GEOPHYSICS, VOL. 78, NO. 4 (JULY-AUGUST 2013); P. V157–V169, 13 FIGS., 1 TABLE.
10.1190/GEO2012-0209.1

Viterbi sparse spike detection

Samuel P. Brown[1] and Michael S. Thorne[2]

ABSTRACT

Accurate interpretation of seismic traveltimes and amplitudes in the exploration and global scales is complicated by the band-limited nature of seismic data. We discovered a stochastic method to reduce a seismic waveform into a most probable constituent spike train. Model waveforms were constructed from a set of candidate spike trains convolved with a source wavelet estimate. For each model waveform, a profile hidden Markov model (HMM) was constructed to represent the waveform as a stochastic generative model with a linear topology corresponding to a sequence of samples. Each match state in the HMM represented a sample in the model waveform, in which the amplitude was represented by a Gaussian distribution. Insert and delete states allowed the underlying source wavelet to dilate or contract, accounting for nonstationarity in the seismic data and errors in the source wavelet estimate. The Gaussian distribution characterizing each sample's amplitude accounted for random noise. The Viterbi algorithm was employed to simultaneously find the optimal nonlinear alignment between a model waveform and the seismic data and to assign a score to each candidate spike train. The most probable traveltimes and amplitudes were inferred from the alignments of the highest scoring models. The method required no implicit assumptions regarding the distribution of traveltimes and amplitudes; however, in practice, the solution set may be limited to mitigate the nonuniqueness of solutions and to reduce the computational effort. Our analyses found that the method can resolve closely spaced arrivals below traditional resolution limits and that traveltime estimates are robust in the presence of random noise and source wavelet errors. The method was particularly well suited to fine-scale interpretation problems such as thin bed interpretation, tying seismic images to well logs, and the analysis of anomalous waveforms in global seismology.

INTRODUCTION

A problem of fundamental importance in all branches of seismology is accurate traveltime and amplitude picking. However, the problem is generally complicated by a low signal-to-noise ratio (S/N) or constructive/destructive interference effects from band-limited wavelets. Figure 1 provides a simple illustration of this problem. The top three traces show individual arrivals consisting of 25-Hz Ricker wavelets scaled in amplitude and shifted in time. The lowermost trace shows the composite waveform obtained by summing the individual arrivals. Because these are zero phase wavelets, traveltimes and amplitudes can be estimated from picks at the locations of local amplitude maxima or minima at peaks or troughs. The black spikes in Figure 1 depict manually picked traveltimes and amplitudes of the individual arrivals comprising the composite waveform, whereas the gray spikes depict the actual traveltimes and amplitudes. The traveltime is correctly picked for the first arrival, but destructive interference from the second arrival has led to an erroneously low amplitude pick. The second arrival pick has significant time and amplitude errors, and the third arrival pick has slight traveltime and amplitude errors. This simple noise-free example demonstrates how constructive and destructive interference from overlapping wavelets complicates the picking process and can lead to erroneous traveltime and amplitude picks.

Accurate picking requires consideration of constructive and destructive interference from neighboring arrivals and noise. As the separation between individual arrivals decreases, the problem shifts from being one of picking accurate traveltimes and amplitudes to resolving the two separate arrivals. Figure 2 demonstrates the resolution criteria proposed by Rayleigh and Ricker (see Ricker,

Manuscript received by the Editor 13 June 2012; revised manuscript received 27 March 2013; published online 24 June 2013.
[1]Petroleum Geo-Services, Houston, Texas, USA and University of Utah, Department of Geology and Geophysics, Salt Lake City, Utah, USA. E-mail: samuel.brown@pgs.com.
[2]University of Utah, Department of Geology and Geophysics, Salt Lake City, Utah, USA. E-mail: michael.thorne@utah.edu.

1953). Figure 2a defines two metrics for a Ricker wavelet, which determine both resolution criteria. The wavelet breadth b is given as the distance between the maximum absolute amplitudes of the symmetric side lobes, while the temporal resolution limit (TR) (Kallweit and Wood, 1982) is the distance between inflection points on the central lobe. With sufficient separation, individual main-lobe peaks are visible and separate arrivals are easily resolved, as shown in Figure 2b. For two uniform-polarity events of equal amplitude, Rayleigh's limit of resolution is reached when two wavelets are separated by $b/2$, or the peak-to-trough time. As Figure 2c illustrates, this is the minimum separation at which two distinct peaks are still visible in the composite waveform. The resolution limit was extended by Ricker to the smaller distance TR. At this separation, the composite wavelet no longer has two distinct peaks; however, the curvature at the central maximum is zero, resulting in a flat spot, as can be seen in Figure 2d.

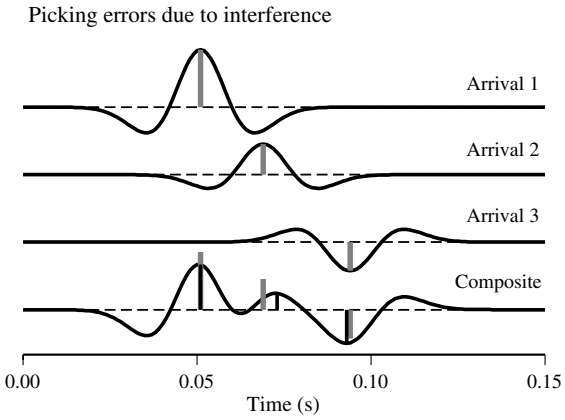

Figure 1. The top three traces show individual 25-Hz Ricker wavelets. The lowermost trace shows a composite waveform consisting of the sum of the individual Ricker wavelets from the upper traces. The gray spikes represent actual arrival times and amplitudes, while the black spikes represent manually picked arrival times and amplitudes.

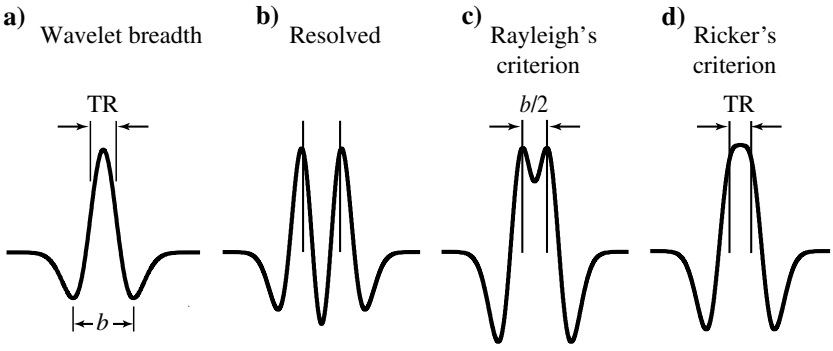

Figure 2. Resolution criteria with decreasing wavelet separation for two uniform-polarity Ricker wavelets. (a) Definition of wavelet breadth b and TR (Kallweit and Wood, 1982) for a Ricker wavelet. The wavelet breadth is the distance between side-lobe troughs, while TR is the distance between inflection points on the central lobe. The separation between arrivals is sufficient that the events in (b) are clearly resolved. (c) Rayleigh's criteria $b/2$ represents the limit of resolvability in which two distinct peaks remain. (d) A zero curvature-induced flat spot allows for event resolution at Ricker's limit TR.

Resolution limits, as well as the effects of constructive and destructive interference on waveform character, can be visualized with a wedge model, as shown in Figure 3. The underlying earth model contains three layers, with impedance varying discretely with depth, such that the absolute value of the zero-offset reflection coefficients at each interface is equal. The thickness of the middle layer decreases from left to right, forming the wedge. The reflectivity caused by impedance contrasts is represented as a function of two-way traveltime by the spikes in Figure 3. Synthetic seismic traces representing an idealized zero-offset acquisition were created by convolving a 25-Hz Ricker wavelet with the reflectivity series. Figure 3a shows a case in which the reflection coefficient is positive at both interfaces, representing an increase in impedance with depth. Figure 3b presents a mixed-polarity wedge model. In this case, a high-impedance wedge is inserted into a whole space, such that the polarity of the top and bottom reflection coefficients is reversed. This type of model would be characteristic of either a low-impedance bed surrounded by identical high-impedance beds, or vice versa, and it is of particular value because this phenomenon is likely to occur in a shale-sand-shale sequence.

In Figure 3a, the transition from well-resolved peaks through Rayleigh's ($b/2$) and Ricker's (TR) resolution limits can clearly be seen from the thickest to thinnest regions of the wedge (from left to right in Figure 3). However, in the mixed-polarity case (Figure 3b) Ricker's (TR) limit has no meaning. Widess (1973) proposes an amplitude-based methodology for resolving thin beds in the mixed-polarity case. Widess notes that at an approximate distance of $\lambda_b/8$, where λ_b is the dominant wavelength, the composite waveform stabilizes into a good approximation of the derivative of the original zero-phase wavelet. Thus, the waveform shape is virtually indistinguishable for beds less than λ_b, but the bed thickness may be inferred from the peak amplitude of the composite wavelet. At thicknesses below the established resolution criteria, the composite waveform for the uniform-polarity and opposite-polarity cases can be visually interpreted as a single arrival; however, the composite wavelet in the opposite-polarity case will approximate the derivative of the original zero-phase pulse with a strong amplitude response dependent on bed thickness (e.g., Figure 3a and 3b). This amplitude information is routinely used by seismic interpreters to construct tuning curves (Bacon et al., 2003) for thin-bed interpretation. In practice, this requires calibration of amplitudes to a known bed thickness. This method is further complicated by the presence of noise and uncertainty in the source wavelet estimate (Widess, 1973).

Figure 4a shows apparent thickness in two-way traveltime and apparent amplitude as a function of true thickness for the synthetic traces pictured in Figure 3a. The solid black line at 45° shows the ideal case where the apparent thickness determined matches the true thickness of the layer. The blue line shows the actual recovered thickness. If one could accurately determine thickness in all cases, the solid blue line should perfectly track the solid black line. However, the apparent thickness increases slightly before Rayleigh's criterion ($b/2$) is reached because of interference between the wavelets. Below the TR limit, the separate arrivals are unresolved

and the composite arrival appears to be an individual arrival. The solid horizontal black line shows the true amplitude of both arrivals, while the red line shows the measured amplitude. For traveltimes at or above Rayleigh's limit, the amplitudes are measured from two separate peaks. Below Rayleigh's limit, only a single peak remains. Destructive interference lowers amplitudes near the resolution criteria, and constructive interference increases amplitudes as the differential traveltime goes to zero.

In Figure 4b, the apparent thickness and amplitude as a function of true thickness for the mixed-polarity wedge model (Figure 3b) are shown. Here, the composite waveforms are comprised of two equal-amplitude, opposite-polarity 25-Hz Ricker wavelets. In this case, it is possible to resolve the two arrivals for all thicknesses; however, the apparent separation is limited to TR. Constructive interference increases amplitudes near the resolution limits, while destructive interference diminishes amplitudes as the differential traveltime goes to zero.

The process of detecting traveltimes and amplitudes of each arrival in a band-limited seismic trace is a form of sparse spike deconvolution. In recent decades, numerous techniques have been developed in this area. Kormylo and Mendel (1983) introduce a maximum-likelihood sparse spike deconvolution algorithm based on state-variable technology, which is capable of estimating a Bernoulli-Gaussian sparse spike train and the correct phase of the seismic wavelet. Wiggins (1978) and Sacchi et al. (1994) propose minimum entropy sparse spike deconvolution methods. Velis (2008) proposes a stochastic sparse spike deconvolution method that incorporates impedance constraints into simulated annealing iterations to find a solution consisting of the least number of spikes, which explains the observed data when convolved with a model wavelet. Phase errors in the wavelet estimate are handled to some extent by computing an optimal phase shift to match the model wavelet to an effective data wavelet. Kaaresen and Taxt (1998) use a Bayesian framework to derive a maximum a posteriori multichannel deconvolution estimate. Their method alternates between steps of wavelet and reflectivity estimation, and it handles continuity between traces by modeling local dependencies. Heimer et al. (2007) use dynamic programming to constrain multichannel blind seismic deconvolution such that reflections must form continuous paths across consecutive traces, representing consistent layers in the earth model. This method is improved upon by Heimer et al. (2009) by incorporating the Viterbi algorithm (Viterbi, 1967) for Markov-Bernoulli random field modeling in place of the previous dynamic programming algorithm, to allow layers to split, merge, and terminate across traces. Dynamic programming methods are also used by Liner and Clapp (2004) to nonlinearly align seismic traces, and the Viterbi algorithm is used by Clapp (2008) to autopick seismic horizons.

In this paper, we propose a stochastic method, using profile hidden Markov models (HMMs) and the Viterbi algorithm (Durbin et al., 1998), to resolve composite waveforms into their constituent spike trains. This effectively poses the sparse spike deconvolution problem as a pattern recognition problem. This method is adapted from the application of the Viterbi algorithm to profile HMMs by Eddy (1995) to construct alignments between, and aid in the identification of, evolutionarily related protein sequences. The Viterbi algorithm is used in the single-channel deconvolution context of picking individual arrival traveltimes and amplitudes. Unlike the other deconvolution algorithms outlined above, the method presented here makes no implicit assumptions about the distributions of traveltimes and amplitudes nor the phase of the wavelet. However, in practice, the method is generally employed within a constrained solution space to mitigate the problem of nonuniqueness and to bound the computation time. While the algorithm does not provide updated estimates of the wavelet, it is relatively insensitive to the types of errors expected in the practice of source wavelet estimation, and it can be applied in cases in which data are insufficient to generate meaningful statistics, such as a single seismic trace. Unlike tuning curves, this method implicitly accounts for noise and errors in the wavelet estimate, and its accuracy degrades gracefully as noise levels and wavelet errors increase.

THEORY

Our goal is to analyze a composite waveform in a data trace specifically to determine the differential traveltimes and relative amplitudes of the individual arrivals. Our ultimate goal is to deconstruct the composite waveform into a spike train, where each spike is aligned on the true arrival time with the correct amplitude of the individual seismic arrivals that went into constructing the composite waveform. To do this, we first define a solution space that consists of a set of unique spike trains, where each spike train is an element of the set containing all possible spike trains in the solution space.

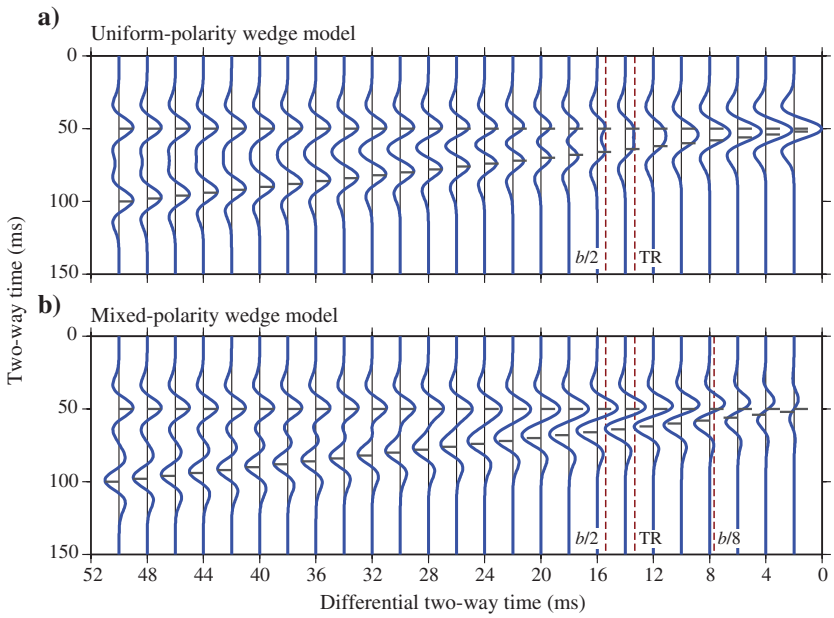

Figure 3. Seismic wedge models that demonstrate the effects of constructive and destructive interference in band-limited seismic traces as the temporal separation of individual arrivals decrease. (a) A seismic wedge model with uniform-polarity reflection coefficients, representing three layers in which acoustic impedance increases with depth. (b) A seismic wedge model with opposite-polarity reflection coefficients, representing a high-impedance layer between two low-impedance layers.

That is, each element consists of a unique collection of differential traveltimes and relative amplitudes that could describe the composite trace. From here on, we refer to one element in our set of spike trains as a model. The goal is to determine which model best describes the composite trace. One approach to solving this problem is to (1) build a reference trace for each point in the solution space by convolving it with a source wavelet, (2) find the best alignment between each reference trace and the data trace using crosscorrelation, and (3) pick the optimal solution through minimizing a residual error norm on the difference between the data trace and the reference traces.

The above algorithm may be a useful method to solve this problem, but it is not without challenges. For example, if an L2-norm is used in the optimization step, the algorithm will not be strongly affected by the presence of Gaussian-distributed random noise, but it does not implicitly handle errors in the wavelet estimate. The approach we propose here is similar to the above algorithm, except that it uses stochastic models to represent the points in our solution space. The proposed algorithm is: (1) Build a profile HMM for each model in which a noise model and a wavelet estimate are assumed at the onset. (2) For each profile HMM, use the Viterbi algorithm to simultaneously find the optimal alignment to the composite waveform and assign a log-likelihood score. Here, the alignment may be nonlinear. (3) Extract traveltimes and amplitudes from the alignment of the highest-scoring HMM.

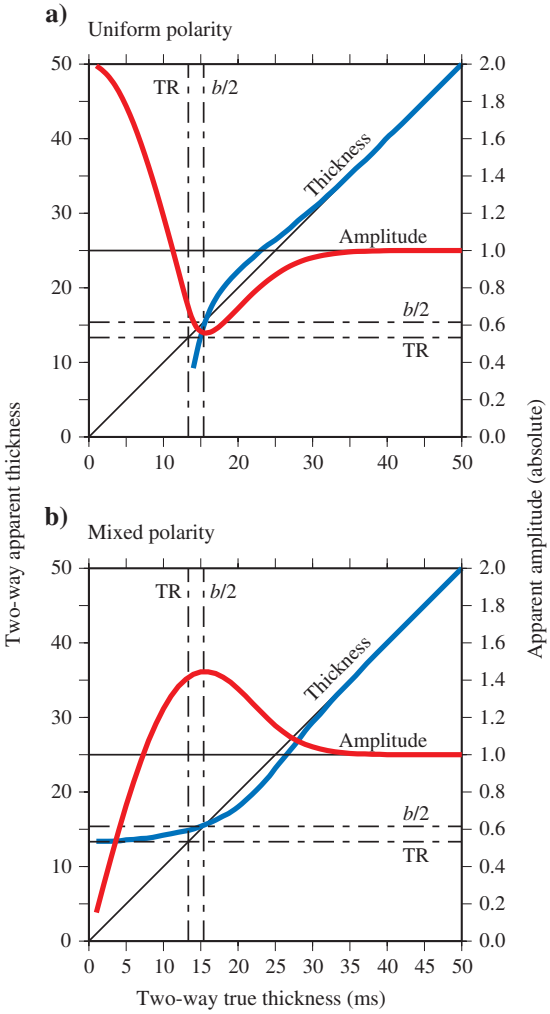

Figure 4. (a) Apparent thickness and amplitude as a function of event separation for two equal-amplitude uniform-polarity 25-Hz Ricker wavelets. The two arrivals are not resolved below the TR. (b) Apparent thickness and amplitude as a function of event separation for two equal-amplitude opposite-polarity 25-Hz Ricker wavelets. In this case, it is possible to detect both arrivals for all separations; however, the apparent thickness is limited to TR.

The first step is to construct a profile HMM for each model, which will be based on a synthetic trace obtained by convolving the model spike train with the assumed wavelet estimate. A profile HMM can be pictured as a finite state automaton such as the one pictured in Figure 5. This is a simplified version of the Plan 7 HMM architecture that Eddy (1995) introduces for analyzing protein sequences. Each of the nodes (diamonds and circles in Figure 5) represents a state, and each of the edges (arrows in Figure 5) represents a possible state transition. Markov models represent a stochastic process for generating sequences, which in this case, will be a sequence of samples constituting a seismic trace. In a Markov model, each state emits a single value or token and the stochastic process is represented by the random set of transitions between states. The model is characterized by its states and by transition probabilities between the states. The underlying stochastic process exhibits the Markov property, which states that the conditional probability distribution of future states depends only on the current state, resulting in a memoryless stochastic process. In an HMM, each state may generate many different tokens based on a state-specific emission probability distribution. The name is derived from the fact that the actual state sequence that produced a given output is hidden and can only be described in terms of probabilities, because many different state sequences can produce the same output. The Viterbi algorithm takes advantage of the Markov property, and it employs dynamic programming to efficiently determine the most likely state sequence for generating a specific output sequence with a specific HMM.

Our goal in using the HMM is to find the best nonlinear alignment between the composite waveform (our data) and a model in our solution space. An alignment can be thought of as a unique path traced through the state machine. The B and E states are for starting and ending alignments; M1, M2, M3, M4, and M5 are match states, and each one is associated with a sample from the model trace. Each match state is characterized by a Gaussian distribution with the amplitude of the associated model trace sample defining the mean. This Gaussian distribution over sample amplitudes represents the implicit random noise model. I1, I2, I3, and I4 are insert states, which allow multiple samples from the data trace to be aligned to a single sample from the model trace. This, in effect, allows the embedded wavelet to dilate. Insert states are also characterized by a Gaussian distribution whose mean is the average of the two surrounding match state means. D2, D3, and D4 are delete states, which allow model samples to be skipped, effectively allowing the embedded wavelet to contract. The dilation and contraction facilitated by the insert and delete states introduce a nonlinearity in alignment between the model and data that is not present in deterministic convolutional models.

To demonstrate the alignment procedure, we refer to Figure 6. Figure 6a shows a composite waveform derived from a candidate model from our set of all possible models that may explain the data trace shown in Figure 6b. One possible way to align the model and data trace is shown in Figure 6c. Consulting Figure 5, we start at the begin state (B) and proceed through the state machine as follows:

For this HMM, the first state is always a match state (M1), so we match the first sample from the model trace to the first sample from the data trace. From the M1 state, transitions exist to the I1, M2, and D2 states. In this example, the transition goes to the M2 state, so we match the second sample from the model trace to the second sample of the data trace. In a similar manner, we next transition through the M3, M4, and M5 states, matching the corresponding samples from the model trace to the data trace. From the M5 state, we can only transition to the end state (E) and we are finished. The end result shown in Figure 6c just matches each sample from the model trace to the data trace.

Another possible alignment is shown in Figure 6d. We start at the begin state, and as before, we transition to the M1 state matching the first sample from the data trace to the first sample of the model trace. Then we transition to the M2 state matching the second sample from the data trace to the second sample of the model trace. From the M2 state, we then transition to the D3 state, which means that the third sample of the model trace is discarded. The M4 state follows, which matches the third data sample to the fourth model sample. The fourth data sample is also aligned to the fourth model sample through the I4 state. Finally, the M5 state is traversed, matching the fifth data and model samples, and the end state (E) once again concludes the alignment.

The alignment in Figure 6e is similar to the one in Figure 6d. In this case, the D2 state is traversed rather than the D3 state. From visual inspection, the alignment of Figure 6d does a better job matching amplitudes than the alignment in Figure 6e. When comparing the alignments from Figure 6c and 6d, it is clear that the amplitudes are more closely matched for the four match states traversed in Figure 6d than they are in Figure 6c. For this reason, the alignment of Figure 6d may be considered the best qualitatively. The quantitative selection of the best alignment, however, will depend on the HMM parameterization as described below.

The next step in our procedure is to use the Viterbi algorithm to find the optimal alignment between a composite waveform in a data trace and a candidate model, which can be described by a sequence of states traced through the HMM. The Viterbi algorithm uses dynamic programming to determine the solution in $\mathcal{O}(N \times M)$ operations, where M is the number of states in the HMM (for example, in Figure 5, this HMM has 14 states) and N is the number of samples in a data trace.

The optimal alignment is defined as the alignment that maximizes the following probability:

$$P^{\pi}(\text{data}|\text{model}) = \Pi_i a^{\pi}_{kl} g^{\pi}_{l}(d_i), \qquad (1)$$

where i is a sample index for the data trace, d_i is a sample of the data trace, a_{kl} is a transition probability associated with transitioning from state k to l, and g_l is a state-dependent probability distribution that determines how likely it is to observe the value of d_i in a given state. The state integer π represents a specific state sequence that determines the values of k and l. The optimal state sequence π^* maximizes the probability of fit between the data and the model and determines the optimal alignment.

The matrix a_{kl} determines the probability of transitioning from any given state k to any other state l. For a profile HMM, most values of this matrix are zero, and the nonzero components correspond to the edges in the state machine graph. Typically, transitions to match states are assigned higher probabilities than transitions to insert and delete states, which have the effect of penalizing wavelet

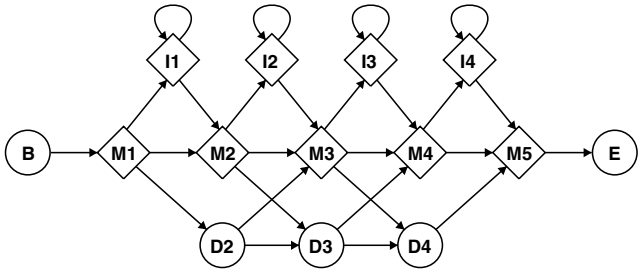

Figure 5. The profile HMM used for waveform analysis pictured as a finite state automaton.

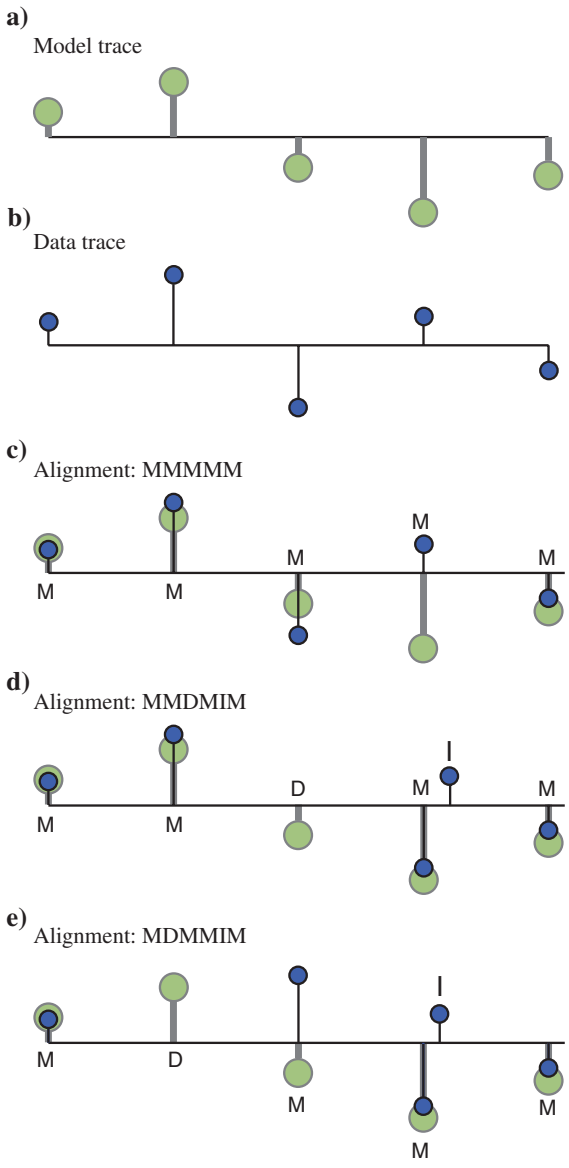

Figure 6. (a) A model trace used to build an HMM represented as a stem plot. (b) A data trace that is to be aligned to the model trace. (c, d, e) Three examples of alignments between the data and model traces. In each panel, the green and blue stems correspond to samples from the model and data trace, respectively.

dilation and contraction. Formal training algorithms exist that can optimize the values of a_{kl} given sufficient training data; however, the existence of adequate training data and overspecialization of the HMM to the training data are problematic. We have found empirically that match state transition probabilities of approximately 0.5 and insert/delete transition probabilities of approximately 0.25 work well in most cases. These probabilities can be fine-tuned if more or less contraction/dilation is desired. There is an interplay between transition and emission probabilities, but having match transitions approximately twice as likely as insert/delete transitions will provide for one or two samples of dilation or contraction where needed, without distorting the underlying wavelet beyond recognition. Each match and insert state has associated with it an emission probability distribution g_l. The mean of each state-dependent emission distribution is set to the amplitude of the associated model trace sample. Representing model trace sample amplitudes with Gaussian distributions accounts for random noise and errors in the wavelet estimate. The standard deviation of the amplitude distributions is set as a run-time parameter. The standard deviations should be inversely proportional to the S/N of the data trace because amplitude differences are penalized more harshly as the standard deviation shrinks.

For computational efficiency, the model and data traces are converted to an integer representation, and discrete probability mass functions are used in the HMM to limit the solution space. Also, computation of equation 1 is subject to numeric underflow because it involves the multiplication of many small probabilities. For this reason, the computations are carried out in a logarithmic space. More specifically, log-odds scores are calculated as a proxy for the probability in equation 1, and are defined as

$$S(\text{data}|\text{model}) = \log\left(\frac{P(\text{data}|\text{model})}{P(\text{data}|W)}\right), \qquad (2)$$

where W is a white-noise model. In this logarithmic computation space, the optimal alignment represents the path through the

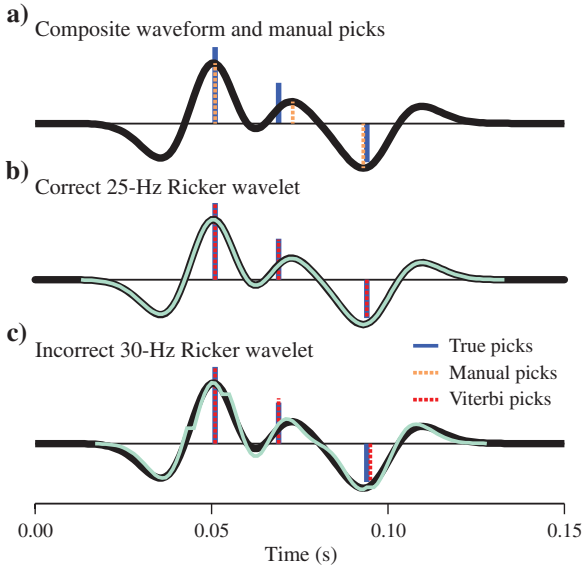

Figure 7. In each panel, the black trace is an example composite waveform (Figure 1) used to demonstrate the VSSD technique. (a) Manual picks for the three arrivals are shown as dashed orange lines, whereas true arrival times and amplitudes are represented by blue lines. (b) Viterbi picks (red lines) using the correct 25-Hz Ricker wavelet. (c) Viterbi picks using a 30-Hz Ricker wavelet. (b and c) Green waveforms represent the means of the emission distributions for all match and insert states traversed in the optimal alignments.

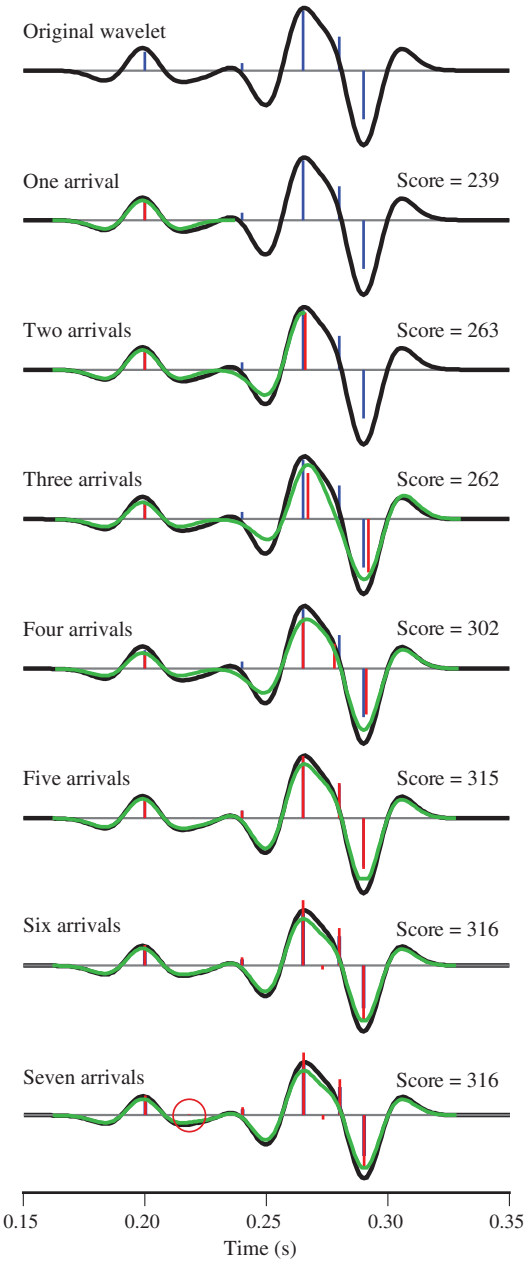

Figure 8. The behavior of the VSSD algorithm for solution spaces characterized by different numbers of arrivals is demonstrated by matching a composite waveform with five arrivals with models comprising one to seven arrivals. The black trace is the composite waveform consisting of five arrivals whose traveltimes and amplitudes are indicated by the blue spikes. Each of the lower panels displays the VSSD results and their log-odds scores for an increasing number of arrivals. In each panel, the red spike indicates the best fit arrival(s) and the green trace represents the modeled trace. Note that the relative amplitudes of the sixth and seventh (red circle) arrivals are very close to zero and that their inclusion does not significantly improve the waveform fit.

HMM that generates the data trace with the highest possible log-odds score. The log-odds score is a ratio of the likelihood of a given model producing a given data trace over the likelihood of a random model producing the data trace.

Referring back to Figure 6, the alignment in Figure 6c may be optimal if the HMM parameterization penalizes dilation and contraction (insert and delete states) more than amplitude errors (amplitude distributions). Conversely, the alignment in Figure 6d may be optimal if amplitude errors are penalized more than dilation and contraction.

Once the optimal log-odds score has been calculated for each model, the model with the highest score is chosen. This brings us to the final step in our procedure, which is to determine the amplitudes and differential traveltimes of our arrivals. The amplitudes are immediately available from the model parameters, and the arrival times of each event can be calculated by analyzing the optimal state path to determine the relative position of match states corresponding to each model spike.

The nonlinear alignment provided by the Viterbi algorithm is the key to this method's ability to work in the presence of noise and with imperfect knowledge of the source wavelet. Through the insert and delete states, the constituent wavelets are allowed to dilate and contract, which compensates for errors in the wavelet phase and amplitude spectra in tandem with random noise when combined with the Gaussian noise model. This ability is crucial to waveform analysis because wavelet estimates are never exact and wavelets in recorded seismograms are inherently nonstationary. This capability also enables interpretation of thin beds in depth-migrated images, in which the dominant wavelength of the embedded wavelet will change as a function of velocity. The construction of a profile HMM using a wavelet estimate and a Gaussian noise model is discussed below.

As an example, we analyze the composite waveform introduced in Figure 1. The composite waveform (black trace, Figure 7a) was created by convolving a 25-Hz Ricker wavelet with the spike train shown in blue. In this case, guided by our prior knowledge, we select an initial solution space consisting of three spikes. That is, for the composite waveform, we find the model that best reproduces the composite waveform using three arrivals. For reference, Figure 7a displays possible manual picks (dashed orange lines) for this composite trace. Figure 7b shows the result obtained through our proposed technique with the correct source wavelet. The dashed red lines show the spike train that the Viterbi process has determined to be the best model. As expected, in this case, we recover the amplitude and relative timing information exactly because we knew exactly how many arrivals we should search for and had perfect knowledge of the source. The green line in Figure 7b represents the aligned model waveform, which comprises the means of all match and insert states traversed in the optimal state path. It is notable that the technique is capable of properly determining the amplitude and timing of these arrivals in which there is strong constructive/destructive interference. However, precise prior knowledge of the source wavelet is usually not available. Figure 7c shows an example using our method with a 30-Hz source wavelet. Even though we use the incorrect wavelet, the traveltime and amplitudes recovered are still better than those picked manually. Note that the green waveform in Figure 7c contains discontinuities due to traversing insert states. This simple example demonstrates the utility of our proposed technique in determining arrival time and amplitude information for a composite wavelet with overlapping arrivals.

The number of arrivals modeled is a key parameter of the candidate solution space. Figure 8 demonstrates the behavior of the VSSD algorithm for different numbers of modeled arrivals. The first trace in Figure 8 illustrates a synthetic waveform composed of five arrivals. The true amplitudes and traveltimes are depicted by the blue lines. The subsequent traces show VSSD results for model spaces containing one to seven arrivals, along with the log-odds scores. When only one arrival is allowed, only a portion of the waveform is matched (the score only reflects the portion of the waveform matched). Allowing for two arrivals increases the portion of the waveform that is matched, but three arrivals are needed to cover the entire duration of the waveform. Note, however, that the log-odds score actually decreases when going from two arrivals to three arrivals. Despite the fact that more match states are traversed, the amplitude discrepancies are larger in the three-arrival result for this example. The four-arrival result increases the fidelity of the fit as well as the log-odds score. The five-arrival result provides a near-perfect fit. Increasing the number of modeled arrivals beyond five in this example does not change the result in a significant way, as the bottom two traces in Figure 8 demonstrate. The sixth and seventh picked arrivals represent spurious low-amplitude events,

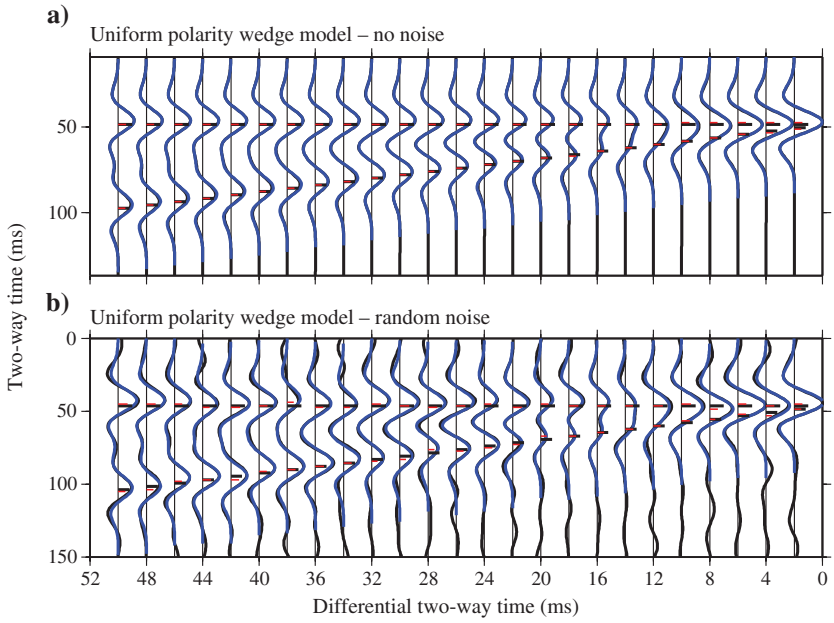

Figure 9. (a) Uniform-polarity wedge model (black curves and spikes) with Viterbi waveforms (blue traces) and picks (red spikes) overlain. The blue curves represent the Viterbi-aligned waveforms corresponding to the means of amplitude distributions for traversed match and insert states. The red spikes represent the amplitudes and traveltimes of the Viterbi-aligned spike train. (b) The same as (a) with the addition of band-limited random noise.

which are essentially fitting quantization errors from the conversion from the floating point to integer representations of the input trace. The amplitude of the seventh event (red circle) is so low that it is not visible in the figure.

Within a small window, an arbitrary number of arrivals may be modeled if the amplitudes for any given arrival are allowed to be zero. As the number of arrivals grows, the solution space grows exponentially, and it may become computationally unfeasible. To mitigate this problem, the method is generally applied in a series of small, overlapping windows in which only a small number of arrivals are modeled. Regularization may be applied by enforcing consistency of the results within the window overlaps. The problem of nonuniqueness can either be mitigated with a priori knowledge about the distribution of expected arrivals (e.g., by using impedance logs) or by penalizing results that do not show lateral consistency as interference patterns evolve across traces. The VSSD method may be used to tie well logs to seismic data in time or in depth because the nonlinear alignment can accommodate velocity dependent dominant wavelengths in the source function. The VSSD method may also be used to extend interpretations away from a well. In what follows, we show how this technique can be applied to determining bed thickness and arrival amplitude for wedge models similar to those in the Introduction.

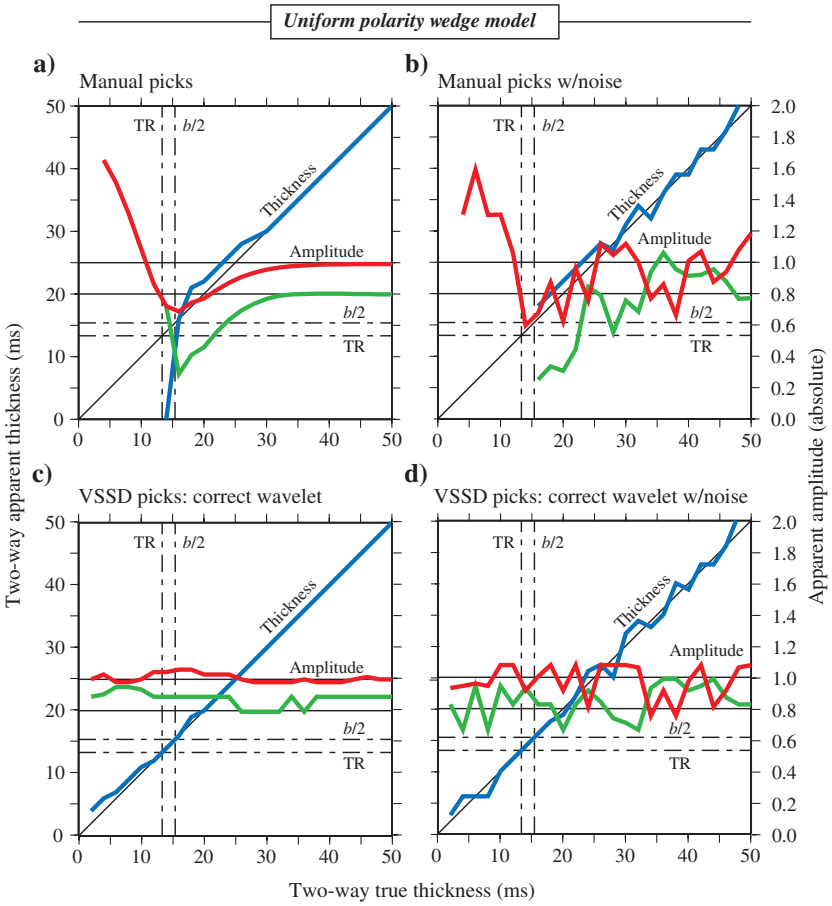

Figure 10. (a) Manual thickness and amplitude picks for the uniform-polarity wedge model. The blue line represents picked bed thickness, the red line represents picked amplitudes for the top interface, and the green line represents picked amplitudes for the lower interface. (b) The same as (a) except with the addition of band-limited random noise. (c) Viterbi thickness and amplitude picks using the correct 25-Hz, 15° phase-rotated wavelet. (d) The same as (c) except with the addition of band-limited random noise.

RESOLUTION TESTS

To examine the effectiveness of the Viterbi technique in determining arrival time and amplitude information for overlapping waveforms, we apply it to synthetic seismograms created for wedge models similar to those introduced in the Introduction. The advantage of analyzing the wedge models is that we can systematically examine amplitudes and differential traveltimes of arrivals through a steady variation in overlap of the arrival wavelets. First, we examine uniform-polarity wedge models, and then we examine the mixed-polarity case.

Uniform-polarity wedge models

A uniform-polarity wedge model is generated with zero-offset reflection coefficients for the upper and lower interfaces set at a constant ratio of 0.8 (see the "Introduction" section for further description on the basic design of the wedge models). That is, if we normalize the reflection coefficient of the upper interface to 1.0, then the lower interface has a normalized reflection coefficient of 0.8. The synthetic traces are constructed through convolution with a 25-Hz Ricker wavelet with a 15° phase rotation. A second set of synthetic traces is constructed with the addition of band-limited random noise with an S/N of 10. The nonuniform reflection coefficients, deviations from zero-phase, and the addition of random noise are intended to simulate real-world conditions more realistically.

Figure 9a shows the noise-free uniform-polarity wedge model waveforms (black traces) and spikes representing the true arrival times and amplitudes (black spikes), overlain by VSSD waveforms (blue traces) and VSSD spikes (red spikes). A common problem in deconvolution is that accurate information about the source time function is not always known. In this example, we use a 25-Hz, 15° phase-rotated wavelet, which is the correct source time function. When using the correct source, the Viterbi method provides almost-perfect traveltime variation all of the way down to the thinnest portion of the wedge. The largest deviations from the true thicknesses are 2 ms (5% of the dominant wavelength) at the smallest separation of the wedge model (at 2 and 4 ms differential two-way times; see Figure 9a). Figure 9b shows the results for the same uniform wedge model with the addition of band-limited random noise. The addition of the random noise increases the misfit between true arrival times (black spikes) and VSSD-picked arrival times (red spikes), but the maximum error has only increased to 3 ms or 7.5% of the dominant wavelength.

In Figure 10a, we show thicknesses and amplitudes from manually picked arrivals for the noise-free case. The blue line represents two-way thickness, while the red and green lines represent amplitudes for the upper and lower interfaces, respectively. Because two separate arrivals are not resolved below TR, the amplitude of the upper interface arrival increases significantly. Figure 10b shows the manual picks for the noisy uniform-polarity wedge model. In this case, the two separate arrivals are only resolved down to $b/2$ and the amplitude of the upper-layer arrival is not accurately recovered even at the largest thicknesses of the wedge model.

In Figure 10c, we summarize results from the VSSD-picked thicknesses and amplitudes, corresponding to the red spikes in Figure 8a. The two events are clearly resolved for all thicknesses, and amplitudes are also reasonably well estimated for all thicknesses. The quality of the thickness estimate degrades slightly in the neighborhood of $b/2$ and TR, but it is still much better than that obtained from manual picks. Figure 10d shows the VSSD picks for the noisy wedge model. Both events remain resolved for all thicknesses. The thickness is also recovered in the thinnest region of the wedge (2-ms two-way thickness). There is more scatter in the amplitude measurements, but there is no spike in amplitude for the upper-layer amplitude as is seen in the manually picked case.

Mixed-polarity wedge models

A mixed-polarity wedge model is generated with zero-offset reflection coefficients for the upper and lower interfaces set at a constant ratio of -0.8. That is, if we normalize the reflection coefficient of the upper interface to 1.0, then the lower interface has a normalized reflection coefficient of -0.8. Once again, the synthetic traces are constructed through convolution with a 25-Hz Ricker wavelet with a 15° phase rotation, and a second set of synthetic traces is constructed with the addition of band-limited random noise.

In Figure 11a, we show thicknesses and amplitudes from manually picked arrivals for the noise-free case. Once again, the blue line represents two-way thickness, while the red and green lines represent amplitudes for the upper and lower interfaces, respectively. Unlike the uniform-polarity case, the two separate arrivals are resolved below TR, but the thickness is overestimated as the two wavelets merge into an approximation of the first derivative of the wavelet (Widess, 1973). Also, the amplitudes of the two events become indistinguishable before $b/2$ is reached. Figure 11b shows the manual picks for the noisy uniform-polarity wedge model. The effect of the noise is primarily manifested in additional scatter in the amplitude picks.

In Figure 11c, we summarize results from the VSSD-picked thicknesses and amplitudes. The two events are clearly resolved for all thicknesses with much greater fidelity below TR than is the case using manual picks. The amplitude picks are nearly perfect except for a small region around Widess' $b/8$ criteria (Widess, 1973), where the combined wavelet approximates the first derivative. Figure 11d shows the VSSD picks for the noisy wedge model. These picks have the same general characteristics as the noise-free picks; however, noise has introduced scatter into the amplitude and thickness estimates.

Sensitivity to errors in the wavelet phase, wavelet frequency, and random noise

Because we would typically not expect to have an exact wavelet with which to perform our analysis, the sensitivity of the VSSD technique to errors in the wavelet estimate is explored with a suite of different Ricker wavelets. To test the method's sensitivity to frequency content, 15 wavelets with peak frequencies equally spaced between 18 and 32 Hz, representing errors of ± 7 Hz, and with the correct 15° phase rotation were used. This range of frequency content represents a 39% increase in the dominant wavelength on the low end and a 22% decrease in the dominant wavelength on the high end. Twenty-one wavelets with the correct peak frequency and phase rotations equally spaced between $-35°$ and 65°, representing phase errors of $\pm 50°$, were used to test the method's sensitivity to

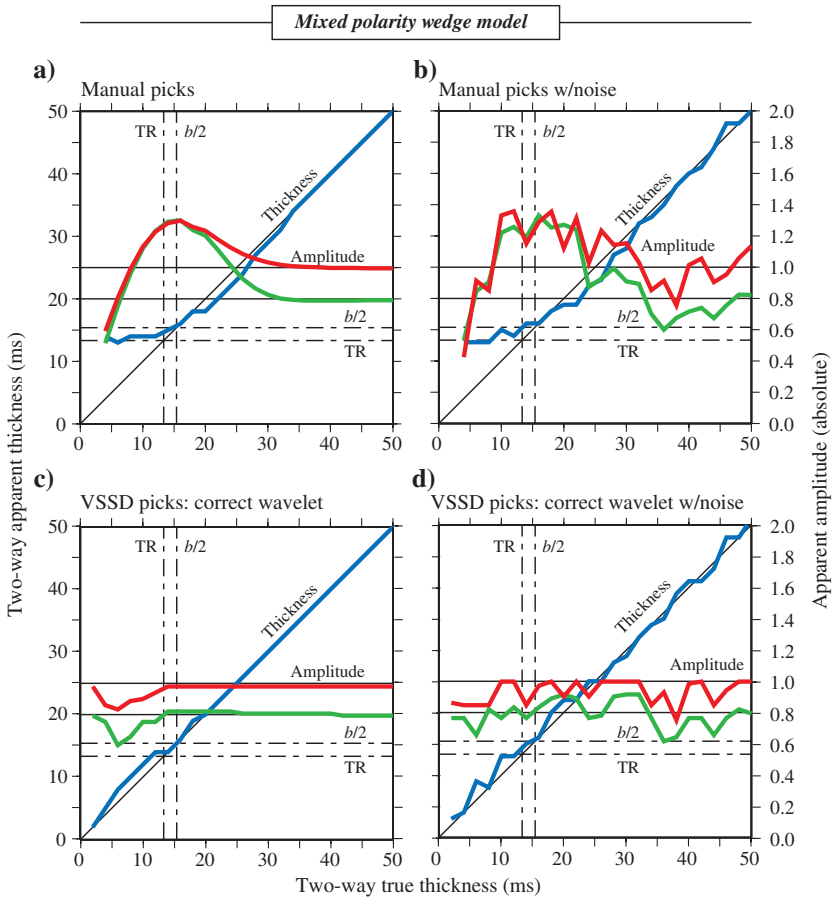

Figure 11. (a) Manual thickness and amplitude picks for the mixed-polarity wedge model. The blue line represents picked bed thickness, the red line represents picked amplitudes for the top interface, and the green line represents picked amplitudes for the lower interface. (b) The same as (a) except with the addition of band-limited random noise. (c) Viterbi thickness and amplitude picks using the correct 25-Hz, 15° phase-shifted wavelet. (d) The same as (c) except with the addition of band-limited random noise.

phase. The VSSD method begins to degrade for errors in phase and frequency content outside of these selected ranges, so that the two events are not consistently resolved for all thicknesses. However, the selected range of parameters provides an ample test suite spanning a larger range of errors than would typically be expected in a wavelet estimate. VSSD analysis was carried out separately for each of the 36 candidate wavelets for the noise-free and noisy uniform-polarity wedge models in Figure 9.

Figure 12 shows the VSSD-determined thicknesses for each of our test cases varying frequency, phase, and additive noise. Results for the 15 test wavelets with varying frequency content are shown in Figure 12a in the absence of noise. The color bar indicates the departure from the correct peak frequency of 25 Hz. The lower-frequency wavelets tend to underestimate the thickness slightly, while the higher-frequency wavelets tend to overestimate the thickness slightly, with a reversal of this general trend in the vicinity of TR and $b/2$ (i.e., between two-way thicknesses 13.33 and 15.38 ms). Wavelet contraction (delete states) tends to pull the center of the individual arrivals together when the frequency content is erroneously low, whereas wavelet dilation (insert states) tends to push the arrivals apart when the frequency content is erroneously high. The reversal of the trend between 12 and 22 ms is due to the effects of destructive interference between main and side lobes. The largest measured errors are 6 ms, with the errors being within 4 ms for all thickness greater than 8 ms. The addition of noise shows the same general trends as displayed in the noise-free model. However, the departure from true thickness is somewhat larger than in the noise-free model, showing that the addition of noise has a larger effect than the frequency content of the wavelet on thickness estimates.

The sensitivity to wavelet phase on thickness measurements is demonstrated in Figure 12b. This shows that thickness estimates are essentially insensitive to phase errors within the ±50° range. One outlier, an 8-ms error corresponding to a 35° phase rotation, indicates the beginning of a departure from acceptable phase errors. For phase rotations larger than ±50°, the wavelet shape has departed significantly enough from the true wavelet such that the algorithm

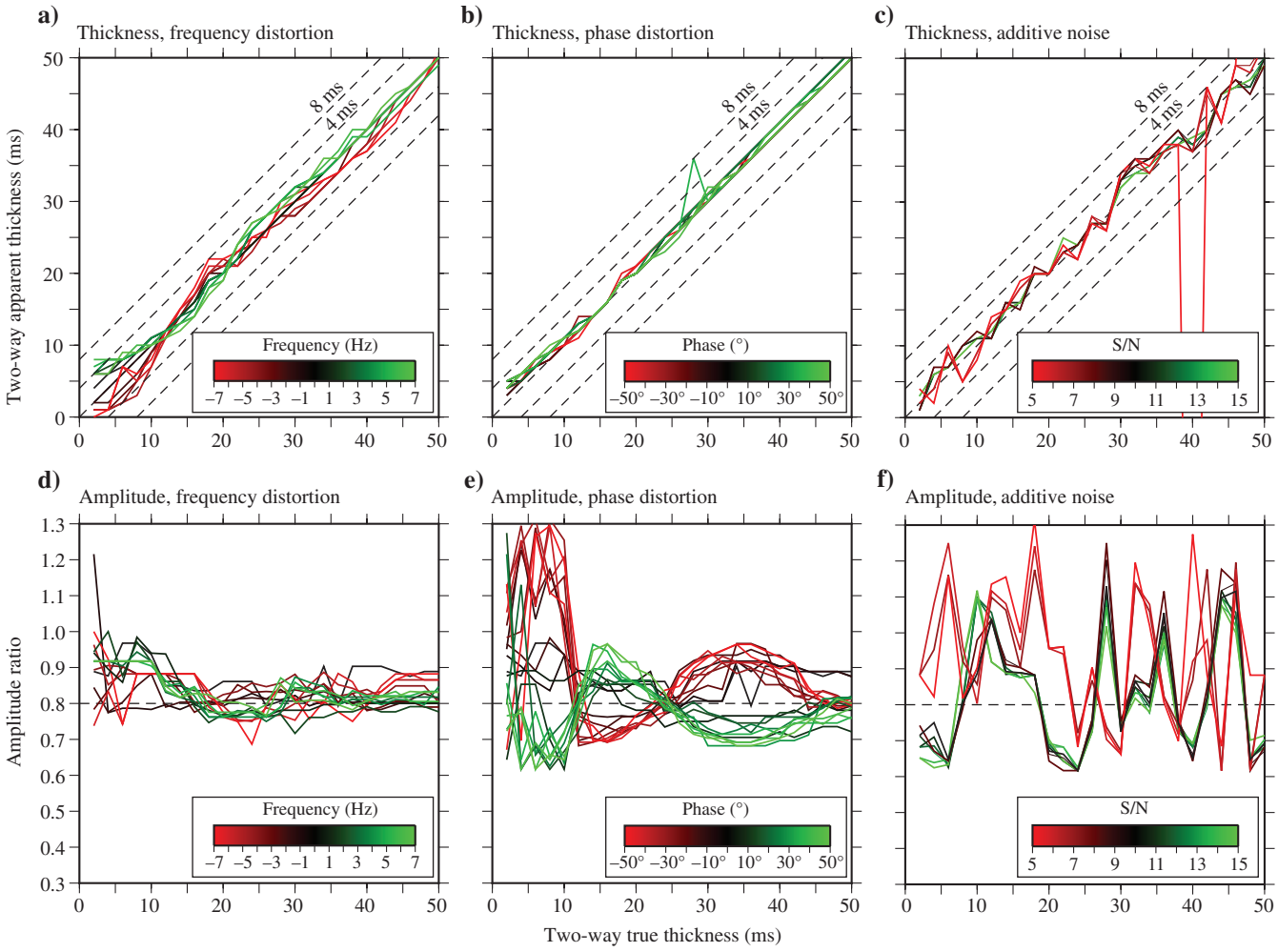

Figure 12. Apparent thickness curves for variations in the dominant frequency or phase of the source wavelet for the uniform-polarity wedge model. Results shown here in (a, b, d, e) are for the noise-free model. (a) Results for wavelets with dominant frequencies ranging from 18 to 32 Hz, representing an error range of ±7 Hz. (b) Results for wavelets with the correct dominant frequency, but with phase errors up to ±50°. (c) Results for wavelets with the correct dominant frequency and phase, but with additive noise with S/Ns from 5 to 15. (d) The same as in (a) except amplitude ratio measurements are displayed. (e) The same as in (b) except amplitude ratio measurements are displayed. (f) The same as in (c) except amplitude ratio measurements are displayed.

cannot accurately account for the effects of constructive and destructive interference across the entire range of differential traveltimes. The addition of noise slightly increases the error in thickness estimates, but it does not introduce a large departure from the noise-free case.

Figure 12d and 12e demonstrates the sensitivity of relative amplitude estimates to frequency content and phase rotation for the noise-free models. The value plotted for each wavelet is the ratio of the two picked amplitudes where the correct amplitude ratio is 0.8. Here, the sensitivities are reversed from those of thickness estimates. That is, when examining the thickness estimates, we observed that the measurements were less sensitive to differences in the wavelet phase than they were to differences in the frequency content. No clear dependence on frequency content can be inferred. That is, errors in amplitude ratio measurements for the noise-free and noisy case show deviations with no systematic trend with respect to wavelet frequency. Yet, phase errors demonstrate clear trends. Negative phase rotations (Figure 12e) tend to overestimate the amplitude ratio, while positive rotations tend to underestimate the ratio. This trend reverses in the 12–22-ms region, as was also noted in the frequency-dependent thickness estimates. The phase dependence of the amplitude estimates can be understood intuitively by considering the manner in which the shape of a wavelet changes as successively larger phase rotations are applied. Negative phase rotations tend to front load the energy in the wavelet, while positive phase rotations tend to back load the energy. For positive phase rotations at thicknesses below 12 ms, the back loading of energy gives more weight to the first arrival as the two wavelets constructively interfere. In the anomalous region between 12 and 22 ms, the opposite is true, as the back-loaded energy in the first arrival undergoes destructive interference. The addition of noise has more devastating effects on amplitude ratio measurements than on thickness estimates.

Figure 12c demonstrates the sensitivity of traveltime picks to different levels of random noise. In this case, the correct wavelet was used and the S/N varied from 5 to 15. The differential traveltime errors are within 4 ms, showing that the traveltimes are generally insensitive to the inclusion of random noise. However, at an S/N of five, a couple of the events are not correctly resolved, as indicated by the red lines that leave the area of the plot. In this case, a visual analysis of the waveforms would also give an erroneous interpretation. Figure 12f shows the sensitivity of amplitude picks to random noise. This result shows that the amplitude estimates are more sensitive to random noise than they are to wavelet errors in either phase or frequency, for all thicknesses. From these observations, one can infer that traveltime estimates are much more robust than amplitude estimates, especially in the presence of noise.

APPLICATION OF VSSD METHOD

We have presented the VSSD method for picking seismic arrival times and amplitudes. In certain cases, the VSSD method can resolve these arrivals below the standard resolution limits. In the previous section, we analyzed the sensitivity of this technique using a variable-thickness wedge model and synthetic seismograms. Here, we show the utility of this method with field data.

Our example demonstrates the application of the method at the exploration scale and comes from the Teapot Dome Oilfield. Teapot Dome is located ~35 miles north of Casper, Wyoming, on Naval Petroleum Reserve No. 3. The field has a long history of production dating back to the early 1900s, and it is currently used as a testing center for emerging technologies. Figure 13a shows an extraction of crossline 123 from a time-migrated 3D seismic cube provided by the Rocky Mountain Oilfield Testing Center (http://www.rmotc.doe.gov/datasets.html). Reflections in the analysis window are bounded by large, consistent arrivals (indicated by gray curves in Figure 13a). However, the central reflections exhibit substantial lateral variation due to subtle variations in bed thickness. The primary observations are as follows: (1) a pair of lower amplitude peaks visible just after the first large peak between inlines 75 and 88 that merges into a single peak as the inline number increases (feature is highlighted with dashed red lines), (2) increasing complexity of the preceding trough as the inline number increases (highlighted with a dashed green line), and (3) the third central peak is clearly visible in the lower inlines, but it diminishes by inline 108 (indicated with a dashed blue line). Here, we use the VSSD method to explore the subtle variations in these arrivals.

To use the VSSD technique, we must first choose an appropriate source time function. A wavelet estimate was obtained by aligning and stacking multiple traces along a set of consistent regional horizons. The results of our analysis are shown in Figure 13b for every

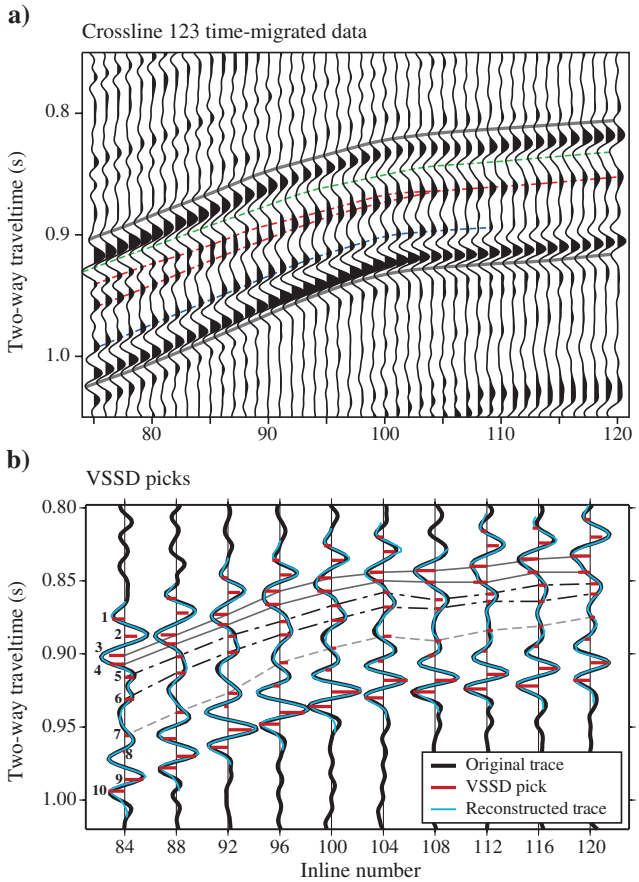

Figure 13. (a) A windowed portion of crossline 123 from the Teapot Dome 3D time-migrated data set. The gray lines indicate the VSSD analysis window. The dashed green, red, and blue lines indicate lateral variations in waveform character of particular interest. (b) VSSD picks (red lines) and reconstructed waveforms (blue curves) for every fourth trace. The VSSD method identified 10 arrivals (numbered on inline 84) with good lateral consistency across the section.

fourth trace from inlines 84–120. The method identifies 10 consistent reflectors that are numbered from 1 to 10 starting from the top and labeled for inline 84 in Figure 13b. One exception is the failure to resolve events 1 and 2 on inline 108. This means that the highest scoring alignment for that trace skipped several match states corresponding to these two arrivals. At either end of the analysis window, there remain constructive and destructive interference from neighboring events that are not considered. It is likely that the preceding events would need to be considered by expanding the analysis window to properly resolve events 1 and 2 on inline 108. The results corresponding to the primary observations of lateral inhomogeneity are discussed below.

- The two peaks highlighted by red dashed lines in Figure 13a correspond to arrivals 5 and 6 in Figure 13b (dashed-dotted line). The two peaks, clearly visible at inline 84, merge into a single arrival as the inline number increases, much like the wedge models from the Introduction. The VSSD picks show a small thinning trend as the inline number increases, but not as much as would be expected from a visual inspection of the waveform.
- Note, however, that the trough that precedes arrivals 5 and 6 (green dashed line in Figure 13a) corresponding to arrivals 3 and 4 in Figure 13b (solid gray lines) shows a thickening trend over this same lateral window. The result implies that the interference of all these arrivals produces a waveform that visually exaggerates the thinning of the bed bounded by arrivals 5 and 6. The combination of these four arrivals consistently matches the data across the range of inlines, indicating sedimentary beds with slight lateral thickness variations.
- The disappearing peak (blue dashed line in Figure 13a) corresponds to the zone of waveform interference between arrivals 7 (dashed gray line in Figure 13b) and 8. On the left side of the section, the peak correlates more strongly with arrival 7. By inline 96, the peak occurs between arrivals 7 and 8. As the inline number increases, the peak itself disappears, yet arrivals 7 and 8 remain with consistent polarities and relatively consistent amplitudes. One exception is the amplitude of arrival 8 on inline 88, which appears to be overestimated. This is likely due to strong constructive interference from arrivals 9 and 10. This amplitude may be better resolved by lengthening the analysis window.

Overall, the VSSD method, applied separately to each individual trace, has produced spike trains that match the data well while demonstrating a good degree of lateral consistency, despite the fact that there were no constraints promoting lateral consistency between traces. This improves our confidence in the VSSD method as a tool for thin-bed interpretation.

DISCUSSION

We have analyzed the sensitivity of the VSSD technique to variations in bed thickness, measuring apparent thickness and the amplitude ratio of arrivals reflecting off the layers. In addition, we added noise and source wavelet frequency and phase distortions. Table 1 summarizes the results from these tests. The apparent trends from Table 1 are

- Thickness estimates are relatively insensitive to errors in wavelet phase.
- When only random noise and frequency content are considered, amplitude estimates tend to be slightly more robust for mixed-polarity events.
- Amplitude estimates tend to be equally sensitive to random noise and large phase errors.
- Thickness estimates are much more robust than amplitude estimates.

Table 1. Results of the sensitivity of the VSSD technique to wavelet errors and random noise.

Uniform-polarity wedge models				Mixed-polarity wedge models			
Noise	Δf (Hz)	$\Delta\varphi$ (°)	Max error (%)	Noise	Δf (Hz)	$\Delta\varphi$ (°)	Max error (%)
Thickness measurements as a percentage of dominant wavelength							
No	0	0	5	No	0	0	5
Yes	0	0	5	Yes	0	0	5
No	±7	0	10	No	±7	0	12.5
Yes	±7	0	15	Yes	±7	0	15
No	0	±50	5	No	0	±50	5
Yes	0	±50	10	Yes	0	±50	7.5
Amplitude ratio measurements							
No	0	0	21	No	0	0	9
Yes	0	0	64	Yes	0	0	27
No	±7	0	25	No	±7	0	15
Yes	±7	0	64	Yes	±7	0	45
No	0	±50	62	No	0	±50	64
Yes	0	±50	64	Yes	0	±50	64

The VSSD technique is a versatile tool that provides accurate timing and amplitude information for seismic arrivals when these arrivals are nearly overlapping. However, this technique is not a blind-deconvolution method. That is, the method requires a priori knowledge about (1) the source wavelet and (2) the number of expected arrivals within the time window of interest.

With respect to the first point, the analyses in the section Resolution Tests primarily show the sensitivity of this technique with respect to errors in our knowledge of the seismic source. In exploration seismology applications, the source function is generally known well enough to be adequate for the VSSD method. For applications in earthquake seismology, we often do not have detailed knowledge of the source time function. However, empirical source time functions can be generated to alleviate this problem. For example, it is possible to stack reference seismic phases to build an empirical source time function (e.g., Thorne and Garnero, 2004).

With respect to the second point, precise knowledge of the number of arrivals within our time window is not necessary. In practice, we know the approximate number of expected arrivals. One can either test different numbers of arrivals, comparing log-odds scores and the quality of the waveform match, or one can allow the amplitudes of arrivals to be zero, which will explicitly search over multiple numbers of arrivals in a single experiment. Blocked impedance models derived from well logs present an excellent source of information to constrain the solution space. This method could be used to perform nonlinear seismic-well ties in either time or depth.

Figures in this paper were drawn using the Generic Mapping Tools (http://gmt.soest.hawaii.edu/; Wessel and Smith, 1998).

CONCLUSIONS

We have presented a stochastic method to decompose individual seismic traces into sparse spike trains describing the traveltimes and relative amplitudes of individual arrivals comprising band-limited signals. The method has been shown to have resolving power below Rayleigh's and Ricker's criteria under appropriate conditions. It has also been shown that estimated traveltimes are relatively insensitive to random noise and errors in the source time function estimate. No assumptions are made regarding the distribution of traveltimes and amplitudes; however, interpreted information can be used to limit the solution space. Application to field data has shown that single-channel sparse spike detection produces results with good lateral consistency. This technique shows promise in targeted exploration scale interpretation, and it may provide an alternative method of tying seismic images to well logs. Further applications may include decomposing overlapping pre- and postcursor teleseismic phases into individual arrivals or improving arrival time measurements in earthquake location or seismic tomography problems.

ACKNOWLEDGMENTS

We thank the Rocky Mountain Oilfield Testing Center (RMOTC) and the United States Department of Energy for providing the Teapot Dome seismic data set. We gratefully acknowledge the University of Utah Center for High Performance Computing (CHPC) for computer resources and support. We would also like to thank five anonymous reviewers for their constructive suggestions and comments. MT and SB were partially supported by National Science Foundation (NSF) grant EAR-0952187, and SB was partially supported by a grant from Total E&P.

REFERENCES

Bacon, M., R. Simm, and T. Redshaw, 2003, 3-D seismic interpretation: Cambridge University Press.

Clapp, R. G., 2008, Lloyd and Viterbi for QC and auto-picking: Stanford Exploration Project, SEP-134.

Durbin, R., S. R. Eddy, A. Krogh, and G. Mitchison, 1998, Biological sequence analysis: Probabilistic models of proteins and nucleic acids: Cambridge University Press.

Eddy, S. R., 1995, Multiple alignment using hidden Markov models: *in* C. J. Rawlings, ed., Proceedings of the Third International Conference on Intelligent Systems for Molecular Biology: AAAI Press, 114–120.

Heimer, A., I. Cohen, and A. A. Vassiliou, 2007, Dynamic programming for multichannel blind seismic deconvolution: 77th Annual International Meeting, SEG, Expanded Abstracts, 1845–1849.

Heimer, A., I. Cohen, and A. A. Vassiliou, 2009, Multichannel seismic modeling and inversion based on Markov-Bernoulli random field: 79th Annual International Meeting, SEG, Expanded Abstracts, 2322–2326.

Kaaresen, K. F., and T. Taxt, 1998, Multichannel blind deconvolution of seismic signals: Geophysics, **63**, 2093–2107, doi: 10.1190/1.1444503.

Kallweit, R. S., and L. C. Wood, 1982, The limits of resolution of zero-phase wavelets: Geophysics, **47**, 1035–1046, doi: 10.1190/1.1441367.

Kormylo, J. J., and J. M. Mendel, 1983, Maximum-likelihood seismic deconvolution: IEEE Transactions on Geoscience and Remote Sensing, **GE-21**, 72–82, doi: 10.1109/TGRS.1983.350532.

Liner, C. L., and R. G. Clapp, 2004, Nonlinear pairwise alignment of seismic traces: Geophysics, **69**, 1552–1559, doi: 10.1190/1.1836828.

Ricker, N., 1953, Wavelet contraction, wavelet expansion, and the control of seismic resolution: Geophysics, **18**, 769–792, doi: 10.1190/1.1437927.

Sacchi, M. D., D. R. Velis, and A. H. Cominguez, 1994, Minimum entropy deconvolution with frequency-domain constraints: Geophysics, **59**, 938–945, doi: 10.1190/1.1443653.

Thorne, M. S., and E. J. Garnero, 2004, Inferences on ultralow-velocity zone structure from a global analysis of SPdKS waves: Journal of Geophysical Research, **109**, 1978–2012, doi: 10.1029/2004JB003010.

Velis, D. R., 2008, Stochastic sparse-spike deconvolution: Geophysics, **73**, no. 1, R1–R9, doi: 10.1190/1.2790584.

Viterbi, A., 1967, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm: IEEE Transactions on Information Theory, **13**, 260–269, doi: 10.1109/TIT.1967.1054010.

Wessel, P., and W.H.F. Smith, 1998, New, improved version of generic mapping tools released: EOS — Transactions of the American Geophysical Union, **79**, 579, doi: 10.1029/98EO00426.

Widess, M. B., 1973, How thin is a thin bed?: Geophysics, **38**, 1176–1180, doi: 10.1190/1.1440403.

Wiggins, R. A., 1978, Minimum entropy deconvolution: Geoexploration: International Journal of Mining and Technical Geophysics and Related Subjects, **16**, 21–35, doi: 10.1016/0016-7142(78)90005-4.