Full Waveform Inversion for Seismic Velocity and Anelastic Losses in Heterogeneous Structures

4 authors, including:

Aysegul Askan

Middle East Technical University

31 PUBLICATIONS **110** CITATIONS

SEE PROFILE

Volkan Akcelik

Carnegie Mellon University

36 PUBLICATIONS **472** CITATIONS

SEE PROFILE

Jacobo Bielak

Carnegie Mellon University

189 PUBLICATIONS **2,535** CITATIONS

SEE PROFILE

Full Waveform Inversion for Seismic Velocity and Anelastic Losses in Heterogeneous Structures

by Aysegul Askan[*], Volkan Akcelik, Jacobo Bielak, and Omar Ghattas

Abstract   We present a least-squares optimization method for solving the nonlinear full waveform inverse problem of determining the crustal velocity and intrinsic attenuation properties of sedimentary valleys in earthquake-prone regions. Given a known earthquake source and a set of seismograms generated by the source, the inverse problem is to reconstruct the anelastic properties of a heterogeneous medium with possibly discontinuous wave velocities. The inverse problem is formulated as a constrained optimization problem, where the constraints are the partial and ordinary differential equations governing the anelastic wave propagation from the source to the receivers in the time domain. This leads to a variational formulation in terms of the material model plus the state variables and their adjoints. We employ a wave propagation model in which the intrinsic energy-dissipating nature of the soil medium is modeled by a set of standard linear solids. The least-squares optimization approach to inverse wave propagation presents the well-known difficulties of ill posedness and multiple minima. To overcome ill posedness, we include a total variation regularization functional in the objective function, which annihilates highly oscillatory material property components while preserving discontinuities in the medium. To treat multiple minima, we use a multilevel algorithm that solves a sequence of subproblems on increasingly finer grids with increasingly higher frequency source components to remain within the basin of attraction of the global minimum. We illustrate the methodology with high-resolution inversions for two-dimensional sedimentary models of the San Fernando Valley, under *SH*-wave excitation. We perform inversions for both the seismic velocity and the intrinsic attenuation using synthetic waveforms at the observer locations as pseudoobserved data.

Introduction

Seismic waveform inversion has become a major topic of interest in earthquake seismology and in geotechnical engineering due to the need for developing accurate earth models and for gaining a better understanding of subsurface structures for engineering applications. In its general form, given observed seismological data, seismic waveform inversion aims to recover the source parameters or the material properties of the crustal model, or both.

Modern inversion methodologies in seismology date back to the pioneering work of Backus and Gilbert (1967, 1970) on the determination of global structure from free oscillations of the Earth. In engineering, the emphasis has been on near-surface characterization, starting in the 1950s and 1960s with simple experimental and interpretive methods (Jones, 1958). Later, Stokoe and coworkers developed the spectral analysis of surface waves (SASW) method for ima-

ging shear-wave velocity, an inversion-based methodology for flat-layered deposits that makes use of the dispersion curve for the fundamental Rayleigh mode (Stokoe *et al.*, 1994). This work has been developed further by others (Gucunski and Woods, 1992; Rix *et al.*, 2000; Guzina and Lu, 2002). In earthquake seismology, the interest varies from whole Earth inversion (Jordan and Anderson, 1974; Lerner-Lam and Jordan, 1983; Ishii *et al.*, 2002) to imaging in local basins, ranging from determining just the boundary shape of the basin structure (Aoi, 2002) to tomography images of the Earth's crust in seismic-prone regions (Chen, 2005; Tromp *et al.*, 2005; Zhao *et al.*, 2005). Recently, Chen, Zhao, and Jordan (2007) performed full three-dimensional tomography for the crustal structure of the Los Angeles region using a real data set. In a subsequent article, Chen, Jordan, and Zhao (2007) compared the computational efficiency of two different formulations of the structural inverse problem: the scattering-integral and adjoint-wave-field methods. Lately, Tape *et al.* (2007) presented a finite frequency tomography technique based on adjoint methods.

[*]Present Address: Department of Civil Engineering, Middle East Technical University, Ankara, 06531, Turkey.

In exploration geophysics, seismic imaging has been in use for over 20 yr, where inversion proceeds mainly from travel-time (Baig *et al.*, 2003; Sheng and Schuster, 2003; Montelli *et al.*, 2004) or full waveform (Pratt, 1999) approaches, either in the time domain (Tarantola, 1984) or in the frequency domain (Pratt and Worthington, 1990; Liao and McMechan, 1996).

With few exceptions (Tarantola, 1988; Rix *et al.*, 2000; Hicks and Pratt, 2001), seismic inversion is generally confined to the seismic velocities, using models for which the intrinsic attenuation is not taken into consideration. This approach relies on data sets that consist primarily of measurements made on direct body waves that are relatively insensitive to the anelastic attenuation of the material (Chen, 2005). On the other hand, anelastic attenuation can significantly affect earthquake ground motion in many cases, especially if surface wave effects are important relative to the body waves. For instance, Olsen *et al.* (2003) found it necessary to assign values of Q as low as 10 in the near-surface low-velocity sediments in the Los Angeles basin in order for synthetic seismograms of their simulation of the 1994 Northridge earthquake to match observations. This underscores the importance of characterizing the intrinsic attenuation of the crust, especially of the low-velocity surficial sediments, in addition to the seismic velocities.

In this article, we present a nonlinear least-squares adjoint full waveform inversion method that is capable of estimating discontinuous distributions of shear-wave velocity and intrinsic attenuation in large heterogeneous earthquake-excited basins. For simplicity, we present the problem in the context of two-dimensional sedimentary valleys subjected to *SH*-wave excitation and provide an illustration for the San Fernando Valley. The components of the method include a full-anelastic waveform-based output least-squares objective, a total variation regularization, a finite-element parametrization of the anelastic moduli (as well as displacement and adjoint displacement fields), a multilevel grid/frequency continuation, an adjoint-based gradient and Hessian computation, an inexact matrix-free (Gauss)–Newton–Krylov minimization, a limited memory quasi-Newton preconditioning, a backtracking line search globalization, and a distributed memory parallel implementation.

Anelasticity and Memory Variables

In order to represent the anelastic losses in the crust, in this section we first consider a mechanical damping model for uniform, massless, linear anelastic fibers and review a common approach for reducing the stress-strain relationship to differential form through the introduction of internal auxiliary variables. We then discuss a procedure for expressing the parameters that enter into the governing differential equations for the auxiliary, or memory, variables in terms of the quality factor, Q. In the next section we discuss how to select Q such that it and the wave velocity will represent approximately the nonlinear behavior of the constitutive material

under strong-motion excitation. The application to a continuum is presented subsequently.

Suppose a mass m is attached to one end of a uniform, massless, linear anelastic fiber of unstretched length L and is subjected to an axial force $f(t)$. The fiber is fixed at the other end, has a cross-sectional area A, has uniform axial stress $\sigma(t)$ throughout its length and cross section, and has axial strain $\epsilon(t) = u(t)/L$, where $u(t)$ is the (small) displacement of the mass, at each instant t. The equation of motion of the mass then is

$$mL\frac{d^2\epsilon}{dt^2} = -\sigma A + f. \tag{1}$$

The simplest constitutive model for the fiber is Hooke's law, $\sigma = E\epsilon$, but this is not dissipative, that is, it does not have an internal damping (or attenuation) mechanism. Thus, we consider dissipative models of the following form. Let ϵ^t denote the strain history at time t, $\epsilon^t(\tau) = \epsilon(t - \tau)$. We then assume there is a linear functional $\mathcal{F}$ such that

$$\sigma(t) = (\epsilon^t). \tag{2}$$

Biot (1958) proposed a simple mechanism made up of N Maxwell elements with a single spring element, all in parallel, as shown in Figure 1a to represent this functional. He also suggested a suitable singular distribution factor for the springs in the Maxwell elements as $N \to \infty$. The Biot model can be equivalently represented as a set of standard linear solids (SLSs) in parallel, as shown in Figure 1b. Caughey (1962) showed that, if $N \to \infty$, the resulting hysteresis loss per cycle under sinusoidal steady-state excitation then takes the form of an inverse tangent function of ω/α, in which ω is the angular excitation frequency and $\alpha = \varphi/c$ is the angular relaxation frequency. This is tantamount to having a quality factor, Q, that becomes nearly independent of frequency for even small ω, provided α is small.

In the limit, as $N \to \infty$, the functional $\mathcal{F}$ for the Biot model can be expressed as that for a viscoelastic model with a singular kernel (Bielak and MacCamy, 1989):

$$\mathcal{F}(\epsilon^t) = k\epsilon(t) + \int_0^\infty g(\tau)[\epsilon(t) - \epsilon(t - \tau)]d\tau, \tag{3}$$

in which $k = \sum_{i=1}^N k_i$ is the effective constant of the single spring element in Biot's model in Figure 1b and $g(\tau) = \tau^{-1}e^{-\alpha\tau}$.

From (1), (2), and (3) it is clear, due to the correspondence principle of viscoelasticity, that in order to solve for ϵ in the frequency domain it suffices to consider an elastic material for which the effective elastic properties are frequency dependent. In the time domain, the problem is more challenging, as the stress at a given instant depends on the complete history of the strain, and (1) becomes an integrodifferential equation. A number of approaches have been proposed for incorporating viscoelastic constitutive relations such as (3) into the time-domain solution framework, without having to evaluate the convolution integral. Originally, Day and
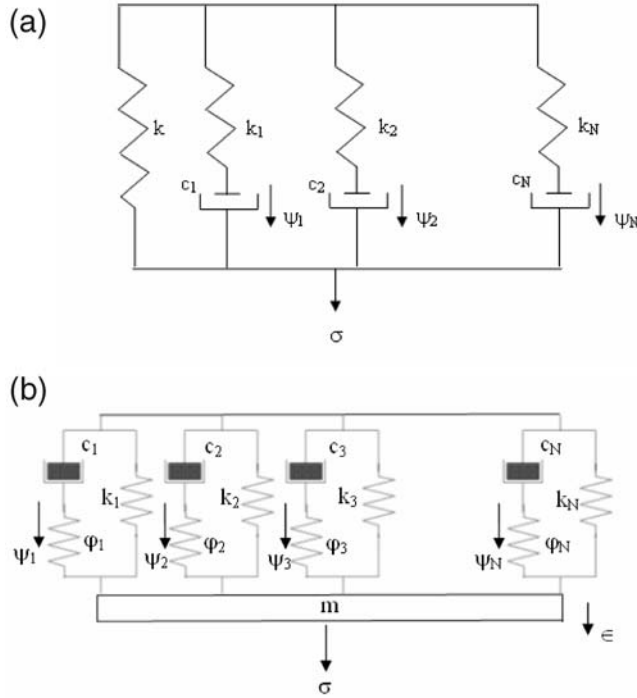
Figure 1.   (a) Biot's linear hysteretic damping model; (b) SLS rheological model.

Minster (1984) proposed expanding the Laplace-transformed viscoelastic modulus into a rational function using Padé approximations, thus giving rise to a set of ordinary differential equations (ODEs) in terms of internal memory variables, in addition to the original state variables. Similar results have been obtained by Emmerich and Korn (1987) by forming a rational function based on a rheological model of the Maxwell body. Liu *et al.* (1976) used relaxation mechanisms to model the viscoelastic stress-strain relationship, while Carcione *et al.* (1988) introduced memory variables as solutions to the ODEs that represent the motion of the relaxation mechanism; Robertsson *et al.* (1994) and Blanch *et al.* (1995) used SLSs as relaxation mechanisms. Moczo *et al.* (2006) have recently shown the equivalence of most of these approaches.

We consider directly the Biot SLS rheological model shown in Figure 1b. Let ϵ be the common strain experienced by all of the springs with stiffness k_j, and let ψ_j be the strain of each spring with stiffness φ_j; c_j is the dashpot constant of the jth damping element. With this notation, the stress σ can be expressed as

$$\sigma = \left(\sum_{j=1}^{N} k_j \right) \epsilon + \sum_{j=1}^{N} \varphi_j \psi_j. \tag{4}$$

In addition, for the stress to be the same in the spring and the dashpot of each Maxwell element, ϵ and ψ_j are related by

$$\frac{\partial \psi_j}{\partial t} + \alpha_j \psi_j = \epsilon, \qquad j = 1, 2, \dots, N, \tag{5}$$

where $\alpha_j = \varphi_j / c_j$ is the angular relaxation frequency of the jth SLS, with corresponding relaxation time $2\pi/\alpha_j$.

To eliminate ψ_j, we take Fourier transforms of both (4) and (5), solve for $\bar{\psi}_j$ from (5), and substitute into the Fourier-transformed equation of (4) to obtain

$$\bar{\sigma}(\omega) = \bar{F}(\omega) \bar{\epsilon}(\omega), \tag{6}$$

where

$$\bar{F}(\omega) = \sum_{j=1}^{N} k_j \left\{ \frac{\eta_j \alpha_j \omega i + \alpha_j^2 + \omega^2 (1 + \eta_j)}{\alpha_j^2 + \omega^2} \right\}. \tag{7}$$

In these equations, an overbar denotes Fourier transform, and $\eta_j = \varphi_j / k_j$ is the relative stiffness of the two springs in each SLS.

We use (7) and the definition of Q^{-1}, in terms of the complex relaxation modulus

$$Q^{-1}(\omega) = \frac{\mathrm{Im}\,[\bar{F}(\omega)]}{\mathrm{Re}\,[\bar{F}(\omega)]}, \tag{8}$$

to obtain a formula for $Q^{-1}(\omega)$ in terms of the parameters η_j and α_j:

$$Q^{-1}(\omega) = \frac{\sum_{j=1}^{N} \eta_j \alpha_j \omega / (\alpha_j^2 + \omega^2)}{1 + \sum_{j=1}^{N} \eta_j \omega^2 / (\alpha_j^2 + \omega^2)}. \tag{9}$$

For future use, it is convenient to relate the mechanical properties α_j and η_j of the rheological model to the quality factor Q through approximate, simple, predictive frequency-independent relationships. To this effect, we minimize the integral of the difference squared between a target Q^{-1} and the value computed from (9) with respect to α_j and η_j as follows:

$$\min_{\alpha_j, \eta_j} \frac{1}{2} \int_{\omega_{\min}}^{\omega_{\max}} [Q^{-1}(\omega, \alpha_j, \eta_j) - Q_{\text{target}}^{-1}]^2 d\omega.$$

The minimization is performed over a frequency range of interest, from 0.1 to 1.0 Hz. The goal is to find α_j and η_j that correspond to a frequency-independent Q^{-1}. For a single SLS and a Q_{target}^{-1} of 0.1, we found $Q^{-1}(\omega)$ as shown in Figure 2. This figure also shows the corresponding results of the least-squares minimization for two SLSs. Clearly, the addition of a second SLS results in a significant improvement in approximating a frequency-independent Q^{-1}. However, an additional SLS requires the introduction of a second memory function ψ_2. Each memory function requires additional storage of half that required for the displacements, using standard second-order accurate time discretization. This is because the differential equation (1) that governs the evolution of ϵ is second order in time, whereas the corresponding equations (5) for ψ_j are only first order. For a second-order approximation in time, using central differences, (1) requires
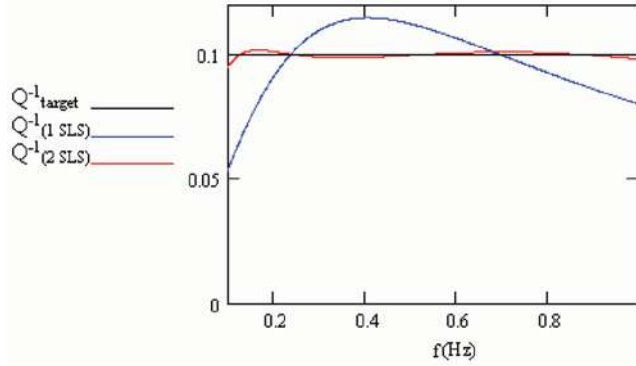
Figure 2. Q^{-1} for one and two SLS mechanisms versus the corresponding Q^{-1}_{target}.

that two values of ϵ, at t and $t - \Delta t$, be stored in order to calculate ϵ at $t + \Delta t$. By contrast, using the step-by-step numerical solution procedure in Day (1998), only the value of ψ_j at t need be stored to obtain its value at $t + \Delta t$. This means that using two memory variables in our forward wave propagation problem would double the required computer memory with respect to that for the corresponding problem without dissipation. On the other hand, the additional computer time needed to handle the memory variables is small, as no spatial derivatives are involved in the calculation of ψ_j, because the governing equations (5) for these variables are uncoupled spatially. In order to limit the number of memory variables per state variable to only one per node, Day (1998), Day and Bradley (2001), and Graves and Day (2003) have devised a highly efficient coarse graining method without sacrificing accuracy in representing Q^{-1}_{target}.

Here, our main objective is to present an optimization method for inverting for both the seismic velocity and Q^{-1}. To simplify the presentation and the calculations, we will consider a single SLS. Further, we will seek simple expressions for approximating η and α as functions of Q^{-1}. For this purpose, we minimize the least-squares difference between $Q^{-1}(\omega, \alpha, \eta)$ and Q^{-1}_{target}, from 0.1 to 1.0 Hz and for values of Q^{-1}_{target} ranging from 0.005 to 0.4. Through a curve fit based on the results of these minimizations, we obtained approximate linear relationships as follows:

$$\eta(Q^{-1}) = 3.447Q^{-1} - 0.041, \tag{10a}$$

$$\alpha(Q^{-1}) = 3.529Q^{-1} + 2.543. \tag{10b}$$

We will use these values of η and α in our applications. A more accurate approximation for η and α could be found by introducing higher degree polynomials in Q^{-1}. Also, adding a second or more memory variables for increased accuracy introduces no additional conceptual complexity, as the only independent dissipation variable in our model is Q^{-1}.

A Model for Relating the Quality Factor to the Shear-Wave Velocity

In general, there is no explicit physical relationship between Q and the shear-wave velocity V_S. In practice, however, one expects them to be strongly correlated. Thus, as a first step in our study, we will restrict Q to be a prescribed function of V_S, as done by Olsen *et al.* (2003).

Anelastic attenuation, here represented by Q, is relatively difficult to estimate in soil media, particularly in soft sedimentary layers for which ground motion is strongly affected by the amount of damping. Common techniques within current engineering practice for recovering the shear-wave velocity and damping ratio profiles include surface wave tests (e.g., Rix *et al.*, 2000; Lai *et al.*, 2002). These techniques use experimental phase velocity dispersion curves and experimental attenuation curves for inversions of the velocity and damping ratio profiles. In addition, there are extensive borehole data that can provide information on the variations of Q related to the variations in the soil properties. In particular, there are studies relating the quality factor to shear-wave velocity (Wiggens *et al.*, 1978). However, most existing studies are based on frequencies higher than our frequency range of interest.

In the present study, we wish to use a simple yet physically validated relationship between the wave velocity and Q. It is well known that shear modulus and material damping are nonlinear functions of strain (Seed and Idriss, 1970). At very small strain levels, shear-wave velocity V_S is directly related to the small-strain shear modulus μ_{max} through

$$\mu_{\text{max}} = \rho V_S^2, \tag{11}$$

where ρ is the mass density of soil.

At moderate to higher strain levels, the secant shear modulus μ is used frequently in engineering to represent the average soil stiffness. The ratio μ/μ_{max} is called the dynamic modulus reduction ratio. On the other hand, the damping ratio, ζ (roughly, $2Q^{-1}$ at resonance) at extremely low strain levels is a constant value and is referred to as the small-strain damping ratio, ζ_{min}. At higher strain levels, due to the nonlinearity in the stress-strain relationship, the damping ratio increases with increasing strain amplitude.

Many empirical studies have been conducted to characterize the factors that influence the ratio μ/μ_{max} and ζ of soils (Richart *et al.*, 1970; Seed and Idriss, 1970; Vucetic and Dobry, 1991; Ishibashi and Zhang, 1993; Vucetic *et al.*, 1998; Stokoe *et al.*, 2004). The most important factors that influence the dynamic behavior of soil properties are found to be the shear strain γ, the effective confining stress σ', and the plasticity index *PI*. There are several databases in North America providing test data on these variables. Based on statistical analyses of these existing databases, Zhang *et al.* (2005) derived predictive equations for the shear modulus reduction ratio and the damping ratio. Although the databases used in the analyses are from South Carolina, North Carolina, and Alabama soils, we concluded that the derived

relationships are valid for regions with various similar soft soil conditions.

Following the hyperbolic models suggested earlier (e.g., Hardin and Drnevich, 1972; Pyke, 1993; Stokoe *et al.*, 1999) for the shear modulus reduction ratio and the strain level, Zhang *et al.* (2005) expressed μ/μ_{max} as

$$\frac{\mu}{\mu_{max}} = \frac{1}{[1 + (\gamma/\gamma_r)^\kappa]}, \tag{12}$$

where γ_r is a reference strain defined as the value of strain where the μ/μ_{max} ratio ceases to be flat or, as defined by other authors, $\gamma_r = \tau_{max}/G_{max}$. Here, κ is a curve-fitting parameter that depends on the plasticity index, PI, of the soil as

$$\kappa = 0.0021PI + 0.834. \tag{13}$$

A common approach for modeling the damping ratio is to relate it to the strain level. Zhang *et al.* (2005) expressed the damping ratio ζ in percent in terms of a polynomial function of the modulus reduction ratio μ/μ_{max} and a minimum damping ratio ζ_{min} as follows:

$$\zeta - \zeta_{min} = 10.6(\mu/\mu_{max})^2 - 31.6(\mu/\mu_{max}) + 21.0. \tag{14}$$

Equation (14) is different from those of previous quadratic models (Ishibashi and Zhang, 1993) through the addition of ζ_{min}, which makes the damping measure $\zeta - \zeta_{min}$ independent of cyclic frequency. In addition, (14) suggests that ζ is equal to ζ_{min} when $\mu = \mu_{max}$ at small strains. When the value of μ/μ_{max} is close to zero at very large strains, the maximum predicted $\zeta - \zeta_{min}$ value is 21%.

Using the preceding model, we develop a simple procedure for relating Q to V_S as follows: To derive a physical relationship between the Q and shear-wave velocity relationship, we apply (14) in our forward wave propagation simulations using a cross section of the San Fernando Valley and a prescribed kinematic source, to be described in a later section. We start with a target shear modulus profile defined by μ_{max} from the Southern California Earthquake Center (SCEC) Los Angeles basin velocity model (Magistrale *et al.*, 2000) and $\zeta_{min} = 1\%$. We then solve the forward wave propagation problem and record the corresponding strain values. We define a nominal strain at each point in space as 60% of the maximum value of the entire shear strain history at that node (Kramer, 1996). Depending on the nominal strain levels, we reduce the shear modulus through (12) with $\gamma_r = 2 \times 10^{-5}$ and a low plasticity level, $PI = 10\%$. Using (14), we obtain a new ζ and repeat the cycle of forward runs until the strain levels remain unchanged.

In order to establish a relationship between Q and V_S, we plot Q versus V_S at a large number of points within the two-dimensional sedimentary valley, for the final iteration of shear modulus reduction cycles, as shown in Figure 3, and

obtain a curve of the following form to fit the data:

$$Q = a \arctan\left(\frac{V_S}{V_{S_{ref}}}\right). \tag{15}$$

Here $V_{S_{ref}}$ is the shear-wave velocity value where the Q versus V_S relationship ceases to be linear, $V_{S_{ref}} = 1000$ m/sec in the present model, and we find $a = 25.5$ with a least-squares best fit. The plot of Q from (15) for these values of a and $V_{S_{ref}}$ is also shown in Figure 3. This is the model that we will use for our forward and inverse solutions.

Forward Viscoelastic Wave Propagation Problem

As discussed earlier, we will describe the methodology for the viscoelastic seismic wave propagation problem in two dimensions only with a single generalized SLS to represent the anelastic behavior of the constitutive material. The extension to multiple SLSs is immediate, along the lines of (4) and (5). Because we wish to invert for material properties within a portion of the Earth, we require appropriately positioned absorbing boundaries to account for the truncated exterior. For simplicity, we use dashpot absorbing boundaries applied on the truncation surface. Thus, the problem is defined over a finite domain Ω with boundary Γ; Γ consists of a horizontal free surface Γ_{FS} and the truncation boundary Γ_{AB} on which a simple absorbing boundary condition is imposed to limit the occurrence of spurious reflections from the outgoing waves. The system is initially at rest.

The corresponding governing partial differential equations (PDEs) and ODEs, boundary conditions, and initial conditions are given by

$$\rho\frac{\partial^2 u}{\partial t^2} - \nabla \cdot [\mu\nabla(u + \eta v)] = f(x, t) \quad \text{in } \Omega \times [0, T], \tag{16a}$$

$$\frac{\partial v}{\partial t} + \alpha v = \frac{\partial u}{\partial t} \quad \text{in } \Omega \times [0, T], \tag{16b}$$

$$\mu\nabla(u + \eta v) \cdot n = 0 \quad \text{on } \Gamma_{FS} \times [0, T], \tag{16c}$$

$$\mu\nabla(u + \eta v) \cdot n = \sqrt{\rho\mu}\frac{\partial u}{\partial t} \quad \text{on } \Gamma_{AB} \times [0, T], \tag{16d}$$

$$u = \frac{\partial u}{\partial t} = 0, \quad v = 0 \quad \text{in } \Omega \quad \text{at } t = 0. \tag{16e}$$

In these equations, $u(x, t)$ is the antiplane displacement field, $v(x, t)$ is the corresponding memory variable, $\rho(x)$ is the density, $\mu(x)$ is the elastic shear modulus, $\alpha(x)$ is the relaxation frequency, $\eta(x)$ is the spring constant of the rheological model, $f(x, t)$ is the body force vector representing the earthquake source, and $[0, T]$ is the time interval of interest. Equations (16a) and (16b) are the governing equations of the viscoelastic model; (16c) enforces the traction-free condition on the free surface; (16d) is the absorbing boundary condition to limit spurious wave reflections from the truncated boundary surfaces; and (16e) is the initial condition.
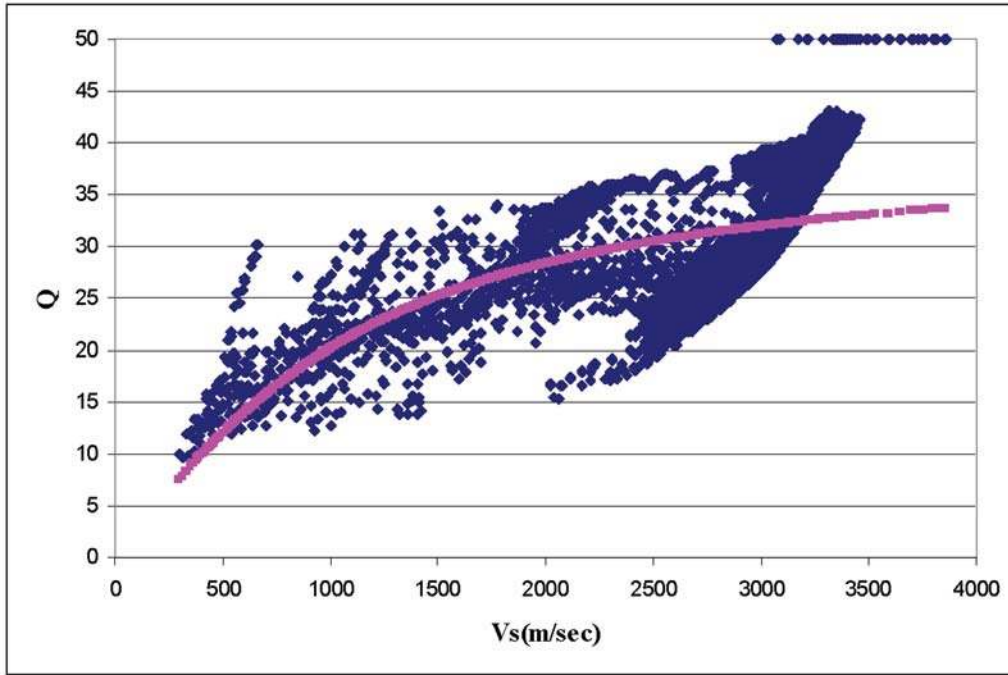
Figure 3. Q versus V_S relationship data from the last iteration (the pink curve is the best-fit curve given in equation 15).

Inverse Viscoelastic Wave Propagation Problem

The inverse problem we consider is to determine the distribution of the seismic shear-wave velocity V_S and the quality factor Q that characterize an earth model within a region of interest, based on the observed displacement seismograms $u^*(t)$ at a finite number of receivers corresponding to the points x_j in the domain. We consider a single event, for which the source body force f is assumed to be known. In addition, the density of the material is taken as known, because its variability is typically much less pronounced than that of the other system parameters. The parameters α and η are expressed in terms of Q^{-1} through (10). From now on we refer to $\mu(x)$ and $Q^{-1}(x)$ as the material model and to $u(x, t)$ and $v(x, t)$ as the state variables.

We formulate the inverse problem as a nonlinear least-squares parameter estimation problem with PDE and ODE constraints. We seek to find the material model that minimizes, over the period $t = 0$ to T, the L^2-norm difference between the observed state and that predicted by the coupled PDE–ODE model of viscoelastic antiplane wave propagation at N_R receiver locations. The optimization problem is

$$\min_{u,v,\mu} \frac{1}{2} \sum_{j=1}^{N_R} \int_0^T \int_\Omega [u^* - u]^2 \delta(x - x_j) d\Omega dt$$

$$+ \beta \int_\Omega (\nabla \mu \cdot \nabla \mu + \varepsilon)^{1/2} d\Omega, \qquad (17)$$

subject to (16).

The first term in the objective function is the misfit between observed and predicted states; the second term represents a regularization functional on the material field μ with regularization parameter β; and the constraints are the equations corresponding to our forward initial-boundary value problem for anelastic two-dimensional wave propagation.

In the most general case, the solution for the optimization problem can be obtained by searching for the optimum values of the state variables (u, v) and the inversion variables (μ, Q^{-1}) that satisfy the first and second optimality conditions (Nocedal and Wright, 1999). We will initially suppose that Q^{-1} is related to V_S through (15), with a and $V_{S_{\text{ref}}}$ assuming the values $a = 25.5$ and $V_{S_{\text{ref}}} = 1000$ m/sec obtained earlier. The inverse problem is then reduced to finding the single material field μ. For this reason, the objective functional (17) includes a regularization term for only μ.

In the absence of regularization, the problem as formulated in the preceding discussion is ill posed (Symes, 1990). Discretization of the inverse problem leads to Hessian matrices that are rank deficient and ill conditioned. Rank deficiency and ill conditioning occur because of the insensitivity of the objective functional to high-frequency material property perturbations (which is due to the band-limited observations). We treat this rank deficiency with a regularization function added to the least-squares waveform misfit function. Several regularization functionals are commonly used (Vogel, 2002). The well-known Tikhonov regularization, which employs the L^2 norm of the gradient of the material model, eliminates the null space of the Hessian operator by penalizing oscillatory components of the inversion field μ.

The regularization effect becomes more pronounced with larger values of the regularization parameter β. A successful choice of β is essential because values that are too large oversmooth the μ field while values that are too small leave the Hessian matrix rank deficient. Tikhonov regularization smooths the discontinuities in the material model μ and is therefore not appropriate for material fields with sharp interfaces and other discontinuities. Consequently, in this work, we use total variation (TV) regularization, which is the L^1 norm of the gradient of the material model. It is particularly effective in recovering the material interfaces because it smooths only along the interfaces and not normal to them. The addition of the small parameter ε in (17) makes the TV regularization functional differentiable when $\nabla\mu = 0$, and effectively adds a small amount of smoothing across the interfaces.

We note that the waveform inversion problem presented here is inherently nonlinear due to the strong nonlinear dependence of waveform to wave velocity, and this nonlinearity becomes worse as contrast in wave velocity increases. In this work, we use regularization as a form of model conditioning to reduce the nonlinearity. However, in the case of inversions using actual observations, data conditioning in addition to model regularization could be required to achieve convergence.

To derive the optimality conditions for the optimization problem (17), we introduce the Lagrangian functional in what follows. Our procedure follows closely that of Akcelik *et al.* (2002, 2003). The Lagrangian functional $\mathcal{L}$ combines the original regularized least-squares objective functional with an inner product of the residual of the governing equations with Lagrange multipliers λ and ϕ (also known as adjoint or dual variables) to incorporate the PDE–ODE constraints, that is,

$$
\mathcal{L}(u, v, \mu, \lambda, \phi)
$$

$$
= \frac{1}{2} \sum_{j=1}^{N_R} \int_0^T \int_\Omega [u^* - u]^2 \delta(x - x_j) d\Omega dt
$$

$$
+ \beta \int_\Omega (\nabla\mu \cdot \nabla\mu + \varepsilon)^{1/2} d\Omega
$$

$$
+ \int_0^T \int_\Omega \lambda \left\{ \rho \frac{\partial^2 u}{\partial t^2} - \nabla \cdot [\mu\nabla(u + \eta v)] - f \right\} d\Omega dt
$$

$$
+ \int_0^T \int_{\Gamma_{FS}} \lambda \{ \mu\nabla(u + \eta v) \cdot n \} d\Gamma_{FS} dt
$$

$$
+ \int_0^T \int_{\Gamma_{AB}} \lambda \left\{ \mu\nabla(u + \eta v) \cdot n - \sqrt{\rho\mu} \frac{\partial u}{\partial t} \right\} d\Gamma_{AB} dt
$$

$$
+ \int_0^T \int_\Omega \phi \left\{ \frac{\partial v}{\partial t} + \alpha v - \frac{\partial u}{\partial t} \right\} d\Omega dt. \tag{18}
$$

Here, λ and ϕ are adjoint variables for the PDE and ODE constraints, respectively.

The first order necessary condition for optimality states that the first variation of the Lagrangian with respect to u, v, μ, λ, and ϕ should vanish at the optimum:

$$
\left\{ \begin{array}{c} \delta_u \\ \delta_v \\ \delta_\mu \\ \delta_\lambda \\ \delta_\phi \end{array} \right\} (u, v, \mu, \lambda, \phi) = 0. \tag{19}
$$

This set of equations is known as the Karush–Kuhn–Tucker (KKT) conditions. We now give the strong form of the KKT conditions. Variations of the Lagrangian with respect to λ and ϕ yield the following two sets of equations, which are referred to as state problems for u and v, respectively:

$$
\rho \frac{\partial^2 u}{\partial t^2} - \nabla \cdot \mu\nabla(u + \eta v) = f \quad \text{in } \Omega \times [0, T],
$$

$$
\mu\nabla(u + \eta v) \cdot n = 0 \quad \text{in } \Gamma_{FS} \times [0, T],
$$

$$
\mu\nabla(u + \eta v) \cdot n = \sqrt{\rho\mu} \frac{\partial u}{\partial t} \quad \text{in } \Gamma_{AB} \times [0, T],
$$

$$
u = \frac{\partial u}{\partial t} = 0 \quad \text{in } \Omega \quad \text{at } t = 0, \tag{20}
$$

and

$$
\frac{\partial v}{\partial t} + \alpha v - \frac{\partial u}{\partial t} = 0 \quad \text{in } \Omega \times [0, T],
$$

$$
v = 0 \quad \text{in } \Omega \quad \text{at } t = 0. \tag{21}
$$

Thus, we recover the original constraint equations and boundary conditions for the state variables. Next, taking variations of the Lagrangian with respect to the state variables u and v leads to adjoint problems for λ and ϕ, respectively:

$$
\rho \frac{\partial^2 \lambda}{\partial t^2} - \nabla \cdot \mu\nabla\lambda = -\sum_{j=1}^{N_R} (u^* - u)\delta(x - x_j) \quad \text{in } \Omega \times [0, T],
$$

$$
\mu\nabla\lambda \cdot n = 0 \quad \text{on } \Gamma_{FS} \times [0, T],
$$

$$
\mu\nabla\lambda \cdot n = -\sqrt{\rho\mu} \frac{\partial \lambda}{\partial t} \quad \text{on } \Gamma_{AB} \times [0, T],
$$

$$
\lambda = \frac{\partial \lambda}{\partial t} = 0 \quad \text{in } \Omega \quad \text{at } t = T, \tag{22}
$$

and

$$
\frac{\partial \phi}{\partial t} - \alpha\phi = -\eta\nabla \cdot \mu\nabla\lambda \quad \text{in } \Omega \times [0, T],
$$

$$
\phi = 0 \quad \text{in } \Omega \quad \text{at } t = T. \tag{23}
$$

We note that the adjoint problem governing λ has a similar form to the state equation for u; however, it is a terminal value problem and the source function is the misfit between the observed displacements and the corresponding simulated values. There is a similar analogy between the state equation for the displacement v and the adjoint equation for ϕ. Equations 22 and 23 represent final value problems, in

which the functions λ, $\partial\lambda/\partial t$, and ϕ are known at $t = T$, and the problem is to determine the adjoint variables λ and ϕ in $[0, T]$. Notice that the terms $\sqrt{\rho\mu}(\partial\lambda/\partial t)$ and $\alpha\phi$ have opposite signs to their counterparts in (20) and (21). At first glance, it may appear that the solution for λ and ϕ could then be unstable. To verify that this is not the case, it suffices to introduce a change of variable $t' = t - T$, which renders problems (22) and (23) into initial value problems with exactly the same operations acting on λ and ϕ as those in (20) and (21) for u and v. Only the right-hand sides are different, and these do not affect the stability of the problem.

The final KKT condition is derived by taking the variation of the Lagrangian with respect to the shear modulus μ, yielding the following equation for the material field μ:

$$
-\beta\nabla\cdot\left[\frac{\nabla\mu}{(\nabla\mu\cdot\nabla\mu+\varepsilon)^{1/2}}\right] + \int_0^T \nabla\lambda\cdot\nabla[u+\eta(\mu)v]dt
$$
$$
-\int_0^T D_\mu\eta(\mu)v\nabla\cdot(\mu\nabla\lambda)dt
$$
$$
+\int_0^T \phi D_\phi\alpha(\mu)vdt = 0 \quad \text{in } \Omega,
$$
$$
\nabla\mu\cdot n = 0 \quad \text{on } \Gamma_{\text{FS}},
$$
$$
\frac{\nabla\mu}{|\nabla\mu|}\cdot n = \frac{1}{2}\int_0^T \sqrt{\frac{\rho}{\mu}}\lambda\frac{\partial u}{\partial t}dt \quad \text{on } \Gamma_{\text{AB}}, \tag{24}
$$

where the operators $D_\mu\alpha(\mu)$ and $D_\mu\eta(\mu)$ represent the first derivatives of the functionals $\alpha(\mu)$ and $\eta(\mu)$, respectively, with respect to μ.

Because the KKT system is a coupled nonlinear system of equations, it must be solved with an iterative method. We next discuss issues related to the design of the optimization algorithm.

Optimization Method

A discussion of the issues faced in solving large-scale optimization problems that are governed by PDEs can be found in Akcelik *et al.* (2006). Here, we give a brief discussion of the methods we use to solve the optimality system (20) to (24). There are at least two main choices we face in designing a suitable algorithm: (a) the solution technique for computing the search direction and (b) the method for finding the step length. In making these decisions, the trade-off between resolution and computing time is of the essence.

At each iteration, given an estimate of the solution, the search direction can be computed using methods that utilize the derivatives of the objective function. The simplest of such gradient-based methods is the steepest descent (SD) method, which is based on a linear model of the objective function and uses the negative gradient as the search direction. This method is simple to implement and requires little work per iteration beyond computation of the gradient. But it often exhibits slow convergence even for mildly ill-conditioned problems. Another common approach for computing the search

direction is the Newton's method, which is based on a quadratic model of the objective function. This method is locally quadratically convergent and often takes just a few iterations once it is close to a minimum. We use variants of Newton's method to determine the search direction in this work. For solving the linear system of equations that arises at each Newton iteration, we use the conjugate gradient (CG) method as discussed in the next section.

Newton's method is locally convergent and therefore requires some form of globalization, such as a line search or trust region methods. Trust region methods define a region around the current estimate in which the model of the objective function is trusted. These methods estimate the step length and the search direction simultaneously by minimizing the model over the trust region. Line search methods first compute a direction of descent and then choose a step length in that direction so that the objective function is decreased. We use a backtracking line search technique in this work and choose the step length to satisfy the Armijo condition (Nocedal and Wright, 1999).

The Newton step for the solution of the KKT optimality system (20) to (24) is given by the following indefinite linear system of equations:

$$
\begin{bmatrix}
\delta_{uu}^2\mathcal{L} & \delta_{uv}^2\mathcal{L} & \delta_{u\mu}^2\mathcal{L} & \delta_{u\lambda}^2\mathcal{L} & \delta_{u\phi}^2\mathcal{L} \\
\delta_{vu}^2\mathcal{L} & \delta_{vv}^2\mathcal{L} & \delta_{v\mu}^2\mathcal{L} & \delta_{v\lambda}^2\mathcal{L} & \delta_{v\phi}^2\mathcal{L} \\
\delta_{\mu u}^2\mathcal{L} & \delta_{\mu v}^2\mathcal{L} & \delta_{\mu\mu}^2\mathcal{L} & \delta_{\mu\lambda}^2\mathcal{L} & \delta_{\mu\phi}^2\mathcal{L} \\
\delta_{\lambda u}^2\mathcal{L} & \delta_{\lambda v}^2\mathcal{L} & \delta_{\lambda\mu}^2\mathcal{L} & 0 & 0 \\
\delta_{\phi u}^2\mathcal{L} & \delta_{\phi v}^2\mathcal{L} & \delta_{\phi\mu}^2\mathcal{L} & 0 & 0
\end{bmatrix}
\begin{Bmatrix}
\bar{u} \\ \bar{v} \\ \bar{\mu} \\ \bar{\lambda} \\ \bar{\phi}
\end{Bmatrix} =
$$
$$
-\begin{Bmatrix}
\delta_u\mathcal{L} \\ \delta_v\mathcal{L} \\ \delta_\mu\mathcal{L} \\ \delta_\lambda\mathcal{L} \\ \delta_\phi\mathcal{L}
\end{Bmatrix}. \tag{25}
$$

Here, the $\delta^2\mathcal{L}$ operator refers to the second variation of the Lagrangian with respect to the indicated variables, the overbars on the optimization variables indicate the Newton direction in each of these variables, and the right-hand side is the negative of the gradient of the Lagrangian. The coefficient matrix of this system is known as the KKT matrix.

In a full space approach, one solves the preceding system simultaneously for $\bar{u}$, $\bar{v}$, $\bar{\mu}$, $\bar{\lambda}$, and $\bar{\phi}$. This approach requires solution of a large indefinite linear system discretized over space-time, which is not feasible for a large-scale problem of the type we consider. On the other hand, the reduced space utilizes a block elimination to reduce the system to one that contains only the material field unknowns. Thus, it does not require storage of the time-dependent variables and associated entries of the space-time matrix in (25). In addition, the reduced Hessian is often better conditioned than the full Hessian and is positive definite near a minimum. Consequently, in this work we use the reduced space approach.

We first compute the Newton direction for the material field, $\bar{\mu}$, by solving the reduced version of the Newton equa-

tions (25). We do this as follows: At each Newton iteration, given a shear modulus estimate μ, we first solve for u and v through state equations (20) and (21). With u and v known, we solve for λ and ϕ through adjoint equations (22) and (23). Knowing $\bar{u}$, $\bar{v}$, $\bar{\lambda}$, and $\bar{\phi}$, we calculate the negative of the reduced gradient $g_\mu = -\delta_\mu \mathcal{L}$. Then, in order to obtain the reduced Newton system for $\bar{\mu}$, we initially solve for $\bar{u}$ and $\bar{v}$ from the last two rows of equation (25) (which are linearized state equations). Then from the first two rows, which are linearized adjoint equations, we solve for $\bar{\lambda}$ and $\bar{\phi}$. Finally, we arrive at the reduced Newton system for $\bar{\mu}$ corresponding to the middle row of the KKT matrix. This block elimination is equivalent to

$$W_\mu \bar{\mu} = -g_\mu, \qquad (26)$$

where W_μ is the Schur complement of $\delta^2_{\mu\mu}\mathcal{L}$ in the matrix of (25) and is known as the reduced Hessian. We note that the second-order optimality condition states that the reduced Hessian is at least positive semidefinite at the solution, and therefore the reduced Hessian is guaranteed to be positive only near a minimum. Because global convergence requires a direction of descent of each Newton iteration, and this can be achieved via a positive-definite approximation of the reduced Hessian, we choose to make a Gauss–Newton approximation, in which we ignore the terms that depend on adjoint variables λ and ϕ in the KKT matrix. It can be shown that this leads to a positive-definite reduced Hessian approximation.

Because of storage and computational limitations, we cannot factor the reduced Hessian matrix, or even construct it, which is of the order of the degrees of freedom of the shear modulus discretization. Instead, we solve the reduced Newton system for the search direction, $\bar{\mu}$, with the conjugate gradient method, which does not require the construction of the reduced Hessian matrix but only its application to vectors. Computing the matrix-vector product at each CG iteration requires solution of the state and adjoint problems (Askan, 2006).

The convergence of the CG algorithm is related to the square root of the condition number of the reduced Hessian. Clustering of the eigenvalue spectrum also leads to faster convergence.

To address possible ill conditioning of the reduced Hessian, we use a preconditioner in the form of a limited memory Broyden–Fletcher–Goldfarb–Shanno (BFGS) method, which is generated using information from the CG iterations (Morales and Nocedal, 2000). Using this preconditioner, we typically obtain a reduction in the total number of CG iterations of a factor of 8 relative to the unpreconditioned iterations.

Finally, we terminate the CG iterations early to avoid oversolving, in the style of an inexact Newton–CG method (Eisenstat and Walker, 1996).

The performance of the line search method depends on successful choices of the step length α as well as the search direction $\bar{\mu}_k$. Here, we use a backtracking line search, beginning with unit step length and decreasing it until the objective function is reduced sufficiently according to the Armijo condition (Nocedal and Wright, 1999).

The final difficulty to address is the presence of many local minima. It is well known for inverse wave propagation problems that multiple minima exist for the least-squares objective functional in the direction of the low wavenumber components of the material field, and these local minima have attraction basin diameters that shrink with increasing frequency of propagating waves (Symes, 1990). Using this observation, we employ a multilevel continuation technique, in which we minimize the objective functional over a sequence of increasingly finer discretizations of the material field. We begin by solving the inverse problem on a coarse material grid and prolong the solution to the next finer grid, continuing the refinement until the material grid matches the wave propagation grid. The converged solution on each grid is used as an initial guess for the next scale. In our experience, this multilevel grid continuation strategy keeps the sequence of minimizers within the basin of attraction of the global minimum, provided the seismic source contains sufficient low-frequency energy. This strategy is similar to that employed by Bunks et al. (1995).

We note that it is possible to use multilevel inversion for regularization purposes as well if one stops in the early stages of the multilevel algorithm. However, the quality of inverted result with this kind of regularization will not be as good as the TV regularization that we use in our formulation. Thus, we show in our results that the TV works effectively as a regularization method, due to the discontinuous (rather, piecewise constant) nature of wave velocity. The purpose of the multilevel algorithm here is only to lead the optimizer to the global minimum.

We have discretized the preceding continuous KKT system in space with a standard Galerkin-type finite-element scheme and a central differences scheme for temporal discretization. We have implemented the described inversion strategy using the parallel numerical library PETSC (Balay et al., 2001).

In the next section, we demonstrate the performance of our methodology through inversions for the shear-wave velocity in a two-dimensional viscoelastic medium.

Example: Synthetic Two-dimensional Viscoelastic Waveform Inversion

To illustrate the feasibility of the methodology described in the preceding sections for finely parameterized models, we apply the viscoelastic inversion algorithm to reconstruct the target two-dimensional shear velocity and attenuation profiles shown in Figure 4. The model under consideration comprises a portion of a vertical cross section of the Los Angeles basin. Figure 4a depicts the spatial distribution of the shear-wave velocity derived from the SCEC seismic velocity model (Magistrale et al., 2000), with an assigned minimum

threshold velocity of 715 m/sec. The vertical line in the model represents the trace of a causative strike-slip fault running perpendicular to the valley. The target velocity profile has a significant contrast of 4.9 between the smallest and largest seismic velocities, with a corresponding contrast in the shear modulus of 25.

In our numerical experiments, we use waveforms synthesized from the target profile on the free surface as pseudoobserved data. To model the source, we apply a uniform *SH* kinematic dislocation, with a triangular slip rate, with a rise time t_0 and total slip u_0, as shown in Figure 5. The model parameters, such as the overall dimensions, timestep, number of receivers, and regularization parameter, are listed in Table 1. We add 10% Gaussian random noise to the synthetic seismograms at the receivers to represent errors in the observations. The wave propagation simulations are performed on a 64×64 grid. The material model is initially taken to be homogeneous; that is, we assume nothing is known about the target. Thus, the optimization grid starts with a single finite element of constant value and becomes uniformly finer at each stage of the multilevel process until it matches the wave propagation grid.

Figure 4b shows the target distribution of Q throughout the profile. These values, as well as those for the seismic velocity in Figure 4a, were obtained iteratively using the approximate procedure described earlier for taking into consideration the nonlinear material behavior, starting with the lossless SCEC velocity model as the initial guess.

Single-Variable Inversion for Shear-Wave Velocity

We initially perform only single-variable inversions, that is, for the material field μ, using the relationship (15) for the intrinsic attenuation. Figure 6 shows a sequence of material fields found by solving the inverse problem on a sequence of increasingly finer material grids. The shear modulus μ is approximated as a piecewise bilinear function on each regular grid, as are the state fields u and v on the

64×64 grid. This inversion was performed with 65 receivers that were distributed uniformly with an even spacing of 0.54 km on the free surface and with a grid capable of resolving around 10×10 wavelengths in both the length and depth directions. The high fidelity of the results is noteworthy: even fine features of the size of a fraction of a wavelength are accurately resolved.

In Figure 7, we compare the inverted ground velocity time history to the target ground velocity time history recorded at a receiver located 8.75 km from the left end of the domain. We observe that even before the final stage of the multilevel grid continuation, the corresponding seismograms are close to the target waveform although the shear-wave velocity recovery at those stages is still poor. This is a manifestation of the ill posedness of the seismic waveform inversion problem.

To assess the accuracy of the inverted waveforms at nonreceiver locations, we solve again the same problem but now using only 33 receivers from among the original 65, leaving out every other one. We examine the inverted waveforms at 32 grid points on the surface, of which 16 correspond to receivers and the other 16 to nonreceivers. We then compare the inverted waveforms with the corresponding target waveforms. The top portion of Figure 8 shows the locations of select grid points and the lower portion the displacement and velocity seismograms registered at the 32 grid points. There is a very close match between the target and inverted response at all grid points, including the nonreceiver locations.

Table 2 displays the numerical performance of the inversion algorithm in terms of the number of Gauss–Newton and CG iterations for the consecutive stages of the multilevel algorithm. The results show that the numbers of CG and Gauss–Newton iterations do not increase with the grid size beyond 8×8 grid, but actually decrease.

The intermediate inversion results in Figures 6 and 7 are shown mainly to illustrate the convergence steps of our mul-
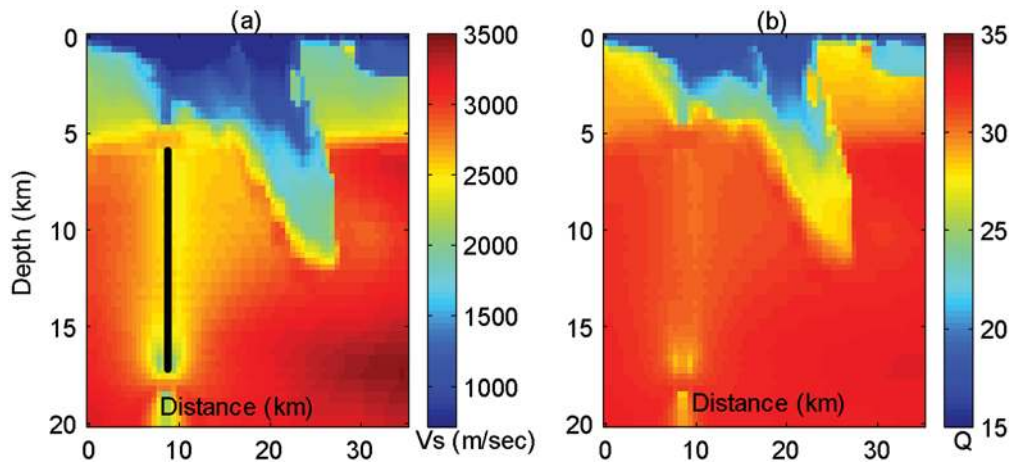

Figure 4. (a) Target shear-wave velocity profile; (b) target damping model.
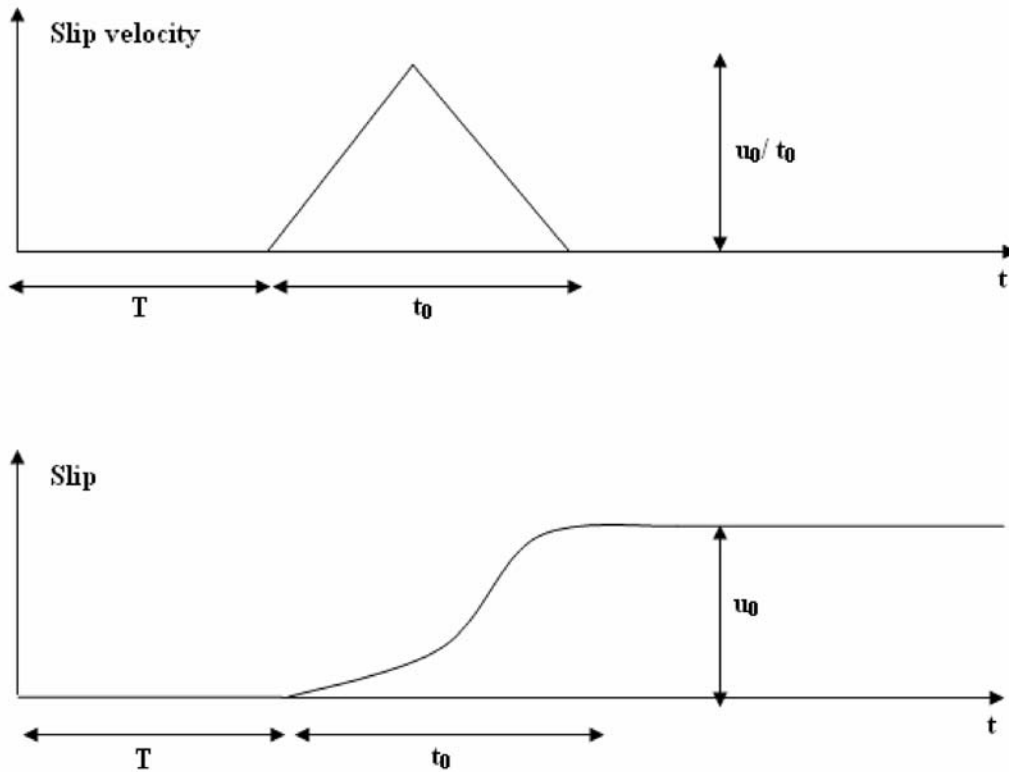
Figure 5.   Seismic source model.

tilevel inversion method. The sequence of optimization problems is solved automatically starting from a homogeneous initial guess, without intervention. It should be noted that alternative procedures for material inversion and imaging using full waveforms have been developed by others working in various fields of application, both in the time domain (Tarantola, 1984; Shipp and Singh, 2002) and in the frequency domain (Pratt, 1999; Sirgue and Pratt, 2004). An important attribute of our approach is that, due to the preconditioned Gauss–Newton-CG iterations, the number of iterations required is independent of the problem size (in both state and material fields) and is thus scalable to large problems.

Simultaneous Inversion for Shear-Wave Velocity and Intrinsic Attenuation

In the previous sections, we presented the solution technique and reported results of inversions performed for the shear modulus, or equivalently, the seismic velocity. We prescribed the viscoelastic parameters in terms of this single material variable and treated them as known quantities throughout the domain. In this section, we enlarge the set of inversion parameters so that we perform inversion simultaneously for the velocity and a viscoelastic parameter. The new inversion variable is the parameter a that enters into the nonlinear relationship between Q and V_S in (15). This parameter gives a direct measure of damping included in the soil model.

The methodology and solution technique for the two-variable optimization problem remain the same as before. However, the reduced Newton system now includes information due to the derivative of the Lagrangian with respect to the new inversion variable in addition to the original one. In this section, we present the modification in the solution technique and report results of inversions for the two variables. Later we discuss the effects of the additional inversion variable on the algorithmic and numerical performance of the inversion algorithm.

Recall from the section "Inverse Viscoelastic Wave Propagation Problem" that the relationship between Q and V_S in our approach is given in a nonlinear form as follows:

$$Q = a \arctan\left(\frac{V_S}{V_{S\mathrm{ref}}}\right), \qquad (27)$$

Table 1
Source and Model Parameters

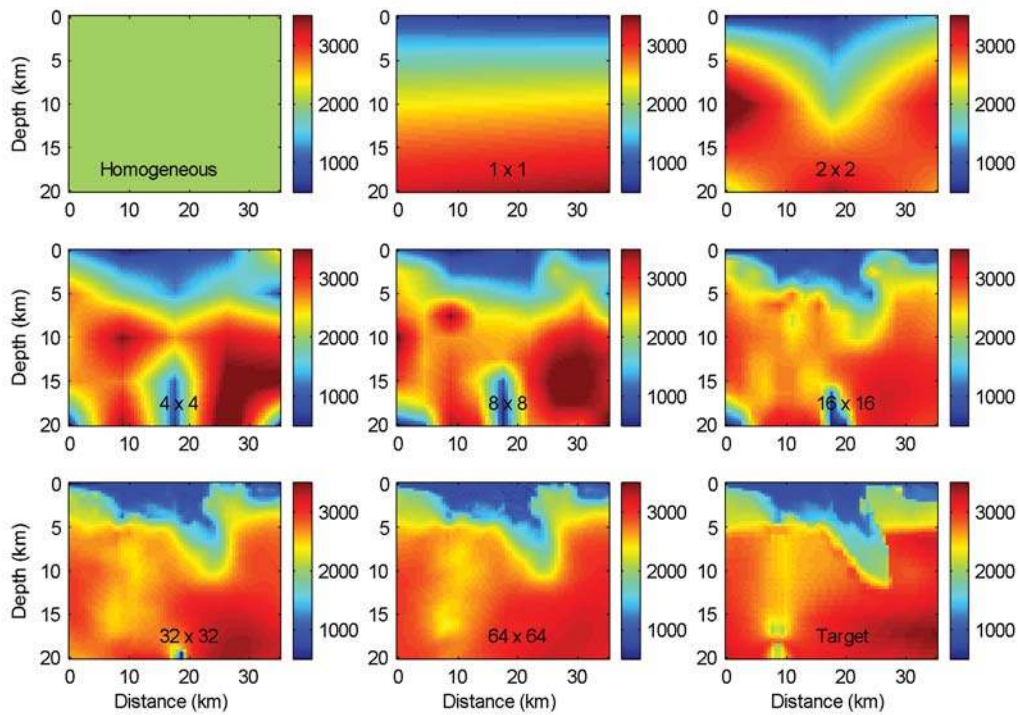| Domain size (km)	35×20
Grid dimensions (km)	0.54×0.31
Timestep Δt (sec)	0.06
Simulation time t_D (sec)	30
Minimum velocity (m/sec)	715
Maximum velocity (m/sec)	3470
Number of receivers	65
Regularization parameter, β	0.001

Figure 6. Inverted shear velocity profiles for the consecutive stages of the multilevel inversion algorithm (V_S in m/sec). The last image (bottom right-hand corner) shows the target profile.
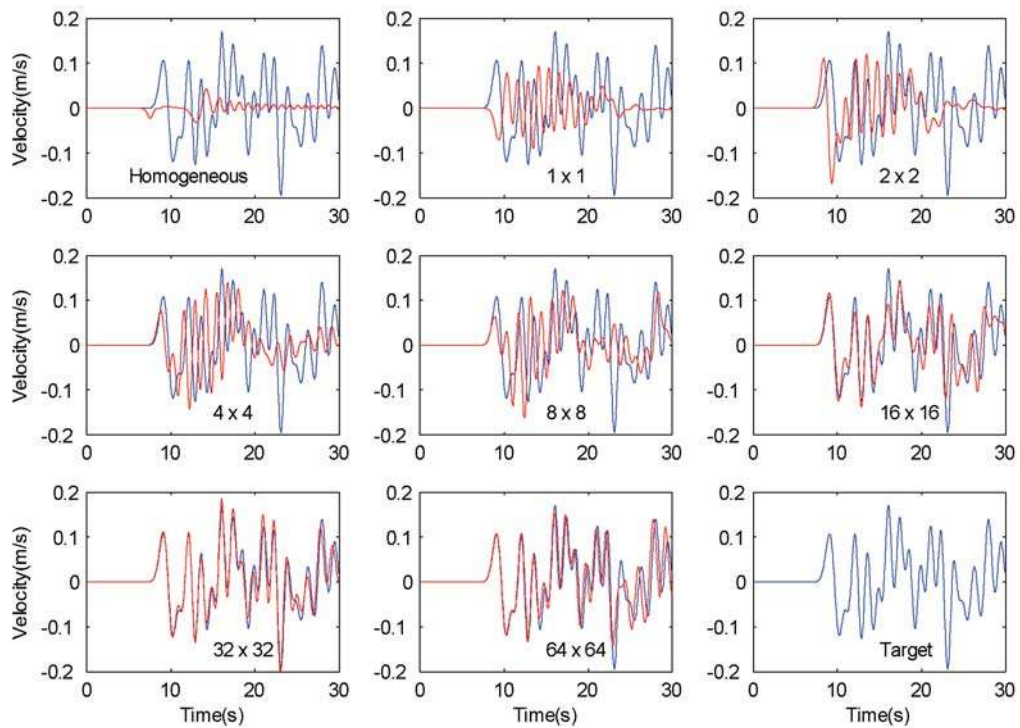

Figure 7. Target versus inverted ground velocity time history for the consecutive stages of the multiscale inversion algorithm at a location 8.75 km from the left end of the domain (target time history, blue line; inverted time history, red line).
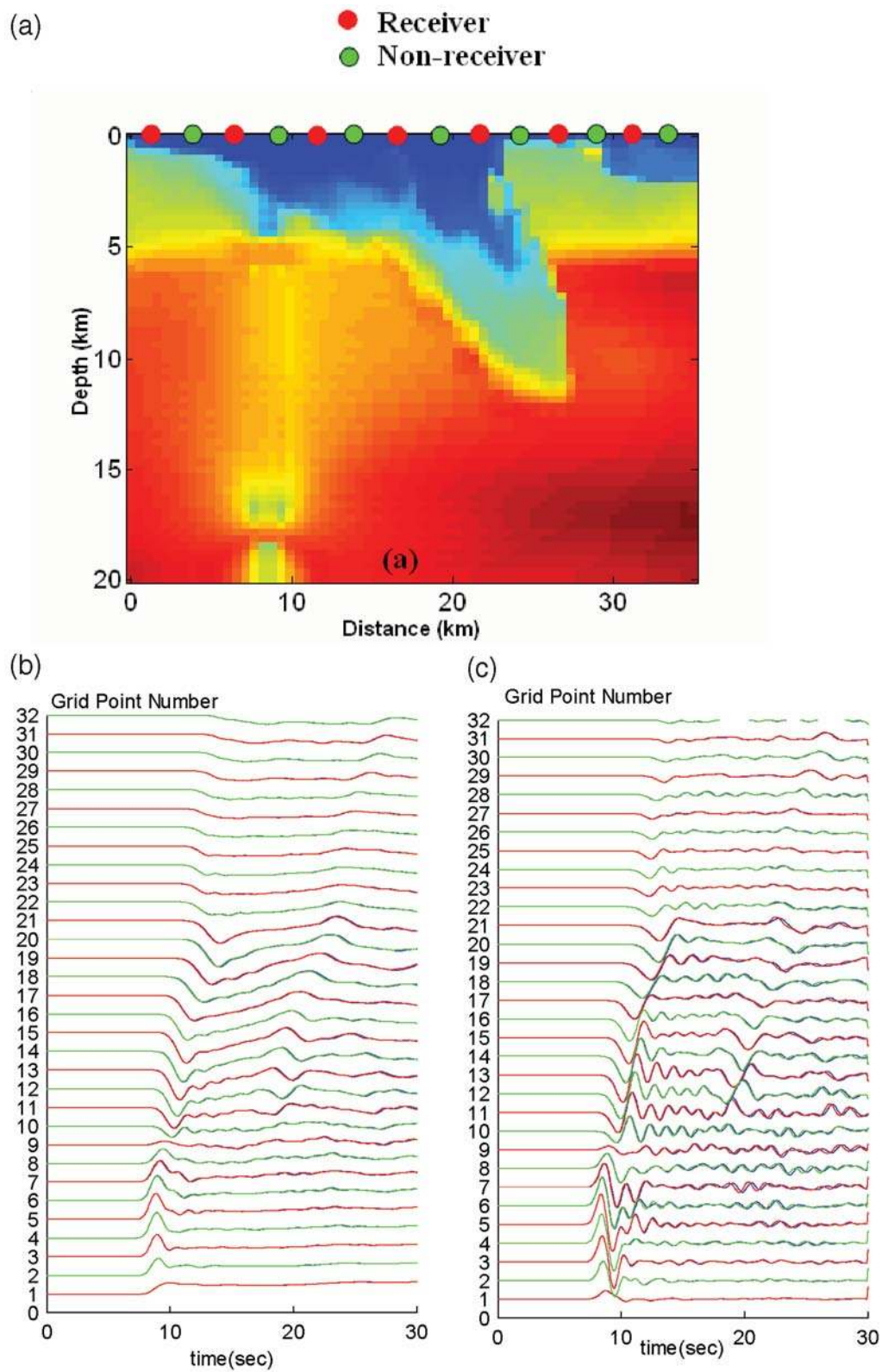
Figure 8. (a) Receiver and nonreceiver grid points on the surface shown only partially for clarity. (b) Target and inverted displacement time histories. (c) Target and inverted velocity time histories (receiver, red curve; nonreceiver, green curve; inverted time histories, blue curve).

where the parameter a is found to be 25.5 for the shear strain range exhibited by the basin soils. In the sections "Inverse Viscoelastic Wave Propagation Problem" and "Example: Synthetic Two-dimensional Viscoelastic Waveform Inversion," we constrained Q to satisfy (27) with $a = 25.5$ both for the forward and inverse wave propagations. Here, in the forward problem, Q is assigned the values obtained from (27) with $a = 25.5$ as before, but in the inverse problem, we release the parameter a by regarding it as an independent inversion variable. Using results from this inversion, we are able to compare the inverted values of Q with the corresponding target values.

To avoid the challenge of inverting simultaneously for both the velocity and the parameter a at every level, we initially let the parameter a have its value fixed to the target value $a = 25.5$ at the first seven stages of the multilevel algorithm, where the only inversion variable is velocity. We then incorporate a as an independent inversion variable once we reach the finest optimization grid.

The solution technique is the same as for one-variable inversion; however, the solution space increases with the addition of the new variable. Including the parameter as an inversion variable in the solution, the optimality conditions and the Newton update for the new full space system are as follows:

$$\begin{Bmatrix} \delta_u \mathcal{L} \\ \delta_v \mathcal{L} \\ \delta_\mu \mathcal{L} \\ \delta_a \mathcal{L} \\ \delta_\lambda \mathcal{L} \\ \delta_\phi \mathcal{L} \end{Bmatrix} (u, v, \mu, a, \lambda, \phi) = 0, \qquad (28)$$

$$\begin{bmatrix} \delta_{uu}^2 \mathcal{L} & \delta_{uv}^2 \mathcal{L} & \delta_{u\mu}^2 \mathcal{L} & \delta_{ua}^2 \mathcal{L} & \delta_{u\lambda}^2 \mathcal{L} & \delta_{u\phi}^2 \mathcal{L} \\ \delta_{vu}^2 \mathcal{L} & \delta_{vv}^2 \mathcal{L} & \delta_{v\mu}^2 \mathcal{L} & \delta_{va}^2 \mathcal{L} & \delta_{v\lambda}^2 \mathcal{L} & \delta_{v\phi}^2 \mathcal{L} \\ \delta_{\mu u}^2 \mathcal{L} & \delta_{\mu v}^2 \mathcal{L} & \delta_{\mu\mu}^2 \mathcal{L} & \delta_{\mu a}^2 \mathcal{L} & \delta_{\mu\lambda}^2 \mathcal{L} & \delta_{\mu\phi}^2 \mathcal{L} \\ \delta_{au}^2 \mathcal{L} & \delta_{av}^2 \mathcal{L} & \delta_{a\mu}^2 \mathcal{L} & \delta_{aa}^2 \mathcal{L} & \delta_{a\lambda}^2 \mathcal{L} & \delta_{a\phi}^2 \mathcal{L} \\ \delta_{\lambda u}^2 \mathcal{L} & \delta_{\lambda v}^2 \mathcal{L} & \delta_{\lambda\mu}^2 \mathcal{L} & \delta_{\lambda a}^2 \mathcal{L} & \delta_{\lambda\lambda}^2 \mathcal{L} & \delta_{\lambda\phi}^2 \mathcal{L} \\ \delta_{\phi u}^2 \mathcal{L} & \delta_{\phi v}^2 \mathcal{L} & \delta_{\phi\mu}^2 \mathcal{L} & \delta_{\phi a}^2 \mathcal{L} & \delta_{\phi\lambda}^2 \mathcal{L} & \delta_{\phi\phi}^2 \mathcal{L} \end{bmatrix} \begin{Bmatrix} \bar{u} \\ \bar{v} \\ \bar{\mu} \\ \bar{a} \\ \bar{\lambda} \\ \bar{\phi} \end{Bmatrix} =$$

$$- \begin{Bmatrix} \delta_u \mathcal{L} \\ \delta_v \mathcal{L} \\ \delta_\mu \mathcal{L} \\ \delta_a \mathcal{L} \\ \delta_\lambda \mathcal{L} \\ \delta_\phi \mathcal{L} \end{Bmatrix}. \qquad (29)$$

Because $\alpha = \alpha(Q)$ and $\eta = \eta(Q)$, where $Q = Q(\mu, a)$, the parameters α and η are functions of μ and a. Similar to the one-variable case, after the block elimination, the final

Table 2

Number of CG and Gauss–Newton Iterations for Consecutive Stages of the Multilevel Algorithm

Grid	CG Iterations (Gauss–Newton Iterations)
Homogeneous	59 (59)
1×1	41 (29)
2×2	58 (33)
4×4	151 (59)
8×8	385 (67)
16×16	262 (53)
32×32	213 (34)
64×64	196 (27)

reduced Gauss–Newton system can be written as

$$\begin{bmatrix} W_{\mu\mu} & W_{\mu a} \\ W_{a\mu} & W_{aa} \end{bmatrix} \begin{Bmatrix} \bar{\mu} \\ \bar{a} \end{Bmatrix} = - \begin{Bmatrix} \delta_\mu \\ \delta_a \end{Bmatrix}. \qquad (30)$$

The step length is calculated using a backtracking line search, as before, and both variables are updated using the same step length α as follows:

$$\mu_{k+1} = \mu_k + \alpha_k \bar{\mu}_k, \qquad a_{k+1} = a_k + \alpha_k \bar{a}_k.$$

Next, we present results for the two-variable inversion problem. We use the target shear-wave velocity profile described in the previous section, 65 receivers, TV regularization, and the multilevel technique. The value of a_{target} for the forward problem is 25.5, and the initial guess for a is taken to be 22.0. Concurrent inversion for a and μ yields $a_{\text{inverted}} = 25.57$, and the shear velocity profiles for the consecutive stages of the multilevel algorithm are depicted in Figure 9. The inverted waveforms are compared to the target waveforms at a receiver (located at $x = 8.75$ km) in Figure 10. As before, to study the effect of receiver density on inversion results, we run the concurrent inversion problem with 33 receivers. Figure 11 shows the displacement and velocity time histories at several grid points including receivers and nonreceivers. There is a close match between the target and inverted response at all grid points including the nonreceiver locations as before.

Through the inversion of the viscoelastic parameter, we recover a damping model. To see how well the inverted damping model matches the target damping model; in Figure 12 we plot the inverted damping model through the relationship

$$Q_{\text{inverted}} = a_{\text{inverted}} \arctan\left[\frac{(V_S)_{\text{inverted}}}{V_{S\text{ref}}}\right]$$

and compare it to Q_{target}, calculated as

$$Q_{\text{target}} = a_{\text{target}} \arctan\left[\frac{(V_S)_{\text{target}}}{V_{S\text{ref}}}\right].$$
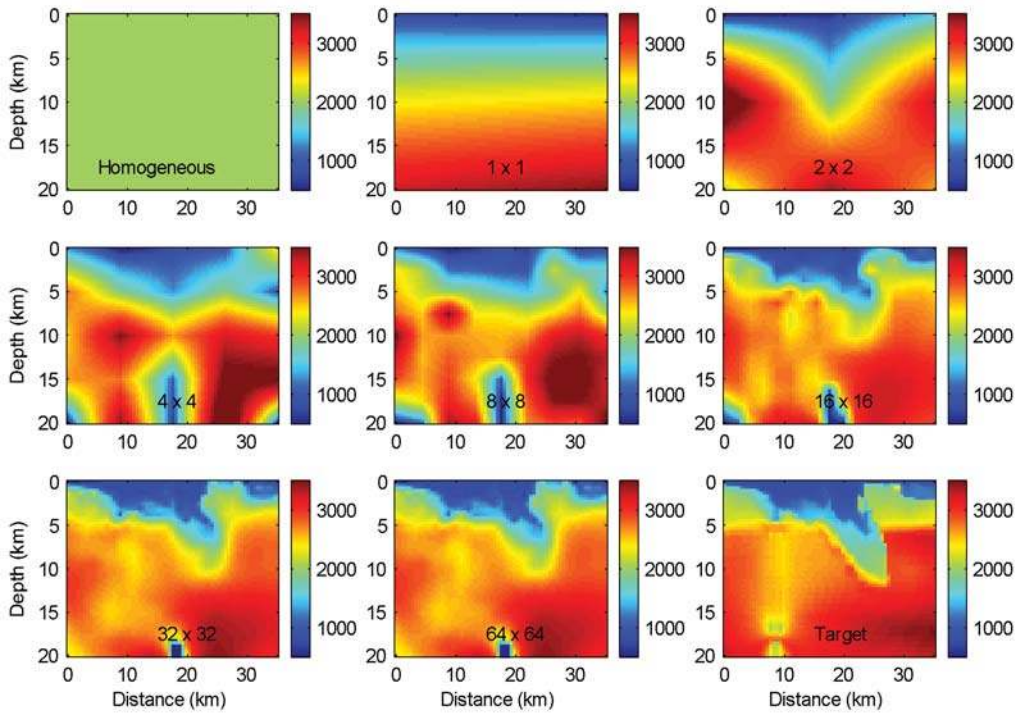
Figure 9.   Inverted shear velocity profiles for the consecutive stages of the multiscale inversion algorithm (V_S in m/sec).
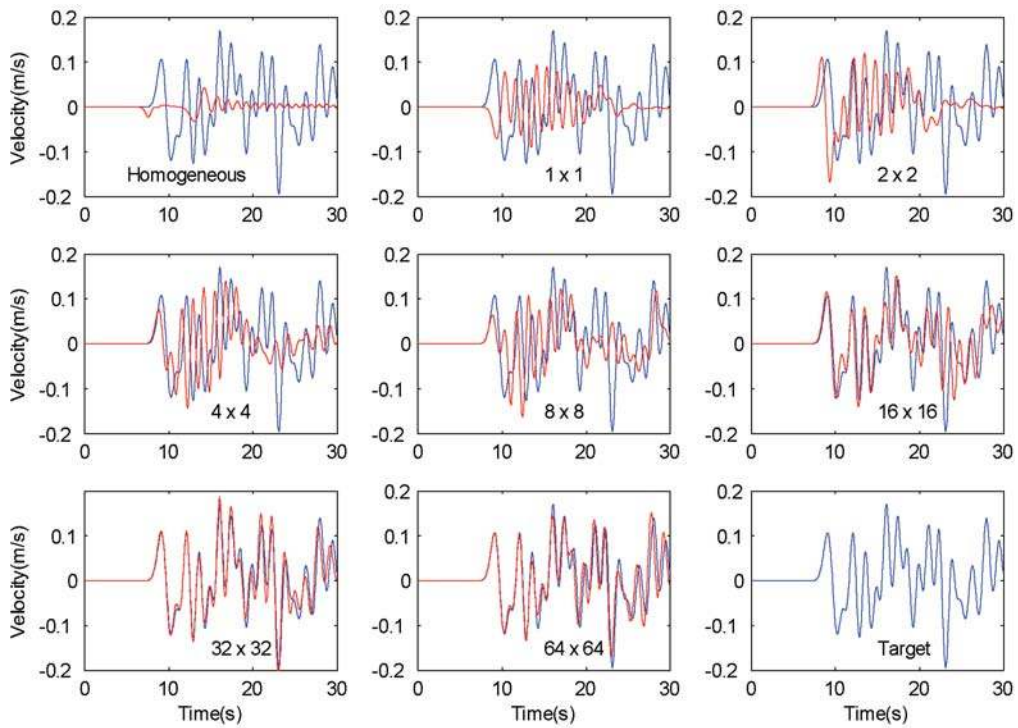

Figure 10.   Target versus inverted ground velocity time history for the consecutive stages of the multiscale inversion algorithm with selected parameters for the two-variable inversion at a location 8.75 km from the left end of the domain (target time history, blue line; inverted time history, red line).
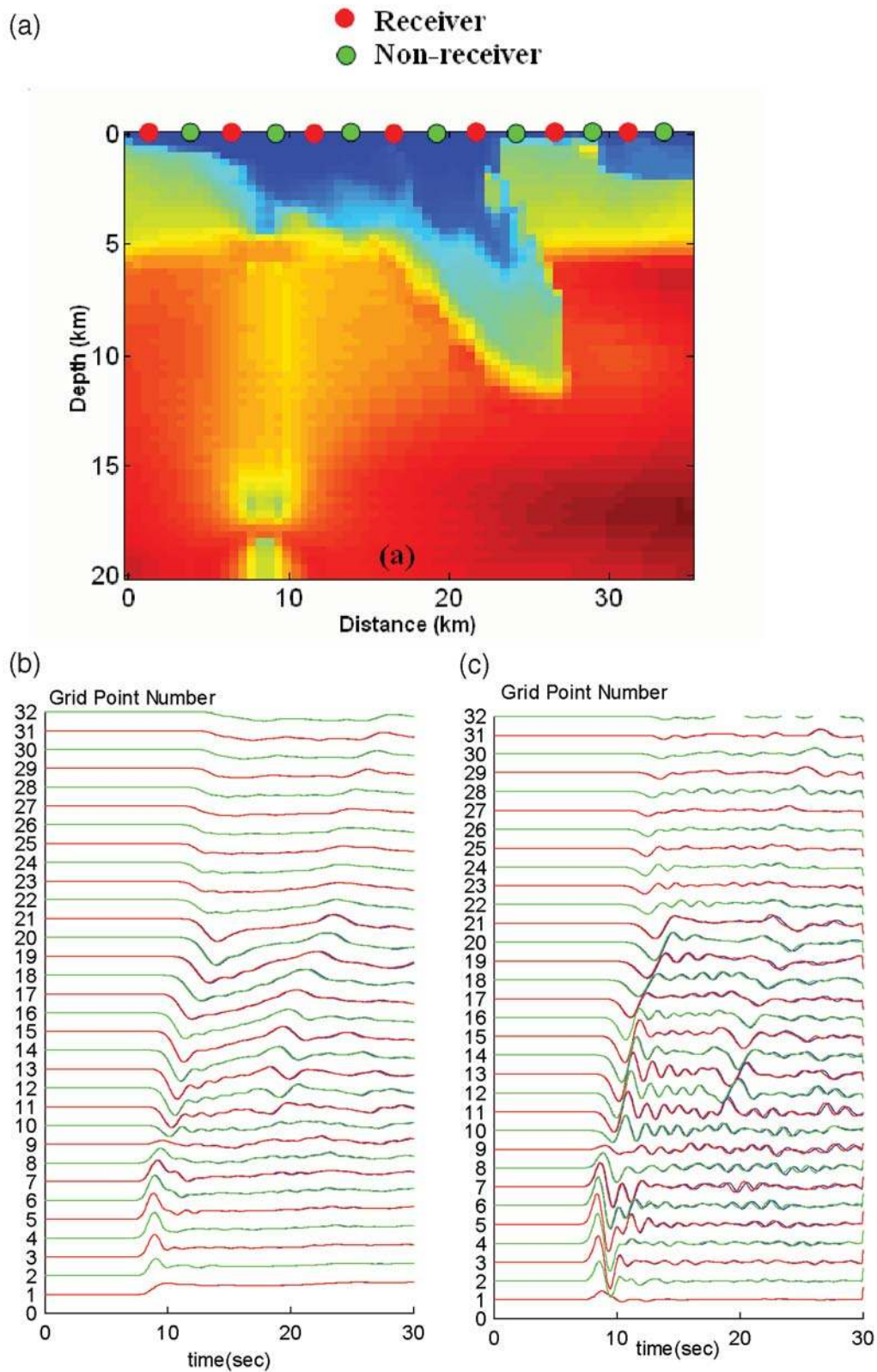
Figure 11. (a) Receiver and nonreceiver grid points on the surface. (b) and (c) Target and inverted displacement and velocity time histories for the two-variable inversion (receiver, red curve; nonreceiver, green curve; inverted time histories, blue curve).
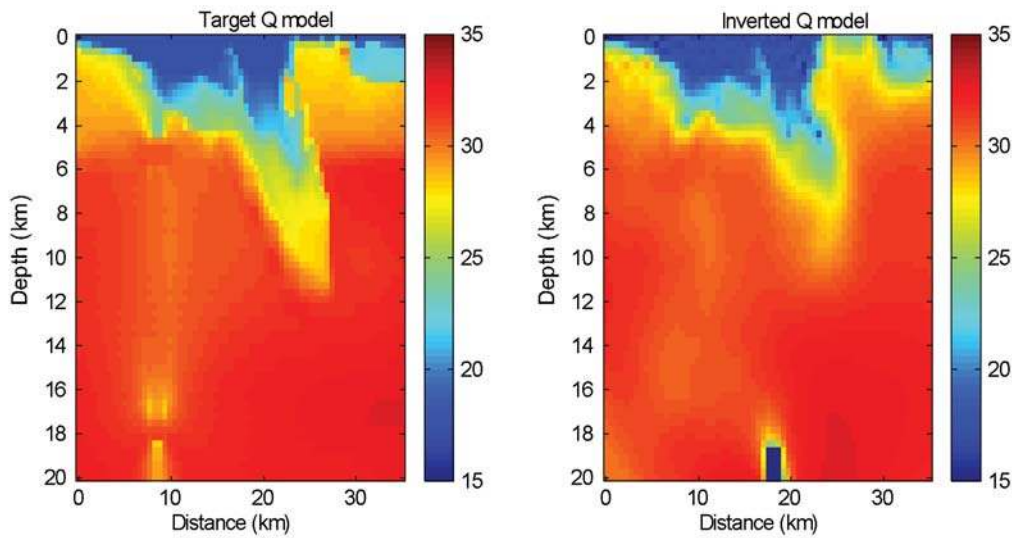
Figure 12.   Comparison of target and inverted Q models.

By comparing Figures 6 and 9, we see that the new inversion with a treated as an independent inversion variable towards the end of the multilevel inversion process leads eventually to the same inverted model for the velocity as the original one, in which a is assigned to the target value from the beginning. This is mainly due to the fact that we incorporate the second inversion variable only after convergence has been attained for the first one, though with an initial guess for a that differs from the target value by about 15%.

Table 3 compares the numerical performance of the two-variable inversion algorithm with the one-variable case. We note that $\beta = 10^{-3}$ in both cases. The numerical performance of the algorithm until the finest grid is not affected. In the last grid, the number of iterations increased by 20%. This does not cause a substantial increase in the overall computing time.

Summary

In this article, we presented an inverse wave propagation algorithm to determine seismic velocity and intrinsic attenuation of soil media. We formulated our inverse problem as a least-squares minimization problem, constrained by the forward viscoelastic wave propagation initial-boundary value problem. To model the intrinsic attenuation effects, we used a model consisting of SLSs. We initially related the mechanical properties to the attenuation level, Q^{-1}, using the complex frequency-dependent relaxation modulus concept. Subsequently, we derived predictive equations between the mechanical properties of the SLS and Q^{-1} and linked Q^{-1} to the shear-wave velocity through a semiempirical model.

To overcome the inherent challenges of inverse wave propagation problems posed by ill conditioning and multiple minima, we employed regularization and multilevel grid continuation, respectively. To solve the large-scale nonlinear optimization problem, we used a Gauss–Newton-CG solution technique, which does not require the Hessian matrix explicitly. A key difference between other seismic inversion approaches that use time-domain waveform misfits and ours is that the former require that the initial guess be located in the neighborhood of the global minimum. Our approach does not impose this requirement. (Another important attribute of our approach is that it is scalable, making it amenable to parallel computation).

We used our inversion method to perform synthetic inversions for a two-dimensional seismic velocity profile of a sedimentary basin model from the San Fernando Valley under *SH*-wave excitations using a single SLS rheological model. We initially performed inversions for the shear-wave velocity taking intrinsic attenuation as a known quantity. Then, permitting Q^{-1} to vary as another inversion variable, we obtained inversion results for shear-wave velocity and intrinsic attenuation together. Overall, the results show that the multilevel limited memory BFGS preconditioned Gauss–Newton-CG method with TV regularization proves to be effective for viscoelastic inversion.

Table   3

Numerical Performance of the
Two-Variable Inversion

	CG Iterations (Gauss–Newton Iterations)	
Grid	One-Variable	Two-Variable
Homogeneous	59 (59)	59 (59)
1×1	41 (29)	41 (29)
2×2	58 (33)	58 (33)
4×4	151 (59)	151 (59)
8×8	385 (67)	385 (67)
16×16	262 (53)	262 (53)
32×32	213 (34)	213 (34)
64×64	196 (27)	232 (49)

In summary, the main contributions of this work are as follows:

1. This study is a successful attempt to invert simultaneously for elastic and viscoelastic properties in wave propagation problems.
2. We have derived a realistic, nonlinear, semiempirical predictive relationship between Q and V_S for the problem-specific shear strain range. It is possible to derive similar predictive relationships that could be used as material property models for other forward or inverse wave propagation problems of interest.
3. Based on our results, it is possible to generate attenuation models. This, in turn, will lead to more precise damping models in next-generation ground-motion modeling studies.
4. Using several numerical remedies, we have succeeded in performing inversions with shear-wave velocity contrasts up to five ($V_{S_{max}}/V_{S_{min}} = 5$). This issue is essential because in the case of real-data inversions, the contrast in actual material property profiles can be very high.
5. Our technique remains to be validated, of course, with actual observations. From the results of the idealized investigations we presented, it appears that the proposed large-scale methodology will provide a useful basis for possible further real-data inversions in basins that exhibit complex geotechnical and seismic features.

Acknowledgments

References

Akcelik, V., G. Biros, and O. Ghattas (2002). Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation, *Proceedings of ACM/IEEE SC2002*.

Akcelik, V., G. Biros, O. Ghattas, J. Hill, D. Keyes, and B. van Bloemen Waanders (2006). Parallel algorithms for PDE-constrained optimization, in *Frontiers of Parallel Computing*, M. Heroux, P. Raghaven and H. Simon (Editors), SIAM, Philadelphia, Pennsylvania.

Akcelik, V., J. Bielak, G. Biros, I. Epanomeritakis, A. Fernandez, O. Ghattas, E. J. Kim, J. Lopez, D. O'Hallaron, T. Tu, and J. Urbanic (2003). High-resolution forward and inverse earthquake modeling on terascale computers, *Proceedings of ACM/IEEE SC2003*.

Aoi, S. (2002). Boundary shape waveform inversion for estimating the depth of three-dimensional basin structures, *Bull. Seismol. Soc. Am.* **92,** 2410–2418.

Askan, A. (2006). Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures, *Ph.D. thesis*, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Backus, G., and F. Gilbert (1967). Numerical application of a formalism for geophysical inverse problems, *Geophys. J. R. Astr. Soc.* **13,** 247–276.

Backus, G., and F. Gilbert (1970). Uniqueness in the inversion of inaccurate gross Earth data, *Philos. Trans. R. Soc. A* **266,** 123–192.

Baig, A. M., F. A. Dahlen, and S.-H. Hung (2003). Traveltimes of waves in random 3-D media, *Geophys. J. Int.* **153,** 1–16.

Balay, S., K. Buschelman, W. D. Gropp, D. Kaushik, M. Knepley, L. C. Mcinnes, B. F. Smith, and H. Zhang (2001). PETSC user manual, http://www.mcs.anl.gov/petsc (last accessed March 2007).

Bielak, J., and R. C. MacCamy (1989). On the strength of mechanical and thermal damping in linear materials, *Quart. Appl. Math.* **47,** 555–570.

Biot, M. A. (1958). Linear thermodynamics and the mechanics of solids, *Proceedings of the 3rd U.S. National Congress of Applied Mechanics*, ASME, New York, 1–18.

Blanch, J. O., J. O. A. Robertsson, and W. W. Symes (1995). Modeling of a constant Q: methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique, *Geophysics* **60,** 176–184.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995). Multiscale seismic waveform inversion, *Geophysics* **50,** 1457–1473.

Carcione, J. M., D. Kosloff, and R. Kosloff (1988). Wave propagation simulation in a linear viscoelastic medium, *Geophys. J.* **95,** 597–611.

Caughey, T. K. (1962). Vibration of dynamic systems with linear hysteretic damping (linear theory), *Proceedings of the 4th US National Congress of Applied Mechanics*, ASME, New York, 87–97.

Chen, P. (2005). A unified methodology for seismic waveform analysis and inversion, *Ph.D. Thesis*, University of Southern California, Los Angeles, California.

Chen, P., T. H. Jordan, and L. Zhao (2007). Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int.* **170,** 175–181.

Chen, P., L. Zhao, and T. H. Jordan (2007). Full 3D tomography for crustal structure of the Los Angeles region, *Bull. Seismol. Soc. Am.* **97,** 1094–1120.

Day, S. M. (1998). Efficient simulation of constant Q using coarse-grained memory variables, *Bull. Seismol. Soc. Am.* **88,** no. 4, 1051–1062.

Day, S. M., and C. R. Bradley (2001). Memory efficient simulation of anelastic wave propagation, *Bull. Seismol. Soc. Am.* **91,** 520–531.

Day, S. M., and J. B. Minster (1984). Numerical simulation of attenuated wave fields using a Padè approximate method, *Geophys. J. R. Astr. Soc.* **78,** 105–118.

Eisenstat, S. C., and H. F. Walker (1996). Choosing the forcing terms in an inexact Newton method, *SIAM J. Sci. Comput.* **17,** 16–32.

Emmerich, H., and M. Korn (1987). Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics* **52,** 1252–1264.

Graves, R. W., and S. M. Day (2003). Stability and accuracy analysis of coarse-grain viscoelastic simulations, *Bull. Seismol. Soc. Am.* **93,** 283–300.

Gucunski, N., and R. D. Woods (1992). Numerical simulation of the SASW test, *Soil Dyn. Earthq. Eng.* **11,** 213–227.

Guzina, B. B., and A. Lu (2002). Coupled waveform analysis in dynamic characterization of lossy solids, *J. Eng. Mech.* **128,** 392–402.

Hardin, B. O., and V. P. Drnevich (1972). Shear modulus and damping in soils, *J. Soil Mech. Found. Div.* **98,** no. 7, 667–692.

Hicks, G., and R. G. Pratt (2001). Reflection waveform inversion using local descent methods: estimating attenuation and velocity over a gas-sand deposit, *Geophysics* **66,** 598–612.

Ishibashi, I., and X. J. Zhang (1993). Unified dynamic shear moduli and damping ratios of sand and clay, *Soils Found.* **33,** no. 1, 182–191.

Ishii, M., J. Tromp, A. M. Dziewonski, and G. Ekström (2002). Joint inversion of normal-mode and body-wave data for inner-core anisotropy: 1. Laterally homogeneous anisotropy, *J. Geophys. Res.* **107,** no. B12, 2379, doi 10.1029/2001JB000712.

Jones, R. B. (1958). In-situ measurement of the dynamic properties of soil by vibration methods, *Geotech.* **8,** no. 1, 1–21.

Jordan, T. H., and D. L. Anderson (1974). Earth structure from free oscillations and travel times, *Geophys. J. R. Astr. Soc.* **36,** 411–459.

Kramer, S. L. (1996). *Geotechnical Earthquake Engineering*, Prentice Hall, Upper Saddle River, New Jersey.

Lai, C. G., G. J. Rix, S. Foti, and V. Roma (2002). Simultaneous measurement and inversion of surface wave dispersion and attenuation curves, *Soil Dyn. Earthq. Eng.* **22,** 923–930.

Lerner-Lam, A. L., and T. H. Jordan (1983). Earth structure from fundamental and higher-mode waveform analysis, *Geophys. J. R. Astr. Soc.* **75,** 759–797.

Liao, Q., and G. A. McMechan (1996). Multifrequency viscoacoustic modeling and inversion, *Geophysics* **61,** no. 5, 1371–1378.

Liu, H., D. L. Anderson, and H. Kanamori (1976). Velocity dispersion due to anelasticity: implications for seismology and mantle composition, *Geophys. J. R. Astr. Soc.* **47,** 41–58.

Magistrale, H., S. Day, R. W. Clayton, and R. Graves (2000). The SCEC Southern California reference three-dimensional seismic velocity model version 2, *Bull. Seismol. Soc. Am.* **90,** S65–S76.

Moczo, P., J. O. A. Robertsson, and L. Eisner (2006). The finite-difference time-domain method for modelling of seismic wave propagation, in *Advances in Wave Propagation in Heterogeneous Earth*, Ru-Shan Wu and Valerie Maupin (Editors), Advances in Geophysics, R. Dmowska (Series Editor), Vol. **48,** Elsevier Academic Press, New York, 421–516.

Montelli, R., G. Nolet, G. Masters, F. A. Dahlen, and S.-H Hung (2004). Global P and PP traveltime tomography: rays versus waves, *Geophys. J. Int.* **158,** no. 2, 637–654.

Morales, J. L., and J. Nocedal (2000). Automatic preconditioning by limited memory quasi-Newton updating, *SIAM J. Optim.* **10,** no. 4, 1079–1096.

Nocedal, J., and S. J. Wright (1999). *Numerical Optimization*, Springer, New York.

Olsen, K. B., S. M. Day, and C. R. Bradley (2003). Estimation of Q for long-period waves in the Los Angeles basin, *Bull. Seismol. Soc. Am.* **93,** no. 2, 627–638.

Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics* **64,** no. 3, 888–901.

Pratt, R. G., and M. H. Worthington (1990). Inverse theory applied to multi-source cross-hole tomography. Part I: Acoustic wave-equation method, *Geophys. Prospect.* **38,** 287–310.

Pyke, R. M. (1993). Modeling of dynamic soil properties, Guidelines for Determining Design Bases Ground Motions, Appendix 7.A, Electric Power Research Institute, Palo Alto, California, 7.A-1–7.A-90.

Richart, F. E., J. R. Hall, and R. D. Woods (1970). *Vibrations of soils and foundations*, Prentice Hall, Englewood Cliffs, New Jersey.

Rix, G. J., C. G. Lai, and A. W. Spang (2000). In situ measurements of damping ratio using surface waves, *J. Geotechnol. Geoenviron. Eng.* **126,** 472–480.

Robertsson, J. O. A., J. O. Blanch, and W. W. Symes (1994). Viscoelastic finite-difference modeling, *Geophysics* **59,** 1444–1456.

Seed, H. B., and I. M. Idriss (1970). Rep. No. EERC 70-10, Soil moduli and damping factors for dynamic response analysis, Earthquake Engineering Research Center, Berkeley, California.

Sheng, J., and G. T. Schuster (2003). Finite-frequency resolution limits of wave path traveltime tomography for smoothly varying models, *Geophys. J. Int.* **152,** 669–676.

Shipp, R. M., and S. C. Singh (2002). Two-dimensional full waveform inversion of wide-aperture marine seismic streamer data, *Geophys. J. Int.* **151,** 325–344.

Sirgue, L., and R. G. Pratt (2004). Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics* **69,** 231–248.

Stokoe, K. H., II, M. B. Darendeli, R. D. Andrus, and L. T. Brown (1999). Dynamic soil properties: laboratory, field and correlation studies, in *Proc. 2nd Int. Conf. on Earthquake Geotechnical Engineering*, Vol. 3, Lisbon, Portugal, 811–845.

Stokoe, K. H., II, S. G. Wright, J. Bay, and J. Roësset (1994). Characterization of geotechnical sites by SASW method, in *Geophysical Characterization of Sites*, R. D. Woods (Editor), Oxford/IBH Publishing, New York/New Delhi, 15–25.

Stokoe, K. H., II, M. B. Darendeli, R. B. Gilbert, F. Y. Menq, and W. K. Choi (2004). Development of a new family of normalized modulus reduction and material damping curves, in *Proc. NSF/PEER Int. Workshop on Uncertainties in Nonlinear Soil Properties and Their Impact on Modeling Dynamic Soil Response*, University of California at Berkeley.

Symes, W. W. (1990). Velocity inversion: a case study in infinite-dimensional optimization, *Math Program.* **48,** 71–102.

Tape, C., Q. Liu, and J. Tromp (2007). Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves, *Geophys. J. Int.* **168,** 1105–1129.

Tarantola, A. (1984). Inversion of Seismic Reflection Data in the Acoustic Approximation, *Geophysics* **49,** no. 8, 1259–1266.

Tarantola, A. (1988). Theoretical background for the inversion of seismic waveforms including elasticity and attenuation, *Pure Appl. Geophys.* **128,** 365–399.

Tromp, J., C. Tape, and Q. Liu (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.* **160,** 195–216.

Vogel, C. (2002). Total variation regularization chapter, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, Pennsylvania.

Vucetic, M., and R. Dobry (1991). Effect of soil plasticity on cyclic response, *J. Geotech. Eng.* **117,** no. 1, 89–107.

Vucetic, M., G. Lanzo, and M. Doroudian (1998). Damping at small strains in cyclic simple shear test, *J. Geotechnol. Geoenviron. Eng.* **124,** no. 7, 585–594.

Wiggens, R. A., G. A. Frazier, J. Sweet, and R. Apsel (1978). Modeling strong motions from major earthquakes, *Proc. of the 2nd International Conference on Microzonation for Safer Construction: Research and Application*, San Francisco, California, 693–700.

Zhang, J., R. D. Andrus, and C. H. Juang (2005). Normalized shear modulus and material damping ratio relationships, *J. Geotechnol. Geoenviron. Eng.* **131,** no. 4, 453–464.

Zhao, L., T. H. Jordan, K. B. Olsen, and P. Chen (2005). Fréchet kernels for imaging regional Earth structure based on three-dimensional reference models, *Bull. Seismol. Soc. Am.* **95,** no. 6, 2066–2080.

Computational Seismology Laboratory
Department of Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
aaskan@andrew.cmu.edu
  (A.A., J.B.)

Stanford Linear Accelerator Center
Stanford University
Menlo Park, California 94025
volkan@slac.stanford.edu
  (V.A.)

Jackson School of Geosciences
Department of Mechanical Engineering
Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712
omar@ices.utexas.edu
  (O.G.)