Phase and amplitude tracking for seismic event separation

Citation: Li, Yunyue Elita, and Laurent Demanet. "Phase and Amplitude Tracking for Seismic Event Separation." GEOPHYSICS 80, no. 6 (November 2015): WD59–WD72. © 2015 Society of Exploration Geophysicists

As Published: http://dx.doi.org/10.1190/GEO2015-0075.1

Publisher: Society of Exploration Geophysicists

Persistent URL: http://hdl.handle.net/1721.1/100553

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

GEOPHYSICS, VOL. 80, NO. 6 (NOVEMBER-DECEMBER 2015); P. WD59–WD72, 18 FIGS., 3 TABLES.
10.1190/GEO2015-0075.1

Phase and amplitude tracking for seismic event separation

Yunyue Elita Li[1] and Laurent Demanet[1]

ABSTRACT

We have developed a method to decompose seismic records into atomic events, each defined by a smooth phase function and a smooth amplitude function. This decomposition is intrinsically nonlinear and calls for a nonconvex least-squares optimization formulation, along the lines of full-waveform inversion. To overcome the lack of convexity, we have developed an iterative refinement-expansion scheme to initialize and track the phase and amplitude for each atomic event. For short, we called the method phase tracking. The initialization is carried out by applying multiple signal classification to a few seed traces in which events can be separated and identified by their arrival times and amplitudes. We then construct the initial solution at the seed traces using linear phase functions from the arrival times and constant amplitude functions, assuming the medium is mostly dispersion free. We refine this initial solution to account for dispersion and imperfect knowledge of the wavelet at the seed traces by fitting the observed data using a gradient descent method. The resulting phase and amplitude functions are then carefully expanded across the traces in an adequately smooth way to match the whole data record. We have evaluated the proposed method on two synthetic records and a field record. Because the parametrization of the seismic events is physically meaningful, it also enables a simple form of bandwidth extension of the observed shot record to unobserved low and high frequencies. We tested this procedure on the same shot records. Bandwidth extension is in principle helpful to initialize full-waveform inversion with frequency sweeps and enhanced its resolution.

INTRODUCTION

In this paper, we address the problem of decomposing a seismic record into elementary, or atomic, components corresponding to individual wave arrivals. Letting t stand for time and x represent the receiver location, we seek to decompose a shot profile d into a small number r of atomic events v_j as

$$d(t,x) \simeq \sum_{j=1}^{r} v_j(t,x). \tag{1}$$

Each v_j should consist of a single wavefront — narrow yet band limited in t, but coherent across different x — corresponding to an event of direct arrival, reflection, refraction, or corner diffraction.

In the simplest convolutional model, we would write $v_j(t,x) = a_j(x)w(t-\tau_j(x))$ for some wavelet w, amplitude $a_j(x)$, and time shift $\tau_j(x)$. In the Fourier domain, this model would read

$$\hat{v}_j(\omega,x) = \hat{w}(\omega)a_j(x)e^{i\omega\tau_j(x)}. \tag{2}$$

This model fails to capture frequency-dependent dispersion and attenuation effects, phase rotations, inaccurate knowledge of the source wavelet, and other distortion effects resulting from near resonances. To restore the flexibility to encode such effects without explicitly modeling them, we consider instead throughout this paper the following expression:

$$\hat{v}_j(\omega,x) = \hat{w}(\omega)a_j(\omega,x)e^{ib_j(\omega,x)}, \tag{3}$$

where the amplitudes a_j and the phases b_j are smooth in x and ω, and b_j deviates little from an affine (linear + constant) function of ω. If the guess $\hat{w}(\omega)$ for the wavelet is not entirely inaccurate, the amplitude $a_j(\omega,x)$ is meant to compensate for it.

Finding physically meaningful, smooth functions a_j and b_j to fit a model such as equations 1 and 3 is a hard optimization problem. Its nonconvexity is severe: It can be seen as a remnant, or cartoon, of the difficulty of full-waveform inversion from high-frequency data. We are unaware that an authoritative solution to either problem has been proposed in the geophysical imaging community.

Manuscript received by the Editor 8 February 2015; revised manuscript received 2 June 2015; published online 1 September 2015.

[1] Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. E-mail: yunyueli@mit.edu; demanet@gmail.com.

Many methods have been proposed to pick individual seismic events, such as autoregression (AR) filters (Leonard and Kennett, 1999) (close in spirit to the matrix pencil method [Hua and Sarkar, 1990]), crosscorrelations (Cansi, 1995), wavelets (Zhang et al., 2003), neural networks (Gentili and Michelini, 2006), etc. These papers mostly address the problem of picking isolated arrivals time, not parameterizing interfering events across traces. Some data processing methods operate by finding local slope events, such as plane-wave annihilation (Fomel, 2002). This idea have been used to construct prediction filters for localized wavelet-like expansion methods (Fomel and Liu, 2010), which in turn allow us to solve problems such as trace interpolation in a convincing manner. It is plausible that concentration or clustering in an appropriate waveletlike domain could be the basis for an algorithm of event separation. Separation of variables in moveout-corrected coordinates has also been proposed to identify dipping events, such as in Raoult (1983) and Blias (2007). Smoothness criteria along reflection events have been proposed for separating them from diffraction events, such as in Fomel et al. (2007). These traditional methods fail for the event decomposition problem as previously stated:

- Because of cycle-skipping, the gradient descent quickly converges to uninformative local minimums.
- Because data sets do not often have useful low frequencies, multiscale sweeps cannot be seeded to guide gradient descent iterations toward the global minimum.
- Because the events are intertwined by possibly destructive interference, simple counterexamples show that greedy "event removal" methods such as matching pursuit cannot be expected to succeed in general.
- Because wavefront shapes are not known in advance, linear transforms such as the slant stack (Radon), velocity scan, wavelets/curvelets, or any other kind of nonadaptive filters, do not suffice by themselves. The problem is intrinsically nonlinear.

The contribution of this paper is the observation that tracking in x and ω, in the form of careful growth of a trust region, can satisfactorily mitigate the nonconvexity of a simple nonlinear least-squares cost function, yielding favorable decomposition results on some synthetic shot profiles and some field data. We have not been able to deal with the nonconvexity of this cost function in any other way than by tracking.

Seismic records from field experiments contain many types of events resulting from the interaction between the source and the complex subsurface. Separating these events has been a long-standing challenge in seismic data processing. The successful resolution of this question would have many implications:

- It would improve our ability to discriminate events based on the different physical processes that generated them: P-S wave separation, primary-multiple separation, reflection-diffraction separation, simultaneous recording separation, etc.
- It automates the "traveltime picking" operation; hence, this may help prepare a data set for traveltime tomography.
- A description of the data set in terms of phases and amplitudes is adequate for interpolation of missing samples. It is also the proper domain in which model reduction should be performed in the high-frequency regime.
- Perhaps most importantly, it enables a not entirely inaccurate extrapolation to high and low frequencies not present in the data set, with a possible application to seeding frequency sweeps in full-waveform inversion.

This article is organized as follows: We first define the objective function and derive its gradient. We point out that the objective function is severely nonconvex, and we explain an explicit initialization scheme with multiple signal classification (MUSIC) and the expansion and refinement scheme for phase and amplitude tracking. Finally, we demonstrate our separation algorithm on two synthetic records and a field record. We illustrate the potential of event separation for extrapolation to unobserved frequencies in the last two examples.

METHOD

Cost function and its gradient

We consider the nonlinear least-squares optimization formulation with a Tikhonov-like cost function:

$$J(\{a_j, b_j\}) = \frac{1}{2}\|\hat{u}(\omega,x) - \hat{d}(\omega,x)\|_2^2 + \lambda\sum_j \|\nabla_\omega^2 b_j(\omega,x)\|_2^2 + \mu\sum_j \|\nabla_x b_j(\omega,x)\|_2^2 + \gamma\sum_j \|\nabla_{\omega,x} a_j(\omega,x)\|_2^2, \tag{4}$$

where $\hat{d}$ is the measured data in the frequency domain, ∇_k and ∇_k^2, with $k = \omega, x$, respectively, denote first-order and second-order partial derivatives, and $\nabla_{\omega,x}$ denotes the full gradient. The prediction $\hat{u}$ is

$$\hat{u}(\omega,x) = \sum_{j=1}^{r} \hat{w}(\omega) a_j(\omega,x) e^{ib_j(\omega,x)}, \tag{5}$$

with the wavelet $\hat{w}(\omega)$ assumed known to a certain level of accuracy. (Mild phase and amplitude inaccuracies in $\hat{w}(\omega)$ are, respectively, compensated by $b_j(\omega,x)$ and $a_j(\omega,x)$.) The constants λ, μ, and γ are chosen empirically.

It is important to regularize with $\nabla_\omega^2 b_j(\omega,x)$ rather than $\nabla_\omega b_j(\omega,x)$, so as to penalize departure from dispersion-free linear phases rather than to penalize large traveltimes.

The cost function is minimized using a gradient descent method within a growing trust region. The gradients of equation 4 with respect to a_j and b_j are computed as follows:

$$\frac{\partial J}{\partial a_j} = \frac{1}{2}\left(\frac{\partial u}{\partial a_j}\delta\hat{u}^* + \delta\hat{u}\frac{\partial u^*}{\partial a_j}\right) + 2\gamma\nabla_{\omega,x}^2 a_j,$$
$$\frac{\partial J}{\partial b_j} = \frac{1}{2}\left(\frac{\partial u}{\partial b_j}\delta\hat{u}^* + \delta\hat{u}\frac{\partial u^*}{\partial b_j}\right) + 2\lambda\nabla_\omega^2 \cdot \nabla_\omega^2 b_j + 2\mu\nabla_x^2 b_j, \tag{6}$$

where

$$\frac{\partial\hat{u}}{\partial a_j} = \hat{w}e^{-ib_j}, \quad \frac{\partial\hat{u}^*}{\partial a_j} = \hat{w}^* e^{ib_j},$$
$$\frac{\partial\hat{u}}{\partial b_j} = -i\hat{w}a_j e^{-ib_j}, \quad \frac{\partial\hat{u}^*}{\partial b_j} = i\hat{w}^* a_j e^{ib_j},$$
$$\delta\hat{u} = \hat{u} - \hat{d}. \tag{7}$$

The notation $\nabla^2_{\omega,x}$ refers to the Laplacian in (ω, x) and $*$ denotes complex conjugation. All the derivatives in the regularization terms are discretized by centered second-order accurate finite differences.

The inverse problem defined by the objective function in equation 4 is clearly nonconvex, in large part due to the oscillatory nature of seismic data. Matching oscillations pointwise generically runs into the cycle-skipping problem. Lack of convexity is also found in phase-retrieval problems (Demanet and Hand, 2014; Gholami, 2014), although in a less severe form.

Initialization

We initialize the iterations by making use of an explicit solution of the minimization problem in a very confined setting where

- we pick a single seed trace x in which the events of interests are well-separated;
- we pick a subset of seed frequencies ω around the dominant frequency of the source wavelet;
- after deconvolving the source wavelet, we assume a simplified model in which the amplitudes are constant, and the phases are linear in ω. As a result, we momentarily return to the convolutional model,

$$\hat{d}(\omega) \simeq \sum_{j=1}^{r} a_j e^{i\omega\tau_j}, \tag{8}$$

of equation 2 to locally approximate equation 3.

In this situation, the problem reduces to a classical signal processing question of identification of sinusoids, i.e., identification of the traveltimes τ_j and amplitudes a_j. There exist at least two high-quality methods for this task: the matrix pencil method of Hua and Sarkar (1990) and the MUSIC algorithm (Schmidt, 1986; Biondi and Kostov, 1989; Kirlin and Done, 1999). We choose the latter for its simplicity and robustness. (It would be a mistake to use either ℓ_1 minimization, or iterative removal pursuit algorithms, for this sinusoid identification subproblem.)

Assume for the moment that the number r of events is known, although we address its determination later in the paper. The MUSIC algorithm only needs data $\{\hat{d}(\omega_i)\}$ at $m = 2r + 1$ frequencies to determine the arrival times and amplitudes for r different events. In practice, the number m of frequencies may be taken to be larger than $2r + 1$ if robustness to noise is a more important concern than the lack of linearity of the phase in ω. In either case, we sample the data $\hat{d}(\omega_i)$ on a grid of spacing $\Delta\omega$ around the dominant frequency ω_f of the source wavelet, where the signal-to-noise ratio (S/N) is relatively high.

The variant of the MUSIC algorithm that we use in this article requires building a Toeplitz matrix, whose columns are constructed by translations of the data samples as

$$\mathbf{T} = \begin{bmatrix} \hat{d}(\omega_f) & \hat{d}(\omega_{f-1}) & \cdots & \hat{d}(\omega_{f-k+1}) \\ \hat{d}(\omega_{f+1}) & \hat{d}(\omega_f) & \cdots & \hat{d}(\omega_{f-k+2}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{d}(\omega_{f+k-1}) & \hat{d}(\omega_{f+k-2}) & \cdots & \hat{d}(\omega_f) \end{bmatrix}, \tag{9}$$

where $k = (m - 1)/2$. After a singular value decomposition (SVD) of $\mathbf{T}$, we separate the components relative to the r largest singular values, from the others, to get

$$\mathbf{T} = \mathbf{U}_s\Sigma_s\mathbf{V}_s^T + \mathbf{U}_n\Sigma_n\mathbf{V}_n^T. \tag{10}$$

We interpret the range space of $\mathbf{U}_s$ as the signal space and the range space of $\mathbf{U}_n$ as the noise space, hence the choice of indices. The orthogonal projector onto the noise subspace can be constructed from $\mathbf{U}_n$ as

$$\mathbf{P}_n = \mathbf{U}_n\mathbf{U}_n^T. \tag{11}$$

We then consider a quantitative measure of the importance of any given arrival time $t \in [0, \frac{2\pi}{\Delta\omega}]$, via the estimator function

$$\alpha(t) = \frac{1}{\|\mathbf{P}_n e^{i\boldsymbol{\omega}t}\|}. \tag{12}$$

In the exponent, $\boldsymbol{\omega}$ is a vector with k consecutive frequencies on a grid of spacing $\Delta\omega$.

In practice, the number of individual events is a priori unknown. The determination of this number is typically linked to the extraneous knowledge of the noise level, by putting to zero the eigenvalues of T below some threshold ϵ, chosen so that the resulting error vector has a magnitude that matches the predetermined level.

In the noiseless case, the estimator function in equation 12 has r sharp peaks that indicate the r arrival times τ_j for the r events. In the noisy case, or in the case when the phases b_j are nonlinear in ω, the number of identifiable peaks is a reasonable estimator for r, and the locations of those peaks are reasonable estimators of τ_j, such that the signal contains r phases locally of the form $\omega\tau_j +$ a constant. In field applications, this procedure should be applied in a trust region Ω, which contains a handful of nearby traces. The consistent arrival times $\tau_j(x)$ across traces provide a more robust estimation for the coherent events. Once the traveltimes τ_j are found, the amplitudes a_j follow from solving the small, overdetermined system in equation 8. Each complex amplitude is further factored into a positive amplitude and a phase rotation factor, and the latter is absorbed into the phase function. We summarize the initialization procedure in Algorithm 1.

Algorithm 1. Initialization with MUSIC.

```
Select a trust region $\Omega$ and a threshold level $\epsilon$
for x $\in \Omega$ do
    Build the Toeplitz matrix $\mathbf{T}$ and SVD
    Build the projection matrix $\mathbf{P}_n$ with threshold $\epsilon$
    Compute the estimator $\alpha(t)$
    Pick the peaks of $\alpha(t)$ to determine $r(x)$ arrival times $\tau_j(x)$
    Solve for the amplitude $a_j(x)$
end for
Find the consistent $\tau_j(x)$, $a_j(x)$ and determine $r$
```

Tracking by expansion and refinement

The phases and amplitudes generated by the initialization give local seeds that need to be refined and expanded to the whole record:

- The refinement step is the minimization of equation 4 with the data misfit restricted to the current trust region Ω. In other words, we use the result of initialization in the form of equation 2 as an initial guess and upgrade it to take the more accurate form (equation 3).
- The expansion step consists in growing the region Ω to include the neighboring frequency samples and extending the solution smoothly to the neighboring traces.

These steps are nested rather than alternated: The inner refinement loop is run until the value of J levels off, before the algorithm returns to the outer expansion loop. The expansion loop is itself split into an outer loop for (slow) expansion in x and an inner loop for (rapid) expansion in ω. The nested ordering of these steps is crucial for convergence to a meaningful minimizer.

A simple trick is used to speed up the minimization of J in the complement of Ω, where only the regularization terms are active: $a_j(\omega, x)$ is extended constantly in x and ω, whereas $b_j(\omega, x)$ is extended constantly in x and linearly in ω. These choices correspond to the exact minimizers of the Euclidean norms of the first and second derivatives of a_j and b_j, respectively, with zero-boundary conditions on the relevant derivatives at the endpoints. This trick may be called preconditioning the regularization terms.

Algorithm 2. Phase and amplitude tracking algorithm.

```
Input: observed data $\tilde{d}(\omega, x)$, estimated wavelet $w(\omega)$, trust region $\Omega = [x_{beg} \quad x_{end}]$
MUSIC: run Algorithm 1 to obtain $r$, $a_j(x)$, and $\tau_j(x)$
    initialize $a_j^0(\omega, x) \leftarrow \Re(a_j(x))$, and
        $b_j^0(\omega, x) \leftarrow \omega\tau_j(x) + \Im(a_j(x))$
for $i = 0, 1, \dots N-1$ do
    Forward modeling: obtain $\tilde{u}(\omega, x)$
    Objective function evaluation: compute $J$ within the range of $\Omega$
    if $J < \epsilon$ then
        if $\Omega$ covers the whole record
            Converge: Output $a_j(\omega, x)$ and $b_j(\omega, x)$
        else
            Expand: $\Omega \leftarrow [x_{beg} - \Delta x x_{end} + \Delta x]$
            $a_j(\omega, x_{beg} - \Delta x) \leftarrow a_j(\omega, x_{beg})$, $b_j(\omega, x_{beg} - \Delta x) \leftarrow b_j(\omega, x_{beg})$
            $a_j(\omega, x_{end} + \Delta x) \leftarrow a_j(\omega, x_{end})$, $b_j(\omega, x_{end} + \Delta x) \leftarrow b_j(\omega, x_{end})$
        end if
    else
        Compute gradient: $\frac{\partial J}{\partial a_j}$ and $\frac{\partial J}{\partial b_j}$
        Update model: $a_j^{i+1} \leftarrow a_j^i - \beta \frac{\partial J}{\partial a_j}$, $b_j^{i+1} \leftarrow b_j^i - \beta \frac{\partial J}{\partial b_j}$
    end if
end for
```

At the conclusion of this main algorithm, the method returns r phases $b_j(\omega, x)$ that are approximately linear in ω and approximately constant in x; and r amplitudes $a_j(\omega, x)$ that are approximately constant in x and ω such that

$$\hat{d}(\omega, x) \simeq \sum_{j=1}^{r} \hat{w}(\omega) a_j(\omega, x) e^{i b_j(\omega, x)}. \tag{13}$$

We summarize the procedure of the refinement and expansion in Algorithm 2.

Further in the paper, we refer to the respective plots of b_j and a_j as functions of ω and x as "phase spectra" and "amplitude spectra."

Application: Frequency extrapolation

Frequency extrapolation, also known as bandwidth extension, is a tantalizing test of the quality of a representation such as equation 8. A least-squares fit is first performed to find the best constant approximations $a_j(\omega, x) \simeq \alpha_j(x)$ and the best affine approximations $b_j(\omega, x) \simeq \omega\beta_j(x) + \phi_j(x)$ from values of ω within a useful frequency band. These phase and amplitude approximations can be evaluated at values of ω outside this band, to yield synthetic flat-spectrum atomic events of the following form:

$$\alpha_j(x) e^{i(\omega\beta_j(x) + \phi_j(x))}. \tag{14}$$

These synthetic events can be further multiplied by a high- or low-pass wavelet, and summed up, to create a synthetic data set. This operation is the seismic equivalent of changing the pitch of a speech signal without speeding it up or slowing it down.

Extrapolation to high frequencies should benefit high-resolution imaging, whereas extrapolation to low frequencies should help avoid the cycle-skipping problem that full-waveform inversion encounters when the low frequencies are missing from the data.

Notice that extrapolation to zero frequency is almost never accurate using such a simple procedure: More physical information is required to accurately predict zero-frequency wave propagation. The exception is free-space wave propagation in 3D, where the Green's function has a linear phase and constant amplitude in ω, hence completely predictable from two samples in ω.

NUMERICAL EXAMPLES

In this section, we demonstrate the tracking method on two synthetic data records and one field data record. The first synthetic example illustrates the workflow behind the iterations. In the second synthetic example, we test our algorithm on a noisy seismic record obtained by finite-difference modeling, with application to frequency extrapolation.

Synthetic example: Three-event separation

In this example, we create an artificial three-event seismic record by convolving a Ricker wavelet along the traveltime curves of one direct arrival event, one water-bottom reflection event, and one reflection event from a deeper layer. The dominant frequency of the Ricker wavelet is 20 Hz. The amplitude of the Ricker wavelet decreases with increasing offset. After generating the data in the time domain, we bring the data to the frequency domain and use the band between 5 and 35 Hz for inversion.

The reference seismic profile is shown in Figure 1a. The events are well-separated when the offset is small. As the offset grows, the three events cross each other. Moreover, the water-bottom reflection event approaches the water-column direct arrival as its asymptote. Therefore, both events become almost perfectly aligned at large offsets over 2.5 km. Separating these crossing and overlapping events is a somewhat challenging task to any automatic event tracker.

To initialize the method, we choose three seed traces where the events are well-separated. We train the MUSIC algorithm with only seven frequencies around the dominant frequency of the Ricker wavelet for each trace. Figure 2a shows one seismic trace at $x = 200$ m. The MUSIC estimator with many trial arrival times is shown in Figure 2b. The peaks of the MUSIC estimator accurately pinpoint the reference arrival times of the events.

With these three seed traces, we run the expansion-refinement tracking algorithm explained earlier. The inverted seismic profile is shown in Figure 1b. The kinematics (phases) and the amplitudes are almost perfectly resolved.

Figure 3 shows the comparison between the inverted atomic events in Figure 3a–3c with the reference atomic events in Figure 3d–3f. The atomic events were separated with high accuracy. Figure 4 shows the phase spectra (Figure 4a–4c) and the amplitude spectra (Figure 4d–4f) for the reference atomic events. In comparison, we show the phase spectra (Figure 5a–5c) and the amplitude spectra (Figure 5d–5f) for the inverted atomic events in Figure 5. The phase spectra for all the events are accurately recovered. However, the amplitude spectra for the first two events are not perfectly recovered due to the ambiguities at large offsets.

Synthetic example: Shallow Marmousi model

In this example, we test our algorithm on a more realistic shot gather. The shot gather is generated from a shallow part of the Marmousi model using finite-difference modeling. The finite-difference scheme is second-order accurate in time and fourth-order accurate in space. We use a 20-Hz Ricker wavelet as the source wavelet. The receiver spacing is 40 m.

As a preprocessing stop, we remove the direct arrival from the shot record because it has the strongest amplitude that would overwhelm the record. We apply an automatic gain control to the remaining events so that the amplitudes on the record is more balanced. We mute the later arrivals in the data so that we resolve a limited number of events at a time. Finally, we add 30% independent and identically distributed Gaussian noise to the data within the bandwidth between 7 and 40 Hz.

Figure 6a shows the shot record with six clear events after preprocessing. For inversion of the atomic events, we consider the data

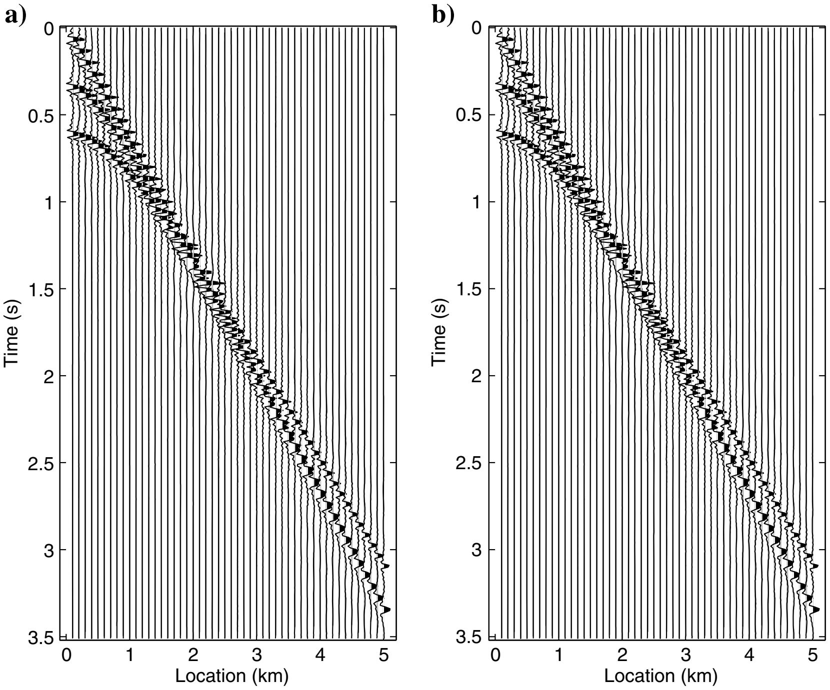


Figure 1. Comparison of the reference seismic profile is shown in panel (a) and the inverted seismic profile is shown in panel (b). Figures are clipped to the same value. The inversion perfectly reconstructs the seismic data.

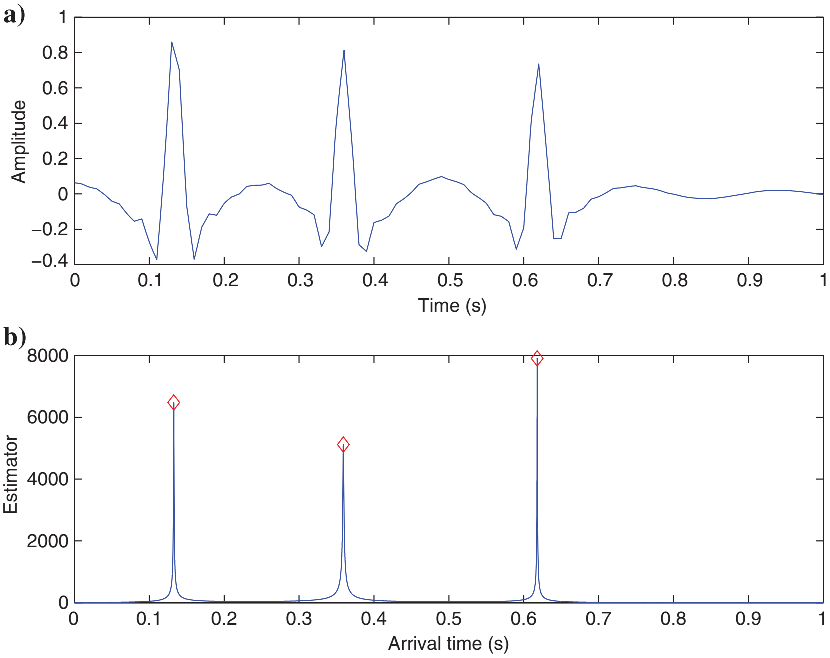


Figure 2. (a) Comparison of the seismic trace at $x = 200$ m and (b) its MUSIC estimator (blue curve) and the picked traveltimes (red diamonds). The MUSIC estimator accurately identifies the arrival time of each event.

only between 7 and 40 Hz. We choose three seed traces at approximately $x = 1.5$ km where the events are well-separated. We apply the MUSIC algorithm using 13 frequencies around the dominant frequency. Figure 7a shows one example of the seed trace at $x = 1.5$ km. The MUSIC estimator accurately determines the arrival times of the six different events, as shown in Figure 7b.

Figure 6b shows the inverted shot profile using the proposed tracking algorithm. The inversion has clearly separated the six strong reflection events from the record, removed the severe random noise, and improved the coherence of the amplitude along the events. This demonstrates the robustness of the method. Figure 8 shows the corresponding well-separated atomic reflection events.

Figures 9 and 10 show the phase and amplitude spectra for each atomic event. As expected, the phase and amplitude spectra are smooth in ω and x. We allow the amplitudes of an event to drop to zero when the average estimated energy of this event at that trace is less than 1% of that at the seed traces. Hence, we allow the events to be terminated in the middle of the section. At each trace x, we estimate the parameters $\alpha_j(x), \beta_j(x)$, and $\phi_j(x)$ and determine the approximated phase spectrum for each event. With this parametrization, we now extrapolate the data to frequencies within [1, 90] Hz.

The atomic events after frequency extrapolation are shown in Figure 11. Compared with the atomic events in Figure 8, the waveform is much more compact thanks to the increased bandwidth, whereas the traveltime information is kept intact. Figure 7 compares the seed trace after frequency extrapolation (Figure 7c) with the original seed trace (Figure 7a). The waveforms are more compact.

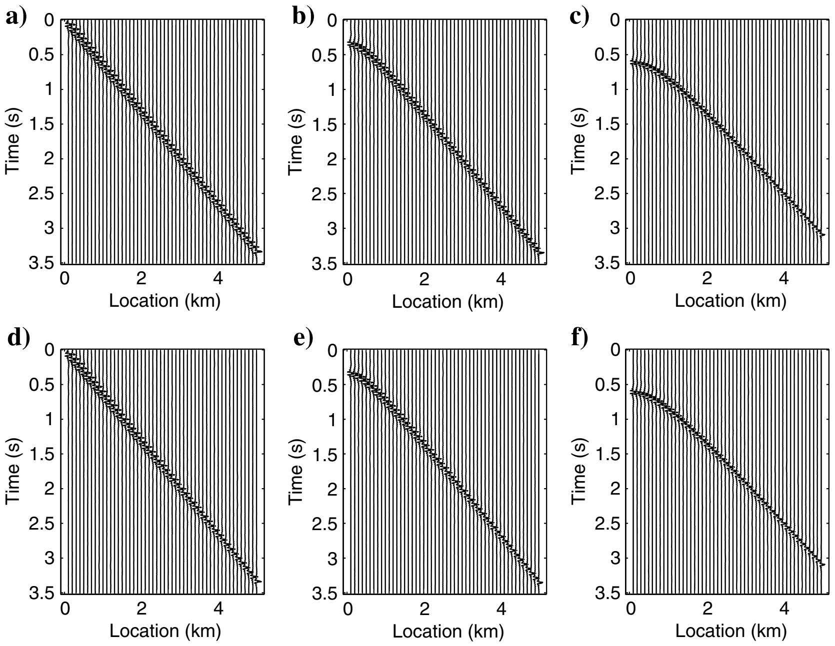


Figure 3. Comparison of (a-c) the inverted atomic events with (d-f) the true atomic events. Figures are clipped to the same value. The inversion accurately reconstructs the atomic events.

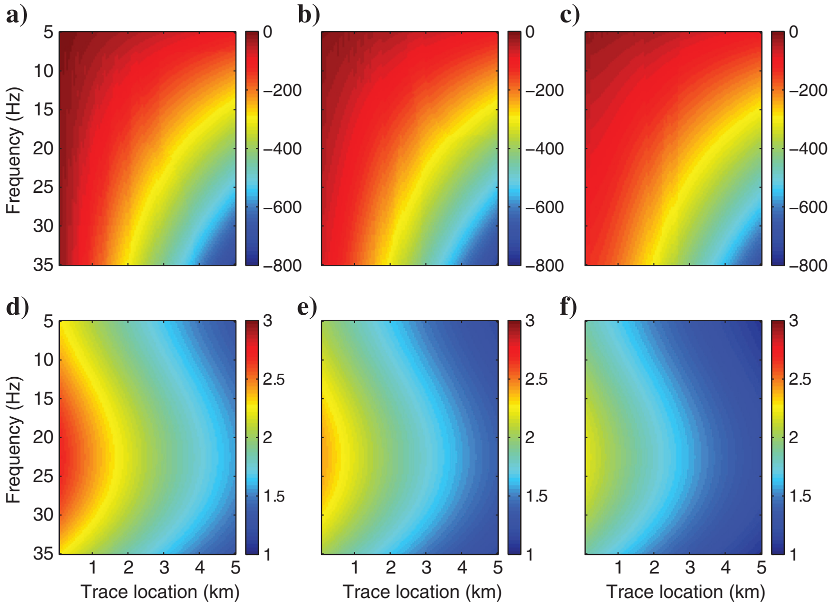


Figure 4. (a-c) Phase spectra and (d-f) amplitude spectra of the true atomic events.

The high-resolution data could be used for broadband high-resolution seismic imaging (Fu et al., 2014).

To evaluate the accuracy of low-frequency extrapolation, we model the seismic record using a broadband source wavelet, whose amplitude spectrum is mostly flat between 1 and 7 Hz. We then compare the modeled data with the data obtained by frequency extrapolation within the same bandwidth in Figure 12. Although the amplitudes differ between the two records, the phase function from the frequency extrapolation represents a reasonable estimation of the modeled phase function.

There are a few reasons to explain the mismatch between the extrapolated low-frequency data and the modeled low-frequency data. First, the unmodeled scattering events contribute to the low frequencies and overlap with the low-frequency signal from the modeled reflection events. Second, there are numerical dispersion effects on the modeled low-frequency data, whereas the extrapolated record is constructed using a nondispersive assumption. Finally, the noise in the original record may introduce extra amplitude and phase perturbations, which do not appear on the noiseless modeled low-frequency record. All of these effects will manifest themselves in the field applications.

Field data example

In this example, we test our algorithm on a shallow field record acquired onshore in China (Wang et al., 2013). Figure 13a shows

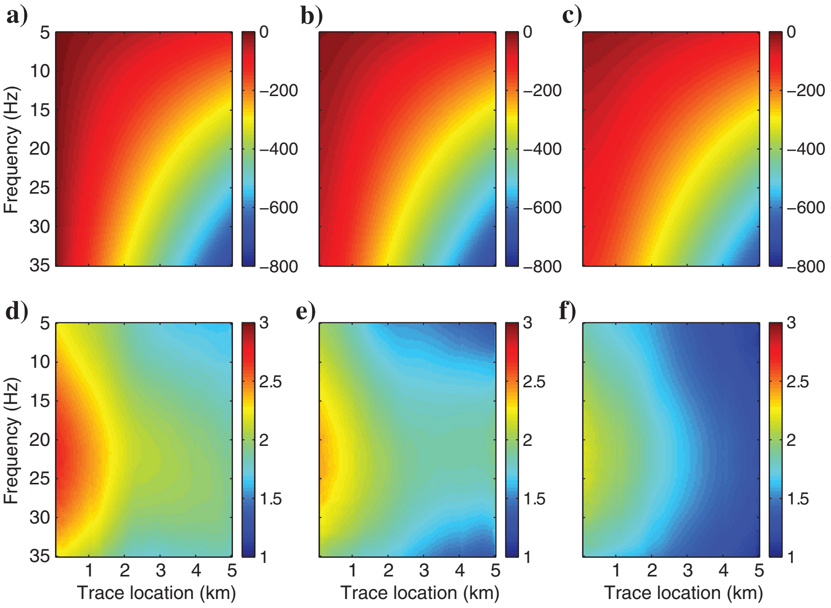


Figure 5. (a-c) Phase spectra and (d-f) amplitude spectra of the inverted atomic events. Plots in this figure are clipped at the same value as the plots in Figure 4, respectively. Notice the amplitudes are well-reconstructed before the three events cross (offset less than 2 km). After the crossing point, the direct arrival and the water-bottom reflection are perfectly overlapped, causing the amplitude spectra (panels d and e) to be not perfectly resolved due to the ambiguity. Nonetheless, the corresponding phase spectra (panels a and b) are well-reconstructed.

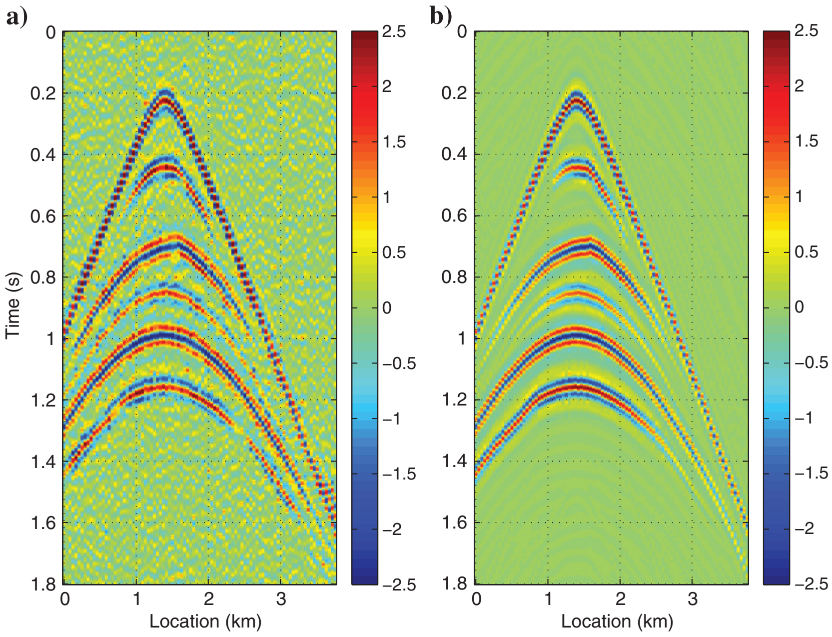


Figure 6. (a) Comparison of the noisy shot record and (b) the inverted shot record. The inverted data have clearly reconstructed the strong reflection events and removed the severe random noise in the data. Amplitude and phase are more coherent along each reflection event.

early events of the record. Strong receiver statics can be observed from the shifting phases across nearby traces. Strong amplitude variations are also observed from this record. Figure 13b shows the average power spectrum of the record. The data have a large S/N between [16, 85] Hz. However, the low-frequency components (less than 10 Hz) are completely missing from the data.

For the inversion of event separation, we further applied an inner mute to the data and band-pass the data between [16, 70] Hz. Figure 14a shows the resulting shot record and we only use this band-limited record to invert for the atomic events. We choose three seed traces at $x = -0.01$, 0, 0.01, and we identify seven events from the coherent phase picks after applying MUSIC on each trace. Clearly, the algorithm does not identify all the events in the record (as pointed out by the arrows in Figure 15a). Instead, only the strong coherent events have been selected.

The phase and amplitude tracking algorithm resolves seven atomic events (Figure 16), and the reconstructed shot record is shown in Figure 14b. The inverted record has been cleaned up

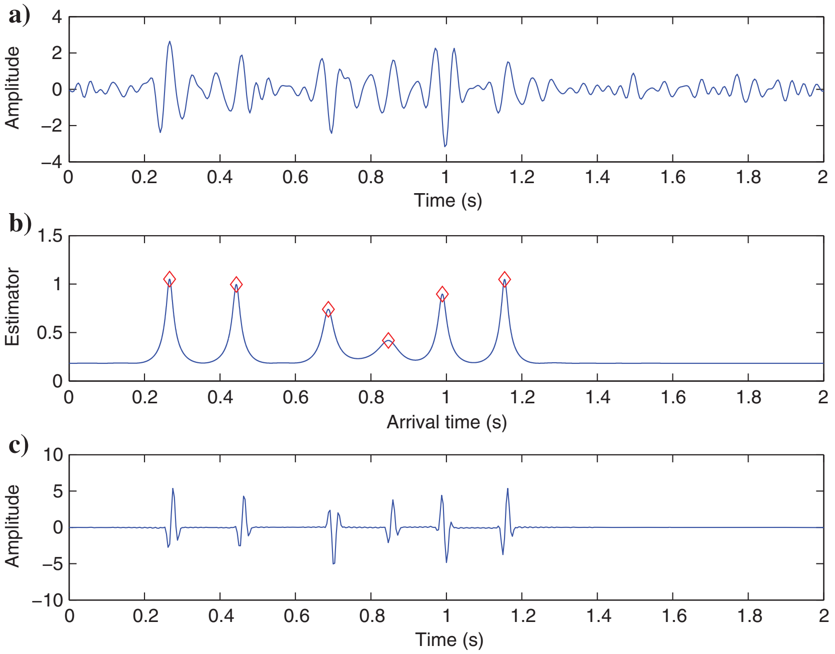


Figure 7. (a) Comparison of a narrow-band ([7, 40] Hz) seismic trace at $x = 1.5$ km, (b) its MUSIC estimator (blue curve) and the picked traveltimes (red diamonds), and (c) the trace after frequency extrapolation to [1, 90] Hz.

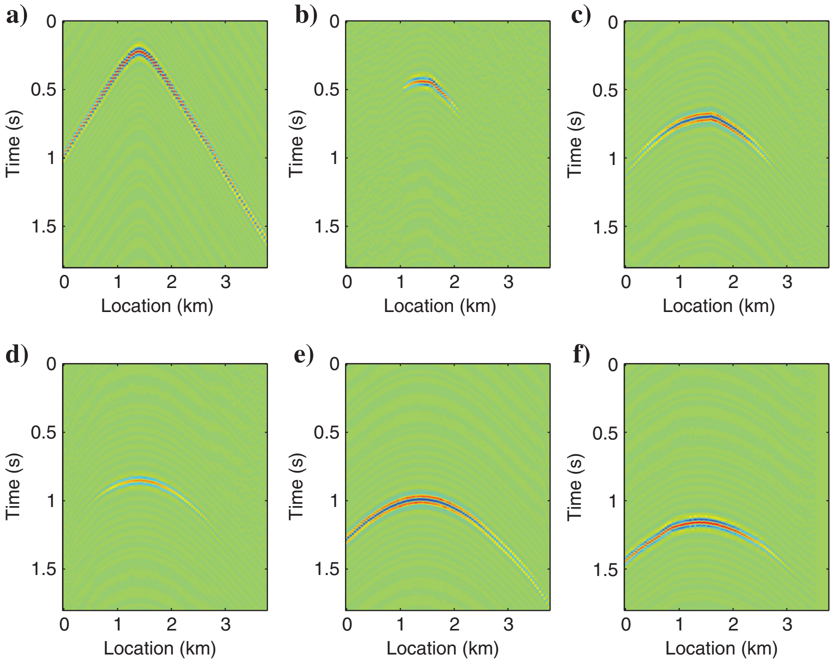


Figure 8. The six well-separated atomic events. Severe noise has been removed from each event. The amplitude and phase are more coherent for each event.

significantly with fewer weak events and overlapping events. With the smoothness constraints, the amplitude and phase of each event across traces are more coherent.

Figure 17 compares the shot record in its full bandwidth with the inverted shot record after frequency extrapolation to [0.5, 120] Hz. Compared with the field record, the arrival times of the resolved events are preserved in the extrapolated record. However, the recorded data appear to contain more events than the extrapolated record. We can explain this observation from two aspects. First, our algorithm indeed does not model all events in the recorded data. Moreover, due to the missing low frequencies, each event has sidelobes in the recorded data, whereas in the extrapolated data, each selected event has a more compact support in time.

Figure 18a shows the shot record after a low-pass filter at 16 Hz. To make a fair comparison, we band-pass the extrapolated shot record (Figure 17b) between 10 and 16 Hz and plot the resulting record in Figure 18b. The field record appears very noisy due to the low S/N, the receiver statics, and the strong amplitude variations across traces. In contrast, the inverted record after extrapolation reserves the phase information with improved amplitude coherence across traces. Figure 15b compares one extrapolated trace with the corresponding recorded trace within the low-frequency bandwidth. After amplitude normalization, the amplitude and phase are matched quite well when most events are successfully identified in the later arrival time. However, the box highlights the discrepancy between the extrapolated and the recorded trace in earlier time due to the

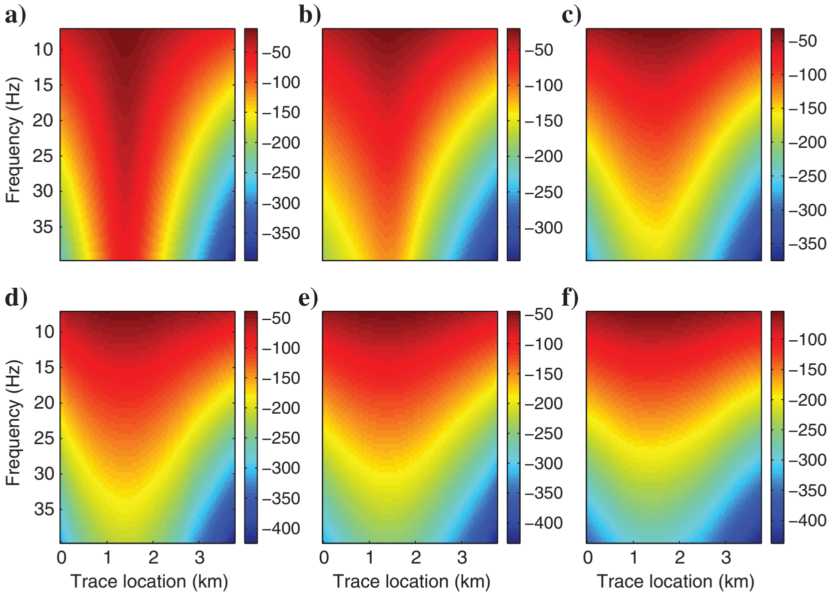


Figure 9. Phase spectrum of each atomic event is smooth along frequency and space axes, as required by the optimization.

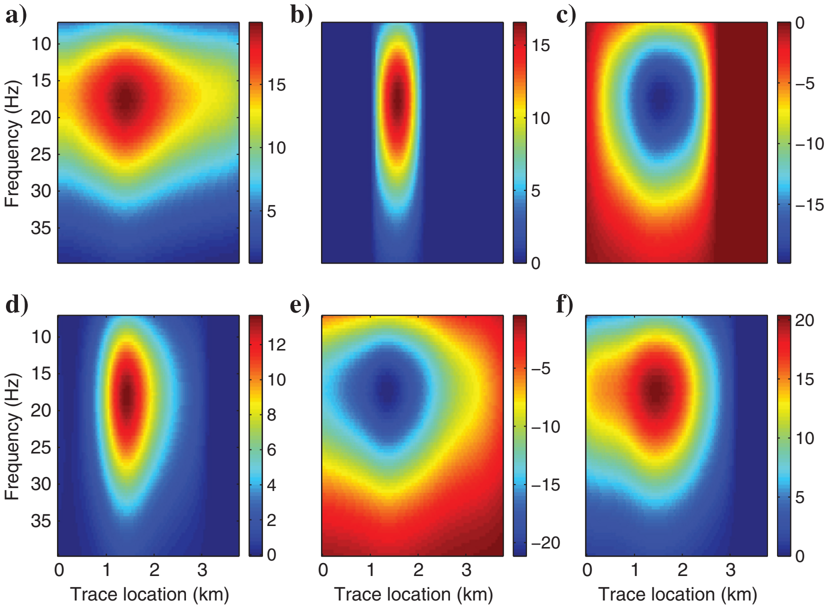


Figure 10. Amplitude spectrum of each atomic event is smooth along frequency and space axes, as required by the optimization.

unmodeled events. The extrapolated data can be further improved with a second pass of phase-and-amplitude tracking on the data residual to pick up the remaining coherent events. Figure 18c shows the inverted shot record after extrapolation to [0.5, 10] Hz. The recorded data within this bandwidth have little meaningful signal, whereas the extrapolated record may provide reliable phase information at these low frequencies.

ASSUMPTIONS AND LIMITATIONS

The fundamental assumption for our tracking algorithm to define seismic events is that the acoustic or elastic wave equations are mostly dispersion free, i.e., that they give rise to solutions with mostly linear phases in the frequency domain. Therefore, in the cases of surface waves and body waves in a severely dispersive medium, our algorithm will have difficulty identifying and tracking the events.

Obviously, frequency extrapolation to the unobserved bandwidth is purely based on the dispersion-free assumption. We have observed phase errors at the extrapolated frequencies due to numerical dispersion in the synthetic test, which naturally indicates potential phase errors due to real dispersion effects in the earth. However, as long as the dispersion is mild, the extrapolated low-frequency data could be a close approximation to the true low-frequency response.

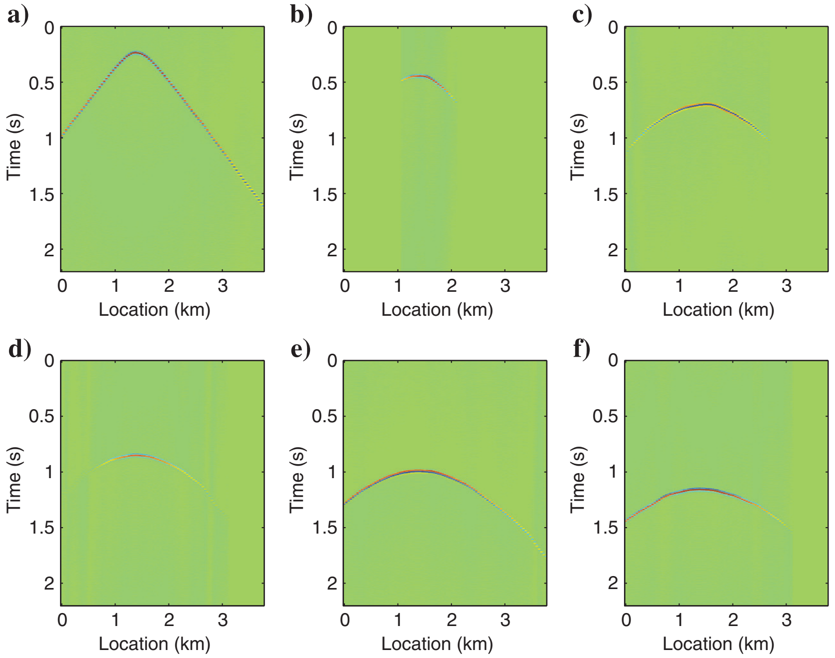


Figure 11. The same six atomic events as in Figure 8 after frequency extrapolation. The waveform is much more compact thanks to the increased bandwidth, whereas the traveltime information is kept intact.

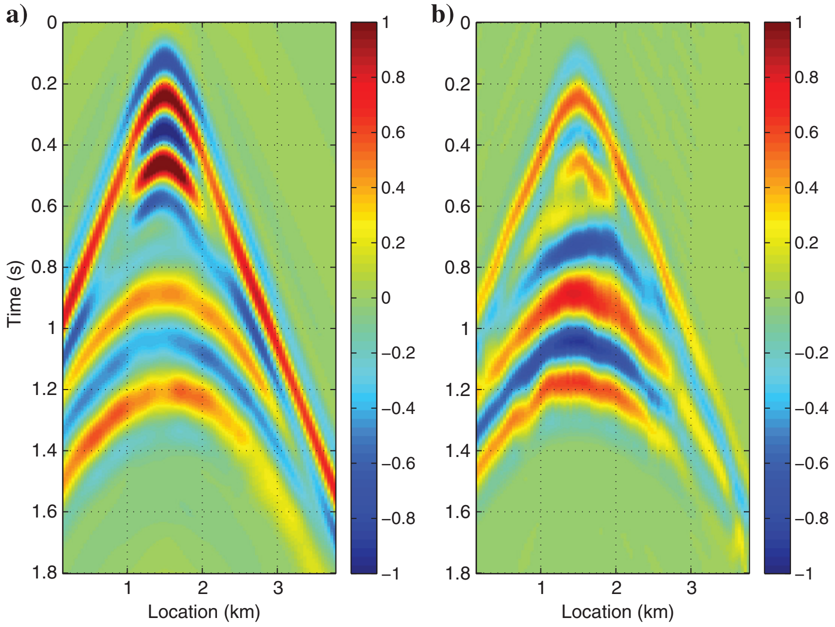


Figure 12. (a) Comparison of the shot record modeled with a low-frequency broadband ([1, 7] Hz) wavelet and (b) the shot record after frequency extrapolation to [1, 7] Hz. Although the amplitudes differ between the two records, the phase function from the frequency extrapolation represents a reasonable estimation of the modeled phase function.

Hence, fitting the extrapolated low-frequency waveforms could bring full-waveform inversion closer to the global minimum.

Extrapolation to low frequencies is also limited by the inability of the phase-amplitude model to capture the physics of zero-frequency radiation due to the generic presence of a scattering pole in Green's function at $\omega = 0$.

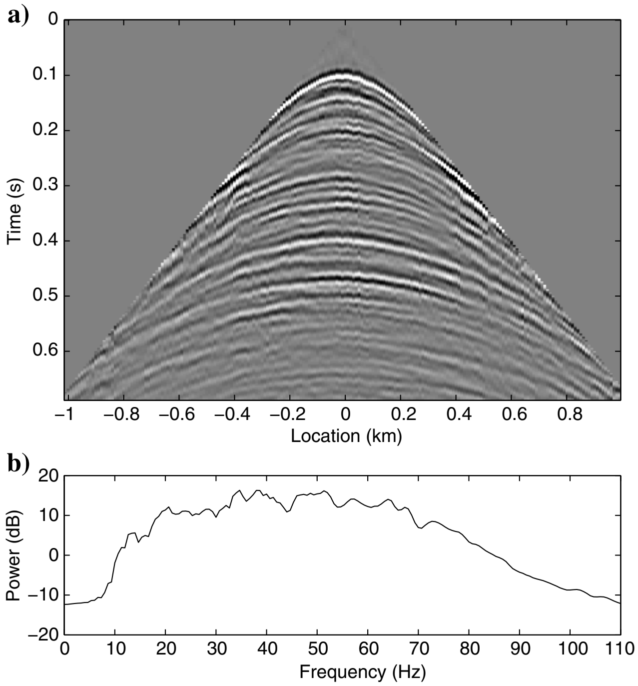


Figure 13. A field data record is shown in panel (a), and its average power spectrum is shown in panel (b). The data show strong receiver statics (shifty phases across traces) and amplitude variations. The data have a large S/N between [16, 85] Hz. The low-frequency components (less than 10 Hz) are missing from the data.

The success of the proposed iterative method for the nonconvex optimization formulation hinges on a delicate interplay between the refinement of phases and amplitudes within a trust region, versus slowly growing this trust region. If the region is grown too fast, the iterations will converge to an undesirable local minimum.

The refinement and expansion depend nonlinearly on the regularization parameters, the determination of which is empirical and can be tedious. We replace the slow converging regularization term, $\lambda\sum_j\|\nabla_\omega^2 b_j(\omega,x)\|_2^2$, by implicitly solving the regularization problem. The rest of the regularization terms are solved explicitly. We let the partial gradient of b_j at the kth iteration be

$$\delta b_j^k = \frac{1}{2}\left(\frac{\partial u}{\partial b_j}\delta\hat{u}^* + \delta\hat{u}\frac{\partial u^*}{\partial b_j}\right) + 2\mu\nabla_x^2 b_j. \qquad (15)$$

We solve the following implicit system to obtain the phase functions for the $(k+1)$th iteration:

$$b_j^{k+1} = b_j^k - \frac{\lambda}{\beta}\nabla^T\nabla b_j^{k+1} - \beta\delta b_j^k. \qquad (16)$$

Hence, the updated phase functions are

$$b_j^{k+1} = \left(1 + \frac{\lambda}{\beta}\nabla^T\nabla\right)^{-1}(b_j^k - \beta\delta b_j^k). \qquad (17)$$

This trick allows us to use large λ values during the inversion. Numerical tests suggest that the inversion results remain stable with respect to the regularization parameters when/where the S/N is high. In generally noisy cases, the solutions can be significantly different with different regularization parameters. The values for each parameter in different numerical tests are shown in Table 1.

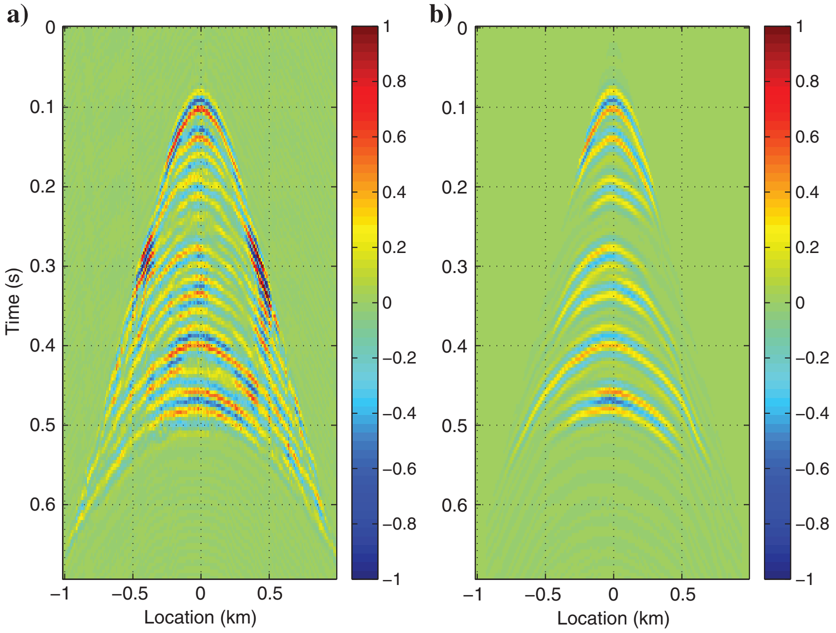


Figure 14. (a) Comparison of the band-limited ([16, 70] Hz) shot record and (b) the inverted shot record. The inverted data have clearly reconstructed the strong reflection events.

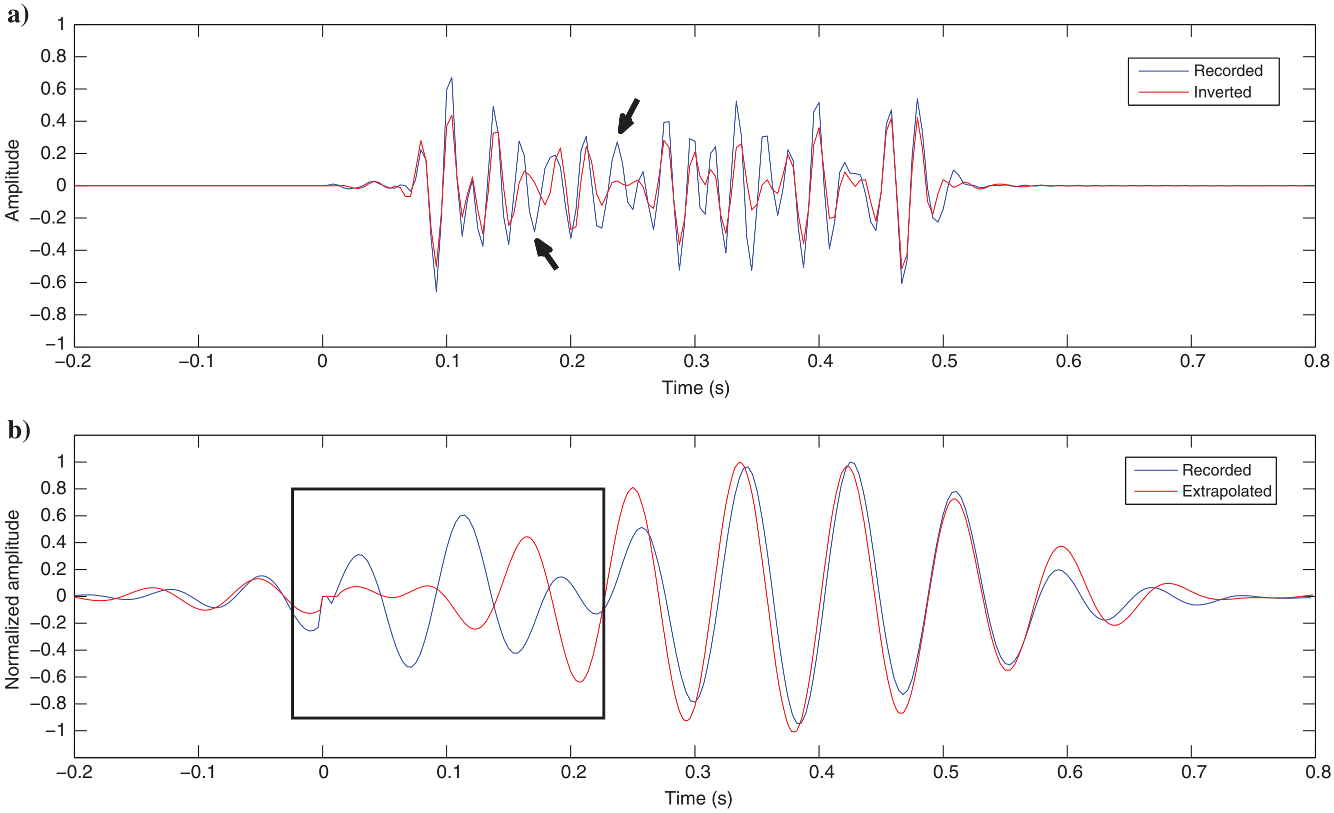
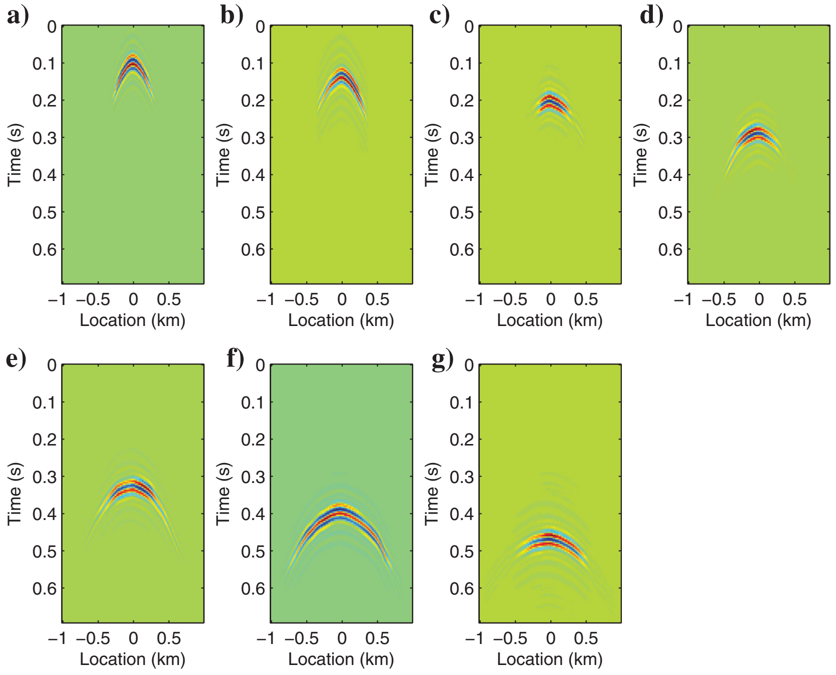


Figure 15. High-frequency ([16, 70] Hz) band-limited seismic trace in (a) and low-frequency (10, 16 Hz) band-limited seismic trace in (b). In the top plot, the inverted trace (red) matches the recorded trace (blue) quite well for the selected events. The arrows point out two events that have not been selected by the inversion at the early arrival times. In the bottom plot, we compare the extrapolated trace (red) with the recorded trace (blue). The (normalized) amplitude and phase are matched quite well when most events are successfully identified in the later arrival time. However, the box highlights the discrepancy between the extrapolated and the recorded traces in earlier time due to the unmodeled events.

Figure 16. Seven well-separated atomic events.



The most important limitation of the tracking method is the deterioration of the accuracy in phase (hence the traveltime) estimation in the presence of unmodeled or splitting events. Our solution for the determination of the number of events further depends on the availability of a few seed traces where the MUSIC estimator correctly categorizes them. We could potentially analyze the variability of the solutions with different regularization parameters to determine the presence of unmodeled or splitting events, which requires further investigation.

Table 1. Regularization parameters used in each test.

Case	λ	μ	γ
Three-event separation	1.0	1.0	$2e-3$
Shallow Marmousi	1.0	$1e+1$	$2e-3$
Field data	1.0	$4e+2$	$3e-3$

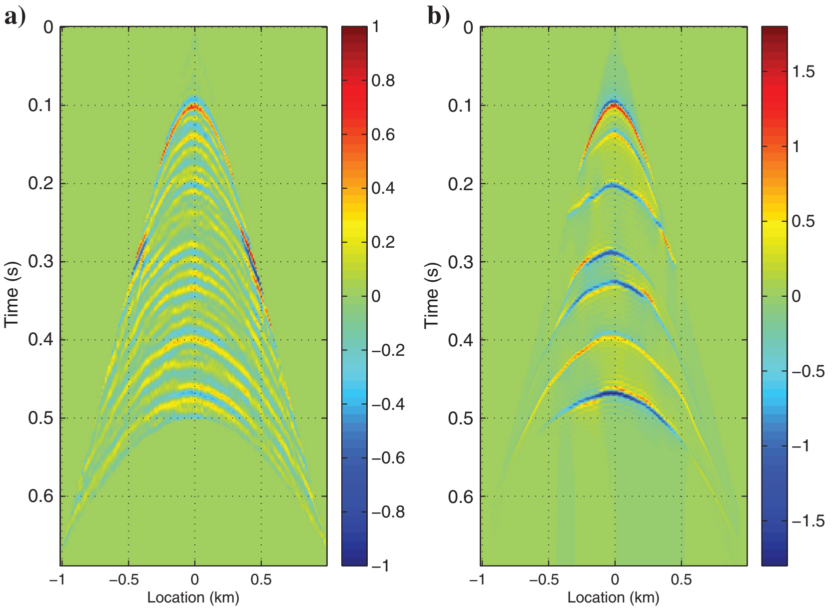


Figure 17. Comparison of the full bandwidth shot record in panel (a) and the inverted shot record after frequency extrapolation to [0.5, 120] Hz in panel (b). Kinematic information of the selected events is well-preserved by extrapolation. The recorded data appear to contain more events than the extrapolated data. The reason can be twofold: First, our tracking algorithm indeed does not model all the events in the recorded data. Moreover, due to the missing low frequencies, each event has sidelobes in the recorded data, whereas in the extrapolated data, each selected event has a more compact support.

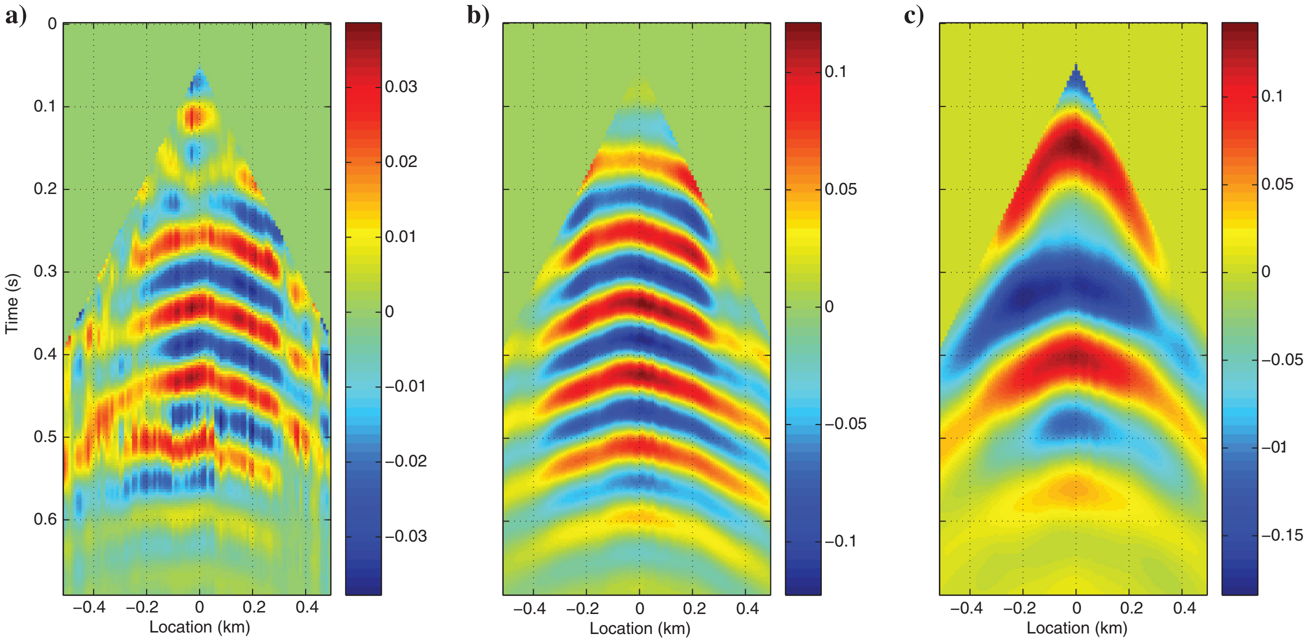


Figure 18. (a) Comparison of the shot record after a low-pass filter at 16 Hz, (b) the inverted record after frequency extrapolation to [10, 16] Hz, and (c) the inverted record after frequency extrapolation to [0.5, 10] Hz. The field record appears to be very noisy due to low S/N, the receiver statics, and the strong amplitude variations across traces. In contrast, the extrapolated record within the same bandwidth reserves the phase information with improved amplitude coherence across traces. The recorded data less than 10 Hz have little meaningful signal, whereas the extrapolated record may provide reliable phase information at these low frequencies.

CONCLUSIONS

We propose a data-driven method for decomposing seismic records into individual, atomic events that correspond to isolated arrivals. The only piece of physical information needed to define these events is the fact that the acoustic or elastic wave equations are mostly dispersion free. The explicit handle on each atomic event enables us to approximate the high- and low-frequency seismic response beyond the observed frequency bandwidth. The numerical examples demonstrate the robustness of the method and the potential of using the extrapolated low-frequency data for the initialization of full-waveform inversion.

ACKNOWLEDGMENTS

This project was funded by Total S.A. L Demanet is also grateful to Air Force Office of Scientific Research, Office of Naval Research, and National Science Foundation for funding.

REFERENCES

Biondi, B., and C. Kostov, 1989, High-resolution velocity spectra using eigenstructure methods: Geophysics, **54**, 832–842, doi: 10.1190/1.1442712.

Blias, E., 2007, VSP wavefield separation: Wave-by-wave optimization approach: Geophysics, **72**, no. 4, T47–T55, doi: 10.1190/1.2744124.

Cansi, Y., 1995, An automatic seismic event processing for detection and location: The P.M.C.C. method: Geophysical Research Letters, **22**, 1021–1024, doi: 10.1029/95GL00468.

Demanet, L., and P. Hand, 2014, Stable optimizationless recovery from phaseless linear measurements: Journal of Fourier Analysis and Applications, **20**, 199–221, doi: 10.1007/s00041-013-9305-2.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960, doi: 10.1190/1.1527095.

Fomel, S., E. Landa, and M. T. Taner, 2007, Poststack velocity analysis by separation and imaging of seismic diffractions: Geophysics, **72**, no. 6, U89–U94, doi: 10.1190/1.2781533.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame: Geophysics, **75**, no. 3, V25–V38, doi: 10.1190/1.3380591.

Fu, H., L. Gan, and R. Clapp, 2014, Scaling reverse time migration performance through reconfigurable data flow engines: IEEE Micro, **34**, 30–40, doi: 10.1109/MM.2013.111.

Gentili, S., and A. Michelini, 2006, Automatic picking of P and S phases using a neural tree: Journal of Seismology, **10**, 39–63, doi: 10.1007/s10950-006-2296-6.

Gholami, A., 2014, Phase retrieval through regularization for seismic problems: Geophysics, **79**, no. 5, V153–V164, doi: 10.1190/geo2013-0318.1.

Hua, Y., and T. Sarkar, 1990, Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise: IEEE Transactions on Acoustics, Speech, and Signal Processing, **38**, 814–824, doi: 10.1109/29.56027.

Kirlin, L., and W. Done, 1999, Covariance analysis for seismic signal processing: SEG, Geophysical Developments Series 8.

Leonard, M., and B. L. N. Kennett, 1999, Multi-component autoregressive techniques for the analysis of seismograms: Physics of the Earth and Planetary Interiors, **113**, 247–264, doi: 10.1016/S0031-9201(99)00054-0.

Raoult, J. J., 1983, Separation of a finite number of dipping events: Theory and applications: 53rd Annual International Meeting, SEG, Expanded Abstracts, 277–279, doi: 10.1190/1.1893839.

Schmidt, R., 1986, Multiple emitter location and signal parameter estimation: IEEE Transactions on Antennas and Propagation, **34**, 276–280, doi: 10.1109/TAP.1986.1143830.

Wang, L., J. Gao, W. Zhao, and X. Jiang, 2013, Enhancing resolution of nonstationary seismic data by molecular-Gabor transform: Geophysics, **78**, no. 1, V31–V41, doi: 10.1190/geo2011-0450.1.

Zhang, H., C. Thurber, and C. Rowe, 2003, Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings: Bulletin of the Seismological Society of America, **93**, 1904–1912, doi: 10.1785/0120020241.