

Full-waveform inversion via source-receiver extension

Guanghai Huang¹, Rami Nammour², and William Symes¹

ABSTRACT

Full-waveform inversion produces highly resolved images of the subsurface and quantitative estimation of seismic wave velocity, provided that its initial model is kinematically accurate at the longest data wavelengths. If this initialization constraint is not satisfied, iterative model updating tends to stagnate at kinematically incorrect velocity models producing suboptimal images. The source-receiver extension overcomes this “cycle-skip” pathology by modeling each trace with its own proper source wavelet, permitting a good data fit throughout the inversion process. Because source wavelets should be constant (or vary systematically) across a shot gather, a measure of source trace dependence, for example, the mean square of the signature-deconvolved wavelet scaled by time lag, can be minimized to update the velocity

model. For kinematically simple data, such measures of wavelet variance are mathematically equivalent to traveltime misfit. Thus, the model obtained by source-receiver extended inversion is close to that produced by traveltime tomography, even though the process uses no picked times. For more complex data, in which energy travels from source to receiver by multiple raypaths, Green’s function spectral notches may lead to slowly decaying trace-dependent wavelets with energy at time lags unrelated to traveltime error. Tikhonov regularization of the data-fitting problem suppresses these large-lag signals. Numerical examples suggest that this regularized formulation of source-receiver extended inversion is capable of recovering reasonably good velocity models from synthetic transmission and reflection data without stagnation at suboptimal models encountered by standard full-waveform inversion, but with essentially the same computational cost.

INTRODUCTION

Full-waveform inversion (FWI) estimates subsurface structure with high precision by minimizing the differences between the synthesized and recorded data in the least-squares sense (Tarantola, 1984; Virieux and Operto, 2009). However, the domain of convexity of the FWI objective function for velocity estimation is generally quite small, on the order of a wavelength in diameter, and iterative optimization methods starting further from the global minimizer may stagnate at physically meaningless apparent optima. The root cause of this behavior is the tendency of the predicted data to be out of phase with, or even orthogonal to, the recorded data in large regions of model space (cycle skipped), and therefore very far away in the mean-square sense. This problem may be avoided to some extent by a combination of initial model accuracy, high signal-to-noise ratio at the lowest recorded frequencies (achieved in some surveys, not in others), and data fitting in expanding frequency bands, from low to high (Bunks et al., 1995; Pratt, 1999; Pratt and Shipp, 1999; Sirgue

and Pratt, 2004; Virieux and Operto, 2009). The stagnation problem is somewhat less severe for fitting of refracted energy (diving or transmitted waves), as noted already by Gauthier et al. (1986).

This paper describes an alternative to least-squares data fitting and illustrates its behavior with several synthetic examples. The basis of this approach is the source-receiver extension of constant density acoustic modeling: An independent source trace is provided for each data trace. An appropriate choice of the source trace fits the data trace, for any choice of velocity field. Because the mean-square error is small for all such extended models (velocity plus source traces), some other objective must take over the role of fit error in standard FWI and drive the extended model toward a physical (non-extended) model that explains the data, hence solving the FWI problem. All traces in a common-source gather should share the same source, so it is natural to penalize the deviation from a common source. Several penalty functions for measuring this deviation are available. Because only one extended source matches the data for each velocity model, any such penalty is implicitly a function of

Manuscript received by the Editor 8 June 2016; revised manuscript received 7 November 2016; published online 21 March 2017.

¹Rice University, Department of Computational and Applied Mathematics, Houston, Texas, USA. E-mail: gh Huang@rice.edu; symes@caam.rice.edu.

²Total E&P R&T, Houston, Texas, USA. E-mail: rami.nammour@total.com.

© 2017 Society of Exploration Geophysicists. All rights reserved.

velocity, and it may be used as an objective function in an iterative optimization algorithm. If the iteration succeeds in driving the extension penalty to zero, then the process achieves an FWI: The iterates fit the data well throughout the inversion process and converge to a physical model.

The main objective of this paper is to demonstrate two facts about source-receiver extended waveform inversion. First, for transmission data with single arrivals (i.e., a unique raypath between each source and receiver), the objective function of source-receiver extended waveform inversion is approximately proportional to the mean-square traveltimes error. It follows that under these circumstances, the iterative inversion as envisioned in the preceding paragraph succeeds in finding a (physical) model that fits the data, that is, achieves an FWI. This process produces roughly the same model as does traveltimes tomography: In effect, it performs tomography with waveform data. The inversion converges from initial models that are hopelessly cycle skipped for standard data-domain FWI, as does traveltimes tomography.

Second, for transmission data with multiple arrivals (i.e., more than one distinct raypath connecting at least some source-receiver pairs), the connection between traveltimes tomography and source-receiver waveform inversion is broken, and the source-receiver objective function exhibits the same tendency to stagnate at nonoptimal solutions as does FWI. This observation applies in particular to inversion of diving wave data: If the diving waves are triplicated, source-receiver extended inversion is no more likely to converge to a useful velocity model than is FWI. We offer an explanation of this phenomenon and numerical examples that demonstrate it. In brief, multiple arrivals with certain amplitude relations among the branches lead to slowly decaying components in the extended source, contributing energy at large lags having nothing to do with the overall traveltimes error between the predicted and observed data traces.

We suggest a partial remedy for this misbehavior: From the spectral point of view, it arises from small eigenvalues of the normal operator or Hessian, equivalent to notches in Green's function spectrum. As is routine with other ill-conditioned inverse problems, one can suppress the effect of small eigenvalues by Tikhonov regularization (Engl et al., 1996) — in other words, prewhitening deconvolution by the Green's function. Provided with an adequate regularization weight parameter, this regularized source-receiver extended inversion tends to recover the convexity exhibited by the unregularized method in the single-arrival case, and it converges to an approximation to the global minimizer. A large regularization weight is required to achieve this goal, causing substantial data misfit (up to 50% in several of our examples). We illustrate all of these claims with numerical examples. The regularization weights used in the examples were chosen by means of a discrepancy principle, that is, setting them, so that the initial data misfit using the regularized extended model is within limits proportional to the initial physical model misfit. The necessary proportion of the initial misfit appears to be substantial — we have used 20%–50% in our examples.

Although our theoretical developments pertain to transmission data, the source-receiver extension applies also to the reflection configuration. For completeness, we include an example in which a salt lens is recovered from reflection data via source-receiver inversion. Rather heavy regularization seems to be required in this case as well.

The source-receiver extension concept is not new: Song and Symes (1994), Symes (1994), Plessix et al. (2000), Plessix (2000), Pratt and Symes (2002), and Luo and Sava (2011) investigate data

fitting via source-receiver extension to enhance the convergence of FWI. Warner and Guasch (2014, 2016) use a very similar approach as part of adaptive waveform inversion (AWI) and shows its capacity to enlarge the domain of attraction for FWI and its practicality for application to contemporary 3D field surveys. Of several possible choices of penalty, we use the dispersion about zero lag suggested in several of these works (Plessix, 2000; Luo and Sava, 2011; Warner and Guasch, 2014, 2016). The relation between source-receiver extended waveform inversion and traveltimes tomography is explained by Song and Symes (1994) in the context of crosswell waveform tomography. Symes (1994) shows that this link is generally broken when multiple arrivals are present in transmission data, using an argument based on causal deconvolution. Neither of these older works actually implemented source-receiver extended inversion, as we do here. Plessix et al. (2000) and Plessix (2000) apply the source-receiver extension approach to field crosswell data. They observe the effect described by Symes (1994), and avoided it by using source-receiver pairs with sufficiently different depths to avoid most guided wave energy. Tikhonov regularization is used in many inversion algorithms, including some of the source-receiver type mentioned above, but our use of regularization specifically to control the slowly decaying energy produced by multiple arrivals seems to be new. Several authors (Luo and Sava, 2011; Warner and Guasch, 2014, 2016) have applied source-receiver extension methods to reflection configurations.

It should be understood that many factors could lead to failure of source-receiver extension to yield an objective with a large domain of convexity about its global minimizer. To name just a few such factors, reflected waves in transmission data, or multiply reflected waves for reflection data, or out-of-plane reflections for 2D data, or shear or converted waves are all potentially capable of causing failure. However, we emphasize that even in the complete absence of these other factors, with data that actually arises in transmission through a slowly varying background, the presence of multiple energetic arrivals is sufficient to derail this approach. As our third example underlines, this conclusion applies particularly to diving wavefields (the principal data of conventional FWI), which may easily contain multiple arrivals from a localized source.

In the following pages, we first review the theoretical foundation of source-receiver extended waveform inversion, and we explain the structure of the algorithm. Then, we illustrate its behavior with four 2D numerical examples. The first two are set in an idealized crosswell geometry, and they illustrate the capability of source-receiver extended waveform inversion to provide tomographic-quality solutions, the obstacle to velocity updating posed by multiple arrivals, and a partial remedy through Tikhonov regularization. The other two examples use surface acquisition geometry. The first of these is a pure diving wave problem, with no reflections. The target model generates multiple raypaths in the diving wavefield, with the same damaging effect on velocity updating via source-receiver extended waveform inversion as in the previous crosswell example. Tikhonov regularization, with a regularization weight chosen by trial and error, suffices for these examples to restore convergence to a useful model. Finally, we include an example of surface data inversion via source-receiver extension, in which a model of a salt lens embedded in sediments is recovered from reflection data. All examples begin with homogeneous or simple layered initial guesses. In all cases, attempted FWI fails, whereas the regularized source-receiver extended inversion succeeds. In the “salt” example, the inclusion

emerges from the background without any special effort. All except the first example require regularization, with the regularization weight being estimated by trial and error.

We end with a discussion of several obvious or not-so-obvious capabilities and limitations of source-receiver extended waveform inversion.

THEORY

The acoustic model of seismic wave propagation treats the excess pressure field u as the solution of the acoustic-wave equation:

$$\frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = \delta(\mathbf{x} - \mathbf{x}_s) f(t), \quad (1)$$

$$u = 0, \quad t \ll 0. \quad (2)$$

The energy source is modeled here as a point isotropic radiator with source pulse $f(t)$. The forward-modeling operator maps the source to the data traces, presumed to be perfect pressure measurements, and depends on the wave velocity:

$$S[v]f(\mathbf{x}_r, t; \mathbf{x}_s) = u(\mathbf{x}_r, t; \mathbf{x}_s). \quad (3)$$

In equations 1 and 3, the source and receiver locations, \mathbf{x}_s and \mathbf{x}_r , respectively, define the acquisition geometry of the survey.

The standard least-squares inversion, or FWI, problem is: Given $f(t)$ and $d(\mathbf{x}_r, t; \mathbf{x}_s)$, determine $v(\mathbf{x})$, so as to minimize

$$J_{\text{FWI}}[v] = \frac{1}{2} \|S[v]f - d\|^2. \quad (4)$$

The vertical bars denote the mean square or l^2 -norm squared, in other words, the sum over all active values of \mathbf{x}_r , t , and \mathbf{x}_s , possibly scaled by cell volume or other factors.

As mentioned in the ‘‘Introduction’’ section, the function defined in equation 4 is difficult to minimize directly, so we will explore an alternative approach to its minimization through models that violate at least some of the modeling assumptions made above.

Source-receiver extended modeling

The source-receiver extension introduces a trace-dependent source function $\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s)$ to replace $f(t)$. The extended acoustic system is

$$\frac{1}{v^2} \frac{\partial^2 \tilde{u}}{\partial t^2} - \nabla^2 \tilde{u} = \delta(\mathbf{x} - \mathbf{x}_s) \tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s), \quad (5)$$

$$\tilde{u} = 0, \quad t \ll 0, \quad (6)$$

and the extended forward-modeling operator is defined by sampling the extended pressure field \tilde{u} as before:

$$\tilde{S}[v]\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s) = \tilde{u}(\mathbf{x}_r, t; \mathbf{x}_s). \quad (7)$$

Because the extended source is trace dependent, it is straightforward to fit the data, which is not the case in general with the nonextended source unless the data kinematics are well-predicted by the velocity. We denote by $G[v](\mathbf{x}_r, t; \mathbf{x}_s)$ the causal Green’s function of the acoustic-wave equation, that is, the solution of the system 1, 2 with $f(t) = \delta(t)$. Then

$$\tilde{S}[v]\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s) = G[v](\mathbf{x}_r, t; \mathbf{x}_s) *_t \tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s). \quad (8)$$

Equation 8 shows that computing the source-receiver extended forward map involves minimal expense beyond that of the nonextended forward map defined in equation 3. Because $G[v] = S[v]\delta(t)$, computing the action of $\tilde{S}[v]$ requires computing the action of $S[v]$, followed by one additional convolution per output trace.

If we denote by $(G[v]*)^{-1}$ a convolution inverse to $G[v]$, then

$$\tilde{S}[v](G[v]*)^{-1}d(\mathbf{x}_r, t; \mathbf{x}_s) = d(\mathbf{x}_r, t; \mathbf{x}_s). \quad (9)$$

That is, $\tilde{S}[v]^{-1} = (G[v]*)^{-1}$. This operation assumes that Green’s function has a convolution inverse, of course.

In this paper, we will assume that all traces are defined on the same time interval $[0, T]$. To accommodate timing errors associated with erroneous velocities, we extend the time interval for the (extended) source and the data to $[-T, T]$, padding the data with zeros for $t < 0$. Then, we regard all functions as periodic in T of period $2T$, and use circulant convolution, which of course in the Fourier domain amounts to multiplication. So, the convolution inverse of Green’s function is simply its reciprocal in the Fourier domain, which is a priori available only if the Fourier transform has no zeros. Sometimes this is the case, sometimes not, as will be illustrated below.

Assuming that $\tilde{S}[v]$ is invertible, the extended source that explains the data d is (exactly)

$$\tilde{f}[v] = \tilde{S}[v]^{-1}d. \quad (10)$$

Extended source-receiver waveform inversion

The extended source constructed in equation 10 is mostly likely unphysical, in which the physics defined at the beginning of this section required that sources are uniform across all traces; that is, $\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s) = f(t)$. Because the data can be fit (assuming deconvolvable Green’s function) for any velocity, it is only this requirement that provides velocity updates. A simple way to quantify failure of the extended source to match a physical one uses an annihilator, that is, an operator A that produces a zero result when applied to a physically consistent source. Two possible choices are differentiation of the source with respect to \mathbf{x}_s and \mathbf{x}_r (Song and Symes, 1993; Pratt and Symes, 2002), or forcing the convolution quotient of the extended and target (nonextended) sources to resemble the delta function, for example, by penalizing the second moment of the squared signal, also known as dispersion about zero lag (Plessix et al., 1999; Luo and Sava, 2011; Warner and Guasch, 2014, 2016). We will use the second option in the work reported below because it is somewhat simpler to implement. Specifically,

$$A\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s) = t(f^\dagger) * \tilde{f}, \quad (11)$$

where f^\dagger is an approximate inverse, or shaping filter, for the known (common to all traces) source wavelet f . It satisfies

$$f^\dagger * f = \chi, \quad (12)$$

in which χ is an approximate (band-limited) delta function, for example, a zero-phase band-pass filter.

If we apply A to the v -dependent extended source $\tilde{f}[v]$ defined in equation 10, we obtain an index of velocity correctness: If the veloc-

ity model and data are kinematically compatible, then all of the inverted sources $\tilde{f}[v](\mathbf{x}_r, t; \mathbf{x}_s)$ are approximately the same as the source $f(t)$; hence, signature deconvolution should yield an approximate delta function at zero lag, and that is in turn nearly annihilated by multiplication by t . Kinematic disagreement between the model and data should lead to larger $A\tilde{f}$.

We capture this idea in an objective function:

$$J[v] = \frac{1}{2} \|A\tilde{f}[v]\|^2. \quad (13)$$

In effect, the model over which this objective is to be optimized includes the velocity $v(\mathbf{x})$ and the extended source function $\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s)$; hence, it is an extended model. The optimization is treated as a nested problem, with \tilde{f} determined as a function of v via deconvolution (equation 10); then v is determined by minimizing $J[v]$ (equation 13).

Remark 1

The nested design of the optimization problem defined in equations 10, 11, and 13 is essential, not merely a computational convenience. Minimization of alternative objective functions of $v(\mathbf{x})$, $\tilde{f}(\mathbf{x}_r, t; \mathbf{x}_s)$, such as the penalty function:

$$J_\alpha[v, \tilde{f}] = \frac{1}{2} \|A\tilde{f}[v]\|^2 + \frac{\alpha^2}{2} \|\tilde{S}[v]\tilde{f} - d\|^2, \quad (14)$$

over v and \tilde{f} jointly, turns out to be very inefficient, with (in principle) arbitrarily slow convergence. The cause of this misbehavior is the very different sensitivities of J_α to v and \tilde{f} . A change in \tilde{f} changes J by an amount proportional to the mean square of \tilde{f} . In contrast, a change in v results in a change in traveltimes, therefore, in general a shift in the events in $\tilde{S}[v]\tilde{f}$. The rate of change of J_α with respect to v therefore involves the derivative (in t) of \tilde{f} , which is roughly speaking bigger than \tilde{f} in mean square by a factor of the maximum frequency. Therefore, relatively high-frequency data, desirable from the point of view of model resolution, yield very different sensitivities to \tilde{f} and v , that is, ill conditioning of the Hessian of J_α . Objectives with ill-conditioned Hessians cause local optimization algorithms to perform poorly (see, e.g., Nocedal and Wright [1999] for information on this point). A problem more amenable to solution via local optimization can be recovered by optimizing J_α first over \tilde{f} to create a reduced objective depending only on v , which is then optimized over v — that is, a nested optimization, similar to that defined in equations 10 and 13. See Symes (2015) for an explanation, and Huang and Symes (2015b) for an explicit illustration of the performance contrast between nested and non-nested optimization for a different model extension. We will not use the penalty function J_α in the work reported here.

Relation with traveltime tomography

Consider for the moment 3D wave propagation through slowly varying velocity fields v , for which the associated ray field connects each source-receiver pair with a unique ray. “Slowly varying” means smooth on the wavelength scale, lacking embedded reflectors, dominated by relatively low spatial frequencies. For such smooth single-arrival models, geometric acoustics provides asymptotic Green’s function approximation:

$$G[v](\mathbf{x}_r, t; \mathbf{x}_s) \approx a[v](\mathbf{x}_r; \mathbf{x}_s) \delta(t - \tau[v](\mathbf{x}_r; \mathbf{x}_s)), \quad (15)$$

where a is the geometric amplitude and τ is the traveltime, for the ray between \mathbf{x}_s and \mathbf{x}_r ; of course, both depend on v .

Assume that the data d are the image under S of a physical model v_* , f , with v_* producing a single arrival, that is, enjoying the asymptotic Green’s function approximation 15. Then

$$\begin{aligned} d(\mathbf{x}_r, t; \mathbf{x}_s) &= S[v_*]f(\mathbf{x}_r, t; \mathbf{x}_s) \\ &\approx a[v_*](\mathbf{x}_r, \mathbf{x}_s) f(t - \tau[v_*](\mathbf{x}_r, \mathbf{x}_s)). \end{aligned} \quad (16)$$

Assuming that the trial velocity v also produces single arrivals, the extended source $\tilde{f}[v]$ defined by equation 10 is

$$\tilde{f}[v] \approx \frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} f(t - (\tau[v_*](\mathbf{x}_r, \mathbf{x}_s) - \tau[v](\mathbf{x}_r, \mathbf{x}_s))). \quad (17)$$

From equations 11 to 13

$$\begin{aligned} J[v] &\approx \sum_{\mathbf{x}_r, \mathbf{x}_s} \left(\frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} \right)^2 \\ &\quad \times \int dt t^2 (\chi(t - (\tau[v_*](\mathbf{x}_r, \mathbf{x}_s) - \tau[v](\mathbf{x}_r, \mathbf{x}_s))))^2 \\ &= \sum_{\mathbf{x}_r, \mathbf{x}_s} \left(\frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} \right)^2 \\ &\quad \times \int dt (t + (\tau[v_*](\mathbf{x}_r, \mathbf{x}_s) - \tau[v](\mathbf{x}_r, \mathbf{x}_s)))^2 \chi(t)^2. \end{aligned} \quad (18)$$

Because χ is assumed zero phase (symmetric about $t = 0$), this

$$\begin{aligned} &= \sum_{\mathbf{x}_r, \mathbf{x}_s} \left(\frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} \right)^2 \int dt t^2 \chi(t)^2 \\ &\quad + \sum_{\mathbf{x}_r, \mathbf{x}_s} \left(\frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} \right)^2 (\tau[v_*](\mathbf{x}_r, \mathbf{x}_s) - \tau[v](\mathbf{x}_r, \mathbf{x}_s))^2 \\ &\quad \times \int dt \chi(t)^2. \end{aligned} \quad (19)$$

Because χ is an approximate delta, its square is concentrated near $t = 0$, so the second moment of its square (the integral in the first summand in equation 19) is small. We conclude that

$$\begin{aligned} J[v] &\approx \left(\int dt \chi(t)^2 \right) \sum_{\mathbf{x}_r, \mathbf{x}_s} \left(\frac{a[v_*](\mathbf{x}_r, \mathbf{x}_s)}{a[v](\mathbf{x}_r, \mathbf{x}_s)} \right)^2 \\ &\quad \times (\tau[v_*](\mathbf{x}_r, \mathbf{x}_s) - \tau[v](\mathbf{x}_r, \mathbf{x}_s))^2. \end{aligned} \quad (20)$$

That is, $J[v]$ is approximately a weighted mean of the square traveltime error. The weights are the squared amplitude ratios.

Assuming that the trial velocities v are kept safely away from creating caustics, the ratio of amplitudes between trial and target (v_*) velocities lies between two positive numbers, the exact values of which depend on precisely what is meant by “safely away.” If J approaches its minimum value (zero) through a sequence of such velocity models, then the (unweighted) mean-square traveltime

error must approach zero also, and vice versa. Thus, global minimization of J is equivalent to global minimization of the standard mean-square traveltimes tomography objective. Similar arguments show that the gradient of J is small only when the gradient of the traveltimes tomography objective is small, and that the Hessian of J at the exact solution ($v = v_*$) dominates the Hessian of the traveltimes tomography objective (for details, see [Huang and Symes, 2015a](#)). That is, there is a domain in model space, whose size is independent of the bandwidth of the data, in which minimization of J determines the same aspects of the velocity model as does traveltimes tomography.

It is possible to make the foregoing statements mathematically precise: That is, the approximations indicated above are all in the same sense (high-frequency asymptotics), and sets of models “safely” far from generating caustics can be precisely characterized ([Song, 1994](#)). It is also possible to justify precisely the same conclusions for 2D propagation. These refinements are beyond the scope of this paper, but they are consistent in every way with the proposition that minimization of the objective J defined here is effectively equivalent to traveltimes tomography for models that generate only single arrivals.

Effect of multiple arrivals

The story changes dramatically if the data exhibit multiple energetic arrivals, even if in other ways it conforms to the limitations mentioned in the last section, that is, transmission through a slowly varying material model. As pointed out already in [Symes \(1994\)](#), with multiple energetic arrivals, it is possible that the (unregularized) source-receiver extended waveform inversion objective may be as nonconvex as the FWI objective. Consequently, the source-receiver extended waveform inversion algorithm can fail to produce kinematically accurate velocity estimates.

We present here a different viewpoint from [Symes \(1994\)](#), who based his discussion on causal deconvolution. As mentioned earlier, our work uses circulant deconvolution (Fourier division) instead. From this point of view, the pathology is a by-product of spectral notches that may develop in the Green’s function. Although this argument does not yield the precise conclusion about nonconvexity, it is simpler and indicates another manifestation of the underlying pathology.

For the generic source-receiver pair (not lying precisely on a caustic), the 3D Green’s function takes the asymptotic form:

$$G[v](\mathbf{x}_r, t; \mathbf{x}_s) \approx \sum_{i=0}^N a_i[v](\mathbf{x}_r; \mathbf{x}_s) H^{\nu_i} \delta(t - \tau_i[v](\mathbf{x}_r; \mathbf{x}_s)), \quad (21)$$

where a_i is the geometric amplitude and τ_i is the traveltimes, for the i th ray connecting \mathbf{x}_s and \mathbf{x}_r . Here, H is the Hilbert transform, and its power ν_i is either zero or one and is related to the Maslov index of the ray in the ray field emanating from the source.

The $S[v]^{-1}$ amounts to deconvolution by $G[v]$. However, the Fourier transform of G is an exponential sum; hence, it may have zeros or approximate zeros (or, as they are known in this literature, notches). The simplest cartoon example, actually relevant to a synthetic example presented later in this paper, is

$$G(t) \approx a(\delta(t) + \delta(t - \Delta t)). \quad (22)$$

Then, $Sf = G * f = 0$ if $f(t) = \cos(\pi t/\Delta t)$; that is, the Fourier transform of G vanishes at odd multiples of $1/\Delta t$.

If G has literal zero Fourier components, as in the cartoon example, then equation 9 may not have a solution, that is, G does not have a convolution inverse, and if equation 9 has a solution, it is not unique. More likely to occur are very small values of the Fourier transform for which the corresponding sinusoid will be vastly over-emphasized in the solution of equation 9. If the data d are noise-free data from the same model, that does not matter because the small Fourier component of G is already part of the data. However, if the trial model is not very close to the model used to generate the data, then the corresponding data component is likely to be large, resulting in a large sinusoidal contribution at the notch frequency and receiver location to the extended source. Because the sinusoid is nondecaying, $\bar{f}[v]$ defined in equation 10 acquires energy at time lags that have nothing to do with the overall traveltimes difference between the trial model and the global minimizer. Thus, the connection between the traveltimes error and the value of J is broken in this case.

The more refined analysis presented by [Symes \(1994\)](#) actually shows that the domain of convexity of source-receiver extended inversion generally has diameter proportional to a wavelength, similar to FWI: Convergence requires that the initial model predict data event times to within a half-wavelength.

Regularized source-receiver extended inversion

To make an estimation of the extended source \bar{f} robust against spectral zeros of Green’s function, we replace equation 9 with the regularized least-squares problem: Choose \bar{f} to minimize

$$\frac{1}{2} \|\bar{S}[v]\bar{f} - d\|^2 + \frac{\epsilon}{2} \|\bar{f}\|^2. \quad (23)$$

The minimizer (again denoted $\bar{f}[v]$) solves the normal equation

$$(\bar{S}[v]^T \bar{S}[v] + \epsilon^2 I) \bar{f}[v](\mathbf{x}_r, t; \mathbf{x}_s) = \bar{S}[v]^T d(\mathbf{x}_r, t; \mathbf{x}_s), \quad (24)$$

and it depends on v , d , and ϵ . The matrix of the operator on the left side of equation 24 is the autocorrelation of G , “prewhitened” by the addition of ϵ^2 . The effect of the prewhitening is to uniformly increase the entire power spectrum of G , thus making the matrix on the left side of equation 24 better conditioned. Of course, the normal equation 24 is again a convolution equation, and it can be solved exactly by means of the Fourier transform.

We define the regularized source-receiver extension objective $J[v]$ again by equation 13, but now with $\bar{f}[v]$ the solution of equation 24.

We expect the modified $J[v]$ to oscillate less than the nonregularized version discussed above: Regularization reduces the very large components in the extended source, arising from the spectral notches, proportionally more than other components that presumably contribute energy at time lags more proportional to the overall traveltimes error between the trial and target models. It is easy to see from low-dimensional matrix analogs that we cannot expect the regularization parameter ϵ to be particularly small.

We use a version of the discrepancy principle (control of the data residual) to set ϵ . We compute the relative initial data residual e_0 and the relative extended initial data residual \bar{e}_0 using the initial velocity v_0 :

$$e_0 = \frac{\|S[v_0]f - d\|}{\|d\|}, \quad \bar{e}_0 = \frac{\|\bar{S}[v_0]\bar{f}[v_0] - d\|}{\|d\|}, \quad (25)$$

in which f is the known source wavelet and \bar{f} is the extended source function, estimated by solving equation 24 with the initial velocity. We adjusted ε by trial-and-error until the ratio \bar{e}_0/e_0 lay in the range of 0.2–0.5. The (somewhat arbitrary) bounds are chosen to ensure that ε is large enough that the residual is substantially larger than zero (the expected value for $\varepsilon = 0$), but small enough that the data are substantially better fit than is possible with a physical source ($\bar{f} = f$). Thus, ε is large enough to suppress the notch contributions to some extent, but small enough to ensure that the larger events in the data must be fit to some extent. This method for selecting ε has only a heuristic justification, but has functioned well in the examples shown below and in other similar examples not shown here.

Gradient computation

We use a gradient-based method to minimize J . Appendix A shows how to compute this gradient. We define \bar{w} to be the solution of

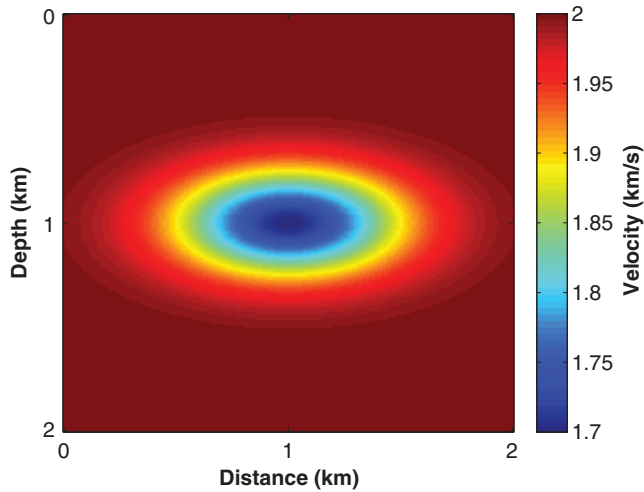


Figure 1. First crosswell example: target velocity model with slow Gaussian anomaly. The lowest velocity is 1.7 km/s.

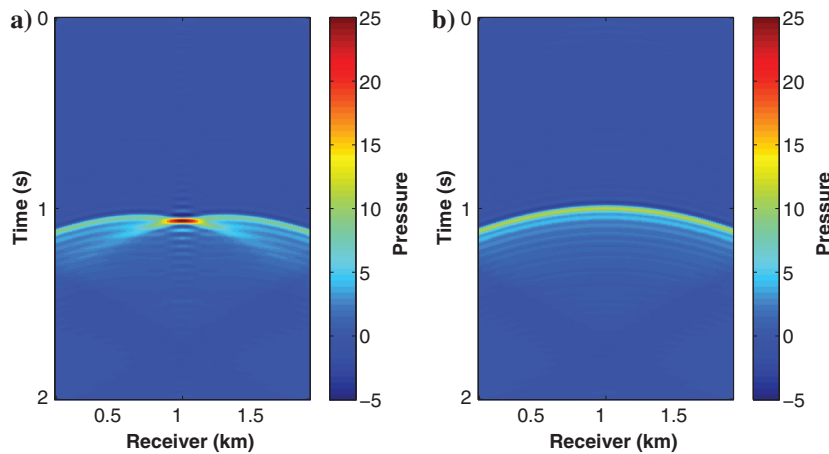


Figure 2. First crosswell example: comparison of shot gathers for the center shot $z_s = 1$ km: (a) target data and (b) simulated data using initial velocity.

$$(\bar{S}[v]^T \bar{S}[v] + \varepsilon^2 I) \bar{w} = A^T A \bar{f} \quad (26)$$

and \bar{r} by

$$\bar{r} = \bar{w} \star (d - \bar{S}[v] \bar{f}) - \bar{f} \star \bar{S}[v] \bar{w}, \quad (27)$$

where \star denotes crosscorrelation. Then

$$\nabla J[v] = (D\bar{S}[v]\delta)^T \bar{r}, \quad (28)$$

in which $(D\bar{S}[v]\delta)^T$ is the well-known impulsive reverse time migration operator (see Appendix A for details).

Computational cost

The computational cost of source-receiver extended waveform inversion is comparable with the cost of standard least-squares FWI. Each function value $J[v]$ requires computation of the Green's function (therefore one modeling step), a deconvolution (solution of equation 24), and some vector algebra. A gradient evaluation $\nabla J[v]$, as defined in the last section, adds another deconvolution (equation 26), and two more convolutions (equation 27) and some more vector algebra, followed by reverse time migration (equation 28). Thus, each step of a gradient-based optimization is of roughly the same cost as a step of the same algorithm to FWI. Concerning a comparison of total costs over the entire iteration, all that can be said in general is that an algorithm that produces a useful approximation in a reasonable number of steps is in a real sense infinitely cheaper than one that does not.

FWI as postprocess

In principle, the principal goal of inversion is the production of an explanatory physical model that fits the data. The regularization level required to induce convergence of source-receiver extended inversion from seriously wrong initial models has a negative effect on data fit. However, in the examples to follow, regularized source-receiver inversion produces kinematically accurate velocity models, while relaxing the data fit constraint of the unregularized algorithm. Its output should be acceptable input to conventional FWI, and the latter algorithm should reduce the remaining data misfit. Therefore, we shall in most cases follow inversion via source-receiver extension with FWI initialized on the final source-receiver model.

NUMERICAL EXAMPLES

We present two sets of numerical examples illustrating the performance of source-receiver extended waveform inversion in comparison with FWI. The first set uses an idealized crosswell geometry, with the source locations on one face of the rectangular scattering domain, the receivers on the opposite face. The second set mimics surface acquisition, with sources and receivers on the same face of the rectangular domain.

We use the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm with a backtracking line search to assure compliance with the weak version of Wolfe's conditions for

global convergence to a stationary point (Nocedal and Wright, 1999). For each example, we will apply LBFGS to solve the full-bandwidth FWI problem and to minimize the source-receiver extended waveform inversion objective defined in equation 13, and the FWI objective (equation 4). Forward modeling is implemented in the frequency domain, using a nine-point (fourth order and cross shaped) stencil to approximate the Helmholtz operator, and a direct matrix solver. The least-squares problems defined in equations 24 and 26 are diagonal in the Fourier domain because of our use of circulant deconvolution, so these are solved to machine precision as well. The objective function and gradient are therefore computed to machine precision, at the discrete level.

Crosswell acquisition geometry

The first two examples use idealized crosswell acquisition, that is, sources lie on one boundary face of a rectangular scattering region, and receivers are on the opposite face. The first example shows that source-receiver extended waveform inversion without regularization (i.e., $\epsilon = 0$) may converge to a reasonable solution of inverse problem when FWI fails to do so. In this example, the data exhibit only a single arrival (although this is partly as a result of wavefront healing). Our previous analysis (Huang and Symes, 2015a) explains the behavior of source-receiver extended waveform inversion in this case: When raypaths from the source to receiver are unique, source-receiver extended waveform inversion is equivalent to least-squares traveltime tomography, and it delivers a comparable solution. The second example shows that unregularized source-receiver extended waveform inversion may fail in the same way as FWI if multiple arrivals are present in the data with significant energy. We observe that this phenomenon may be understood as ill-posedness of the inner (source-estimation) problem, and that Tikhonov regularization can restore apparent convergence to a tomographic-quality solution.

Weak low-velocity lens

The target velocity model for the first example is a Gaussian low-velocity anomaly embedded in a constant background velocity $v_0 = 2$ km/s (Figure 1), i.e.,

$$v(x, z) = 2 - 0.7e^{-\frac{(x-1)^2}{0.5^2} - \frac{(z-1)^2}{0.25^2}} \text{ km/s.} \quad (29)$$

The sources and receivers are placed at $x_s = 0.01$ km and $x_r = 1.99$ km, respectively. The 39 shots are evenly spaced between $z_s = 0.05$ and 1.95 km, and 199 receivers are located from $z_r = 0.01$ to 1.99 km with $\Delta z_r = 0.01$ km.

For display purposes, we synthesize a time-domain solution of the target problem from frequency-domain fields, and we show the recorded data for the center shot at $z_s = 1$ km in Figure 2a. The effective time-domain source is a boxcar 0.5–30 Hz band-pass filter. We use the constant velocity $v_0 = 2$ km/s as the initial model. The simulated data in Figure 2b using the initial velocity for the centered shot show a traveltime error larger than a half-wavelength at the median frequency of 12 Hz for the central data traces of

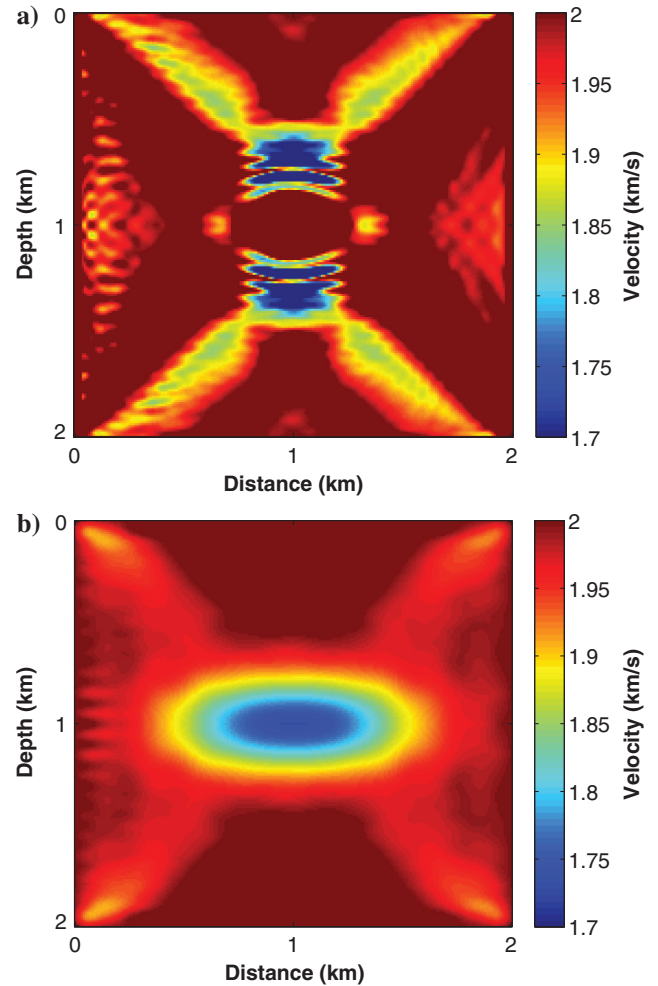


Figure 3. First crosswell example: inverted velocity after 21 LBFGS iterations with 3–20 Hz band-pass data by (a) FWI and (b) source-receiver extended waveform inversion ($\epsilon = 0$). Initial velocity is $v_0 = 2$ km/s in both cases.

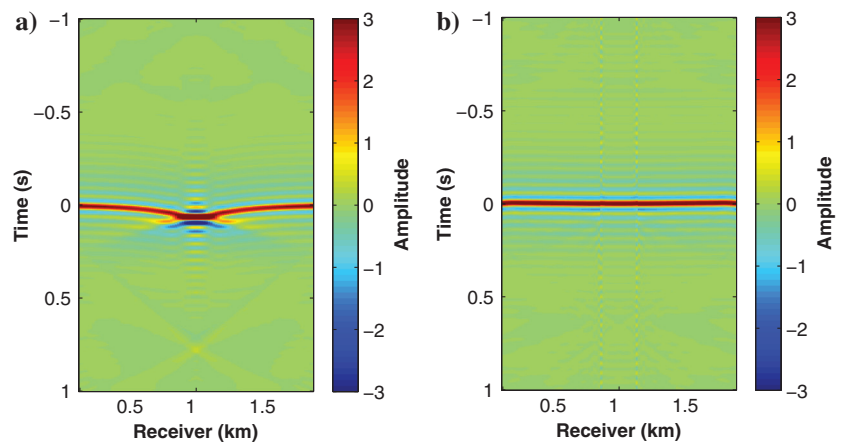


Figure 4. First crosswell example: extended sources for central shot $z_s = 1$ km using (a) initial velocity and (b) inverted velocity by source-receiver extended waveform inversion.

the recorded data. For inversion, we use the data frequency band from 3 to 20 Hz. After 21 LBFGS iterations, FWI is stuck in a physically meaningless solution, whereas source-receiver extended waveform inversion produces a reasonable estimate of v (see Figure 3a and 3b for the detailed result). To further confirm the kin-

ematic accuracy of the source-receiver extended waveform inversion solution, we display the extended sources at the initial velocity (Figure 4a) and the final inverted velocity (Figure 4b).

We can understand this example on basis of theory presented in the last section. Examination of Figure 2a reveals that a caustic likely is present in the ray field of this example, but that finite frequency has “healed” it and presented an effective single arrival. Therefore, we would expect the source-receiver extended inversion to converge even from an initial model that is cycle skipped for FWI to a model of tomographic quality, and indeed it does.

Strong low-velocity lens

The second example (Figure 5) keeps the shape velocity anomaly, but makes it stronger:

$$v(x, z) = 2 - 0.6e^{-\frac{(x-1)^2}{0.5^2} - \frac{(z-1)^2}{0.25^2}}. \quad (30)$$

Now, the lowest velocity of this model is 1.4 km/s. We use the same source and receiver geometry and the data frequency band as the first example. Synthetic data (Figure 6a–6c) for three shot positions $z_s = 0.1, 0.5, 1$ km show energetic later arrivals. For comparison, Figure 7a–7c shows simulated data with initial velocity $v_0 = 2$ km/s at the same shot positions.

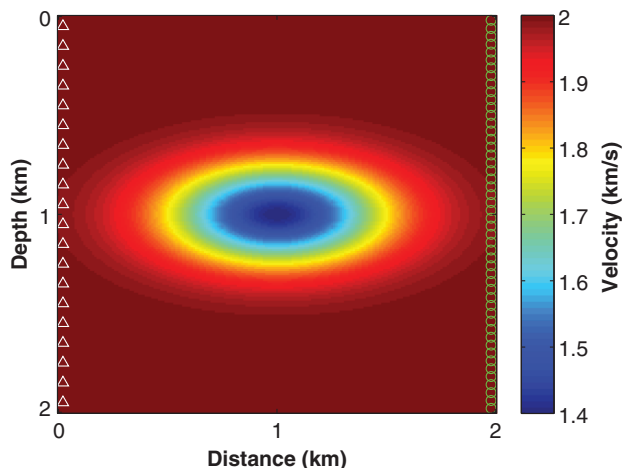


Figure 5. Second crosswell example: target velocity model with slow Gaussian anomaly. The lowest velocity is 1.4 km/s.

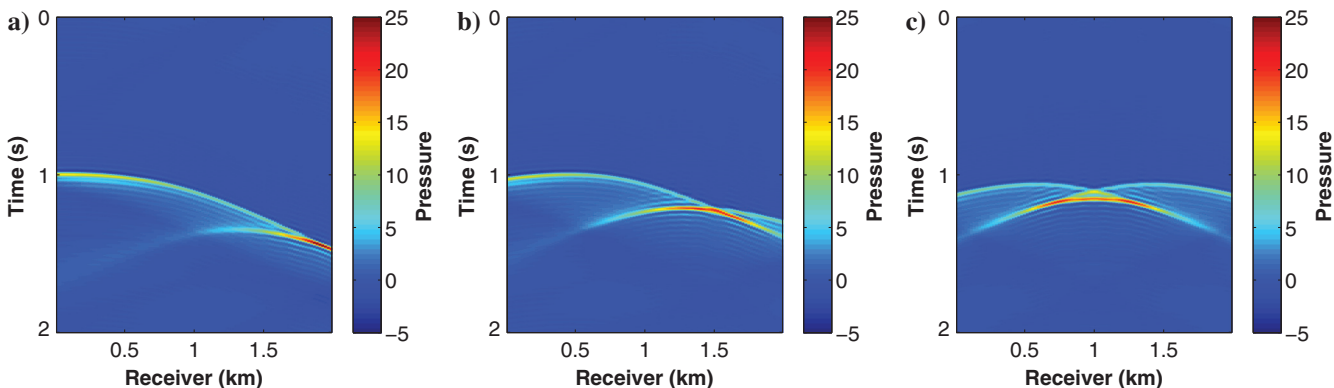


Figure 6. Second crosswell example: shot gathers of recorded data for shot at (a) $z_s = 0.1$ km, (b) $z_s = 0.5$ km, and (c) $z_s = 1$ km.

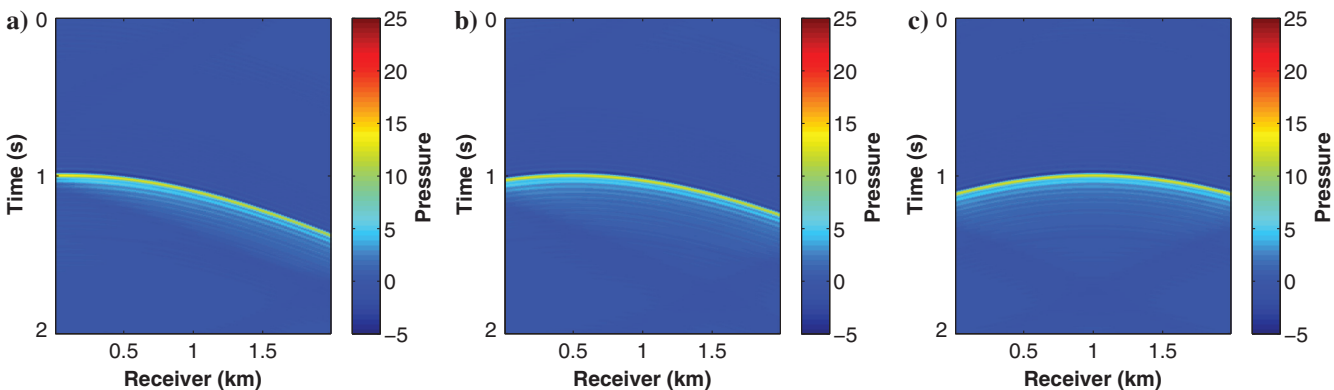


Figure 7. Second crosswell example: shot gathers of simulated data using initial velocity v_0 for shot at (a) $z_s = 0.1$ km, (b) $z_s = 0.5$ km, and (c) $z_s = 1$ km.

The misfit of traveltimes in the initial data is even more severe than was the case in the first example, and FWI indeed fails as we expect (Figure 8a). However, source-receiver extended waveform

inversion without regularization also fails to produce a kinematically accurate velocity after 100 LBFSGS iterations (Figure 8b), for the reasons explained in the “Theory” section.

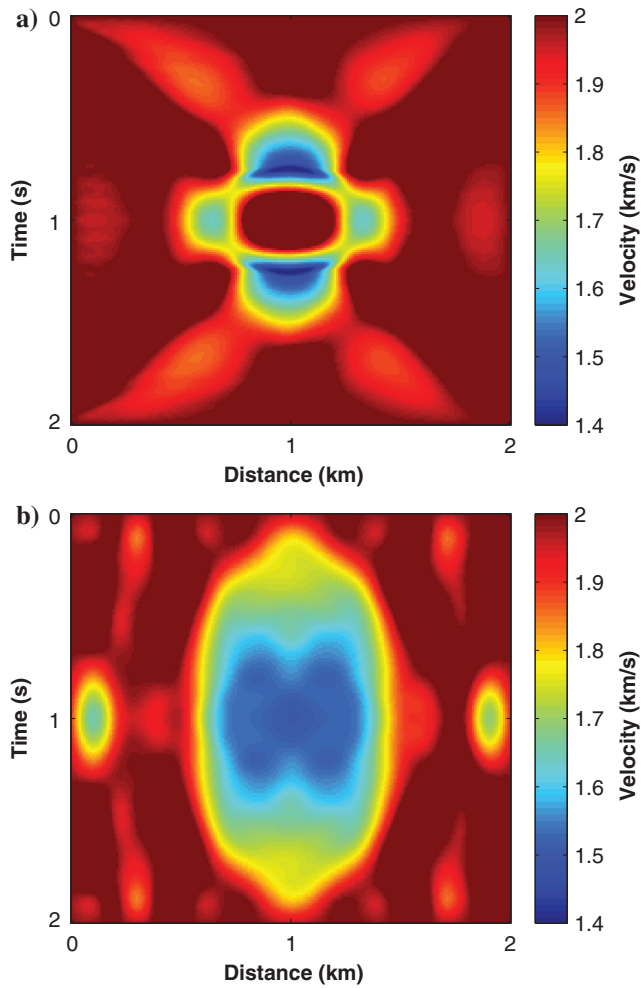


Figure 8. Second crosswell example: inverted velocity after 100 LBFSGS iterations with 3–20 Hz data by (a) FWI and (b) unregularized source-receiver extended waveform inversion ($\epsilon = 0$). The initial velocity is $v_0 = 2$ km/s in both cases.

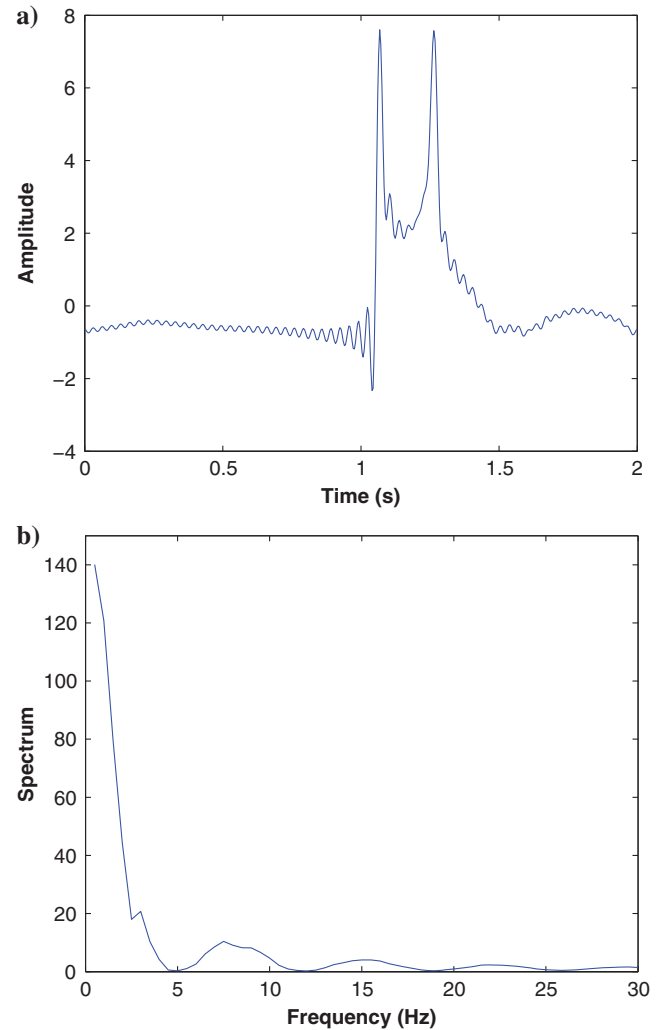


Figure 10. Second crosswell example: (a) data trace at $z_r = 0.55$ km for shot at $z_s = 1$ km and (b) spectrum of normal operator $\bar{S}^T \bar{S}$ of trace (a).

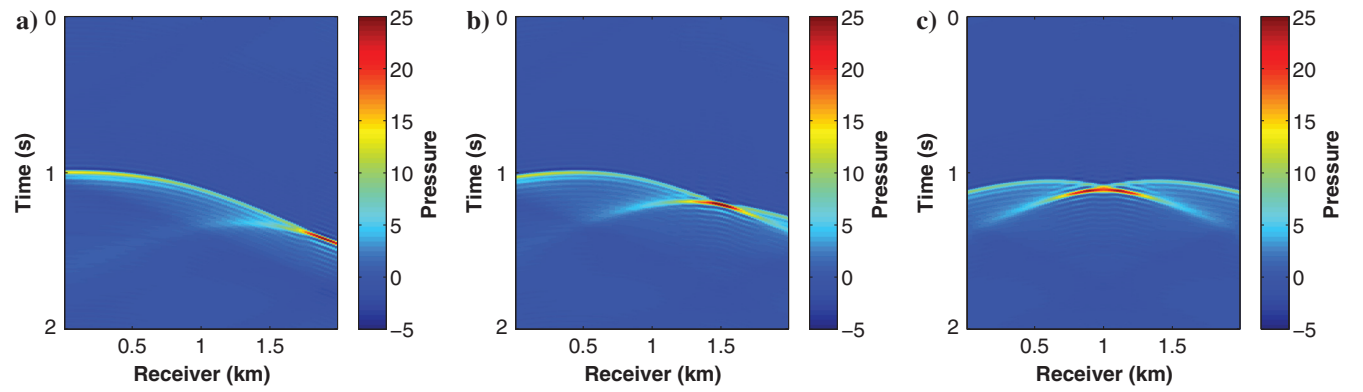


Figure 9. Second crosswell example: shot gathers of simulated data using trial velocity $v_t = 0.8v + 0.2v_0$ for shot at (a) $z_s = 0.1$ km, (b) $z_s = 0.5$ km, and (c) $z_s = 1$ km.

Downloaded 09/14/17 to 128.42.226.198. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

To illustrate the mechanism of failure, we simulated data in Figure 9a–9c with the trial velocity $v_r = 0.8v + 0.2v_0$, which is close to the target velocity model. We can see there are still quite obvious triplications present in the simulated data. Taking a single trace $z_r = 0.55$ km (Figure 10a) for the center shot gather at $z_s = 1$ km, for example, we plot the spectrum of the normal operator $\bar{S}^T \bar{S}$, which is the same as the power spectrum of Green's function (Figure 10b). For traces with multiple energetic arrivals, the spectrum oscillates and almost vanishes at several frequencies, suggesting the existence of an effective numerical null space of the normal operator.

To illustrate the effect of regularization, we plot the extended source, estimated by solving equation 24 with the trial velocity v_r , for regularization parameters $\epsilon = 10^{-6}, 0.5, 5$ (Figure 11a–11i). Oscillatory and nondecaying traces are present in the extended source function $\tilde{f}[v_r]$ for $\epsilon = 10^{-6}$, but they are suppressed by increasing ϵ .

The objective function for different values of $\epsilon = 10^{-6}, 0.5, 5$ on a line segment in model space, between the initial and target velocities, appears in Figure 12a–12c. Parameter $\alpha = 0$ corresponds to

the target velocity, $\alpha = 1$ to the initial velocity. These plots show that the region of convexity of source-receiver extended waveform inversion objective is quite small for small ϵ ; however, it expands to include the initial velocity for a large enough ϵ .

As mentioned before, the arrival time error in the initial 3–20 Hz data (Figure 7a–7c) is too large to permit successful FWI, starting with $v_0 = 2$ km/s (Figure 8a). We find that $\epsilon = 5$ satisfies the discrepancy criterion articulated in the “Theory” section. With this choice, 15 LBFGS iterations produce the velocity estimate in Figure 13a. The position and shape of Gaussian anomaly are well-resolved. In addition, the extended source functions are almost focused on the zero-lag time after regularized source-receiver extended waveform inversion (Figure 14a–14c), indicating that the final velocity estimate is kinematically accurate. In view of its kinematic accuracy, the final estimate from source-receiver extended waveform inversion should be a usable initial estimate for FWI: Indeed, 25 LBFGS iterations of full-bandwidth FWI produce a quite accurate inversion (Figure 13b).

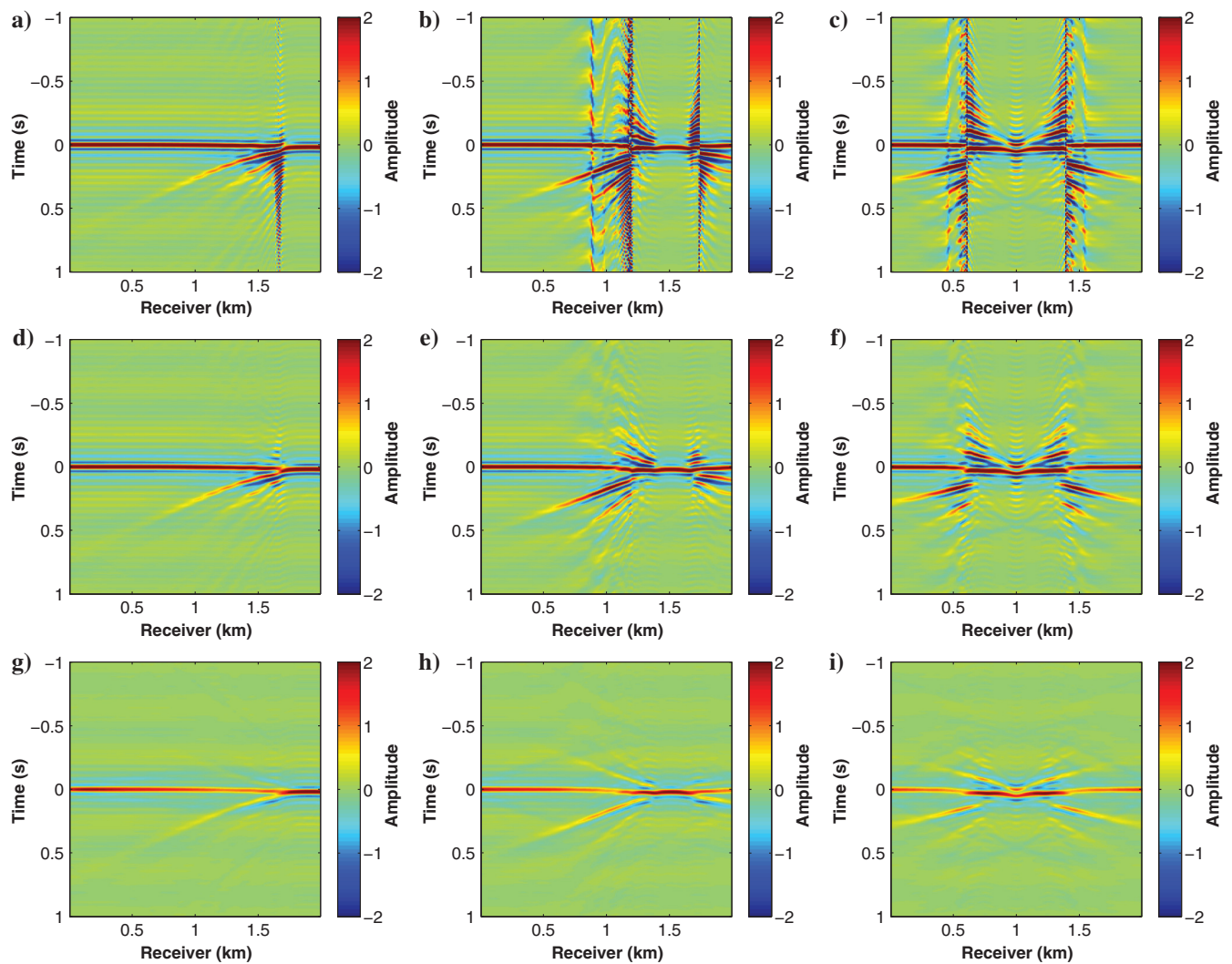


Figure 11. Second crosswell example: plots of regularized extended source functions for three regularization parameters $\epsilon = 10^{-6}, 0.5, 5$ for shot at $z_s = 0.1, 0.5, 1$ km, respectively. (a) $\epsilon = 10^{-6}, z_s = 0.1$ km; (b) $\epsilon = 10^{-6}, z_s = 0.5$ km; (c) $\epsilon = 10^{-6}, z_s = 1$ km; (d) $\epsilon = 0.5, z_s = 0.1$ km; (e) $\epsilon = 0.5, z_s = 0.5$ km; (f) $\epsilon = 0.5, z_s = 1$ km; (g) $\epsilon = 5, z_s = 0.1$ km; (h) $\epsilon = 5, z_s = 0.5$ km; and (i) $\epsilon = 5, z_s = 1$ km.

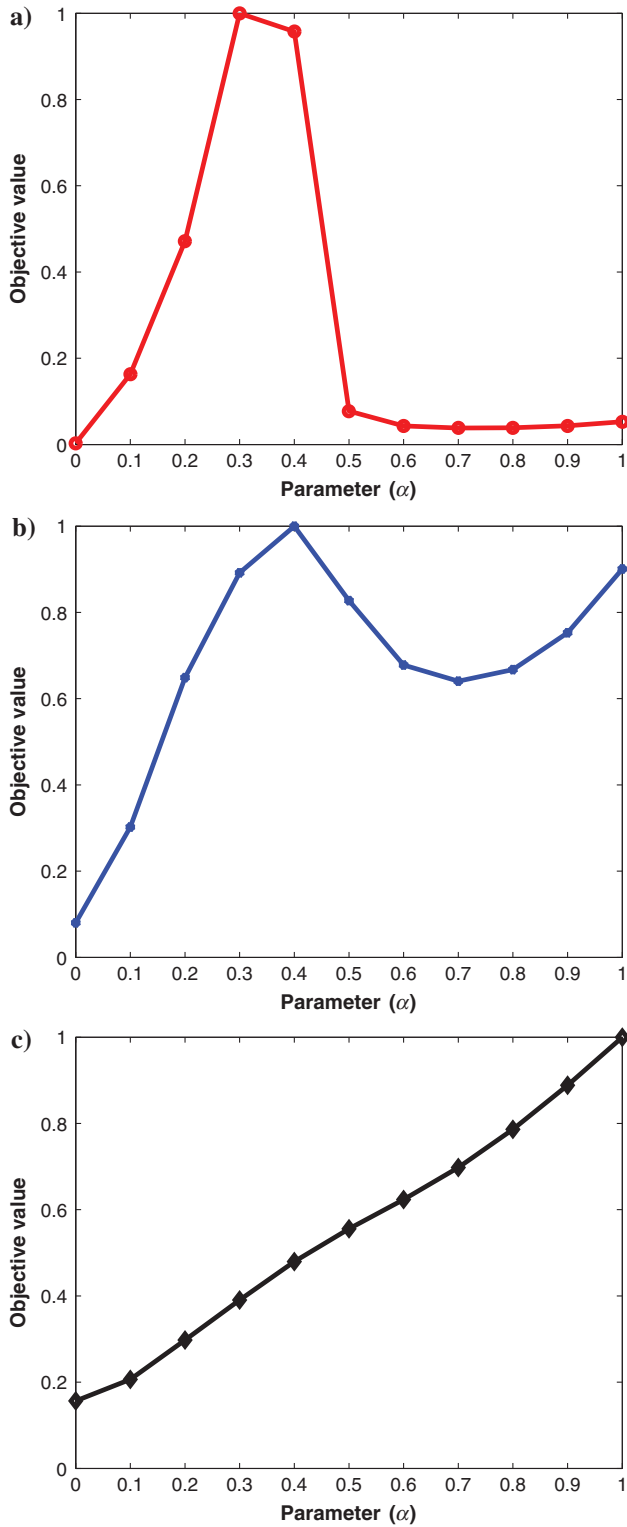


Figure 12. Second crosswell example: objective function evaluated at velocities $(1 - \alpha)v + \alpha v_0$, $0 \leq \alpha \leq 1$, for various choices of regularization parameter ε : (a) $\varepsilon = 10^{-6}$, (b) $\varepsilon = 0.5$, and (c) $\varepsilon = 5$.

Surface acquisition geometry

In the remainder of this section, we apply regularized source-receiver extended waveform inversion to waveform inversion to two examples with surface acquisition geometry; that is, the sources and receivers are separated by a hyperplane from the scattering region.

Diving wave inversion

The predominant contemporary use of FWI is to invert diving wave energy (Virieux and Operto, 2009; Vigh et al., 2013). This example examines the use of source-receiver extended waveform inversion, with and without regularization, for a model generating diving waves with triplications. The model is smooth on the wavelength scale, hence, transparent: The data consist only of direct and diving waves, with no reflections.

The target model (Figure 15a) consists of a low-velocity Gaussian anomaly embedded in a linearly increasing background velocity. The 100 receivers are placed at a depth of $z_r = 0.04$ km from

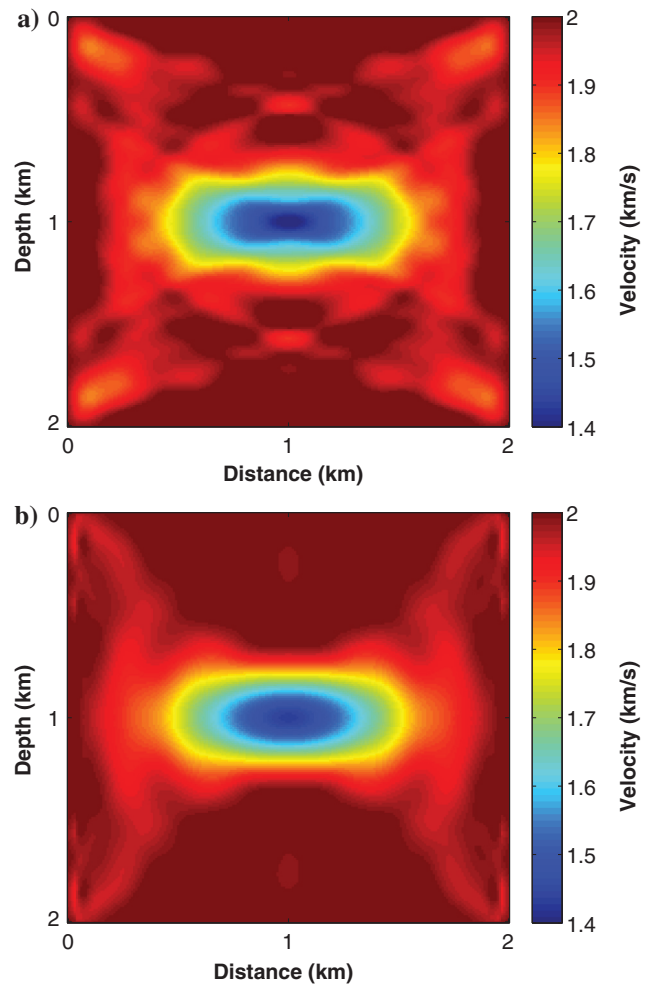


Figure 13. Second crosswell example: (a) inverted velocity by regularized source-receiver extended waveform inversion ($\varepsilon = 5$) after 15 iterations and (b) FWI result after 25 iterations using the regularized source-receiver extended waveform inversion in (a) as initial velocity.

$x_r = 0.04$ to 7.96 km with $\Delta x_r = 0.08$ km. The 67 shots are placed at a depth of $z_s = 0.08$ km from $x_s = 0.04$ km to $x_r = 7.96$ km, with $\Delta x_s = 0.12$ km. The frequency band used in the inversion is 5–11 Hz. The choice $\epsilon = 10$ gives $\bar{e}_0 = 6.45\%$ and $e_0 = 20.95\%$, satisfying the criterion explained earlier.

The initial model in the iterative inversion is the linearly increasing background (Figure 15b), which produces diving wave arrivals without triplication.

Comparison of the data for target and initial data in Figure 16a–16c and 16d–16f shows that as expected, first-arrival times differ by well over a cycle, and of course the triplication structure does not appear in the initial data at all. Therefore, one would expect FWI to stagnate far from a useful model estimate, as indeed happens (Figure 17a). Unregularized source-receiver extended waveform inversion ($\epsilon = 10^{-6}$) also fails. This is a pure transmission problem, and precisely the same phenomenon occurs as in the strong lens crosswell example (Figure 17b). On the other hand, regularized source-receiver extended waveform inversion with the choice of penalty weight $\epsilon = 10$ explained above produces a satisfactory model estimation in the same number of iterations (Figure 18). Examination of extended sources at the initial and final regularized source-receiver extended waveform inversion model suggests that the kinematics of the data have been adequately captured in the final model (Figure 19a–19f).

The output of regularized source-receiver extended waveform inversion also performs well as an initial model for FWI because it has already matched the data arrival times. The FWI with 50 iterations of LBFGS, starting at the model shown in Figure 20, produces the slightly more refined model shown in Figure 20. The simulated data (Figure 21a–21c) generated by the final-inverted results show the best match with the recorded data in Figure 16a–16c.

Pluto

We modify a portion of the Pluto model (Stoughton et al., 2001) to create the target model in Figure 22a, mimicking an isolated salt pillow, gridded with a cell of 0.01×0.01 km. The velocity in the salt inclusion is 4.5 km/s; the background medium is layer-like with velocity averaging 2.2 km/s. Source depth is $z_s = 0.02$ km. Sources range from $x_s = 0.06$ to 2.94 km. Fixed-spread receivers range from $x_r = 0.02$ to 2.98 km placed at depth $z_r = 0.04$ km. The maximum time in the recorded data is 4 s, and the frequency band is 4–10 Hz. We choose the regularization parameter $\epsilon = 1$, for

which initial relative errors are $\bar{e}_0 = 3.2\%$ and $e_0 = 6.6\%$, respectively.

Note that in this example, positions of the sources and receivers permit “undershooting,” that is, transmitted and reflected raypaths that transit the region under the inclusion. Therefore, the data should contain adequate kinematic information to determine the velocity throughout the model, except for the poorly illuminated edges. On the other hand, the large velocity contrast between the salt and surrounding “sediments” implies that FWI will likely fail to reconstruct the inclusion from an initial model (Figure 22b) in which it is absent. Indeed, 200 FWI LBFGS iterations method locates the inclusion top, mispositions the bottom, and grossly underestimates the velocity in between (Figure 23).

Regularized source-receiver extended waveform inversion, in contrast, comes much closer to locating the top and bottom of the inclusion and filling it with approximately correct velocity values (Figure 24a). The focus of the extended source estimates at zero time lag is much improved (Figure 25d–25f) over the initial model (Figure 25a–25c). In fact, the model depicted in Figure 24a appears to be a reasonable initial guess for FWI, providing approximately

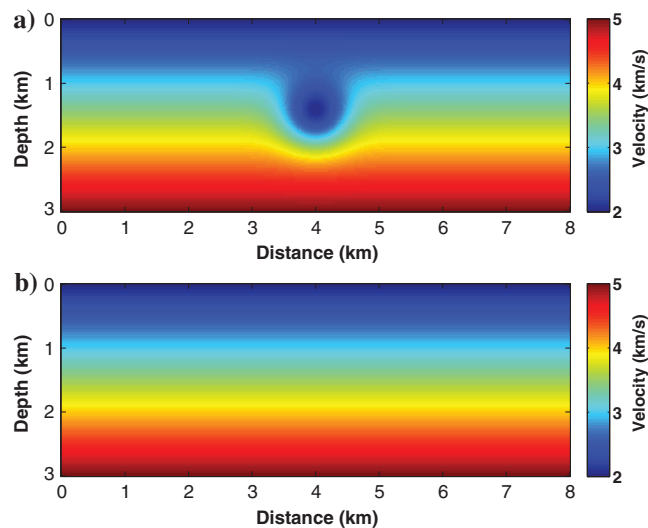


Figure 15. Diving wave example: (a) target velocity model with Gaussian low-velocity lens embedded in the linearly increasing background velocity and (b) $v_0(z)$ linearly increasing initial model.

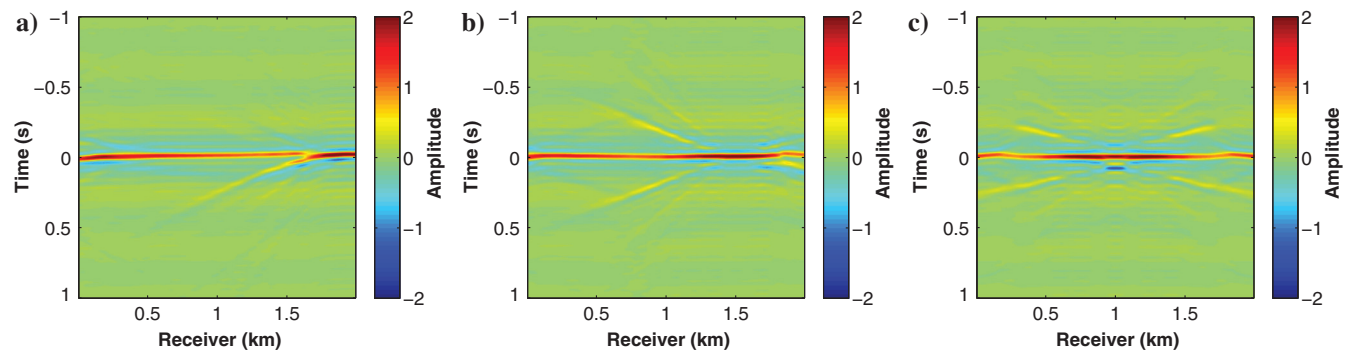


Figure 14. Second crosswell example: extended sources after regularized source-receiver extended waveform inversion for shot at (a) $z_s = 0.1$ km, (b) $z_s = 0.5$ km, and (c) $z_s = 1$ km.

correct kinematics. A further 50 LBFGS iterations of FWI with frequency band 4–16 Hz data results in an accurate reconstruction of the inclusion (Figure 24b).

The final data residual in Figure 26d–26f implies that the inverted model has fit especially the refracted energy in the data quite well, in comparison with the initial data residual (shown in Figure 26a–26c).

DISCUSSION

We have already mentioned in the “Introduction” section that the source-receiver extension used here is the basis for several other algorithms, for example, adaptive waveform inversion (Warner and Guasch, 2014, 2016). Other source extension concepts have also been productive. Waveform reconstruction inversion as described by van Leeuwen and Herrmann (2013) is roughly equivalent to introducing artificial sources “everywhere” (for discussion of other related algorithms, see Huang and Symes, 2016a; Wang and Yingst, 2016). Contrast source inversion, described, for example, by Abubakar et al. (2011), may also be regarded as a source extension approach. Other extended modeling modifications of FWI have been motivated by wave-equation migration velocity analysis (WEMVA) (Biondi and Sava, 2004). These WEMVA-like extensions add parameters to the velocity model itself, for example, subsurface offset (space shift; Shen et al., 2003, 2005; Khoury et al., 2006; Shen and Symes, 2008, 2015; Biondi and Almomin, 2012; Shen, 2012; Weibull and Arntsen, 2014; Lameloise et al., 2015), time shift (Yang and Sava, 2011; Biondi and Almomin, 2014), scattering angle (De Hoop et al., 2003; Shen and Calandra, 2005), shot

coordinates (Symes and Carazzone, 1991; Kern and Symes, 1994; Sun and Symes, 2012; Chauris and Plessix, 2013), and surface offset (Chauris and Noble, 2001; Mulder and ten Kroode, 2002). The common feature in all of these extension-based modifications to FWI is their tendency to produce the same sort of long-wavelength velocity updates as does traveltimes tomography, that is, to extract

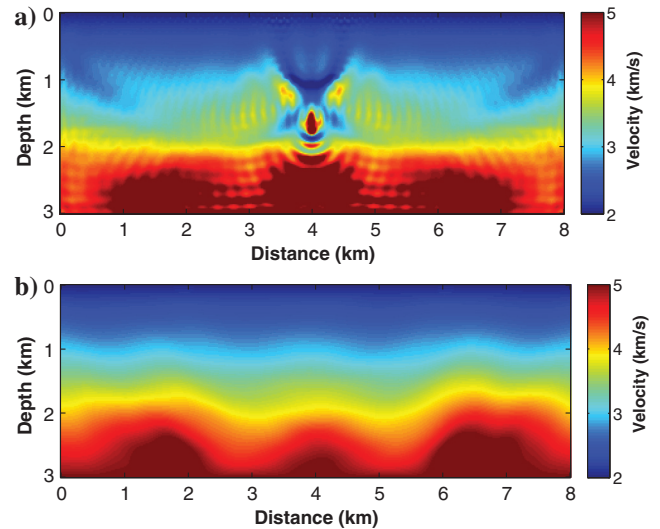


Figure 17. Diving wave example: inverted velocity after 200 iterations using 5–11 Hz data by (a) FWI and (b) unregularized source-receiver extended waveform inversion ($\epsilon = 10^{-6}$).

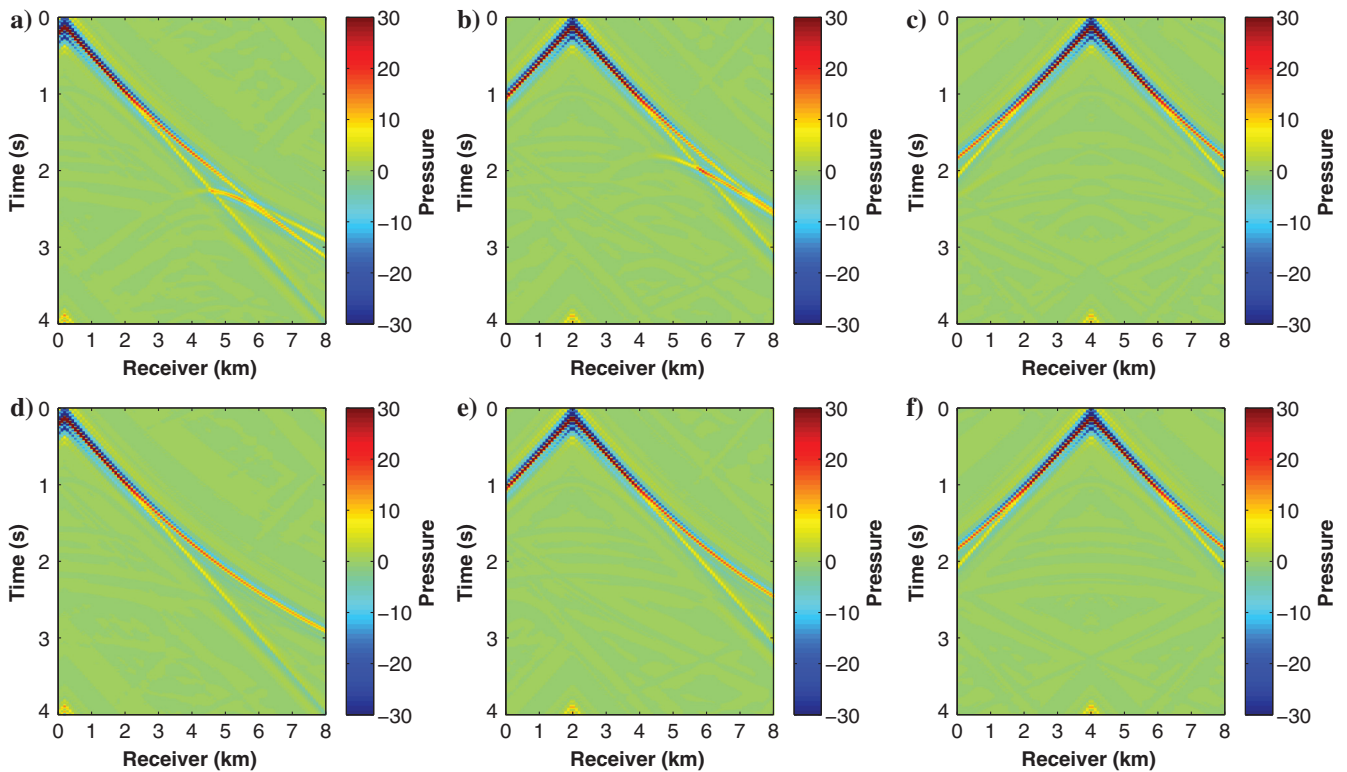


Figure 16. Diving wave example: shot gathers of recorded data for shot at (a) $x_s = 0.2$ km, (b) $x_s = 2$ km, and (c) $x_s = 4$ km. Shot gathers of simulated data by initial velocity for shot at (d) $x_s = 0.2$ km, (e) $x_s = 2$ km, and (f) $x_s = 4$ km.

kinematic information from the data. We have shown here that source-receiver extended inversion accomplishes this goal in the special case of single arrivals. Similar computations show that subsurface offset extended waveform inversion and reflection slope tomography have proportional Hessians at a global solution (Symes, 2014; ten Kroode, 2014).

We have addressed a major difficulty in source-receiver extended waveform inversion, its tendency to develop apparent multimodality in the event that energy arrives in the data along multiple raypaths. This phenomenon presents a real impediment to using source-receiver extended waveform inversion for crosswell tomography, for example because waveguides are quite commonly encountered in that application. Multiple arrivals are also common in diving wavefields, and they can be present for many other reasons. We have advanced an explanation in the context of periodic convolutional modeling,

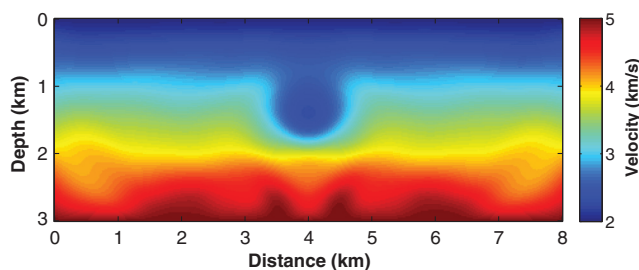


Figure 18. Diving wave example: inverted velocity using 5–11 Hz data after 200 iterations by regularized source-receiver extended waveform inversion ($\epsilon = 10$).

namely, the presence of spectral zeros and near zeros of the Hessian and partly recovered convexity by Tikhonov regularization, a standard technique for selecting solutions of ill-posed problems. Our justification for rather heavy-handed regularization is only heuristic, but it does seem to be effective in many settings, including cartoon examples of crosswell tomography and diving wave FWI, as the examples show.

The other extension methods mentioned above show strong parallels with the observations made here. In particular, the failure of convexity caused by multiple arrivals echoes the similar failure of the surface extension in the presence of source wavefield caustics (Nolan and Symes, 1996; Stolk and Symes, 2004; Symes, 2008). For instance, this better analyzed pathology arises from the impossibility of distinguishing arrivals by midpoint slowness in individual offset gathers. A consequence is the failure of the inner problem to

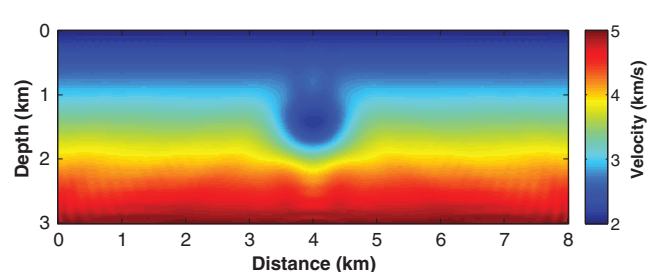


Figure 20. Diving wave example: inverted velocity by 50 FWI iterations, beginning with following regularized source-receiver extended waveform inversion velocity (Figure 18).

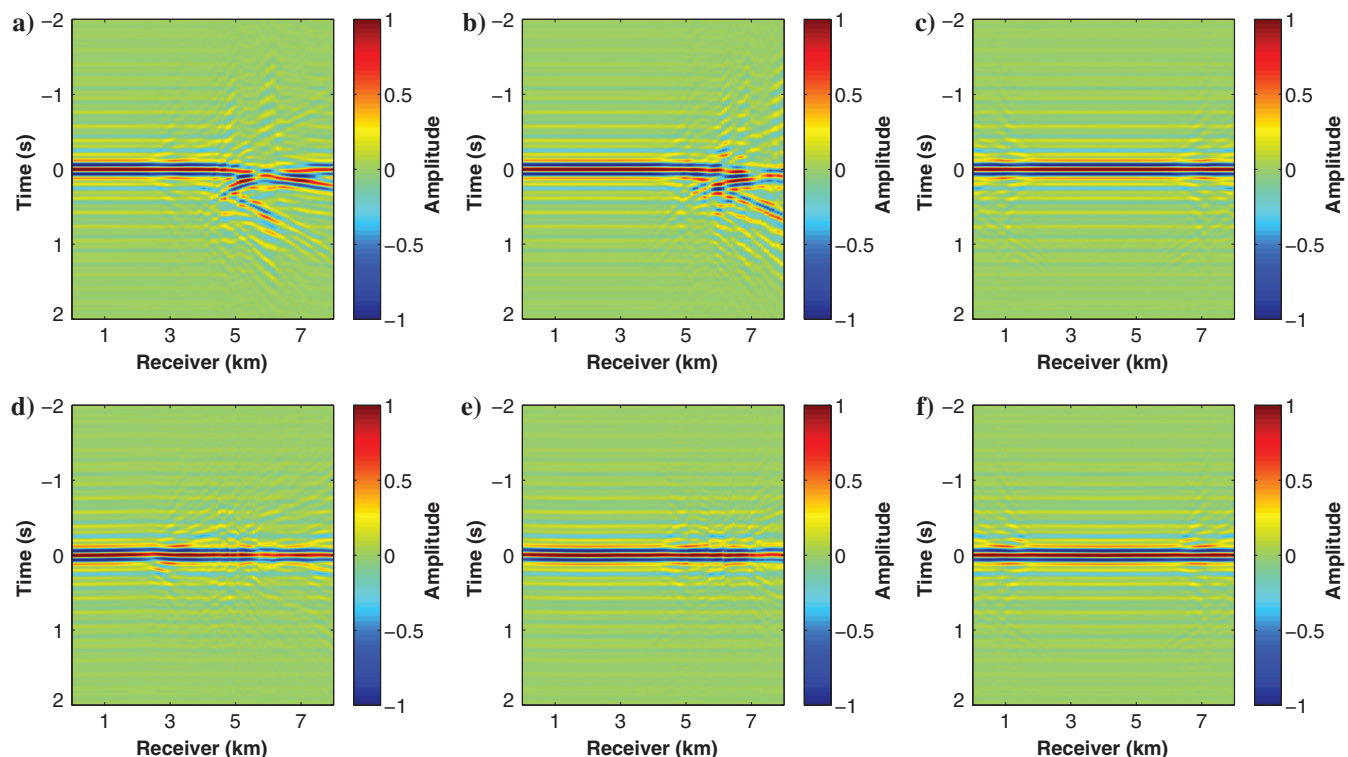


Figure 19. Diving wave example: top row: extended sources by initial model for shot at (a) $x_s = 0.2$ km, (b) $x_s = 2$ km, and (c) $x_s = 4$ km; bottom row: extended sources by inverted velocity after regularized source-receiver extended waveform inversion for shot at (d) $x_s = 0.2$ km, (e) $x_s = 2$ km, and (f) $x_s = 4$ km.

have a unique solution. Just so, arrival slowness cannot be inferred from a single trace, so the source-receiver extension does not permit separation of arrivals by slowness, and consequently the inverse $(\tilde{S}[v])^{-1}$ does not, in general, exist. For medium-based extensions, the remedy is to use so-called subsurface offset instead of surface offset, leading to a well-posed inner problem and consequently a smooth and stable reduced objective (Shen et al., 2003, 2005; Khoury et al., 2006; Shen and Symes, 2008; Stolk et al., 2009; Symes, 2014; ten Kroode, 2014). Other source extensions, in contrast to source-receiver, allow similar conclusions to be drawn and lead to extended inversion algorithms that avoid cycle skipping and maintain data fit, and they do not require strong regularization (Huang and Symes, 2016b).

In the examples presented here, we have used a very large number of LBFGS iterations (hundreds in several examples). The need for so many iterations may be linked to the rather small amount of data used (in almost all cases, 5–10 frequencies), which of course also makes the iterations rather inexpensive. We believe that such computational largesse is appropriate for a study designed to explore the source-receiver extended waveform inversion concept. However, it certainly begs the question: Can reasonable results be obtained in a more reasonable number of RTM applications, say $O(10)$?

Our use of circulant convolution to model the relation between source and data traces is computationally convenient. In particular, it enables an inexpensive and machine-precise solution of the normal equation 24 because the matrices involved are diagonal. Because our examples are both computed and inverted entirely in the frequency domain, this is an appropriate methodology. However, this methodology is commonly termed as an inverse crime; that is, the same tools are used for modeling as for inversion. Field data come in the form of time-domain traces, with finite duration and are not necessarily amplitude decaying. A time cutoff is necessary, either implicitly or explicitly, and that operation does not commute with convolution. Therefore, an appropriate version of problem 24 applicable to field data will need a different solution mode, either the Gaussian elimination or the Levinson algorithm, or (more likely) an iterative solver, such as conjugate gradient iteration.

An inexact solve of the inner problem via an iterative method would bring to the fore another difficulty. The velocity gradient (equation 28) involves reverse time migration, the transpose operator of the derivative of the modeling operator $S[v]$ with respect to velocity. The $S[v]$ amounts to convolution with Green's function $G[v]$. In

the simplest (single-arrival) case, ignoring amplitudes, $G[v] = \delta(t - \tau[v])$. The derivative of $S[v]$ with respect to v is the convolution with the derivative of $G[v]$ with respect to v : From the chain rule, the directional derivative in the direction δv in velocity model space is

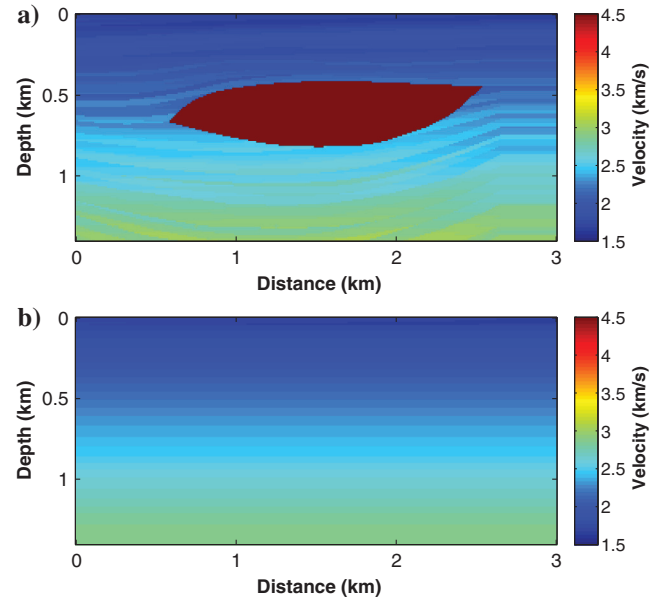


Figure 22. Pluto example: (a) target velocity and (b) 1D initial velocity.

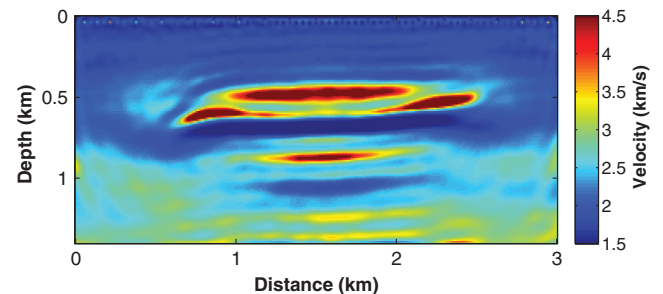


Figure 23. Pluto example: inverted result by FWI method after 200 iterations using 4–10 Hz data.

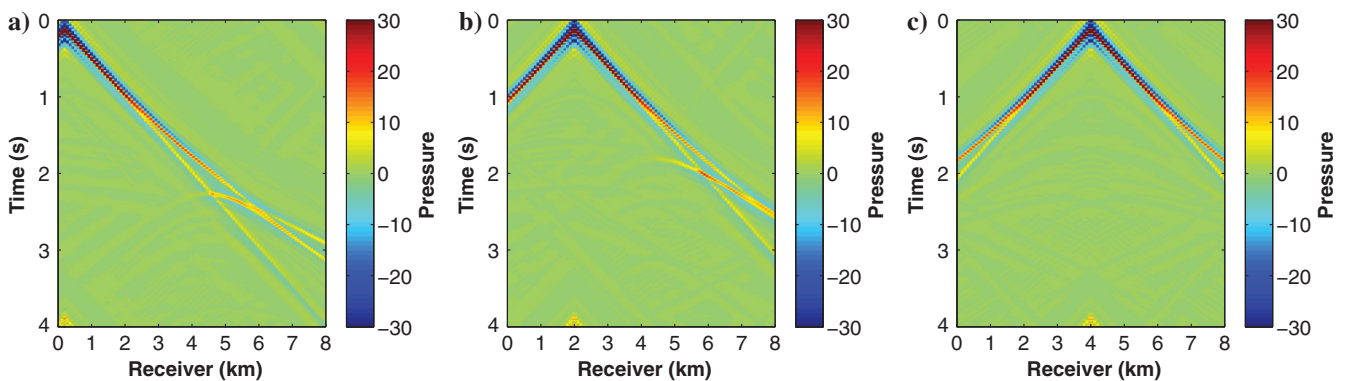


Figure 21. Diving wave example: shot gathers of simulated data with final source-receiver extended waveform inversion + FWI inverted result (Figure 20) for shot at (a) $x_s = 0.2$ km, (b) $x_s = 2$ km, and (c) $x_s = 4$ km.

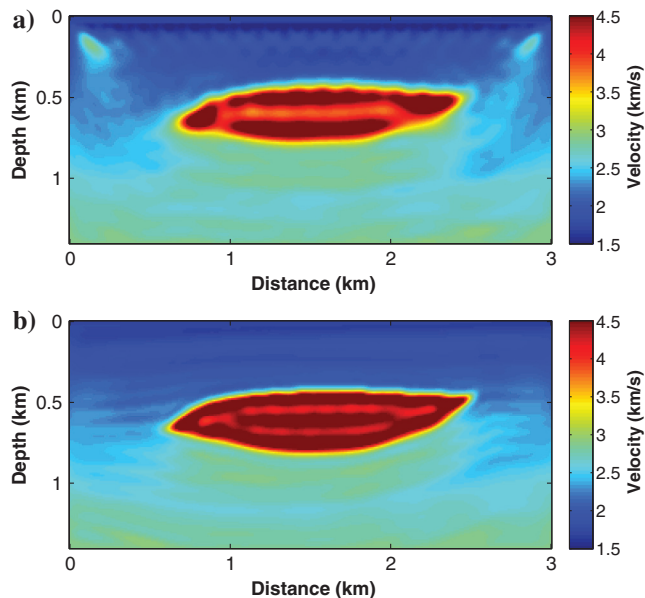


Figure 24. Pluto example: (a) inverted result produced by regularized source-receiver extended waveform inversion after 200 iterations using 4–10 Hz data and (b) FWI result after 50 iterations using 4–16 Hz data with regularized source-receiver extended waveform inversion velocity in (a) as initial estimate.

$$DG[v]\delta v = -\frac{d\delta}{dt}(t - \tau[v])D\tau[v]\delta v. \quad (31)$$

That is, the derivative of $\bar{S}[v]$ involves convolution with a multiple of $d\delta/dt$ (in this simple case), that is, a shift of the derivative of the input trace, whereas $\bar{S}[v]$ shifts the (undifferentiated) input trace. The Fourier transforms of the two results differ by a factor of frequency. The same is true for the migration operator, that is, the transpose of $D\bar{S}[v]$, and the conclusion remains true in the high-frequency asymptotic sense with the proper amplitudes and without the single-arrival assumption. Iteratively, computing the extended source $\bar{f}[v]$ (solution of equation 24) and auxiliary field \bar{w} (solution of equation 26), using (for example) the conjugate gradient algorithm, results in small mean-square errors, as follows from standard theory (Nocedal and Wright, 1999; Golub and van Loan, 2012). However, the derivative of a root-mean-square (rms) small error trace is not necessarily rms small: It can be as large by a factor of the highest frequency present with significant energy. Therefore, the convergence of the gradient obtained by iterative approximation of the expression in equation 28 can be arbitrarily slower (with sufficient bandwidth) than the convergence of the extended source. See Huang (2016) for an explicit illustration of this effect using a different extension, and see Symes (2015) for recent theoretical progress in using special accelerators for iterative solution of the inner problem (equations 24 and 26) to accelerate convergence of the gradient.

Note that the use of circulant convolution and the consequent machine-precision solution of the inner problem by regularized de-

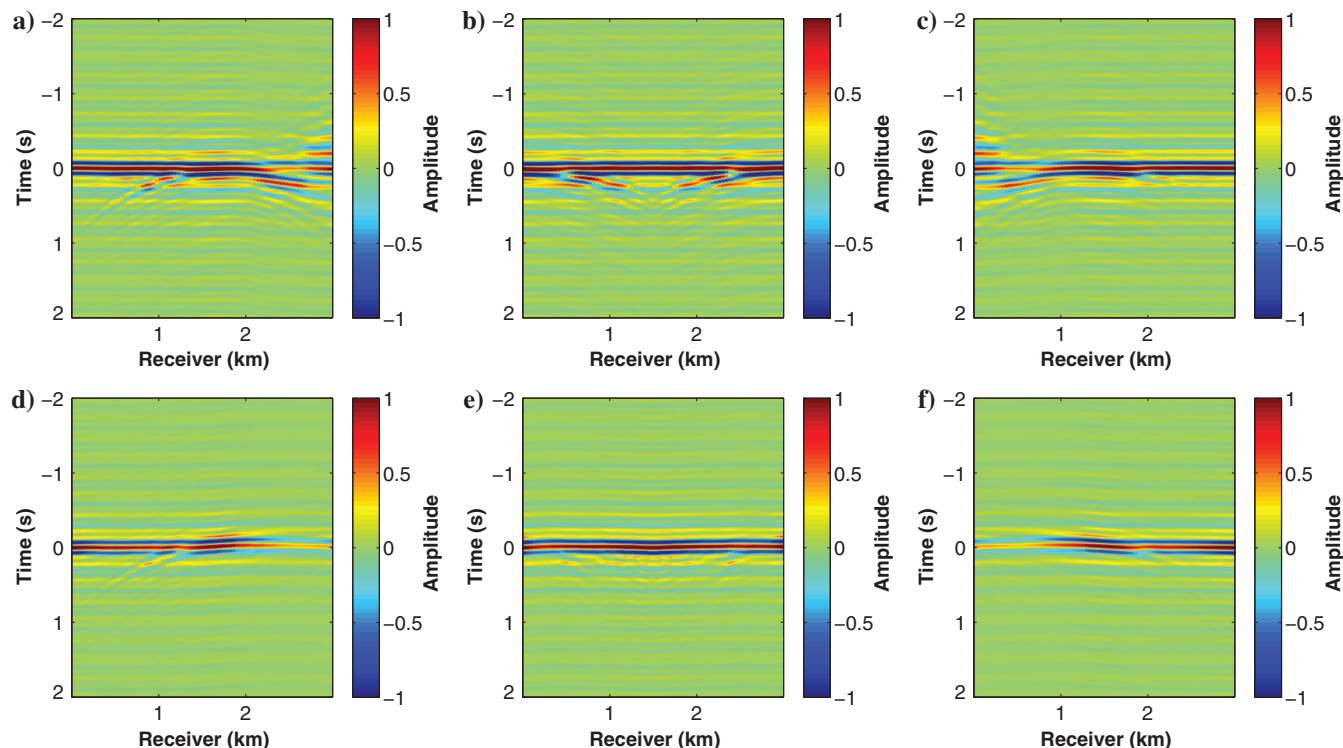


Figure 25. Pluto example: top row: regularized extended source functions with initial velocity for shot at (a) $x_s = 0.06$ km, (b) $x_s = 1.5$ km, and (c) $x_s = 2.94$ km; bottom row: regularized extended source functions with regularized source-receiver extended waveform inversion velocity for shot at (d) $x_s = 0.06$ km, (e) $x_s = 1.5$ km, and (f) $x_s = 2.94$ km.

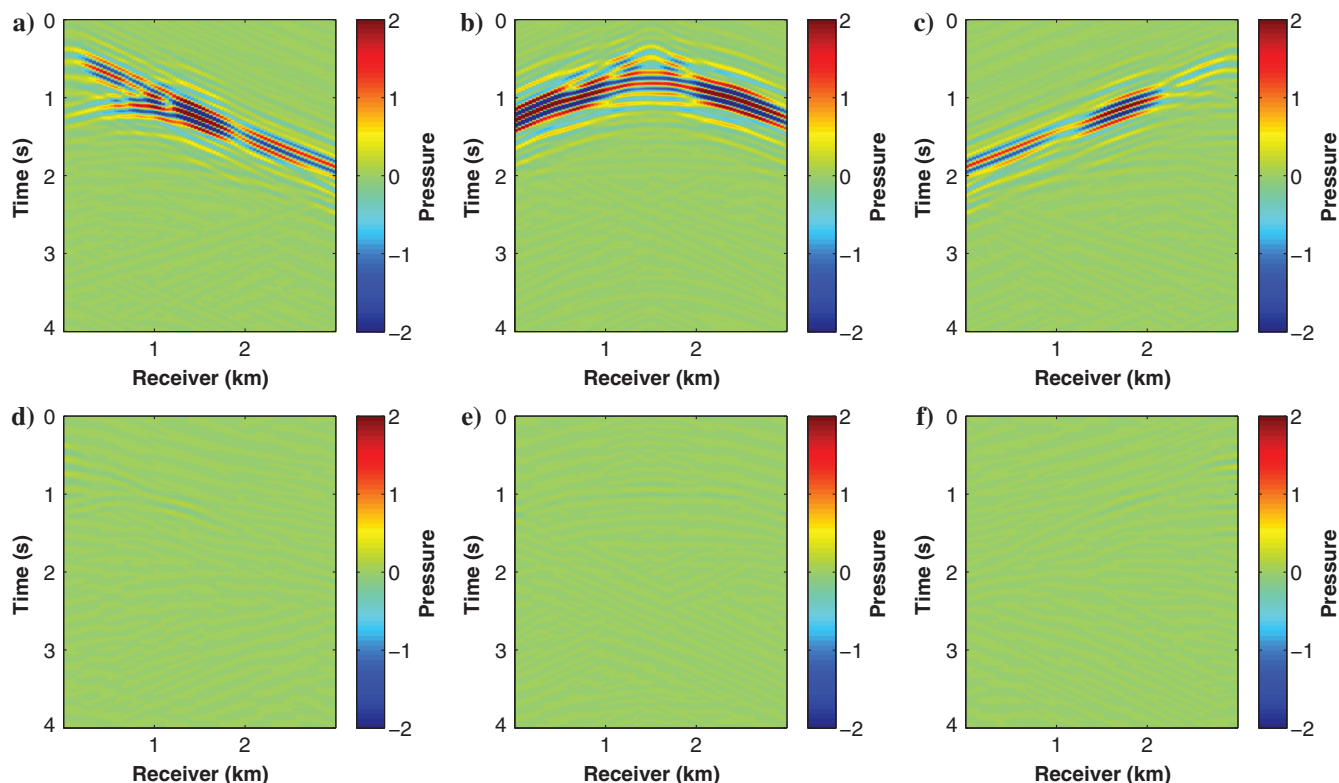


Figure 26. Pluto example: (a) data residual between recorded data and data simulated with initial velocity for shot at (a) $x_s = 0.06$ km, (b) $x_s = 1.5$ km, and (c) $x_s = 2.94$ km. (b) Data residual between recorded data and simulated data using source-receiver extended waveform inversion + FWI velocity (Figure 24b) for shot at (d) $x_s = 0.06$ km, (e) $x_s = 1.5$ km, and (f) $x_s = 2.94$ km.

convolution eliminates the source of the outer gradient inaccuracy just described.

The sharp-eyed reader familiar with Warner's AWI algorithm (Warner and Guasch, 2014, 2016) will notice that our definition of the objective function (equation 13) lacks the normalization by the mean square of \bar{f} used in AWI. There are excellent reasons to normalize such an objective. The principal reason is the amplitude trade-off that occurs in reflection: Data fit to a reflection are not affected, to first order, by scaling the reflectivity up and the source down by the same factor. Therefore, one would not expect an effective velocity update in the reflection case without normalization of the source. The reader might well ask how we got away with it. The answer is that transmitted waves do not suffer from this scale ambiguity, and all of our examples contained transmitted waves. In particular, the data in the reflection-dominated example (Pluto) contained direct wave energy because the boundary conditions used here are absorbing in all directions, so the dipole effect of the free surface is absent. Clearly, this device is not one to rely upon, and a modification of our algorithm to normalize J following the model of Warner and Guasch (2014, 2016) (and many previous works on joint source-model estimation, such as Minkoff and Symes, 1997) is indicated. As pointed out by Warner and Guasch (2014, 2016), the additional computational expense from normalization is minimal.

CONCLUSION

We have presented an extended modeling approach to overcome the tendency of FWI to stagnate at uninformative models. The key

ingredient in this approach is the addition of parameters to the model, in the form of unphysical source parameters, which allow the model to fit the data at every stage of the inversion. Specifically, we allow the source pulse to depend on the source and receiver coordinates. Other choices of source extension are possible (see below). To eliminate the unphysical additional parameters and recover an FWI solution, we have imposed a penalty on the variance of the extended source across the source and receiver. Because the extended source is uniquely determined by the data and the velocity model, the penalty is a function of the velocity.

We have analyzed the use of this penalty as an inversion objective function. We find that in the simplest case, single-arrival transmission data, this objective is closely related to the mean-square traveltime error, so that its minimization yields a tomographic-quality result, even starting from initial models that would be grossly cycle skipped for straightforward FWI. For more complex data with multiple arrivals, the link to traveltime inversion is much more tenuous, and strong regularization is required to produce global convergence. In either case, an iterative local optimization method applied to source-receiver extended waveform inversion produces a sequence of models approaching a kinematically accurate model, even starting with a grossly inaccurate initial velocity estimate. The approach is close enough that convergent FWI iteration can be started from a source-receiver extended waveform inversion solution. Thus, in a sense, the combination source-receiver extended waveform inversion + FWI appears to globalize the convergence of FWI, at least in some cases.

ACKNOWLEDGMENTS

The authors express their gratitude to assistant editor V. Socco, associate editor A. Baumstein, reviewer M. Warner, and two anonymous reviewers for their insightful and helpful comments. G. Huang and W. Symes are grateful to TOTAL E&P USA and to the sponsors of The Rice Inversion Project for their support. We thank TOTAL E&P USA for permission to publish this study. We acknowledge the Rice University Computing Support Group for providing HPC resources that have contributed to the research results reported here.

APPENDIX A

COMPUTATION OF GRADIENT

In this section, we show how to compute the gradient of the inversion velocity analysis objective function via trace-based extension. The directional derivative of the objective function J with respect to velocity v in the direction (velocity perturbation) δv is given by

$$DJ[v]\delta v = \left\langle A^T A \bar{f}[v], D\bar{f}[v]\delta v \right\rangle. \quad (\text{A-1})$$

Using the normal equation 24, we have

$$\begin{aligned} & (\bar{S}[v]^T \bar{S}[v] + \varepsilon^2 I) D\bar{f}[v]\delta v \\ &= (D\bar{S}[v]\delta v)^T d - D(\bar{S}[v]^T \bar{S}[v] + \varepsilon^2 I) \bar{f}[v] \\ &= (D\bar{S}[v]\delta v)^T (d - \bar{S}[v] \bar{f}[v]) - \bar{S}[v]^T (D\bar{S}[v]\delta v) \bar{f}[v]. \end{aligned} \quad (\text{A-2})$$

Let us introduce an auxiliary extended source $\bar{w}[v](\mathbf{x}_r, t; \mathbf{x}_s)$, such that

$$(\bar{S}[v]^T \bar{S}[v] + \varepsilon^2 I) \bar{w}[v] = A^T A \bar{f}[v]. \quad (\text{A-3})$$

Then, by combining equations A-2 and A-3

$$\begin{aligned} DJ[v]\delta v &= \left\langle \bar{w}[v], (D\bar{S}[v]\delta v)^T (d - \bar{S}[v] \bar{f}[v]) \right\rangle \\ &\quad - \left\langle \bar{w}[v], \bar{S}[v]^T (D\bar{S}[v]\delta v) \bar{f}[v] \right\rangle \\ &= \left\langle (D\bar{S}[v]\delta v) \bar{w}[v], d - \bar{S}[v] \bar{f}[v] \right\rangle \\ &\quad - \left\langle \bar{S}[v] \bar{w}[v], (D\bar{S}[v]\delta v) \bar{f}[v] \right\rangle. \end{aligned} \quad (\text{A-4})$$

Note that $(D\bar{S}[v]\delta v) \bar{f}[v](\mathbf{x}_r, t; \mathbf{x}_s) = \bar{f}[v] * (D\bar{S}[v]\delta v) \delta_t(\mathbf{x}_r, t; \mathbf{x}_s)$, and that $\left\langle \bar{S}[v] \bar{w}[v], \bar{f}[v] * (D\bar{S}[v]\delta v) \delta_t \right\rangle = \left\langle \bar{f}[v] \star \bar{S}[v] \bar{w}[v], (D\bar{S}[v]\delta v) \delta_t \right\rangle$, where \star denotes crosscorrelation and δ_t is the delta function of t for every trace. A similar transformation of the other term above shows that

$$DJ[v]\delta v = \left\langle \bar{r}[v], (D\bar{S}[v]\delta v) \delta_t \right\rangle, \quad (\text{A-5})$$

where the “residual” $\bar{r}[v](\mathbf{x}_r, t, \mathbf{x}_s)$ is (note that this is a trace-by-trace computation)

$$\bar{r}[v] = \bar{w}[v] \star (d - \bar{S}[v] \bar{f}[v]) - \bar{f}[v] \star \bar{S}[v] \bar{w}[v]. \quad (\text{A-6})$$

Therefore,

$$\nabla J[v] = (D\bar{S}[v](\cdot) \delta_t)^T \bar{r}[v]. \quad (\text{A-7})$$

The transpose of the impulsive Born simulation operator $D\bar{S}[v](\cdot) \delta_t$ is the well-known impulsive reverse time migration operator, that is, the zero-lag crosscorrelation between the incident Green function G and the adjoint field with the residual $\bar{r}[v]$ as the adjoint source.

REFERENCES

- Abubakar, A., G. Pan, M. Li, L. Zhang, T. Habashy, and P. van den Berg, 2011, Three-dimensional seismic full-waveform inversion using the finite-difference contrast source inversion method: *Geophysical Prospecting*, **59**, 874–888.
- Biondi, B., and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-0275.1](https://doi.org/10.1190/segam2012-0275.1).
- Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: *Geophysics*, **79**, no. 3, WA129–WA140, doi: [10.1190/geo2013-0340.1](https://doi.org/10.1190/geo2013-0340.1).
- Biondi, B., and P. Sava, 2004, Wave-equation migration velocity analysis. I: Theory and II: Subsalt imaging examples: *Geophysics*, **52**, 593–623.
- Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, **60**, 1457–1473, doi: [10.1190/1.1443880](https://doi.org/10.1190/1.1443880).
- Chauris, H., and M. Noble, 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization applications synthetic and real data sets: *Geophysical Journal International*, **144**, 14–26, doi: [10.1046/j.1365-246x.2001.00279.x](https://doi.org/10.1046/j.1365-246x.2001.00279.x).
- Chauris, H., and R.-E. Plessix, 2013, Differential waveform inversion: A way to cope with multiples?: 75th Annual International Conference and Exhibition, EAGE, Extended Abstracts, WS06–F06.
- De Hoop, M. V., S. Brandsberg-Dahl, and B. Ursin, 2003, Seismic velocity analysis in the scattering-angle/azimuth domain: *Geophysical Prospecting*, **51**, 295–314, doi: [10.1046/j.1365-2478.2003.00370.x](https://doi.org/10.1046/j.1365-2478.2003.00370.x).
- Engl, H., M. Hanke, and A. Neubauer, 1996, Regularization of inverse problems: Kluwer Academic Publishers.
- Gauthier, O., A. Tarantola, and J. Virieux, 1986, Two-dimensional nonlinear inversion of seismic waveforms: *Geophysics*, **51**, 1387–1403, doi: [10.1190/1.1442188](https://doi.org/10.1190/1.1442188).
- Golub, G. H., and C. F. van Loan, 2012, Matrix computations, 4th ed.: Johns Hopkins University Press.
- Huang, G., and W. Symes, 2015a, Full waveform inversion via matched source extension: 85th Annual International Meeting, SEG, Expanded Abstracts, 1320–1325.
- Huang, G., and W. Symes, 2016a, Matched source waveform inversion: Space-time extension: 86th Annual International Meeting, SEG, Expanded Abstracts, 1426–1431.
- Huang, G., and W. Symes, 2016b, Matched source waveform inversion: Volume extension: 86th Annual International Meeting, SEG, Expanded Abstracts, 1364–1368.
- Huang, Y., 2016, Born waveform inversion in shot coordinate domain: Ph.D. thesis, Rice University.
- Huang, Y., and W. Symes, 2015b, Born waveform inversion via variable projection and shot record model extension: 85th Annual International Meeting, SEG, Expanded Abstracts, 1326–1331.
- Kern, M., and W. Symes, 1994, Inversion of reflection seismograms by differential semblance analysis Algorithm structure and synthetic examples: *Geophysical Prospecting*, **99**, 565–614.
- Khoury, A., W. Symes, P. Williamson, and P. Shen, 2006, DSR migration velocity analysis by differential semblance optimization: 76th Annual International Meeting, SEG, Expanded Abstracts, 2450–2454.
- Lameloise, C.-A., H. Chauris, and M. Noble, 2015, Improving the gradient of the image-domain objective function using quantitative migration for a more robust migration velocity analysis: *Geophysical Prospecting*, **63**, 391–404, doi: [10.1111/1365-2478.12195](https://doi.org/10.1111/1365-2478.12195).
- Luo, S., and P. Sava, 2011, A deconvolution-based objective function for wave-equation inversion: 81st Annual International Meeting, SEG, Expanded Abstracts, 2788–2792.
- Minkoff, S. E., and W. Symes, 1997, Full waveform inversion of marine reflection data in the plane-wave domain: *Geophysics*, **62**, 540–553, doi: [10.1190/1.1444164](https://doi.org/10.1190/1.1444164).
- Mulder, W., and F. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: *Geophysics*, **67**, 1184–1191, doi: [10.1190/1.1500380](https://doi.org/10.1190/1.1500380).
- Nocedal, J., and S. Wright, 1999, Numerical optimization: Springer Verlag.
- Nolan, C. J., and W. Symes, 1996, Imaging and coherency in complex structure: 66th Annual International Meeting, SEG, Expanded Abstracts, 359–363.

- Plessix, R.-E., 2000, Automatic cross-well tomography an application of the differential semblance optimization to two real examples: *Geophysical Prospecting*, **48**, 937–951, doi: [10.1046/j.1365-2478.2000.00218.x](https://doi.org/10.1046/j.1365-2478.2000.00218.x).
- Plessix, R.-E., Y.-H. de Roeck, and G. Chavent, 1999, Waveform inversion of reflection seismic data for kinematic parameters by local optimization: *SIAM Journal on Scientific Computation*, **20**, 1033–1052, doi: [10.1137/S1064827596311980](https://doi.org/10.1137/S1064827596311980).
- Plessix, R.-E., W. Mulder, and F. ten Kroode, 2000, Automatic cross-well tomography by semblance and differential semblance optimization theory and gradient computation: *Geophysical Prospecting*, **48**, 913–935, doi: [10.1046/j.1365-2478.2000.00217.x](https://doi.org/10.1046/j.1365-2478.2000.00217.x).
- Pratt, R., 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](https://doi.org/10.1190/1.1444597).
- Pratt, R., and R. Shipp, 1999, Seismic waveform inversion in the frequency domain. Part 2: Fault delineation in sediments using crosshole data: *Geophysics*, **64**, 902–914, doi: [10.1190/1.1444598](https://doi.org/10.1190/1.1444598).
- Pratt, R., and W. Symes, 2002, Semblance and differential semblance optimization for waveform tomography a frequency domain approach: *Journal of Conference Abstracts*, **7**, 183–184.
- Shen, P., 2012, An RTM based automatic migration velocity analysis in image domain: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-0214.1](https://doi.org/10.1190/segam2012-0214.1).
- Shen, P., and H. Calandra, 2005, One-way waveform inversion within the framework of adjoint state differential migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 1709–1712.
- Shen, P., and W. Symes, 2008, Automatic velocity analysis via shot profile migration: *Geophysics*, **73**, no. 5, VE49–VE60, doi: [10.1190/1.2972021](https://doi.org/10.1190/1.2972021).
- Shen, P., and W. Symes, 2015, Horizontal contraction in image domain for velocity inversion: *Geophysics*, **80**, no. 3, R95–R110, doi: [10.1190/geo2014-0261.1](https://doi.org/10.1190/geo2014-0261.1).
- Shen, P., W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 2249–2252.
- Shen, P., W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2135–2139.
- Sirgue, L., and G. Pratt, 2004, Efficient waveform inversion and imaging a strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248, doi: [10.1190/1.1649391](https://doi.org/10.1190/1.1649391).
- Song, H., 1994, On a transmission inverse problem: Ph.D. thesis, Rice University.
- Song, H., and W. Symes, 1993, A differential semblance formulation of crosswell tomography: Technical Report TR92-22, Department of Computational and Applied Mathematics, Rice University.
- Song, H., and W. Symes, 1994, Inversion of crosswell tomography via differential semblance optimization: Technical Report, The Rice Inversion Project (Annual Report).
- Stolk, C. C., M. V. de Hoop, and W. Symes, 2009, Kinematics of shot-geophone migration: *Geophysics*, **74**, no. 6, WCA18–WCA34, doi: [10.1190/1.3256285](https://doi.org/10.1190/1.3256285).
- Stolk, C. C., and W. Symes, 2004, Kinematic artifacts in prestack depth migration: *Geophysics*, **69**, 562–575, doi: [10.1190/1.1707076](https://doi.org/10.1190/1.1707076).
- Stoughton, D., J. Stefani, and S. Michell, 2001, 2-D elastic model for wavefield investigations of subsalt objectives, deep water Gulf of Mexico: 71st Annual International Meeting, SEG, Expanded Abstracts, 1269–1272.
- Sun, D., and W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: [10.1190/segam2012-1190.1](https://doi.org/10.1190/segam2012-1190.1).
- Symes, W., 1994, Inverting waveforms in the presence of caustics: Technical Report, The Rice Inversion Project (Annual Report).
- Symes, W., 2008, Migration velocity analysis and waveform inversion: *Geophysical Prospecting*, **56**, 765–790, doi: [10.1111/gpr.2008.56.issue-6](https://doi.org/10.1111/gpr.2008.56.issue-6).
- Symes, W., 2014, Seismic inverse problems recent developments in theory and practice: *Proceedings of the Inverse Problems — From Theory to Application*, Institute of Physics, 2–5.
- Symes, W., 2015, Algorithmic aspects of extended waveform inversion: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, WS05–A01.
- Symes, W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663, doi: [10.1190/1.1443082](https://doi.org/10.1190/1.1443082).
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- ten Kroode, F., 2014, A Lie group associated to seismic velocity estimation: *Proceedings of the Inverse Problems — From Theory to Application*, Institute of Physics, 142–146.
- van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: *Geophysical Journal International*, **195**, 661–667, doi: [10.1093/gji/ggt258](https://doi.org/10.1093/gji/ggt258).
- Vigh, D., K. Jiao, W. Huang, N. Moldoveanu, and J. Kapoor, 2013, Long-offset-aided full-waveform inversion: 13th International Congress of the Brazilian Geophysical Society & EXPOGEF, 294–297.
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC127–WCC152, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Wang, C., and D. Yingst, 2016, Full-waveform inversion with the reconstructed wavefield method: 86th Annual International Meeting, SEG, Expanded Abstracts, 1237–1241.
- Warner, M., and L. Guasch, 2014, Adaptive waveform inversion: Theory: 84th Annual International Meeting, SEG, Expanded Abstracts, 1089–1093.
- Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: *Geophysics*, **81**, no. 6, R429–R445, doi: [10.1190/geo2015-0387.1](https://doi.org/10.1190/geo2015-0387.1).
- Weibull, W., and B. Arntsen, 2014, Anisotropic migration velocity analysis using reverse time migration: *Geophysics*, **79**, no. 1, R13–R25, doi: [10.1190/geo2013-0108.1](https://doi.org/10.1190/geo2013-0108.1).
- Yang, T., and P. Sava, 2011, Wave-equation migration velocity analysis with time-shift imaging: *Geophysical Prospecting*, **59**, 635–650, doi: [10.1111/j.1365-2478.2011.00954.x](https://doi.org/10.1111/j.1365-2478.2011.00954.x).