On the Retrieval of Forward-Scattered Waveforms From Acoustic Reflection and Transmission Data With the Marchenko Equation

Joost van der Neut, Joeri Brackenhoff, Giovanni Meles, Lele Zhang, Evert Slob, and Kees Wapenaar

***Abstract*— A Green's function in an acoustic medium can be retrieved from reflection data by solving a multidimensional Marchenko equation. This procedure requires *a priori* knowledge of the initial focusing function, which can be interpreted as the inverse of a transmitted wavefield as it would propagate through the medium, excluding (multiply) reflected waveforms. In practice, the initial focusing function is often replaced by a time-reversed direct wave, which is computed with help of a macro velocity model. Green's functions that are retrieved under this (direct-wave) approximation typically lack forward-scattered waveforms and their associated multiple reflections. We examine whether this problem can be mitigated by incorporating transmission data. Based on these transmission data, we derive an auxiliary equation for the forward-scattered components of the initial focusing function. We demonstrate that this equation can be solved in an acoustic medium with mass density contrast and constant propagation velocity. By solving the auxiliary and Marchenko equation successively, we can include forward-scattered waveforms in our Green's function estimates, as we demonstrate with a numerical example.**

***Index Terms*— Acoustic propagation, acoustic signal processing, acoustic waves.**

I. Introduction

IT HAS been shown that Green's function between a horizontal acquisition surface and an arbitrary location $\mathbf{x}$ inside an unknown lossless acoustic medium can be retrieved from a single-sided reflection response by solving a multidimensional Marchenko equation [1]. This insight has led to numerous applications in the field of applied geophysics, see [2] for an overview. Besides knowledge of the single-sided reflection response at an acquisition surface, the Marchenko methodology requires access to the source signature, which can sometimes be retrieved from the recorded data [3] and an initial estimate of the transmitted wavefield as it would propagate from the acquisition surface to $\mathbf{x}$ in the absence of (multiple) reflections. Typically, this initial estimate is obtained from a macro model of the propagation velocity [4]. The phase [5], [6] and amplitude [7] of the initial estimate can be updated within the Marchenko framework.

In theory, the wavefield that is used to initialize the Marchenko scheme should include all forward-scattered waveforms [8], [9], which are those waveforms that do not (ever) change vertical direction while propagating from the (horizontal) acquisition boundary to $\mathbf{x}$. In practice, we typically compute the initial focusing function in a smooth macro model, which does not contain sharp contrasts. Consequently, forward-scattered waveforms and their associated multiples will not be accurately reconstructed [10], which can harm the (Marchenko) imaging process [11]. Another problem is posed by thin-layered structures, generating multiple reflections with short periods that cannot be resolved due to the finite frequency content of the data [12]. This problem can be mitigated (at least to some extent) by enforcing energy conservation and minimum-phase conditions [13]–[15].

The Marchenko scheme can also be applied to ultrasonic data [16]–[18], opening new ways for biomedical applications, especially below objects with strong contrast such as the human skull [19]. Hence, the Marchenko equation might be tailored to supplement common biomedical modalities, such as transcranial wavefield focusing [20], brain imaging [21], or transcranial photoacoustics [22]. Remarkably, some of the acquisition designs that are common for these applications allow the collection of auxiliary transmission data. As the transmission response bears an imprint of the desired forward-scattered waveforms, these data could be key to improve the initial wavefield estimate, which is needed to solve the Marchenko equation. In this article, we elaborate on this idea. Unlike the Marchenko scheme mentioned in [23], which has been proposed recently for closed-boundary data, we make a sharp distinction between reflection and transmission data. We start with a brief derivation of the Marchenko equation for reflection data. By modifying the derivation slightly, we find an auxiliary equation for the recorded transmissions. Our aim

Manuscript received March 17, 2022; accepted March 24, 2022. Date of publication March 31, 2022; date of current version April 26, 2022. This work was supported by the European Research Council under Grant 742703. *(Corresponding author: Joost van der Neut.)*

Joost van der Neut, Evert Slob, and Kees Wapenaar are with the Department of Geoscience and Engineering, Delft University of Technology, 2600 Delft, The Netherlands (e-mail: j.r.vanderneut@tudelft.nl).

Joeri Brackenhoff is with the Department of Earth Sciences, ETH Zürich, 8092 Zürich, Switzerland.

Giovanni Meles is with the Institute of Earth Sciences, University of Lausanne, 1015 Lausanne, Switzerland.

Lele Zhang is with the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430079, China.

Digital Object Identifier 10.1109/TUFFC.2022.3163906

is to resolve forward-scattered waveforms from this equation, in order to improve the initial wavefield that is used in the Marchenko equation. In this article, we consider a medium with density contrast and constant propagation velocity. By solving the auxiliary (transmission-based) equation and the (reflection-based) Marchenko equations successively, forward-scattered waveforms and their associated multiples can be included in our Green's function estimates, as we demonstrate numerically. As an example of a potential application for our methodology, we consider Marchenko-based solutions of inverse source problems, which are key in photoacoustic imaging [24]. Applications in media with significant velocity contrast are more challenging and require additional research to be conducted.

II. Marchenko Equation for Reflection Data

In this section, we briefly review a recent derivation of the Marchenko-type representation by [25] for Green's function retrieval from reflection data. Consider the configuration in Fig. 1. Let $\mathbf{x} = (x_1, x_2, x_3)$ be a location in 3-D space, where the x_3-axis is pointing downward. Volume $\mathbb{D}$ is bounded by the horizontal boundaries $\partial\mathbb{D}_U$ and $\partial\mathbb{D}_L$, which are located at depth levels $x_{3,U}$ and $x_{3,L}$, respectively. A lossless acoustic medium is characterized by the propagation velocity $c(\mathbf{x})$ and the mass density $\rho(\mathbf{x})$. We emphasize that all of our methodologies could be extended to include a free surface at the upper boundary [26] or, more generally, to allow arbitrary medium properties above this level [27].

An acoustic pressure field $p(\mathbf{x}, t)$ can be expressed as a function of space $\mathbf{x}$ and time t. This field can be transformed to the frequency domain by the Fourier transform

$$p(\mathbf{x}, \omega) = \int_{-\infty}^{+\infty} p(\mathbf{x}, t) e^{i\omega t} \mathrm{d}t \tag{1}$$

where ω is the angular frequency. Wave propagation is assumed to obey the acoustic wave equation

$$\mathcal{L}p = i\omega q \tag{2}$$

with $q(\mathbf{x}, \omega)$ being a volume-injection rate density source. Furthermoe, operator $\mathcal{L}$ is defined as

$$\mathcal{L} = \partial_i \frac{1}{\rho} \partial_i + \frac{\omega^2}{\rho c^2} \tag{3}$$

where ∂_i is the spatial derivative in the i-direction and Einstein's summation convention applies. Let volume $\mathbb{D}$ be source-free such that $\forall \mathbf{x} \in \mathbb{D} : q(\mathbf{x}, \omega) = 0$. We assume that the wavefield is recorded at $\partial\mathbb{D}_U$, where it can be decomposed into downgoing constituents p^+ and upgoing constituents p^-, such that $p = p^+ + p^-$. It has been shown that the wavefield at any location $\mathbf{x} \in \mathbb{D}$ may then be expressed as [25]

$$p(\mathbf{x}, \omega) = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, \omega) p^-(\mathbf{x}_U, \omega) \mathrm{d}\mathbf{x}_U + \int_{\partial\mathbb{D}_U} F_U^\star(\mathbf{x}, \mathbf{x}_U, \omega) p^+(\mathbf{x}_U, \omega) \mathrm{d}\mathbf{x}_U. \tag{4}$$

In this representation, superscript $\star$ denotes complex conjugation and it is assumed that evanescent waves at and above $\partial\mathbb{D}_U$ can be neglected. Furthermore, F_U is a so-called focusing function, which focuses at the upper boundary and obeys wave (2) with $q = 0$. This function is subject to the focusing condition [25]

$$F_U(\mathbf{x}, \mathbf{x}_U, \omega)|_{x_3 = x_{3,U}} = \delta(\mathbf{x}_H - \mathbf{x}_{H,U}) \tag{5}$$

where F_U is upgoing at and above $\partial\mathbb{D}_U$. In (5), δ is a (2-D) Dirac delta distribution and $\mathbf{x}_H = (x_1, x_2)$ denotes the horizontal coordinates. For the wavefield p, we assume that a vertical dipole source is located at $\mathbf{x}'_U \in \partial\mathbb{D}'_U$, just above $\partial\mathbb{D}_U$ (see Fig. 1). This results in the (dipole) Green's function

$$\Gamma(\mathbf{x}, \mathbf{x}'_U, \omega) = \frac{-2}{i\omega\rho'_U} \partial'_{3,U} G(\mathbf{x}, \mathbf{x}'_U, \omega) \tag{6}$$

where $\partial'_{3,U}$ is a vertical partial derivative applied at $\mathbf{x}'_U$ and ρ'_U is the density at $\partial\mathbb{D}'_U$. In (6), $G(\mathbf{x}, \mathbf{x}'_U, \omega)$ is Green's function of a monopole source at $\mathbf{x}'_U$, evaluated at $\mathbf{x}$, obeying (2) with $q = \delta(\mathbf{x} - \mathbf{x}'_U)$. When $\mathbf{x}'_U$ approaches $\partial\mathbb{D}_U$ in the limit $\epsilon = x_{3,U} - x'_{3,U} \to 0$ (see Fig. 1), it can be deduced that the downgoing part of the dipole response obeys [28]

$$\lim_{x'_{3,U} \to x_{3,U}} \Gamma^+(\mathbf{x}, \mathbf{x}'_U, \omega)\Big|_{x_3 = x_{3,U}} = \delta(\mathbf{x}_H - \mathbf{x}'_{H,U}). \tag{7}$$

When we substitute $p(\mathbf{x}, \omega) = \Gamma(\mathbf{x}, \mathbf{x}'_U, \omega)$ into (4) and apply (7), it follows that:

$$\Gamma(\mathbf{x}, \mathbf{x}'_U, \omega) = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, \omega) \Gamma^-(\mathbf{x}_U, \mathbf{x}'_U, \omega) \mathrm{d}\mathbf{x}_U + F_U^\star(\mathbf{x}, \mathbf{x}'_U, \omega). \tag{8}$$

In this representation, $\Gamma^-(\mathbf{x}_U, \mathbf{x}'_U, \omega)$ can be interpreted as the (upgoing) reflection response of the medium recorded at $\mathbf{x}_U$, stemming from a dipole source at $\mathbf{x}'_U$. We wish to express this result in the time domain with help of the inverse Fourier transform, which is defined for an arbitrary wavefield as the inverse of (1), that is [29]

$$p(\mathbf{x}, t) = \frac{1}{\pi} \Re \left[\int_0^\infty p(\mathbf{x}, \omega) e^{-i\omega t} \mathrm{d}\omega \right] \tag{9}$$

where it is assumed that $p(\mathbf{x}, t)$ is real-valued and $\Re$ denotes the real part. With help of these definitions, (8) can be rewritten

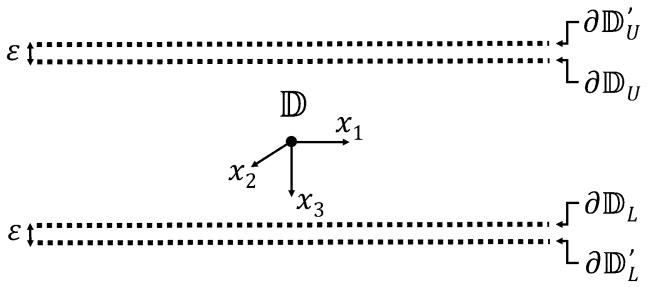


Fig. 1. Configuration: volume $\mathbb{D}$ is enclosed by horizontal boundaries $\partial\mathbb{D}_U$ and $\partial\mathbb{D}_L$ (both extending infinitely in the lateral directions). The coordinate system is also indicated. The medium is nonreflective above $\partial\mathbb{D}_U$. Vertical dipole sources are located at an upper boundary $\partial\mathbb{D}'_U$, at an infinitesimal distance $\epsilon \to 0$ above $\partial\mathbb{D}_U$, and at the lower boundary $\partial\mathbb{D}'_L$, at an infinitesimal distance $\epsilon \to 0$ below $\partial\mathbb{D}_L$. Receivers are located at $\partial\mathbb{D}_U$.

in the time domain as

$$\Gamma_U = (\mathcal{R}_U + \mathcal{Z})F_U. \tag{10}$$

In this expression, we have $\Gamma_U = \Gamma(\mathbf{x}, \mathbf{x}'_U, t)$, while $\mathcal{R}_U$ is an operator for multidimensional convolution with the reflection response $\Gamma^-(\mathbf{x}_U, \mathbf{x}'_U, t)$ at the upper boundary, obeying

$$\mathcal{R}_U F_U = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, t) * \Gamma^-\big(\mathbf{x}_U, \mathbf{x}'_U, t\big)\mathrm{d}\mathbf{x}_U \tag{11}$$

where $*$ denotes temporal convolution. Furthermore, $\mathcal{Z}$ is an operator for time reversal.

Our goal is to retrieve the focusing function from (10). To achieve this goal, we require some prior knowledge about Green's function $\Gamma(\mathbf{x}, \mathbf{x}'_U, t)$. More specifically, we assume that, for each $(\mathbf{x}, \mathbf{x}'_U)$-pair, the traveltime $t_{Ud}(\mathbf{x}, \mathbf{x}'_U)$ of the first (or direct) arrival of this Green's function can be estimated from a macro velocity model [4] (in case of a triplicated wave, t_{Ud} is the traveltime of the first onset [30]). Based on these traveltimes, we design a window operator Θ_U (also referred to as a projector [31]) that removes all arrivals at $t \geq t_{Ud} - t_\epsilon$ (note that our window is not symmetric in time, in contrast to various previous publications). Here, subscript U refers to the upper boundary, where the window operator is applied. In our formulation, a small additional time shift t_ϵ has been included to account for the finite frequency content of the data. In practice, we typically choose t_ϵ as half the temporal support of the source wavelet [32]. Based on causality, we assume that

$$\Theta_U \Gamma_U = 0. \tag{12}$$

In various publications on geophysical applications of the Marchenko equation, the medium is assumed to be layered with moderately curved interfaces [1], [2]. Under these conditions, the focusing function consists of a time-reversed direct wave, which is timed at $-t_{Ud}$, and a coda, which is timed thereafter. A common interpretation is that the direct wave focuses at $\mathbf{x}$ when injected into the medium from the upper boundary [28], while the coda is associated with all (primary and multiple) reflections that are generated between this boundary and $\mathbf{x}$. In media with increasing propagation velocity (which are common in geophysical settings), problems arise at long offsets, due to incorrect handling of refracted waves and postcritical reflections [9], [33]. In the presence of sharp discontinuities in the lateral direction, such as point diffractors, the focusing function contains additional forward-scattered components (i.e., waveforms that have not altered their vertical propagation direction between the upper boundary and $\mathbf{x}$) that are (partly) timed before $-t_{Ud}$ [8], [10], [11]. To allow these (unknown) components in our formulation, we formally partition the focusing function in an initial focusing function F_{Ui}, containing all waveforms in the interval $(-\infty, -t_{Ud}+t_\epsilon]$, and a coda F_{Um}, containing all waveforms in the interval $(-t_{Ud} + t_\epsilon, \infty)$. With help of these definitions, we may write

$$F_U = F_{Ui} + F_{Um}. \tag{13}$$

When the operators $\mathcal{Z}$ and Θ_U are applied successively to the focusing function, it follows from these definitions that:

$$\Theta_U \mathcal{Z} F_U = \Theta_U \mathcal{Z} F_{Um}. \tag{14}$$

When we substitute (13) into (10) and apply operator Θ_U (from the left) to both sides of the result, it follows with help of (12) and (14) that:

$$-\Theta_U \mathcal{R}_U F_{Ui} = \Theta_U(\mathcal{R}_U + \mathcal{Z})F_{Um}. \tag{15}$$

This result is generally known as the Marchenko equation. If the initial focusing function F_{Ui} is known, (15) can be solved for the coda F_{Um}. It is common practice to approximate the initial focusing function F_{Ui} for any relevant $(\mathbf{x}, \mathbf{x}'_U)$-pair by a time-reversed direct wave F_{Ud}, which is typically computed in an approximate macro velocity model [4]. We refer to this practice as the direct-wave approximation. Under this approximation, we ignore additional waveforms F_{Ua} that are (mostly) related to forward scattering, which we define formally as

$$F_{Ua} = F_{Ui} - F_{Ud}. \tag{16}$$

Substitution of $F_{Ui} = F_{Ud} + F_{Ua}$ into (15) yields

$$-\Theta_U \mathcal{R}_U (F_{Ud} + F_{Ua}) = \Theta_U(\mathcal{R}_U + \mathcal{Z})F_{Um}. \tag{17}$$

By assuming $F_{Ua} = 0$ (i.e., the direct-wave approximation), the coda F_{Um} can be resolved from (17) by linear inversion. A common strategy for the inversion is to rewrite the equation by a Neumann series expansion, which is guaranteed (at least in 1-D for infinite frequency content) to converge as long as the spectral radius of operator $\mathcal{R}_U$ is less than one [31], [34]. However, a variety of alternative numerical solvers might be employed [31], [34]–[36]. For the construction of operator $\mathcal{R}_U$, we require access to a complete, well-sampled reflection response [37], [38] and sufficient aperture [39]. Modifications of the methodology have been proposed to allow for gaps in the acquisition design [40] and imperfect sampling [41]. Once the coda F_{Um} is resolved, the complete focusing function can be constructed with (13), and eventually, Green's function follows from (10). In this section, we have derived a Marchenko equation for (vertical) dipole Green's functions, see (6). However, the theory can be modified for the retrieval of monopole Green's functions, see [25].

III. Auxiliary Equation for Transmission Data

It is well known that forward-scattered waveforms cannot be accurately retrieved under the direct-wave approximation [8]–[11]. Ideally, we would like to add the additional components F_{Ua} to F_{Ud}, prior to solving the Marchenko equation (17). We show in the following that, when auxiliary transmission data are available, some of these components can be recovered. We acquire these data by placing additional dipole sources at the lower boundary $\partial\mathbb{D}'_L$ (which is located just below $\partial\mathbb{D}_L$, see Fig. 1). We define their associated (dipole) Green's functions as

$$\Gamma\big(\mathbf{x}, \mathbf{x}'_L, \omega\big) = \frac{-2}{i\omega\rho'_L}\partial'_{3,L} G\big(\mathbf{x}, \mathbf{x}'_L, \omega\big) \tag{18}$$

with $\mathbf{x}'_L \in \partial\mathbb{D}'_L$, $\partial'_{3,L}$ denoting the vertical partial derivative at $\mathbf{x}'_L$ and ρ'_L being the density at $\partial\mathbb{D}'_L$. Here, $G(\mathbf{x}, \mathbf{x}'_L, \omega)$ is a monopole Green's function, obeying (2) with $q = \delta(\mathbf{x} - \mathbf{x}'_L)$.

When we substitute $p = \Gamma(\mathbf{x}, \mathbf{x}'_L, \omega)$ into (4), it follows $\forall \mathbf{x} \in \mathbb{D}$ that:

$$\Gamma\big(\mathbf{x}, \mathbf{x}'_L, \omega\big) = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, \omega)\Gamma^-\big(\mathbf{x}_U, \mathbf{x}'_L, \omega\big)\mathrm{d}\mathbf{x}_U \tag{19}$$

where we used the fact that the medium is nonreflective above $\partial\mathbb{D}_U$ such that $\Gamma^+(\mathbf{x}, \mathbf{x}'_L, \omega)|_{x_3=x_{3,U}} = 0$. In this representation, $\Gamma^-(\mathbf{x}_U, \mathbf{x}'_L, \omega)$ can be interpreted as the (upgoing) transmission response of the medium recorded at $\mathbf{x}_U$, stemming from a dipole source at $\mathbf{x}'_L$. After inverse Fourier transformation, (19) can be compactly rewritten as

$$\Gamma_L = \mathcal{T}_{LU} F_U \tag{20}$$

where $\Gamma_L = \Gamma(\mathbf{x}, \mathbf{x}'_L, t)$. Furthermore, $\mathcal{T}_{LU}$ is an operator for multidimensional convolution with the transmission response, obeying

$$\mathcal{T}_{LU} F_U = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, t) * \Gamma^-\big(\mathbf{x}_U, \mathbf{x}'_L, t\big)\mathrm{d}\mathbf{x}_U. \tag{21}$$

Once more, we assume that the traveltimes $t_{Ld}(\mathbf{x}, \mathbf{x}'_L)$ of the first (or direct) arrivals of $\Gamma(\mathbf{x}, \mathbf{x}'_L, t)$ can be estimated from a macro velocity model such that an operator Θ_L can be constructed, which mutes all arrivals at $t \geq t_{Ld} - t_\epsilon$ (where t_ϵ is a small time shift, as defined earlier, and subscript L denotes the lower boundary, where the window operator is applied). Akin to (12), causality leads to the assumption that

$$\Theta_L \Gamma_L = 0. \tag{22}$$

When we apply operator Θ_L to (20), it follows straight from (22) that F_U should be in the nullspace of operator $\Theta_L \mathcal{T}_{LU}$:

$$\Theta_L \mathcal{T}_{LU} F_U = 0. \tag{23}$$

We may substitute $F_U = F_{Ud} + F_{Ua} + F_{Um}$ into (23) and rewrite the result strategically as

$$-\Theta_L \mathcal{T}_{LU}(F_{Ud} + F_{Um}) = \Theta_L \mathcal{T}_{LU} F_{Ua}. \tag{24}$$

We refer to (24) as our auxiliary equation for transmission data. In this article, we investigate whether this equation can be solved in a medium with density contrast and constant propagation velocity. At first glance, the left-hand side of (24) seems to depend on both F_{Ud} and F_{Um}. However, since the waveforms in F_{Um} are timed after the time-reversed direct wave, they reside mainly in the nullspace of $\Theta_L \mathcal{T}_{LU}$. Therefore, we assume that $-\Theta_L \mathcal{T}_{LU} F_{Um} \approx 0$. Consequently, (24) can be rewritten/approximated as

$$-\Theta_L \mathcal{T}_{LU} F_{Ud} = \Theta_L \mathcal{T}_{LU} F_{Ua}. \tag{25}$$

In Fig. 2(a), we show that the presence of forward-scattered waveforms that are generated below $\mathbf{x}$ yield $-\Theta_L \mathcal{T}_{LU} F_{Ud} = 0$. On the other hand, when forward-scattered waveforms are generated above $\mathbf{x}$, we find $-\Theta_L \mathcal{T}_{LU} F_{Ud} \neq 0$, as shown in Fig. 2(b). Based on (25), these data should match $\Theta_L \mathcal{T}_{LU} F_{Ua}$, resulting in a linear inverse problem that can be solved for F_{Ua}, for instance, by LSQR [42]. Since F_{Ui} and F_{Ud} are only allowed to be nonzero at $(-\infty, -t_{Ud} + t_\epsilon]$, so does $F_{Ua} = F_{Ui} - F_{Ud}$, which we enforce during the inversion by restricting the unknown quantity to this time interval.

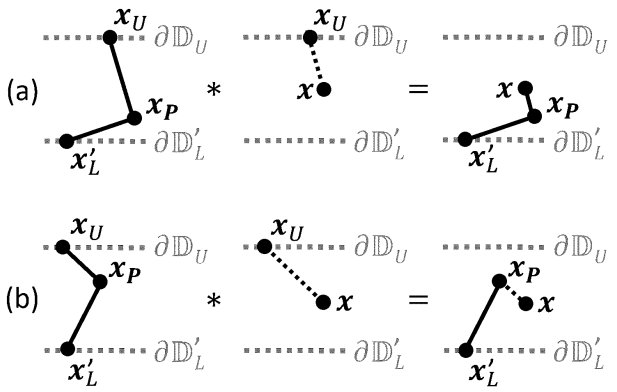


Fig. 2. Let $\mathbf{x}_P$ be a scattering point in a medium with constant propagation velocity, generating a forward-scattered event with traveltime $t_d(\mathbf{x}_U, \mathbf{x}_P) + t_d(\mathbf{x}_P, \mathbf{x}'_L)$ (where $\mathbf{x}'_L$ is a source at $\partial\mathbb{D}'_L$ and $\mathbf{x}_U$ is a receiver at $\partial\mathbb{D}_U$). (a) Situation where $x_{3,P} > x_3$. When the transmission data are convolved with the direct focusing function $F_{Ud}(\mathbf{x}, \mathbf{x}_U, t)$, a physical event is generated with traveltime $t_d(\mathbf{x}_P, \mathbf{x}'_L) + t_d(\mathbf{x}_P, \mathbf{x})$. This process is illustrated in the figure, where the solid and dashed black lines indicate causal and acausal raypaths, respectively (the dashed gray lines designate $\partial\mathbb{D}_U$ and $\partial\mathbb{D}'_L$, while the black dots indicate specific locations). From the triangle inequality, it follows that $t_d(\mathbf{x}_P, \mathbf{x}'_L) + t_d(\mathbf{x}_P, \mathbf{x}) \geq t_d(\mathbf{x}, \mathbf{x}'_L)$. Hence, the event maps at or after the direct arrival of $\Gamma_L = \Gamma(\mathbf{x}, \mathbf{x}'_L, t)$, resulting in $\Theta_L \mathcal{T}_{LU} F_{Ud} = 0$. (b) Situation where $x_{3,P} < x_3$. Now, the generated event is a nonphysical arrival with traveltime $t_d(\mathbf{x}_P, \mathbf{x}'_L) - t_d(\mathbf{x}_P, \mathbf{x})$. From the triangle inequality, it follows that $t_d(\mathbf{x}_P, \mathbf{x}'_L) \leq t_d(\mathbf{x}_P, \mathbf{x}) + t_d(\mathbf{x}, \mathbf{x}'_L)$ or $t_d(\mathbf{x}_P, \mathbf{x}'_L) - t_d(\mathbf{x}_P, \mathbf{x}) \leq t_d(\mathbf{x}, \mathbf{x}'_L)$. Hence, the event maps before or at the direct arrival of $\Gamma_L = \Gamma(\mathbf{x}, \mathbf{x}'_L, t)$, resulting in $\Theta_L \mathcal{T}_{LU} F_{Ud} \neq 0$ as long as $\mathbf{x}_P$, $\mathbf{x}$ and $\mathbf{x}'_L$ are not collinear. In the collinear case, F_{Ua} overlaps with F_{Ud} (i.e., the propagation direction is not altered by scattering at $\mathbf{x}_P$) and cannot be recovered.

Forward-scattered components can only be resolved if they are kinematically separated from the direct wave such that they reside outside the nullspace of $\Theta_L \mathcal{T}_{LU}$. The dependence of operator Θ_L on t_ϵ reveals that the separation of waveforms that can be recovered from waveforms that cannot be recovered is intimately related to the frequency content of the data. This observation also means that the transmission loss that forward-scattering imposes on the direct wave should be formally included in our definition of F_{Ud}. In practice, we neglect these effects by approximating F_{Ud} in a macro velocity model, which could lead to amplitude mismatches in the retrieved wavefields.

Once F_{Ua} is resolved, we may evaluate the Marchenko equation again with the (conventional) workflow that was described in the previous section. However, this time we include F_{Ua} in the left-hand side of (17) such that we can retrieve (estimates of) waveforms in the coda F_{Um} beyond the direct-wave approximation.

IV. Numerical Example

In this section, we apply the proposed methodology to ultrasonic data from a 2-D numerical experiment in the 0–150-kHz frequency range. In Fig. 3, we show our synthetic model, which contains two rectangular density contrasts. These contrasts have been intentionally designed such that their corners (labeled as A, B, ..., H) and vertical interfaces (labeled as AB, CD, EF, and GH) generate forward-scattered waveforms that are not handled well by the conventional Marchenko methodology, which we want to improve on by incorporating

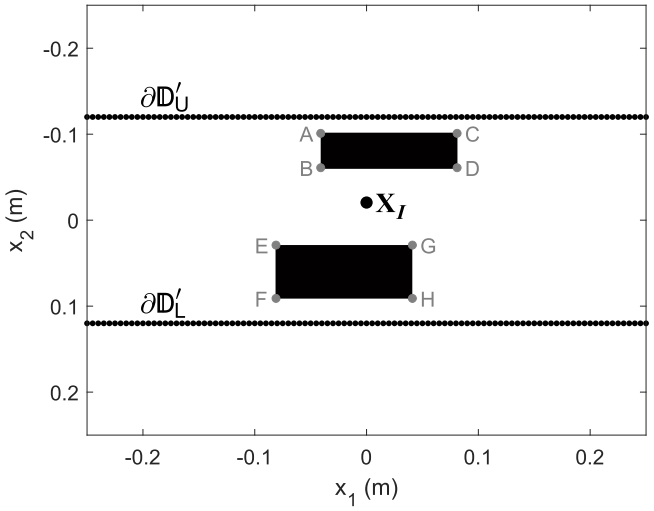


Fig. 3. Configuration for the 2-D numerical experiment. The medium has a constant velocity of 1500 m/s and a background density $\rho_{\text{white}} = 1000$ kg/m^3. The two black rectangles are density anomalies with $\rho_{\text{black}} = 2000$ kg/m^3. The gray dots (labeled as A, B, . . . , H) denote sharp corners that generate forward-scattered waveforms. The upper boundary $\partial\mathbb{D}'_U$ contains 101 coinciding vertical force (dipole) sources and pressure receivers with a spacing of 5 mm. The lower boundary $\partial\mathbb{D}'_L$ contains 101 vertical force (dipole) sources with a spacing of 5 mm. Additional taper zones have been added to these arrays at the intervals $x_1 \in [-0.5\text{ m}, -0.25\text{ m})$ and $x_1 \in (-0.25\text{ m}, 0.5\text{ m}]$ to eliminate truncation artifacts of the spatial integrals. The black dot marks the specific location $\mathbf{x}_I = (0, -0.0205\text{ m})$, where we retrieve our focusing function and Green's function.

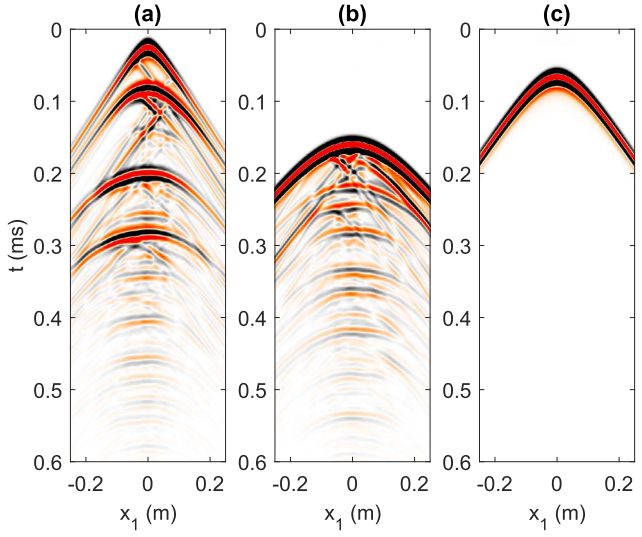


Fig. 4. (a) Reflection response: $\forall \mathbf{x}'_U \in \partial\mathbb{D}'_U : \Gamma^-(\mathbf{x}_0, \mathbf{x}'_U, t)$. (b) Transmission response: $\forall \mathbf{x}'_L \in \partial\mathbb{D}'_L : \Gamma^-(\mathbf{x}_0, \mathbf{x}'_L, t)$. In both panels, we have set $\mathbf{x}_0 = (0, -0.120\text{ m})$ as a reference location. (c) Analytic dipole Green's function in a homogeneous background medium: $\forall \mathbf{x}'_U \in \partial\mathbb{D}'_U : \Gamma_0(\mathbf{x}_I, \mathbf{x}'_U, t)$. All panels are convolved with the wavelet and clipped at 10% of the maximum amplitude.

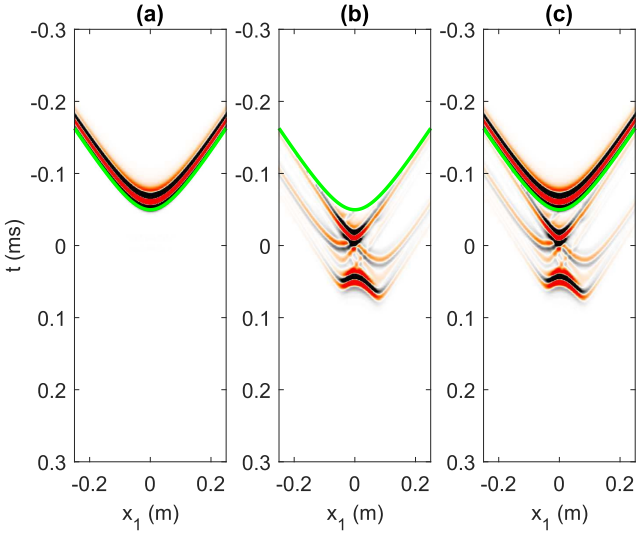


Fig. 5. (a) Direct focusing function F_{Ud}. (b) Coda of the focusing function F_{Um} as retrieved by solving the Marchenko equation under the direct-wave approximation $F_{Ua} = 0$. (c) Updated focusing function F_U, as obtained by adding the retrieved coda F_{Um} to F_{Ud}. All panels are clipped at 10% of the maximum amplitude of the direct wave.

auxiliary transmission data. Our aim is to retrieve F_U, Γ_U and Γ_L at the specific location $\mathbf{x}_I = (0, -0.0205\text{ m})$, which is shown in Fig. 3.

As a wavelet, we take the second derivative of a Gaussian function (also known as a Ricker wavelet) with a 50-kHz peak frequency (for the truncation operators, we choose $t_\epsilon = 20\ \mu$s). Our traces consist of 1024 time samples with $dt = 3(1/3)\ \mu$s. Reflection and transmission data are generated by solving an interface integral equation [43], for which we discretize the mass density model on a spatial grid with a spacing of 2 mm. In Fig. 4, we show sections of these data, as well as an analytic Green's function Γ_0 that we computed in a background homogeneous medium.

For the estimation of the direct focusing function, we would like to reverse the analytic Green's function Γ_0 from 4(c) in time. Unfortunately, this procedure does not take care of transmission loss, which leads to an unacceptable amplitude error. Although methodologies exist to predict the transmission losses from the data [7] (even in angle-dependent mode), we mitigate this problem here with help of a single (angle-independent) scaling factor α, which is determined by

$$\alpha(\mathbf{x}_I) = \frac{\int\int_{\partial\mathbb{D}_U} \Gamma_0(\mathbf{x}_I, \mathbf{x}_U, t) G_0(\mathbf{x}_U, \mathbf{x}_I, t) \mathrm{d}\mathbf{x}_U \mathrm{d}t}{\int\int_{\partial\mathbb{D}_U} \Gamma_0(\mathbf{x}_I, \mathbf{x}_U, t) G(\mathbf{x}_U, \mathbf{x}_I, t) \mathrm{d}\mathbf{x}_U \mathrm{d}t}. \quad (26)$$

In the numerator of (26), we focus Green's function G_0 in the background medium at $\mathbf{x}_I$, while in the denominator, we do the same with Green's function G that is computed in the actual medium. The amplitude ratio between both focusing processes determines our scaling factor α, which can be interpreted as an estimate of the average inverse transmission loss for propagation between $\partial\mathbb{D}'_U$ and $\mathbf{x}_I$. By this procedure, we find $\alpha = 1.1141$ for the focal point that is specified in Fig. 3. The direct focusing function is then estimated as $F_{Ud} \approx \alpha \mathcal{Z} \Gamma_0$ and is shown in Fig. 5(a).

First, we solve Marchenko (17) in a conventional manner, using reflection data only. We do so by evaluating the first ten terms of its associated Neumann series under the direct-wave approximation $F_{Ua} = 0$, see Fig. 5(b). The retrieved coda F_{Um} is added to F_{Ud} and the result is shown in Fig. 5(c).

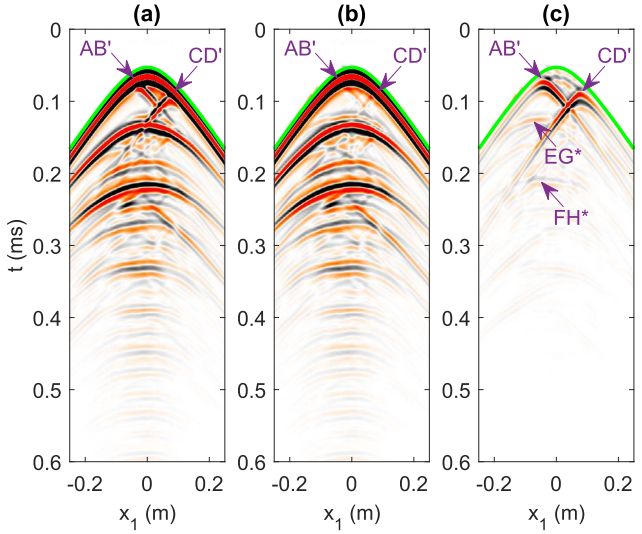


Fig. 6. (a) Reference Green's function Γ_U, as obtained by forward modeling. (b) Same Green's function, as retrieved by solving Marchenko equation under the direct-wave approximation $F_{Ua} = 0$. (c) Difference between Fig. 6(b) and (a). The green curve marks the traveltime of $t_{Ud} - t_\epsilon$. All panels are clipped at 10% of the maximum amplitude of the direct wave.

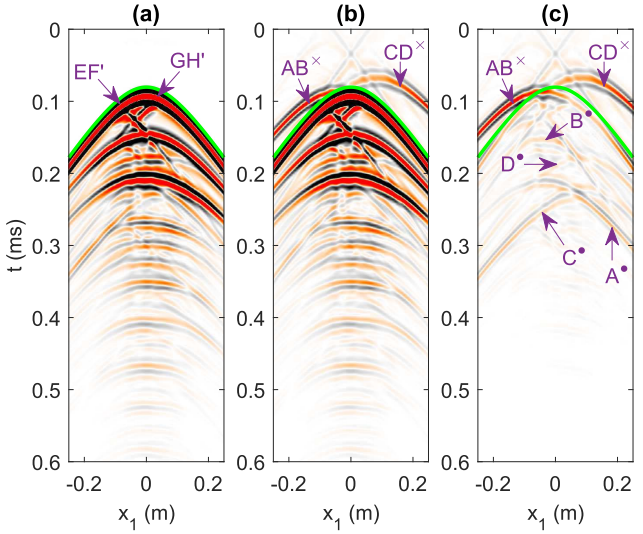


Fig. 7. (a) Reference Green's function Γ_L, as obtained by forward modeling. (b) Same Green's function, as retrieved by solving Marchenko equation under the direct-wave approximation $F_{Ua} = 0$. (c) Difference between Fig. 7(b) and (a). The green curve marks the traveltime of $t_{Ld} - t_\epsilon$. All panels are clipped at 10% of the maximum amplitude of the direct wave.

For our reference, we compute the desired Green's function Γ_U directly from the model parameters by solving an interface integral equation [43], see Fig. 6(a). In Fig. 6(b), we show the same Green's function, as retrieved by the Marchenko equation under the direct-wave approximation $F_{Ua} = 0$ [which is obtained by substituting the retrieved focusing function from Fig. 5(c) into (10)]. The difference between the retrieved and reference Green's functions is given in Fig. 6(c). Events AB′ and CD′, which are shown in Fig. 6, relate to forward-scattered waveforms that have not been retrieved accurately (where $'$ is used to indicate forward-scattered waveforms). These events are mainly generated by the vertical interfaces AB and CD (as shown in Fig. 3). A similar statement can be made about the (weaker) events EG* and FH*, which seems to be related to the reflections of event AB′ at the horizontal interfaces EG and FH in Fig. 3 (where * is used to indicate reflected waveforms).

Green's function Γ_L can be estimated by applying the transmission operator to the retrieved focusing function F_U, see (20). In Fig. 7, we compare the result of this procedure with a reference Green's function that we found by solving an interface integral equation [43]. Events EF′ and GH′ relate to forward-scattered waveforms that have been retrieved accurately. These events are mainly generated by the vertical interfaces EF and GH (as shown in Fig. 3). Based on Fig. 2(a), it is clear that these waveforms map in the interval $[t_{Ld} - t_\epsilon, \infty)$ (i.e., below the green curve in the figure). Hence, they will be in the nullspace of the operator $\Theta_L \mathcal{T}_{LU}$. On the other hand, events AB$^\times$ and CD$^\times$ are artifacts that are generated by forward-scattered waveforms from the vertical interfaces AB and CD in Fig. 3 (where $^\times$ is used to indicate nonphysical events). We confirm that these artifacts map mostly in the interval $(-\infty, t_{Ld} - t_\epsilon)$ (i.e., above the green curve in the figure), as we already predicted in Fig. 2(b). Hence, $\Theta_L \mathcal{T}_{LU} F_U \neq 0$, which violates (23). This observation can be exploited to retrieve the missing components of the

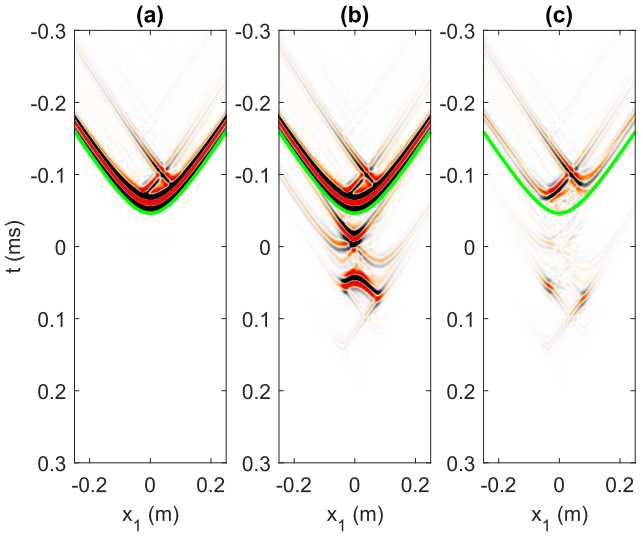


Fig. 8. (a) Initial focusing function F_{Ui} as obtained by adding F_{Ua} [which is retrieved by least-squares inversion of (25)] to F_{Ud}. (b) Focusing function F_U as retrieved by solving the Marchenko equation with $F_{Ui} = F_{Ud} + F_{Ua}$. (c) Difference between Figs. 8(b) and 5(c). Furthermore, all settings are similar as in Fig. 5.

focusing function F_{Ua} from the auxiliary equation, as we do shortly. Furthermore, we have indicated A$^\bullet$, B$^\bullet$, C$^\bullet$, and D$^\bullet$ in Fig. 7(c). We interpret these as diffractions (indicated by $^\bullet$) from the corners A, B, C, and D in Fig. 3, which have not been retrieved accurately.

Next, we show how the missing components of the focusing function can be resolved from the transmission data. To achieve this, we solve (25) for F_{Ua} on the interval $(-\infty, -t_{Ud} + t_\epsilon]$ by 20 iterations of LSQR [42]. We add our solution to F_{Ud} and show the result in Fig. 8(a). Then, we use F_{Ud} and F_{Ua} to invert (17) for F_{Um} by evaluating the first ten terms of its associated Neumann series. This results in a renewed estimate of the focusing function

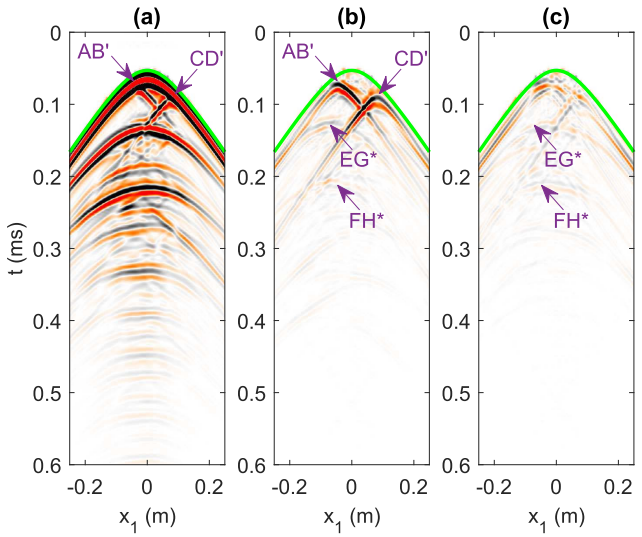


Fig. 9. (a) Green's function Γ_U, as retrieved by solving the auxiliary equation and the Marchenko equation successively. (b) Difference between Figs. 9(a) and 6(b). (c) Difference between Figs. 9(a) and 6(a). Furthermore, all settings are similar as in Fig. 6.

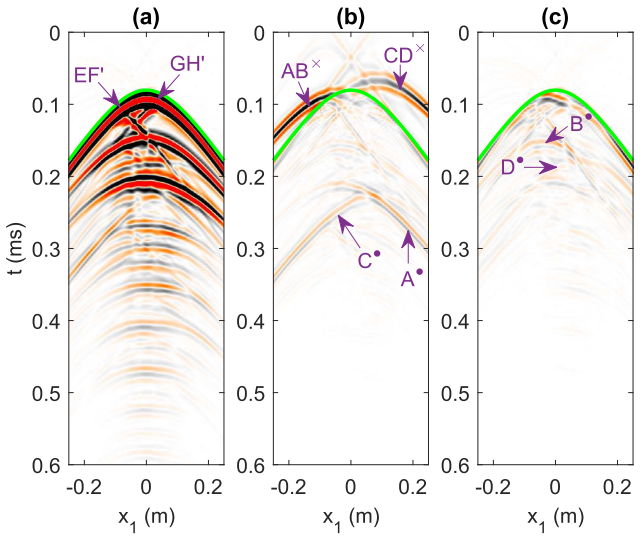


Fig. 10. (a) Green's function Γ_L, as retrieved by solving the auxiliary equation and the Marchenko equation successively. (b) Difference between Figs. 10(a) and 7(b). (c) Difference between Figs. 10(a) and 7(a). Furthermore, all settings are similar as in Fig. 7.

$F_U = F_{Ud} + F_{Ua} + F_{Um}$, see Fig. 8(b). In Fig. 8(c), we show the difference between Figs. 8(b) (obtained with transmissions) and 5(c) (obtained without transmissions). Several events can be observed in this figure, not only on the interval $(-\infty, -t_{Ud} + t_\epsilon]$ (i.e., updates of F_{Ua}) but also on $(-t_{Ud} + t_\epsilon, \infty)$ (i.e., updates of F_{Um}). These events help us to improve the retrieval of forward-scattered waveforms and their associated multiple reflections in Green's functions Γ_U and Γ_L, as we demonstrate next.

In Fig. 9(a), we show Green's function Γ_U, as obtained by substituting our renewed estimate of the focusing function into (10). In Fig. 9(b), we show the difference between Figs. 9(a) (obtained with transmissions) and 6(b) (obtained without transmissions). As indicated in this figure, the forward-scattered waveforms that were indicated as AB′ and CD′, as well as EG* and FH* (relating to high-order scattering), can be recognized. In Fig. 9(c), we show the difference between Fig. 9(a) and the reference Green's function in Fig. 6(a). By comparing this result with Fig. 6(c), we see that the forward-scattered waveforms (AB′ and CD′) and some of their associated multiples (EG* and FH*) have been better resolved. We also observe that our result is not optimal, which we assume to be attributed to the incorrect amplitude spectrum of the direct focusing function (where transmission losses have not been handled well) and finite aperture.

We may also retrieve Green's function Γ_L by substituting our renewed estimate of the focusing function into (20). The result of this operation (including the associated difference plots) is shown in Fig. 10. Comparing this result with Fig. 7, it is clear that the artifacts AB$^\times$ and CD$^\times$ have been suppressed, as enforced by the inversion of (25). We also observe that the diffractions A$^\bullet$ and C$^\bullet$ have been retrieved better, which is not so much the case for the diffractions B$^\bullet$ and D$^\bullet$. The latter difference might be attributed to finite aperture. More specifically: the Fresnel zones [44] that are required for the retrieval of B$^\bullet$ and D$^\bullet$ are more extended and closer to the edges of the array compared to the those of A$^\bullet$ and C$^\bullet$. Consequently, they seem to suffer more from the finite aperture of the acquisition array.

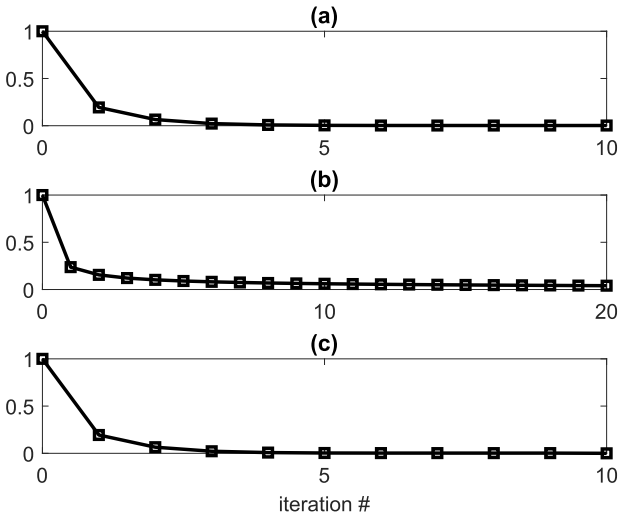


Fig. 11. Frobenius norm of the residual (normalized by the residual at iteration 0) when solving (a) the Marchenko equation under the direct-wave approximation $F_{Ua} = 0$, (b) the auxiliary equation and (c) the Marchenko equation with an updated initial focusing function $F_{Ui} = F_{Ud} + F_{Ua}$.

In Fig. 11(a), we provide the convergence curve of the Marchenko equation, when solved under the direct-wave approximation. An equivalent curve of the auxiliary equation can be found in Fig. 11(b). In Fig. 11(c), we show the convergence curve when solving the Marchenko equation after updating the initial focusing function. Note that this curve is not too different from Fig. 11(a).

V. Application to Inverse Source Problems

As a potential application of the proposed method, we focus our attention on inverse source problems, which are common in photoacoustic imaging [24]. We consider the configuration

in Fig. 1 and assume that a distribution of sources is located inside the volume $\mathbb{D}$. We characterize this source distribution as $q_e(\mathbf{x}, t)$, where subscript e stands for "experiment." An important assumption is that all sources are ignited at $t = 0$, with a zero-phase wavelet (which might have to be enforced in practice by additional preprocessing). Let $\forall \mathbf{x}_U \in \partial\mathbb{D}_U : p_e(\mathbf{x}_U, t)$ be the recorded wavefield at the upper boundary. Our goal is to reconstruct the initial pressure distribution $\forall \mathbf{x} \in \mathbb{D} : p_e(\mathbf{x}, t = 0)$ from these recordings. A common strategy is to propagate the data back into the volume with Green's functions from a (typically smooth) macro model [45]. However, a range of alternative solutions exists [46]. When reflection and transmission data are available, we may compute focusing functions by the methodology that we have derived in this article and rely on (4) for the required wavefield reconstruction process. Unfortunately, (4) can only be applied to a wavefield p if its associated source distribution q is zero throughout the volume $\mathbb{D}$, which is obviously not the case for $p = p_e$ and $q = q_e$. To overcome this problem, we symmetrize the wavefield by the following operation:

$$p_h(\mathbf{x}, t) = p_e(\mathbf{x}, t) + p_e(\mathbf{x}, -t). \tag{27}$$

Here, subscript h stands for "homogeneous," referring to the fact that p_h is a solution to the homogeneous (=source-free) wave equation [16], [47]; i.e., p_h satisfies (2) with $q = 0$. Consequently, we can use (4) to reconstruct p_h throughout volume $\mathbb{D}$, given our recorded data at the upper boundary $\partial\mathbb{D}_U$. To facilitate this process, we substitute $p = p_h$ into (4) and rewrite the result in the time domain as

$$p_h(\mathbf{x}, t) = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, t) * p_h^-(\mathbf{x}_U, t) \mathrm{d}\mathbf{x}_U + \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, -t) * p_h^+(\mathbf{x}_U, t) \mathrm{d}\mathbf{x}_U. \tag{28}$$

Next, we realize that $\forall \mathbf{x}_U \in \partial\mathbb{D}_U : p_h^-(\mathbf{x}_U, t) = p_e(\mathbf{x}_U, t)$ and $p_h^+(\mathbf{x}_U, t) = p_e(\mathbf{x}_U, -t)$. Substitution into (28) yields

$$p_h(\mathbf{x}, t) = \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, t) * p_e(\mathbf{x}_U, t) \mathrm{d}\mathbf{x}_U + \int_{\partial\mathbb{D}_U} F_U(\mathbf{x}, \mathbf{x}_U, -t) * p_e(\mathbf{x}_U, -t) \mathrm{d}\mathbf{x}_U. \tag{29}$$

We can utilize this result to reconstruct the symmetrized data p_h throughout the volume $\mathbb{D}$ from the recordings p_e at the boundary $\partial\mathbb{D}_U$. To retrieve the initial pressure field, we may evaluate

$$p_e(\mathbf{x}, t = 0) = \frac{1}{2} p_h(\mathbf{x}, t = 0) \tag{30}$$

where we used (27). In the following, we will demonstrate the potential of the proposed strategy with a simple synthetic example, based on the configuration that we presented earlier in Fig. 3. We choose the following (point) source distribution: $q_e(\mathbf{x}, t) = \delta(\mathbf{x} - \mathbf{x}_I) S(t)$, where $\mathbf{x}_I = (0, -0.0205 \text{ m})$ and $S(t)$ is the second derivative of a Gaussian wavelet with a 50 kHz peak frequency, as in our previous example. We compute the response $\forall \mathbf{x}_U \in \partial\mathbb{D}_U : p_e(\mathbf{x}_U, t)$ of this source distribution by forward modeling. Then, we reconstruct the symmetrized wavefield $\forall \mathbf{x} \in \mathbb{D} : p_h(\mathbf{x}, t)$ by (29), with help of focusing functions that we retrieve from reflection and transmission data. Finally, we evaluate $(1/2) p_h(\mathbf{x}, t = 0)$ to estimate the initial pressure field $p_e(\mathbf{x}, t = 0)$, following (30).

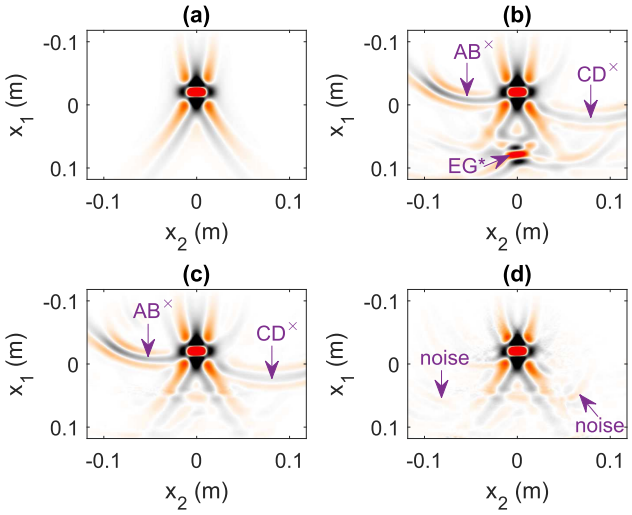


Fig. 12. Reconstructed initial pressure field $\forall \mathbf{x} \in \mathbb{D} : p_e(\mathbf{x}, t = 0)$. In panel (a) (which serves as a reference for the other panels), data have been computed in a homogeneous background medium and F_{Ud} acted as our focusing function. In panels (b)–(d), data have been computed in the heterogeneous medium of Fig. 3. (b) F_{Ud} acted as our focusing function. (c) The focusing function is obtained by solving the Marchenko equation under the direct-wave approximation $F_{Ua} = 0$. (d) The focusing function is obtained by solving the auxiliary equation (with $F_{Um} = 0$) and the Marchenko equation successively. All panels are clipped at 10% of the maximum amplitude.

For reference, we start with data $\forall \mathbf{x}_U \in \partial\mathbb{D}_U : p_e(\mathbf{x}_U, t)$ that was computed in a homogeneous medium with $c = 1500$ m/s and $\rho = 1000$ kg/m^3 (i.e., without mass density contrast). We reconstruct the initial pressure distribution from these data, where we let F_{Ud} act as our focusing function. The result of this procedure is shown in Fig. 12(a). Since the model has not generated any scattered or reflected waveform, F_{Ud} is identical to the complete focusing function (i.e., both F_{Ua} and F_{Um} are zero in this case). Consequently, our initial result cannot be improved by any of the approaches that we have discussed in this article. Next, we repeat the exercise for data that were computed in the heterogeneous medium of Fig. 3, but we still choose F_{Ud} as our focusing function for the evaluation of (29). The image that is obtained by this procedure is shown in Fig. 12(b). Compared to Fig. 12(a), we observe two kinds of artifacts. Artifact EG* is related to reflections from the horizontal interface EG (see Fig. 3). This phenomenon can be understood as follows: the initial radiation from $\mathbf{x}_I$ has reflected at the interface EG; this reflection was recorded at $\partial\mathbb{D}_U$ and has been backpropagated by F_{Ud}, generating a mirror image below $\mathbf{x}_I$. Artifacts AB$^\times$ and CD$^\times$ are related to forward scattering. These phenomena can be understood as follows: the initial radiation from $\mathbf{x}_I$ has interacted with the (corners and vertical interfaces of) the density contrast above $\mathbf{x}_I$, generating forward-scattered waveforms. These waveforms have been backpropagated by F_{Ud} (where forward-scattered components have not been accounted for), resulting in additional events/artifacts (all intersecting at $\mathbf{x}_I$). Next, we solve

the Marchenko equation under the direct-wave approximation $F_{Ua} = 0$ and we use the retrieved focusing functions to evaluate (29). This leads to the image in Fig. 12(c). Note that the reflection-based artifact EG* has been removed by this procedure, but the artifacts from forward scattering remain. Finally, we compute the focusing functions by solving the auxiliary equation and the Marchenko equation successively, leading to the image in Fig. 12(d). This time, both types of artifacts have been suppressed significantly. It is also observed that some weaker artifacts have emerged in the image, which is indicated as noise in Fig. 12(d). To quantify the quality of our results, we introduce a global normalized root-mean-square error as

$$E = \sqrt{\frac{\int_{\mathbb{D}} |I(\mathbf{x}) - I_0(\mathbf{x})|^2 \mathrm{d}\mathbf{x}}{\int_{\mathbb{D}} |I_0(\mathbf{x})|^2 \mathrm{d}\mathbf{x}}}. \tag{31}$$

Here, I is our image and I_0 is the associated reference image in Fig. 12(a). For the image I in Fig. 12(b), we find $E = 0.246$. This number is reduced significantly to $E = 0.079$ for the image in Fig. 12(c), and down to $E = 0.059$ in case of Fig. 12(d). This reduction of E suggests that the result has improved. In our numerical example, we have chosen a monopole point source at $\mathbf{x}_I$ as our unknown source distribution. However, the linearity of the equations allows for arbitrary source distributions and mechanisms, as has been demonstrated in equivalent geophysical problems [48]. It is remarkable that—apart from the propagation velocity—no medium parameters are required for the reconstruction.

VI. Discussion

In this article, we have proposed to partition the focusing function at an arbitrary location $\mathbf{x} \in \mathbb{D}$ as

$$F_U = F_{Ud} + F_{Ua} + F_{Um}. \tag{32}$$

Here, F_{Ud} is the direct focusing function, while F_{Ua} and F_{Um} contain all other waveforms in the intervals $(-\infty, -t_{Ud} + t_\epsilon]$ and $(-t_{Ud} + t_\epsilon, \infty)$, respectively. We have assumed that F_{Um} resides in the nullspace of $\Theta_L \mathcal{T}_{LU}$, such that it can be excluded from the auxiliary equation, while F_{Ua} is not in this nullspace, and hence can be recovered. The latter assumption is motivated by Fig. 2(b), where we illustrate that forward-scattered waveforms which are generated above $\mathbf{x}$ are not in the nullspace of $\Theta_L \mathcal{T}_{LU}$. To investigate the validity of the abovementioned assumptions (especially in media with additional velocity contrast), it would be highly valuable to model F_U directly from the medium parameters, for instance by depth extrapolation [49].

From the illustration in Fig. 2(b), it is clear that forward-scattered waveforms can only be retrieved from the auxiliary equation if they are sufficiently delayed (with respect to the direct wave). On a similar note, we have learned before that reflected waveforms can only be retrieved from the Marchenko equation if they are sufficiently delayed [12]. It has been shown that the latter problem can be mitigated by introducing an augmented focusing function [13]–[15]. It might be worthwhile to investigate if a similar strategy could also be applied to the auxiliary equation.

In our study, we have not investigated the effects of finite spatial aperture, which can have a detrimental effect on Marchenko-based Green's function retrieval [38], [39]. In our numerical examples, potential problems have been circumvented by choosing long spatial arrays and applying significant spatial tapers. With shorter arrays, it is well-understood that particular components of Green's functions cannot be retrieved when the stationary points of the underlying integrals are not properly evaluated [10].

Further, we have chosen to demonstrate the validity of our workflow for 2-D wave propagation only. Various publications have emerged recently on the implementation of the Marchenko equation for 3-D wave propagation problems [50]–[52]. Building on these developments, we prospect that a 3-D implementation of the auxiliary equation should also be feasible.

We emphasize that our methodology has been derived under the assumption that the medium is lossless. By utilizing two-sided (rather than single-sided) reflection data and solving an alternative system of Marchenko equations, Green's functions can also be retrieved in dissipative media [53]. A potential direction of further research is to include losses in our formulation of the auxiliary equation, such that forward-scattered waveforms and their associated multiples can also be retrieved in dissipative media. Such an approach might be beneficial not only for acoustic problems but also for related applications that involve electromagnetic wave propagation, where losses typically cannot be neglected [54].

Last but not least, our work may have an impact on elastodynamic Green's function retrieval in solid media (in those fortunate cases where auxiliary transmission data are available). Although the underlying representations of the Marchenko equation are well established for elastodynamic wave propagation [55], [56], it remains challenging to retrieve elastodynamic focusing functions in practice, since they overlap partly with their associated Green's functions in the time-space domain (due to the different velocities of P- and S-waves) and their forward-scattered components are generally unknown [57]. It seems likely that these problems can be mitigated (at least to some extent) with help of auxiliary transmission data, by extending the theory from our paper to the elastodynamic case.

VII. Conclusion

Marchenko-type focusing functions are useful tools to allow Green's function retrieval from single-sided reflection data. We can partition the focusing function in an intial focusing function and a coda. The initial focusing function contains the inverse direct wave and all preceding waveforms, which are typically associated with forward scattering. The coda contains all events after the inverse direct wave, which are typically associated with (primary and multiple) reflections. Given the initial focusing function, we can retrieve the coda from reflection data by solving a multidimensional Marchenko equation. In practice, the initial focusing function is often

approximated by a time-reversed direct wave, which is generally computed in a macro velocity model. Under this (direct-wave) approximation, forward-scattered components of Green's function and their associated multiples cannot be successfully recovered. In this article, we have proposed to mitigate this problem by incorporating additional transmission data. We derived an auxiliary equation, which can be used to resolve the missing components of the initial focusing function from the transmission data by least-squares inversion. Once this is done, we can use our updated initial focusing function and the reflection data to solve the Marchenko equation beyond the direct-wave approximation. This procedure can be used to retrieve forward-scattered constituents of Green's function in a medium with density contrast and constant propagation velocity, as we have demonstrated numerically. Implementing this methodology in a medium with velocity contrast is more challenging and requires additional research to be conducted.

REFERENCES

[1] K. Wapenaar, F. Broggini, E. Slob, and R. Snieder, "Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Green's function retrieval, and their mutual relations," *Phys. Rev. Lett.*, vol. 110, no. 8, Feb. 2013, Art. no. 084301.

[2] K. Wapenaar *et al.*, "Marchenko redatuming, imaging, and multiple elimination and their mutual relations," *Geophysics*, vol. 86, no. 5, pp. WC117–WC140, Sep. 2021.

[3] C. Mildner, F. Broggini, K. de Vos, and J. O. A. Robertsson, "Accurate source wavelet estimation using Marchenko focusing functions," *Geophysics*, vol. 84, no. 6, pp. Q73–Q88, Nov. 2019.

[4] F. Broggini, R. Snieder, and K. Wapenaar, "Data-driven wavefield focusing and imaging with multidimensional deconvolution: Numerical examples for reflection data with internal multiples," *Geophysics*, vol. 79, no. 3, pp. WA107–WA115, May 2014.

[5] C. Mildner, F. Broggini, J. O. A. Robertsson, D.-J. van Manen, and S. Greenhalgh, "Target-oriented velocity analysis using Marchenko-redatumed data," *Geophysics*, vol. 82, no. 2, pp. R75–R86, Mar. 2017.

[6] X.-C. Chen *et al.*, "Marchenko imaging based on self-adaptive traveltime updating," *Appl. Geophys.*, vol. 17, no. 1, pp. 81–91, Mar. 2020.

[7] C. Mildner, M. Dukalski, P. Elison, K. De Vos, F. Broggini, and J. O. A. Robertsson, "True amplitude-versus-offset Green's function retrieval using augmented Marchenko focusing," in *Proc. 81st EAGE Conf. Exhib.*, 2019, pp. 1–5.

[8] I. Vasconcelos, K. Wapenaar, J. van der Neut, C. Thomson, and M. Ravasi, "Using inverse transmission matrices for Marchenko redatuming in highly complex media," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2015, pp. 5081–5086.

[9] L. Diekmann and I. Vasconcelos, "Focusing and Green's function retrieval in three-dimensional inverse scattering revisited: A single-sided Marchenko integral for the full wave field," *Phys. Rev. Res.*, vol. 3, no. 1, Mar. 2021, Art. no. 013206.

[10] J. van der Neut, I. Vasconcelos, and K. Wapenaar, "On Green's function retrieval by iterative substitution of the coupled Marchenko equations," *Geophys. J. Int.*, vol. 203, no. 2, pp. 792–813, Nov. 2015.

[11] D. Vargas, I. Vasconcelos, Y. Sripanich, and M. Ravasi, "Scattering-based focusing for imaging in highly complex media from band-limited, multicomponent data," *Geophysics*, vol. 86, no. 5, pp. WC141–WC157, Sep. 2021.

[12] E. Slob, K. Wapenaar, F. Broggini, and R. Snieder, "Seismic reflector imaging using internal multiples with marchenko-type equations," *Geophysics*, vol. 79, no. 2, pp. S63–S76, Mar. 2014.

[13] M. Dukalski, E. Mariani, and K. de Vos, "Handling short-period scattering using augmented Marchenko autofocusing," *Geophys. J. Int.*, vol. 216, no. 3, pp. 2129–2133, Mar. 2019.

[14] P. Elison, M. S. Dukalski, K. de Vos, D. J. van Manen, and J. O. A. Robertsson, "Data-driven control over short-period internal multiples in media with a horizontally layered overburden," *Geophys. J. Int.*, vol. 221, no. 2, pp. 769–787, May 2020.

[15] H. Peng, M. Dukalski, P. Elison, and I. Vasconcelos, "Data-driven multiple suppression for laterally varying overburden with thin beds," in *Proc. 82nd EAGE Annu. Conf. Exhib.*, 2021, pp. 1–5.

[16] K. Wapenaar, J. Brackenhoff, J. Thorbecke, J. van der Neut, E. Slob, and E. Verschuur, "Virtual acoustics in inhomogeneous media with single-sided access," *Sci. Rep.*, vol. 8, no. 1, p. 2497, Dec. 2018.

[17] T. Cui, T. S. Becker, D.-J. van Manen, J. E. Rickett, and I. Vasconcelos, "Marchenko redatuming in a dissipative medium: Numerical and experimental implementation," *Phys. Rev. A, Gen. Phys.*, vol. 10, no. 4, Oct. 2018, Art. no. 044022.

[18] C. A. da Costa Filho, K. Tant, A. Curtis, A. Mulholland, and C. M. Moran, "Using laboratory experiments to develop and test new Marchenko and imaging methods," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2018, pp. 4352–4356.

[19] G. A. Meles, J. van der Neut, K. W. A. van Dongen, and K. Wapenaar, "Wavefield finite time focusing with reduced spatial exposure," *J. Acoust. Soc. Amer.*, vol. 145, no. 6, pp. 3521–3530, Jun. 2019.

[20] G. Clement and K. Hynynen, "A non-invasive method for focusing ultrasound through the human skull," *Phys. Med. Biol.*, vol. 47, no. 8, pp. 1219–1236, Apr. 2002.

[21] L. Guasch, O. Calderón Agudo, M.-X. Tang, P. Nachev, and M. Warner, "Full-waveform inversion imaging of the human brain," *NPJ Digit. Med.*, vol. 3, no. 1, p. 28, Dec. 2020.

[22] J. Poudel, S. Na, L. V. Wang, and M. A. Anastasio, "Iterative image reconstruction in transcranial photoacoustic tomography based on the elastic wave equation," *Phys. Med. Biol.*, vol. 65, no. 5, Mar. 2020, Art. no. 055009.

[23] M. S. R. Kiraz, R. Snieder, and K. Wapenaar, "Focusing waves in an unknown medium without wavefield decomposition," *JASA Exp. Lett.*, vol. 1, no. 5, May 2021, Art. no. 055602.

[24] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1097–1110, Jun. 2013.

[25] K. Wapenaar, R. Snieder, S. de Ridder, and E. Slob, "Green's function representations for Marchenko imaging without up/down decomposition," *Geophys. J. Int.*, vol. 227, no. 1, pp. 184–203, Jun. 2021.

[26] S. Singh, R. Snieder, J. van der Neut, J. Thorbecke, E. Slob, and K. Wapenaar, "Accounting for free-surface multiples in Marchenko imaging," *Geophysics*, vol. 82, no. 1, pp. R19–R30, Jan. 2017.

[27] M. Ravasi, "Rayleigh-marchenko redatuming for target-oriented, true-amplitude imaging," *Geophysics*, vol. 82, no. 6, pp. S439–S452, Nov. 2017.

[28] K. Wapenaar, J. Thorbecke, J. van der Neut, E. Slob, and R. Snieder, "Review paper: Virtual sources and their responses—Part II: Data-driven single-sided focusing," *Geophys. Prospecting*, vol. 65, no. 6, pp. 1430–1451, Nov. 2017.

[29] R. Bracewell, *The Fourier Transform and its Applications*. New York, NY, USA: McGraw-Hill, 1999.

[30] K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini, E. Slob, and R. Snieder, "Green's function retrieval from reflection data, in absence of a receiver at the virtual source position," *J. Acoust. Soc. Amer.*, vol. 135, no. 5, pp. 2847–2861, May 2014.

[31] M. Dukalski and K. de Vos, "Marchenko inversion in a strong scattering regime including surface-related multiples," *Geophys. J. Int.*, vol. 212, pp. 760–776, Oct. 2017.

[32] J. Thorbecke, E. Slob, J. Brackenhoff, J. van der Neut, and K. Wapenaar, "Implementation of the Marchenko method," *Geophysics*, vol. 82, no. 6, pp. WB29–WB45, Nov. 2017.

[33] L. Zhang, J. Thorbecke, K. Wapenaar, and E. Slob, "Transmission compensated primary reflection retrieval in the data domain and consequences for imaging," *Geophysics*, vol. 84, no. 4, pp. Q27–Q36, Jul. 2019.

[34] E. Slob and L. Zhang, "Unified elimination of 1D acoustic multiple reflection," *Geophys. Prospecting*, vol. 69, no. 2, pp. 327–348, Feb. 2021.

[35] R. S. Santos *et al.*, "An application of the Marchenko internal multiple elimination scheme formulated as a least-squares problem," *Geophysics*, vol. 86, no. 5, pp. WC105–WC116, Sep. 2021.

[36] M. Ravasi and I. Vasconcelos, "An open-source framework for the implementation of large-scale integral operators with flexible, modern high-performance computing solutions: Enabling 3D Marchenko imaging by least-squares inversion," *Geophysics*, vol. 86, no. 5, pp. WC177–WC194, Sep. 2021.

[37] X. Jia, A. Guitton, and R. Snieder, "A practical implementation of subsalt Marchenko imaging with a Gulf of Mexico data set," *Geophysics*, vol. 83, no. 5, pp. S409–S419, Sep. 2018.

[38] H. Peng, I. Vasconcelos, Y. Sripanich, and L. Zhang, "An analysis of acquisition-related subsampling effects on Marchenko focusing, redatuming, and primary estimation," *Geophysics*, vol. 86, no. 5, pp. WC75–WC88, Sep. 2021.

[39] Y. Sripanich and I. Vasconcelos, "Effects of aperture on Marchenko focussing functions and their radiation behaviour at depth," *Geophys. Prospecting*, vol. 67, no. 2, pp. 443–454, Feb. 2019.

[40] C. Haindl, M. Ravasi, and F. Broggini, "Handling gaps in acquisition geometries—Improving marchenko-based imaging using sparsity-promoting inversion and joint inversion of time-lapse data," *Geophysics*, vol. 86, no. 2, pp. S143–S154, Mar. 2021.

[41] J. van IJsseldijk and K. Wapenaar, "Adaptation of the iterative Marchenko scheme for imperfectly sampled data," *Geophys. J. Int.*, vol. 224, no. 1, pp. 326–336, Nov. 2020.

[42] C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," *ACM Trans. Math. Softw.*, vol. 8, no. 1, pp. 43–71, Mar. 1982.

[43] P. M. van den Berg, *Forward and Inverse Scattering Algorithms Based on Contrast Source Integral Equations*. Hoboken, NJ, USA: Wiley, 2021.

[44] K. Wapenaar, D. Draganov, R. Snieder, X. Campman, and A. Verdel, "Tutorial on seismic interferometry: Part 1—Basic principles and applications," *Geophysics*, vol. 75, Sep. 2010, Art. no. 75A195.

[45] M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 71, no. 1, Jan. 2005, Art. no. 016706.

[46] J. Poudel, Y. Lou, and M. A. Anastasio, "A survey of computational frameworks for solving the acoustic inverse problem in three-dimensional photoacoustic computed tomography," *Phys. Med. Biol.*, vol. 64, no. 14, Jul. 2019, Art. no. 14TR01.

[47] M. L. Oristaglio, "An inverse scattering formula that uses all the data," *Inverse Problems*, vol. 5, no. 6, pp. 1097–1105, Dec. 1989.

[48] J. Brackenhoff, J. Thorbecke, and K. Wapenaar, "Monitoring of induced distributed double-couple sources using Marchenko-based virtual receivers," *Solid Earth*, vol. 10, no. 1, pp. 1301–1319, 2019.

[49] P. Elison, M. Dukalski, and D. Van Manen, "Quality control of under-constrained Marchenko equation solvers in complex media using reference focusing functions," in *Proc. 82nd EAGE Annu. Conf. Exhib.*, 2021, pp. 1–5.

[50] A. Lomas and A. Curtis, "Marchenko methods in a 3-D world," *Geophys. J. Int.*, vol. 220, no. 1, pp. 296–307, Jan. 2020.

[51] M. Staring and K. Wapenaar, "Three-dimensional Marchenko internal multiple attenuation on narrow azimuth streamer data of the Santos Basin, Brazil," *Geophys. Prospecting*, vol. 68, no. 6, pp. 1864–1877, 2020.

[52] J. Brackenhoff, J. Thorbecke, G. Meles, V. Koehne, D. Barrera, and K. Wapenaar, "3D Marchenko applications: Implementation and examples," *Geophys. Prospecting*, vol. 70, no. 1, pp. 35–56, Jan. 2022.

[53] E. Slob, "Green's function retrieval and Marchenko imaging in a dissipative acoustic medium," *Phys. Rev. Lett.*, vol. 116, no. 16, Apr. 2016, Art. no. 164301.

[54] B. Yang and E. C. Slob, "Marchenko inversion of GPR data for a 1D dissipative medium," *Prog. Electromagn. Res. M*, vol. 102, pp. 65–79, 2021.

[55] K. Wapenaar and E. Slob, "On the Marchenko equation for multicomponent single-sided reflection data," *Geophys. J. Int.*, vol. 199, no. 3, pp. 1367–1371, Dec. 2014.

[56] C. A. da Costa Filho, M. Ravasi, A. Curtis, and G. A. Meles, "Elastodynamic Green's function retrieval through single-sided Marchenko inverse scattering," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 6, Dec. 2014, Art. no. 063201.

[57] C. Reinicke, M. Dukalski, and K. Wapenaar, "Comparison of monotonicity challenges encountered by the inverse scattering series and the Marchenko demultiple method for elastic waves," *Geophysics*, vol. 85, no. 5, pp. Q11–Q26, Sep. 2020.

Joost van der Neut received the Ph.D. degree *(cum laude)* in applied geophysics from the Delft University of Technology (TU Delft), Delft, The Netherlands, in 2012.

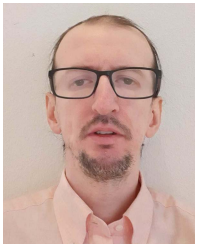
In 2013, he received a Veni Personal Grant from the Dutch Research Council (NWO). In 2017, he started his postdoctoral studies at the Department of Applied Sciences, TU Delft, where he also continued with another postdoctoral position at the Department of Geosciences and Engineering in 2021. His research interests include wave propagation, imaging, and inversion.

Dr. van der Neut received the Best Student Presentation Award (SEG, 2009), the Best Presentation Award (SEG, 2010), the J. Clarence Karcher Award (SEG, 2015), and the Best Paper Award (Geophysical Prospecting, 2016).

Joeri Brackenhoff received the B.Sc. degree *(cum laude)* in applied Earth sciences from the Delft University of Technology, Delft, The Netherlands, in 2014, the M.Sc. degree *(cum laude)* in applied geophysics from the IDEA League, Delft, in 2016, and the Ph.D. degree in applied geophysics from the Delft University of Technology in 2021.

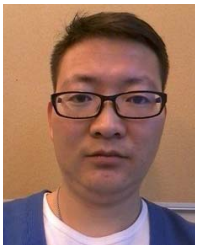
Since 2021, he has been a Postdoctoral Fellow with the Seismology and Wave Physics Group, ETH Zürich, Zürich, Switzerland. His research interests include the study of wave physics for the purpose of seismic imaging and inversion.

Dr. Brackenhoff is a Board Member of the European Association of Geoscientists and Engineers (EAGE) Local Chapter Netherlands and a member of the Society of Exploration Geophysicists (SEG), EAGE, and American Geophysical Union (AGU).



Giovanni Meles received the M.S. degree in physics from the Università Statale di Milano, Milan, Italy, in 2004, and the Ph.D. degree from the Swiss Federal Institute of Technology (ETH Zürich), Zürich, Switzerland, in 2011, under the supervision of A. Green, S. Greenhalgh, and J. van der Kruk.

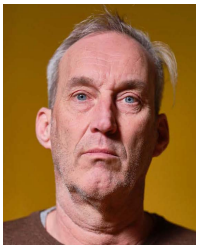
In 2007, he moved to Zürich. From September 2011 to December 2016, he was with the Edinburgh Interferometry Project at The University of Edinburgh, Edinburgh, U.K., where he conducted research in close contact with A. Curtis. In February 2017, he became part of the K. Wapenaar's group at the Delft University of Technology (TU Delft), Delft, The Netherlands, working mainly on Marchenko plane-wave methods and focusing strategies. Since December 2020, he has been working with N. Linde from the University of Lausanne, Lausanne, Switzerland. His research interests include GPR/seismic imaging, full-waveform inversion, and focusing strategies.



Lele Zhang received the Ph.D. degree in applied sciences from the Delft University of Technology (TU Delft), Delft, The Netherlands, in 2019.

Since the beginning of 2020, he has worked as a Postdoctoral Research Fellow with the Department of Geosciences and Engineering, TU Delft. His research interests include signal processing, wavefield propagation, and migration.

Dr. Zhang has organized several international workshops and conference sessions and was a guest editor for one special journal issue on multiple reflections in *Geophysics*.



Evert Slob received the Ph.D. degree *(cum laude)* in applied sciences from the Delft University of Technology (TU Delft), Delft, The Netherlands, in 1994.

From 1995 to 2011, he was a Research Fellow, an Assistant, and an Associate Professor, and since 2011, he has been a Professor with the Department of Geosciences and Engineering, TU Delft. He is a coauthor of *Introduction to Controlled-Source Electromagnetic Methods: Detecting Subsurface Fluids* published in 2019. He is a co-editor of *Seismoelectric Exploration: Theory, Experiments, and Applications* published in 2021. His research interests include wavefield processing, imaging, and inversion.

Dr. Slob is a member of EAGE and SEG. He received the Reginald Fessenden Award from SEG in 2020. He was the General Chair of the 10th International Conference on GPR in 2004. He was the SEG Editor from 2013 to 2015. He served as the Director of Education of Civil Engineering and Geosciences from 2014 to 2018. He has organized more than a dozen international workshops and conference sessions and was a guest editor for 12 special journal issues on different topics, including GPR and hydrogeophysics.

Kees Wapenaar received the Ph.D. degree *(cum laude)* in applied physics from the Delft University of Technology, Delft, The Netherlands, in 1986.

From 1986 to 1999, he was a Postdoctoral Research Fellow and an Associate Professor with the Department of Applied Physics, Delft University of Technology. In 1999, he was appointed as the Antoni van Leeuwenhoek Professor with the Department of Geoscience and Engineering, where he has been the Head of the Geophysics and Petrophysics Research Group since 2002. His research interests include wave theory and its applications in seismic imaging and interferometric methods.

Prof. Wapenaar received the SEG's Virgil Kauffman Gold Medal in 2010, the EAGE's Conrad Schlumberger Award in 2013, and the European Research Council (ERC) Advanced Grant from the European Union in 2016.