Data-driven wavefield reconstruction and focusing for laterally-varying velocity models, spatially-extended virtual sources, and inaccurate direct arrivals

Filippo Broggini\textsuperscript{1}, Roel Snieder\textsuperscript{1} & Kees Wapenaar\textsuperscript{2}

\textsuperscript{1} Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401, USA
\textsuperscript{2} Department of Geotechnology, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

ABSTRACT
Seismic interferometry allows one to create a virtual source inside a medium, assuming a receiver is present at the position of the virtual source. We discuss a method that creates a virtual source inside a medium from reflection data measured at the surface, without needing a receiver inside the medium and, hence, presenting and advantage over seismic interferometry. In addition to the reflection data, an estimate of the direct arrivals is required. However, no information about the medium is needed. We analyze the proposed method for a simple configuration using physical arguments based on the stationary-phase method and show that the retrieved virtual-source response correctly contains the multiples due to the inhomogeneous medium. We illustrate the method with numerical examples in lossless acoustic media with laterally-varying velocity and density and take into consideration finite acquisition aperture and spatially-extended virtual sources. We examine the reconstructed wavefield when a macro model is used to estimate the direct arrivals.

Key words: focusing, wavefield reconstruction, stationary-phase, inverse scattering, finite aperture, virtual source, seismic interferometry, multiples

1 INTRODUCTION
We propose and discuss a new approach to reconstruct the response to a virtual source inside a medium. Controlled-source interferometric methods (Schuster, 2009; Bakulin and Calvert, 2006) allow us to retrieve such a response without the need to know the medium parameters, but these methods require a receiver at the location of the virtual source in the subsurface and assume the medium is surrounded by sources. Our new approach removes the constraint of having a receiver at the virtual source location and is based on an extension of the 1D theory proposed previously by Broggini et al. (2011, 2012); Broggini and Snieder (2012). They show that, given the reflection response of a 1D layered medium, it is possible to reconstruct the response to a virtual source inside the medium, without having a receiver at the virtual source location and without knowing the medium. The insights for this method were based on the work of Rose (2002). An essential element of this approach is to use a complicated incident wave that is designed to collapse onto a point inside the medium at a specified location and time.

Wapenaar et al. (2011a) made a first attempt to generalize to 3D media the method of Broggini et al. (2011, 2012) and Broggini and Snieder (2012). They used physical arguments to propose an iterative scheme that transforms the reflection response of a three-dimensional medium into the response to a virtual source inside the unknown medium. Apart from the reflection data measured at the surface, the proposed method also requires an estimate of the direct arrivals between the virtual source location and the acquisition surface. These arrivals are a key element of the method because they specify the location of the virtual source in the subsurface. For this reason, the proposed method is not fully model-independent. However, a model that relates the direct arrival to a virtual source position is simpler than a model that correctly handles the multiples.
In the proposed method, the reflection data contributes to the multiple-scattering part of the virtual-source response.

As in seismic interferometry (Wapenaar et al., 2005; Curtis et al., 2006; Schuster, 2009), our goal is to retrieve the response to a virtual source inside an unknown medium, removing the imprint of a complex subsurface. This is helpful in situations where waves have traversed a strongly inhomogeneous overburden (e.g., in subsalt imaging, Sava and Biondi, 2004). In this paper, we demonstrate that the requirement of having an actual receiver inside the medium can be circumvented, presenting an advantage over seismic interferometry.

In the next section, following the work of Wapenaar et al. (2012), we analyze the iterative scheme for a simple 2D configuration with three parallel dipping reflectors characterized by variable-density and constant-velocity. We use physical arguments based on the stationary-phase method to show that the method converges and allows for the reconstruction of the wavefield originating from the virtual source location. In the subsequent section, we present numerical examples in lossless acoustic media with variable velocity and density for more complex configurations: a layered medium with varying density and velocity and a model with a syncline reflector. We discuss the influence of errors in the estimate of the first arrivals on the reconstructed wavefield. Such errors arise if a macro model (a routine product of velocity analysis) is used to compute the first arriving waveforms when such data are not available with other approaches, e.g., check shots or microseismic events. We then examine how the finite acquisition aperture and the limitations of the modeling code affect the results. While seismic interferometry usually deals with virtual point sources, this method allows us also to reconstruct the response to a spatially-extended virtual source. This feature has a potential application in beam migration (Gray et al., 2009). Finally, we show that the proposed scheme implicitly reconstructs the incident field that focuses the wavefield in time and space at the virtual source location.

2 STATIONARY-PHASE ANALYSIS

We discuss the proposed iterative scheme for a simple two-dimensional configuration. We use a geometrical approach to the method of stationary phase to solve the Rayleigh-like integrals, which yield the reflected response to an arbitrary incident field. We explain each step of the iterative procedure and emphasize the physical arguments that support our expectations for the method to converge to the virtual-source response.

2.1 Configuration

We consider a configuration of three parallel dipping reflectors in a lossless, constant-velocity, variable-density medium (Figure 1). We choose a constant velocity medium because the response obeys simple analytical expressions. The proposed iterative scheme is, however, not restricted to constant velocity media as we show in the next sections. We denote spatial coordinates as \( x = (x, z) \). The acquisition surface is located at \( z = 0 \) m and is transparent (i.e., the upper half-space has the same medium parameters as the first layer). The first dipping reflector obeys the relation \( z = z_1 - ax \) with \( z_1 = 2000 \) m and \( a = 1/3 \). The black dot denotes the position of the virtual source, with coordinates \( x_{VS} = (x_{VS}, z_{VS}) = (475, 3425) \) m. The second and third reflector are parallel to the first one, so that all mirror images of the virtual source lie on a line perpendicular to the reflectors. This line obeys the relation \( z = z_1 + x/a \). The first, second, and third reflectors cross this line at \( x_1 = (z_1, z_1), x_2 = (x_2, z_2), \) and \( x_3 = (x_3, z_3), \) respectively. The relative position of the reflectors is chosen such that the Euclidean distance between the pairs \( (x_1, x_2) \) and \( (x_3, x_2) \) is 1000 m. The velocity of the medium is constant and equal to \( c = 2000 \) m/s. The densities in the four layers are \( \rho_1 = \rho_3 = 1000 \) kg/m\(^3\), \( \rho_2 = 5000 \) kg/m\(^3\), and \( \rho_4 = 4000 \) kg/m\(^3\), respectively. The reflection coefficients for downgoing waves at the three interfaces are \( r_n = (\rho_{n+1} - \rho_n)/(\rho_{n+1} + \rho_n) \), where \( n = 1, 2, 3 \) denotes the layer. The reflection coefficients for upgoing waves are \(-r_n\). The transmission coefficients for downgoing (+) and upgoing (−) waves are \( t_n^{\pm} = 1 \pm r_n \). Since the velocity is constant in this particular configuration, the reflection and transmission coefficients hold not only for normal incidence but for all the angles of incidence. Note that the large contrast between the density of the layers causes strong multiple reflections.

2.2 Primary arrivals

We introduce the Green’s function \( G(x, x_s, t) \) as a solution to the wave equation \( LG = -\rho \delta(x - x_s) \partial_t \partial_t^\circ \), with \( L = \rho \nabla \cdot (\rho^{-1} \nabla) - c^{-2} \partial_t^2 \). Defined in this way, the Green’s function is the response to an impulsive point source of volume injection rate at \( x_s \) (de Hoop, 1995). Using the Fourier convention \( \hat{F}(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-j\omega t) dt \), the frequency domain Green’s function \( \hat{G}(x, x_s, \omega) \) obeys the equation \( \hat{L} \hat{G} = -j\omega \rho \delta(x - x_s) \), with \( L = \rho \nabla \cdot (\rho^{-1} \nabla) + \omega^2/c^2 \). Here \( j \) is the imaginary unit and \( \omega \) denotes the frequency domain. We write \( \hat{G} = \hat{G}^d + \hat{G}^s \), where superscripts \( d \) and \( s \) stand for direct and scattered waves, respectively. As mentioned in the introduction, we need an estimate of the direct arrivals. For the constant velocity model of Figure 1, the high-frequency approximation of the Fourier transform of the direct Green’s function \( \hat{G}^d(x, x_{VS}, t) \)
Data-driven wavefield reconstruction and focusing 185

2.3 Reflection response

To retrieve the virtual-source response \(G(x, x_{VS}, t)\), we need the reflection response at the surface \(R(x_R, x_S, t) + s(t)\) in addition to an estimate of the direct arrivals. We assume that the acquisition surface is transparent, hence the reflection response does not include any surface-related multiples. For this purpose, \(R(x_R, x_S, t) + s(t)\) can be obtained from reflection data measured at \(z = 0\) after surface-related multiple elimination (Verschuur et al., 1992). Following Wapenaar and Berkhout (1993) (equation 15a), the reflection response can be derived from a Rayleigh-type integral:

\[
\hat{p}^{-}(x_R, \omega) = \frac{2}{j \omega p_1} \int_{-\infty}^{\infty} \left[ \partial G^s(x_R, x, \omega) \hat{p}^+(x, \omega) \right] d\omega,
\]

where \(\hat{p}^+\) and \(\hat{p}^-\) are the Fourier transform of the downgoing and upgoing wavefields, respectively. Hence, in the frequency domain, we define the reflection response in terms of the scattered Green's function \(\hat{G}^s\) via

\[
\hat{R}(x_R, x_S, \omega) \delta(x) = \frac{2}{j \omega p_1} \hat{G}^s(x_R, x_S, \omega) \delta(\omega),
\]

for \(z_R = z_S = 0\) and after multiplying both sides by the spectrum of the source wavelet \(s(t)\). Note that, apart from a factor of \(-2\), \(\hat{R}(x_R, x_S, \omega) \delta(x)\) represents the pressure at \(x_R\) due to a force source at \(x_S\) or, via reciprocity, the particle velocity at \(x_S\) due to a point source of volume injection rate \(\frac{\partial}{\partial t} s(t)\) at \(x_R\) (Berkhout, 1982).

2.4 Initiating the iterative process

We define the initial incident downgoing wavefield at \(z = 0\) as the time-reversed version of the direct arrivals at the recording surface excited by the virtual source in Figure 2. Hence, the initial wavefield is \(p_0^d(x, t) = G^d(x, x_{VS}, -t) + s(t)\) and is shown in Figure 3 with the label 1. The subscript 0 of \(p_0^d(x, t)\) indicates the initial wavefield (or the 0th iteration). In Figure 3, we also define two traveltime curves indicated by the solid black lines. The upper curve follows directly after the initial incident wavefield \(p_0^d(x, t)\) and the lower curve is defined as the time-reversal of the upper curve (hence, it is placed just before the direct arrival of Figure 2). These two curves allow us to define a window function

\[
w(x, t) = 1 \quad \text{between the solid black lines of Figure 3,}
\]

\[
w(x, t) = 0 \quad \text{elsewhere.}
\]

This window function is a key component of the iterative scheme.

The reflected upgoing wavefield \(p_0^u(x, t)\) is obtained by convolving the downgoing incident wavefield \(p_0^d(x, t)\) with the deconvolved reflection response and integrating...
correspond to the events of 2000. The reflection response only until 2.5 s. The solid black lines denote the time of the direct arrivals and its time-reversed counterpart. These lines are repeated in the subsequent figures.

over the source positions:

\[
p_0^{-}(x_R, t) = \int_{-\infty}^{\infty} \left[ R(x_R, x, t) * p_0^+(x, t) \right]_{z=0} dx, \quad (5)
\]

for \( z_R = 0 \). Equation (5) is the time-domain version of the Rayleigh integral described by equation (2). We discuss and solve this integral with geometrical arguments based on the method of stationary phase (a detailed mathematical derivation is given by Wapenaar et al., 2012). Figure 4 shows a number of stationary rays for different receiver positions. These rays are said to be stationary because the rays of the incident field (converging in \( x_{VS} \)) and of the reflection response (for the first reflector) have the same direction (see the appendix in Wapenaar et al., 2010). With simple geometrical arguments, it follows that these rays cross each other at the mirror image of the virtual source with respect to the first reflector, i.e., at \( x_{VS}^{(1)} \). The traveltimes of the convolution product are given by the length of the rays from \( x_{VS}^{(1)} \) to the surface, divided by the velocity. Hence, it is as if the response of the first reflector to the initial incident field originates from a source at \( x_{VS}^{(1)} \). This response is thus equal to \( r_{1}G^{d}(x_R, x_{VS}^{(1)}, t) \ast s(t) \) and is shown as the event with label 2 in Figure 3. Following similar stationary-phase arguments, the response of the second reflector to the initial downdogging field apparently originates from a mirror image of the virtual source with respect to the second reflector, i.e., at \( x_{VS}^{(2)} \). This response is equal to \( r_{2}G^{d}(x_R, x_{VS}^{(2)}, t) \ast s(t) \) (see the event with label 3 in Figure 3). However, the multiple reflected responses to the initial incident field also apparently originate from mirror images of the virtual source, all located along the line \( z = z_1 + x/a \) (see Figure 1). We derive these responses with the method of stationary phase, hence they are free of artifacts. However, later in this manuscript, we show numerical examples and we discuss finite aperture artifacts, tapering effects, error in the estimate of the direct arrivals, etc.

2.5 Iterative process

We now discuss an iterative scheme, which uses the \((k-1)\)th iteration of the reflection response \( p_{k-1}^{-}(x, t) \) to create the \( k \)th iteration of the incident field \( p_k^{+}(x, t) \). The objective is to iteratively update the incident field in such a way that, within the upper and lower solid black lines shown in Figure 3, the field is anti-symmetric in time. The meaning of this criterion will be evident in the next section, where we show how to reconstruct \( G(x, x_{VS}, t) \). The method requires a combination of time reversal and windowing and the \( k \)th iteration of the incident field is defined by

\[
p_k^{-}(x, t) = p_0^{-}(x, t) - w(x, t)p_{k-1}^{-}(x, -t), \quad \text{for } x \text{ at } z = 0,
\]

(6)

where the time window \( w(x, t) \) is defined by equation (4). The reflection response is then obtained using equation (5), which we rewrite here as

\[
p_k^{+}(x_R, t) = \int_{-\infty}^{\infty} \left[ R(x_R, x, t) * p_k^{-}(x, t) \right]_{z=0} dx, \quad (7)
\]

for \( x \) and \( x_R \) at \( z = 0 \). The first and second iteration of the incident and reflected field are shown in Figures 5 and 6, respectively. The events of \( p_k^{-}(x, t) \) labeled 2, 3, 4 in Figure 6 correspond to the events of \( p_1^{-}(x, t) \) labeled 4, 5, 6 in Figure 5 (multiplied by \(-1\)).

For this particular configuration, the \( k \)th iteration of the incident field (for \( k > 2 \)) is similar to \( p_2^{-}(x, t) \) and is composed by four events, as shown in Figure 6 for \( t < 0 \). The events labeled 1 and 4 remain unchanged in the iterative process. The other two events (labeled 2 and 3) correspond to \(-A_k^{(2)}G^{d}(x_R, x_{VS(S)}, t) \ast s(t) \) and \(-A_k^{(3)}G^{d}(x_R, x_{VS(S)}, -t) \ast s(t) \), respectively. The coefficient \( A_k^{(2)} \) varies at each iteration and is equal to the...
and shows the thirtieth iteration and, within the time window, all events inside the time window cancel each other.

We show the reflection response only until 2.5 s.

The numbers identify the first seven events in the total field.

We show the reflection response only until 2.5 s.

The partial sum of the geometric series \( a + ax + ax^2 + ax^3 + ax^4 + \cdots + ax^n \) where \( a = t_1^r r_2 t_2^r t_1 \) and \( x = r_2^t \). The sum of the series converges because \( x < 1 \) and yields

\[
\sum_{k=0}^{\infty} ax^k = \frac{a}{1 - x} = r_2 t_2^t t_1^r, \tag{8}
\]

where we used that \( 1 - r_2^t = t_1^r t_1^t \). The coefficient \( A_k^{(3)} \) of the third event (labeled 3 in Figure 6) is equal to \( -r_1 A_k^{(2)} \) and, hence, it converges to \( -r_1 r_2 t_2^r t_1^r \). Figure 7 shows the thirtieth iteration and, within the time window \( w(x, t) \), the wavefield is antisymmetric in time. This is the result we predicted when we described the iterative method. Note that the antisymmetry was the design criterion for the iterative scheme. Wapenaar et al. (2012) show that, for their simple configuration with one dipping layer, convergence is reached after one iteration.

2.6 Wavefield reconstruction from the virtual source

After showing that the method converges to the desired result, we define \( p_k(x, t) \) as the superposition of the \( k \)th version of the incident and reflected wavefields: \( p_k(x, t) = p_k^i(x, t) + p_k^r(x, t) \). Figures 3, 5, 6 and 7 show \( p_k(x, t) \) for \( k = 0, 1, 2, \) and 30, respectively. For brevity, we define \( p(x, t) = p_{30}(x, t) \).

We remind the reader that, within the solid black lines, the total field at \( z = 0 \) is antisymmetric in time and this particular feature was the design criterion for the iterative scheme. Consequently, if we stack the total field and its time-reversed version, i.e., \( p(x, t) + p(x, -t) \), all events inside the time window cancel each other.

![Figure 5](image1.png)

Figure 5. First iteration of the incident wavefield \((t < 0)\) and its reflection response \((t > 0)\), both measured at \( z = 0 \). The numbers identify the first six events in the total field. We show the reflection response only until 2.5 s.

![Figure 6](image2.png)

Figure 6. Second iteration of the incident wavefield \((t < 0)\) and its reflection response \((t > 0)\), both measured at \( z = 0 \). The numbers identify the first seven events in the total field. We show the reflection response only until 2.5 s.

![Figure 7](image3.png)

Figure 7. Thirtieth iteration of the incident wavefield \((t < 0)\) and its reflection response \((t > 0)\), both measured at \( z = 0 \). Within the solid black lines, the total field is antisymmetric in time and this particular feature was the design criterion for the iterative scheme. The method has converged to the final result.

![Figure 8](image4.png)

Figure 8. Thirtieth iteration. Superposition of the total field and its time-reversed version after the method has converged. Here, unlike the previous figures, we show the wavefield for the time interval \(-4 < t < 4 \) s.
as shown in Figure 8. Note that \( p(x, t) + p(x, -t) \) also obeys the wave equation because we consider a lossless medium. The causal part of this superposition corresponds to \( p^+(x, t) + p^+(x, -t) \) and the anti-causal part is equal to \( p^-(x, t) + p^-(x, -t) \), as shown in Figure 8 for \( t < 0 \) and \( t > 0 \), respectively. From a physical point of view, time-reversal changes the propagation direction. Hence, it follows that the causal part propagates upward at \( z = 0 \) and the anti-causal part propagates downward at \( z = 0 \). The first event of the causal part of Figure 8 has the same arrival time as the direct arrival and matches the 2000.

For the total wavefield, we obtain

\[
G(x, x_{VS}, t) = \sum_{r} \rho^{(r)} G^{(r)}(x, x_{VS}(t), t) \ast s(t),
\]

with the virtual source position and its mirror images shown in Figure 1. For the configuration of Figure 1, expression (9) is proportional to the wavefield \( G(x, x_{VS}, t) \ast s(t) \) originated from the virtual source and recorded at the surface (with \( t_1^2 t_2^2 \) as the coefficient of proportionality). The directly modeled response to the virtual source is shown in Figure 9 and matches the causal part of the field shown in Figure 8. This is also illustrated in Figure 10, where we superposed the traces \( t_1^2 t_2^2 G(x_{R0}, x_{VS}, t) \ast s(t) \) and \( p^-(x_{R0}, t) + p^+(x_{R0}, -t) \) where \( x_{R0} = (0, 0) \).

For the total wavefield, we obtain

\[
p(x, t) + p(x, -t) = t_1^2 t_2^2 G_h(x, x_{VS}, t) \ast s(t),
\]

where \( G_h(x, x_{VS}, t) = G(x, x_{VS}, t) + G(x, x_{VS}, -t) \).

Note that \( G(x, x_{VS}, t) \) obeys the same wave equation as \( G(x_{R}, x_{VS}, t) \), i.e.,

\[
\frac{1}{c^2} \frac{\partial^2}{\partial t^2} G(x_{R}, x_{VS}, t) = -\rho_0 \delta(x - x_{VS}) \frac{\partial^2}{\partial t^2} \delta(t) = \rho_0 \delta(x - x_{VS}) \frac{\partial^2}{\partial t^2} \delta(t).
\]

Hence, \( G_h \) obeys the homogeneous equation \( L_{G_h} = -\rho_0 \delta(x - x_{VS}) \frac{\partial^2}{\partial t^2} + \rho_0 \delta(x - x_{VS}) \frac{\partial^2}{\partial t^2} \delta(t) \). This is in agreement with the fact that \( p(x, t) + p(x, -t) \) has been constructed without introducing a singularity (i.e., a real source) at \( x_{VS} \). \( G_h \) is called the homogeneous Green’s function, after Porter (1970) and Oristaglio (1989) (but note that these authors take the difference instead of the sum of the causal and acausal Green’s functions because of a different definition of the source in the wave equation).

Finally, we introduce a visual way to show the convergence of the proposed method, compute the energy of \( p_h(x, t) + p_h(x, -t) \) within the time window \( w(x, t) \), and plot it versus the number of iterations:

\[
\text{Energy}(k) = \sum_{t} \sum_{x} w(x, t) [p_h(x, t) + p_h(x, -t)]^2.
\]

The quantity is shown in Figure 11 and the energy inside the time window clearly converges to zero, as confirmed by Figure 8. Note that this procedure is expected to converge because in each iteration the reflected energy is smaller than the incident energy. We consider the proposed method as a correction scheme that minimizes the energy inside the time window \( w(x, t) \).

3 NUMERICAL EXAMPLES

In this section, we show numerical examples for two-dimensional lossless acoustic media with variable velocity and density. We consider two different configurations: two parallel dipping reflectors and a syncline.
model. We require that the wavefield focuses at a specific location, and hence the proposed method cannot be totally independent of knowledge about the medium. The iterative scheme requires the reflection response of the medium measured at the surface, complemented with independent information about the primary arrivals originated from the focusing location, to focus the acoustic wavefield inside the medium. The primary arrival wavefront can be estimated or measured in various ways: by forward modeling using a macro model, directly from the data by the Common Focusing Point method (CFP) (Thorbecke, 1997) when the virtual source is located at an interface, from micro-seismic events (Artman et al., 2010), or from borehole check shots. We consider a transparent acquisition surface (i.e., the upper half-space has the same medium parameters as the first layer), so that the data do not include surface-related multiples. In Section 2, we used wiggle plots because the response obeys analytical expressions and hence these kinds of plots allowed us to clearly show the waveform of the events. Here, we use raster plots to display the wavefields because they are the standard visualization method used to show seismic images.

### 3.1 Two dipping reflectors

We first consider a configuration of two parallel dipping reflectors, whose velocity is shown in Figure 12a. The density profile (not shown) has a similar behavior as the velocity profile and the densities of the three layers are \( \rho_1 = 1000 \text{ kg/m}^3 \), \( \rho_2 = 5000 \text{ kg/m}^3 \), and \( \rho_3 = 7000 \text{ kg/m}^3 \), from the top to bottom layer. The contrast between the densities is responsible for strong multiples. As in the previous section, we first consider a virtual point source and, with the next example, we discuss the iterative method for a spatially-extended source. To start the iterative scheme, we use a macro model to compute the direct arrivals originating from the virtual source indicated by the black dot in Figure 12b. Here, we use a smooth version of the true velocity model, as the macro model, as shown in Figure 12b. Smooth models are used in standard practice to test the validity of imaging techniques and are a product of velocity analysis. The initial incident field \( p_0^\omega(x, t) \) is displayed in Figure 13a. Following the procedure explained in Section 2, we inject \( p_0^\omega(x, t) \) into the actual medium to compute the initial response \( p_0^\omega(x, t) \). This can be accomplished by convolving the incident field with the reflection responses measured at the surface (deconvolved for the source wavelet \( s(t) \)) and summing over the sources. Due to limited aperture, we apply a spatial tapering to the incident field to avoid edge-diffractions. Then we apply the window function \( w(x, t) \) of equation 4 and construct the first iteration of the updated incident wavefield \( p_1^\omega(x, t) \) (see Figure 13b). After computing \( p_1(x, t) \), we build the field \( p_1(x, t) + p_1(x, -t) \) to reconstruct the response originating from the virtual source location. The causal part of this field is shown in Figure 14a. Label (1) indicates an incomplete cancellation of the events inside the window \( w(x, t) \). This is due to numerical limitations of the modeling code and not to a lack of convergence: it is caused by numerical dispersion and the influence of spatial-tapering on the initial incident
field. Label (2) points to the imprint of finite acquisition aperture on the amplitude of the reconstructed events. For this particular configuration, we achieve the final result after one iteration because further iterations will not cause any changes in the incident field. The amplitude of the data shown in Figure 14 are clipped to 70% of the maximum amplitude and this enhances the features indicated by the labels (1) and (2). The comparison between the two panels of Figures 14 shows that it is possible to reconstruct the full response to a virtual source inside the medium, including all internal multiples, only from it first arrivals and the reflection response of the medium measured at the surface. This is fascinating because it allows one to obtain the same virtual source response as with seismic interferometry (including all scattering effects) with a starting smooth model only, but without the need for a receiver at the virtual source location.

This proposed scheme also implicitly constructs the incident field that collapses the wavefield onto the virtual source inside the medium and such a field corresponds to the last iteration of \( p_k^1 (x, t) \). To show the physics of the focusing process, we simulate the propagation of the updated incident field \( p_1^1 (x, t) \) inside the actual medium. Figure 15 displays six snapshots extracted from this simulation. Panel d clearly shows that the wavefield collapses to a focus at the location indicated by the black dot in Figure 12a; whereas the other events cancel in this panel.

We repeat the same steps for the same velocity model but we now use a different virtual source, i.e., two planar parallel line sources. The left square inset (bottom right corner) of Figure 12a shows the shape of the extended virtual source considered in this second example. The two planar and parallel line sources replace the point source used in the first example. Figure 16 shows the first two iterations of the incident field. The same considerations we formulated for the example with the point source are also valid for this different spatially-extended virtual source. Also, this example reconstructs the full response to the virtual sources inside the medium starting from the direct arrivals obtained with a macro velocity model, as shown in Figure 17. In Figure 18, we show the snapshots extracted from two simulations using the two different spatially-extended sources, shown in the square insets (bottom right corner) of Figure 12a. In both cases, the proposed method focuses the wavefield at the virtual source location.

### 3.2 Configuration with a syncline

In this section, we examine a configuration with a syncline reflector, whose velocity is shown in Figure 19a. The density profile (not shown) has a similar behavior and the densities of the three layers are \( \rho_1 = 6500 \text{ kg/m}^3 \), \( \rho_2 = 1000 \text{ kg/m}^3 \), and \( \rho_3 = 7000 \text{ kg/m}^3 \), from
the top to bottom layer. The wavefield associated with this model is more complicated than the ones shown in Section 3.1; in fact the direct arriving wavefront contains a triplication. To start the iterative scheme, we compute the direct arrivals originating from the virtual source using the macro model of Figure 19b. This is a smooth version of the velocity model in Figure 19a. The initial incident field $p_+^0(x,t)$ is displayed in Figure 20a. Note that, due to the smoothing, the triplications are almost not present in this field (i.e., the time-reversed version of the direct arrivals). As in Section 3.1, we compute the initial response $p_0^i(x,t)$ injecting $p_+^0(x,t)$ into the actual medium, apply the window function $w(x,t)$, and construct the first iteration of the updated incident wavefield $p_1^i(x,t)$ (see Figure 20b). This field is more complex than the field associated with the model considered in the previous section.

We form the field $p_1(x,t)+p_1(x,-t)$ to reconstruct the response originating from the virtual source location. The causal part of this field is shown in Figure 21a. Labels (1) and (2) indicate an incomplete cancellation of events. As in the previous section, this is due to numerical limitations of the modeling code: numerical dispersion and the influence of spatial-tapering on the initial incident field. Also, for this configuration, we achieve the final result after one iteration. As in the previous examples, the amplitude of the data shown in Figure 21 are clipped to 70% of the maximum amplitude and this enhances the features indicated by the labels (1) and (2). The comparison between the two panels of Figures 21 shows that it is possible to reconstruct the full response to a virtual source inside the medium, including all multiples, using the direct arrivals computed using a smooth model.

We simulate the propagation of the updated incident field $p_1^i(x,t)$ inside the medium of Figure 19a. Figure 22 displays six snapshots extracted from this simulation. Panel d shows that the wavefield collapses to a focus at the location indicated by the black dot in Figure 19a. We note that the approximate direct arrivals (computed with the smooth model) used to start the iterative process causes an imperfect virtual source as indicated by the artifacts around the virtual source. We see that the complex wavefield propagating inside the syncline (panels a,b,c,e, and f) annihilates at the focusing time as shown in panel d.

Finally, we repeat similar steps for the same velocity model but we now use a different virtual source, i.e., a $\wedge$-shaped source (as in the previous section). The square inset (bottom right corner) in Figure 19a shows

---

**Figure 15.** Time snapshots extracted from the propagating wavefield when the field of Figure 13b is used as a source. Panel d shows the wavefield focused at the location indicated by the black dot in Figure 12a. Time is increasing from panel a to f.

**Figure 16.** (a) Initial incident wavefield $p_+^i(x,t)$. (b) First iteration of the incident wavefield $p_1^i(x,t)$. This initial field will focus the wavefield at the virtual source location in Figure 12a at $t = 0$, when the two planar and parallel line sources replace the virtual source indicated by the black dot.
the shape of the spatially-extended virtual source considered here. The ∧-shaped source replaces the black dot used previously. In contrast to the first example of this section, we compute the first arrivals using the true model instead of the smooth macro model. We do this to show that the quality and degree of focusing at the virtual source location depends on the first arrivals. Since the direct arrivals are now exact, Figure 25d shows that the wavefield focuses well on the target source. The proposed scheme produces a better result with fewer artifacts in this situation (compared to Figure 22d) because we have a correct estimate of the first arriving wavefront. Furthermore, the reconstructed response shown in Figure 24a looks almost identical to the directly-modeled response shown in Figure 24b.

4 CONCLUSIONS

We proposed a generalization to two dimensions of the model-independent wavefield reconstruction method of Broggi et al. (2011, 2012). Unlike the one-dimensional method, which uses the reflection response only, the proposed multi-dimensional extension requires, in addition to the reflection response, independent information about the first arrivals.

The proposed data driven procedure yields the response to a virtual source (Figures 14a, 17a, 21a, 24a) and reconstructs correct internal multiples, without needing a receiver at the virtual source location and without needing detailed knowledge of the medium. The method requires (1) the direct arriving wave front at the surface originated from a virtual source in the subsurface, and (2) the reflection impulse responses for all source and receiver positions at the surface. The direct arriving wave front can be obtained by modeling in a macro model, from microseismic events (Artman et al., 2010), from borehole check shots, or directly from the data by the CFP method (Berkhout, 1997) when the virtual source is located at an interface. In our numerical examples, we used a smooth version of the true model to compute the direct arrivals. The required reflection impulse responses are obtained from seismic reflection data after surface-related multiple elimination (Verschuur et al., 1992) and deconvolution for the source wavelet. Furthermore, because the proposed method is non-recursive, the reconstruction of internal
Data-driven wavefield reconstruction and focusing

Figure 19. (a) True velocity model with a syncline reflector. (b) Smooth velocity model used to model the first arrivals between the virtual source location and the acquisition surface at $z = 0$. In both panels, the black dot represents the location of the extended virtual source used in the first example of Section 3.2. The square inset (bottom right corner) shows the shape of the extended virtual source considered in the second example of Section 3.2. The ∧-shaped source replace the black dot in the second example.

multiples will not suffer from error magnification, unlike other imaging methods that aim at internal multiple suppression (Berkhout and Verschuur, 2005; Luo et al., 2011).

For a simple configuration with planar dipping reflectors, the stationary-phase analysis gives insight into the mechanism of the two-dimensional iterative scheme and confirms that the method converges to the virtual-source response. Following the physical arguments in section 2, we produced numerical examples to test the method in various configurations and showed that the full response to the virtual source can be successfully reconstructed. We showed that the method is not limited to point sources, but it competently handles spatially-extended virtual sources. This feature permits one to create and steer oriented beams originating at depth under a complex overburden that generates strong multiples. This beamforming process could have a potential use in beam migration techniques (Gray et al., 2009). Errors in the estimated first arrivals (due to a smooth macro model) cause defocusing and a mislocalization of the virtual source (as in standard imaging algorithms). Such errors, however, do not affect the handling of the internal multiples and do not deteriorate their reconstruction. This reconstruction is handled by the ac-

Figure 20. (a) Initial incident wavefield $p^+_0(x, t)$. (b) First iteration of the incident wavefield $p^+_1(x, t)$. This initial field will focus the wavefield at the virtual source location (black dot) in Figure 19a at $t = 0$.

Figure 21. (a) Causal part of the superposition of the total field and its time-reversed version, $p_1(x, t) + p_1(x, -t)$. Label (1) indicates an incomplete cancellation of the events inside the window $w(x, t)$. Label (2) shows the effect of edge diffraction. (b) Directly-modeled full response to the virtual source (black dot) shown in Figure 19a.
Figure 22. Time snapshots extracted from the propagating wavefield when the field of Figure 20b is used as a source. Panel d shows the wavefield focused at the location indicated by the black dot in Figure 19a. Time is increasing from panel a to f.

virtual medium through the reflection data measured at the surface (which includes all the information about the medium itself). Note that also the virtual source method (Bakulin and Calvert, 2006) does not optimally focus the wavefield at the virtual source location, but Wapenaar et al. (2011b) show that the focusing can be improved by applying multi-dimensional deconvolution.

ACKNOWLEDGMENTS

The authors thank the members of the Center for Wave Phenomena for their constructive comments. This work was supported by the sponsors of the Consortium Project on Seismic Inverse Methods for Complex Structures at the Center for Wave Phenomena.

REFERENCES


Figure 23. (a) Initial incident wavefield $p^i(x, t)$. (b) First iteration of the incident wavefield $p^i_1(x, t)$. This initial field will focus the wavefield at the virtual source location in Figure 19a at $t = 0$, when the $\wedge$-shaped virtual source replaces the black dot.

Figure 24. (a) Causal part of the superposition of the total field and its time-reversed version, $p_1(x, t) + p_1(x, t^-)$. Label (1) indicates an incomplete cancellation of the events inside the window $w(x, t)$. (b) Directly-modeled full response to the $\wedge$-shaped virtual source shown in Figure 19a.
Figure 25. Time snapshots extracted from the propagating wavefield when the field of Figure 23b is used as a source. Panel d shows the wavefield focused at the location indicated by the ∨-shaped symbol in Figure 19a. Time is increasing from panel a to f.