## Full-waveform adjoint tomography in a multiscale perspective

*Yanhua O. Yuan* and Frederik J. Simons, Princeton University; Ebru Bozdağ†, Université de Nice, Sophia-Antipolis*

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

### SUMMARY

We discuss an algorithm to regularize elastic waveform inversions using wavelet-based constructive approximations of the data, synthetic and observed, in models that evolve as part of a gradient-based iterative scheme relying on forward and adjoint modeling carried out with a spectral-element method. For an elastic Marmousi model we show how our wavelet-based multiscale waveform inversion proceeds successively from large to small scales in the seismograms, with a progressive increase of the complexity of the resulting model. We explore the sensitivity of surface waves in imaging shallow structure. To circumvent cycle skipping we designed an envelope-misfit function within a wavelet-multiscale framework. We test our approach in a toy model in preparation for inversions at full complexity.

### INTRODUCTION

In seismic tomography, the distance between observed and synthetic seismograms can be measured as picked or crosscorrelated travel times, amplitude anomalies, or via waveform subtraction (Tarantola, 1984; Nolet, 1987; Luo and Schuster, 1991; Dahlen and Baig, 2002). If they converge (Gauthier et al., 1986; Mora, 1987; Bunks et al., 1995), full waveform-difference inversions yield higher-resolution images. Nonlinearity leads to local minima in the objective function (Alkhalifah and Choi, 2012), especially when the starting model is far from the target, or when it contains details of great complexity.

Yuan and Simons (2014) developed a wavelet-based multiscale approach to waveform inversion. Wavelet decomposition has advantages over Fourier filtering: flexibility in basis selection, efficiency of signal representation, convergence and misfit reduction. Yuan and Simons (2014) applied wavelet-scale decomposition of data generated in an elastic Marmousi model, to implement a multiscale scheme that works successively from coarse to finer scales, retrieving smooth background structure before heterogeneities of great complexity. They did not consider surface waves but removed them before inversion.

Surface waves are important to constrain shallow structure, and provide corrections for deep imaging. The challenge in the inversion of surface-wave waveforms lies in cycle skipping. To alleviate this problem, we developed an envelope-based objective function (Bozdağ et al., 2011) to measure oscillatory surface waves as part of a multiscale strategy.

### Multiscale Waveform Adjoint Tomography

Adjoint methods (e.g. Tarantola, 1984, 1986) allow much choice to measure the distance between predicted and observed data. The expression of the misfit gradient or kernel is unchanged — only the adjoint source function has to be adjusted.

*Waveform adjoint method:* Waveform-difference tomography solves the full elastic wave-propagation problem in heterogeneous media, explaining all the available recorded information. Upon convergence, waveform tomography reveals more structural information than traveltimes (Luo and Schuster, 1991).

The *waveform-difference misfit* function $\chi(\mathbf{m})$ in a model $\mathbf{m}$ is the sum of the residuals between synthetics $\mathbf{s}(\mathbf{x}_r,\mathbf{x}_s,t;\mathbf{m})$ and observations $\mathbf{d}(\mathbf{x}_r,\mathbf{x}_s,t)$, over all sources $s$ at $\mathbf{x}_s$ and receivers $r$ at $\mathbf{x}_r$, over some time window $T$ (Tromp et al., 2005):

$$\chi(\mathbf{m}) = \frac{1}{2}\sum_{s,r}\int_0^T \left\|\mathbf{s}(\mathbf{x}_r,\mathbf{x}_s,t;\mathbf{m})-\mathbf{d}(\mathbf{x}_r,\mathbf{x}_s,t)\right\|^2 dt. \quad (1)$$

Gradient-based methods require the derivative of the misfit, conveniently expressed in the form of a volume integral of a sensitivity kernel against model perturbations. The misfit kernels relate to data misfit via the zero-lag crosscorrelations of the adjoint and forward wavefields. The adjoint wavefield can be calculated numerically by running the forward model with adjoint sources at the receivers instead of earthquake sources.

The *waveform adjoint source* can be written as:

$$\mathbf{f}^\dagger(\mathbf{x},t) = \sum_r \left[\mathbf{s}(\mathbf{x},t;\mathbf{m})-\mathbf{d}(\mathbf{x},t)\right]\,\delta(\mathbf{x}-\mathbf{x}_r). \quad (2)$$

The adjoint wavefield is obtained by back-projecting the time-reversed residuals between predicted and observed waveforms at receiver $\mathbf{x}_r$. The gradient of the misfit function is the opposite update direction in steepest-descent optimization.

*Wavelet-based multiscale approach:* We combat nonlinearity in waveform inversion via wavelet transformation. Working successively from long to short wavelengths is a powerful strategy to approach the global minimum (Nolet et al., 1986). For long-wavelength measurements, the number of local minima is reduced, and the inversion problem faster to converge to the global solution (de Hoop et al., 2012), or to a local minimum in its neighborhood (Bunks et al., 1995; Brossier et al., 2009).

Instead of using full-resolution seismograms, we apply a wavelet transform to break down the seismograms to different multiresolution levels $j$, which yields the subbands $\mathbf{s}_j(\mathbf{x}_r,\mathbf{x}_s,t;\mathbf{m})$ and $\mathbf{d}_j(\mathbf{x}_r,\mathbf{x}_s,t)$. The multiscale *waveform-difference misfit* function $\chi_j$ at a resolution level $j$ is defined as

$$\chi_j(\mathbf{m}) = \frac{1}{2}\sum_{s,r}\int_0^T \left\|\mathbf{s}_j(\mathbf{x}_r,\mathbf{x}_s,t;\mathbf{m})-\mathbf{d}_j(\mathbf{x}_r,\mathbf{x}_s,t)\right\|^2 dt. \quad (3)$$

The *multiscale waveform adjoint source* can be expressed as:

$$\mathbf{f}_j^\dagger(\mathbf{x},t) = \sum_r \left[\mathbf{s}_j(\mathbf{x},t;\mathbf{m})-\mathbf{d}_j(\mathbf{x},t)\right]\,\delta(\mathbf{x}-\mathbf{x}_r). \quad (4)$$

We start from a certain maximum decomposition level, and the updated solution from the quasi-linear problem at the large scale serves as a starting point, closer to the global target, for subsequent inversions at smaller scales.

© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0816.1

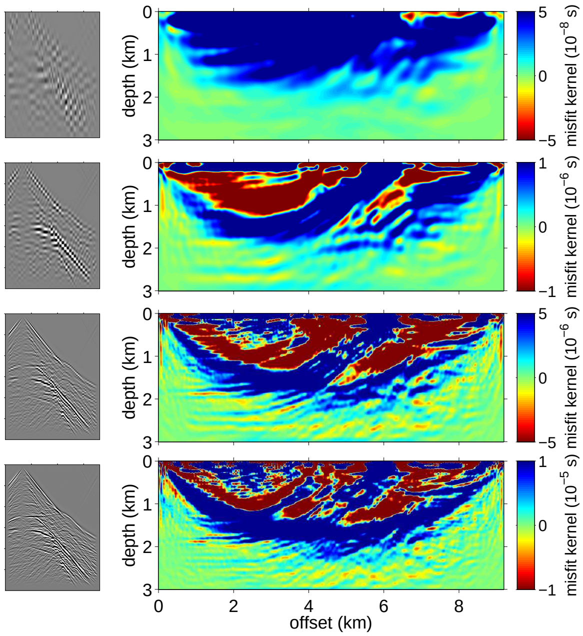


Figure 1: (*Left*:) Wavelet-subspace representations of processed shot gathers, at different levels. (*Right*:) Corresponding waveform-adjoint kernels showing increased complexity.

*Choice of wavelets and decomposition parameters:* The basis best suited depends on the data under consideration. Wavelet expansions under which synthetics and observations have a high degree of similarity are preferred. Computational cost of wavelet analysis and synthesis should be considered, especially for massive data processing. The "best" decomposition depth should provide a good starting point at which the synthetic seismograms are close to the corresponding observations in the subspace defined by the wavelet basis at the maximum scale. Successive reconstructions are terminated when they resemble the input according to a misfit convergence criterion.

*Numerical experiments:* We illustrate the performance of our method on data generated in the Marmousi model (Versteeg, 1993), converted to an elastic model as a Poisson solid. The maximum frequency modeled is 25 Hz, using a Ricker-wavelet source. We applied free-surface conditions at the top and Perfectly Matching Layer absorbing boundary conditions (Festa and Nielsen, 2003) on the remaining three sides of the model domain. Band-pass and dip-filtering in Seismic Un*x (Stockwell, 1999) removed surface waves .

We apply the Daubechies (1988) D12 wavelet transform (with six vanishing moments) to seismic data in the time domain. Subspace decompositions of one shot gather are shown in the left column of Figure 1. The right column shows the corresponding shear-wave speed misfit kernels, for all shots and all stations. They capture the discrepancy between the current model and the target at each of the wavelet scales in the seismograms. With decreasing scale in data space, there is an increase in complexity in structure and an increase in the number of local minima in the associated misfit contours. Our stable multiscale scheme for waveform inversion works successively from large-scale data fitting to small-scale explaining, progressively revealing coarse, then detailed heterogeneities.

Figure 2 shows the sequence of normalized data residual and model norms as the algorithm progresses. The left panel shows the overall rms misfit within the scale levels of the approximation for the iterations, on a log scale. This residual norm is decreased by the adjoint modeling within each scale until convergence. When switching to the seismograms at the next level, there is an uptick in the residual norm due to the inclusion of extra detail in the seismogram. The black line in the middle panel shows the evolution of the rms misfit for all scales without (*black line*) and with (*red line*) surface waves. Since surface waves were not considered in our inversions, the behavior of the latter curve is much more erratic throughout the iterations. The model-norm evolution is in the rightmost panel, separately for the compressional ($\alpha = V_P$), shear wavespeed ($\beta = V_S$), and their combination.

The Marmousi *P*-wave speed model is in Figure 3 (*top left*). To get an initial model far from the target, we smoothed the target model with an isotropic Gaussian kernel, see Figure 3 (*bottom left*). The final *P*- and *S*-wave speed models are shown in Figure 3 (*right column*) after 301 iterations of the multiscale waveform-difference adjoint modeling. The upper part of the Marmousi model has been very well recovered. The lower part suffers from lower resolution, due to insufficient ray coverage.

**Multiscale Envelope Inversion of Surface Waves**

We use a toy model (Figure 4, *left*) to illustrate the challenges of surface waves in waveform inversion. The $400 \times 100$ m model has constant $V_P = 2000$ m/s and density $\rho = 1000$ kg/m$^3$. The shear-wave model consists of a homogeneous background ($V_{S_0}$=800 m/s) and an anomaly in a curved 10 m-thick layer ($V_{S_1} = 1000$ m/s). We use a 40 Hz Ricker vertical source at $x = 50$ m, 0.5 m below the surface, and a receiver at $x = 350$ m, the same depth as the source. The modeled surface waves (Figure 5, *top left*) are strongly dispersive, while the synthetic surface waves (Figure 5, *bottom left*) in the uniform model of $V_S = 900$ m/s are not.

*Cycle skipping of surface waves:* This may occur when an adequate initial model is not available. Figure 5 (*left column*) shows a clear discrepancy between the predicted and target surface-wave waveforms. Figure 6 (*top*) shows the waveform-difference misfit contour (Solano, 2013) with respect to the two shear-wave speeds $V_{s_0}$ and $V_{s_1}$. Cycle skips cause waveform inversions to converge to secondary minima. To combat this, we work with the waveform envelopes, which in Figure 5 are shown together with their associated waveforms.

Figure 5 (*right column*) shows the surface-wave waveforms and envelopes projected onto scale 8 using D12 wavelets. The envelopes are very consistent with each other, and the corresponding 2-D misfit contours (Figure 6, *bottom*) display a wide convergence basin devoid of the numerous local minima present in the original waveform misfit contour plot. The white circles in Figure 6 track the update path after one iteration starting from a uniform model denoted by the red circles.

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

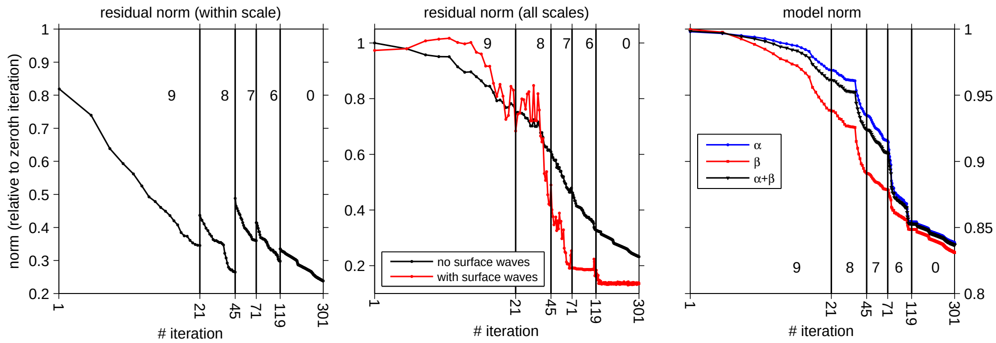


Figure 2: Residual norm evolution, with the scales within which the adjoint optimization is being conducted noted. At the marked points, additional (lower) scales of the seismograms are introduced. Data norms are normalized to the initial-model residual norm, within the scale of the approximation *(left)*, considering the full-resolution seismograms with and without surface waves *(middle)*, and in the model space *(right)*, for the compressional and shear wavespeed portions of the Marmousi model, separately and together.

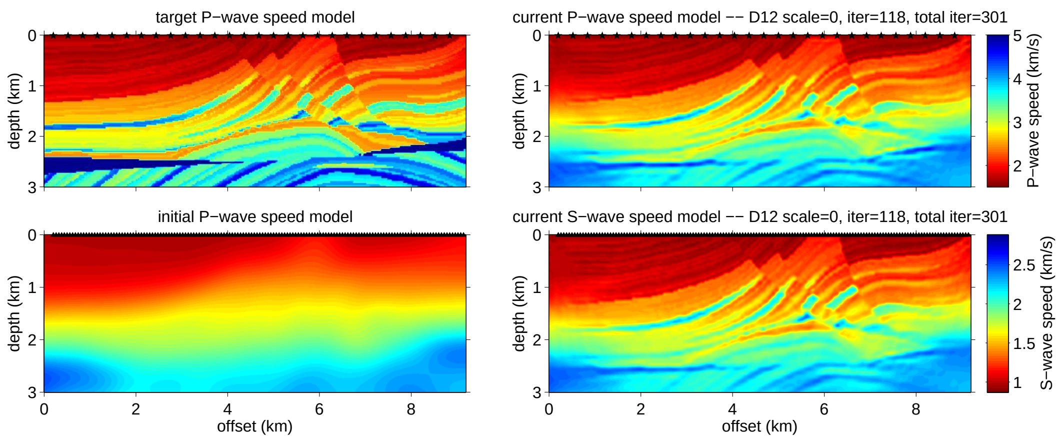


Figure 3: (*Left*:) The compressional-wavespeed Marmousi model (*top*), and its isotropic Gaussian smoothed version (*bottom*) which is our starting model. Sources are marked by stars and receivers by triangles. (*Right*:) Final $V_P$ and $V_S$ models obtained after 301 multiscale waveform-difference adjoint inversion steps, using all the available multiscale information in the seismograms.

*Envelopes:* The analytic signal of a real-valued signal $x(t)$ can be expressed as (Claerbout, 1992):

$$x_a(t) = x(t) + i\mathscr{H}\{x(t)\} = E(t)e^{i\phi(t)}, \tag{5}$$

where $\mathscr{H}\{x(t)\}$ is the Hilbert transform of the real signal $x(t)$; $\phi(t)$ and $E(t)$ stand for the instantaneous phase and the instantaneous amplitude (or envelope) of the analytic signal:

$$\phi(t) = \arctan\frac{\mathscr{H}\{x(t)\}}{x(t)},$$
$$E(t) = \sqrt{x^2(t) + \mathscr{H}^2\{x(t)\}}. \tag{6}$$

Instead of focusing on oscillatory phases, inversions based on envelopes are able to reduce the nonlinearity of waveforms.

*Envelope-based adjoint method:* Our least-squares *envelope-difference misfit* function of observed $\mathbf{d}(\mathbf{x}_r, \mathbf{x}_s, t)$ and synthetic $\mathbf{s}(\mathbf{x}_r, \mathbf{x}_s, t; \mathbf{m})$ data is inspired by Bozdağ et al. (2011):

$$\chi(\mathbf{m}) = \frac{1}{2}\sum_{s,r}\int_0^T \left\| \mathbf{E}^{syn}(\mathbf{x}_r, \mathbf{x}_s, t; \mathbf{m}) - \mathbf{E}^{obs}(\mathbf{x}_r, \mathbf{x}_s, t) \right\|^2 dt. \tag{7}$$

The associated adjoint source can be expressed as:

$$\mathbf{f}^{\dagger}(\mathbf{x}, t) = \sum_r \left[\mathbf{E}^{ratio}\,\mathbf{s} - \mathscr{H}\{\mathbf{E}^{ratio}\,\mathscr{H}\mathbf{s}\}\right]\delta(\mathbf{x} - \mathbf{x}_r), \tag{8}$$

where $\mathbf{E}^{ratio}$ captures the difference of current predicted and target envelopes:

$$\mathbf{E}^{ratio}(\mathbf{x}, t; \mathbf{m}) = \frac{\mathbf{E}^{syn}(\mathbf{x}, t; \mathbf{m}) - \mathbf{E}^{obs}(\mathbf{x}, t)}{\mathbf{E}^{syn}(\mathbf{x}, t; \mathbf{m})}. \tag{9}$$

The adjoint source will be re-transmitted from all stations $\mathbf{x}_r$ simultaneously to generate the *adjoint wavefield*, which illuminates the discrepancy of the observed and predicted envelopes.

*Synthetic experiment:* We carry out multiscale envelope-based inversions using the toy model shown in Figure 4 starting from a homogeneous model of $V_S = 900$ m/s. We use 39 vertical sources with a 40 Hz Ricker wavelet located at 0.5 m below the surface with 10 m horizontal spacing between 10 m and 390 m, and 400 receivers spaced 1 m apart at the source depth. The final estimated model after 44 iterations is shown in the right panel of Figure 4.

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0816.1

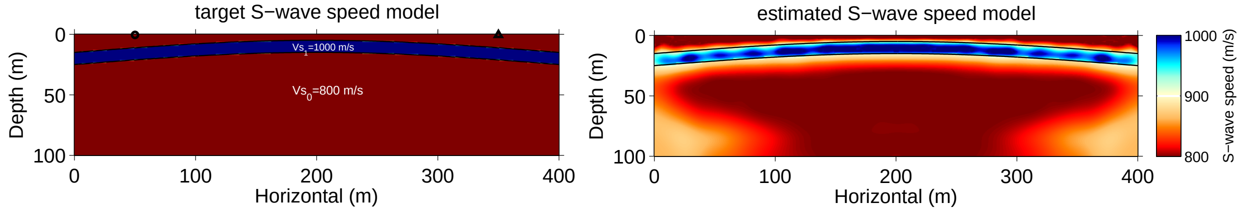


Figure 4: (*Left*:) Shear-wave speed model of a homogeneous background of 800 m/s and an anomalous layer of 1000 m/s. The circle denotes a source and the triangle marks a receiver. (*Right*:) Estimated $V_S$ model using the multiscale envelope approach.

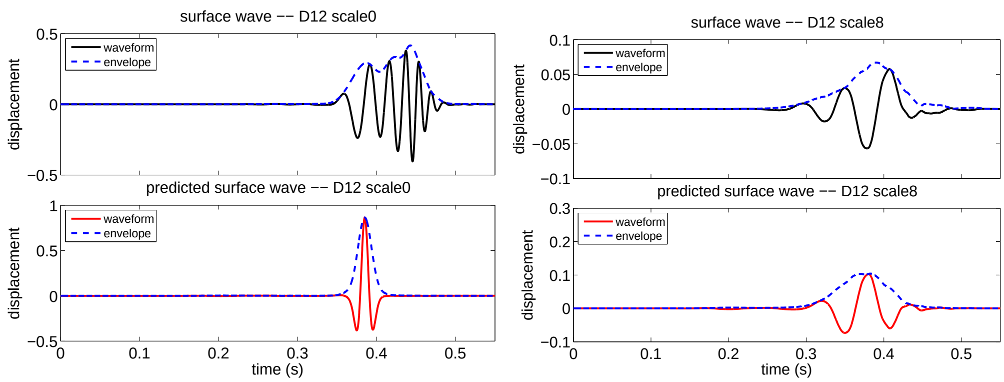


Figure 5: (*Top*:) Modeled surface-wave waveforms and envelopes with the shear-wave speed model using the source-receiver pair marked in Figure 4 left, at full-resolution (*left*) and scale 8 using D12 (*right*). (*Bottom*:) Predicted surface-wave waveforms and envelopes at full-resolution (*left*) and scale 8 (*right*), with a homogeneous shear-wave speed model of 900 m/s.

## CONCLUSIONS

We have formalized a multiscale approach for full-waveform adjoint tomography based on the wavelet transform. We work progressively from large-scale data fitting to finer-scale more detailed explanations. Progressive refinements in data space result in increasing complexity in the tomographic model updates. The use of surface waves brings its own challenges in the form of cycle skipping, which we combat with a special treatment using multiscale envelopes. Based on our successful numerical experiments of waveform inversions of body and surface waves, we have now designed a complete scheme for true "full-waveform inversion", without prior separation of body and surface waves, making the explanation of all available information in the seismogram possible to great detail.

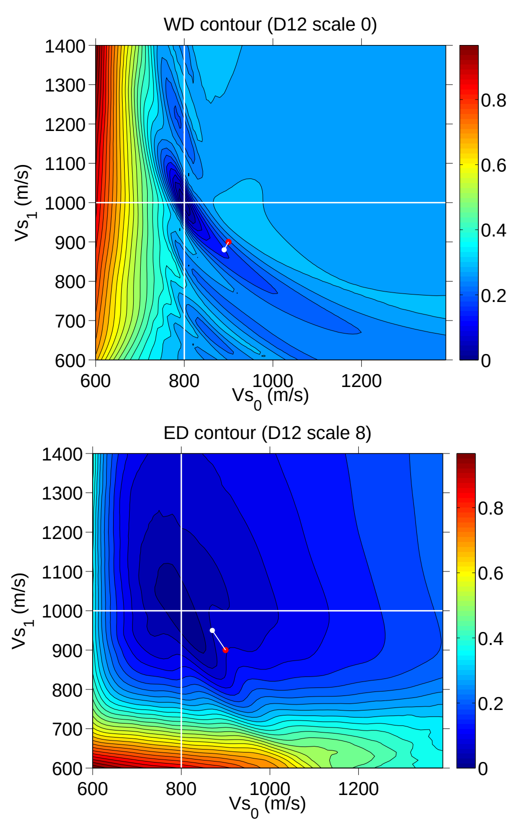


Figure 6: Misfit contours with respect to the shear-wave speeds $V_{S_0}$ and $V_{S_1}$ using waveform-difference (WD) (*top*) and multiscale envelope-difference (ED) (*bottom*) measurements. The intersecting white lines represent the toy model and the red circle denotes a uniform model of $V_S = 900$ m/s. The red circle denotes the estimated model after one iteration using the respective measurements, and the line connecting the red and white circles shows the first update direction.

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/


DOI http://dx.doi.org/10.1190/segam2014-0816.1

http://dx.doi.org/10.1190/segam2014-0816.1

**EDITED REFERENCES**

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

**REFERENCES**

Alkhalifah, T., and Y. Choi, 2012, Taming waveform inversion nonlinearity through phase unwrapping of the model and objective functions: Geophysical Journal International, **191**, 1171–1178, http://dx.doi.org/10.1111/j.1365–246X.2012.05699.x.

Bozdag, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full-waveform inversion based on instantaneous phase and envelope measurements: Geophysical Journal International, **185**, no. 2, 845–870, http://dx.doi.org/10.1111/j.1365-246X.2011.04970.x.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, no. 6, WCC105–WCC118, http://dx.doi.org/10.1190/1.3215771.

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

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell.

Dahlen, F. A., and A. Baig, 2002, Frechet kernels for body-wave amplitudes: Geophysical Journal International, **150**, no. 2, 440–466, http://dx.doi.org/10.1046/j.1365-246X.2002.01718.x.

Daubechies, I., 1988, Orthonormal bases of compactly supported wavelets: Communications on Pure and Applied Mathematics, **41**, no. 7, 909–996, http://dx.doi.org/10.1002/cpa.3160410705.

de Hoop, M. V., L. Qiu, and O. Scherzer, 2012, Local analysis of inverse problems: Hölder stability and iterative reconstruction: Inverse Problems, **28**, no. 4, 045001, http://dx.doi.org/10.1088/0266-5611/28/4/045001.

Festa, G., and S. Nielsen, 2003, PML absorbing boundaries: Bulletin of the Seismological Society of America, **93**, no. 2, 891–903, http://dx.doi.org/10.1785/0120020098.

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

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–663, http://dx.doi.org/10.1190/1.1443081.

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

Nolet, G., 1987, Seismic tomography: Reidel.

Nolet, G., J. van Trier, and R. Huisman, 1986, A formalism for nonlinear inversion of seismic surface waves: Geophysical Research Letters, **13**, no. 1, 26–29, http://dx.doi.org/10.1029/GL013i001p00026.

Solano, C. A. P., 2013, Two-dimensional near-surface seismic imaging with surface waves: alternative methodology for waveform inversion: Ph.D. thesis, Mines ParisTech, Fontainebleau.

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Stockwell, J. W. Jr., 1999, The CWP/SU: Seismic Un*x package: Computers & Geosciences, **25**, no. 4, 415–419, http://dx.doi.org/10.1016/S0098-3004(98)00145-9.

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

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903, http://dx.doi.org/10.1190/1.1442046.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, no. 1, 195–216, http://dx.doi.org/10.1111/j.1365-246X.2004.02453.x.

Versteeg, R. J., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, **58**, 873–882, http://dx.doi.org/10.1190/1.1443471.

Yuan, Y. O., and F. J. Simons, 2014, Multiscale adjoint waveform-difference tomography using wavelets: Geophysics, **79**, no. 3, WA1–WA17, http://dx.doi.org/10.1190/geo2013-0383.1.

Downloaded 09/12/14 to 128.112.66.66. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

© 2014 SEG

DOI http://dx.doi.org/10.1190/segam2014-0816.1