Full reciprocity-gap waveform inversion enabling sparse-source acquisition

Florian Faucher[1], Giovanni Alessandrini[2], Hélène Barucq[3], Maarten V. de Hoop[4], Romina Gaburro[5], and Eva Sincich[2]

ABSTRACT

The quantitative reconstruction of subsurface earth properties from the propagation of waves follows an iterative minimization of a misfit functional. In marine seismic exploration, the observed data usually consist of measurements of the pressure field, but dual-sensor devices also provide the normal velocity. Consequently, a reciprocity-based misfit functional is specifically designed, and it defines the full reciprocity-gap waveform inversion (FRgWI) method. This misfit functional provides additional features compared to the more traditional least-squares approaches, in particular, in that the observational and computational acquisitions can be different. Therefore, the positions and wavelets of the sources from which the measurements are acquired are not needed in the reconstruction procedure and, in fact, the numerical acquisition (for the simulations) can be chosen arbitrarily. Based on 3D experiments, FRgWI is shown to behave better than full-waveform inversion in the same context. It allows for arbitrary numerical acquisitions in two ways: when few measurements are given, a dense numerical acquisition (compared to the observational one) can be used to compensate. However, with a dense observational acquisition, a sparse computational one is shown to be sufficient, for instance, with multiple-point sources, hence reducing the numerical cost. FRgWI displays accurate reconstructions in both situations and appears more robust with respect to crosstalk than least-squares shot stacking.

INTRODUCTION

The full-waveform inversion (FWI) method has been developed extensively in the past few decades for quantitative recovery of subsurface earth media in seismic exploration. The concept of FWI is to minimize, with respect to the earth parameters, a misfit criterion defined from the simulations of wave propagation and the measured seismograms (i.e., the "full waveform"). The method was introduced originally by Bamberger et al. (1977, 1979) for the 1D wave equation, and was extended by Lailly (1983) and Tarantola (1984, 1987b). The method first was used with time-domain wave propagation, and the frequency-domain formulation of FWI, which requires a Fourier transform of the original time-dependent seismic traces, was established by Pratt et al. (1996, 1998).

In marine seismic, the data usually consist of measurements of the pressure field from hydrophones, but new devices, dual sensors, have been deployed recently and also give access to the vertical velocity (Carlson et al., 2007; Tenghamn et al., 2007). This additional information has shown advantages of reducing noise for image processing (see Whitmore et al., 2010; Rønholt et al., 2015), and it has motivated new analysis, such as in the work of Alessandrini et al. (2018, 2019), and a specific numerical methodology by

Faucher (2017) and Zhong and Liu (2018) for the seismic inverse problem. In our work, we implement a misfit functional dedicated to dual-sensor data and demonstrate its efficiency for the recovery of subsurface physical parameters.

FWI relies on an iterative minimization of the misfit functional, which, in the traditional least-squares approach, is the L^2 difference between the observations and the simulations. In addition to the numerical challenges induced by the large-scale domain, one main difficulty of FWI (which is a nonlinear and ill-posed inverse problem) is that the misfit functional suffers from local minima, in particular when low-frequency data are missing and/or in the absence of a priori information, as has been observed by, for example, Gauthier et al. (1986), Luo and Schuster (1991), Bunks et al. (1995), and Faucher et al. (2020a). To mitigate the "cycle-skipping" effect, selection of increasing frequency content in the data is commonly used (Bunks et al., 1995; Sirgue and Pratt, 2004). In our work, we follow the frequency-domain formulation, in which this approach is natural, and we further use increasing sequential frequency, as advocated by Faucher et al. (2020a).

Several alternatives to least-squares have been studied to enhance the convexity of the misfit functional, such as logarithmic function, mentioned by Tarantola (1987a), that is particularly appropriate in complex-frequency FWI (Shin and Min, 2006; Shin and Cha, 2008; Faucher, 2017). Comparisons of misfit using the phase and amplitude of signals are carried out in the work of Shin et al. (2007), Bednar et al. (2007), and Pyun et al. (2007). Fichtner et al. (2008) study the use of a misfit based upon the phase, correlation, or the envelope; the latter also is advocated by Bozdağ et al. (2011), and it is shown to be efficient for the reconstruction of attenuation properties in global earth seismology by Karaoğlu and Romanowicz (2017). Brossier et al. (2010) study the use of the L^1 criterion. The optimal transport distance is considered in the work of Métivier et al. (2016) and Yang et al. (2018) and is shown to improve the misfit functional convexity. To avoid the local minima, one also can rely on a specific model parameterization such as the migration-based traveltime method that decomposes the velocity model into a (smooth) background profile and a reflectivity in the data space (Clément et al., 2001; Barucq et al., 2019). This approach is shown to increase the size of the attraction basins for background velocity reconstruction by Barucq et al. (2019).

In the case in which Cauchy data are available on the boundary, that is, a field and its normal derivative (e.g., the pressure and the normal derivative of the pressure, which actually relates to the normal velocity; see Appendix A) the reconstruction can use a combination of the Dirichlet and Neumann traces, as in the work of Kohn and Vogelius (1985) and Colton and Haddar (2005) in inverse scattering, or Alessandrini et al. (2018) for seismic. This approach is labeled as the "reciprocity-gap" in the literature (de Hoop and de Hoop, 2000; Colton and Haddar, 2005). Lipschitz stability can be obtained for the partial Cauchy data inverse problem, in particular in the seismic context with surface measurements, as proved by Alessandrini et al. (2018, 2019). This result further holds for piecewise-linear parameters, used for the numerical experiments of Faucher (2017) and Alessandrini et al. (2019). In the context of acoustic waves governed by the Euler's equations, we shall see that Cauchy data are in fact equivalent to the pressure field and the normal velocity, that is, dual-sensor data. In Appendix A, we further connect the boundary measurements to the volume properties of the media. Hence, the minimization of the misfit functional defined from surface data is connected to the recovery of the volume medium parameters. In our work, we use the reciprocity-gap formulation and further combine all sources, defining the full reciprocity-gap waveform inversion (FRgWI) framework for seismic imaging using dual-sensor data.

In the time-harmonic formulation, the reciprocity-gap results in a misfit functional in which the observations and simulations are multiplied. Although our work is conducted in the frequency domain, it can similarly be carried out in the time domain where it relates to the family of correlation misfits. Correlation-based misfit functionals have been studied in seismic tomography, for instance, by Van Leeuwen and Mulder (2010) and in the work of Choi and Alkhalifah (2011) and Zhang et al. (2016), although using a single type of measurements (i.e., the acoustic pressure fields). The use of the velocity fields is studied in the time domain by Zhong and Liu (2019), assuming that all directional velocities are available and convolving the same fields. A fundamental difference with these earlier studies, and in particular with Zhong and Liu (2019), is that in our work, we correlate different fields (i.e., pressure with velocity). This is the essence of the reciprocity gap, and it is crucial for relating surface measurements to global model reconstruction using Green's identity (Appendix A). In addition, our framework is naturally designed for the normal velocity, in accordance with the dual-sensor devices. The works of Menke and Levin (2003) and Bodin et al. (2014) in seismology also combine observations and simulations, but using a cross-convolution formula made of the vertical and horizontal components of the waves. More generally, approaches based upon a specific filtering of the data, such as in the work of Warner and Guasch (2016) and Guasch et al. (2019), also rely on a minimization in which the fields are not directly the quantities compared.

The main feature of FRgWI is to allow different observational and numerical acquisitions (Faucher, 2017; Alessandrini et al., 2019; Zhong and Liu, 2019). Minimal information regarding the observational sources is required: the source function and the source positions are not needed to conduct the reconstruction. This is due to the definition of the misfit functional, which does not compare an observation with a simulation directly, but products of an observation with a simulation. This flexibility in the choice of numerical probing sources opens up many perspectives and, in particular, to use denser or sparser computational acquisitions than the given observational one. The use of sparse acquisition relates to the shot-stacking approach for data decimation, which sums several single point-source data using the linearity of the wave equation. It has been used early in seismic applications to reduce the computational cost by Mora (1987). It is based upon the redundancy of information in the data, but it has a major drawback in that it is difficult to avoid crosstalk between the encoded sources (Krebs et al., 2009; Zhang et al., 2018). It has motivated several works to efficiently assemble the encoded sources in seismic, that is, source blending (Berkhout, 2008). We mention, for instance, the random combination of sources changing with iteration by Krebs et al. (2009), the approach based upon compressive sensing by Li et al. (2012), while wavelet encoding is used by Zhang et al. (2018). In our applications, we will see that FRgWI, by using arbitrary numerical sources, can work with multiple-point sources (for computational or observational acquisition) while being naturally robust to crosstalk. In fact, it is not exactly crosstalk in this context because the key of FRgWI is that the measurements are not modified; that is, they still are

tested independently, one by one, with respect to the numerical simulations.

In this paper, we study the seismic inverse problem associated with time-harmonic waves using dual-sensor data. We first state the mathematical problem and define the two misfit functionals that we analyze for the iterative reconstruction: the traditional least-squares difference and the reciprocity-based functional. Next, we detail the features provided by the reciprocity-gap formulation. We carry out 3D reconstruction experiments: first with a layered medium, in which we compare the performance of reciprocity-gap and FWI. We further demonstrate the efficiency of FRgWI with simultaneous point sources to reduce the computational cost and its robustness compared to traditional shot stacking. Finally, we carry out a larger scale experiment including salt domes using the SEG Advanced Modeling Corporation (SEAM) benchmark. In particular, we highlight that FRgWI performs equally well in the case of acquisitions that are sparse for observations/dense for simulations and when acquisitions are dense for observations/sparse for simulations. Our experiments are carried out using data acquired in the time domain, and, even though our reconstruction algorithm is conducted in the frequency domain, which might present some scale limitations for field configurations, we believe that our method can be implemented with the time-domain wave equation. This can be more appropriate for larger scales, and it requires only a slight modification to our model. A discussion about it is carried out in the "Conclusion" section.

TIME-HARMONIC SEISMIC INVERSE PROBLEM WITH DUAL-SENSOR DATA

We work with time-harmonic wave propagation for the identification of the physical parameters in a seismic context. The quantitative reconstruction is conducted using an iterative minimization of a misfit functional. We first give the misfit as the traditional L^2 difference and further design the reciprocity-gap version of the functional, which combines pressure and normal velocity data.

Acoustic wave equation, forward problem from dual sensors

We consider a 3D domain $\Omega \subset \mathbb{R}^3$ with boundary Γ. The propagation of waves in acoustic media is represented with the scalar pressure and vectorial velocity fields that satisfy Euler's equations (Kirsch, 1996; Colton and Kress, 1998; de Hoop and de Hoop, 2000):

(Problem 1)

$$\begin{cases} -i\omega\rho(\boldsymbol{x})\boldsymbol{v}(\boldsymbol{x}) = -\nabla p(\boldsymbol{x}), & \text{in } \Omega, & \text{(1a)} \\ -i\omega\kappa(\boldsymbol{x})^{-1}p(\boldsymbol{x}) = -\nabla\cdot\boldsymbol{v}(\boldsymbol{x}) + f(\boldsymbol{x}), & \text{in } \Omega, & \text{(1b)} \\ p(\boldsymbol{x}) = 0, & \text{on } \Gamma_1 \,(\text{Free surface}), & \text{(1c)} \\ \partial_\nu p(\boldsymbol{x}) - \frac{i\omega}{c(\boldsymbol{x})}p(\boldsymbol{x}) = 0, & \text{on } \Gamma_2 \,(\text{ABC}). & \text{(1d)} \end{cases}$$

The frequency is denoted by ω, f is the (scalar) interior (harmonic) source term, and ∂_ν denotes the normal derivative. The propagation is characterized by the two physical parameters of the medium: density ρ and bulk modulus κ. In addition, the velocity c is defined such that

$$c(\boldsymbol{x}) = \sqrt{\frac{\kappa(\boldsymbol{x})}{\rho(\boldsymbol{x})}}. \tag{2}$$

For boundary conditions, we follow a geophysical context in which the boundary is separated into two: $\Gamma = \Gamma_1 \cup \Gamma_2$. We denote by Γ_1 the interface between the air and the acoustic medium, in which a free-surface boundary condition holds, equation 1c. On the other part of the boundary Γ_2, we implement an absorbing boundary condition (ABC), Engquist and Majda (1977), equation 1d, to ensure that waves reaching the boundary are not reflected back into the domain. It corresponds to the fact that the domain Ω is a numerical restriction of the earth.

The dual-sensor devices have been recently introduced in marine seismic exploration, and they allow the recording of the pressure field and the vertical velocity (Carlson et al., 2007; Tenghamn et al., 2007). Consequently, we define the forward problem (which maps the parameters to the data) $\mathcal{F}_\omega^{(f)}$ at frequency ω for a source f such that

$$\mathcal{F}_\omega^{(f)}(m) = \{p_\omega^{(f)}|_\Sigma, v_{\nu,\omega}^{(f)}|_\Sigma\}. \tag{3}$$

The model parameters are referred to by $m = (\kappa, \rho)$, the normal velocity by v_ν, and Σ corresponds with the (discrete) set of receivers' location:

$$p|_\Sigma = \{p(\boldsymbol{x}_1), \ldots, p(\boldsymbol{x}_{n_{\text{rcv}}})\}, \tag{4}$$

where $\boldsymbol{x}_i$ is the position of the ith receiver for a total of n_{rcv} receivers. For notation, we introduce the restriction operator $\mathcal{R}$, which reduces the fields to the set Σ, such that

$$\mathcal{R}(p) = p|_\Sigma, \qquad \mathcal{R}(v_\nu) = v_\nu|_\Sigma. \tag{5}$$

Misfit functionals

The inverse problem aims the quantitative reconstruction of the subsurface medium parameters (κ and ρ) from data measured at the receivers' location. Using the pressure and vertical velocity measurements, we denote the data at frequency ω for a source f by

$$d_\omega^{(f)} = \{d_{\omega,p}^{(f)}, d_{\omega,v}^{(f)}\}, \tag{6}$$

where $d_{\omega,p}^{(f)}$ and $d_{\omega,v}^{(f)}$ are vectors of $\mathbb{C}^{n_{\text{rcv}}}$ and they respectively refer to the pressure and normal velocity records, following equation 3. We further denote by $d_\omega^{(f)}(\boldsymbol{x}_i)$ the data recorded for the source f at the ith receiver.

In the following, we omit the frequency index and space dependency for the sake of clarity (in the experiments, we use an increasing sequential frequency as suggested by Faucher et al., 2020a). We introduce two misfit functionals that evaluate the difference between the observations and simulations.

First, the functional $\mathcal{J}_{L2}$, which follows the traditional least squares, is

$$\mathcal{J}_{L2}(m) = \frac{1}{2}\sum_{i=1}^{n_{\text{src}}}\|\mathcal{R}(p^{(f_i)}) - d_p^{(f_i)}\|_2^2 + \frac{\eta}{2}\|\mathcal{R}(v_\nu^{(f_i)}) - d_v^{(f_i)}\|_2^2, \tag{7}$$

where η is a scaling factor to adjust between the amplitudes of the pressure and velocity. In our applications, it is taken such that $\|d_p\| = \eta\|d_v\|$.

Second, we define an alternative misfit functional based upon the reciprocity gap, motivated by Green's identity:

$$\mathcal{J}_r(m) = \frac{1}{2} \sum_{i=1}^{n_{src}^{obs}} \sum_{j=1}^{n_{src}^{sim}} \| d_v^{(f_i)T} \mathcal{R}(p^{(g_j)}) - d_p^{(f_i)T} \mathcal{R}(v_{\nu}^{(g_j)}) \|_2^2,$$

$$= \frac{1}{2} \sum_{i=1}^{n_{src}^{obs}} \sum_{j=1}^{n_{src}^{sim}} \left\| \sum_{k=1}^{n_{rcv}} (d_v^{(f_i)}(\boldsymbol{x}_k) p^{(g_j)}(\boldsymbol{x}_k) - d_p^{(f_i)}(\boldsymbol{x}_k) v_{\nu}^{(g_j)}(\boldsymbol{x}_k)) \right\|_2^2,$$

$$(8)$$

where T denotes the transpose. The misfit functional is motivated by the Green's identity and has been introduced in the context of inverse scattering, from Cauchy data, as mentioned in the "Introduction" section (Kohn and Vogelius, 1985; Colton and Haddar, 2005) and used with partial seismic data by Alessandrini et al. (2019). In Appendix A, we justify the formulation of the misfit functional $\mathcal{J}_r$ using variational formulation of problem 1, and note that the mathematical foundation of the reciprocity-gap formulation relies naturally on the normal velocity. Therefore, it perfectly matches the dual-sensor data, and it does not necessitate the specific directional components (v_x, v_y, or v_z).

Iterative reconstruction procedure

The reconstruction procedure follows an iterative minimization of the selected misfit functional. Least-squares formulations such as equation 7 are traditionally referred to as the FWI method in seismic (because one makes use of the full data seismograms); see the review of Virieux and Operto (2009). Consequently, we shall refer to the minimization of equation 8 as FRgWI. In any case, the iterative minimization follows successive updates of the physical models, such that, at iteration l, the new model is given by $m_{l+1} = m_l - \alpha_l s_l$, where α is

Algorithm 1: Iterative reconstruction procedure: For the minimization of equation 7, the computational acquisition (source positions and wavelet) must follow the one used to generate the measurements. For the minimization of equation 8, the user can prescribe a computational acquisition by means of the positions and wavelet of the probing sources g_j in equation 8.

Initialization: starting models $m_1 = (\kappa_1, \rho_1)$ and measurements d_p, d_v.

for $\omega_i = \omega_1, \ldots, \omega_{N_\omega}$ **do**
  for $j = 1, \ldots, n_{iter}$ **do**
    set $l := (i-1)n_{iter} + j$;
    solve problem 1 using current models m_k and frequency ω_i;
    compute the misfit functional, equation 7 or equation 8;
    compute the gradient of the misfit functional using the adjoint-state method;
    compute the search direction, s_l, using limited-BFGS;
    compute the step length, α_l, using line-search method;
    update the model: $m_{l+1} = m_l - \alpha_l s_l$.
  end
end

the scalar step size typically computed via a line-search method (Nocedal and Wright, 2006) and s is the search direction.

The search direction depends on the gradient of the cost function, which is computed using the adjoint-state method. The method has its foundation in the body of work of Lions (1971) and is reviewed for seismic application by Plessix (2006). Application with complex fields is further described by Barucq et al. (2018, 2019). The adjoint-state method for the reciprocity-gap functional is briefly reviewed in Appendix B (Alessandrini et al., 2019). In our implementation, the search direction depends only on the gradient of the misfit functional, and we use the limited-BFGS algorithm (Nocedal, 1980; Nocedal and Wright, 2006). We review the steps for the iterative minimization in Algorithm 1.

Discretization with the hybridizable discontinuous Galerkin method

In the numerical implementation, the pressure and velocity fields must be computed to feed the misfit functional. We discretize problem 1 using the hybridizable discontinuous Galerkin (HDG) method, which is specifically designed for first-order problems, and avoid oversize linear systems. Indeed, the global matrix using HDG is composed only of the degrees of freedom associated with the trace of the pressure field, that is, only the degrees of freedom on the faces of the elements of the discretization mesh (Cockburn et al., 2009; Griesmaier and Monk, 2011; Bonnasse-Gahot et al., 2017). Then, local (small) systems are solved to calculate the volume solutions of the pressure and the velocity fields, computed with similar accuracy. We refer to Faucher and Scherzer (2020) for the numerical implementation associated with problem 1.

With more traditional discretization methods such as continuous Galerkin, finite differences of internal penalty discontinuous Galerkin methods, one has to create a linear system whose size is the total number of degrees of freedom for all unknowns (i.e., the pressure and the three components of the velocity), possibly leading to large linear systems. Alternatively, one can solve only for the scalar pressure field and postprocess the solution to obtain the velocity, but the computed velocity loses one order of accuracy compared to the discretized pressure field due to the derivative in equation 1a. In the HDG method, we obtain the pressure and velocity fields with the same accuracy, whereas the global linear system contains only the degrees of freedom of a scalar unknown (Faucher and Scherzer, 2020). In addition, only the degrees of freedom on the faces of the elements are taken into account, hence removing all interior ones. Consequently, the HDG method has been shown to be more efficient (i.e., less memory consumption for the matrix factorization due to the smaller global matrix) compared to other approaches by Kirby et al. (2012) and Bonnasse-Gahot et al. (2017), for discretization.

In the context of large-scale seismic applications, the use of HDG is crucial (particularly in the frequency domain) to efficiently compute the pressure and the velocity because it eventually leads to linear systems with sizes that are not larger than for second-order wave problems. For the resolution of the subsequent linear system, we use a direct solver (Amestoy et al., 2001; Liu et al., 2020) for the multiple right-sides feature, in particular, the solver MUMPS, designed for sparse matrices (Amestoy et al., 2006). Therefore, the matrix factorization of the forward problem is reused for the backward one that

serves to compute the gradient of the misfit functional depicted in Appendix B.

FEATURES OF FRgWI AND DATA ACQUISITION

Following Faucher (2017) and Alessandrini et al. (2019), the fundamental feature of the reciprocity-gap misfit functional equation 8 compared to equation 7 is that the set of computational sources is separated from the observational ones (respectively, g_j and f_i in equation 8). It implies that (1) we do not have to know the observational source positions (the location of the f_i in equation 8) for the reconstruction algorithm. (2) We do not have to know the observational source signatures (wavelet). Consequently, and contrary to the least-squares misfit such as equation 7, minimal information is required regarding the observational acquisition (only the positions of the receivers). The recovery of the source wavelet in parallel to the iterative minimization procedure usually performs well in seismic exploration, using the method prescribed by Pratt (1999); see remark 1. However, incorrect knowledge of the position of the sources is a strong difficulty, which can lead to the failure of the reconstruction procedure. Namely, FRgWI increases robustness by being free of such considerations. In addition, (3) the set of computational sources can be arbitrarily taken; hence, they can differ from the observational one (respectively, n_{src}^{obs} and n_{src}^{sim} in equation 8).

Therefore, it provides high flexibility regarding the choice of computational sources. It is important to notice that whatever computational sources are selected, the observed data still are independently tested one by one against the simulations; that is, the measurements are not modified in any way.

Data acquisition

In our experiments, we consider the source term for the wave propagation (f or g) to be a Ricker function in time and a delta-Dirac function in space. In the frequency domain, it translates to a delta-Dirac function δ (with value equal to the discrete Fourier transform of the time-domain signal). We refer to a point source when the source is localized at position x_k:

$$f_k = \delta(x - x_k), \qquad \text{point source in } x_k. \qquad (9)$$

We further refer to the multiple point source when it is composed of n_{pt} point sources, such that

$$f_j = \sum_{k=1}^{n_{pt}} \delta(x - x_{j,k}^{(f)}), \qquad (10)$$

$$\text{multiple point source.}$$

The observational setup uses a fixed lattice of receivers and sources located slightly above, as illustrated in Figure 1a. Furthermore, the position of the receivers remains the same for all sources. This configuration is consistent with the TopSeis acquisition system for marine seismic developed by Compagnie Générale de Géophysique and Lundin Norway AS. This principle has been recently deployed in the field, showing benefits for near-offset coverage. Namely, it consists in

having two boats: one that moves and carries the sources and one that carries the receivers and remains fixed. Similarly, the use of ocean-bottom node networks is appropriate.

For the reconstruction using FRgWI, we can use arbitrary numerical sources, which do not have to coincide with the observational acquisition. Here, we investigate the two following configurations.

Dense point-source computational acquisition for sparse multiple-point measurements

In the case when the observational acquisition is composed of few sources (e.g., to reduce the cost of field experiments), FRgWI can use a dense coverage of computational point sources to enhance the sensitivity. The sources for the measurements can be multiple points, for example, if several air guns are excited at the same time. In our experiments, the choice of positions for the multiple points follows a structured decomposition: adjacent sources are grouped, with a fixed distance in x and y between each (Figure 1b). It appears more natural to use such a structured decision; for example, the boat will carry the air-gun sources all together. In addition, we have observed in our experiments that this structured partition gives better performance than using a random combination of sources.

Sparse multiple-point computational acquisition for dense point-source measurements

In the case when the observational acquisition is composed of many (point) sources (as is routine in exploration seismic), FRgWI can instead use multiple-point sources in the computational acquisition to reduce the computational cost. It consists of a sparsification of the observational acquisition.

To compare with the traditional FWI approach, one can use the linearity of the wave equation and the multiple-point source can be assimilated to the well-known shot-stacking approach, which rewrites the misfit functional equation 7 such that

$$\mathcal{J}_{L2}^{stack}(m) = \sum_{j=1}^{n_{stack}} \frac{1}{2} \left\| \sum_{k=1}^{n_{pt}} \mathcal{R}(p^{(f_{j,k})}) - \sum_{k=1}^{n_{pt}} d_p^{(f_{j,k})} \right\|^2$$

$$+ \frac{\eta}{2} \left\| \sum_{k=1}^{n_{pt}} \mathcal{R}(v_{\nu}^{(f_{j,k})}) - \sum_{k=1}^{n_{pt}} d_v^{(f_{j,k})} \right\|^2, \qquad (11)$$

where we have a total of n_{stack} multiple-point sources, each composed of n_{pt} points.
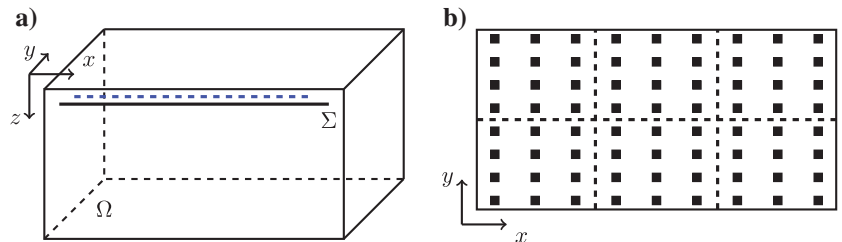

Figure 1. (a) Domain of acquisition: all sources are positioned at the same depth (the blue dashed line) and the receivers are slightly below (the black line, Σ). The receivers remain fixed for all sources according to the Topseis acquisition. (b) The (x, y) source plane (all sources are at the same depth z): every black square corresponds to a point source. The multiple-point sources are defined by taking all points in a specified subdomain, indicated with the dashed line, that is, using a structured decomposition.

Therefore, we will investigate in the following numerical sections different situations in which one of the acquisitions (observational or computational) is composed of point sources whereas the other is made of multiple-point sources. The main advantage of FRgWI is that it does not require any modification of the data, which are tested one by one against the computational acquisition in equation 8. However, the shot-stacking version of FWI requires the summation of data in equation 11, which results in the possible loss of information, that is, crosstalk (Zhang et al., 2018).

NUMERICAL EXPERIMENT 1: LAYERED MEDIUM

In this section and the next, we carry out 3D experiments of geophysical reconstruction to study the performance of FRgWI. In this first test, we compare with the traditional least-squares functional, and we use multiple-point sources to probe the robustness of arbitrary source positions; we also test the combination of sparse and dense acquisitions.

Velocity model

We consider a 3D acoustic medium (provided by Statoil) of $2.55 \times 1.45 \times 1.22$ km^3 in the x-, y-, and z-axes, respectively, (i.e., z is depth). We assume a constant density $\rho = 1000$ kg m^{-3}, and the subsurface velocity is pictured in Figure 2. It consists of different geophysical layers, with nonmonotone model variations, from 1500 to 5200 ms^{-1}. The first 500 m are mostly constant with a velocity of approximately 1600 m s^{-1}.

For the iterative reconstruction, we start with a 1D velocity profile pictured in Figure 3. This initial guess does not encode any a

priori information, and it has only an increasing velocity with the depth. Moreover, the range of values is lower compared to the true medium of Figure 2.

Time-domain data with noise

We work with time-harmonic wave propagation, but we follow a seismic context where measurements are obtained in the time domain. Furthermore, we incorporate white Gaussian noise in the synthetic seismograms to have a more realistic setup. The observational acquisition consists of 160 point sources, excited one by one. They are positioned at the depth of $z = 10$ m, and form a regular array such that sources are every 160 m along the x-axis and every 150 m along the y-axis. There is a total of 1376 receivers, which are located at a fixed depth of 100 m, every 60 m along the x-axis, and every 50 m along the y-axis. Note that the receivers remain at the same position for all sources (Figure 1).

We generate the time-domain seismograms (we have used the parallel time-domain code Hou10ni [see https://team.inria.fr/magique3d/software/hou10ni/]; it relies on internal penalty discontinuous Galerkin discretization [whereas we use HDG]. Also, the meshes are different between the time-domain modeling and the harmonic inversion) and incorporate noise with a signal-to-noise ratio (S/N) of 10 dB. Then, we proceed with the discrete Fourier transform to feed algorithm 1. In Figure 4, we picture the noisy time-domain pressure trace for a single point source, and the corresponding Fourier-transform that we use for the reconstruction. We respect the seismic constraint that the low frequencies are not available from the time-domain data (because these are not contained in the active source), and we work only with data from 5 to 15 Hz frequency. In Figure 5, we picture 2D time-space sections of the traces for the pressure and normal velocity, and we illustrate the effect of the added noise (10 dB S/N in our experiments).

Performance comparison of the mist functionals

We first compare the performance of both misfit functionals $\mathcal{J}_{L2}$ and $\mathcal{J}_r$ in the same context: The numerical acquisition is taken to be the same as the observational one (which is anyway mandatory for $\mathcal{J}_{L2}$). Therefore, we take $n_{src}^{sim} = n_{src}^{obs}$ and the same set for f and g in equation 8. We use algorithm 1, using sequential frequency progression from 5 to 15 Hz; more precisely, we use $\{5, 6, 7, 8, 9, 10, 12, 15\}$ Hz data and 30 minimization iterations per frequency. In Figures 6 and 7, we show the final reconstruction after 15 Hz iterations when minimizing $\mathcal{J}_{L2}$ and $\mathcal{J}_r$, respectively. For the discretization, we use a mesh of approximately 75,000 tetrahedra and polynomials of order three to five (depending on the frequency) for accuracy (note that the mesh differs from the one used to generate the time-domain data).

Remark 1

For the minimization of $\mathcal{J}_{L2}$ equation 7, one has to use the same source wavelet for the observations and simulations. Because the observational source wavelet is not precisely known, we need to reconstruct the source during the iterative process as well. We use the update formula given by Pratt (1999) for the iterative point-source reconstruction. However, this can induce additional difficulty in cases in which the source characteristic is not well re-

Figure 2. The velocity model with layers of $2.55 \times 1.45 \times 1.22$ km^3. For visualization, we extract sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).
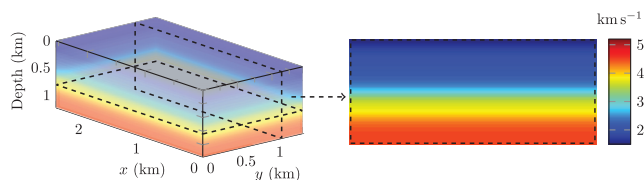

Figure 3. Starting model for the reconstruction of the subsurface velocity of Figure 2. It consists in a 1D profile of increasing speed with depth, with values from 1500 to approximately 4000 m s^{-1}. The horizontal section is for $y = 1125$ m (right).

covered or, more important, when the source positions are not precisely known.

In this experiment, we use the same acquisition context for the two choices of misfit functional, to observe the intrinsic differences between the two methods. We observe the following:

1) Both approaches provide a good reconstruction, where the layers of velocity appear (see the vertical section), and the correct velocity values are retrieved.
2) The FRgWI provides slightly better results, in particular for the parts that are near the boundaries (see the vertical and horizontal sections of Figures 6 and 7). This can be explained as FRgWI formulation equation 8 tests every simulation source with each observational one (the two sums in equation 8), and it somehow compensates for the limited illumination of the boundary zones.
3) Regarding the computational time, both approaches are similar (the only difference is the two sums in the misfit functional equation 8, which is a computationally efficient operation compared to the matrix factorization and linear system resolution).

Therefore, in this first test case, we have demonstrated the efficiency of FRgWI, which produces a better reconstruction compared to the traditional approach, without incurring any increase of computational time. But more important, the reciprocity-gap offers flexibility for the choice of computational acquisition, which we now study.

Comparison of multiple-point sources FRgWI with shot stacking

The main feature of FRgWI is to enable the use of arbitrary probing sources for the numerical simulations (g in equation 8). Here, we investigate the use of multiple-point sources to reduce the computational cost. The observational acquisition is composed of 160 point sources; that is, each source function $f_i(\boldsymbol{x})$ in equations 7 and 8 corresponds to a delta-Dirac function in $\boldsymbol{x}_i^{(f)}$, according to equation 9. For the computational acquisition (g in equation 8), we now consider n_{stack} multiple-point sources, each composed of n_{pt} points (equation 10).

We assume that the multiple-point sources all have the same number of points. The positions of the multiple points are taken to coincide with the observational acquisition, and we consider a group of sources in a structured partition of the original configuration, as illustrated in Figure 1. We have

$$\boldsymbol{x}_{j,\bullet}^{(g)} = \{\boldsymbol{x}_i^{(f)}\}_{i=i_1}^{i_2}, \quad \text{with } i_1 = (j-1)n_{\text{pt}} + 1 \quad \text{and}$$
$$i_2 = i_1 + n_{\text{pt}} - 1, \tag{12}$$

where $\boldsymbol{x}_j$ corresponds with the multiple point source, composed of the point sources located in $\boldsymbol{x}_i$. The FWI counterpart is to use shot stacking, which requires summation of the observed data, according to equation 11.

We perform the iterative reconstruction using $n_{\text{stack}} = 5$ multiple-point sources, each of them composed of $n_{\text{pt}} = 32$ points. The reconstruction using the shot-stacking version of FWI equation 11
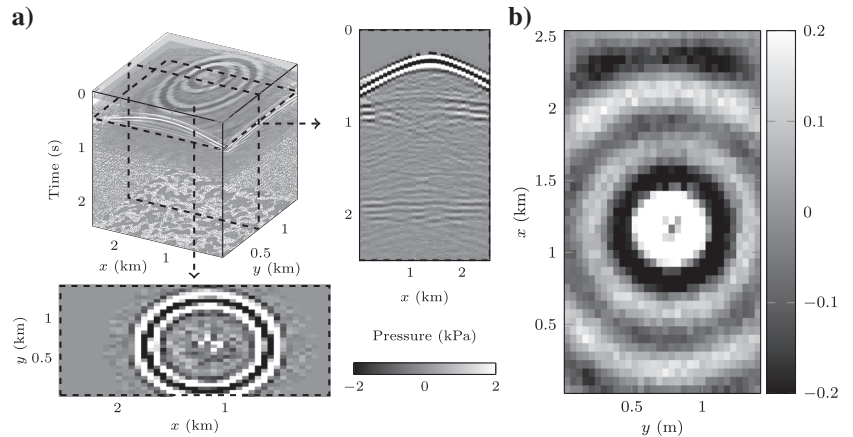

Figure 4. (a) Seismic traces in time including noise and (b) the corresponding real part Fourier transform at 5 Hz frequency. These are associated with a single point source located in $(x_s, y_s, z_s) = (1160, 780, 10)$ m, for the velocity model of Figure 2. The measurements are obtained from 1376 receivers positioned at a depth of $z = 100$ m.
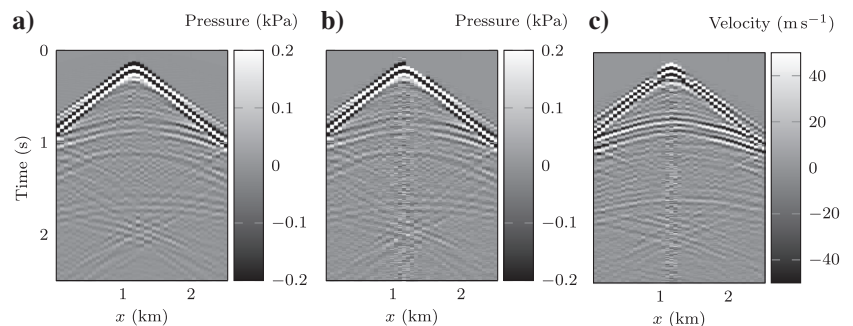

Figure 5. Time-domain (a) pressure trace without noise, (b) pressure trace with 10 dB S/N, and (c) normal velocity trace with 10 dB S/N. These correspond with a line of receivers at a fixed $y = 695$ m, for a single point source located in $(x_s, y_s, z_s) = (1160, 780, 10)$ m for the velocity model shown in Figure 2. For our experiment, we apply 10 dB S/N to the synthetic data and use the Fourier transform of the noisy traces. The available frequency ranges from 5 to 15 Hz (see Figure 4).
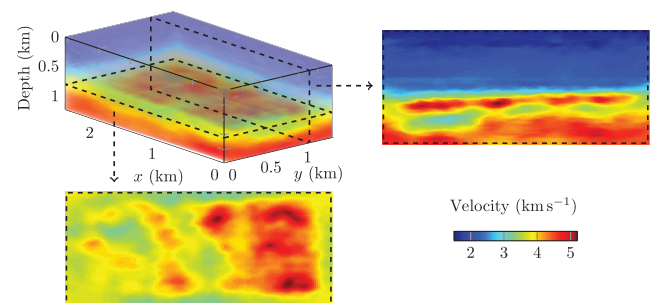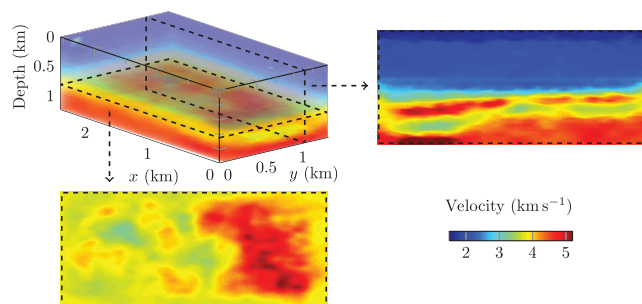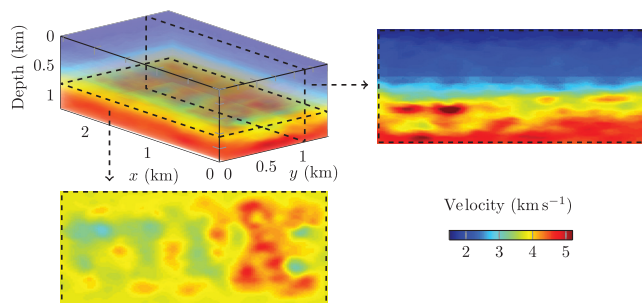

Figure 6. Velocity reconstruction from the minimization of $\mathcal{J}_{L2}$ in equation 7, starting with the model of Figure 3. The reconstruction uses 30 iterations per frequency between 5 and 15 Hz, with pressure and normal velocity data. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).

is shown Figure 8, and the result using FRgWI (i.e., where only the computational acquisition is changed) is shown Figure 9.

We observe that FRgWI performs better than the shot stacking in FWI, in this experiment of relatively small scale:

1) The vertical and horizontal sections shown in Figure 9 are actually very close from the reconstruction obtained using the same computational and observational acquisition (Figure 7). The pattern of increasing and decreasing velocity is well captured with the appropriate values.
2) However, the FWI with shot stacking loses accuracy, in particular the vertical section in Figure 8 where the nonmonotonic variation is not even recovered.

Testing each of the observational sources independently with the arbitrary computational ones, the FRgWI is robust with multiple-point sources and allows a sharp decrease in the number of numerical sources for the simulations.

Remark 2

The results using the shot stacking with FWI could be improved by using a more advanced strategy of shot blending as mentioned in the "Introduction" section (Krebs et al., 2009; Li et al., 2012). We illustrate here that even a naive approach (i.e., simple to implement numerically) of shot stacking is sufficient for the FRgWI to produce satisfactory results. Also, note that we have tried to use a random selection of positions for the multiple point, but it was not performing as well as the structured criterion. This would, however, require more testing to confirm.

We can further reduce the number of computational sources. The reconstruction using FRgWI with $n_{stack} = 2$, $n_{pt} = 80$ is shown in Figure 10. In Figure 11, we show the reconstruction using only one source: $n_{stack} = 1$, $n_{pt} = 160$. We see that FRgWI still behaves quite well, in particular for two sources the reconstruction carries similar accuracy than before: The vertical structures are discovered accurately, with pertinent variations of velocity. For the reconstruction using only one source, Figure 11, we see some deterioration in the recovered layers, but it remains more accurate than the shot-stacking FWI using five sources (Figure 8).

The essence of FRgWI is that it does not modify the observational acquisition and instead it probes every measured seismogram independently. It increases the robustness of the reconstruction procedure, and it allows for the design of arbitrary computational acquisitions to reduce the computational cost. In this experiment, we have used a naive approach of shot summation, which requires no effort for its implementation and already shows accurate reconstructions.
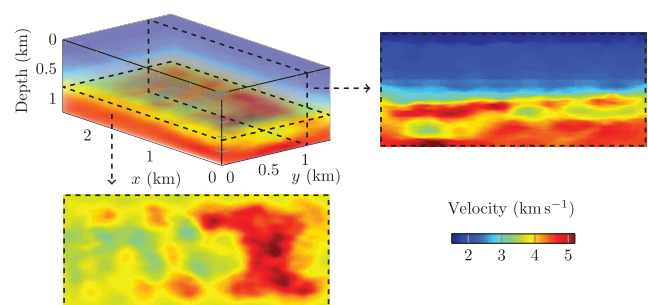

Figure 7. Velocity reconstruction from the minimization of $\mathcal{J}_r$ in equation 8, starting with the model of Figure 3. The reconstruction uses 30 iterations per frequency between 5 and 15 Hz, with pressure and normal velocity data. Here, the numerical acquisition for FRgWI follows the observational one (the number of sources and position). For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).

Figure 9. Velocity reconstruction from the minimization of $\mathcal{J}_r$, starting with the model of Figure 3. The numerical acquisition (g in equation 8) uses $n_{stack} = 5$ sources composed of $n_{pt} = 32$ (compare equation 10). The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).
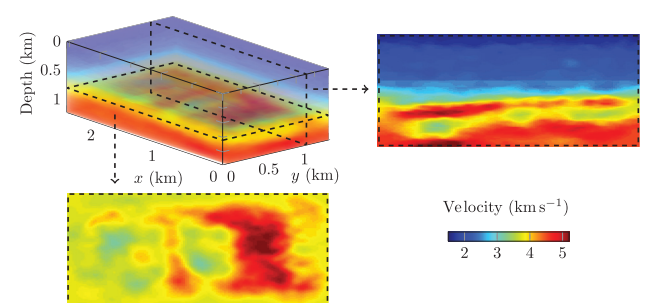

Figure 8. Velocity reconstruction from the minimization of $\mathcal{J}_{L2}^{stack}$ in equation 11, starting with the model of Figure 3. It uses $n_{stack} = 5$ multiple-point sources composed of $n_{pt} = 32$ (compare equation 11). The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).
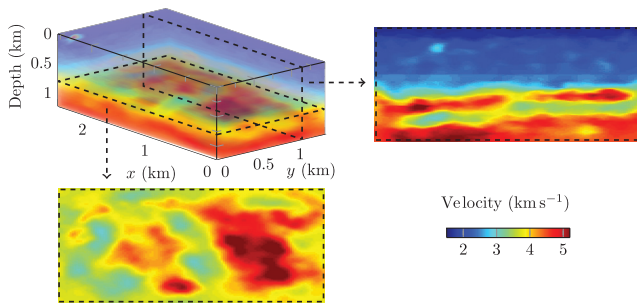

Figure 10. Velocity reconstruction from the minimization of $\mathcal{J}_r$, starting with the model of Figure 3. The numerical acquisition (g in equation 8) uses $n_{stack} = 2$ sources composed of $n_{pt} = 80$ (equation 10). The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).
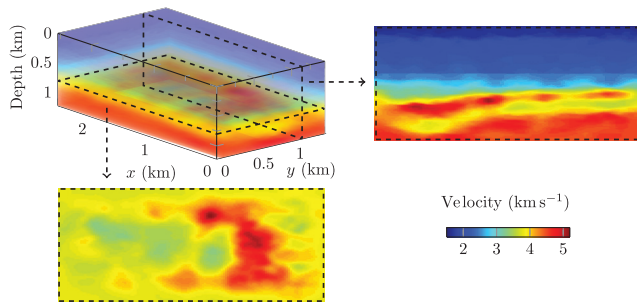
Sparse observational acquisition

We now investigate a different situation: when the measurements are obtained from only a few set of multiple-point sources. This happens, for example, in marine seismic when an array of air guns are excited simultaneously. It consequently speeds up the field acquisition and reduces the amount of data obtained, hence reducing the cost of the field acquisition.

Hence, we now consider that the measurements are obtained from five multiple-point sources. We follow the configuration of the previous subsection but inverting the computational and observational acquisitions: it is now f in equation 8 that is represented with a multiple-point sources equation 12 whereas the computational sources g consist of 160 single point sources. It means that the observed measurements are acquired from the physical phenomenon corresponding with multiple-point sources. Note that from this type of measurement, the FWI reconstruction coincides with the shot-stacking result of Figure 8. In Figure 12, we show the reconstruction in which the observations are obtained from $n_{stack} = 5$ multiple-point sources and the computational acquisition consists of 160 single point sources. In Figure 13, we show the reconstruction in which the measured data result from one multiple point source, composed of 160 points (i.e., all sources are excited at the same time) whereas the numerical acquisition remains with 160 point sources.

We observe that the reconstructions obtained from a drastically reduced set of measurements, resulting from multiple-point sources, retrieve the appropriate velocity variation. It displays similar accuracy compared to the use of single-point source measurements with multiple-point source simulations (Figures 9 and 11 for five and one computational sources, respectively). Namely, it seems that FRgWI is insensitive to which acquisition is made sparse (i.e., with multiple-point sources).

NUMERICAL EXPERIMENT 2: SALT-BODY SEAM MODEL

In this experiment, we consider a 3D velocity model encompassing salt domes. The velocity is extracted from the seismic SEAM benchmark, and is $7 \times 6.5 \times 2.1$ km^3. The velocity model is depicted in Figure 14, and it varies from 1500 to 4800 m s^{-1}. In this experiment, the density is heterogeneous, and it is pictured in Figure 15. Compared to a previous test case, it is of a different nature (salt domes), it is larger, and it has a heterogeneous density. Therefore, we expect this numerical experiment to be more challenging.

Time-domain data

Similarly to the previous experiment, we work with noisy time-domain data and compute their Fourier transform to generate the harmonic data used in the iterative reconstruction. There is a total of 272 point sources in the observational acquisition, which is located at 10 m in depth. The sources are placed on a 2D plane with 400 m between each source along the x- and y-axes. For multiple point acquisition, we follow the structured combination illustrated in Figure 1b. In this experiment, we generate the time-domain data and we use a 20 dB S/N. The resulting seismograms are illustrated in Figure 16 for a single source, where we show the pressure and vertical velocity measurements. The data are acquired by a total of 2805 receivers. For the reconstruction, we perform a Fourier transform of the time-domain noisy data, and we use frequency content from 2 to 7 Hz.

Reconstruction using FRgWI

For the reconstruction, we start with initial guesses that correspond to the smooth version of the true models; they are shown in Figures 17 and 18 for the starting velocity and density, respectively.
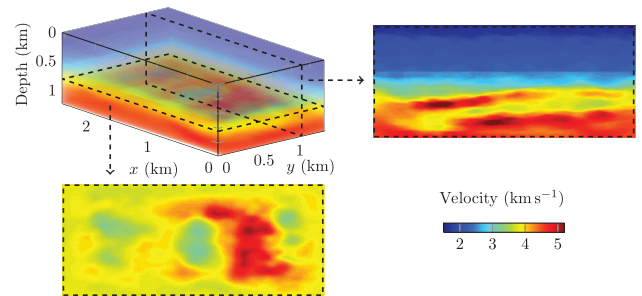

Figure 11. Velocity reconstruction from the minimization of $\mathcal{J}_r$, starting with the model of Figure 3. The numerical acquisition (g in equation 8) uses $n_{stack} = 1$ source composed of $n_{pt} = 160$ (equation 10). The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).

Figure 12. Velocity reconstruction from the minimization of $\mathcal{J}_r$, starting with the model of Figure 3. The observational acquisition (f in equation 8) uses $n_{stack} = 5$ sources composed of $n_{pt} = 32$ (equation 10) whereas the computational acquisition (g in equation 8) uses 160 point sources. The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).

Figure 13. Velocity reconstruction from the minimization of $\mathcal{J}_r$, starting with the model of Figure 3. The observational acquisition (f in equation 8) uses $n_{stack} = 1$ source composed of $n_{pt} = 160$ (equation 10) whereas the computational acquisition (g in equation 8) uses 160 point sources. The reconstruction uses 30 iterations per frequency between 5 and 15 Hz. For visualization, we picture sections at a fixed depth of $z = 800$ m (bottom) and for $y = 1125$ m (right).

We actually only focus on the reconstruction of the velocity model, and we keep the density as its initial representation of Figure 18. The density is known to be more poorly resolved than the velocity because of the lack of sensitivity in the data (Jeong et al., 2012), but it should not prevent us from recovering the velocity (Faucher, 2017).

We use the FRgWI algorithm with the minimization of equation 8 using sequential frequency content from 2 to 7 Hz. We perform 30 iterations per frequency for a total of 180 iterations. We analyze the two following configurations.

1) We first use the dense observational acquisition composed of 272 point sources and multiple-point sources for the computational acquisition (g in equation 8). The multiple-point sources use $n_{pt} = 8$ points, resulting in $n_{stack} = 34$ sources. The reconstructed velocity after the iterations at 7 Hz is shown in Figure 19.
2) Next, we consider the opposite situation: we assume a sparse observational acquisition of $n_{stack} = 34$ sources each composed of $n_{pt} = 8$ points. Here, the computational acquisition is made dense, with 272 point sources. The recovered velocity is shown in Figure 20.

The minimization performs well in both situations: The salt dome is appearing with appropriate values (approximately $4500\,\text{ms}^{-1}$). The upper boundary of the salt is recovered accurately (compare the right sides of Figures 19 and 20), whereas the deepest parts
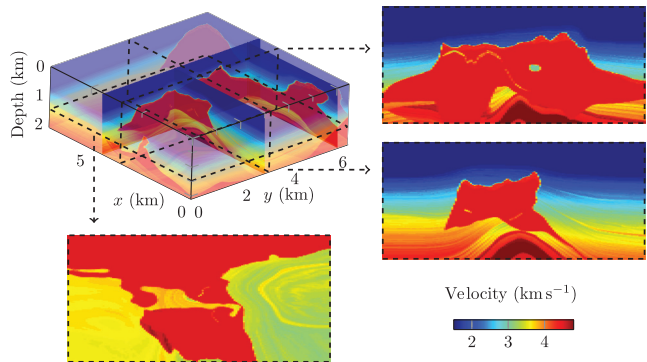
are more difficult to obtain. It is possible that the high velocity contrast prevents us from imaging below the salt. The horizontal section (the bottom of Figures 19 and 20) also shows the appropriate pattern, in particular with the decrease of velocity in the middle, and only the parts near the boundary are slightly less accurate. It confirms the results of our first experiment that FRgWI is relatively insensitive to whether an observational or computational acquisition is taken as sparse or dense: the reconstructions display similar resolution for the two cases. In terms of numerical cost, sparse computational acquisition reduces the computational time but, in the frequency domain with the use of direct solver, this gain remains marginal. To further improve the reconstruction, we shall include a regularization parameter in the minimization, for example, with the total variation approach, which can indeed be supported by our misfit functional; see also Faucher et al. (2020b).

Figure 16. Time domain (a) pressure trace and (b) normal velocity trace with 20 dB S/N. These correspond to a line of receivers at a fixed $y = 3.1$ km, for a single-point source located in $(x_s, y_s, z_s) = (3.4, 3.4, 0.01)$ km for the velocity and density models given in Figures 14 and 15. For our experiment, we apply 20 dB S/N to the synthetic data and use the Fourier transform of the noisy traces. The available frequency ranges from 2 to 7 Hz.

Figure 14. The velocity model of $7 \times 6.5 \times 2.1$ km³. For visualization, we extract sections at a fixed depth of $z = 1.5$ km (bottom), for $x = 3.5$ km (top right), and $y = 3.25$ km (bottom right).

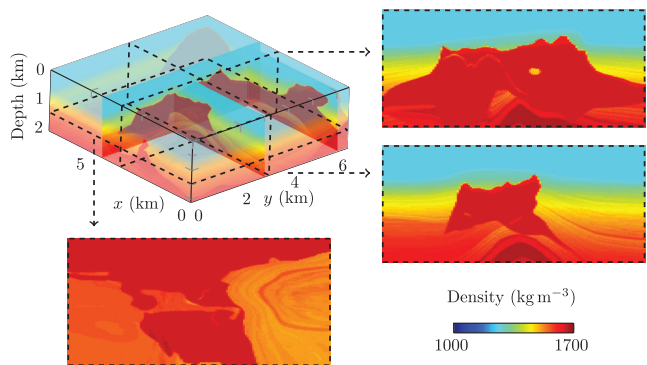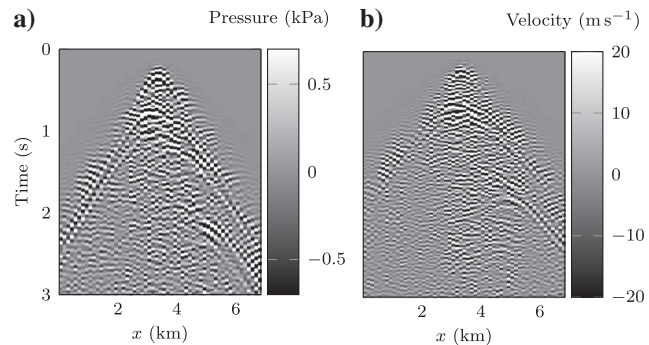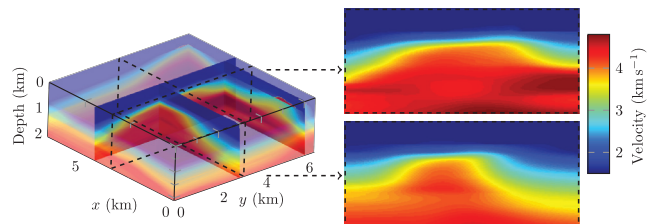Figure 17. Starting velocity model for the reconstruction of the model of Figure 14. We picture vertical sections at a fixed $x = 3.5$ km (top right) and $y = 3.25$ km (bottom right).

Figure 15. The density model of $7 \times 6.5 \times 2.1$ km³. For visualization, we extract sections at a fixed depth of $z = 1.5$ km (bottom), for $x = 3.5$ km (top right), and $y = 3.25$ km (bottom right).
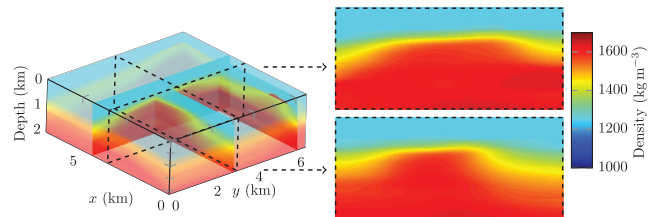

Figure 18. Starting density model for the reconstruction of the model of Figures 14 and 15. We picture vertical sections at a fixed $x = 3.5$ km (top right) and $y = 3.25$ km (bottom right).

ANALYSIS OF RESULTS AND PERSPECTIVES

We have implemented the iterative minimization method based upon the reciprocity-gap functional in the frequency domain, with the HDG discretization to efficiently address the first-order system. To probe the method, we have carried out two 3D experiments: one with a velocity made of layers and another with a medium containing a salt dome, extracted from the SEAM benchmark. First, our experiments have shown that the FRgWI method performs better than FWI in the same configuration, that is, when we keep the observational acquisition for the numerical one, in particular by improving the illumination on the sides. These promising results motivate the implementation of the method for larger, field benchmarks, to probe the scalability of FRgWI. In this case, the time-domain formulation should be considered to continue the performance analysis that we have started.

Flexibility in the computational acquisition

The main feature of FRgWI is its robustness regarding the absence of information on the observational acquisition, and the consequent freedom in the computational one. This is due to the misfit functional equation 8 that works with a product of a datum with a simulation, and that relates to the correlation-based family of methods. For instance, the wavelet used for the computational source does not have to match the observational one, as long as there is an overlap in their frequency contents, and we have not observed any other requirement on the choice of the source function. The use of less traditional sources such as plane waves is theoretically possible, and it should be investigated in the future.

Our main result comes from the positions of the sources used for the simulations that can be chosen arbitrarily. This flexibility is in the core of the misfit functional, which does not directly compare an observation with a simulation but tests the product of any observation with any simulation, hence enforcing the domain illumination. It is crucial to note that the measurements are not modified and are independently tested one by one against each of the (chosen) computations. In our work, we stay with illumination from one side (i.e., simulations and observations are generated near the surface), motivated by Alessandrini et al. (2019), which demonstrates the stability of the method. This choice also is motivated from the derivation of the variational formulation that we give in Appendix A, to avoid the appearance of additional integrals (see the discussion at the end of Appendix A). The use of probing sources from another part of the domain (e.g., deeper) has to be carefully investigated, and it is part of our ongoing research.

Using sparse acquisitions

The use of multiple-point sources allows for the reduction of the computational time or of the field acquisition process. Here, the key is that only one of the two acquisitions is reduced, whereas the other remains dense. Once again, because every datum in the dense acquisition is tested with respect to each datum of the sparse one, it appears to overall produce the proper illumination of the domain. This is possible due to the misfit formula, which offers high flexibility, contrary to more traditional approaches in which both acquisitions must be the same.

We have experimented with the two configurations (the sparse computational acquisition with the dense observational one or

the opposite), and it results in similar accuracy in the reconstruction. Because FRgWI independently tests all combinations of an observation with a simulation, it further appears robust with respect to shot stacking. With FRgWI, we do not exactly observe the crosstalk effect because the measurements are not modified before entering the misfit; instead, we observe some inherent difficulties when the acquisition is too sparse. This is unavoidable, for instance, if the original field data have been obtained from a few sources only: we cannot artificially enhance the information contained in the measurements. However, the difficulties that come from the summation of data (i.e., having multiple-point sources in a sparse acquisition) and that results in crosstalk are diminished by the misfit functional, which takes the product with each datum in the dense acquisition.

In the context of shot stacking, there is a "reference" solution, which is the use of the nonsummed fields. Here, we raise the
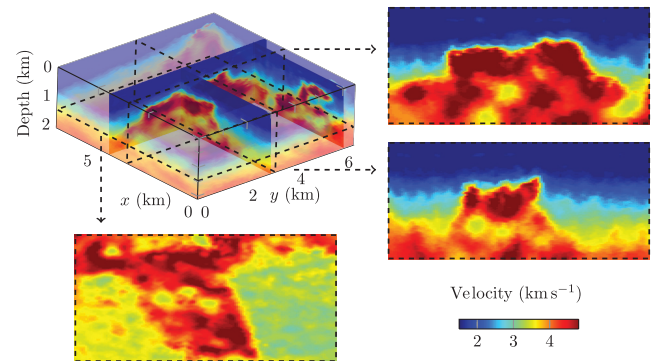

Figure 19. Velocity reconstruction of the SEAM model from the minimization of $\mathcal{J}_r$, starting with the models of Figure 17. It uses a dense observational acquisition of 272 point sources, whereas the numerical acquisition (g in equation 8) uses $n_{stack} = 34$ sources composed of $n_{pt} = 8$ points. The reconstruction follows 30 iterations per frequency between 2 and 7 Hz. For visualization, we picture sections at a fixed depth of $z = 1.5$ km (bottom), for $x = 3.5$ km (top right), and $y = 3.25$ km (bottom right).
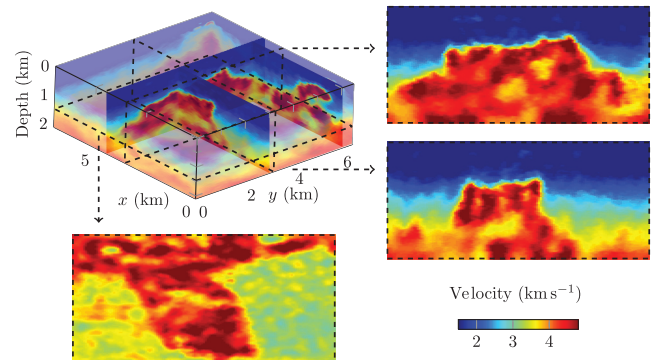

Figure 20. Velocity reconstruction of the SEAM model from the minimization of $\mathcal{J}_r$, starting with the models of Figure 17. It uses a sparse observational acquisition of $n_{stack} = 34$ sources composed of $n_{pt} = 8$ points, whereas the numerical acquisition (g in equation 8) uses a dense array of 272 point-sources. The reconstruction follows 30 iterations per frequency between 2 and 7 Hz. For visualization, we picture sections at a fixed depth of $z = 1.5$ km (bottom), for $x = 3.5$ km (top right), and $y = 3.25$ km (bottom right).

following question for future applications of the FRgWI method: How do we optimally select the numerical acquisition to make the best use of the observed measurements?

Data coverage

It is important to remember that the reciprocity-gap misfit functional relates, in terms of receivers, to the discretization of a surface integral, as depicted in Appendix A. In our experiments, we have a dense array of receivers, according to the standards in exploration geophysics. In the case in which the surface is not properly covered, we certainly need to appropriately weight the records, using, for instance, existing techniques from earth seismology, such as the one prescribed by Montagner et al. (2012). Similarly, the use of specific weight (e.g., from quadrature rules) could give a first indication on how to select the computational acquisition to ensure the equal illumination of the domain with minimal points.

Toward practical applications

Our reconstruction algorithm has been carried out in the frequency domain, in which it is difficult to address the scales encountered in field applications due to the operations of linear algebra (mainly the memory required for the matrix factorization). Although the efficiency of the frequency domain is improved by the use of HDG discretization (to reduce the size of linear systems) and by new techniques developed for the solver (Liu et al., 2020), time- or hybrid-domain inversion should be implemented to further probe the method in field studies. In this case, the multiplication of fields in the misfit functional equation 8 turns into convolution products, maintaining the same features regarding the flexibility of the numerical acquisition. We do not expect a difference in the performance of the method when implemented in the time domain for the situation we have studied here, especially if one follows the (usual) progression of increasing frequency content in the data, as introduced by Bunks et al. (1995). However, it is crucial to experiment on larger field configurations to further evaluate the method compared to the existing techniques. Here, the impact of the (limited) data coverage has to be carefully investigated.

CONCLUSION

We have performed waveform inversion using a reciprocity-gap misfit functional, which is adapted to the dual-sensor devices that capture the pressure field and the normal velocity of the waves. In this case, the measurements can be seen as Cauchy data, motivating our choice of misfit from the point of view of inverse scattering theory. Our method enables the use of different acquisitions for the observations and the simulations, by relying on a product of data. We have investigated its performance with sparse acquisitions made of multiple-point sources and have demonstrated its efficiency compared to the traditional FWI. We have used a sparse acquisition either for the observational one (to reduce the cost of the field acquisition) or for the computational one (to reduce the numerical cost of the iterative minimization), while not altering the resolution of the reconstruction.

We have implemented the method in the frequency domain and carried out two experiments of different natures. These represent the first steps of the validation, while time- and hybrid-domain inversions now should be considered. The implementation

of the misfit with time signals does not result in any technicality and should not lead to a difference in performance for the scales that we have studied with our frequency-domain setup. However, it is necessary to further probe the behavior of FRgWI in larger practical test cases and to analyze how it competes with state-of-the art methods.

The FRgWI method allows for arbitrary computational source positions that now need to be further analyzed. For instance, the definition of criteria for the design of the computational acquisition and the study of the source positions (e.g., not just close to the near surface area, with quadrature rules to ensure equal illumination) is the subject of our future research. Note also that the concept of FRgWI extends readily to elasticity.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.

APPENDIX A

RECIPROCITY-GAP FORMULA WITH EULER'S EQUATIONS IN SEISMIC

In this appendix, we motivate the reciprocity-gap misfit, constructed from the variational formulation of problem 1. In the context of seismic applications, we assume that the data (i.e., pressure and normal velocity) are acquired on a line Σ, which is slightly underneath the surface Γ_1. The domain is decomposed into above and below the area from the line of receivers, such that $\Omega = \Omega^+ \cup \Omega^-$, illustrated in Figure A-1.

Let us first rewrite problem 1 introducing continuity conditions at Σ, using exponent $+$ and $-$ for the fields that are in Ω^+ and Ω^-, respectively, such that

$$p|_{\Omega^+} = p^+, \quad p|_{\Omega^-} = p^-, \tag{A-1}$$

and similarly for the velocity. Problem 1 is equivalent to

$$\begin{cases} -i\omega\rho(\boldsymbol{x})\boldsymbol{v}(\boldsymbol{x}) = -\nabla p(\boldsymbol{x}), & \text{in } \Omega^+ \cup \Omega^-, & \text{(A-2a)} \\ i\omega\kappa(\boldsymbol{x})^{-1} p(\boldsymbol{x}) = -\nabla \cdot \boldsymbol{v}(\boldsymbol{x}) + f(\boldsymbol{x}), & \text{in } \Omega^+ \cup \Omega^-, & \text{(A-2b)} \\ p(\boldsymbol{x}) = 0, & \text{on } \Gamma_1, & \text{(A-2c)} \\ \partial_\nu p(\boldsymbol{x}) - \frac{i\omega}{c(\boldsymbol{x})} p(\boldsymbol{x}) = 0, & \text{on } \Gamma_2, & \text{(A-2d)} \\ p^+ = p^-, \partial_\nu p^+ = \partial_\nu p^-, & \text{on } \Sigma, & \text{(A-2e)} \\ \boldsymbol{v}^+ = \boldsymbol{v}^-, \partial_\nu \boldsymbol{v}^+ = \partial_\nu \boldsymbol{v}^-, & \text{on } \Sigma. & \text{(A-2f)} \end{cases}$$

We consider two couples $(p_1, \boldsymbol{v}_1)$ and $(p_2, \boldsymbol{v}_2)$ solutions to problem A-2. They are, respectively, associated to two sources, f and g, and two different sets of physical parameters: (κ_1, ρ_1) and (κ_2, ρ_2), respectively. We take $p_k \in H^1(\Omega)$ and $\boldsymbol{v}_k \in (H^1(\Omega))^3$, for $k = 1, 2$, where H refers to the Hilbert space. For the derivation of the reciprocity-gap formula, we write the variational formulation on Ω^- only. The variational formulation of equation A-2 for $(p_1, \boldsymbol{v}_1)$, using for the test functions $(p_2, \boldsymbol{v}_2)$ gives

$$\begin{cases} \int_{\Omega^-} -i\omega\rho_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + \nabla p_1 \cdot \boldsymbol{v}_2 d\Omega^- = 0, \\ \int_{\Omega^-} -i\omega\kappa_1^{-1} p_1 p_2 + (\nabla \cdot \boldsymbol{v}_1) p_2 d\Omega^- = 0, \end{cases} \tag{A-3}$$

where we omit the $-$ exponent in the fields p and $\boldsymbol{v}$ for the sake of clarity. Furthermore, we assume that the source f is supported outside Ω^-, for example, by using delta-Dirac functions in some position $\boldsymbol{x}_f \in \Omega^+$, according to Figure A-1.

We use an integration-by-parts for both equations, and we subtract the first one from the second one to get

$$\int_{\Omega^-} i\omega\rho_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + p_1 \nabla \cdot (\boldsymbol{v}_2) - i\omega\kappa_1^{-1} p_1 p_2 - \boldsymbol{v}_1 \cdot \nabla p_2 d\Omega^-$$
$$= \int_{\partial\Omega^-} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\partial\Omega^-. \tag{A-4}$$

In the volume integral, we replace $(\nabla \cdot \boldsymbol{v}_2)$ and (∇p_2) using the fact that $(p_2, \boldsymbol{v}_2)$ solves equation A-2

$$\int_{\Omega^-} i\omega\rho_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + p_1 \nabla \cdot (\boldsymbol{v}_2) - i\omega\kappa_1^{-1} p_1 p_2 - \boldsymbol{v}_1 \cdot \nabla p_2 d\Omega^-$$
$$= \int_{\Omega^-} i\omega\rho_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + p_1(i\omega\kappa_2^{-1} p_2) - i\omega\kappa_1^{-1} p_1 p_2 - \boldsymbol{v}_1 \cdot (i\omega\rho_2 \boldsymbol{v}_2) d\Omega^-$$
$$= \int_{\Omega^-} i\omega(\rho_1 - \rho_2) \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + i\omega(\kappa_2^{-1} - \kappa_1^{-1}) p_1 p_2 d\Omega^-. \tag{A-5}$$

Next, the integral on the boundary in equation A-4 is decomposed between $\Gamma_2^- = \Gamma_2 \cap \partial\Omega^-$ and Σ such that

$$\int_{\partial\Omega^-} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\partial\Omega^-$$
$$= \int_{\Gamma_2^-} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\Gamma_2^-$$
$$+ \int_{\Sigma} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\Sigma. \tag{A-6}$$

On Γ_2^-, we first use equation 1a to replace $(\boldsymbol{v}_1 \cdot \boldsymbol{\nu})$ and $(\boldsymbol{v}_2 \cdot \boldsymbol{\nu})$, and then the ABC equation 1d

$$\int_{\Gamma_2^-} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\Gamma_2^-$$
$$= \int_{\Gamma_2^-} p_1((i\omega\rho_2)^{-1} \partial_\nu p_2) - ((i\omega\rho_1)^{-1} \partial_\nu p_1) p_2 d\Gamma_2^-$$
$$= \int_{\Gamma_2^-} p_1((c_2\rho_2)^{-1} p_2) - ((c_1\rho_1)^{-1} p_1) p_2 d\Gamma_2^-$$
$$= \int_{\Gamma_2^-} ((c_2\rho_2)^{-1} - (c_1\rho_1)^{-1}) p_1 p_2 d\Gamma_2^-. \tag{A-7}$$

Eventually, injecting the new formulas for the volume and integral equations in equation A-4, we obtain

$$\int_{\Omega^-} i\omega(\rho_1 - \rho_2) \boldsymbol{v}_1 \boldsymbol{v}_2 + i\omega(\kappa_2^{-1} - \kappa_1^{-1}) p_1 p_2 d\Omega^-$$
$$- \int_{\Gamma_2^-} ((c_2\rho_2)^{-1} - (c_1\rho_1)^{-1}) p_1 p_2 d\Gamma_2^-$$
$$= \int_{\Sigma} p_1 (\boldsymbol{v}_2 \cdot \boldsymbol{\nu}) - (\boldsymbol{v}_1 \cdot \boldsymbol{\nu}) p_2 d\Sigma. \tag{A-8}$$

The right side coincides with our choice misfit functional for FRgWI see equation 8 (where we take the square of the expression to have a positive functional), and it equates to zero if and only if $\rho_1 = \rho_2$ and $\kappa_1 = \kappa_2$ (meaning that $c_1 = c_2$) over the whole domain Ω^-. We further refer to Alessandrini et al. (2018, 2019) for stability results. We see that the formula equation A-8 does not involve the sources because we have positioned them above the receivers' line (Figure A-1). In the case in which sources are below, or if one wants to create a numerical acquisition with sources inside the domain, it results in an additional integral in equation A-3, which then has to be included in the misfit.
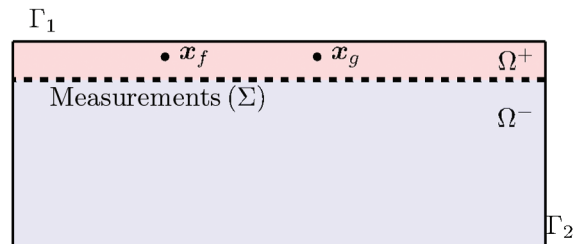

Figure A-1. Illustration of the domain of interest for the Euler's equation A-2, the line of measurements cuts the domain into $\Omega = \Omega^+ \cup \Omega^-$, which are colored in red and blue, respectively. We also consider the sources ($\boldsymbol{x}_f$ and $\boldsymbol{x}_g$) to be located in Ω^+.

APPENDIX B

GRADIENT FORMULATION USING ADJOINT-STATE METHOD

The gradient of the misfit functional is computed using the adjoint-state method, which comes from the work of Lions in optimal control (Lions, 1971), with early applications in Chavent (1974). The method is popular in seismic applications because it avoids the computation of the Jacobian matrix, thus requiring limited numerical effort; it is reviewed in Plessix (2006).

For generality, we consider the misfit functional $\mathcal{J}_\bullet$, which is either $\mathcal{J}_{L2}$ or $\mathcal{J}_r$, and define the constrained minimization problem

$$\min_m \mathcal{J}_\bullet(m) \quad \text{subject to} \quad \sum_i \mathcal{A}(U^{(f_i)}) = F_i,$$
$$\text{for } i \in \{1, \ldots, \mathrm{n_{src}}\}, \tag{B-1}$$

where $\mathcal{A}$ is the linear wave operator corresponding with the Euler's equations of problem 1. We denote by f_i the sources and use the notation $U^{(f_i)} = \{\boldsymbol{v}^{(f_i)}, p^{(f_i)}\} = \{v_x^{(f_i)}, v_y^{(f_i)}, v_z^{(f_i)}, p^{(f_i)}\}$ for the solution associated with the volume source $F_i = \{0, 0, 0, f_i\}$, in accordance with problem 1. For the sake of notation, we first consider a single source and drop the index f_i, the formulation with Lagrangian gives

$$\mathcal{L}(m, \tilde{U}, \tilde{\gamma}) = \mathcal{J}_\bullet + < \mathcal{A}\tilde{U} - F, \tilde{\gamma} >_\Omega . \tag{B-2}$$

Here, $< ., . >_\Omega$ denotes the complex inner product in $L^2(\Omega)$ such that $< a, b >_\Omega = \int_\Omega \overline{a} b \, d\Omega$, with — the complex conjugate.

The first step of the adjoint-state method is to take $\tilde{U} = U$ such that

$$\nabla_m \mathcal{L}(m, \tilde{U} = U, \tilde{\gamma}) = \nabla_m \mathcal{J}_\bullet. \tag{B-3}$$

Then, the adjoint state γ is selected such that the derivative of the Lagrangian with respect to U equates zero; that is, γ is the solution to

$$\mathcal{A}^* \gamma = -\partial_U \mathcal{J}_\bullet, \tag{B-4}$$

where $*$ denotes the adjoint (transposed of the complex conjugate). With this choice of adjoint state, the gradient of the misfit functional (which coincides with the one of the Lagrangian from equation B-3) is

$$\nabla_m \mathcal{J}_\bullet = \nabla_m \mathcal{L}(m, \tilde{U} = U, \tilde{\gamma} = \gamma) = \mathrm{Re}(< \partial_m \mathcal{A}U, \gamma >_\Omega). \tag{B-5}$$

We further refer readers to Faucher (2017) and Appendix A of Barucq et al. (2019, 2018) for more details on the complex-variable adjoint-state method and to Faucher and Scherzer (2020) for the specificity with HDG discretization.

When we incorporate back the different sources, the adjoint problem equation B-4 for the misfit functional $\mathcal{J}_{L2}$ is, for each source f_i in the acquisition,

$$\mathcal{A}^* \begin{pmatrix} \gamma_{v_x}^{(f_i)} \\ \gamma_{v_y}^{(f_i)} \\ \gamma_{v_z}^{(f_i)} \\ \gamma_p^{(f_i)} \end{pmatrix} = -\mathcal{R}^* \begin{pmatrix} \eta(\mathcal{R}(v_\nu^{(f_i)}) - d_v^{(f_i)})\boldsymbol{\nu}_x \\ \eta(\mathcal{R}(v_\nu^{(f_i)}) - d_v^{(f_i)})\boldsymbol{\nu}_y \\ \eta(\mathcal{R}(v_\nu^{(f_i)}) - d_v^{(f_i)})\boldsymbol{\nu}_z \\ \mathcal{R}(p^{(f_i)}) - d_p^{(f_i)} \end{pmatrix} \tag{B-6}$$

adjoint-state problem for $\mathcal{J}_{L2}$.

Here, we use $\boldsymbol{\nu}$ to indicate the normal direction (in the case of a flat surface, i.e., the $x - y$ plane, $\boldsymbol{\nu}_x = \boldsymbol{\nu}_y = 0$ and $\boldsymbol{\nu}_z = 1$). The gradient using all sources is

$$\nabla_m \mathcal{J}_{L2} = \mathrm{Re}\left(\sum_{i=1}^{n_{src}^{obs}} \langle \partial_m \mathcal{A}U^{(f_i)}, \gamma^{(f_i)} \rangle_\Omega \right). \tag{B-7}$$

For the misfit functional $\mathcal{J}_r$ of equation 8, the adjoint-state solves, for each computational source g_j,

$$\mathcal{A}^* \begin{pmatrix} \gamma_{v_x}^{(g_j)} \\ \gamma_{v_y}^{(g_j)} \\ \gamma_{v_z}^{(g_j)} \\ \gamma_p^{(g_j)} \end{pmatrix} = -\sum_{i=1}^{n_{src}^{obs}} \xi_{i,j} \mathcal{R}^* \begin{pmatrix} -\overline{d}_p^{(f_i)}\boldsymbol{\nu}_x \\ -\overline{d}_p^{(f_i)}\boldsymbol{\nu}_y \\ -\overline{d}_p^{(f_i)}\boldsymbol{\nu}_z \\ \overline{d}_v^{(f_i)} \end{pmatrix}$$

adjoint-state problem for $\mathcal{J}_r$,   (B-8)

with the scalar $\xi_{i,j}$ given by

$$\xi_{i,j} = \sum_{k=1}^{n_{rcv}} (d_v^{(f_i)}(\boldsymbol{x}_k) p^{(g_j)}(\boldsymbol{x}_k) - d_p^{(f_i)}(\boldsymbol{x}_k) v_\nu^{(g_j)}(\boldsymbol{x}_k)). \tag{B-9}$$

The gradient is

$$\nabla_m \mathcal{J}_r = \mathrm{Re}\left(\sum_{j=1}^{n_{src}^{sim}} \langle \partial_m \mathcal{A}U^{(g_j)}, \gamma^{(g_j)} \rangle_\Omega \right). \tag{B-10}$$

We refer readers to Alessandrini et al. (2019) for more details on the adjoint-state method using Cauchy data. Therefore, we see that the adjoint state for a single source in the computational acquisition for $\mathcal{J}_r$ encodes information from all measurements because of the reciprocity gap (the sum over i in equation B-8), whereas it takes only the current source with $\mathcal{J}_{L2}$ (equation B-6).

REFERENCES

Alessandrini, G., M. V. de Hoop, F. Faucher, R. Gaburro, and E. Sincich, 2019, Inverse problem for the Helmholtz equation with Cauchy data: Reconstruction with conditional well-posedness driven iterative regularization: ESAIM: Mathematical Modelling and Numerical Analysis, **53**, 1005–1030, doi: 10.1051/m2an/2019009.

Alessandrini, G., M. V. de Hoop, R. Gaburro, and E. Sincich, 2018, Lipschitz stability for a piecewise linear Schrödinger potential from local Cauchy data: Asymptotic Analysis, **108**, 115–149, doi: 10.3233/ASY-171457.

Amestoy, P. R., I. S. Duff, J.-Y. L'Excellent, and J. Koster, 2001, A fully asynchronous multifrontal solver using distributed dynamic scheduling: SIAM Journal on Matrix Analysis and Applications, **23**, 15–41, doi: 10.1137/S0895479899358194.

Amestoy, P. R., A. Guermouche, J.-Y. L'Excellent, and S. Pralet, 2006, Hybrid scheduling for the parallel solution of linear systems: Parallel Computing, **32**, 136–156, doi: 10.1016/j.parco.2005.07.004.

Bamberger, A., G. Chavent, and P. Lailly, 1977, Une application de la théorie du contrôle à un problème inverse de sismique: Annales de Géophysique, **33**, 183–200.

Bamberger, A., G. Chavent, and P. Lailly, 1979, About the stability of the inverse problem in the 1-D wave equation: Journal of Applied Mathematics and Optimisation, **5**, 1–47, doi: 10.1007/BF01442542.

Barucq, H., G. Chavent, and F. Faucher, 2019, A priori estimates of attraction basins for nonlinear least squares, with application to Helmholtz seismic inverse problem: Inverse Problems, **35**, 115004, doi: 10.1088/1361-6420/ab3507.

Barucq, H., F. Faucher, and H. Pham, 2018, Localization of small obstacles from back-scattered data at limited incident angles with full-waveform inversion: Journal of Computational Physics, **370**, 1–24, doi: 10.1016/j.jcp.2018.05.011.

Bednar, J. B., C. Shin, and S. Pyun, 2007, Comparison of waveform inversion — Part 2: Phase approach: Geophysical Prospecting, **55**, 465–475, doi: 10.1111/j.1365-2478.2007.00618.x.

Berkhout, A. J., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938, doi: 10.1190/1.2954035.

Bodin, T., H. Yuan, and B. Romanowicz, 2014, Inversion of receiver functions without deconvolution — Application to the Indian craton: Geophysical Journal International, **196**, 1025–1033, doi: 10.1093/gji/ggt431.

Bonnasse-Gahot, M., H. Calandra, J. Diaz, and S. Lanteri, 2017, Hybridizable discontinuous Galerkin method for the 2D frequency-domain elastic wave equations: Geophysical Journal International, **213**, 637–659, doi: 10.1093/gji/ggx533.

Bozdağ, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements: Geophysical Journal International, **185**, 845–870, doi: 10.1111/j.1365-246X.2011.04970.x.

Brossier, R., S. Operto, and J. Virieux, 2010, Which data residual norm for robust elastic frequency-domain full waveform inversion? Geophysics, **75**, no. 3, R37–R46, doi: 10.1190/1.3379323.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473, doi: 10.1190/1.1443880.

Carlson, D., W. Söllner, H. Tabti, E. Brox, and M. Widmaier, 2007, Increased resolution of seismic data from a dual-sensor streamer cable: 77th Annual International Meeting, SEG, Expanded Abstracts, 994–998, doi: 10.1190/1.2792572.

Chavent, G., 1974, Identification of functional parameters in partial differential equations, *in* Identification of parameters in distributed systems: American Society of Mechanical Engineers, 31–48.

Choi, Y., and T. Alkhalifah, 2011, Source-independent time-domain waveform inversion using convolved wavefields: Application to the encoded multisource waveform inversion: Geophysics, **76**, no. 5, R125–R134, doi: 10.1190/geo2010-0210.1.

Clément, F., G. Chavent, and S. Gómez, 2001, Migration-based traveltime waveform inversion of 2-D simple structures: A synthetic example: Geophysics, **66**, 845–860, doi: 10.1190/1.1444974.

Cockburn, B., J. Gopalakrishnan, and R. Lazarov, 2009, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems: SIAM Journal on Numerical Analysis, **47**, 1319–1365, doi: 10.1137/070706616.

Colton, D., and H. Haddar, 2005, An application of the reciprocity gap functional to inverse scattering theory: Inverse Problems, **21**, 383–398, doi: 10.1088/0266-5611/21/1/023.

Colton, D., and R. Kress, 1998, Inverse acoustic and electromagnetic scattering theory, 2nd ed.: Springer, Applied Mathematical Sciences.

de Hoop, M. V., and A. T. de Hoop, 2000, Wave-field reciprocity and optimization in remote sensing: Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, **456**, 641–682, doi: 10.1098/rspa.2000.0534.

Engquist, B., and A. Majda, 1977, Absorbing boundary conditions for numerical simulation of waves: Proceedings of the National Academy of Sciences, **74**, 1765–1766, doi: 10.1073/pnas.74.5.1765.

Faucher, F., 2017, Contributions to seismic full waveform inversion for time harmonic wave equations: Stability estimates, convergence analysis, numerical experiments involving large scale optimization algorithms: Ph.D. thesis, Universite de Pau et Pays de l'Ardour.

Faucher, F., G. Chavent, H. Barucq, and H. Calandra, 2020a, A priori estimates of attraction basins for velocity model reconstruction by time-harmonic full waveform inversion and data-space reflectivity formulation: Geophysics, **85**, no. 3, R223–R241, doi: 10.1190/geo2019-0251.1.

Faucher, F., and O. Scherzer, 2020, Adjoint-state method for Hybridizable Discontinuous Galerkin discretization: Application to the inverse acoustic wave problem: arXiv preprint arXiv:2002.06366.

Faucher, F., O. Scherzer, and H. Barucq, 2020b, Eigenvector models for solving the seismic inverse problem for the Helmholtz equation: Geophysical Journal International, **221**, 394–414, doi: 10.1093/gji/ggaa009.

Fichtner, A., B. L. Kennett, H. Igel, and H.-P. Bunge, 2008, Theoretical background for continental-and global-scale full-waveform inversion in the time-frequency domain: Geophysical Journal International, **175**, 665–685, doi: 10.1111/j.1365-246X.2008.03923.x.

Gauthier, O., J. Virieux, and A. Tarantola, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Numerical results: Geophysics, **51**, 1387–1403, doi: 10.1190/1.1442188.

Griesmaier, R., and P. Monk, 2011, Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation: Journal of Scientific Computing, **49**, 291–310, doi: 10.1007/s10915-011-9460-z.

Guasch, L., M. Warner, and C. Ravaut, 2019, Adaptive waveform inversion: Practice: Geophysics, **84**, no. 3, R447–R461, doi: 10.1190/geo2018-0377.1.

Jeong, W., H.-Y. Lee, and D.-J. Min, 2012, Full waveform inversion strategy for density in the frequency domain: Geophysical Journal International, **188**, 1221–1242, doi: 10.1111/j.1365-246X.2011.05314.x.

Karaoğlu, H., and B. Romanowicz, 2017, Global seismic attenuation imaging using full-waveform inversion: A comparative assessment of different choices of misfit functionals: Geophysical Journal International, **212**, 807–826, doi: 10.1093/gji/ggx442.

Kirby, R. M., S. J. Sherwin, and B. Cockburn, 2012, To CG or to HDG: A comparative study: Journal of Scientific Computing, **51**, 183–212, doi: 10.1007/s10915-011-9501-7.

Kirsch, A., 1996, An introduction to the mathematical theory of inverse problems: Springer, Applied Mathematical Sciences.

Kohn, R. V., and M. Vogelius, 1985, Determining conductivity by boundary measurements — Part 2: Interior results: Communications on Pure and Applied Mathematics, **38**, 643–667, doi: 10.1002/cpa.3160380513.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, no. 6, WCC177–WCC188, doi: 10.1190/1.3230502.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering: Theory and Application, 206–220.

Li, X., A. Y. Aravkin, T. van Leeuwen, and F. J. Herrmann, 2012, Fast randomized full-waveform inversion with compressive sensing: Geophysics, **77**, no. 3, A13–A17, doi: 10.1190/geo2011-0410.1.

Lions, J. L., 1971, Optimal control of systems governed by partial differential equations: Springer.

Liu, X., Y. Xi, Y. Saad, and M. V. de Hoop, 2020, Solving the three-dimensional high-frequency Helmholtz equation using contour integration and polynomial preconditioning: SIAM Journal on Matrix Analysis and Applications, **41**, 58–82, doi: 10.1137/18M1228128.

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653, doi: 10.1190/1.1443081.

Menke, W., and V. Levin, 2003, The cross-convolution method for interpreting SKS splitting observations, with application to one and two-layer anisotropic earth models: Geophysical Journal International, **154**, 379–392, doi: 10.1046/j.1365-246X.2003.01937.x.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, 2016, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society, **205**, 345–377, doi: 10.1093/gji/ggw014.

Montagner, J.-P., C. Larmat, Y. Capdeville, M. Fink, H. Phung, B. Romanowicz, E. Clévédé, and H. Kawakatsu, 2012, Time-reversal method and cross-correlation techniques by normal mode theory: A three-point problem: Geophysical Journal International, **191**, 637–652, doi: 10.1111/j.1365-246X.2012.05619.x.

Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228, doi: 10.1190/1.1442384.

Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: Mathematics of computation, **35**, 773–773, doi: 10.1090/S0025-5718-1980-0572855-7.

Nocedal, J., and S. J. Wright, 2006, Numerical optimization, 2nd ed.: Springer Series in Operations Research.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365-246X.2006.02978.x.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain — Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901, doi: 10.1190/1.1444597.

Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, doi: 10.1046/j.1365-246X.1998.00498.x.

Pratt, R. G., Z.-M. Song, P. Williamson, and M. Warner, 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 323–340, doi: 10.1111/j.1365-246X.1996.tb07023.x.

Pyun, S., C. Shin, and J. B. Bednar, 2007, Comparison of waveform inversion — Part 3: Amplitude approach: Geophysical Prospecting, **55**, 477–485, doi: 10.1111/j.1365-2478.2007.00619.x.

Rønholt, G., J. Lie, O. Korsmo, B. Danielsen, S. Brandsberg-Dahl, S. Brown, N. Chemingui, A. V. Mavilio, D. Whitmore, and R. Martinez, 2015, Broadband velocity model building and imaging using reflections, refractions and multiples from dual-sensor streamer data: Proceedings of the 14th International Congress of the Brazilian Geophysical Society and EXPOGEF, 1006–1009.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: Geophysical Journal International, **173**, 922–931, doi: 10.1111/j.1365-246X.2008.03768.x.

Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, no. 3, R31–R42, doi: 10.1190/1.2194523.

Shin, C., S. Pyun, and J. B. Bednar, 2007, Comparison of waveform inversion — Part 1: Conventional wavefield vs logarithmic wavefield: Geophysical Prospecting, **55**, 449–464, doi: 10.1111/j.1365-2478.2007.00617.x.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248, doi: 10.1190/1.1649391.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Tarantola, A., 1987a, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier. (Other Versions: 1988, 1994, 1998, 2002).

Tarantola, A., 1987b, Inversion of travel times and seismic waveforms, *in* Seismic tomography: Springer, 135–157.

Tenghamn, R., S. Vaage, and C. Borresen, 2007, A dual-sensor towed marine streamer: Its viable implementation and initial results: 77th Annual International Meeting, SEG, Expanded Abstracts, 989–993, doi: 10.1190/1.2792571.

Van Leeuwen, T., and W. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394, doi: 10.1111/j.1365-246X.2010.04681.x.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: Geophysics, **81**, no. 6, R429–R445, doi: 10.1190/geo2015-0387.1.

Whitmore, N., A. Valenciano, W. Sollner, and S. Lu, 2010, Imaging of primaries and multiples using a dual-sensor towed streamer: 80th Annual International Meeting, SEG, Expanded Abstracts, 3187–3192, doi: 10.1081/22020586.2010.12042051.

Yang, Y., B. Engquist, J. Sun, and B. F. Hamfeldt, 2018, Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion: Geophysics, **83**, no. 1, R43–R62, doi: 10.1190/geo2016-0663.1.

Zhang, Q., W. Mao, H. Zhou, H. Zhang, and Y. Chen, 2018, Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise: Geophysical Journal International, **215**, 1659–1681, doi: 10.1093/gji/ggy366.

Zhang, Q., H. Zhou, Q. Li, H. Chen, and J. Wang, 2016, Robust source-independent elastic full-waveform inversion in the time domain: Geophysics, **81**, no. 2, R29–R44, doi: 10.1190/geo2015-0073.1.

Zhong, Y., and Y. Liu, 2018, Some improvement strategies for the time-domain dual-sensor full waveform inversion: Arabian Journal of Geosciences, **11**, 346–357, doi: 10.1007/s12517-018-3687-z.

Zhong, Y., and Y. Liu, 2019, Source-independent time-domain vector-acoustic full-waveform inversion: Geophysics, **84**, no. 4, R489–R505, doi: 10.1190/geo2018-0304.1.