version of the true model for our direct arrival calculations, assuming that in practise some initial or reference velocity model will be available. The direct arrival signal can be modelled with finite-difference solvers (Figure 7c), or approximated using eikonal solvers to find the travel time at which a scaled source wavelet can be assumed to arrive. In this example, for accuracy, both datasets were created using finite-difference solutions to the acoustic wave equation (in Figure 7b a source was fired at the image point and recorded along the surface receiver array, giving Figure 7c). See van der Neut and Wapenaar (2016) for an alternative solution to solving the Marchenko method if an estimated velocity model is not available.

By iterating the two-dimensional form of equations 1-7 (Wapenaar et al., 2014) the focusing functions and Green's functions are obtained, and the final Green's function estimates are shown in Figure 8b. For simplicity we have not included every component of the estimated Marchenko Green's function (e.g. focusing functions) - for these we refer readers to the accompanying code package within which these figures are included.

[Figure 8 about here.]

To test the accuracy of the Marchenko solution, the calculated Green’s functions in Figure 8b are plotted beside the true solutions in Figure 8a. The latter panel shows a solution computed in the true model in Figure 7a by firing a source at the image point. The estimated signal shows a good match, with negligible errors visible. The errors that are present can be attributed to limited boundary coverage by the acquisition array, errors in the finite difference solution, and windowing artifacts.

In Figure 8 all of the amplitudes have been scaled to values between 1 and -1. This has been done for comparison purposes, as the Marchenko methods implemented in this article cannot estimate true absolute amplitudes of Green’s functions. The primary reason for this is that the direct arrival was approximated at the start of the Marchenko method as $f_0^+(\mathbf{x}_0, \mathbf{x}_i, t) \approx T_d(\mathbf{x}_i, \mathbf{x}_0, -t)$: we do not know the amplitude of the true inverse, so it is impossible to estimate a Marchenko solution with the true absolute amplitude as the initial focusing function estimate is implicit in the final solution – see equation 5.

MARCHENKO CODE PACKAGE

Accompanying this article is a set of well-commented MATLAB codes for two-dimensional Green’s function estimation. The first of these (*CODE_1*) is the code used to create Figure 8 and running it without edits should produce a version of that figure. For this code the inputs are pre-computed and the variables already set, below we discuss the operation of this. However, because that code is inflexible (the image point is fixed) we have included a second code for user experimentation which we introduce at the end of this section. By choosing an appropriate set-up, users should be able to use the second code to reproduce similar outputs of the first – a good exercise for learning and teaching purposes.

Data

Four datasets accompany this MATLAB code, as summarised in Table 1. All of these datasets are stored in the time domain, although for computational efficiency most of the code operations are performed in the frequency domain.

[Table 1 about here.]

The direct arrival, windowing function and true signals are all common shot gathers from a source at the image point with identical dimensions: $i = 2001$ and $j = 188$, where i is the number of time samples and j is the number of surface receivers. In this example the sampling interval was $0.002s$ and the recording time was from $-2s$ to $2s$. The fourth dataset is the reflectivity which has an extra dimension k to account for multiple shot locations, with $i = 2001$, $j = 188$ and $k = 188$.

Codes

The code called *ICCR_marchenko.m*, follows the same workflow introduced earlier in this paper. Algorithm 1 shows the corresponding pseudocode. The equations referred to in Algorithm 1 are the one-dimensional versions given herein, but the code uses the equivalent two-dimensional versions given in Wapenaar et al. (2014). We have not defined a measure of convergence within this code; instead a desired number of iterations is input by the operator, the default being 5.

One additional feature in the MATLAB code has not been discussed above: a spatial taper is applied during each summation (integration) of sources along the acquisition boundary. Its purpose is to account for the limited acquisition coverage: it ensures cancel-

Algorithm 1 Marchenko Green's Function Estimation Pseudocode

```
load  $R$ ,  $T_d$  and  $\theta$ 

calculate  $f_0^+$  and  $f_0^-$  using equations 3 and 4

for  $n$  iterations ( $k=1,2,\dots,n$ ) do

    calculate  $M_k^+$  using equation 6

    calculate  $f_k^-$  using equation 7

end for

calculate  $f_k^+$  using equation 5

calculate  $G^+$  and  $G^-$  using equations 1 and 2 (or 8 and 9)

return  $G^+$  and  $G^-$ 
```

lation of edge effects from the extremities of the acquisition array, as these would otherwise create spurious energy in the solutions (if the boundary was infinite, as assumed in two- or three-dimensional Marchenko theory, this would not be required). The taper takes a half cosine shape at either end of the array and the number of points to be tapered at either end of the array can be varied by the user (the default is 20% of the number of receivers).

A second code in the accompanying package (*CODE_2*), operates using the same fundamentals as the code introduced above, but has been implemented as matrix multiplications for computational efficiency and changed to allow alternative virtual source locations to be used. To the latter end the direct arrival and window are no longer pre-computed. Instead we have included a function that computes these using an eikonal solution (*ICCR_marchenko_eik.mat*). The input seismic data remains the same but there is no longer a true solution available for comparison. This code is set up so the input data can be changed; to do this a measured or calculated reflectivity dataset in the correct form will be required as well as an eikonal solution through an estimated velocity model.

MARCHENKO IMAGING

Marchenko imaging creates images of the subsurface using the estimated Green’s functions. This works in a similar way to conventional imaging algorithms such as reverse time migration (RTM), in the sense that the similarity of two signals is tested by the use of an imaging condition (Claerbout, 1971), and the result is used to identify subsurface inhomogeneities.

Marchenko imaging should in theory be able to be used to calculate more accurate images than standard methods as we have access to more accurate Green’s functions $G^{+/-}$ which account for multiply reflected waves. The imaging condition applied for our implementation of Marchenko imaging is a (zero-lag) cross-correlation between the downgoing Green’s function G^+ and the direct arrival estimate T_d as proposed by da Costa Filho et al. (2015):

$$I_{MI}(\mathbf{x}_i) = \sum_{\mathbf{x}_0} \sum_t T_d(\mathbf{x}_i, \mathbf{x}_0, t) G^+(\mathbf{x}_0, \mathbf{x}_i, t) \quad (10)$$

For comparison purposes we have also implemented an alternative imaging condition to approximate standard methods:

$$I_{RTM}(\mathbf{x}_i) = \sum_{\mathbf{x}_0} \sum_t T_d(\mathbf{x}_i, \mathbf{x}_0, t) G_0^+(\mathbf{x}_0, \mathbf{x}_i, t) \quad (11)$$

where $G_0^+ \approx R \otimes T_d^*$, which is equivalent to a back-propagated wavefield used in RTM methods – see da Costa Filho and Curtis (2016) for further details.

The two signals in equations 10 and 11 should be most similar when the image point \mathbf{x}_i is on a reflector, so the cross-correlation will produce maxima at those points. We note that as with conventional RTM there are several alternative imaging conditions that could be

applied. Comparisons and discussion of these is beyond the scope of this article but more detail is given in da Costa Filho et al. (2015) and Singh and Snieder (2017b).

We imaged the model in Figure 7a using equations 10 and 11, with results shown in Figure 9. Image points were selected on a four meter grid inside the imaged area. Directionally decomposed Green’s functions were calculated at each of these points. The only difference between these Green’s functions and those calculated in Figure 8 is in the direct arrival input, which here was constructed by placing a source wavelet at the travel time calculated using an eikonal solver, rather than calculating T_d using finite difference methods; this saves on computational cost since eikonal solvers require relatively few operations compared to finite-difference methods. This example therefore also illustrates that this approximate method of modelling the direct field can be sufficient for some imaging applications.

[Figure 9 about here.]

The image calculated using equation 10 at each image point is shown in Figure 9a. All interfaces in the subsurface are identified with few artifacts in the solution, despite the presence of internal multiples in the surface acquisition data that is input to the Marchenko method. The clarity of the image is mainly due to the accuracy of the Green’s function estimates G^+ which are calculated by Marchenko methods and used by equation 10. For comparison the solution in Figure 9b is calculated using equation 11 with the approximate G_0^+ shows clear artifacts due to internal multiples.

We have included a final code (*CODE_3*) in the software package which can be used to implement Marchenko imaging using the methods discussed above. There is an increased computational cost associated with implementing this because a Green’s function now needs to be calculated at each and every imaging point. Inside the code the image can be target-

orientated by defining the image point spacing and a limited or targeted subsurface location.

DISCUSSION

The aim of this tutorial is to provide beginners to Marchenko methods with an accessible way to understand the topic, and to allow them to begin to experiment with the methods on synthetic examples using easily understandable and editable MATLAB code. Marchenko methods as introduced in this paper use surface seismic data and an estimate of the subsurface velocities to estimate Green's functions between a subsurface image point and the surface. These estimated signals have a steadily increasing number of applications, including subsurface imaging and seismic redatuming, but also for removing multiples (Meles et al., 2014), constructing primaries (Meles et al., 2016) and calculating Green's function where both the source and receiver are inside the subsurface (Wapenaar et al., 2016; Singh and Snieder, 2017a). Marchenko methods also have applications outside of seismology, for example for imaging using ground penetrating radar (Slob et al., 2014a) or for medical imaging (van der Neut et al., 2017).

The results are promising, and offer a data-driven method that improves on current imaging methods, in particular by correctly predicting the arrival of multiply reflected waves at image points. Given the novelty of these methods there are still aspects that are poorly understood – areas of ongoing research. They include exploring how to apply Marchenko in real, dissipative media with seismic attenuation, the effects on Green's function estimates of velocity model errors and corresponding poor estimates of direct arrivals (both in time and amplitude), the effect of various types of noise in the reflectivity field, and the cost and effort of scaling two-dimensional Marchenko methods to three dimensions.

The accompanying MATLAB codes for the estimation of acoustic Green's functions using two-dimensional Marchenko methods, also comes with an accompanying dataset which can be used to reconstruct Figure 8. The codes can easily be adapted for users to include their own datasets. The data will need to be formatted correctly, the details for which are discussed in the earlier sections of this tutorial and in comments within the code. Datasets of a similar simplicity should achieve similarly positive results.

CONCLUSION

In this article we have introduced Marchenko methods, a set of novel, data-driven techniques that can be applied to seismic redataing and imaging problems. We have shown these methods can accurately estimate directionally decomposed Green's functions from virtual subsurface source locations to surface receiver locations in both one and two dimensions, and this includes the multiply-scattered components of the Green's functions. However, all of the methods we have presented are based on synthetic seismic datasets; extending these methods to more realistic datasets and examples is an area of active research, and with the accompanying MATLAB code readers have the tools to contribute to this.

ACKNOWLEDGMENTS

The authors would like to thank Petrobras and Shell for their sponsorship of the International Centre for Carbonate Reservoirs (ICCR), and for permission to publish this work from the VSP project. We would also like to thank fellow members of the ICCR and members of the Edinburgh Interferometry Project (EIP) for their numerous fruitful discussions. Finally, we would like to thank Jan Thorbecke, Johan Robertsson, two anonymous review-

ers and the Associate Editor Jonas D. De Basabe for their comments which helped improve this article. The data used within this article is generated with the Madagascar open-source software package freely available from www.ahay.org.

REFERENCES

- Baysal, E., D. D. Kosloff, and J. W. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524.
- Behura, J., K. Wapenaar, and R. Snieder, 2014, Autofocus imaging: Image reconstruction based on inverse scattering theory: *Geophysics*, **79**, A19–A26.
- Broggini, F., R. Snieder, and K. Wapenaar, 2012, Focusing the wavefield inside an unknown 1D medium: Beyond seismic interferometry: *Geophysics*, **77**, A25–A28.
- Broggini, F., K. Wapenaar, J. Neut, and R. Snieder, 2014, Data-driven Green’s function retrieval and application to imaging with multidimensional deconvolution: *Journal of Geophysical Research: Solid Earth*, **119**, 425–441.
- Cassereau, D., and M. Fink, 1992, Time-reversal of ultrasonic fields. iii. theory of the closed time-reversal cavity: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, **39**, 579–592.
- Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: *Geophysics*, **36**, 467–481.
- Cui, T., I. Vasconcelos, D.-J. v. Manen, and K. Wapenaar, 2018, A tour of Marchenko redatuming: Focusing the subsurface wavefield: *The Leading Edge*, **37**, 67a1–67a6.
- da Costa Filho, C. A., and A. Curtis, 2016, Attenuating multiple-related imaging artifacts using combined imaging conditions: *Geophysics*, **81**, S469–S475.
- da Costa Filho, C. A., M. Ravasi, and A. Curtis, 2015, Elastic P-and S-wave autofocus imaging with primaries and internal multiples: *Geophysics*, **80**, S187–S202.
- da Costa Filho, C. A., M. Ravasi, A. Curtis, and G. A. Meles, 2014, Elastodynamic Green’s function retrieval through single-sided Marchenko inverse scattering: *Physical Review E*, **90**, 063201.

- Jia, X., A. Guitton, and R. Snieder, 2018, A practical implementation of subsalt Marchenko imaging with a Gulf of Mexico dataset: *Geophysics*, **83**, 1–57.
- Jones, I. F., 2010, Tutorial: Velocity estimation via ray-based tomography: *First Break*, **28**, 45–52.
- Kaelin, B., and A. Guitton, 2006, Imaging condition for reverse time migration, *in* SEG Technical Program Expanded Abstracts 2006: Society of Exploration Geophysicists, 2594–2598.
- Lomas, A., and A. Curtis, 2017, 3d seismic imaging using Marchenko methods: AGU Fall Meeting Abstracts, NS31C–03.
- Marchenko, V. A., 1955, On reconstruction of the potential energy from phases of the scattered waves: *Dokl. Akad. Nauk SSSR*, 695–698.
- Meles, G. A., K. Löer, M. Ravasi, A. Curtis, and C. A. da Costa Filho, 2014, Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry: *Geophysics*, **80**, A7–A11.
- Meles, G. A., K. Wapenaar, and A. Curtis, 2016, Reconstructing the primary reflections in seismic data by Marchenko redatuming and convolutional interferometry: *Geophysics*, **81**, Q15–16.
- Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Ravasi, M., I. Vasconcelos, A. Kritski, A. Curtis, C. A. d. C. Filho, and G. A. Meles, 2016, Target-oriented Marchenko imaging of a North sea field: *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **205**, 99–104.
- Rose, J. H., 2001, single-sided focusing of the time-dependent Schrödinger equation: *Phys-*

- ical Review A, **65**, 012707.
- Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. i. theory: Geophysical Prospecting, **52**, 593–606.
- Singh, S., and R. Snieder, 2017a, Source-receiver Marchenko redatuming: Obtaining virtual receivers and virtual sources in the subsurface: Geophysics, **82**, Q13–Q21.
- , 2017b, Strategies for imaging with Marchenko-retrieved Greens functions: Geophysics, **82**, 1–68.
- Singh, S., R. Snieder, J. Behura, J. van der Neut, K. Wapenaar, and E. Slob, 2015, Marchenko imaging: Imaging with primaries, internal multiples, and free-surface multiples: Geophysics, **80**, S165–S174.
- Singh, S., R. Snieder, J. van der Neut, J. Thorbecke, E. Slob, and K. Wapenaar, 2016, Accounting for free-surface multiples in Marchenko imaging: Geophysics, **82**, R19–R30.
- Slob, E., J. Hunziker, J. Thorbecke, and K. Wapenaar, 2014a, Creating virtual vertical radar profiles from surface reflection ground penetrating radar data: Ground Penetrating Radar (GPR), 2014 15th International Conference on, IEEE, 525–528.
- Slob, E., K. Wapenaar, F. Brogini, and R. Snieder, 2014b, Seismic reflector imaging using internal multiples with Marchenko-type equations: Geophysics, **79**, S63–S76.
- Snieder, R., 2015, Demystifying Marchenko imaging: Presented at the 77th EAGE Conference and Exhibition-Workshops.
- Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, 680–692.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
- Thorbecke, J., E. Slob, J. Brackenhoff, J. van der Neut, and K. Wapenaar, 2017, Imple-

- mentation of the Marchenko method: *Geophysics*, **82**, 1–56.
- van der Neut, J., J. L. Johnson, K. van Wijk, S. Singh, E. Slob, and K. Wapenaar, 2017, A Marchenko equation for acoustic inverse source problems: *The Journal of the Acoustical Society of America*, **141**, 4332–4346.
- van der Neut, J., J. Thorbecke, K. Wapenaar, and E. Slob, 2015a, Inversion of the multidimensional Marchenko equation: Presented at the 77th EAGE Conference and Exhibition 2015.
- van der Neut, J., I. Vasconcelos, and K. Wapenaar, 2015b, On Green’s function retrieval by iterative substitution of the coupled Marchenko equations: *Geophysical Journal International*, **203**, 792–813.
- van der Neut, J., and K. Wapenaar, 2016, Adaptive overburden elimination with the multidimensional Marchenko equation: *Geophysics*, **81**, T265–T284.
- van der Neut, J., K. Wapenaar, J. Thorbecke, E. Slob, and I. Vasconcelos, 2015c, An illustration of adaptive Marchenko imaging: *The Leading Edge*, **34**, 818–822.
- van Manen, D.-J., A. Curtis, and J. O. Robertsson, 2006, Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity: *Geophysics*, **71**, SI47–SI60.
- van Manen, D.-J., J. O. Robertsson, and A. Curtis, 2005, Modeling of wave propagation in inhomogeneous media: *Physical Review Letters*, **94**, 164301.
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.
- Wapenaar, K., 2014, Single-sided Marchenko focusing of compressional and shear waves: *Physical Review E*, **90**, 063202.
- Wapenaar, K., J. Brackenhoff, J. Thorbecke, J. van der Neut, E. Slob, and E. Verschuur,

- 2018, Virtual acoustics in inhomogeneous media with single-sided access: *Scientific reports*, **8**, 2497.
- Wapenaar, K., F. Broggini, E. Slob, and R. Snieder, 2013, Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Greens function retrieval, and their mutual relations: *Physical Review Letters*, **110**, 084301.
- Wapenaar, K., and J. Fokkema, 2006, Greens function representations for seismic interferometry: *Geophysics*, **71**, SI33–SI46.
- Wapenaar, K., and E. Slob, 2014, On the Marchenko equation for multicomponent single-sided reflection data: *Geophysical Journal International*, **199**, 1367–1371.
- Wapenaar, K., J. Thorbecke, and J. van der Neut, 2016, A single-sided homogeneous Green’s function representation for holographic imaging, inverse scattering, time-reversal acoustics and interferometric Green’s function retrieval: *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **205**, 531–535.
- Wapenaar, K., J. Thorbecke, J. Van Der Neut, F. Broggini, E. Slob, and R. Snieder, 2014, Marchenko imaging: *Geophysics*, **79**, WA39–WA57.
- Yilmaz, Ö., 2001, *Seismic data analysis*: Society of Exploration Geophysicists Tulsa, **1**.

LIST OF FIGURES

1	Illustration of the types of singly- or multiply-reflected signals estimated by the Marchenko method.	35
2	Reflectivity and focusing functions of a one-dimensional medium. The medium has a constant velocity (2500m/s) but variable density: dashed lines represent subsurface interfaces between layers of different densities which reflect energy. The reflectivity of the medium in panel (a) shows the location of the wavefield in space (depth) at every time for a single impulsive source (denoted δ) fired at time zero. This can be related to the focusing function in panel (b) where additional components α , β and γ are injected at the surface after the initial source δ : these cancel various reflections in the subsurface to ensure that focusing occurs in the subsurface. In this example the focusing location was chosen to be at 1400m depth (indicated by an arrow, and circled in panel (b)). The decomposed focusing functions (c) are the downgoing f^+ and upgoing f^- (dashed arrows) components at the surface ($depth = 0m$) in panel (b). These diagrams are of a similar form to those presented by Slob et al. (2014b).	36
3	A schematic diagram of ray paths contributing to the initial focusing function estimate f_0^- using equations 3 and 4. The first column shows the inverted direct arrival f_0^+ , approximated by the time reversed direct wave (note that zero time is at the centre of each horizontal axis). As stated in equation 3, this is convolved with the reflectivity, which in column two is decomposed into three primary reflections. The combination of these two events across each row creates the events shown in the right column which are all components of f_0^- . Dashed rays are time-reversed compared to their physical counterparts; solid rays are not time-reversed. Hence, starting at the source point at time zero, a wave in the right column would have the travel time of the solid ray segments minus the travel time along the dashed segments (and is therefore non-physical). These diagrams are of a similar form to those presented by van der Neut et al. (2015b).	37
4	A schematic ray path diagram for the retrieval of the first estimate ($k = 1$) of M_k^{++} using equation 6. The input into this step is $(f_0^-)^*$ given in the first column which is obtained from different rows in the right column of Figure 3 after time reversal. This is convolved again with the reflectivity in column two (equation 6) and produces the results $(f_0^-)^* \otimes R$ in column three. After windowing with θ these are the time reverse of the components that make up the later part of the downgoing focusing function injected in Figure 2b and c (components α , β and γ). This is shown in column 4 which is simply the time reversal of the results in column 3. These diagrams are of a similar form to those presented by van der Neut et al. (2015b).	38
5	Estimated Green's functions from two subsurface image points. Figure (a) is for an image point at 1400m; Figure (b) is for an image point at 850m. Panels 3 and 6 (counting from the top downwards) compare Marchenko and true Green's functions. Panels 1, 2, 4 and 5 show the upgoing and downgoing decomposition of the corresponding total Green's function.	39

6	Wavelets used in two-dimensional finite-difference modelling. Panels (a) and (b) show zero phase, time domain plots of the reflectivity and direct arrival wavelets respectively, and panel (c) compares the amplitude of the frequency spectra of the two wavelets. The reflectivity and direct arrival are shown as solid and dashed lines respectively.	40
7	The true (a) and smoothed (b) subsurface models used for the two-dimensional synthetic example. The subsurface model has a variable velocity (shown) and a proportionate variable density model (densities lie in the range $1000kg/m^3 - 5000kg/m^3$). The surface is spanned by 188 co-located sources and receivers represented by stars and triangles (with every tenth source and receiver plotted). The white circle in (b) marks a chosen subsurface image point at location \mathbf{x}_i . Panel (c) shows an estimate of the direct arrival between the image point and the surface as calculated through the smooth model in panel (b).	41
8	A comparison of Green's functions from image point \mathbf{x}_i in Figure 7b. Panel (a) shows the true solution calculated through the true model in Figure 7a using finite difference methods. Panel (b) shows the Marchenko solution calculated using the methods discussed in the main text. Panel (c) compares trace number 51 (offset= $804m$) taken from panels (a) and (b).	42
9	The Marchenko (a) and reverse time migration (b) images for the subsurface models defined in Figure 7a and 7b. A Green's functions has been estimated every 4 meters and the imaging condition used for each of the images is defined in equations 10 and 11 respectively. The dashed red lines represent the true subsurface heterogeneities.	43

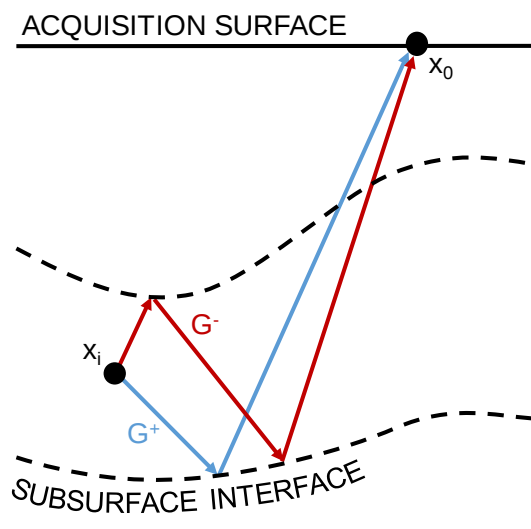


Figure 1: Illustration of the types of singly- or multiply-reflected signals estimated by the Marchenko method.

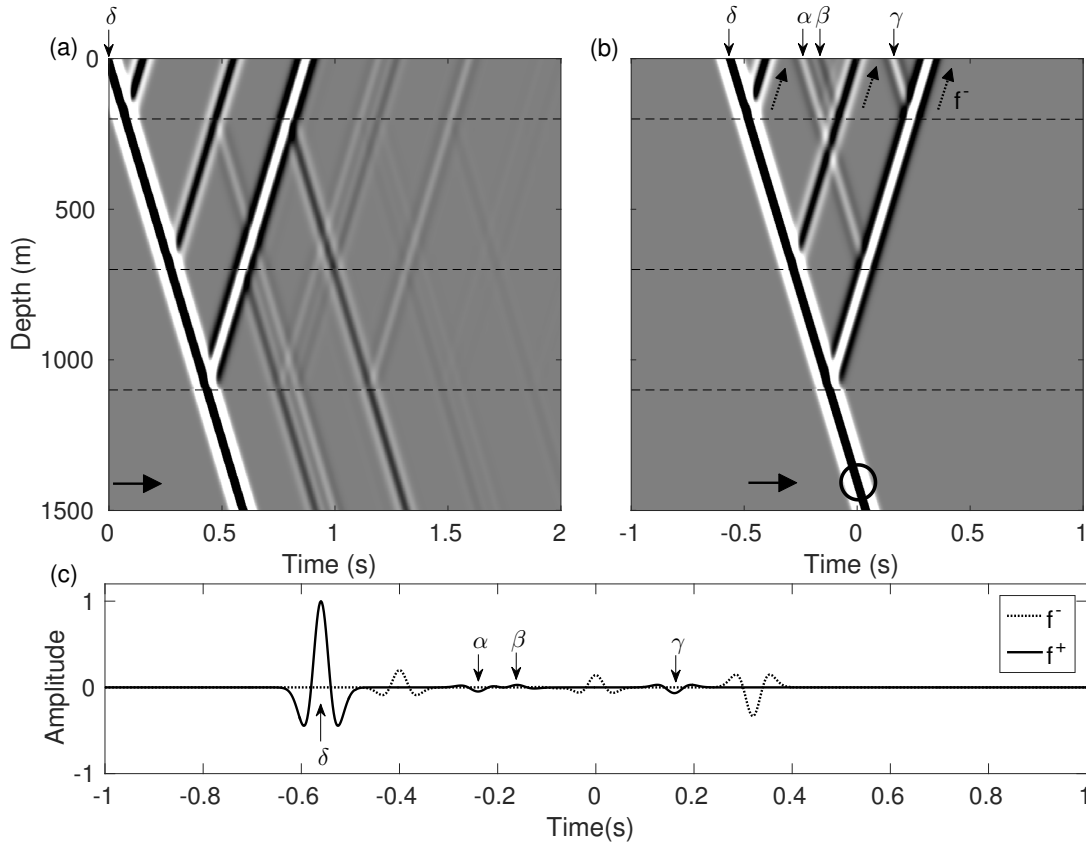


Figure 2: Reflectivity and focusing functions of a one-dimensional medium. The medium has a constant velocity ($2500m/s$) but variable density: dashed lines represent subsurface interfaces between layers of different densities which reflect energy. The reflectivity of the medium in panel (a) shows the location of the wavefield in space (depth) at every time for a single impulsive source (denoted δ) fired at time zero. This can be related to the focusing function in panel (b) where additional components α , β and γ are injected at the surface after the initial source δ : these cancel various reflections in the subsurface to ensure that focusing occurs in the subsurface. In this example the focusing location was chosen to be at $1400m$ depth (indicated by an arrow, and circled in panel (b)). The decomposed focusing functions (c) are the downgoing f^+ and upgoing f^- (dashed arrows) components at the surface ($depth = 0m$) in panel (b). These diagrams are of a similar form to those presented by Slob et al. (2014b).

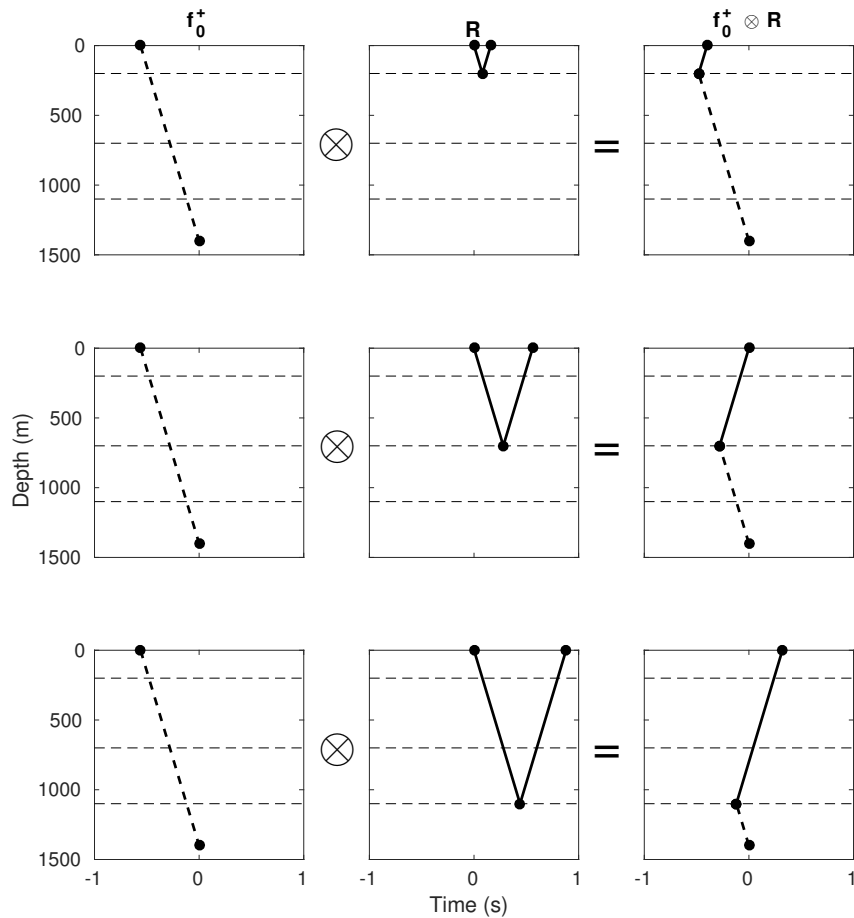


Figure 3: A schematic diagram of ray paths contributing to the initial focusing function estimate f_0^- using equations 3 and 4. The first column shows the inverted direct arrival f_0^+ , approximated by the time reversed direct wave (note that zero time is at the centre of each horizontal axis). As stated in equation 3, this is convolved with the reflectivity, which in column two is decomposed into three primary reflections. The combination of these two events across each row creates the events shown in the right column which are all components of f_0^- . Dashed rays are time-reversed compared to their physical counterparts; solid rays are not time-reversed. Hence, starting at the source point at time zero, a wave in the right column would have the travel time of the solid ray segments minus the travel time along the dashed segments (and is therefore non-physical). These diagrams are of a similar form to those presented by van der Neut et al. (2015b).

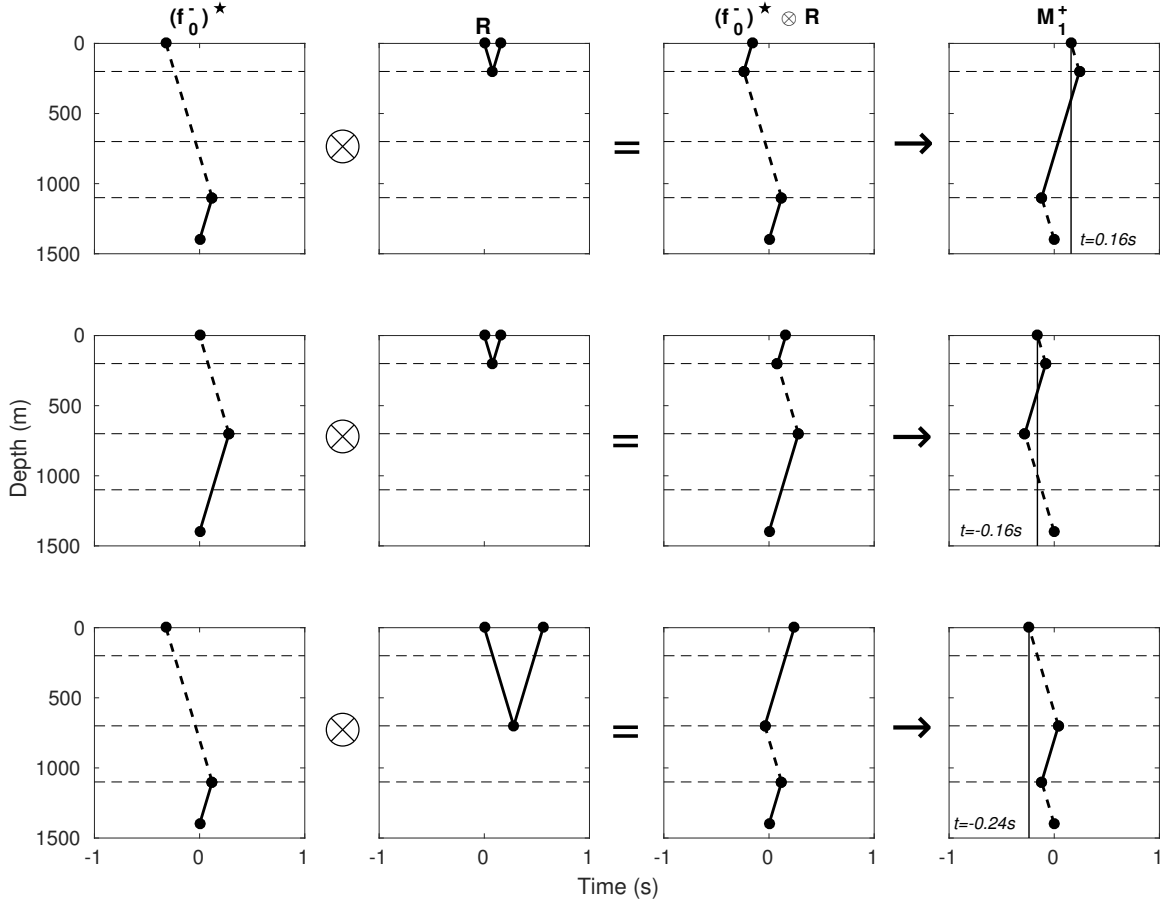


Figure 4: A schematic ray path diagram for the retrieval of the first estimate ($k = 1$) of M_k^{+*} using equation 6. The input into this step is $(f_0^-)^*$ given in the first column which is obtained from different rows in the right column of Figure 3 after time reversal. This is convolved again with the reflectivity in column two (equation 6) and produces the results $(f_0^-)^* \otimes R$ in column three. After windowing with θ these are the time reverse of the components that make up the later part of the downgoing focusing function injected in Figure 2b and c (components α , β and γ). This is shown in column 4 which is simply the time reversal of the results in column 3. These diagrams are of a similar form to those presented by van der Neut et al. (2015b).

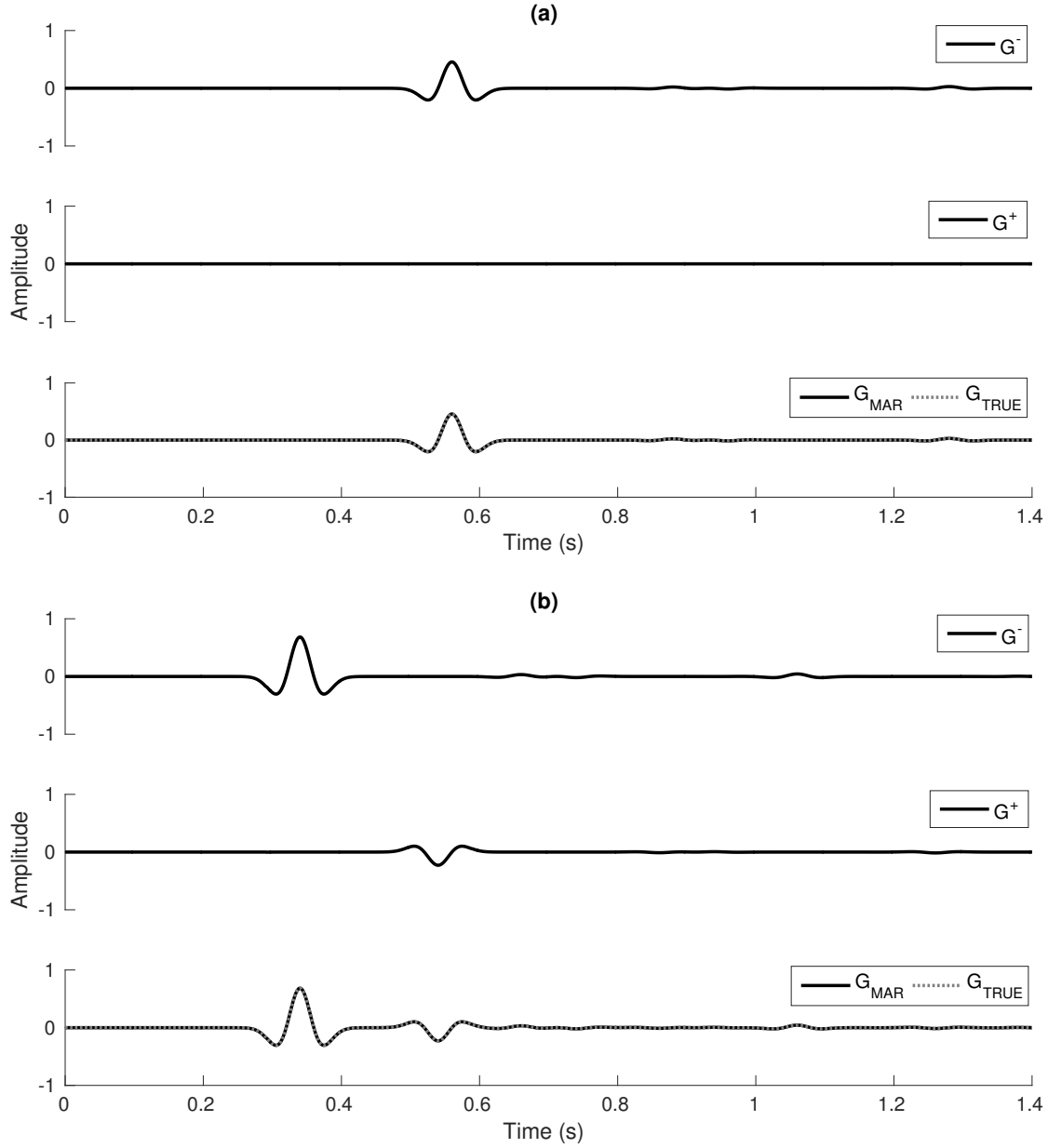


Figure 5: Estimated Green's functions from two subsurface image points. Figure (a) is for an image point at 1400m; Figure (b) is for an image point at 850m. Panels 3 and 6 (counting from the top downwards) compare Marchenko and true Green's functions. Panels 1, 2, 4 and 5 show the upgoing and downgoing decomposition of the corresponding total Green's function.

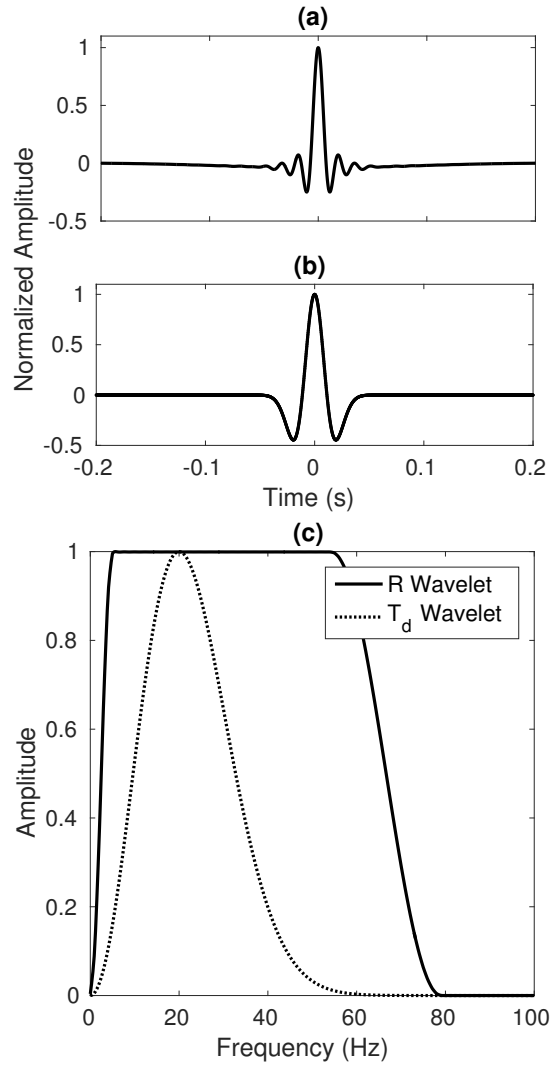


Figure 6: Wavelets used in two-dimensional finite-difference modelling. Panels (a) and (b) show zero phase, time domain plots of the reflectivity and direct arrival wavelets respectively, and panel (c) compares the amplitude of the frequency spectra of the two wavelets. The reflectivity and direct arrival are shown as solid and dashed lines respectively.

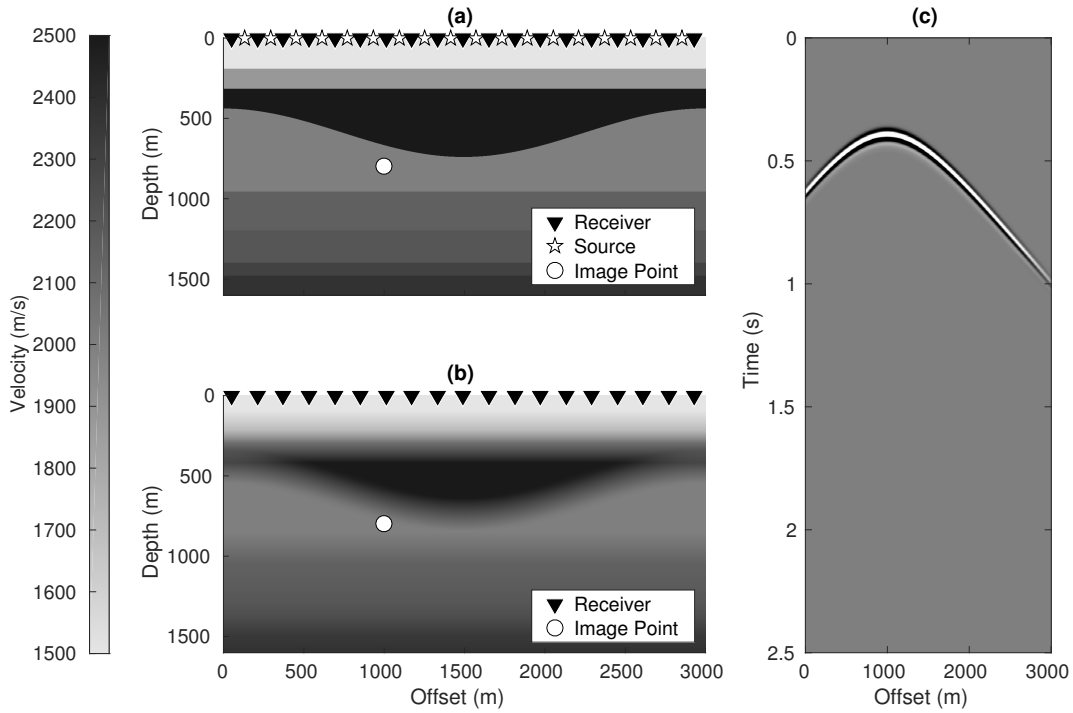


Figure 7: The true (a) and smoothed (b) subsurface models used for the two-dimensional synthetic example. The subsurface model has a variable velocity (shown) and a proportionate variable density model (densities lie in the range $1000\text{kg}/\text{m}^3 - 5000\text{kg}/\text{m}^3$). The surface is spanned by 188 co-located sources and receivers represented by stars and triangles (with every tenth source and receiver plotted). The white circle in (b) marks a chosen subsurface image point at location \mathbf{x}_i . Panel (c) shows an estimate of the direct arrival between the image point and the surface as calculated through the smooth model in panel (b).

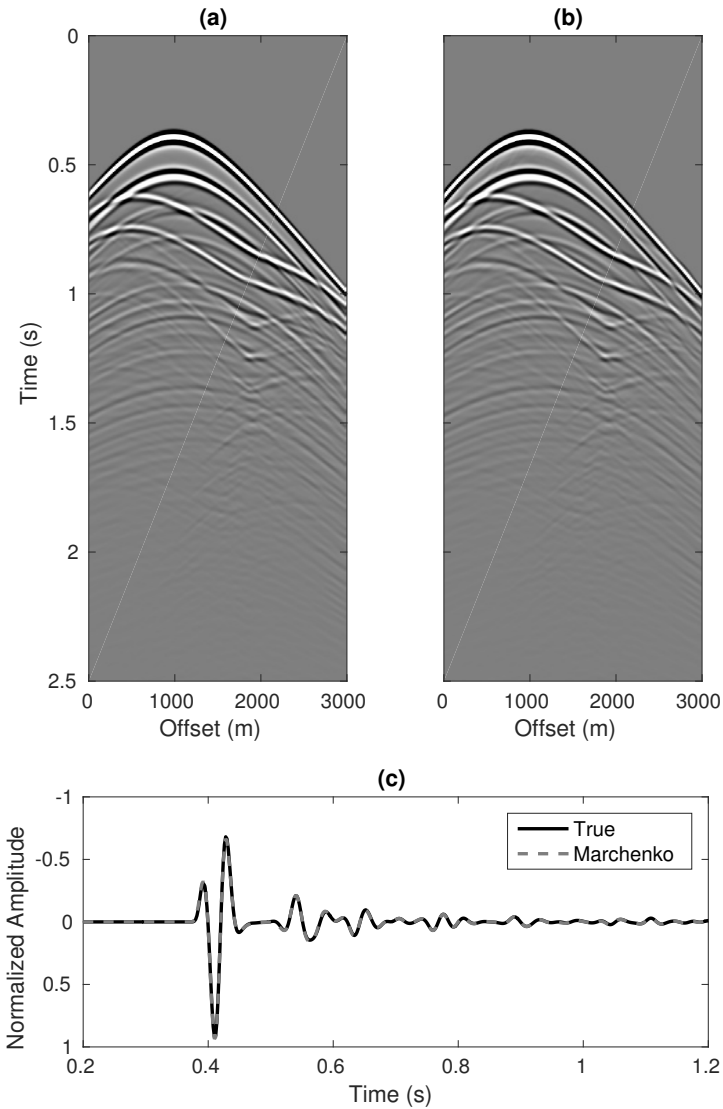


Figure 8: A comparison of Green's functions from image point \mathbf{x}_i in Figure 7b. Panel (a) shows the true solution calculated through the true model in Figure 7a using finite difference methods. Panel (b) shows the Marchenko solution calculated using the methods discussed in the main text. Panel (c) compares trace number 51 (offset=804m) taken from panels (a) and (b).

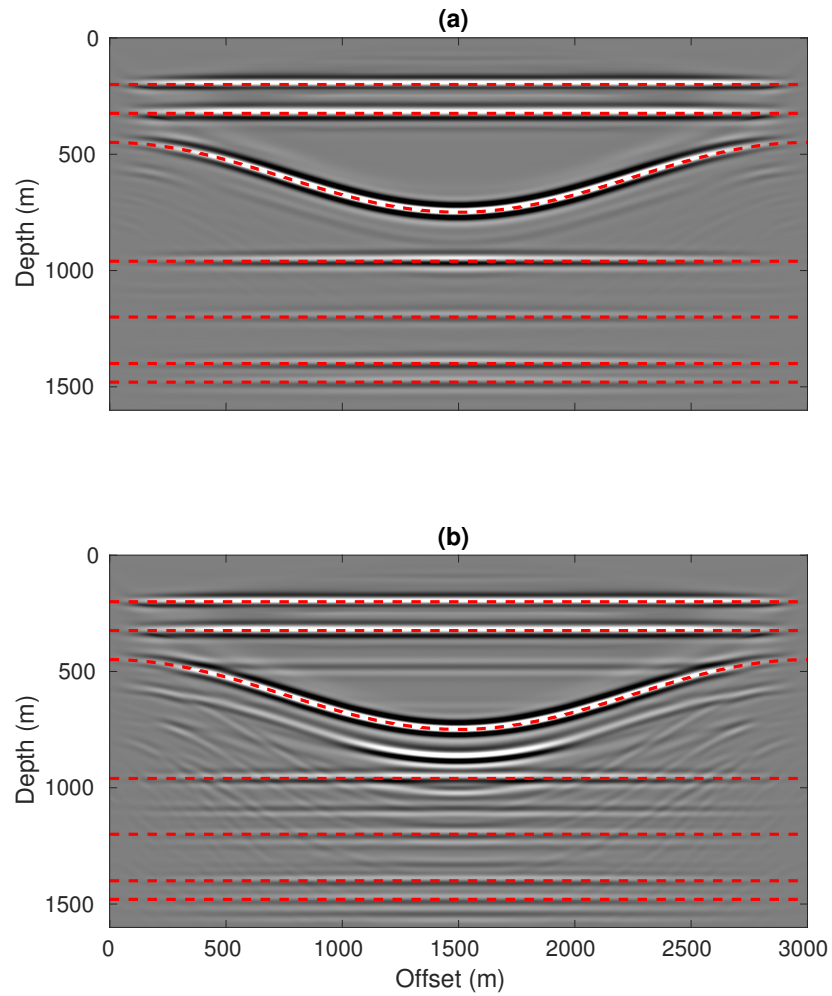


Figure 9: The Marchenko (a) and reverse time migration (b) images for the subsurface models defined in Figure 7a and 7b. A Green's functions has been estimated every 4 meters and the imaging condition used for each of the images is defined in equations 10 and 11 respectively. The dashed red lines represent the true subsurface heterogeneities.

LIST OF TABLES

- 1 Description of the datasets used as inputs to the MATLAB code *ICCR_marchenko.m*. 45

Dataset	Description
<i>ICCR_marchenko_R.mat</i>	The modelled reflectivity: the acquisition boundary is at the top surface ($depth = 0m$) of the model defined in Figure 7a.
<i>ICCR_marchenko_TD.mat</i>	The modelled direct arrival: the signal with a source at the image point and the receivers on the acquisition boundary as shown in Figure 7b.
<i>ICCR_marchenko_theta.mat</i>	A filter designed using the direct arrival <i>ICCR_marchenko_TD.mat</i> , which mutes the signal at times greater than or equal to the direct arrival.
<i>ICCR_marchenko_GT.mat</i>	The modelled true solution for comparison: a real source is located at the image point and receivers are on the acquisition boundary in Figure 7a.

Table 1: Description of the datasets used as inputs to the MATLAB code *ICCR_marchenko.m*.