

Concepts in the Direct Waveform Inversion (DWI) Using Explicit Time-Space Causality

Yingcai Zheng^{1,*} and Zhonghan Liu¹

¹ Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, 77204, USA.

Received 10 October 2018; Accepted (in revised version) 12 September 2019

Abstract. The Direct Waveform Inversion (DWI) is a recently proposed waveform inversion idea that has the potential to simultaneously address several existing challenges in many full waveform inversion (FWI) schemes. A key ingredient in DWI is the explicit use of the time-space causality property of the wavefield in the inversion which allows us to convert the global nonlinear optimization problem in FWI, without information loss, into local linear inversions that can be readily solved. DWI is a recursive scheme which sequentially inverts for the subsurface model in a shallow-to-deep fashion. Therefore, there is no need for a global initial velocity model to implement DWI. DWI is unconditionally convergent when the reflection traveltime from the boundary of inverted model is beyond the finite recording time in seismic data. In order for DWI to work, DWI must use the full seismic wavefield including interbed and free surface multiples and it combines seismic migration and velocity model inversion into one process. We illustrate the concepts in DWI using 1D and 2D models.

AMS subject classifications: 34K29, 34L25, 74J20, 78A46

Key words: Waveform Inversion, full waveform inversion, direct waveform inversion, DWI, causality, 1D inversion, 2D inversion.

1 Introduction

The seismic full waveform inversion (FWI) [1,2] initially formulated in the time domain represents an important conceptual leap whose purpose is to find a subsurface model that can be used to predict observed seismic waveforms in both phase and amplitude, wiggle to wiggle. FWI can also be implemented in the frequency domain [3–6]. Despite some success, FWI mathematical formulation has significant physical limitations. The goal of this paper is to understand the cause of the limitations and propose a new formulation, called direct waveform inversion (DWI), to overcome these limitations. At present, DWI

*Corresponding author. *Email addresses:* yzheng12@uh.edu (Y. Zheng), lyuuu@hotmail.com (Z. Liu)

is by no means perfect but nevertheless it provides promising directions for waveform inversion.

Seismic FWI problem was frequently cast into a global nonlinear mathematical optimization problem where a model is sought to minimize a misfit/objective function defined between the observed data and the model-predicted data [7]. The nonlinearity arises because the change in seismic data in response to the change in the model parameter is not linear. In FWI, one linearizes the problem around a starting model and then computes the local gradient (e.g., the Frechet derivative) of the data perturbation with respect to the model perturbation and updates the model along the gradient direction. The updated model will be the next starting model and this process can be iterated until certain criterion about the misfit is met.

Challenges to implement FWI were almost immediately recognized since the inception of the FWI idea [1,2]. The first one is the initial-model dependence and convergence issue in the nonlinear global optimization. The second one is FWI's apparent lack of ability to recover low-wavenumber (large-scale) strong-contrast model variations.

FWI results strongly depend on the initial model that is usually not an outcome of FWI itself but is provided as an input to FWI. FWI works well if a good initial model in the neighborhood of the true model can be found in the beginning [8–11]. If the initial model is far from the true model, the FWI iteration may converge to a local minimum of the objective function and the global optimization cannot be attainable. Real geological models are likely to be complex. Demanding an initial model that is already close to the true model undercuts the true value of practical implementation of FWI. Fortunately, Kolb et al. [12] showed in numerical examples that if a coarsely smoothed version of the true model is available as the starting model, FWI could converge to the true model. This conclusion had been confirmed more recently [4, 13]. *Can FWI produce its own low-wavenumber initial modeling?* Within the gradient-based FWI theoretical framework, we need low frequency seismic data in order for FWI to recover the low-wavenumber model component [12, 14] because the FWI formulation/approximation is more linear/accurate at low frequencies. Seismic data in exploration settings are bandlimited. However, even without the low frequency data, the low-wavenumber model information is indeed contained in data and we can readily obtain it using many other methods such as traveltime tomography or the normal moveout analysis [15]. The inability of FWI to invert for low-wavenumber model variations relative to a simple starting model (e.g., homogeneous, or linear) shows the deficiency in FWI formulation.

Gradient-based FWI methods do not account for the full physics of wave scattering and propagation. Tarantola [16, p.128] pointed out that the FWI local Frechet gradient amounts to the linear single-scattering Born approximation. Recent work by Wu and Zheng [17] showed a one-to-one correspondence between the n -th order Frechet derivative and the n -th order multiple Born scattering. This means that FWI's dropping high-order functional derivatives is to physically ignore possible multiple scattering among unknown model perturbations/scatterers. This is a significant drawback in the FWI assumption. Wu and Zheng [17] further showed in numerical modeling that including

multiple forward scattering is critical to build correct traveltimes for transmission paths. Both transmitted and reflected waves have transmission paths. Therefore, it is critical to be able to take into account multiple scattering among model perturbations in waveform inversion. (Note that the single scattering here is referred to single interactions between the background wavefield with the unknown model perturbations. However, the background field can be multiply scattered by structure contained in the background model.) The Born approximation is accurate if the following two conditions are satisfied: the data frequency is low or if the size of the perturbation is small compared to the wavelength; and the unknown model perturbation is weak. The first condition can be fulfilled if the low-wavenumber component of the starting model matches the low-wavenumber component of the true model. Within the framework of FWI theoretic, there has been a significant amount of work to obtain a low-wavenumber initial model, which is then used as an input to FWI in sequel. This includes the envelope inversion [18–23], intensity inversion [24], waveform correlation function inversion [25], and Laplace FWI [26] and so on. There is also effort to extrapolate the bandwidth of the data to ultra-low frequency [27,28].

FWI formulation by no means is the only formulation for waveform inversion. There are also other efforts to directly consider the nonlinearity by incorporating multiple scattering in waveform inversion such as the T-matrix approach [29–35], or inverse scattering series [36–39].

Recently, Liu and Zheng [40,41] proposed a reflection direct waveform inversion (DWI) scheme by explicitly exploiting the causality of space-time wavefield in the inversion process as a constraint. As a consequence, *DWI converts a global nonlinear problem into many local linear inversion problems*. DWI does not require a global initial model. DWI first inverts for the shallow model and then recursively inverts for the deeper model. Unlike the FWI iterative process, this recursion process is always convergent and in the meantime it can fit the waveforms. DWI must use all wave types (primary reflections, internal and free surface multiples, etc.). DWI represents a different theoretical framework compared to gradient-based inversion methods. Our focus here is to expand and review the basic ideas of DWI using the time-space causality principle as an explicit constraint so this idea can be developed further by interested people.

2 Basic idea of DWI

2.1 DWI in 1D

The time-space causality concept of DWI can be simply understood in a 1D case along the z direction (Fig. 1). The model can be parameterized as a stack of horizontally stratified layers. We consider the acoustic case. Each layer is homogeneous and characterized by a sound speed, c_i , and a layer thickness, H_i , where the subscript indicates the layer number (Fig. 1). The pressure field is $P(z,t)$ and particle velocity field is $V_z(z,t)$, where z is the depth and t is time. $U(z,t)$ and $D(z,t)$ are upgoing and downgoing pressure fields at z ,

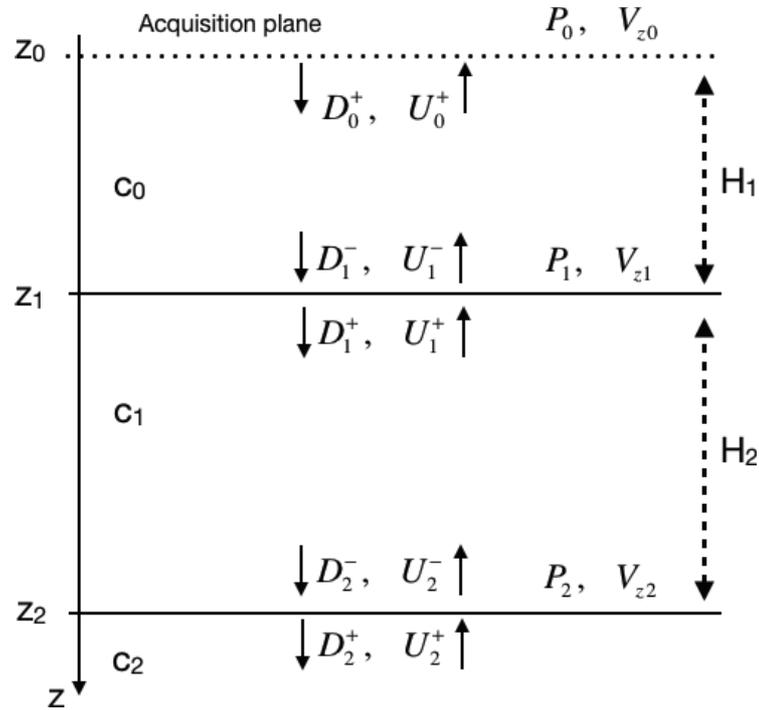


Figure 1: DWI process for a 1D layered model.

respectively. To simplify notations, we use $P_i = P_i(z, t)$, $V_{zi} = V_z(z_i, t)$, $D_i^\pm = D_i(z_i^\pm, t)$, and $U_i^\pm = U_i(z_i^\pm, t)$. z_i^+ and z_i^- represent depths slightly below and above z_i , respectively.

Our starting assumptions are:

- (i) The incident wave is a downgoing plane wave initiated at time zero at the acquisition plane z_0 . The wavelet is an impulse.
- (ii) Both the wave pressure P_0 and the vertical particle velocity V_{z0} are recorded. The recorded waves include all the primary reflections as well as the multiple reflections among all layers.
- (iii) The model parameters for all layers are unknown except the velocity of the top layer is known, c_0 .
- (iv) All layers have the same density, $\rho_i = \rho$.

Having both types of data (pressure and particle velocity) will allow us to separate the wavefield at a depth into upgoing and downgoing waves [42]. In many cases, we may have just one type of data. If only pressure data are available, it can be used to predict the particle velocity data [43–45]. Because we assume the availability of both pressure and particle velocity data, it does not matter whether the acquisition plane is a free surface or not. For now, we regard the first layer as an infinite half space.

For plane waves, we can decompose the pressure-velocity data, $P-V_z$, at a depth level into an upgoing U and a downgoing D pressure wave at the same depth via the following two relations:

$$P = D + U, \quad I \cdot V_z = D - U, \quad (2.1)$$

where I is the acoustic impedance, defined as the product of sound speed and density.

Now we describe the DWI inversion steps:

- (1) Using the recorded data, P_0 and V_{z0} , we can obtain the upgoing and downgoing pressure at z_0 , U_0^+ and D_0^+ , respectively;
- (2) If there is no free surface, there will be only one pulse in $D_0^+(t)$ which is the direct wave at time zero. However, the upgoing wave $U_0^+(t)$ may have many arrivals due to wave interaction with multiple layer interfaces. We know that the earliest arrival in $U_0^+(t)$ must be from the nearest reflector, which is at z_1 . We can certainly pick its traveltime, t_1 and amplitude, A_1 . Reflections from any other depths must come later than t_1 and this is the **time-space causality principle** which states a correspondence between *the earliest arrival in time* and *the nearest reflector in space*;
- (3) Since we know c_0 and have picked t_1 , we can compute the distance H_1 to obtain the depth of the first reflector z_1 : $H_1 = c_0 t_1 / 2$ and $z_1 = z_0 + H_1$;
- (4) We extrapolate the upgoing and downgoing pressure waves from z_0 to depth z_1^- and get U_1^- and D_1^- . D_1^- is a time-delayed version of D_0^+ and the amount of time delay is H_1/c_0 . On the other hand, U_1^- is a time-advanced version of which means we need to shift U_0^+ to the negative time direction to get U_1^- and the amount of time shift is $-H_1/c_0$. The amplitude ratio $A_1 = U_1^-(t_1)/D_1^-(t_1)$ is related to the impedance contrast between the two layers: $A_1 = (c_1 - c_0)/(c_0 + c_1)$ so we can obtain the velocity c_1 for the next layer;
- (5) From U_1^- and D_1^- and using Eq. (2.1), we can obtain the $P-V_z$ data at depth z_1^- , P_1 and V_{z1} . The $P-V_z$ data are continuous across the boundary so we can pass the $P-V_z$ data to depth z_1^+ ;
- (6) At the very top of the second layer z_1^+ , we can decompose P_1-V_{z1} data into upgoing and downgoing pressure waves, $U_1^+(t)$ and $D_1^+(t)$, using c_1 that has already been obtained in step (4);
- (7) Unlike $D_0^+(t)$ which has a single seismic event, the downgoing $D_1^+(t)$ now may have a train of events because the interface at z_1 can constantly reflect waves downward. The upgoing $U_1^+(t)$ should be composed of many seismic events too;
- (8) We can regard the downgoing $D_1^+(t)$ as an input signal and the upgoing $U_1^+(t)$ as an output. We then deconvolve $D_1^+(t)$ from $U_1^+(t)$ and obtain a reflection response of the medium ($z > z_1$ part and with a transparent boundary condition at z_1) due

to a downgoing impulsive plane wave initiated at z_1^+ . Here $D_1^+(t) * R(t) = U_1^+(t)$ where $*$ is the convolution operation. In the deconvolution, the response $R(t)$ must be causal, $R(t) = 0$ for $t < 0$. Now, we can loop this process recursively from step (4) to step (8) with a stopping criterion in the next step (9);

- (9) Because the seismic data recording time length is finite and each depth extrapolation (Step (4)) will shift the upgoing wave $U(t)$ to the negative time direction. At sufficiently large depth z , the upgoing wave will be all zero and this is the stopping criterion for DWI.

We can see that DWI does not need a global initial model to start with. It integrates imaging and inversion as a single step and it is recursive and always convergent. If the layers have different densities, we need to use several plane waves of different incident angles to resolve elastic properties of the underlying layers.

2.2 DWI in 2D and 3D models

In two- or three-dimension models, the DWI idea is essentially the same however the implementation is considerably more complex. We need to make two changes.

First, instead of using downgoing/upgoing waves in the 1D case, we use outgoing $W(\mathbf{x}, t)$ and ingoing $N(\mathbf{x}, t)$ waves (Fig. 2), respectively. The sense of outgoing or ingoing is with respect to a closed boundary. The acquisition plane S_0 with an outward normal

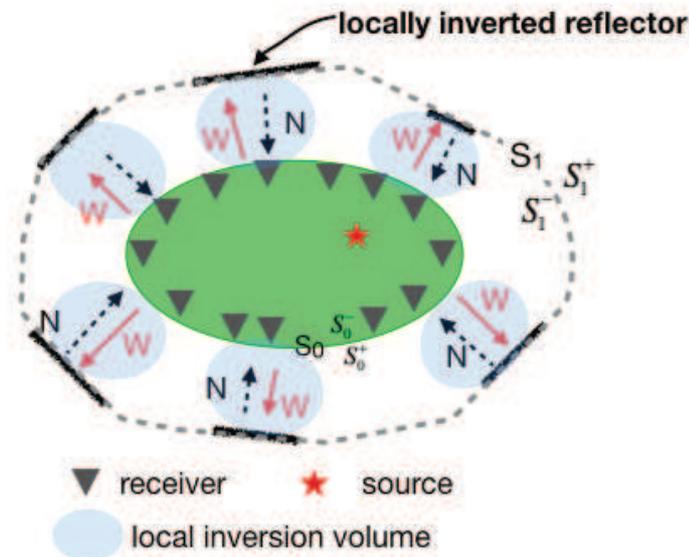


Figure 2: DWI process for a 2D (or 3D) model.

\mathbf{n} is closed and on S_0 we record both pressure $u(\mathbf{x}, t)$ and $\partial u(\mathbf{x}, t) / \partial n$ which is related to the particle velocity along the outward normal direction at $\mathbf{x} \in S_0$.

Based on the extinction theorem (46), we can compute the outgoing W and ingoing N waves on S_0 :

$$W(\mathbf{x}, t) = - \lim_{\varepsilon \rightarrow +0} \iint_{S_0} \left[u(\mathbf{x}', t) * \frac{\partial G(\mathbf{x} + \varepsilon \mathbf{n} | \mathbf{x}', t)}{\partial n(\mathbf{x}')} - G(\mathbf{x} + \varepsilon \mathbf{n} | \mathbf{x}', t) * \frac{\partial u(\mathbf{x}', t)}{\partial n(\mathbf{x}')} \right] dS_0(\mathbf{x}'), \quad (2.2)$$

$$N(\mathbf{x}, t) = - \lim_{\varepsilon \rightarrow +0} \iint_{S_0} \left[u(\mathbf{x}', t) * \frac{\partial G(\mathbf{x} - \varepsilon \mathbf{n} | \mathbf{x}', t)}{\partial n(\mathbf{x}')} - G(\mathbf{x} - \varepsilon \mathbf{n} | \mathbf{x}', t) * \frac{\partial u(\mathbf{x}', t)}{\partial n(\mathbf{x}')} \right] dS_0(\mathbf{x}'), \quad \mathbf{x} \in S_0, \quad (2.3)$$

where $G(\mathbf{x} | \mathbf{x}', t)$ is the Green's function between a receiving point \mathbf{x} and a source point \mathbf{x}' , the symbol $*$ is the time convolution, $\mathbf{n}(\mathbf{x}')$ is the outward normal at $\mathbf{x}' \in S_0$.

Secondly, to use the time-space causality, we need a pair of collocated source and receiver. We have obtained $W(\mathbf{x}, t)$ and $N(\mathbf{x}, t)$ for any point $\mathbf{x} \in S_0$. By studying the time lapse between ingoing and outgoing waves, we can obtain the position of the local reflector (Fig. 2). All reflectors form a surface S_1 . Since the velocity in the region between S_0 and S_1 is known, we can extrapolate the fields, W and N , from S_0 to S_1^- (interior side of S_1). The time-space causality stipulates that: the first/earliest pulse in the ingoing (reflected) wave $N(\mathbf{x} \in S_1, t)$ must be produced by the first/earliest pulse of $W(\mathbf{x} \in S_1, t)$ due to a reflector at $\mathbf{x} \in S_1$ (Fig. 2). The ratio of the two pulse amplitudes yields the velocity at the other side the reflector in S_1^+ . Up to this point, we have figured out the location of the reflector S_1 and the velocity on the other side (or the exterior) of S_1 . We can then recombine the outgoing and ingoing waves to pressure-particle-velocity data on S_1^- and pass them continuously to S_1^+ using boundary conditions. We can repeat this process and the domain of the inverted model is recursively enlarged until we exhausted the data due to their finite recording time.

3 Numerical examples

In this section, we show two numerical examples to illustrate the effectiveness of the DWI approach. The first example is for a 1D horizontally layered medium with plane wave incidence and reflection. The second one is for a 2D irregular piece-wise homogenous layered medium with a point source (i.e., a line source in 2D if embedded in the 3D space).

3.1 DWI for a 1D layered medium

Our first test model is a layered medium (Fig. 3). The layers are horizontally stratified and within each layer the medium is homogeneous. The layer thickness is 3 m. The medium density is constant throughout the model space. The model has a free surface

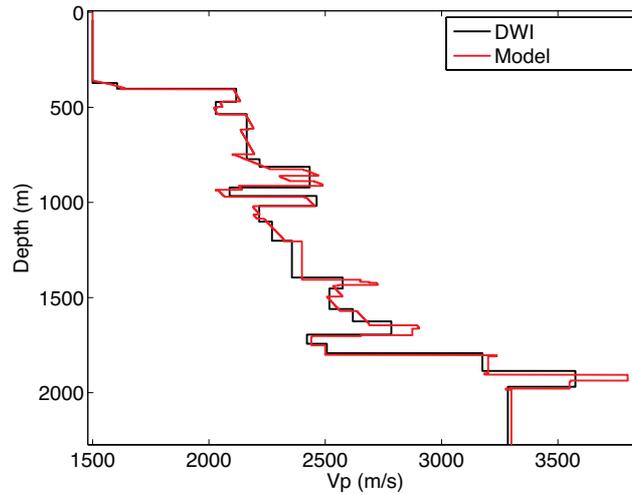


Figure 3: Comparison between the true model (red line) and the DWI inverted model (black line).

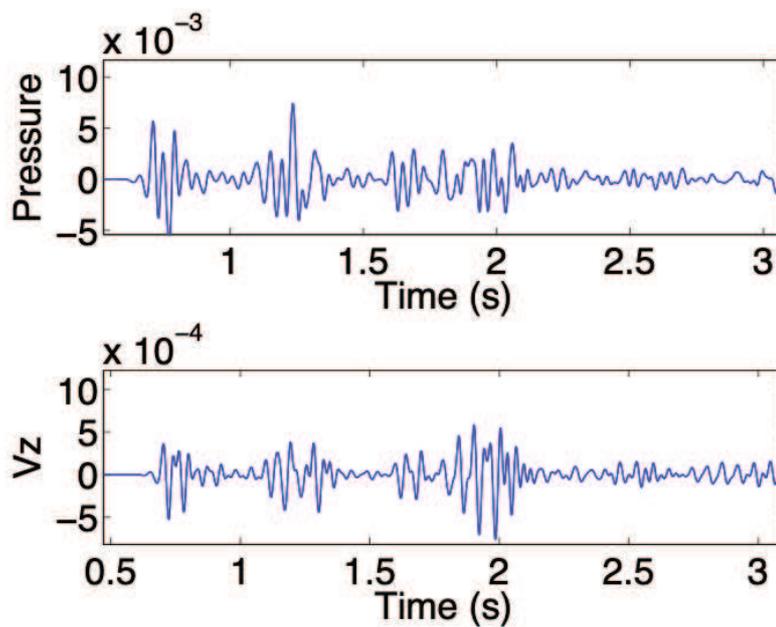


Figure 4: Recorded pressure, P , and vertical particle motion data, V_z .

boundary condition. The plane wave source depth is 20 m. The source wavelet is a Ricker of 20 Hz central frequency. Therefore, the wavelength is about 100 m, which is large compared to the 3-m layer thickness. The receiver is at 40 m depth, recording both pressure and vertical-component particle velocity waveforms (Fig. 4). The recorded waveforms

are complex due to the free-surface and internal multiples (Fig. 4). However, DWI can successfully recover the model sound speed profile (Fig. 3). In fact, DWI captures all physics and DWI inversion must use all wave types.

Because DWI uses the reflectivity model to invert velocities, DWI approximates a gradient layer with a constant-velocity layer. Therefore, DWI can invert for major discontinuities in the true model (Fig. 3). Because of this limitation of using reflectivity inversion, DWI is not able to produce perfect result for a model with many thin layers (3m in this case for wavelength ~ 100 m). Considering the results are obtained using only one plane wave, we expect the results can be improved with more plane waves of different incident angles. The other limitation is that if there are no physical reflections, DWI cannot produce an inversion result. However, given that DWI is a highly efficient algorithm, the DWI result can be used as a good starting model for FWI. This idea will be tested in future.

3.2 DWI for a 2D irregularly layered medium

The second test model is a layered model with irregular layers (Fig. 5a). The maximum velocity contrast is 200%. We used a 2-D acoustic finite difference method to model the full seismic data to be used in DWI. The model grid size is 5m and time step is 1ms. The density is constant throughout the model. We calculated wavefields for 3 shots. The 3 sources are all at depth 0.1km. Their horizontal locations are $X = 0.5$ km (Shot-1), 1.5km (Shot-2), and 2.5km (Shot-3), respectively. The source wavelet is a Ricker wavelet of 20Hz central frequency. A line of receivers, recording the pressure (P) and particle velocity (V_z),

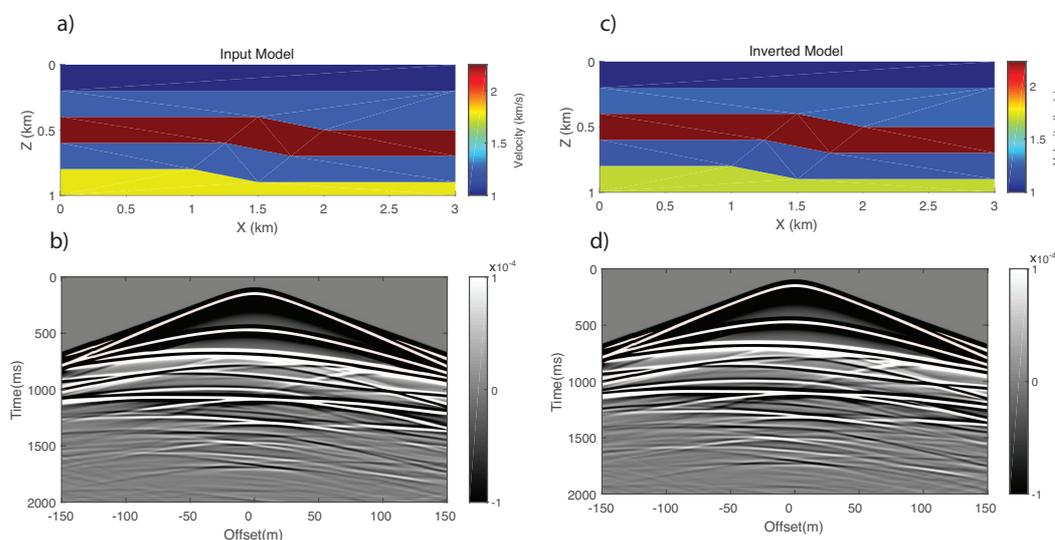


Figure 5: 2D DWI inversion. a) True model; b) recorded data for Shot-2 ($x = 1.5$ km, $z = 0.1$ km); c) DWI inverted model; d) modeled seismic gather using the DWI model in c) for Shot-2. In both b) and d), the direct waves are muted.

are placed in the first layer below the source, at a depth that is 0.19 km above the seafloor (i.e., the bottom of the first layer). The receiver interval is 5m from $X=0$ km to $X=3$ km. Both top and bottom of the model are half spaces and there is no free surface. A sample shot gather for Shot-2 is shown (Fig. 5b) with direct P -waves muted.

We make some remarks about the 2D DWI inversion result. First, 2D DWI again integrates imaging and inversion into a single process using time-space causality. The inverted model (Fig. 5c) contains both layer interfaces and velocity values of the layers. Secondly, the modeled shot gather (Fig. 5d) using the DWI-inverted model (Fig. 5c) is close to the recorded data (Fig. 5b) for Shot-2, where primary reflections, refracted arrivals, and multiples are all modeled.

4 Discussions and relations with other methods

As discussed in Introduction, the gradient-based full waveform inversion (FWI) is just one of many types of waveform inversion strategies. In FWI, interaction between the background wavefield and the unknown scatterer is treated as a single-scattering Born scattering which is valid if the size of the scatterer is small relative to the wavelength (implies low frequency wave) and the velocity perturbation is weak. As a consequence, FWI needs to start from an accurate global low-wavenumber model or needs low-frequency data. This also means that FWI is likely to succeed if the unmodeled scattered field (e.g., the difference between the observed data and the modeled data) is weak and can be modeled by the Born scattering.

In contrast, DWI builds on the explicit use of the time-space causality. The causality allows us to perform waveform inversion locally. It reduces the global nonlinear optimization to many local inversions. Therefore, a global starting model for DWI is not necessary. DWI builds the model in a shallow-to-deep fashion until all recorded data are used due to their finite recording length. It is interesting to note that Kolb, Collino and Lailly [12] found that "progressive downward determination of the velocity distribution" was better in overcoming the "local minimum" issue in FWI. While FWI is likely to succeed when the unmodeled field is weak, DWI works well if the unmodeled scattered field is a strong signal.

We also need to point out that the Marchenko-type methods also employ causality principle [47–58]. It is beyond the scope of the paper to give an in-depth review of the development in Marchenko inversion and imaging, which is an active research field. Instead, we refer interested readers to Kiraz and Nowack [59] for a review regarding the 1D inversion and to Lomas and Curtis [60] for higher dimensions. However, it is worth noting that Marchenko focusing needs an accurate velocity model to estimate the travel-time from the receiver depth to the focal point inside the model whereas DWI does not need a global initial velocity model to start the inversion using one-sided data.

There are many challenging issues in DWI to be solved in future. Since DWI inverts model for the shallow part first, how the error in the shallow part will affect inversion

for deeper structure needs further work to clarify and such error propagation should be model dependent. We may not be able to find the true model but at least we can test if the inverted model is correct or not using modeling and observation. If the observed data have no reflections, DWI will not work because it is built on reflectivity modeling while FWI can still work. In the real Earth, the wavefield is elastic and P and S waves can convert to each other. For P -to- P reflections, the time-space causality principle still holds but the application of the same principle to the S -to- S reflection requires us to isolate S waves which itself can be a challenge. We view DWI as a useful complement to FWI and the DWI output model may be a good input model for FWI to refine the model details.

5 Conclusions

We showed that DWI is a working strategy for performing waveform inversion. The numerical success for 1D and 2D models we have presented here show that DWI is a promising alternative framework to the conventional FWI formalism. In our numerical examples, we have shown that DWI inverts the model in a shallow-to-deep fashion and DWI does not require a global initial model to start with. DWI is unconditionally convergent and it is recursive rather than iterative. DWI combines imaging and inversion into a single step. In future, we will extend the DWI to models with more complex structure.

Acknowledgments

We appreciate the reviews from two anonymous reviewers. Many colleagues and friends have encouraged our development of the idea/concept of DWI. In particular, we want to express our sincere gratitude to Dr. Ru-Shan Wu, Yike Liu, and to Prof. Huawei Zhou for very thought-provoking discussions that have benefited us a lot. Without their encouragement, we would not be able to finish this paper. We are grateful to Professor M.Sen for his generosity to provide the 1d model in Fig. 3 to us. We thank the U.S. National Science Foundation for partially supporting the work (funding numbers: EAR-1388058 and EAR-1621878).

References

- [1] P. Lailly, The Seismic Inverse Problem as a Sequence of Before-Stack Migrations, Conference on Inverse Scattering: Theory and Applications, SIAM, 206-220,(1983).
- [2] A. Tarantola, Inversion of seismic-reflection data in the acoustic approximation. *Geophysics* 49, 1259-1266 (1984).
- [3] Q. B. Liao, G. A. McMechan, Multifrequency viscoacoustic modeling and inversion. *Geophysics* 61, 1371-1378 (1996).
- [4] L. Sirgue, R. G. Pratt, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics* 69, 231-248 (2004).

- [5] R. G. Pratt, R. M. Shipp, Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data. *Geophysics* 64, 902-914 (1999).
- [6] R. G. Pratt, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics* 64, 888-901 (1999).
- [7] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74, WCC1-WCC26 (2009).
- [8] A. Bourgeois, B. F. Jiang, P. Lailly, Linearized inversion: a significant step beyond pre-stack migration. *Geophys. J. Int.* 99, 435-445 (1989).
- [9] P. Mora, Nonlinear twodimensional elastic inversion of multioffset seismic data. *Geophysics* 52, 1211-1228 (1987).
- [10] O. Gauthier, J. Virieux, A. Tarantola, Two-dimensional nonlinear inversion of seismic waveforms - numerical results. *Geophysics* 51, 1387-1403 (1986).
- [11] A. Tarantola, A strategy for nonlinear elastic inversion of seismic-reflection data. *Geophysics* 51, 1893-1903 (1986).
- [12] P. Kolb, F. Collino, P. Lailly, Pre-stack inversion of a 1-d medium. *Proceedings of the Ieee* 74, 498-508 (1986).
- [13] R. G. Pratt, Waveform Tomography - Successes, cautionary tales, and future directions, In Workshop 11 "Full Waveform inversion - current status and perspectives", 70th EAGE Conference and Exhibition (2008).
- [14] C. Bunks, F. M. Saleck, S. Zaleski, G. Chavent, Multiscale seismic waveform inversion. *Geophysics* 60, 1457-1473 (1995).
- [15] P. Mora, Inversion = migration + tomography. *Geophysics* 54, 1575-1586 (1989).
- [16] A. Tarantola, Inverse problem theory and methods for model parameter estimation. (SIAM, 2005), pp. 342.
- [17] R.S. Wu, and Y. Zheng, Non-linear partial derivative and its De Wolf approximation for non-linear seismic inversion, *Geophys. J. Int.*, 196(3), 1827-1843 (2014).
- [18] R.S. Wu, J. Luo, B. Wu, R. S. Wu, J. R. Luo, B. Y. Wu, Ultra-low-frequency information in seismic data and envelope inversion, 83rd Annual International Meeting, SEG, Expanded Abstracts, 3078-3082 (2013).
- [19] R. S. Wu, J. R. Luo, B. Y. Wu, Seismic envelope inversion and modulation signal model. *Geophysics* 79, WA13-WA24 (2014).
- [20] J. R. Luo, R. S. Wu, Seismic envelope inversion: reduction of local minima and noise resistance. *Geophysical Prospecting* 63, 597-614 (2015).
- [21] G. X. Chen, R. S. Wu, S. C. Chen, Reflection multiscale envelope inversion. *Geophysical Prospecting*, 66, 1258-1271 (2018).
- [22] Y. Q. Wang, R. S. Wu, G. X. Chen, Z. M. Peng, Seismic modulation model and envelope inversion with smoothed apparent polarity. *Journal of Geophysics and Engineering* 15, 2278-2286 (2018).
- [23] E. Bozdog, J. Trampert, J. Tromp, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophys. J. Int.* 185, 845-870 (2011).
- [24] Y. Liu, B. He, H. Lu, Z. Zhang, X. Xie, and Y. Zheng (2018), Full intensity waveform inversion, *Geophysics*, 86(6), R649-R658 (2018).
- [25] T. van Leeuwen, W. A. Mulder, Velocity analysis based on data correlation. *Geophysical Prospecting* 56, 791-803 (2008).
- [26] W. Ha, C. Shin, Why do Laplace-domain waveform inversions yield long-wavelength results? *Geophysics* 78, R167-R173 (2013).
- [27] Y. E. Li, L. Demanet, Full-waveform inversion with extrapolated low-frequency data. *Geo-*

- physics 81, R339-R348 (2016).
- [28] X.-B. Xie, Recover certain low-frequency information for full waveform inversion, 83rd Annual International Meeting, SEG, Expanded Abstracts, 1053-1057 (2013).
 - [29] M. Jakobsen, B. Ursin, Full waveform inversion in the frequency domain using direct iterative T-matrix methods. *Journal of Geophysics and Engineering* 12, 400-418 (2015).
 - [30] M. Jakobsen, R. Wu, T-matrix representation of the De Wolf series for modeling and inversion in strongly scattering media, 85th Annual International Meeting, SEG, Expanded Abstracts, 1280-1285 (2015).
 - [31] R. S. Wu, B. F. Wang, C. H. Hu, Renormalized nonlinear sensitivity kernel and inverse thin-slab propagator in T-matrix formalism for wave-equation tomography. *Inverse Problems* 31 (2015).
 - [32] R.-S. Wu, B. Wang, M. Jakobsen, Green's function and T-matrix reconstruction using surface data for direct nonlinear inversion, 85th Annual International Meeting, SEG, Expanded Abstracts, 1286-1291 (2015).
 - [33] B.-F. Wang, R.-S. Wu, X.-H. Chen, W.-K. Lu, Non-linear parameter estimation method based on T-matrix. *Chinese Journal Geophysics* 59, 2257-2265 (2016).
 - [34] M. Jakobsen, R.-S. Wu, Multiscale T-matrix completion method for FWI in the absence of a good starting model, 87th Annual International Meeting, SEG, Expanded Abstracts, 1274-1279 (2017).
 - [35] M. Jakobsen, R. S. Wu, Accelerating the T-matrix approach to seismic full-waveform inversion by domain decomposition. *Geophysical Prospecting* 66, 1039-1059 (2018).
 - [36] R. Eftekhar, H. Hu, Y. Zheng, Convergence acceleration in scattering series and seismic waveform inversion using nonlinear Shanks transformation. *Geophys. J. Int.* 214, 1732-1743 (2018).
 - [37] R. Schlottmann, Direct waveform inversion via iterative inverse propagation, 57th Annual International Meeting, SEG, Expanded Abstracts, 2146-2150 (2006).
 - [38] J. Yao, A. C. Lesage, B. G. Bodmann, F. Hussain, D. J. Kouri, One dimensional acoustic direct nonlinear inversion using the Volterra inverse scattering series. *Inverse Problems* 30 (2014).
 - [39] J. Yao, A. C. Lesage, B. G. Bodmann, F. Hussain, D. J. Kouri, Inverse scattering theory: Inverse scattering series method for one dimensional non-compact support potential. *Journal of Mathematical Physics* 55 (2014).
 - [40] Z. Liu, and Y. Zheng, Direct waveform inversion, 85th Annual International Meeting, SEG, Expanded Abstracts, 1268-1273 (2015).
 - [41] Z. Liu, Y. Zheng, Applications of the direct-waveform inversion on 2D models, 87th Annual International Meeting, SEG, Expanded Abstracts, 1687-1691 (2017).
 - [42] M. M. Backus, Water reverberations—their nature and elimination. *Geophysics* 24, 233-261 (1959).
 - [43] L. Amundsen, Wavenumber-based filtering of marine point-source data. *Geophysics* 58, 1335-1348 (1993).
 - [44] Y. Zheng, Retrieving the exact Green's function by wavefield crosscorrelation. *J. Acoust. Soc. Am.* 127, EL93-EL98 (2010).
 - [45] L. Amundsen, B. G. Secest, B. Arntsen, Extraction of the normal component of the particle velocity from marine pressure data. *Geophysics* 60, 212-222 (1995).
 - [46] J. Zhang, A. Weglein, Extinction theorem deghosting method using towed streamer pressure data: analysis of the receiver array effect on deghosting and subsequent free surface multiple removal, 56th Annual International Meeting, SEG, Expanded Abstracts, 2095-2098 (2005).
 - [47] R. G. Newton, Inverse Scattering Iv. Three Dimensions - Generalized Marchenko Construc-

- tion with Bound-States, and Generalized Gelfand-Levitan Equations. *Journal of Mathematical Physics* 23, 594-604 (1982).
- [48] K. Wapenaar, F. Broggini, E. Slob, R. Snieder, Three-Dimensional Single-Sided Marchenko Inverse Scattering, Data-Driven Focusing, Green's Function Retrieval, and their Mutual Relations. *Physical Review Letters* 110 (2013).
- [49] C. A. da Costa, M. Ravasi, A. Curtis, G. A. Meles, Elastodynamic Green's function retrieval through single-sided Marchenko inverse scattering. *Phys Rev E* 90 (2014).
- [50] K. Wapenaar, J. Thorbecke, J. van der Neut, F. Broggini, E. Slob, and R. Snieder, Marchenko imaging. *Geophysics* 79, WA39-WA57 (2014).
- [51] G. A. Meles, K. Loer, M. Ravasi, A. Curtis, C. A. da Costa, Internal multiple prediction and removal using Marchenko autofocusing and seismic interferometry. *Geophysics* 80, A7-A11 (2015).
- [52] K. Wapenaar, J. van der Neut, E. Slob, On the role of multiples in Marchenko imaging. *Geophysics* 82, A1-A5 (2017).
- [53] R. L. Nowack, M. S. R. Kiraz, Virtual Greens Functions Using Seismic Interferometry and Marchenko Redatuming. *Seismol. Res. Lett.* 89, 613-619 (2018).
- [54] J. H. Rose, 'Single-sided' autofocusing of sound in layered materials. *Inverse Problems* 18, 1923 (2002).
- [55] J. H. Rose, Time Reversal, Focusing and Exact Inverse Scattering, in *Imaging of Complex Media with Acoustic and Seismic Waves*, M. Fink, W. A. Kuperman, J.-P. Montagner, A. Tourin, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002), pp. 97-106.
- [56] J. A. Ware, K. Aki, Continuous and Discrete Inverse Scattering Problems in a Stratified Elastic Medium. I. Plane Waves at Normal Incidence. *J. Acoust. Soc. Am.* 45, 911-921 (1969).
- [57] R. Burridge, The Gelfand-Levitan, the Marchenko, and the Gopinath-Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse-response problems. *Wave Motion* 2, 305-323 (1980).
- [58] F. Broggini, R. Snieder, K. Wapenaar, Focusing the wavefield inside an unknown 1D medium: Beyond seismic interferometry. *Geophysics* 77, A25-A28 (2012).
- [59] M. S. R. Kiraz, R. L. Nowack, Marchenko redatuming and imaging with application to the Frio carbon sequestration experiment. *Geophys. J. Int.* 215, 1633-1643 (2018).
- [60] A. Lomas, A. Curtis, An introduction to Marchenko methods for imaging. *Geophysics* 84, F35-F45 (2018).