Data-driven control over short-period internal multiples in media with a horizontally layered overburden

Journal Article

Author(s):
Elison, Patrick; Dukalski, Marcin S.; De Vos, Koos; van Manen, Dirk-Jan; Robertsson, Johan O.A.

Data-driven control over short-period internal multiples in media with a horizontally layered overburden

P. Elison [1], M.S. Dukalski,[2] K. de Vos,[3] D.J. van Manen[1] and J.O.A. Robertsson[1]

[1]*ETH Zürich, Institute of Geophysics, Zürich, Switzerland. E-mail: patrick.elison@erdw.ethz.ch*
[2]*Aramco Overseas Company B.V., Informaticalaan 6-12, 2628 ZD, Delft, The Netherlands*
[3]*Shell Global Solutions International B.V., Grasweg 31, 1031 HW Amsterdam, The Netherlands*

SUMMARY

Short-period internal multiples, resulting from closely spaced interfaces, may interfere with their generating (bandlimited) primaries, and hence they pose a long-standing challenge in their prediction and removal. A recently proposed method based on the Marchenko equation enables removal of the entire overburden-related scattering by means of calculating an inverse transmission response. However, the method relies on time windowing and can thus be inexact in the presence of short-period internal scattering. In this work, we present a detailed analysis of the impact of band-limitation on the Marchenko method. We show the influence of an incorrect first guess, and that adding multidimensional energy conservation and a minimum phase principle may be used to correctly account for both long- and short-period internal multiple scattering. The proposed method can currently only be solved for media with a laterally invariant overburden, since a multidimensional minimum phase condition is not well understood for truly 2-D and 3-D media. We demonstrate the virtue of the proposed scheme with a complex acoustic numerical model that is based on sonic log measurements in the Middle East. The results suggest not only that the conventional scheme can be robust in this setting, but that the 'augmented' Marchenko method is superior, as the latter produces a structural image identical to one where the finely layered overburden is missing. This is the first demonstration of a data-driven method to account for short-period internal multiples beyond 1-D.

Key words: Theoretical Seismology; Wave propagation; Wave Scattering and Diffraction; Inverse Theory; Image processing.

1 INTRODUCTION

Most seismic imaging algorithms of reflection data still rely on a single-scattering assumption, making multiply scattered energy often an undesired signal component. The resulting interference of shallow-borne multiples with deeper primary reflections gives rise to errors in structural and quantitative interpretation. For this reason, multiple scattering events need to be predicted and removed from the reflection data. The prediction step requires understanding of the multiple generation (i.e. which reflectors caused multiple scattering) and can be carried out by either modelling (given accurate and detailed knowledge of the shallower medium), or in a data-driven fashion (using the reflection data itself and little to no other medium information).

In practice, successful strategies have been developed for the removal of free-surface-related multiples, while this is not the case for internally generated multiples. *A priori* knowledge of the medium required for most available algorithms make it a considerable challenge. Several layer-stripping approaches have been proposed, including the work of Jakubowicz (1998), who proposed to predict internal multiples by a data correlation-convolution sequence of three primary reflections, identified and extracted from the seismic data itself. The surface-related multiple elimination (SRME) method (Verschuur *et al.* 1992) has been extended for internal multiples (Berkhout & Verschuur 1997) and relies on downward extrapolation of data to reflective subsurface interfaces, and the work of Berkhout & Verschuur (2005) relies on the identification of primaries in the data. Another range of internal-multiple attenuation approaches build up on inverse scattering theory (e.g. Ware & Aki 1969), which do not require identification of individual reflectors or events. The so-called inverse scattering series (ISS, Weglein *et al.* 1997, and further developments of it) require the source wavelet only. Similarly, methods using the Marchenko equation (Agranovich & Marchenko 1963) treat the whole overburden

as a multiple generator and calculate a so-called focusing function [i.e. an inverse transmission response, see recent 2-D and 3-D work by Broggini *et al.* (2012), Slob *et al.* (2014), Wapenaar *et al.* (2014) and van der Neut & Wapenaar (2016)]. Such approaches remove the *entire* overburden (including related multiples *and* primaries). They are therefore no demultiple methods in a classical sense (and not as such referred to in this work). These Marchenko methods are particularly advantageous over layer-stripping methods in geological settings characterized by highly scattering overburdens. In such cases the latter would not only be more costly (accounting for many generation mechanisms independently), but could also introduce an accumulation of amplitude errors cascading down, potentially giving rise to multiple mispredictions in the deeper section of the medium.

All of the above internal demultiple methods rely on some kind of temporal windowing. Such a windowing approach ultimately always introduces artefacts if the separation between the multiple generators is smaller than the inverse dominant frequency f_D^{-1} (dominant period T_D). In this case adjacent reflectors create overlapping band-limited primary reflections as well as so-called *short period* internal multiples (SPIMs). Appropriately accounting for SPIMs is critical not only for structural analysis, but also for amplitude-versus-offset (AVO) investigations and for quantitative interpretation (QI), as SPIMs can lead to perceived intrinsic attenuation, which leads to quality factor (Q) overestimation and incorrect fluid content assessment. Using demultiple methods that require time windowing is no longer possible when the two primary reflections interfere and thus cannot be separated without *a priori* information. As shown by Slob *et al.* (2014) all methods making use of the Marchenko-equation suffer from a similar problem, as the fidelity of the solutions depends on the knowledge of the solution defined on the first, wavelet-wide, interval on its support. This is true also for variations of the algorithm that rely on partial prior knowledge of the solution, such as overburden elimination introduced by van der Neut & Wapenaar (2016), Marchenko double-focusing as used by Staring *et al.* (2017, 2018) or the work by Zhang *et al.* (2018). Assuming the solution defined on this interval to be trivial in the presence of SPIMs leads to erroneous solutions, and this is the problem that we intend to solve.

In this work we show how *all* (i.e. both long- and short-period) internal multiples can successfully be handled in 2-D or 3-D media with horizontally layered overburdens. At the heart of the algorithm lies verification of energy conservation (see, e.g. Resnick *et al.* 1986) and a minimum phase condition (Sherwood & Trorey 1965; Claerbout 1968; Ware & Aki 1969; Rickett 2001) of the Marchenko solution(s)—arguments which are both symptoms of incorrect solutions as well as means to correct both their amplitude and the phase spectra, respectively. This is a generalization of the work of Dukalski *et al.* (2019). To the best of our knowledge these aspects have not been combined in a multidimensional setting elsewhere. We call this *augmented* Marchenko. It lifts the divide between long- and short-period internal multiples that is typically enforced in other algorithms. However, since the innovation of this algorithm lies in the incorporation of short-period internal multiples in the Marchenko scheme, we refer to this as the *SPIM problem*.

Our method is purely physics-based and does not add any additional assumptions or *a priori* knowledge to the 'conventional' Marchenko method. However, because the *implementation* of the algorithm presented here is limited by decoupled ray parameters in the overburden, it is (for now) only valid for horizontally layered overburdens. The structural complexity of the medium *below* the focusing depth level does not affect the effectiveness of our method.

This still makes it the first method to account for SPIM that is completely data-driven. It is suitable for geological settings with layered characteristics, such as the Middle East.

The SPIM problem is inherently linked to the band-limited nature of measured data. It is therefore indispensable to carefully analyse the representation theorems for the use of band-limited data. The attempt to solve the band-limited Marchenko equations results in a range of additional assumptions that must be considered. Only a detailed analysis of the conventional and augmented Marchenko scheme in a band-limited setting allows us to solve this problem and point out ways for general 2-D and 3-D media. Since such an analysis is crucial for this work and has not yet been presented in detail elsewhere, we dedicate Section 2 to this analysis. If readers are only interested in handling SPIMs for horizontally layered media, we refer them directly to Section 3. Throughout this paper we assume that the source signature used for deconvolution is known. We recognize that this is an outstanding practical challenge, however it is an independent problem than the one discussed in this paper and henceforth will not be further discussed.

Fig. 1 demonstrates the impact of SPIM and the improvements that result from the method described in this work. The figure illustrates injection of focusing functions into a layered medium with thin layers, which give rise to SPIM. Such an injection ideally produces a single focus at the focusing position (marked with a black cross), after which no more events should follow from the medium above. The conventional Marchenko method provides blurred functions [panel (a)] that do not properly account for SPIM. Injecting it into the medium [panels (c), (e) and (g)] hence yields a blurred focus [panel (e)], and many overburden-related internal multiples that are not properly handled remain and reach the focusing position after the actual focusing [panel (g)]. In contrast to that, the Augmented Marchenko method provides an improved (i.e. deblurred) focusing function [panel (b)], which correctly interacts with the oberburden [panel (d)] to produce a 'clean' focus [panel (f)]. We will show in this paper that improved focusing, as shown by these examples, provides similar improvements on seismic imaging. This is achieved in a completely data-driven way, and does not require detailed knowledge of the medium.

This article is structured as follows: We use Section 2 for a detailed analysis of the augmented Marchenko equations and the impact of band-limitation. We start in Section 2.2 by presenting acoustic representation theorems, an energy conservation criterion and a minimum phase argument, which are the three basic ingredients for the augmented Marchenko method. In the same section we describe why the problem is best approached inside the *overburden elimination* setting, rather than by Marchenko *redatuming* the sources and receivers below the scattering overburden; in combination with the minimum phase argument the former setting is the more natural choice. In Section 2.3, we derive the coupled Marchenko equations from the representation theorems using separability assumptions. In Section 2.4, the shortcomings of band-limited data are summarized. Ultimately, Section 2.5 is used to discuss the SPIM problem of the coupled Marchenko equations, and how it may be solved for general 2-D and 3-D media when these equations are complemented by energy conservation and a minimum phase argument (i.e. by using augmented Marchenko).

Section 3 is dedicated to the SPIM problem in media with layered overburden, in which the minimum phase criterion can be enforced in the linear Radon domain. The section is divided into two parts, which describe the correction of the amplitude spectrum (Section 3.1) and of the phase spectrum (Section 3.2) of the focusing function. The workflow is summarized in Fig. 4. Additional details
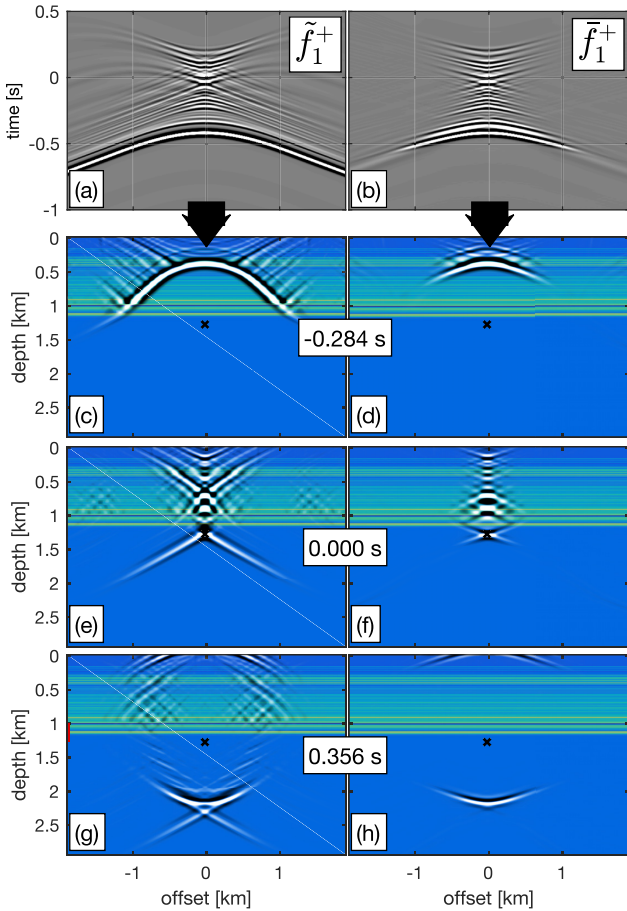
Figure 1. Comparison of blurred and deblurred (i.e. incorrect and correct) focusing, demonstrating the innovation of this work. Panels (a) and (b) show blurred and deblurred downgoing focusing functions $\tilde{f}_1^+$ and $\bar{f}_1^+$, respectively. The panels below show snapshots of simulations in which these focusing functions are injected into the reference medium [compare with Fig. 2(c) and Section 2.1], taken at time steps *before* [panels (c) and (d)], *just at* [panels (e) and (f)] and *after* focusing [panels (g) and (h)]. Note the positive upwards time axis in panels (a) and (b), illustrating how injection of the focusing functions 'pushes' them into the medium. Panels (a) and (b) are clipped to 8 per cent of the maximum amplitude, all other panels are clipped to 2 per cent of the maximum amplitude in the respective simulation.

about implementation aspects are given in Appendix A, where we investigate some important caveats, which are critical in successfully implementing this scheme in settings it is suitable for. We present results for a complex synthetic, based on a well-log from the Middle East in Section 4. The work is concluded in Section 5, where we also give a short overview of future research questions.

2 ELEMENTARY RELATIONS FOR THE AUGMENTED MARCHENKO METHOD WITH BAND-LIMITATION

In this section we present all the necessary 2-D and 3-D relations used in the augmented Marchenko demultiple algorithm (Section 2.2). We show how the coupled Marchenko equations are derived and solved (Section 2.3), we highlight the implications of band-limitation (Section 2.4), and ultimately present the effects of SPIM on the augmented Marchenko equations (Section 2.5).
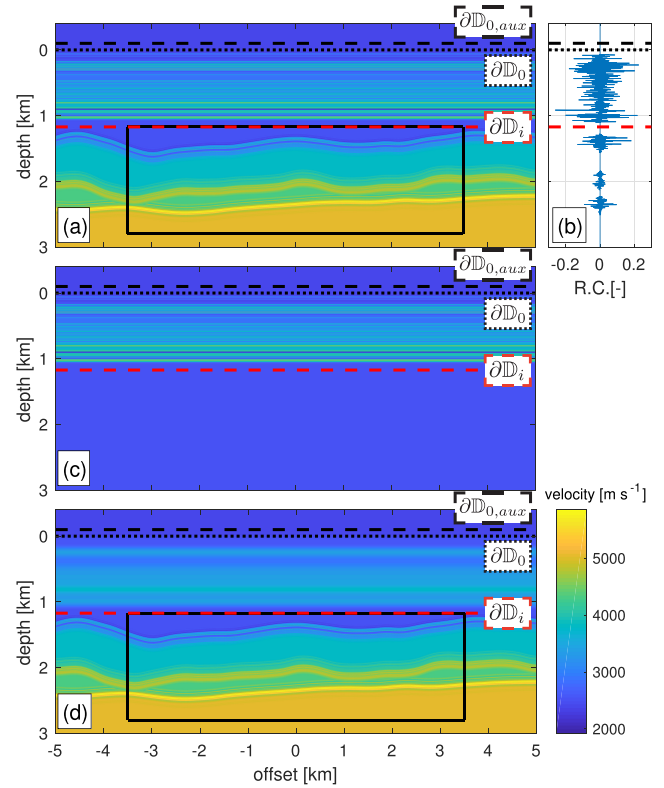

Figure 2. Synthetic 2-D acoustic velocity model inspired by the Middle East. (a) Entire model; (b) 1-D reflection coefficient taken in the model centre; (c) reference model (with reflection-free underburden); (d) target model (with smooth overburden). Surface $\partial \mathbb{D}_0$ denotes the original acquisition surface at $x_0 = (x_\perp, 0)$ (black doted line), which has been redatumed 100 m upwards to the auxiliary surface $\partial \mathbb{D}_{0,aux}$ (black dashed line, see explanation in Section 2.3). The overburden-target separating boundary $\partial \mathbb{D}_i$ is illustrated with a red dashed line.
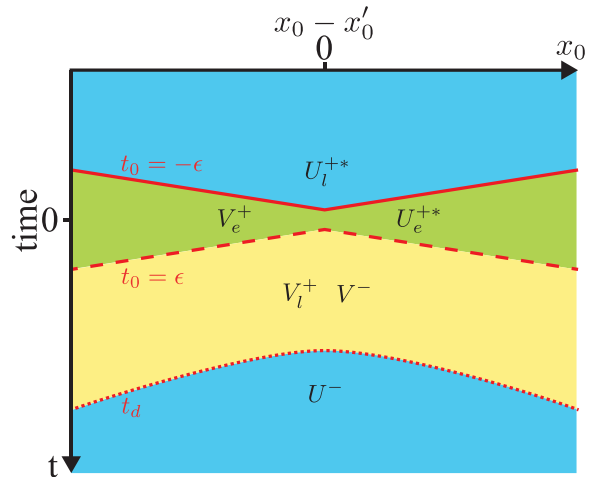
Figure 3. Schematic illustration of an exemplary spatio-temporal support of V (yellow) and U (blue) for a single shot- or receiver gather. They overlap within the green area.

2.1 Conventions and notation

We assume a dissipationless medium supporting propagation of acoustic waves only. We denote spatial coordinates of the complete medium with $x_n = (x_\perp, z_n)$ with depth z_n located on a horizontal surface $\partial \mathbb{D}_n$. In particular, we will consider two special cases with

$x_0 = (x_\perp, 0) \in \partial\mathbb{D}_0$ (acquisition surface) and $x_i = (x_\perp, z_i) \in \partial\mathbb{D}_i$ (a horizontal overburden-target separating boundary, often denoted as focusing or *redatuming* level).

The medium containing only the overburden, with a homogeneous half-space below $\partial\mathbb{D}_i$, is called a *reference* medium (or simply *overburden*). Similarly, the medium containing a smooth, contrast-free overburden and the target (i.e. the medium below $\partial\mathbb{D}_i$) will be called *target*-medium. Lastly, the medium is assumed homogeneous at and above $\partial\mathbb{D}_0$ (i.e. the free surface is absent). In practice this means that surface-related multiples have to be removed. Fig. 2 shows the 2-D velocity model that is used as synthetic example in Section 4 and illustrates the reference medium, the target medium as well as horizontal surfaces $\partial\mathbb{D}_0$ and $\partial\mathbb{D}_i$.

We use superscripts $+$ and $-$ to denote down- or up-decomposed pressure-normalized one-way wavefields, respectively (Wapenaar 1998). This normalized choice can occasionally make the notation slightly cumbersome, so to further simplify we introduce a short hand notation $\partial_{z,k} f(x) = \partial_z f(x_\perp, z)|_{z \to z_k}$, that is, the value of a vertical derivative of $f(x)$ evaluated at $z = z_k$. For functions of more than one variable, we will use superscripts [1] ([2]) to denote if the derivative is taken over the first (second) spatial coordinate. Further notational simplifications will be made known in the course of the article. In order to implement our scheme, we will need to work across a number of domains. To keep the argument clear and consistent we summarize all the notation in Table 1. The sign convention of the Fourier transforms is used as defined in Wapenaar & Berkhout (1989,p. 67), and similar to these authors we consider only positive frequencies, that is $\omega \geq 0$.

We use the definition of the Green's function (e.g. see Wapenaar *et al.* 2014) due to a monopole volume injection rate source at x' and a receiver at x

$$\left[\rho(x) \nabla \frac{1}{\rho(x)} \nabla + \frac{\omega^2}{c(x)^2} \right] G(x, x', \omega)$$
$$= -i\omega\rho(x)\delta(x - x') . \tag{1}$$

The (complete) medium is characterized by an appropriately scaled reflection (pressure) wavefield $R^\cup(x_0'', x_0, \omega) = [\frac{1}{2}i\omega\rho(x_0)]^{-1}\partial_{z,0}^{[1]} G^-(x, x_0'', \omega)$, recorded using a (regularly spaced) grid of colocated sources and receivers at $\partial\mathbb{D}_0$. Source or receiver signatures are not considered for this quantity, however this will become important later in this work.

We speak of *broad-band* wavefields if they contain frequencies from zero to infinity (or zero to f_{Nyquist} for discrete signals), and of *band-limited* wavefields otherwise (we will always make the presence of the wavelet explicit). This paper focuses on temporal band-limitation aspects and the spatial sampling is always assumed to be sufficiently dense to avoid spatio-temporal aliasing when computing Rayleigh integrals.

2.2 Acoustic representation theorems

Marchenko equation-based methods [as derived by Slob *et al.* (2014) and Wapenaar *et al.* (2014)] rely on a focusing wavefield $F_1(x_i', x_i, \omega)$, defined in the reference medium [see illustration in Fig. 2(c)], whose down- and up-propagating components obey

$$\partial_{z,i}^{[2]} F_1^+(x_0', x, \omega) = -\frac{1}{2}i\omega\rho(x_i')\delta(x_\perp - x_\perp') , \tag{2}$$
$$F_1^-(x', x_i, \omega) = 0 \quad \forall z > z_i.$$

The interpretation of these equations is that F_1 collapses to a downward radiating source at x_i and has no upward propagating component below [compare with Fig. 1]. The focusing field can be used in the reciprocity theorems of correlation and convolution type (Wapenaar & Grimbergen 1996) to derive the so-called representation theorems (Slob *et al.* 2014; Wapenaar *et al.* 2014) relating the focusing functions $F_1^\pm$ and the Green's functions $G^\pm$ via the reflection response $R^\cup$

$$G^-(x_i', x_0'', \omega) + F_1^-(x_0'', x_i', \omega)$$
$$= \int dx_0 R^\cup(x_0'', x_0, \omega) F_1^+(x_0, x_i', \omega) , \tag{3}$$

$$-G^{+*}(x_i', x_0'', \omega) + F_1^+(x_0'', x_i', \omega)$$
$$= \int dx_0 R^{\cup*}(x_0'', x_0, \omega) F_1^-(x_0, x_i', \omega) . \tag{4}$$

The first index denotes the measurement location, the second index is the the focal point location and the superscript * denotes complex conjugation.

Assuming knowledge of the reflection response $R^\cup$, the goal of Marchenko redatuming techniques is to reduce this set of two equations with four unknowns into a system where solutions $F_1^\pm$ can be found. This is done by applying appropriate temporal mutes in the time domain which remove (most of) $G^\pm$ and leave the overlaps between G^{+*} and F_1^+, or G^- and F_1^-. By making some assumptions about these overlaps (see Wapenaar *et al.* (2014) and the discussion in Section 2.3), solutions $F_1^\pm$ are found, eqs (3–4) are used to determine $G^\pm$ and these are used to invert an Amundsen (2001)-type relation

$$G^-(x_i, x_0'', \omega)$$
$$= \int dx_i' R^\cup_{\text{Target}}(x_i, x_i', \omega) G^+(x_i', x_0'', \omega) , \tag{5}$$

also known as multidimensional deconvolution (MDD). $R^\cup_{\text{Target}}(x_i, x_i', \omega)$ denotes a Green's function due to a dipole source (at x_i') and monopole receiver (at x_i) measured in the target medium and hence called *target data*.[1] The process of going from $R^\cup$ to $R^\cup_{\text{Target}}$ can be interpreted as a *form* of demultiple, as the latter is void of any multiple scattering due to the overburden. It is not a demultiple process in a classical sense though, as it removes *all* events due to the overburden, including primary reflections. As will become clear later in this article, the success of this approach is contingent on knowing the overlaps of G^{+*}/F_1^+ and G^-/F_1^-; in particular the former, which is challenging in the presence of closely spaced reflectors. This paper, and especially Section 3, focuses on obtaining these overlaps correctly in a data-driven way.

To further simplify notation and bring it more in tune with a numerical implementation, we represent each of the fields in eqs (3–4) by an $n_0 \times n_i \times n_f$ array (and $n_0 \times n_0 \times n_f$ for $R^\cup$), where n_0 and n_i represent the number of sources and receivers. Time translational invariance (Noether 1918) of the representation theorems states that they hold independently for every frequency component (subscript f). We therefore represent eqs (3–4) by matrix multiplications for every frequency slice (hence we suppress ω in further notation and refer to Table 1 for representations in different domains) and we get

$$G^- + F_1^- = R^\cup \ F_1^+ dx_0 , \tag{6}$$

[1]This data set has previously been referred to as a redatumed response (Wapenaar *et al.* 2014) or one measured in an overburden-subtracted medium (van der Neut & Wapenaar 2016). Here we prefer to take a more unifying approach, by referring to this data set as target data.
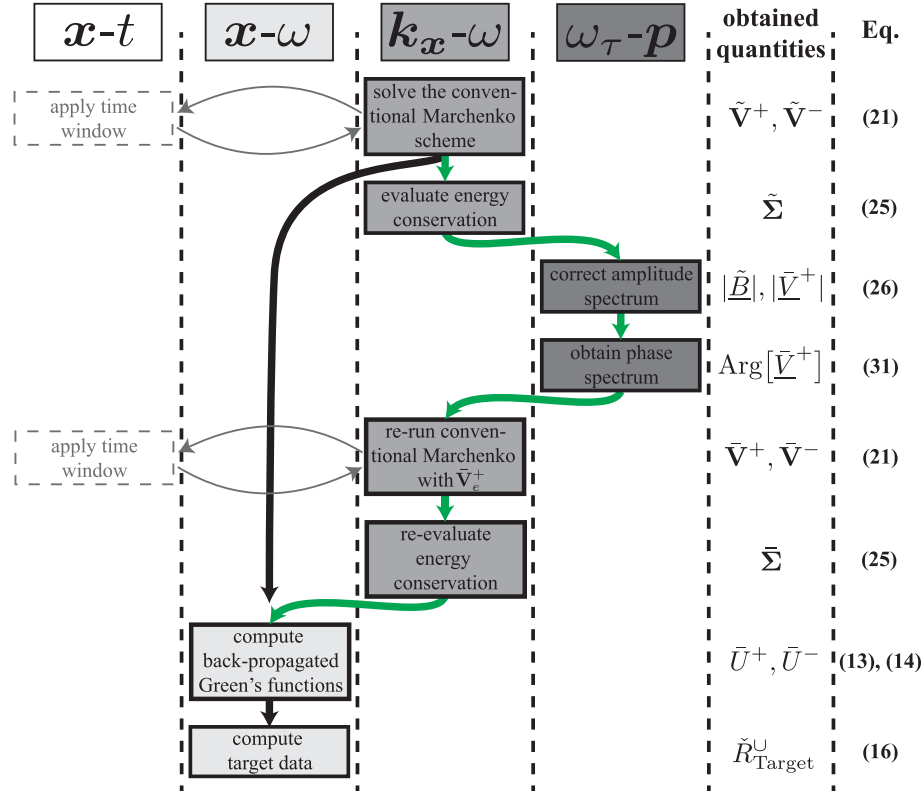
Figure 4. Workflow of the augmented Marchenko method including SPIM for media with a laterally invariant overburden as described in this section. Black arrows illustrate the workflow of the conventional Marchenko method, green arrows illustrate additional steps of the augmented scheme. Important additional aspects regarding the implementation are summarized in Appendix A.

$$- G^{+*} + F_1^+ = R^{\cup *} F_1^- \, d\mathbf{x}_0 \,. \tag{7}$$

We leave the integration measure explicit next to every matrix product as it is not only an essential scaling but also acts as reminder over which surface the integration is being taken. The form of eqs (6–7) also notably shows the freedom to multiply from the left or the right by some compatible (frequency dependent) matrix, which is what had been used previously by Singh *et al.* (2015) and Ravasi (2017) from the left, and van der Neut & Wapenaar (2016) from the right.

In the reference medium the reciprocity theorem of correlation type can be used to obtain

$$\left[F_1^{+\dagger} \partial_{z,0}^{[1]} F_1^+ + F_1^{-\dagger} \partial_{z,0}^{[1]} F_1^- \right] \rho_0^{-1} d\mathbf{x}_0 = -\frac{i\omega^2}{4} \rho_i I_{MM} \,, \tag{8}$$

with the density ρ_0 and ρ_i at depth levels z_0 and z_i, respectively, and where superscript $\dagger$ denotes the conjugate transpose and

$$I_{MM}(\mathbf{x}_i', \mathbf{x}_i) = I_{MM}(\mathbf{x}_i' - \mathbf{x}_i)$$
$$= \frac{1}{(2\pi)^N} \int d\mathbf{k}_x \, e^{i\mathbf{k}_x \cdot (\mathbf{x}_{\perp,i}' - \mathbf{x}_{\perp,i})} k_{z,i}^{-1} \,, \tag{9}$$

where $N = \dim[\mathbf{k}_x]$, $k_{z,i}^2 = \omega^2 c(\mathbf{x}_i)^{-2} - |\mathbf{k}_x|^2$ and $\mathbf{k}_x$ are the wavenumber components in the horizontal plane (x, y) perpendicular to the z-direction. We have used subscripts MM to denote the monopole-receiver and monopole-source character, respectively (and we will use subscript D to denote a dipole source/receiver-character). Expressing the definition of I_{MM} in eq. (9) in the wavenumber–frequency domain allows us to present it in a form that is valid in any number of dimensions. We have assumed

$\partial_{\mathbf{x}_{\perp,0}} \rho = \partial_{\mathbf{x}_{\perp,i}} \rho = 0$ because $\partial \mathbb{D}_0$ and $\partial \mathbb{D}_i$ should be scattering-free. On the right-hand side of eq. (8) we used the defining property of the focusing function in eq. (2). Eq. (8) states that net in- and outgoing energy of the focusing fields along $\partial \mathbb{D}_0$ is equal to an appropriately scaled 'identity', one whose *vertical derivative* $\partial_{z,i}^{[1]} I_{MM}(\mathbf{x}_i', \mathbf{x}_i) \equiv I_{DM}(\mathbf{x}_i', \mathbf{x}_i)$ is a zero time and a zero-displacement acoustic propagator. In other words, this is a homogeneous medium impulse response due to a monopole source at and dipole receivers along $\partial \mathbb{D}_i$. This is an extension of the 1-D energy conservation statement that can be found in, for example, Resnick *et al.* (1986). Note that the densities $\rho(\mathbf{x}_0)$ and $\rho(\mathbf{x}_i)$ assure that the focusing functions focus to a unit strength downward radiating source, consistent with eq. (1). Mildner *et al.* (2019) have shown how eq. (8) can be used to obtain two-way direct transmission amplitude information.

Lastly, in the reference medium we have that $G^-(\mathbf{x}_i, \mathbf{x}_0') = 0$ and $G^+(\mathbf{x}_i, \mathbf{x}_0') = T^+(\mathbf{x}_i, \mathbf{x}_0')$ (a transmission response between a monopole source at $\mathbf{x}_0'$ and a monopole receiver at $\mathbf{x}_i$) and hence reciprocity theorem of convolution type yields

$$\left[\partial_{z,0}^{[2]} T^+ \right]^T F_1^+ \rho_0^{-1} d\mathbf{x}_0 = \frac{\omega^2}{4} \rho_i^{-1} I_{MM} \,, \tag{10}$$

meaning that focusing and transmission fields are each others *inverses* in the sense of eq. (10), where the additional pre-factors and derivatives are a consequence of the up-/downpressure normalization choice. Note that due to our definition of the Green's function [eq. (1)] the transmission $T^+(\mathbf{x}_i, \mathbf{x}_0')$ in eq. (10) could also be replaced by the equivalent $T^-(\mathbf{x}_0', \mathbf{x}_i)$.

In 1-D, eq. (10) has a well known interpretation: when measured from the onset of the signal, transmissions (T^+) and their inverses

Table 1. Notation convention for different wavefields used in this work.

Notation	Meaning	Example
Lowercase	x-t-domain	f_1^-, g^+
Uppercase	x-ω-domain	F_1^-, G^+, V^+, U^-, W
Bold uppercase	k_x-ω-domain	$\boldsymbol{F}_1^+, \boldsymbol{V}^-, \boldsymbol{B}$
Lowercase, underline	τ-$\boldsymbol{p}$-domain	$\underline{f}_{-1}^+, \underline{v}^-$
Uppercase, underline	ω_τ-$\boldsymbol{p}$-domain	$\underline{F}_{-1}^+, \underline{V}^-, \underline{B}$
Lowercase bold	Vectors	$\boldsymbol{x}_0, \boldsymbol{x}_i', \boldsymbol{k}_\perp, \boldsymbol{p}$
Overbar	Wavelet-dressed	$\overline{f}_1^+, \overline{V}^-, \overline{F}_{-1}^+, \overline{V}^+$
Check	Band-limited	$\check{R}^\cup$
Tilde	Blurred with B	$\tilde{f}_1^+, \tilde{U}^+, \underline{\tilde{f}}_1^+, \underline{\tilde{V}}^-$

$\left(F_1^+\right)$ form a minimum phase couple (Sherwood & Trorey 1965; Claerbout 1968; Ware & Aki 1969). This means that they are always stable and causal 1-D responses[2] with stable and causal inverses. As a consequence of that, the amplitude spectrum dictates the phase spectrum (and vice versa) through the Kolmogorov relation (Claerbout 1976; Skingle *et al.* 1977). Eq. (10) is a manifestation of the same fact, but now in a multidimensional sense. Since the minimum phase condition is unaffected by scalar multiplication, the densities in eq. (10) can be ignored and the factor of ω^2 can be absorbed by redefining the Green's function source term and hence too can be ignored.

A possible way to generalize minimum-phase properties to a multidimensional setting is to propagate the entire transmission- and focusing functions F_1^+ and T^+ by the direct transmission event T_d^+ or the direct inverse transmission T_d^{+inv}, respectively.[3] For a justification of this argument further remarks on minimum phase objects beyond 1-D are summarized in Appendix B. Consequently, the propagated *minimum phase* transmission $\mathcal{T}^+$ and inverse transmission V^+ fields are defined as

$$V^+(\boldsymbol{x}_0', \boldsymbol{x}_0) = \int d\boldsymbol{x}_i \, F_1^+(\boldsymbol{x}_0', \boldsymbol{x}_i) T_d^+(\boldsymbol{x}_i, \boldsymbol{x}_0), \quad (11)$$

$$\mathcal{T}^+(\boldsymbol{x}_0', \boldsymbol{x}_0) = \int d\boldsymbol{x}_i \, T^+(\boldsymbol{x}_0', \boldsymbol{x}_i) T_d^{+inv}(\boldsymbol{x}_i, \boldsymbol{x}_0), \quad (12)$$

where T_d^+ and T_d^{+inv} are appropriately-scaled dipole-source *dipole-receiver* propagators (though we omit subscripts DD to make for a more compact notation). The direct arrival events of V^+ and $\mathcal{T}^+$ (potentially including forward scattering, but no internal multiples) collapses onto I_{DM} (or its derivative depending on the Green's function definition), evaluated at $\partial\mathbb{D}_0$. Definition of V^+ is akin to the one found in van der Neut & Wapenaar (2016), except that we are using pressure-normalized wavefield separation, and we are suggesting to account for forward scattering and 2-D effects such as triplications etc. We also prefer to use the term 'propagated' instead of 'back-projected' to imply a temporal shift and to avoid alluding to dimensionality reduction.

Seeing how V^+ (rather than F_1^+) appears to be the more natural candidate to investigate, we define the propagated equivalents of $F_1^\pm$ and $G^\pm$ as

$$V^- = F_1^- T_d^+ d\boldsymbol{x}_i, \quad (13)$$
$$U^{+*} = G^{+*} T_d^+ d\boldsymbol{x}_i,$$
$$U^- = G^- T_d^+ d\boldsymbol{x}_i,$$

[2]Impulse responses are always full bandwidth, while signals (measured responses) can be band-limited.

[3]In many not-too-complex cases this should be the inverse of the first event in either of the response.

and multiply both sides of eqs (6–7) from the right by $T_d^+ d\boldsymbol{x}_i$. This results in

$$U^- + V^- = R^\cup \, V^+ d\boldsymbol{x}_0, \quad (14)$$

$$-U^{+*} + V^+ = R^{\cup*} V^- d\boldsymbol{x}_0, \quad (15)$$

which are the propagated equivalents of eqs (6–7). Note that U^+ is propagated in the same direction in time as $\mathcal{T}^+$ [see eq. (12)], but in opposite direction as U^-.

Eqs(14–15) can be further multiplied by a frequency-dependent scalar (i.e. a band-limited wavelet) such that wavefields $V^\pm$ and $U^\pm$ can be replaced with their band-limited equivalents $\overline{V}^\pm$ and $\overline{U}^\pm$. Moreover, we wish to point out that eqs (14–15) *still* retain the freedom of being multiplied by an arbitrary frequency dependent matrix from either side—a property that we will make use of in Section 2.5.

Using the same freedom, we multiply the energy conservation relation [eq. (8)] from the left with $T_d^{+\dagger} d\boldsymbol{x}_i'$ and from the right with $T_d^+ d\boldsymbol{x}_i$ and get a modified energy conservation relation

$$\left[V^{+\dagger}\partial_{z,0}^{[1]} V^+ + V^{-\dagger}\partial_{z,0}^{[1]} V^- \right] \rho_0^{-1} d\boldsymbol{x}_0 \quad (16)$$
$$= -\frac{i\omega^2}{4} \rho_i T_d^{+\dagger} I_{MM} T_d^+ d\boldsymbol{x}_i d\boldsymbol{x}_i'.$$

In this propagated formalism we can still obtain $R_{\text{Target}}^\cup$, however now with sources and receivers at $\partial\mathbb{D}_0$

$$U^-(\boldsymbol{x}_0, \boldsymbol{x}_0'', \omega)$$
$$= \int d\boldsymbol{x}_0' \, R_{\text{Target}}^\cup\left(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega\right) U^+(\boldsymbol{x}_0', \boldsymbol{x}_0'', \omega) \quad (17)$$

such that $R_{\text{Target}}^\cup\left(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega\right)$ and $R_{\text{Target}}^\cup\left(\boldsymbol{x}_i, \boldsymbol{x}_i', \omega\right)$ [from eq. (5)] contain the same set of events, however, with different timing, moveouts and amplitudes. Relating this to the example model used in this work, the former is recorded at $\partial\mathbb{D}_0$ of Fig. 2(d), whereas the the latter is recorded at $\partial\mathbb{D}_i$ in a model that contains only the target, but is homogeneous above (not shown). The former is a particularly convenient data set for quality control of the Marchenko method, as the target primaries are identical to those in the input $R^\cup$, but with a transmission correction, compare Figs 5(a) and (b). Appendix C elaborates on that transmission correction using a simple 1-D example. Using wavelet-dressed wavefields $\overline{U}^\pm$ instead of their broadband equivalents $U^\pm$ in eq. (17) allows to recover $\check{R}_{\text{Target}}^\cup$, a band-limited version of the target medium reflection response $R_{\text{Target}}^\cup$.

2.3 Deriving and solving the coupled set of Marchenko equations

To extract the propagated focusing functions $V^\pm$ we introduce separability conditions. They aim at muting (parts of) $U^\pm$ from the representation theorems given in eqs (14–15) and therefore reduce the number of unknowns. This is achieved by means of a projector Φ (i.e. a window operator). Due to an overlap in $V^\pm$ and $U^\pm$, the equations must be physically constrained by providing that overlap on input. A schematic illustration of an exemplary spatio-temporal support of $V^\pm$ and $U^\pm$ in Fig. 3 is used to depict the separability assumptions.

We define a projector Φ which by design preserves V^- according to $\Phi_{t_0}\left[V^-\right] = V^-$ and is given by

$$\Phi_{t_0} = \begin{cases} 1, & \text{if } t_0(\boldsymbol{x}_0', \boldsymbol{x}_0) < t < t_d(\boldsymbol{x}_0', \boldsymbol{x}_0) \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$
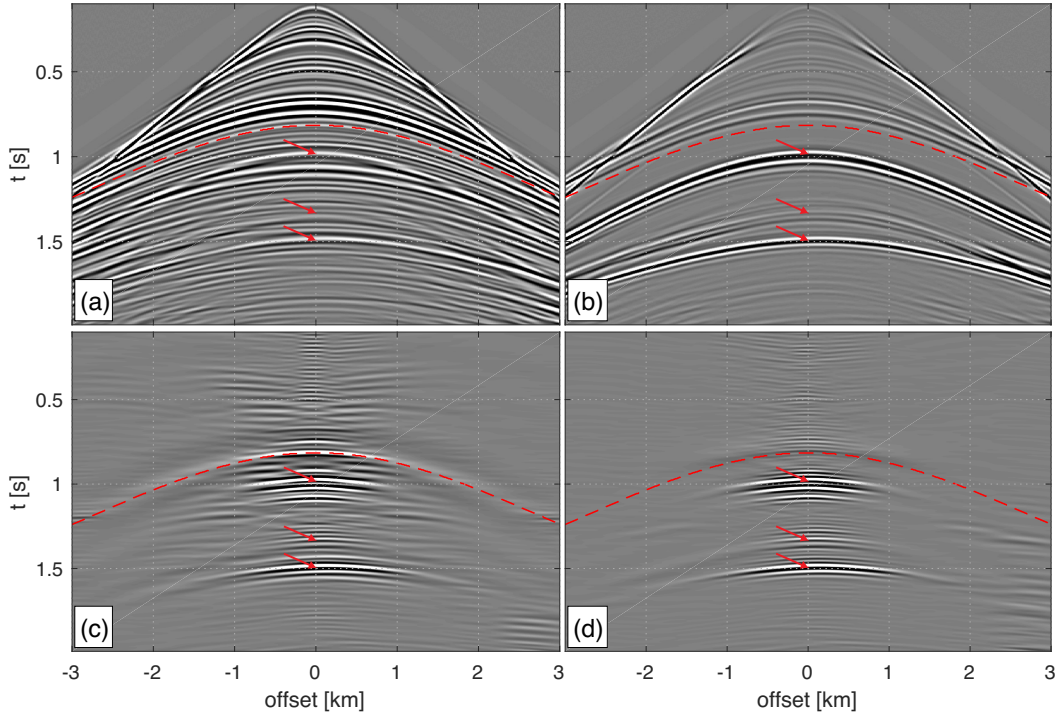
Figure 5. Central shot of modelled and Marchenko-derived reflection responses, all dressed with a 20 Hz Ricker wavelet. (a) modelled data of the entire model; (b) modelled target data; (c) target data obtained from blurred Marchenko results; (d) target data obtained from deblurred Marchenko results. Reflection data $\bar{r}^{\cup}(\boldsymbol{x}_0'', \boldsymbol{x}_0)$ in panel (a) corresponds to the model in Fig. 2(a), $\bar{r}_{\text{Target}}^{\cup}(\boldsymbol{x}_0'', \boldsymbol{x}_0)$ in panels (b) and (d) and $\tilde{r}_{\text{Target}}^{\cup}(\boldsymbol{x}_0'', \boldsymbol{x}_0)$ in panel (c) correspond to Fig. 2(d). The Marchenko-derived data [panels (c) and (d)] have been dip-filtered in k_x-ω-domain to reduce linear artefacts. Panels (a) and (b) are displayed with a t^2-gain, all panels are clipped to 15 per cent of the maximum amplitude.

The bounds of the projector are marked with red lines in Fig. 3. The late bound $t_d(\boldsymbol{x}_0'', \boldsymbol{x}_0')$ (dotted line) is the travel-time of the direct arrival between sources and receivers at $\partial\mathbb{D}_0$ via a focusing position at $\partial\mathbb{D}_i$ (similar to a two-way travel-time). The early bound t_0 is defined by a (band-limited) spatio-temporal delta function. It is illustrated at negative and at positive times in Fig. 3 (solid and dashed line), because its definition is dependent on a choice that is further explained in the following.

We introduce a subdivision of $U^+ = U_e^+ + U_l^+$ and $V^+ = V_e^+ + V_l^+$ into their 'early' and 'late' parts (denoted with subscripts e and l), respectively. By construction, the two early parts U_e^{+*} and V_e^+ always temporally overlap over some offset dependent interval $|t| \leq \varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0')$, illustrated by the green area in Fig. 3. For the solution of eqs (14–15) this overlap needs to be provided as input and is naturally dependent on the bandwidth of these fields. It can be treated in two ways: either by providing U_e^{+*} on input and using $t_0 = -\varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0')$ for the projector Φ_{t_0} (i.e. the solid red line in Fig. 3), or by setting $t_0 = \varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0')$ (dashed line) and providing V_e^+ instead. Here we chose the latter, and defer the discussion on the former to Appendix D.

Another separability assumption requires that V^- is separated from U^- through t_d (see the dotted red line in Fig. 3). This assumption might break down in the presence of diving waves in high velocity gradients, but also if the level $\partial\mathbb{D}_i$ is sandwiched between two closely spaced reflectors such that their (band-limited) primary reflections overlap. We will assume that any diving waves or steeply dipping events have been removed, and that the target is separated from the overburden by a sufficiently thick layer in which $\partial\mathbb{D}_i$ can be placed. Consequently, in the following V^- and U^- are assumed to be well separated.

We apply Φ_{t_0} (given by the yellow area in Fig. 3) to both sides of eqs (14) and (15)

$$X^- + V^- = \Phi_{t_0}\left[R^{\cup} \ V^+ d\boldsymbol{x}_0\right], \tag{19}$$

$$\Phi_{t_0}\left[-U^{+*} + V^+\right] = \Phi_{t_0}\left[R^{\cup*}V^- d\boldsymbol{x}_0\right], \tag{20}$$

where $X^- \equiv X^-(\boldsymbol{x}_0, \boldsymbol{x}_0') = \Phi_{t_0}\left[U^-(\boldsymbol{x}_0, \boldsymbol{x}_0')\right]$. Due to the second separability assumption introduced above $X^- = 0$, and our choice for the first separability assumption reduces the left-hand side of eq. (20) to V_l^+. Combining the two resulting equations with $V^+ = V_e^+ + V_l^+$ yields

$$V^+ - \Phi_\varepsilon\left[R^{\cup*}\Phi_\varepsilon\left[R^{\cup}V^+ d\boldsymbol{x}_0\right]d\boldsymbol{x}_0\right] = V_e^+, \tag{21}$$

which can be solved in several ways (Slob *et al.* 2014; van der Neut & Wapenaar 2016; Dukalski & de Vos 2017). For most purposes where free-surface multiples are absent (as assumed in this work) a Neumann series expansion

$$V^+ = \sum_{k=0}^{\infty} \Omega_k\left[V_e^+\right] \quad \text{and} \quad V^- = \Phi_\varepsilon\left[R^{\cup}V^+ d\boldsymbol{x}_0\right],$$

$$\text{with } \Omega_k[A] = \Phi_\varepsilon\left[R^{\cup*}\Phi_\varepsilon\left[R^{\cup}\Omega_{k-1}[A]d\boldsymbol{x}_0\right]d\boldsymbol{x}_0'\right]$$
$$\text{and } \Omega_0[A] = A \tag{22}$$

yields the quickest convergence (Dukalski & de Vos 2017). For a broadband reflection response $R^{\cup}$ with infinitesimal shot and receiver sampling ($d\boldsymbol{x}_0 \to 0$), $\varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0') \to 0$ and the initial term V_e^+ in the series above is given by I_{DM}. Note that, in practice (i.e. in a band-limited setting and in the presence of SPIM), V_e^+ is often unknown. It is the purpose of this work to show how to correct for an erroneous estimate of V_e^+, and thus obtain the true solution.

Note that Φ_{t_0} is defined in the time domain, therefore its action on frequency domain arrays is implicitly understood as a sequence of an inverse Fourier transform, temporal mute application (with appropriately chosen tapers) and a forward Fourier transform. Note also that t_0 is the time smaller than the onset of signal in V^-, which should be the same as the onset of signal in $R^\cup$. In the case of a very shallow first reflector (e.g. a water bottom), where $t_0(\boldsymbol{x}_0'', \boldsymbol{x}_0')$ is closer to $t = 0$ than the width of a wavelet, one should first sufficiently extrapolate the source and receiver locations in $R^\cup$ upwards, using a homogeneous medium Green's function. For short distances a simple static shift applied to each trace would suffice. This has been done for the model illustrated in Fig. 2, for which the original acquisition surface $\partial\mathbb{D}_0$ has been redatumed upwards by 100 m to obtain the auxiliary surface $\partial\mathbb{D}_{0,aux}$.

2.4 Shortcomings introduced by band-limited data

In practice the reflection response $R^\cup$ is band-limited in the wavenumber and frequency domain, here denoted with $\check{R}^\cup$. It enables calculation of the band-limited fields $\overline{V}^\pm$, $\overline{U}^\pm$ and $\check{R}^\cup_{\text{Target}}$, with spatio-temporal bandwidths that do not exceed those of $\check{R}^\cup$. As a result of band-limitation, the initial term $\overline{V}_e^+$ also has a temporal width and $\varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0') \neq 0$. One must therefore use a band-limited expression to constrain the solution of the representation theorems: instead of using I_{DM} as initial term, one must use a band-limited version $\overline{I}_{DM}$, which is dressed with a $2\varepsilon(\boldsymbol{x}_0, \boldsymbol{x}_0)$-wide zero-phase wavelet. This could be a good choice under the assumption that the coda wave information is present in $\overline{V}_l^+$ *only*, and *not* in $\overline{V}_e^+$. This assumption is no longer true in the presence of SPIMs, as the shortest multiple reverberation is less than ε and enters the overlap between $\overline{V}_e^+$ and $\overline{U}^{+*}$. Choosing $\overline{V}_e^+ = \overline{I}_{DM}$ therefore leads to erroneous solutions. This problem is further discussed in the next section.

Band limitation introduces another problem that is independent of the SPIM problem identified above. We have observed that for sufficiently strongly scattering media eq. (22) requires many iterations, which leads to erroneous solutions. We further explain this challenge in Appendix E.

2.5 Failure to estimate V_e^+: consequences and remedies

In order to remedy errors that originate from the incorrect choice of the initial term $\overline{V}_e^+$, we first need to understand what impact it has on the solutions to eq. (21) and secondly find a way to reverse it. To achieve that we use the right-multiplicative freedom of the representation theorems in eqs (3–4). We exploited this freedom already to transform these equations to their modified versions in eqs (14–15) using a *known* operator $T_d(\boldsymbol{x}_i', \boldsymbol{x}_0)$. We now repeat the process on the latter set of equations using an *a priori unknown* operator $B(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$. Due to linearity of eqs (14–15), this transformation replaces variables V^+, V^-, U^- and U^{*+} with their so-called *blurred* equivalents according to the following rule

$$
\begin{aligned}
\tilde{V}^+ &= V^+ B \mathrm{d}\boldsymbol{x}_0\,, \\
\tilde{V}^- &= V^- B \mathrm{d}\boldsymbol{x}_0\,, \\
\tilde{U}^- &= U^+ B \mathrm{d}\boldsymbol{x}_0\,, \\
\tilde{U}^{+*} &= U^{+*} B \mathrm{d}\boldsymbol{x}_0\,.
\end{aligned} \tag{23}
$$

The $V^\pm \to \tilde{V}^\pm$ transformation also does not affect eq. (21) under the assumption that $\Phi_\varepsilon[\tilde{U}^-] = 0$ and $\Phi_\varepsilon[\tilde{V}^-] = \tilde{V}^-$ still holds. This is the case given that B is a location dependent filter with a temporal support identical to the one of the wavelet used to dress $\overline{I}_{DM}$. Solutions are still given by eq. (22) and we again obtain $\tilde{V}^+$ given some initial term $\tilde{V}_e^+$.

This shows that the incorrect *choice* of the initial term $\tilde{V}_e^+ = \overline{I}_{DM}$ can be interpreted as a specific choice of an unknown B. Hence, using $\overline{I}_{DM}$ as initial choice is erroneous, in which case we actually calculate $\tilde{V}^\pm$ instead of $\overline{V}^\pm$. The latter is the desired set of solutions, whereas the former is that solution multidimensionally convolved with some unknown filter $B(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$. Henceforth, this filter is referred to as a *blurring operator*. It is critical to realize that *both* solutions ($\overline{V}^+$ and $\overline{V}^-$) are modified by *the same* operator B, leading not only to an incorrect initial part $\overline{V}_e^+$, but to incorrect amplitude and phase spectra of the entire $\overline{V}$.

To reconstruct fields $\overline{V}^\pm$ from $\tilde{V}^\pm$ we use a data-driven procedure that correctly recovers $\overline{V}^+\overline{V}^{+\dagger}$, that is, the multidimensional auto-correlation of $\overline{V}^+$ (a multidimensional equivalent to a 1-D power spectrum). We then attempt to fix the phase using a minimum phase argument mentioned in Section 2.2.

First, we observe that the energy conservation relation for fields $\tilde{V}^\pm$ is given by

$$
\begin{aligned}
&\left[\tilde{V}^{+\dagger} \partial_{z,0}^{[1]} \tilde{V}^+ + \tilde{V}^{-\dagger} \partial_{z,0}^{[1]} \tilde{V}^- \right] \rho_0^{-1} \mathrm{d}\boldsymbol{x}_0 \equiv \Sigma\,, \quad \text{where} \\
&\Sigma = -\frac{i\omega^2}{4} \rho_i B^\dagger T_d^{+\dagger} I_{MM} T_d^+ B \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{x}_i' \mathrm{d}\boldsymbol{x}_0'' \mathrm{d}\boldsymbol{x}_0'''\,,
\end{aligned} \tag{24}
$$

where the right-hand side is a consequence of eq. (16) and the definition of $\tilde{V}_e^+$ in eq. (23) was used. Here we assumed that $\partial_{z,0}B = 0$ and $\partial_{z,i}B = 0$, which means that B (and hence also the SPIM contribution) does not change over infinitesimal displacements $z_0 \to z_0 + \mathrm{d}z$ and $z_i \to z_i + \mathrm{d}z$, that is, z_0 and z_i are sufficiently (at least half a wavelet width) far away from the nearest reflector. This is the same condition which guarantees $\tilde{X}^- = 0$. Note that since V_e^+ is a function of the location of $\partial\mathbb{D}_i$ (i.e. our choice of overburden target-separation), so is B. This becomes clear in the expression that relates these two quantities $(1 - \Phi_\epsilon)V B \mathrm{d}\boldsymbol{x}_0 = \tilde{V}_e^+$, since regardless of the level choice we always use an identity for $\tilde{V}_e^+$.

In order to recover $\overline{V}^+$ from $\tilde{V}^+$, we can follow a line of research proposed by Reinicke *et al.* (2019). We take the right-hand side of eq. (24), (pseudo-)invert it and multiply with $\tilde{V}^+$ and $\tilde{V}^{+\dagger}$ from the left and right, respectively. This yields the normal product $\overline{V}_1^+ \overline{V}_1^{+\dagger}$, up to factor given by frequency vector, local densities and vertical wavenumbers measured at $\partial\mathbb{D}_0$ and $\partial\mathbb{D}_i$. In order to recover the multidimensional phase, we need to be able to factorize this normal product, such that $\overline{V}^+$ is a wavelet-dressed minimum phase matrix. Performing such factorization uniquely and on band-limited matrices in general (arbitrary overburdens) is a huge outstanding theoretical challenge (Horn & Johnson 1990) and we hope to return to this question in the future. In this work we will present a special case of laterally invariant overburdens where such factorization is straightforward due to an additional conservation law.

3 ACCOUNTING FOR SHORT-PERIOD INTERNAL MULTIPLES OF A HORIZONTALLY LAYERED OVERBURDEN

In this section, we show how the augmented Marchenko equations can be used to alleviate the SPIM problem if the overburden does not vary laterally. To give an overview of the augmented Marchenko scheme, and to guide the reader through the equations in this section, we illustrate a workflow in Fig. 4. Different columns indicate the different domains in which individual steps are performed (compare

with Table 1). The thick black arrow illustrates the workflow of the conventional Marchenko method, and the green arrows illustrate the additional steps introduced in this section. In addition to the theory outlined in this section, we give insights into the implementation details in Appendix A.

For laterally invariant overburdens we can exploit symmetry properties in either the wavenumber-frequency domain or the linear Radon domain [also denoted as $\tau - p$ domain] (Deans 1983; Foster & Mosher 1992). In particular the latter is very useful, as Snell's law conserves *frequency-independent* ray parameters $p(z)$ over a sequence of horizontal interfaces, that is,

$$p(z) = \frac{\sin(\phi(z))}{c(z)} = \text{constant}. \tag{25}$$

Here $\phi(z)$ and $c(z)$ are the propagation angle and the wave propagation speed at depth z. This equation states that in a laterally invariant overburden focusing functions can be analysed independently per ray parameter, such that a 2-D or a 3-D field $V^\pm$ decouples into a set of 1-D fields $\underline{v}^\pm$ (note the domain-specific notation).

3.1 Amplitude spectrum correction

Focusing functions $V^\pm$ of laterally invariant overburdens only depend on $x_{\perp,0} - x_{\perp,i}$, a relative horizontal distance between the focusing location and a receiver.[4] Spatial integrals such as those in eq. (8) become convolutions, which can be evaluated by Hadamard (piece-wise) products in k_x-ω domain, and we express wavefields in this domain with capital, bold letters, for example, $\tilde{V}^+(k_{x,n}, \omega)$. We can thus rewrite eq. (24) as

$$4\frac{k_{z,0}k_{z,i}}{\rho_0\rho_i}\frac{1}{\omega^2}\left[|\tilde{V}^+|^2 - |\tilde{V}^-|^2\right] = \tilde{\Sigma}(k_{x,i}, \omega),$$
$$\text{with } \tilde{\Sigma}(k_{x,i}, \omega) = \left|B(k_{x,i}, \omega)\right|^2 \left|T_d(k_{x,i})\right|^2, \tag{26}$$

where we have used $\partial_{z,0}^{[1]}\tilde{V}^\pm = \pm i k_{z,0}\tilde{V}^\pm$ and eqs (2) and (13); $|P|^2$ denotes an element-wise amplitude squared of P. Also note that $\tilde{\Sigma}$ decomposes into a frequency dependent part $\left|B(k_{x,i}, \omega)\right|^2$ and frequency-independent part $\left|T_d(k_{x,i})\right|^2$. The latter term has already been studied by Mildner *et al.* (2019) and accounts for relative amplitudes. It is inconsequential for this work and will hence not be further considered.

Eq. (26) can be used directly for quality control. If it is evaluated using the correct (i.e. not blurred) fields [analogue to eq. (16)] it yields an unblurred Σ which is the input wavelet squared and carries the imprint of the overburden's two-way transmission response in k_x-direction (caused by the T_d^+'s on the right hand side). Since the k_x-dependence is not relevant for this work [it drops out upon deconvolution in eq. (17)], in the following we will neglect the two-way transmission imprint and assume a Σ that is *flat* in k_x-direction. For more information on how to correct for the overburden's transmission effect (i.e. the k_x-dependency of the focusing and Green's functions) we refer the reader to Mildner *et al.* (2019).[5] Otherwise, if eq. (26) is evaluated with the blurred fields, it yields $\tilde{\Sigma}$, which

[4] In case of a laterally varying target the same is not true for Green's functions $G^\pm$ and the reflection response $R^\cup$, because these also contain information about the target.

[5] The imprint of the two-way transmission is introduced on Σ due to the use of *propagated* wavefields [defined in eq. (13)]. If energy conservation is computed using the correct, non-propagated focusing functions [as in eq. (8)] the right-hand side is flat in k_x-direction.

can be used to correct the amplitude spectrum of $|\tilde{V}^+(k_{x,0}, \omega)|$. However, in the wavenumber-frequency domain it can not be used to recover any phase spectra. We therefore continue in the linear Radon domain, where we can exploit minimum phase properties per ray parameter p.

Taking the square root of $\underline{\tilde{\Sigma}}(\omega_\tau, p)$ yields $|\underline{B}(\omega_\tau, p)|$ which relates the blurred and deblurred (propagated) focusing functions as

$$|\underline{\tilde{V}}^+(\omega_\tau, p)||w(\omega)| = |\underline{\overline{V}}^+(\omega_\tau, p)||\underline{B}(\omega_\tau, p)|. \tag{27}$$

This equation can be used to derive the amplitude spectrum of the correct $\overline{V}^+$, for example, by finding a Moore-Penrose pseudo-inverse.

Seeing how densities $\rho(x_0)$ and $\rho(x_i)$ are typically unknown they can be omitted in eq. (26). As a result amplitudes of all reconstructed propagated focusing functions $\overline{V}^\pm$ and the propagated Green's functions $\overline{U}^\pm$ will be off by the same scalar, which will drop out when inverting eq. (17) for $\check{R}^\cup_{\text{Target}}$. Likewise, the multiplication with $k_{z,0}k_{z,i}$ can be omitted if offset-dependent scaling is not important (Mildner *et al.* 2019).

3.2 Phase correction

For media with a horizontally layered overburden the minimum phase condition [eq. (10)] can be rewritten as

$$\underline{T}^+(p, \omega_\tau)\,\underline{\overline{V}}^+(p, \omega_\tau) = 1. \tag{28}$$

The advantage of this is that in 1-D we can uniquely determine the phase given an amplitude spectrum using the Kolmogorov method (Claerbout 1976)

$$\text{Arg}[D(\omega)] = -\mathcal{H}[\log|D(\omega)|], \tag{29}$$

of any minimum phase response $D(\omega)$ via the Hilbert transform $\mathcal{H}$

$$\mathcal{H}[D(\omega)] = D(\omega) * \frac{1}{\pi\omega} \tag{30}$$

where $*$ stands for a convolution. This definition holds for both continuous and discrete signals, and highlights the fact that in order to determine the phase spectrum, the entire amplitude spectrum needs to be known. This generally is an immediate problem for phase spectrum recovery of *band-limited* signals (as opposed to *responses* that by our definition are always broadband), however it is not so in this particular case. To address this problem let us consider a band-limited minimum phase signal $\check{D}$ obtained by convolving a broad-band one $D(\omega)$ with some band-limiting rectangle function $\text{rect}(\omega)$ which is equal to identity on $[\omega_{min}, \omega_{max}]$ and some $\eta \ll 1$ otherwise. Using $\left|\check{D}\right|$ to get the phase spectrum of $\check{D}$ via the Kolmogorov method from eq. (29) and elementary properties of logarithms we obtain

$$\text{Arg}\left[\check{D}(\omega)\right] = \text{Arg}[D(\omega)] - \mathcal{H}[\log[|\text{rect}(\omega)|]], \tag{31}$$

where the second term is the approximate band-limitation artefact which is dependent on the actual value of η. Fortunately, applying $-\mathcal{H}[\log\cdots]$ to both sides of eq. (27) to solve for $-\mathcal{H}\left[\log\left|\overline{V}^+\right|\right]$ yields two such artefacts (one due to $\tilde{V}^+$ on the left and another due to $\underline{B}$ on the right-hand side), which approximately cancel out in expression

$$\text{Arg}\left[\underline{\overline{V}}^+(\omega_\tau, p)\right] = \mathcal{H}\left[\log|\underline{B}(\omega_\tau, p)|\right] - \mathcal{H}\left[\log|\underline{\tilde{V}}^+(\omega_\tau, p)|\right]. \tag{32}$$

This allows reliable recovery of a phase spectrum from a band-limited signal.

4 NUMERICAL EXAMPLE

A complex 2-D velocity model that consists of a laterally invariant overburden and a laterally varying target was kindly provided by Shell, see Fig. 2(a). The overburden, with a thickness of over 1 km, is extremely finely layered (with medium properties varying every 2 m), whereas the target is composed of three 'packages' of reflectors. This lossless acoustic model is inspired by sonic measurements in the Middle East and is discretized on a 2 m × 2 m grid. A corresponding density distribution is obtained from Gardner's relation (Gardner *et al.* 1974), and reflection coefficients (Fig. 2b) highlight the fine layering and relatively high-impedance contrast, which should cause a significant amount of overburden-borne SPIM. Industry standard approaches to handling internal multiples in such a setting would amount to dip-filtering the migrated result, however that would come at the cost of removing the apices of the anticlines in the target. Available internal demultiple elimination methods would fail here because of a sheer number of possible internal multiple generators and because of temporal inseparability of primaries. Layer-stripping approaches would likely introduce cascading down errors in the amplitudes of the demultipled data.

Here we wish to demonstrate that a Marchenko equation-based scheme can perform reasonably well and that the augmented scheme that we introduce additionally handles all orders and all periods of internal multiples properly. The workflow follows the one outline in Section 3, but given the two-dimensionality of the example the vectors $\boldsymbol{k}_x$ and $\boldsymbol{p}$ reduce to scalars k_x and p. We ensure that the model is homogeneous above $z = 0$ (no free surface) and that we can extend the model upwards by 100 m, to ensure that the onset of the first primary reflection (due to a shallow reflector originally around 50 m) arrives after $t = \varepsilon$ (i.e. a vertical upward redatuming of sources and receivers). This is in line with the discussion in Section 2.3, where we wish to ensure that $\Phi_\varepsilon[V^-] = V^-$ is satisfied. We illustrated this with the auxiliary acquisition surface $\partial \mathbb{D}_{0,aux}$ in Fig. 2.

To model synthetic 2-D reflection data we use a conventional second-order staggered-grid finite-difference solver (Virieux 1984) on a computational grid of 10.0 by 3.1 km with 2 m discretization (i.e. 5001 by 1551 grid points). Appropriate data scaling in line with the definitions of Green's function G and $R^\cup$ (as given in Section 2.1) requires that the data are a 2-D Green's function recorded with colocated dipole-sources and monopole receivers. Here we use a fixed-spread acquisition with a grid of 1001 sources and receivers of 10 m spacing, and the data is modelled using a 20 Hz Ricker wavelet. The data is appropriately scaled and deconvolved for the wavelet used in modelling to conform with the definition of $\check{R}^\cup$. Lastly, these data are free from random noise[6] and do not contain surface-related multiples. A shot gather of $\bar{r}^\cup(\boldsymbol{x}_0'', \boldsymbol{x}_0)$ is presented in Fig. 5(a).

We choose the overburden-target separating boundary $\partial \mathbb{D}_i$ at $z = 1270$ m depth (marked with a dashed red line in Fig. 2). This choice ensures that a mute Φ_ε can be defined such that the side

lobes of $\overline{V}^-$ and $\overline{U}^-$ do not overlap, that is, $X^- = 0$ (or $\tilde{X}^- = 0$) in eq. (19). It was suggested by van der Neut & Wapenaar (2016) that this mute can be obtained by picking a primary that is due to a strong reflector in the target. However, considering the shot gather in Fig. 5(a) it is clear that this is a tall order in this setting (we used red arrows to indicate some target primaries). Additionally, there is little evidence supporting the claim that one could use primary reflections due to non-horizontal reflectors to derive said mute.[7] Instead, we use a smoothed (migration) velocity model and an eikonal solver to obtain two way travel times $t_d(\boldsymbol{x}_0, \boldsymbol{x}_0')$ between $\partial \mathbb{D}_0$ and $\partial \mathbb{D}_i$ [shown with dashed red lines in Fig. 6(a)] that are used in the definition of the mute Φ_ε in eq. (18).

We use a temporally and spatially band-limited delta spike as the first term V_e^+ of the Neumann series expansion [see eq. (22)]. It is obtained by defining its triangular-shaped spectrum in the wavenumber-frequency domain (compare with Fig. A1), and dressing it with a 20 Hz Ricker wavelet. Given the fine layering we know that this choice is incorrect and hence the solutions will be the blurred wavefields $\tilde{V}^\pm$, which we will subsequently fix. Since highly scattering overburdens require many iterations (Dukalski & de Vos 2017), we calculate the first 100 terms in eq. (22). Convergence is further analysed in Appendix F. The resulting propagated downgoing focusing function is illustrated in $\boldsymbol{x}$-t-domain ($\tilde{v}^+$) and in τ-p-domain ($\tilde{\underline{v}}^+$) in Figs 6(b) and (e), respectively. At the early end of the time window it can clearly be seen how Φ_ε protects the first event $\tilde{v}_e^+$, which is identical in Figs 6(a) and (b).

The blurred focusing functions $\tilde{v}^+$ and $\tilde{v}^-$ are Fourier transformed to the k_x-ω-domain and used to evaluate energy conservation [eq. (26)]. The result is displayed in Figs 7(a), (e) and (i) in the $\boldsymbol{x}$-t-domain ($\tilde{\sigma}$), k_x-ω-domain ($\tilde{\boldsymbol{\Sigma}}$) and τ-p-domain ($\tilde{\underline{\sigma}}$), respectively. It can be seen in a number of domains that the propagated focusing functions are blurred: in the $\boldsymbol{x}$-t-domain we do not obtain a single spike for $\tilde{\sigma}$, but we observe additional side-lobe energy; in the k_x-ω-domain the spectrum is not flat in k_x-direction but has a peak around $k_x = 0$ which is not centered around the peak frequency of the input wavelet (it is at 15 Hz instead of at 20 Hz); and in τ-p-domain $\tilde{\underline{\sigma}}$ deviates from a straight line and contains too much side-lobe 'ringiness'.

The amplitude and phase spectrum of the blurred, propagated focusing functions need to be corrected, using eq. (27) and eq. (32) [compare also with Fig. 4]. As outlined in Appendix A, the corrected $\overline{V}_e^+$ is isolated with Φ_ε, which is subsequently used to re-evaluate eq. (22). The resulting band-limited (propagated) focusing function is shown in Figs 6(c) and (f). The difference to the previous, blurred equivalents [Figs 6(b) and (e)] is most visible around the first arrival, which clearly differs from our first estimate [Figs 6(a) and (d)]. Additionally, the 'noisy' events beyond $|p| > 1.7 \times 10^{-4} \text{sm}^{-1}$ are not present in the corrected result since they are not contained in $\tilde{\underline{\sigma}}$ [Fig. 7(i)] and consequently many of these linear artefacts are removed. To highlight that all events are affected by the correction process we compare central traces in Fig. 6(g). Although the differences might seem subtle, the impact on imaging is considerable, as we will see later.

The quality of the deblurred focusing functions is controlled by re-evaluating energy conservation, shown in Figs 7(b), (f) and (j). As opposed to the results from the blurred wavefields (left-hand column of the same figure), a significant improvement can be observed: in

---

[6]There is little published work on effects of random noise on the solutions to the Marchenko equation in 1-D or 2-D, but we expect that in the presence of random (incoherent) noise, similar to methods such as SRME or IME, multidimensional convolutions/correlations of noisy input data will stack out incoherent noise.

[7]In Wapenaar *et al.* (2014), the authors always require that the boundary $\partial \mathbb{D}_i$ is horizontal to ensure actual up-down (i.e. normal to $\partial \mathbb{D}_i$) decomposition of the focusing and the Green's functions.
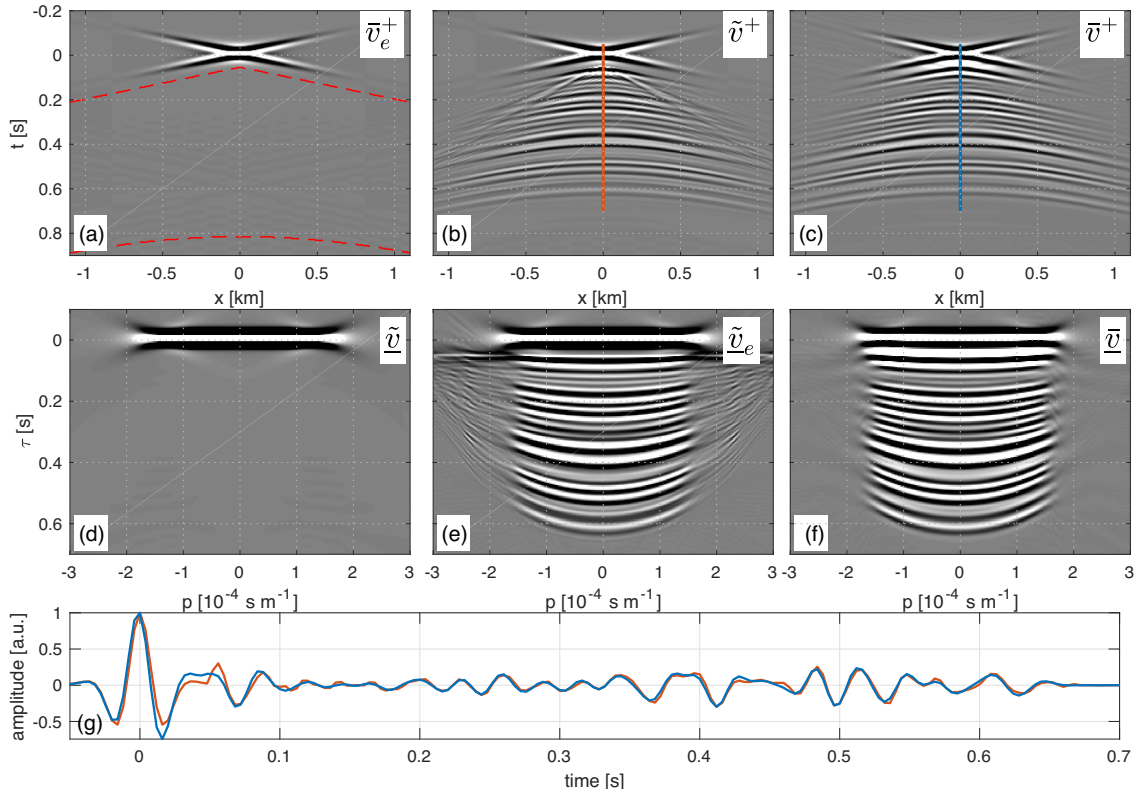
Figure 6. Propagated focusing functions in different domains, before and after deblurring. The top row shows quantities in the x-t-domain, the central row in the τ-p-domain. (a) $\tilde{v}_e^+$ with the projection condition Φ_ε illustrated using dashed red lines, (b) $\tilde{v}^+$, (c) $\bar{v}^+$, (d) $\underline{\tilde{v}}_e^+$, (e) $\underline{\tilde{v}}^+$, (f) $\underline{\bar{v}}^+$, (g) central traces taken from the propagated focusing functions before (red) and after (blue) deblurring. To highlight their difference at late times panels (a)–(c) and (g) are displayed with a t^2-gain; panels (a)–(f) are clipped to 5 per cent of the maximum amplitude.

the x-t-domain energy conservation yields an isolated, band-limited spike, in the k_x-ω-domain the spectrum is flat in k_x-direction, and in the τ-p-domain we observe a flat 'line' without the side-lobe energy that is present in Fig. 7(i). The dominant frequency is shifted towards higher frequencies, such that the peak frequency after correction occurs at the frequency of the input wavelet (20 Hz), as it should. These observations assure us that we have successfully corrected the amplitude spectrum of the focusing functions. However, they do not allow any conclusions about the correctness of the newly obtained phase (though an incorrect amplitude spectrum will always give rise to an incorrect phase spectrum, seeing how the latter is derived from the former). Note that the frequency dependent errors in the amplitude spectrum may actually drop out when obtaining the target medium reflection from multidimensional convolution [see eq. (17)]. Analysis of the phase spectrum is thus important, and will be addressed when quality controlling the migrated reflection data.

The rest of the workflow follows the conventional approach, see the last two steps in Fig. 4. Inverting for the reflection response of the target medium $\check{R}_{\text{Target}}^\cup(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$ is done by solving the normal problem $\bar{U}^- \bar{U}^{+\dagger} = \check{R}_{\text{Target}}^\cup \bar{U}^+ \bar{U}^{+\dagger}$ (Amundsen *et al.* 2001; van der Neut *et al.* 2011) per frequency slice using MATLAB's rdivide functionality. This step requires stabilization of low amplitudes, which we implement by ensuring a denominator with at least 5 per cent of its maximum amplitude. The result for a central shot is shown in Fig. 5(d)

To benchmark the performance of the scheme presented in this paper we compare it to the conventional Marchenko scheme (i.e. the black arrows in Fig. 4). The resulting target-data is illustrated in Fig. 5(c). Whereas the augmented scheme imposes an aperture

limitation (to avoid artefacts introduced by the Radon transform, see Appendix A), this is not the case for the conventional scheme. We therefore use conventional Marchenko solutions with the widest possible aperture [i.e. the aperture shown in Fig. A1(b)]. Hence, the resulting focusing functions (not shown) have a wider aperture than the augmented ones [and wider than the ones shown in Figs 6(b) and (e), which have been artificially limited in aperture for better comparison with panels (c) and (f)].

The fact that Fig. 5(c) does not have a significantly wider aperture than Fig. 5(d) (regardless of using propagated focusing functions with wider aperture), is due to filtering. In order to suppress linear artefacts introduced by multidimensional deconvolution, Marchenko-derived reflection responses have been dip-filtered in the k_x-ω-domain using a filter velocity of $c_{\text{filt}} = 4000$ m s^{-1} (such that all signal is muted for which $|k_x| > f/c_{\text{filt}}$). This results in similar apertures of the Marchenko derived reflection data, which is reduced as opposed to the modelled data in panels (a) and (b). Comparison of panels (c) and (d) shows that both data contain the relevant main primary reflections related to the three 'packages' of reflectors in the target (marked with red arrows), however the deblurred (augmented Marchenko output) data in panel (d) contains less coherent and incoherent noise.

Moreover, in order to show that the deblurred result is an actual improvement and not just different from the conventional one, we need to compute the *reference* target medium reflection response (i.e. the result of the workflow in Fig. 4). This proves to be no easy task: one could model $\check{R}_{\text{Target}}^\cup(\boldsymbol{x}_i, \boldsymbol{x}_i')$, with sources and receivers along $\partial\mathbb{D}_i$, however, this results in illumination angle differences, which are not captured by our acquisition of $R^\cup(\boldsymbol{x}_0, \boldsymbol{x}_0')$
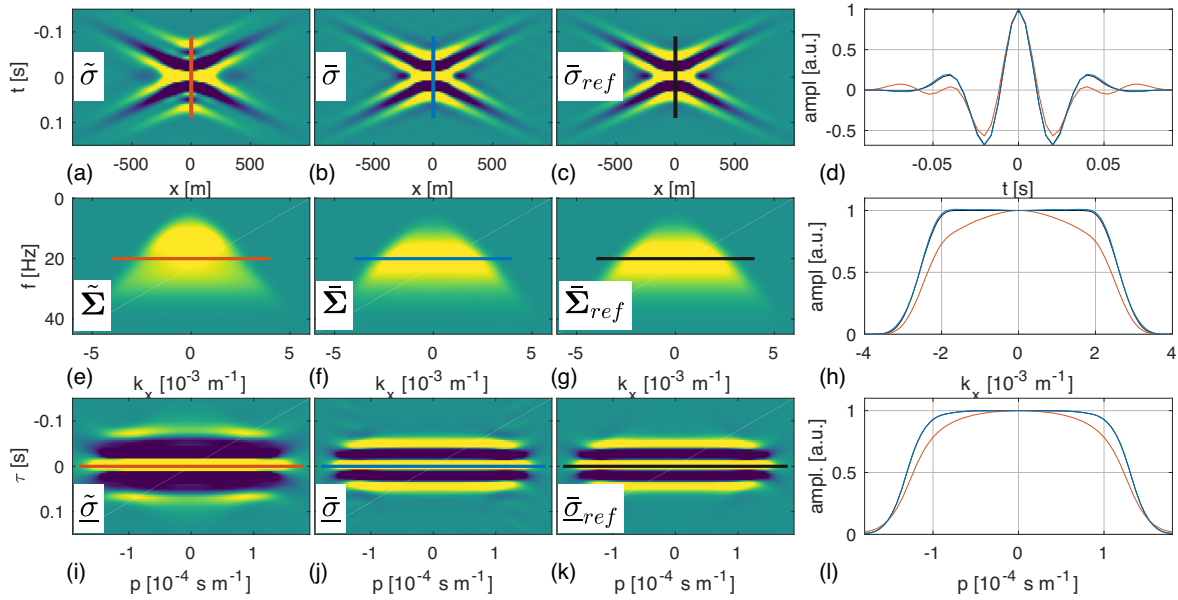
Figure 7. Energy Conservation [eq. (26)] displayed in different domains and evaluated for blurred, band-limited wavefields (left-hand column), for the corrected (deblurred and band-limited) wavefields (second column). The third column illustrates references (for an ideal point source created by the focusing functions). The right-hand column shows individual traces taken from the panels in the first three columns. Panels in the first row are in the *x*-*t*-domain, panels in the second row in the k_x-ω-domain, and panels in the third row in the τ-*p*-domain. Panels in the first, second and third row are clipped to 5, 100 and 8 per cent of their maximum amplitudes, respectively.

(sources and receivers along $\partial\mathbb{D}_0$). It would therefore require artificial post-processing of $\check{R}^{\cup}_{\text{Target}}(\boldsymbol{x}_i, \boldsymbol{x}_i')$ to reduce the illumination angles, which introduces additional artefacts. For these reasons we have opted for an alternative approach which respects the ray content of the data used on input. We smooth the overburden (in slowness) in order to properly replicate the kinematics of the overburden and remove any scattering, and leave the target untouched, resulting in Fig. 2(d). We use the same finite difference modelling scheme as before to obtain $\check{R}^{\cup}_{\text{Target}}(\boldsymbol{x}_0, \boldsymbol{x}_0')$, illustrated in Fig. 5(b). Although very smooth, the overburden still introduced some reflections which also result in some unwanted (very low frequency) overburden-related multiples that will be visible in the final image.

The four reflection data sets described above (Fig. 5) are Kirchhoff depth migrated (e.g., Schneider 1978), and the results are shown in Fig. 8. Note that no scaling or normalization was applied to the redatumed results. Migration of the conventional Marchenko preprocessing result shown in panel (c) is clearly an improvement over the migrated input data [panel (a)], as it removes a lot of coherent internal-multiple noise. Moreover, Fig. 8(c) indicates that conventional Marchenko appears to perform reasonably well in this setting (suggesting some robustness in handling the SPIM-induced errors).

The result produced by the augmented Marchenko scheme in Fig. 8(d) is much more similar to the input model [Fig. 2(a)] than the conventional Marchenko approach shown in panel (c). It is also closer to the reference solution shown in Fig. 8(b) [though it is important to keep in mind that the latter also contains spurious overburden-related artefacts, indicated with arrow (I)]. Comparison of Figs 8(c) and (d) highlights that the former contains both acausal and causal ghosts [events appear to be *smeared out*], marked with arrows (II) and (III) for the first package of events. Neglecting SPIMs also introduces artefacts [arrow (IV)] which may interfere with other events such that, for example, continuous events seem discontinuous [arrows (V)].

To highlight the differences between the two Marchenko results we display vertical traces in Figs 8(e) and (f). The augmented Marchenko results follow the reference traces very closely. Moreover, the conventional Marchenko result appears to have a small overall phase shift, which is not present in the augmented Marchenko result.

In the review process it has been pointed out that a bilinear minimization process that relies on sparsity of the solution (similar to the one presented by Becker *et al.* 2018) *could* be a viable alternative to incorporate SPIMs into the Marchenko method. However, such an inversion process involves inverting an underconstrained problem, and there is no physical proof that dictates that the true solution is sparse. To date, no attempts to account for SPIMs with an algorithm alike the one proposed by Becker *et al.* (2018) have been made, and there is no evidence that such an approach would work in general.

5 CONCLUSION AND OUTLOOK

We have introduced an augmented Marchenko method for 2-D and 3-D media which can correctly account for SPIM in a horizontally layered overburden. Our imaging results for a complex synthetic model, representing the challenges in the Middle East, reveal that conventional Marchenko redatuming removes a great amount of overburden-related internal multiples and thus yields much cleaner images than migration of surface data. Moreover, the augmented Marchenko method correctly accounts for SPIMs, yielding another significant improvement with sharper events and fewer artefacts.

The theory we introduced is valid for general 2-D and 3-D media, however, we are currently only able to solve it for laterally invariant overburdens. This is because, in the latter case, the problem is (irreducibly) represented per ray parameter, where minimum phase recovery is possible using the Kolmogorov method. For general, laterally varying media different ray parameters will be coupled, and hence the entire solution (all coupled rays) will be a multidimensional minimum phase response (as opposed to each ray being
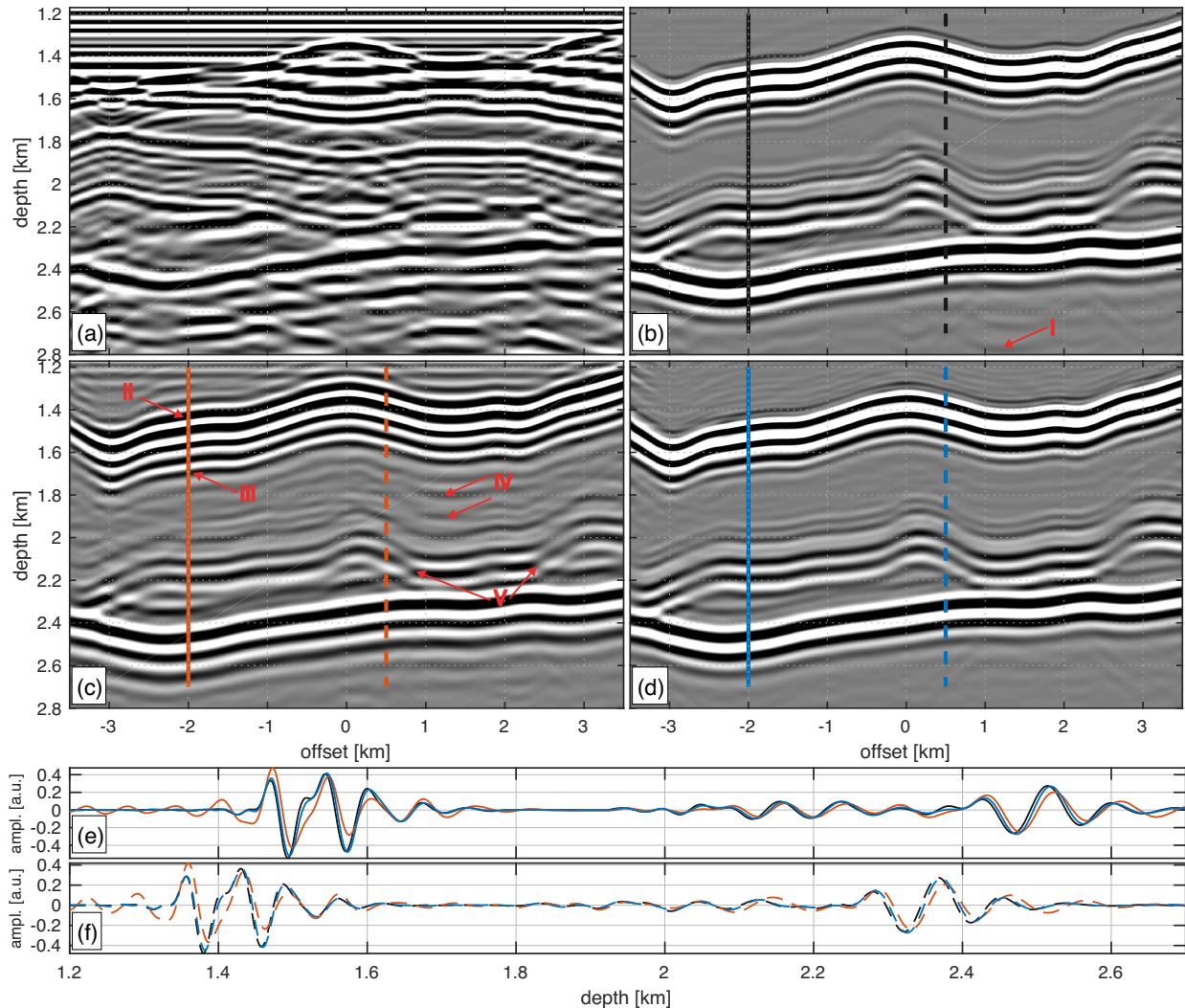
Figure 8. Migration results for the target model, using the offset range marked with a black box in Fig. 2. Panels (a)–(d) correspond to the data for which central shots are shown in Fig. 5. (a) Migration of modelled reflection data computed in the entire model; (b) migration of modelled data corresponding to the target model in Fig. 2(d), used as a reference for panels (c) and (d); (c) migration of target data derived from blurred Marchenko results; (d) migration of target data derived from deblurred Marchenko results; (e) and (f) show traces recorded at −2 km and 0.5 km offset, respectively. Red arrows point at erroneous events and artefacts discussed in the text. Panels (a)–(d) are clipped to 20 per cent of the maximum amplitude.

minimum phase individually). Seeing how in 1.5-D, for geologically relevant models, the correction imposed by energy conservation and minimum phase was rather subtle [but which had significant effects on the migrated image, compare the traces in Fig. 6 (g) with Figs 8(c) and (d)], we first require sensitive tools for quality control (i.e. that allow us to accurately calculate true (propagated) focusing functions in complex media). However, such tools are not available yet. Once a robust methodology is found to address this problem, further work will be necessary to investigate, understand and quantify how quickly the algorithm proposed in this work breaks down on a *range* of synthetics. A generalized solution for more complex media will require a unique matrix factorization and understanding of minimum-phase properties in higher dimension than 1-D. This will be subject to future research.

An initial and possibly simpler approach involves investigating the limitations of a ray-parameter approach for subhorizontally layered media. For media with gently dipping interfaces and without significant horizontal scattering a constant ray parameter approach might still be acceptable.

Apart from accounting for SPIM, extending the conventional Marchenko equations with energy conservation may enable a range of other applications. Mildner *et al.* (2019) have shown how to obtain true amplitude-versus-offset Green's functions (a step that is implicit in our work), and Reinicke *et al.* (2019) introduced ways to deal with overlapping events in elastodynamic Marchenko redatuming. Possible future research questions could involve the use of energy conservation for assessing forward-scattered acoustic waves (such as prismatic events related to salt diapirs) as well as tunneling effects in finely layered media with high velocity contrasts.

ACKNOWLEDGEMENTS

REFERENCES

Agranovich, Z. & Marchenko, V.A., 1963. *The Inverse Problem of Scattering Theory,* Gordon and Breach.

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

Amundsen, L., Ikelle, L.T. & Berg, L.E., 2001. Multidimensional signature deconvolution and free-surface multiple elimination of marine multicomponent ocean-bottom seismic data, *Geophysics,* **66**(5), 1594–1604.

Becker, T., Ravasi, M., van Manen, D., Broggini, F. & Robertsson, J.O.A., 2018. Sparse inversion of the coupled marchenko equations for simultaneous source wavelet and focusing functions estimation, in *Proceedings of the 80th EAGE Conference and Exhibition 2018.*

Berkhout, A. & Verschuur, D., 1997. Estimation of multiple scattering by iterative inversion, Part I: theoretical considerations, *Geophysics,* **62**(5), 1586–1595.

Berkhout, A.J. & Verschuur, D., 2005. Removal of internal multiples with the common-focus-point (CFP) approach. Part 1: explanation of the theory, *Geophysics,* **70**(3), V45–V60.

Broggini, F., Snieder, R. & Wapenaar, C., 2012. Focusing the wavefield inside an unknown 1D medium: beyond seismic interferometry, *Geophysics,* **77**(5), A25–A28.

Claerbout, J.F., 1968. Synthesis of a layered medium from its acoustic transmission response, *Geophysics,* **33**(2), 264–269.

Claerbout, J.F., 1976. *Fundamentals of Geophysical Data Processing,* Blackwell Scientific Publications.

Deans, S.R., 1983. *The Radon Transform and Some of its Applications,* John Wiley & Sons, Inc.

Dukalski, M. & de Vos, K., 2017. Marchenko inversion in a strong scattering regime in the presence of a free surface, *Geophys. J. Int.,* **212,** 760–776.

Dukalski, M., Mariani, E. & de Vos, K., 2019. Handling short-period scattering using augmented Marchenko autofocusing, *Geophys. J. Int.,* **216**(3), 2129–2133.

Foster, D.J. & Mosher, C.C., 1992. Suppression of multiple reflections using the Radon transform, *Geophysics,* **57**(3), 386–395.

Frijlink, M. & Wapenaar, K., 2004. Migration with correction for transmission losses in arbitrary inhomogeneous media, in *SEG Technical Program Expanded Abstracts 2004,* pp. 1029–1032 , Society of Exploration Geophysicists.

Gardner, G.H.F., Gardner, L.W. & Gregory, A.R., 1974. Formation velocity and density—the diagnostic basics for stratigraphic traps, *Geophysics,* **39**(6), 770–780.

Horn, R.A. & Johnson, C.R., 1990. *Matrix Analysis,* Cambridge Univ. Press.

Jakubowicz, H., 1998. Wave equation prediction and removal of interbed multiples, in *Proceedings of the 68th Annual International Meeting, SEG, Expanded Abstracts,* pp. 1527–1530.

Lamoureux, M.P. & Margrave, G.F., 2007. An analytic approach to minimum phase signals, *CREWES Res. Rep,* **19,** https://www.crewes.org/ForOurS ponsors/SlideShows/2007/2007_13_ppt.pdf.

Mildner, C., Dukalski, M., Elison, P., De Vos, K., Broggini, F. & Robertsson, J., 2019. True amplitude-versus-offset green's function retrieval using augmented Marchenko focusing, in *Proceedings of the 81st EAGE Conference and Exhibition 2019.*

Noether, E., 1918. Invariante Variationsprobleme, *Nachr. d. König. Gesellsch. d. Wiss. zu. Göttingen, Math-phys. Klasse,* pp. 235–257.

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

Reinicke, C., Dukalski, M. & Wapenaar, C., 2019. Tackling different velocity borne challenges of the elastodynamic Marchenko method, in *Proceedings of the 81st EAGE Conference and Exhibition 2019.*

Resnick, J., Lerche, I. & Shuey, R., 1986. Reflection, transmission, and the generalized primary wave, *Geophys. J. R. astr. Soc.,* **87,** 349–377.

Rickett, J.E., 2001, Spectral factorization of wavefields and wave operators, *PhD thesis,* Stanford University.

Schneider, W.A., 1978. Integral formulation for migration in two and three dimensions, *Geophysics,* **43**(1), 49–76.

Schonewille, M.A., 2000, Fourier reconstruction of irregularly sampled seismic data, *PhD thesis,* TU Delft.

Sherwood, J. & Trorey, A., 1965. Minimum-phase and related properties of the response of a horizontally stratified absorptive earth to plane acoustic waves, *Geophysics,* **30**(2), 191–197.

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

Skingle, C., Heron, K. & Gaukroger, D., 1977. The application of the Hilbert Transform to System Response Analysis, HM Stationery Office, 1378.

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

Staring, M., Pereira, R., Douma, H., van der Neut, J. & Wapenaar, C., 2017. An adaptive method for source-receiver Marchenko redatuming, in *Proceedings of the 15th International Congress of the Brazilian Geophysical Society,* Rio de Janeiro, Brazil, pp. 1323–1328.

Staring, M., Pereira, R., Douma, H., van der Neut, J. & Wapenaar, C., 2018. Source-receiver Marchenko redatuming on field data using an adaptive double-focusing method, *Geophysics,* **83**(6), S579–S590.

Thomsen, H.R., Broggini, F., Van Manen, D.J., Ravasi, M. & Kritski, A., 2017. Robust Marchenko focusing—calibrating surface reflection with VSP data, *Proceedings of the 79th EAGE Conference and Exhibition 2017,* pp. 12–15.

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

van der Neut, J., Thorbecke, J., Mehta, K., Slob, E. & Wapenaar, K., 2011. Controlled-source interferometric redatuming by crosscorrelation and multidimensional deconvolution in elastic media, *Geophysics,* **76**(4), SA63.

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

Virieux, J., 1984. SH-wave propagation in heterogeneous media: velocity-stress finite-difference method, *Geophysics,* **49**(11), 1933–1942.

Wapenaar, C., 1998. Reciprocity properties of one-way propagators, *Geophysics,* **63**(4), 1795–1798.

Wapenaar, C. & Berkhout, A., 1989. *Elastic Wave Field Extrapolation,* Elsevier.

Wapenaar, C., Draganov, D. & Thorbecke, J., 2003. Relations between codas in reflection and transmission data and their applications in seismic imaging, *Proceedings of the 6th SEGJ International Symposium,* pp. 152–159.

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

Wapenaar, C.P.A. & Grimbergen, J.L.T., 1996. Reciprocity theorems for one-way wavefields, *Geophys. J. Int.,* **127**(1), 169–177.

Ware, J.A. & Aki, K., 1969. Continuous and discrete inverse-scattering problems in a stratified elastic medium. 1. Plane waves at normal incidence, *J. acoust. Soc. Am.,* **45,** 911–921.

Weglein, A.B., Gasparotto, F., Carvalho, P. & Stolt, R., 1997. An inverse-scattering series method for attenuating multiples in seismic reflection data, *Geophysics,* **62**(6), 1975–1989.

Zhang, L., Slob, E., van der Neut, J. & Wapenaar, C., 2018. Artifact-free reverse time migration, *Geophysics,* **83**(5), A65–A68.

Zhou, B. & Greenhalgh, S.A., 2002. Linear and parabolic τ-p transforms revisited, *Geophysics,* **59**(7), 1133–1149.
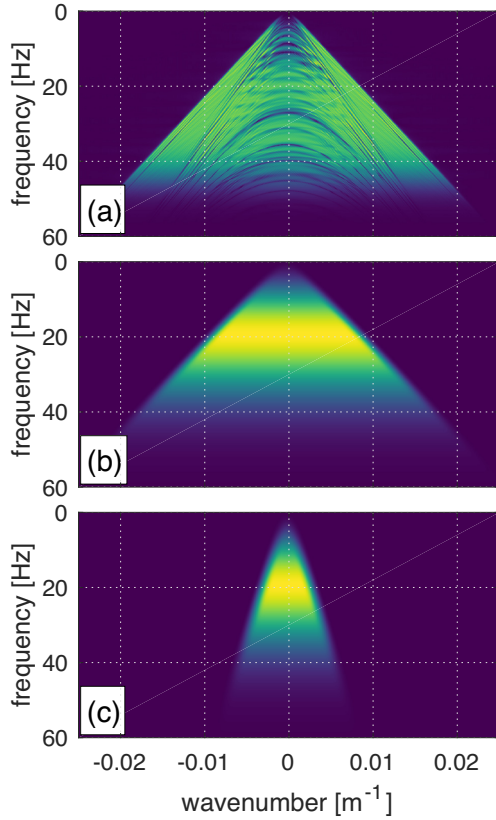
Figure A1. Spectra in k_x-ω-domain. (a) $\check{R}^\cup$, (b) $\tilde{V}_e^+$ with the same aperture as the data in panel (a), (c) $\tilde{V}_e^+$ with a narrow aperture used to create the deblurred examples in Section 4. All panels are clipped to 100 per cent of the maximum amplitude.

APPENDIX A: IMPLEMENTATION ASPECTS OF AUGMENTED MARCHENKO FOR SPIM

A1 Amplitude stabilization for phase retrieval

It has already been recognized by Lamoureux & Margrave (2007) that Hilbert transforms of band-limited signals need to be stabilized around frequencies which are (close to) zero, and that the result can be sensitive to amplitude stabilization for each of the terms on the right-hand side of eq. (32) . Therefore to improve the band-limitation-borne artefact cancellation in this expression, prior to evaluating the logarithm we stabilize the amplitude spectra as follows

$$\left|\underline{D}(\omega_\tau, \boldsymbol{p})\right|_\eta = \begin{cases} \left|\underline{D}(\omega_\tau, \boldsymbol{p})\right| & \text{if } \left|\underline{D}(\omega_\tau, \boldsymbol{p})\right| \geq \eta \\ \eta, & \text{otherwise.} \end{cases} \quad (A1)$$

This yields a significant improvement however it still leaves one with a free parameter of η. If chosen too small the phase error can be very large and vary considerably on $[0, f_{\text{Nyquist}}]$ and make cancellation less efficient. If η is set too large that could mask the relevant signal in $\underline{D}$. Investigations on simple 1-D models have shown that we can express η in terms of the maximum value of the amplitude spectra, based on the overall transmission amplitude from $\partial\mathbb{D}_0$ to $\partial\mathbb{D}_i$: the lower the transmission amplitude (i.e. the more interfaces, or the stronger the scattering in the overburden) the lower the required stabilization relative to the maximum amplitude. For transmission amplitudes of more than 0.6 we suggest a stabilization of 10 per cent, whereas for more complex overburdens with smaller

transmission amplitudes a stabilization of 2 per cent yields good results. For geological settings where SPIMs play a significant role the latter is more likely to be the factor of choice.

Moreover, we have found that ℓ^2-normalizing the amplitude spectra (prior to stabilization with η) balances the two terms on the right-hand side of eq. (32), such that the expression becomes

$$\begin{aligned} \text{Arg}&\left[\underline{\overline{V}}^+(\omega_\tau, \boldsymbol{p})\right] \\ &= \mathcal{H}\left[\log|\underline{B}(\omega_\tau, \boldsymbol{p})|_{\eta,n}\right] - \mathcal{H}\left[\log|\underline{\tilde{V}}^+(\omega_\tau, \boldsymbol{p})|_{\eta,n}\right], \end{aligned} \quad (A2)$$

where the additional subscript n denotes ℓ^2-normalization. This results in more similar band-limitation borne artefacts that cancel each other out more effectively, and thus further improve the efficacy of phase spectra recovery.

A2 Using propagated focusing functions

The use of propagated focusing functions $V^\pm$ over the (conventional) focusing functions $F^\pm$ has practical advantages in addition to the aforementioned mathematical ones. On the one hand we find that fewer linear artefacts originating from the Rayleigh integral are introduced in the propagated focusing functions. Although we have not studied this in detail, this might originate from a better handling of refractions. In conventional Marchenko these artefacts so far had not been identified as a serious cause for concern. However, in our scheme better behaved linear artefacts are an advantage, due to the need for forward and inverse Radon transform (compare with the workflow summarized in Fig. 4).

The same applies to events that are cut off by the projector, which is why we cannot chose a first event with too large aperture. We illustrated this problem in Fig. A1: if we chose the first event $\tilde{V}_e^+$, shown in panel (b), with the same aperture as the reflection data $\check{R}^\cup$ in panel (a), the time domain version will have too steep flanks, and consequently the projector will cut into the first events of the created coda $\tilde{V}_l^+$. Such a truncation of events translates to strong artefacts in the τ-$\boldsymbol{p}$-domain upon forward Radon transform. To prevent this from happening, we have chosen a $\tilde{V}_e^+$ with a relatively narrow aperture [Fig. A1(c)] and therefore gently dipping slopes in the time domain, as seen in its time-domain equivalent in Fig. 6(a). The time domain equivalent of Fig. A1(b) would have much steeper slopes as opposed to the object in Fig. 6(a), and hence would cut into the coda events visible in Fig. 6(b).

Note that the correct aperture for $\tilde{V}_e^+$ is given by its definition $\tilde{V}_e^+ = T_d^{+,inv} T_d^+ \mathrm{d}\boldsymbol{x}_i$. The aperture of this expression is narrower than the input data [Fig. A1(a)], since the aperture of T_d^+ is similar to a primary from $\partial\mathbb{D}_i$ (and hence narrower than that of earlier primaries contained in the data). These considerations might therefore justify choosing a narrow aperture for $\tilde{V}_e^+$ such as the one in Fig. A1(c).

A3 Rerun Marchenko after amplitude and phase correction

After reconstruction of $\tilde{V}^+ \to \overline{V}^+$ (amplitude and phase correction), we extract $\overline{V}_e^+$ from $\overline{V}^+$ with an appropriately chosen time windowing operator $1 - \Phi_\varepsilon$ (i.e. the green region in Fig. 3). This (corrected) early focusing field is subsequently used to re-evalute eq. (22), which together with eq. (19) yields $\overline{V}^-$ (with $X^- = 0$). In theory this is not necessary, because we should have already obtained both amplitude and phase information with eq. (27) and

(32). We find however, that the target medium reflection response obtained that way is closer to the true one.

A4 Definition of the linear Radon transform

Care has to be taken with the linear Radon transform. The forward and its corresponding inverse transform require a pre-factor of $|\omega|/(2\pi)$, and it is a free choice how this factor is 'distributed' over the foward/inverse transform pair. Zhou & Greenhalgh (2002) denote these transforms with $|\omega|/(2\pi)$ applied on either the forward transform or the inverse transform as version 1 and version 2, respectively.

If version 1 of the forward linear Radon transform is used, the fact that the pre-factor $|\omega|/(2\pi)$ has been used on the forward transform gives an added phase distribution to $|\underline{B}|$, and gives rise to an erroneous phase spectrum of $|\underline{\overline{V}}^+|$. Consequently, this problem can be avoided if version 2 of the linear Radon transform is used. Insights into the implementation of the Radon transform can be found in Schonewille (2000).

APPENDIX B: MINIMUM PHASE BEYOND 1-D

The statement 'when measured from the onset of signal' means that in 1-D the wavefields (as well as their inverses) are shifted to start at $t = 0$. In 2-D and 3-D, this shift is offset dependent and hence to get a better understanding on how to extend this concept beyond 1-D in this work a 1.5-D (laterally invariant 2-D or 3-D) medium is considered. For the laterally invariant scenario the problem can be decoupled into a series of 1-D problems, each characterized by some ray parameter $\boldsymbol{p}$, where eq. (10) becomes

$$\underline{T}^+(\boldsymbol{p})\,\underline{F}_1^+(\boldsymbol{p}) = 1 \,\forall\, \boldsymbol{p} < \boldsymbol{p}_{\max}, \tag{B1}$$

and $\boldsymbol{p}_{\max}$ is the ray parameter for the maximally oblique plane wave component measured at $\partial\mathbb{D}_0$. The fields satisfying the minimum phase condition are $\underline{T}^+(\boldsymbol{p})\,e^{-i\omega_\tau\tau_d(\boldsymbol{p})}$ and $\underline{F}_1^+(\boldsymbol{p})\,e^{i\omega_\tau\tau_d(\boldsymbol{p})}$, where the exponent $i\omega_\tau\tau_d(\boldsymbol{p})$ takes care of the $\boldsymbol{p}$-dependent linear phase shift.

Multi-dimensional minimum phase responses are more complex than their 1-D equivalents, and certain conditions need to be better investigated in the future (compare with Rickett 2001; Wapenaar *et al.* 2003; Frijlink & Wapenaar 2004). Possible research questions include if invertibility and stability are always satisfied (e.g., in complex media with large local velocity variations or for different data acquisition modes), and whether the former should be replaced with pseudo-invertability.

APPENDIX C: TRANSMISSION CORRECTION OF THE REDATUMED DATA

In this section, we use a simple 1-D example to demonstrate that the target data $R_{\text{Target}}^{\cup}(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$, obtained from eq. (17) by (multidimensional) deconvolution of the propagated focusing functions $U^{\pm}$, are corrected for overburden-related transmission effects. This can be interpreted that in the process of retrieving $R_{\text{Target}}^{\cup}(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$ from $R^{\cup}(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$, one removes all events from the overburden (including primaries and multiples), and the remaining (underburden-related) events are identical to those in the input data $R^{\cup}(\boldsymbol{x}_0, \boldsymbol{x}_0', \omega)$, but corrected for an angle-dependent transmission coefficient of the overburden.

Assume a 1-D medium with three reflectors R_1 to R_3, which are described by $Z_i = r_i \exp\{i\omega t_i\}$, where r_i is the reflection coefficient of reflector R_i and t_i is the two-way traveltime (Slob *et al.* 2014). The reflection coefficient $R^{\cup}(\omega)$ for such a medium is given by

$$R^{\cup}(\omega) = \frac{Z_1 + Z_2 + Z_3 + Z_3 Z_2^* Z_1}{1 + Z_2 Z_1^* + Z_3 Z_2^* + Z_3 Z_1^*}. \tag{C1}$$

From eq. (22) the propagated focusing functions V^+ and V^- can be derived from eq. (22), given by

$$V^+ = 1 + Z_2 Z_1^*, \quad \text{and}$$
$$V^- = Z_1 + Z_2. \tag{C2}$$

With eqs (14) and (15) the propagated Green's functions $U^{\pm}$ are hence given by

$$U^- = R^{\cup}V^+ - V^- = Z_3 \frac{(1 - |Z_1|^2)(1 - |Z_2|^2)}{1 + Z_2 Z_1^* + Z_3 Z_2^* + Z_3 Z_1^*},$$
$$U^{+*} = -R^{\cup *}V^- + V^+ = \left[\frac{(1 - |Z_1|^2)(1 - |Z_2|^2)}{1 + Z_2 Z_1^* + Z_3 Z_2^* + Z_3 Z_1^*}\right]^*. \tag{C3}$$

Deconvolution of these two expressions yields a target reflection response [see eq. (17)]

$$R_{\text{Target}}^{\cup}(\omega) = \frac{U^-}{U^+} = Z_3, \tag{C4}$$

which is the primary reflection of the third interface void of the overburden's transmission effect [which are contained in the same primary in the input data in eq. (C1).

APPENDIX D: ALTERNATIVE MARCHENKO EQUATION MUTE, INITIAL CONDITION AND ITS CONSEQUENCES

Alternatively to the way we proceeded in Section 2.3 we can apply to both sides a different mute $\Phi_{-\varepsilon}$ (with $t_0 = -\varepsilon$), such that the overlap in eq. (20) is preserved instead of muted away. As a result the system of equations reduces to

$$V^+ - \Phi_{-\varepsilon}\left[R^{\cup *}\Phi_{-\varepsilon}\left[R^{\cup}V^+ d\boldsymbol{x}_0\right]d\boldsymbol{x}_0'\right] = U_e^{+*}, \tag{D1}$$

which can be solved by Neumann series expansion

$$V^+ = \sum_{k=0}^{\infty} \Omega_{k,-\varepsilon}\left[U_e^{+*}\right] \tag{D2}$$

with $\Omega_{k,-\varepsilon}[A] = \Phi_{-\varepsilon}[R^{\cup *}]\Omega_{\varepsilon}[R^{\cup}\Omega_{k-1,-\varepsilon}[A]d\boldsymbol{x}_0']$
and $\Omega_{0,-\varepsilon}[A] = A$.

The solution is now parametrized by the choice of U_e^{+*}, instead of V_e^+ (see eq. (22) for comparison).

Recall the definition from eq. (13)

$$U^{+*} = G^{+*}T_{d,DD}^+ d\boldsymbol{x}_i = \left(T_{d,MM}^{-*} + G_{\text{coda}}^{+*}\right)T_{d,DD}^+ d\boldsymbol{x}_i,$$

where $_{MM}$ and $_{DD}$ stand for monopole-monopole and dipole-dipole source and receiver pair respectively. In a broadband setting (i.e. $\varepsilon \to 0$) U_e^{+*} should be given by $T_{d,MM}^{-*}T_{d,DD}^+$, which in a simple (e.g. laterally invariant) case is just I_{DM} with a ray-parameter dependent (two-way direct transmission) amplitude, that is a non-isotropic zero-distance propagator. However, in more complex media T_d^+ (and hence U_e^+) can contain additional events (in the event of forward scattered acoustic waves or triplications). This problem can

be avoided by propagating eqs (3–4) with $T_{d,DD}^{-\mathrm{inv}}$, and changing the definition of U^+ to $U^+ = G^+ T_{d,DD}^{-\mathrm{inv}} \mathrm{d}\boldsymbol{x}_i$ such that $U_e^+ = I_{DM}$.

Similar to the discussion in Section 2.3, with band-limited $R^\cup$, the $\overline{U}_e^+$ term has a temporal extend and the best guess for it is $\overline{I}_{DM}$. This guess is correct provided that the $\overline{U}^+$ coda $\left(G_{\mathrm{coda}}^{+*} T_{d,DD}^+\right)$ is not contained in $\overline{U}_e^+$ (i.e. in the absence of SPIMs). From the medium point of view this is the same condition as discussed is Section 2.3, just for a different *choice* of t_0.

Following the approach from Section 2.5, we can again use the right-multiplicative freedom and define a different operator $B' \neq B$, such that

$$
\begin{aligned}
\tilde{V}'^+ &= V^+ B' \mathrm{d}\boldsymbol{x}_0 , \\
\tilde{V}'^- &= V^- B' \mathrm{d}\boldsymbol{x}_0 , \\
\tilde{U}'^- &= U^+ B' \mathrm{d}\boldsymbol{x}_0 , \\
\tilde{U}'^{+*} &= U^{+*} B' \mathrm{d}\boldsymbol{x}_0 ,
\end{aligned}
\tag{D3}
$$

and where we can then make the choice $\tilde{U}_e'^{+*} = \overline{I}_{DM}$ in

$$
\tilde{V}'^+ = \sum_{k=0}^{\infty} \Omega_{k,-\varepsilon} \left[\tilde{U}_e'^{+*} \right] .
\tag{D4}
$$

As argued towards the end of Section 2.5, in order to recover $\overline{V}^+ \overline{V}^{+\dagger}$ from $\tilde{V}'^{\pm}$ we use the energy conservation condition

$$
\left[\tilde{V}'^{+\dagger} \partial_{z,0} \tilde{V}'^+ + \tilde{V}'^{-\dagger} \partial_{z,0} \tilde{V}'^- \right] \rho_0^{-1} \mathrm{d}\boldsymbol{x}_0 \equiv \Sigma,
\tag{D5}
$$

where

$$
\Sigma = -\frac{i\omega^2}{4} \rho_i B'^\dagger T_d^{+\dagger} I_{MM} T_d^+ B' \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{x}_i' \mathrm{d}\boldsymbol{x}_0'' \mathrm{d}\boldsymbol{x}_0'''
\tag{D6}
$$

or

$$
\Sigma = -\frac{i\omega^2}{4} \rho_i B'^\dagger T_d^{-\mathrm{inv}\dagger} I_{MM} T_d^{-\mathrm{inv}} B' \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{x}_i' \mathrm{d}\boldsymbol{x}_0'' \mathrm{d}\boldsymbol{x}_0''' ,
\tag{D7}
$$

depending on the propagation operator used in the definition of V and U. As with the $t_0 = \varepsilon$ case we can subsequently invoke the minimum phase property to recover $\overline{V}^+$.

The challenge with the approach using $t_0 = -\varepsilon$ is that the solutions $\tilde{V}'^{\pm}$ can have rather large amplitudes (on account of many auto-correlations in the $|t_0| \leq \varepsilon$ window in $\tilde{V}'^+$), and these can cause round-off errors whose difference [on the left-hand side of eq. (D5)] could lead to numerically induced errors in Σ.

APPENDIX E: WRONG BAND-LIMITED SOLUTION FOR A HIGHLY SCATTERING OVERBURDEN

In this paper we have shown how using band-limited data when solving the Marchenko equation may result in failing to retrieve the correct focusing function (propagated or otherwise). Whereas the reciprocity theorems in eqs (3–4) are exact scattering relations, the band-limited, constrained Marchenko equations are only true under the conditions mentioned in the main body of the article: that $\tilde{V}^-$ does not have any events very close to t_0 (i.e. no extremely shallow reflectors) and that $\tilde{V}^-$ does not overlap with $\tilde{U}^-$. We have also pointed out and presented a solution to the problem of the unknown overlap between $\tilde{V}^+$ and $\tilde{U}^{+*}$.

There is however a third problem with band limitation, and it has to do with solving eq. (21) or eq. (D1) itself. The problem is best illustrated for strongly-scattering overburdens and a Neumann series-type solver (but it is not limited to it) where many terms in the expansion need to be calculated [see eqs (22) or (D2)]. The problem

will not result in convergence (or divergence) issues as observed by Dukalski & de Vos (2017), but will result in convergence to an *incorrect solution*.

We illustrate this with three simple, 1-D examples shown in Figs E1 (a), (f) and (k), with amplitude spectra of reflection responses given in (b), (g) and (l) respectively. Blue and red traces illustrate broadband and band-limited results, respectively. The first example features a simple two reflector overburden with an archetipical two-event focusing function [panel (d)]. The second example features the same overburden, but with increased reflectivities of the two interfaces (unrealistically for seismic applications), such that a greater number of iterations is required [see panel (h) vis-a-vis panel (c)] to converge to the solution. The result should be the same focusing function but with a greater amplitude coda [compare panels (d) and (i)]. We see that if broadband data is used (in combination with the wavelet-dressed first arrival field), the two solutions do defer by an expected amplitude difference of the coda event. However if band-limited data is used instead, the solution not only converges slower, but it also does not converge to a true solution.

In order to show that this could be a relevant problem in a setting such as the Middle East (i.e. for strongly scattering overburdens), we use the third example with *many* interfaces and with realistically strong (and alternating) reflection coefficients, yet sufficiently spaced to avoid short-period scattering issues (see a wavelet used for comparison in purple). Here we observe the same qualitative features as with the simple (unrealistically) high-impedance two-interface model.

We believe that the source of the problem has to do with the convolve-mute-correlate-mute sequence $\Phi_{\pm\varepsilon} \left[\check{R}^{\cup *} \Phi_{\pm\varepsilon} \left[\check{R}^\cup \cdots \right] \right]$, combined with the fact that the amplitude spectrum of the reflection response $\left| \check{R}^\cup(\omega) \right| \approx 1$ for a large number of frequencies. The tempral mute is equivalent to a convolution in the frequency domain, which can wash in the incorrectly scaled amplitudes at the slopes of $\check{R}^\cup$ towards the correctly scaled frequency band. This would suggest that some form of regularization step inside the convolve-mute-correlate-mute sequence could help alleviate the problem.

Spotting that one might be dealing with this problem could be a challenge. However, we have observed that in many (but not necessarily all) cases, when the incorrectly converged solutions are used to evaluate the energy conservation [see panels (e), (j) and (o)] the amplitude spectrum of the energy conservation relation can carry uniquely abrupt and spiky features.

APPENDIX F: CONVERGENCE

In the course of this work we have noticed that understanding the implications of convergence of the Marchenko equation solver [e.g. the series expansions in eqs (22) or (D2)] is very important. There are three aspects that we will outline here: theoretical soundness, interpretation of consecutive terms in the expansion and implications for field data studies.

On the theoretical front, the augmented Marchenko algorithm presented here, in particular the statements surrounding the energy conservation condition, only hold when the Marchenko equation solver converged completely. They do not hold on a term-by-term basis. In theory that means that we need to take the $k \to \infty$-limit in the Neumann series expansion, however for many practical cases fewer than 20 iterations tend to be sufficient.

It was proposed by van der Neut & Wapenaar (2016) to carry out internal demultiple by expanding the conventional Marchenko
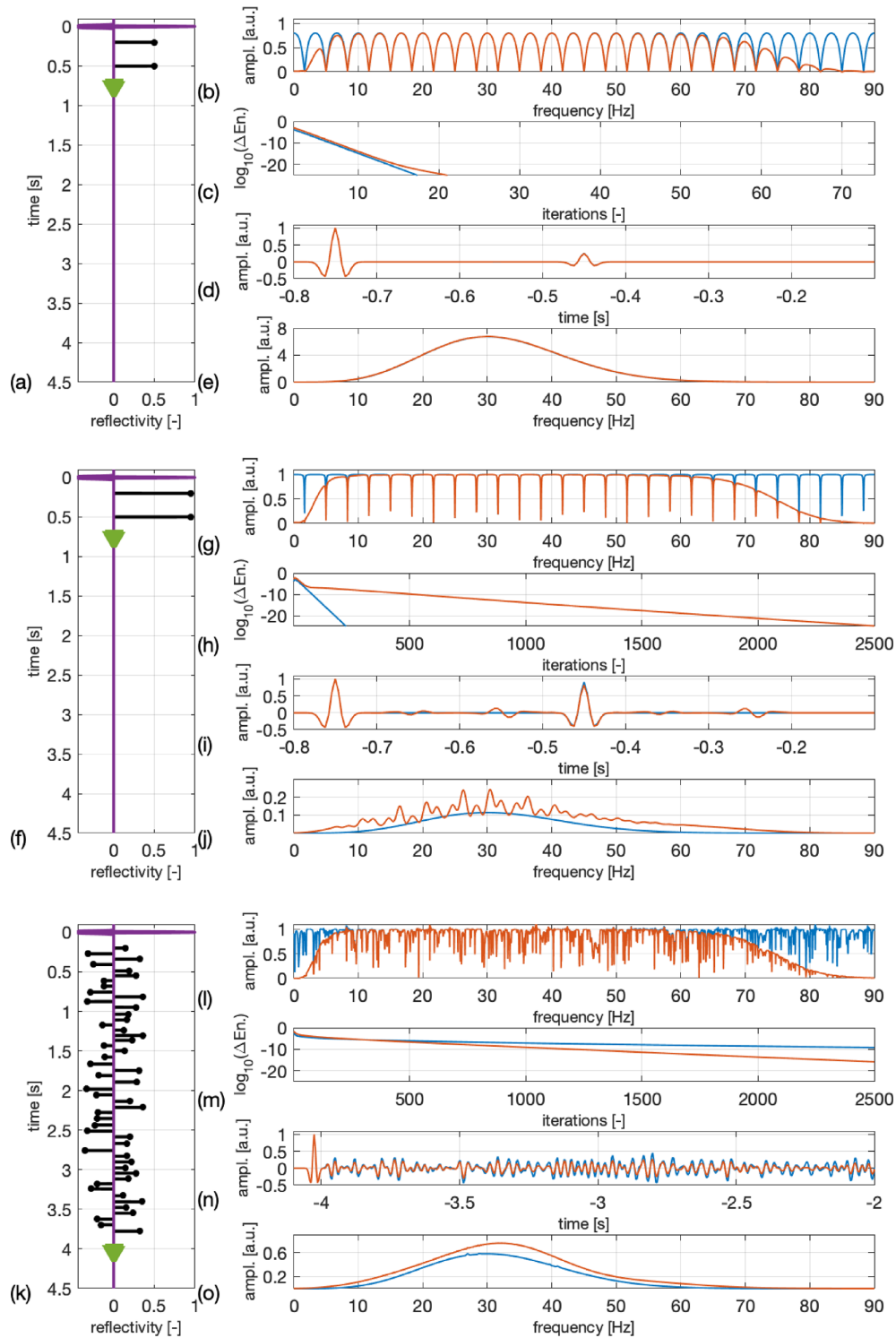
Figure E1. Illustration of 1-D models for which the band-limited iterative Marchenko scheme succeeds or fails to retrieve the correct solution. (a) two-reflector model with interfaces at 0.2 s and 0.5 s and reflectivities of 0.5 (black lines), the focusing position at 0.75 s (green triangle) and the input 30 Hz Ricker wavelet (purple); (b) Amplitude spectrum of the full-bandwidth reflection data $R^{\cup}$(blue) and the band-limited equivalent $\check{R}^{\cup}$(red); (c) Convergence illustrated in terms of the energy that is contributed to the focusing function with consecutive iterations; (d) Focusing functions retrieved using full-bandwidth data (blue) and band-limited data (red); (e) amplitude spectra energy conservation (d). Panels (f)-(j) are equivalent for a two-reflector model with reflectivities 0.95, panels (k)-(o) for a 50-reflector model with randomly placed (but well-separated) interfaces of reflectivities between ±0.36. ampl. = amplitude, En.= Energy, a.u. = arbitrary unit.

method (see Fig. 4) in series. The authors suggested to collect only the first two terms of the Neumann series and append them with a

(time-dependent) matching filter, which is supposed to fix the amplitude mistakes emerging from the series truncation. This approach
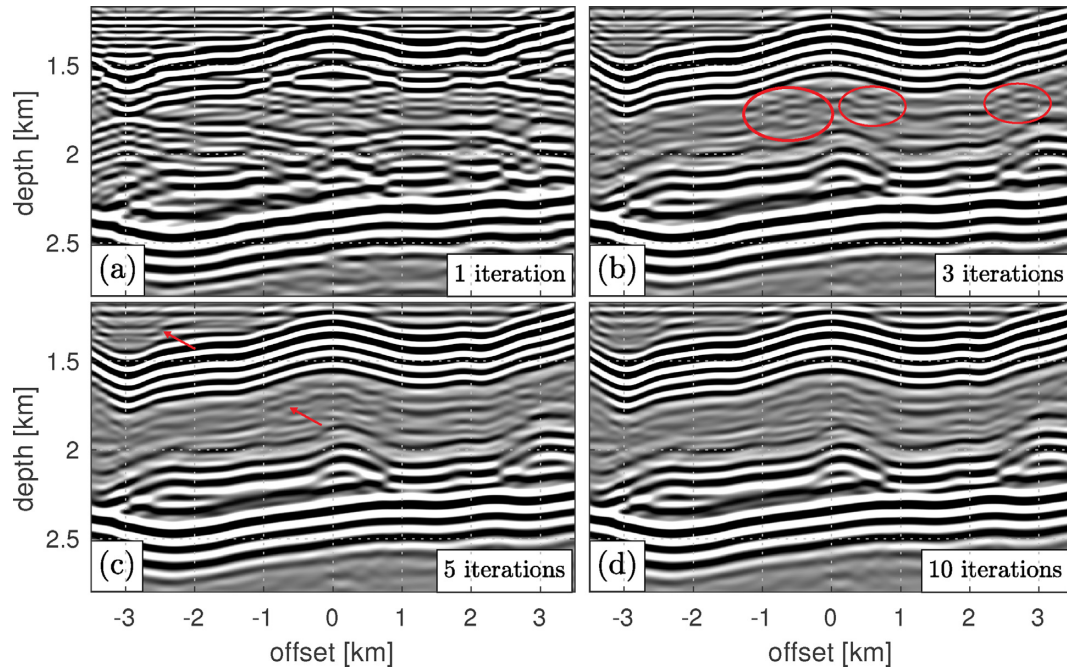
Figure F1. Migration images derived from conventional Marchenko results for different (and insufficient) numbers of iterations in eq. (21). Red circles and arrows are used to denote artefacts.
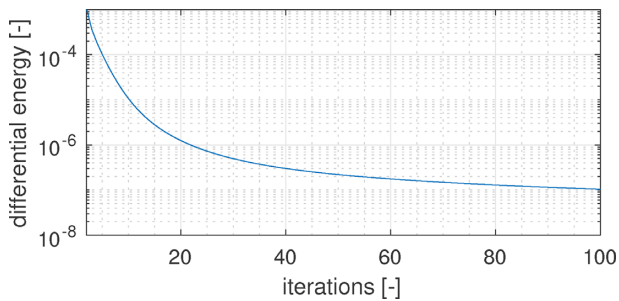

Figure F2. Convergence of the Marchenko scheme. Here the differential energy is defined as the squared difference between the $\tilde{v}^{+}$-wavefields of two subsequent Marchenko iterations.

naturally would no longer work in the presence of SPIMs, as such a series would not be possible due to overlapping events. Any methods alluding to using or assuming the ability to extract particular events of V or U are likely to fail in the setting where internal multiples are a significant challenge. Moreover, in complex geological settings the first few terms of the Neumann series contain complex interference patters, and various 'counter-terms' from these early terms are collectively removed by higher order terms in the expansion. Put differently, in order to remove 'erroneous' contributions from the

lower order terms the higher order terms *must* be included in the series. Otherwise, the early counter-terms may introduce significant amplitude and phase distortions to the solutions, which hamper a matching filter approach like the one proposed by van der Neut & Wapenaar (2016). The impact of using too few iterations in the Neumann series is illustrated in Fig. F1.

When trying to apply Marchenko-equation based methods to field data a correct data scaling is a great practical challenge: if the data is 'underscaled' ('overscaled'), then we will see that the series will converge too quickly (too slowly, or even diverge). Provided that the data set is off by a single scalar, the rate at which the series in eq. (22) converges contains valuable information. Suppose that some knowledge about the geological setting is available, for example access to a well log. Thomsen *et al.* (2017) suggested to improve scaling by matching a VSP signal with the Green's functions G. Our approach relies on comparing convergence rates that can be computed based on a 1-D synthetic model from that well log, and the scaling of the data is subsequently adjusted to match its convergence rate to that of the synthetic example. Said synthetic can be obtained either from a well log, or from regional geological knowledge. A good reference can be found in Fig. F2, where one sees that the *L*2 norm of the update (differential energy) falls to below 1 ppm after 20 iterations – perhaps a good general rule of thumb in the Middle East setting.