JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 109, B02318, doi:10.1029/2003JB002625, 2004

A new nonlinear finite fault inversion with three-dimensional Green's functions: Application to the 1989 Loma Prieta, California, earthquake

Pengcheng Liu and Ralph J. Archuleta[1]

Institute for Crustal Studies, University of California, Santa Barbara, California, USA

Received 11 June 2003; revised 4 November 2003; accepted 26 December 2003; published 27 February 2004.

[1] We present a new procedure to invert for kinematic source parameters on a finite fault. On the basis of the reciprocity relation of the Green's functions, we use a newly developed fourth-order viscoelastic finite-difference algorithm to calculate three-dimensional (3-D) Green's functions (actually the tractions) on the fault. We invert the data for the unknown source parameters at the nodes (or corners) of the subfaults. The source parameters within a subfault area are allowed to vary; this variation is calculated through bilinear interpolation of the four nodal quantities. We have developed a global nonlinear inversion algorithm that is based on simulated annealing methods to solve efficiently for the nodal parameters. We apply this method to the 1989 Loma Prieta, California, *M* 6.9 earthquake for both a 1-D and 3-D velocity structure. We show (1) the bilinear interpolation technique reduces the dependence of inversion results on the subfault size by naturally including the effects of nearby subfaults. (2) While the number of synthetic seismograms that must be computed is greatly increased by the bilinear interpolation, the structure of the inversion method minimizes the actual numbers of computations. (3) As expected, complexity in the velocity structure is mapped into the source parameters that describe the rupture process; there are significant differences between faulting models derived from 1-D and 3-D structural models. *INDEX TERMS:* 1734 History of Geophysics: Seismology; 3230 Mathematical Geophysics: Numerical solutions; 3260 Mathematical Geophysics: Inverse theory; *KEYWORDS:* source process, ground motion, inversion

Citation: Liu, P., and R. J. Archuleta (2004), A new nonlinear finite fault inversion with three-dimensional Green's functions: Application to the 1989 Loma Prieta, California, earthquake, *J. Geophys. Res.*, *109*, B02318, doi:10.1029/2003JB002625.

1. Introduction

[2] Kinematics parameters of the rupture process [*Haskell*, 1964] form the basis of our inferences about the nature of earthquakes and provide a stepping stone for our understanding of earthquake physics. The kinematic parameters obtained from inversion of ground motion waveforms can be used to infer the stress drop distribution [e.g., *Mikumo and Miyatake*, 1995; *Bouchon*, 1997; *Day et al.*, 1998] that in turn can be used as the input for dynamic models [e.g., *Olsen et al.*, 1997; *Nielsen and Olsen*, 2000; *Peyrat et al.*, 2001; *Favreau and Archuleta*, 2003]. The kinematic parameters have been used to infer scaling properties [e.g., *Somerville et al.*, 1999; *Mai and Beroza*, 2002] and as input to finite difference codes in an attempt to determine frictional parameters [e.g., *Ide and Takeo*, 1997]. Of course, the spatial and temporal distribution of source parameters is critical in forward modeling of ground motion that can be used to predict ground motions for engineering design purposes and to study the effects of complex earth structure.

[3] The most common method for determining the kinematic source parameters on a finite fault is to invert the observed ground-motion data. In simple terms this is an automatic procedure that computes synthetic seismograms that are compared with the data [e.g., *Olson and Apsel*, 1982; *Hartzell and Helmberger*, 1982]. The procedure adjusts the kinematic parameters at a predetermined number of points on the fault. The inversion continues until an objective function that measures the difference between the synthetic seismograms and the data for all of the stations reaches a minimum. In the inversion the Green's functions play a critical role because they are essential for computing synthetic seismograms. Of course, the Green's functions depend on an assumed model that includes the geological structure, the elastic parameters, density and attenuation parameters. In some cases, empirical Green's functions have been used to invert for the kinematic parameters [*Hartzell*, 1989]. The finite-fault inversion usually requires dividing the fault into a grid of subfaults where each has a set of parameters to be determined by the inversion. The data

[1]Also at Department of Geological Sciences, University of California, Santa Barbara, California, USA.

0148-0227/04/2003JB002625$09.00

(time histories of ground motion) at a number of different stations are approximated by a linear sum of the synthetics generated by the subfaults. The size of subfault is a factor that affects the inversion. Obviously, we would like as small a subfault as possible to reflect the spatial variation of the rupture, but this leads to a larger set of unknown parameters and instability in the procedure. The fewer the number of subfaults the more overdetermined is the problem; however, the spatial resolution of the rupture process decreases.

[4] Determining only the distribution of slip amplitude is a linear problem. The static slip distribution is inadequate for describing the rupture process; a slip distribution does not radiate. The earthquake described kinematically is the process of going from one slip distribution (generally assumed to be zero) to another distribution of slip. The entire suite of kinematic parameters that describe the spatial distribution of slip amplitude, rupture velocity, rise time, or temporal function of the slip function, on the fault need to be determined. Comparisons of different rupture models based only on the similarity of the slip distribution are inadequate. The temporal quantities, for example, rupture time and rise time, are critical; however these temporal variables are not linearly related to the data [*Archuleta*, 1984]. Determining these temporal variables along with the slip is a nonlinear problem.

[5] Over the last two decades, the finite fault inversion method has been greatly improved. Initially a linear least squares method was used to determine the two components of slip for each subfault [*Hartzell and Helmberger*, 1982; *Olson and Apsel*, 1982]. *Olson and Apsel* [1982] allowed for a non-constant rupture velocity or variable rise time by introducing multiple time windows, i.e., slip in a particular subfault could occur more than once. An interesting feature in their method was that the time windows were centered about a constant rupture velocity; a rupture could be faster or slower than the assumed rupture. *Hartzell and Heaton* [1983] used multiple time windows where the first window in which slip can occur is the time of the assumed rupture velocity. Consequently, any rupture variation can only lead to a slower rupture velocity than the one used to initialize the inversion. *Cohee and Beroza* [1994] provide a synopsis (up to 1994) of different approaches [*Cohee and Beroza*, 1994, and references therein] to invert for the temporal variables. In particular they investigated the differences in source inversions that use single and multiple time windows. Some of their salient results include (1) single-window methods allow larger variations in rupture time with fewer model parameters; (2) multiple-window methods are more flexible in that they allow for spatially variable rise time as well as rupture time with the caveat that the solutions are less stable; (3) both methods work well when the rise time is short compared to the periods of interest; (4) while the slip distribution may be similar for both methods, there are important differences in the rupture propagation models. *Cohee and Beroza* [1994] and *Hartzell* [1989] both found that adding multiple time windows commensurately added to the seismic moment.

[6] *Olson and Anderson* [1988], using synthetic data, took a different approach: invert for the kinematic variables in the frequency domain [*Spudich*, 1980]. In such a method the rupture time and slip function are completely arbitrary and to be determined from the data. To our knowledge this method has not been applied to data. *Cotton and Campillo* [1995] slightly modified this method by assuming a slip distribution that was parameterized by its rise time, each point on the fault could slip only once and the rupture was unilateral. Using this approach they successfully inverted near-source ground motion data from the 1992 Landers earthquake. More recently global inversion methods have been introduced that simultaneously determine slip amplitudes, rupture time, and rise time [e.g., *Hartzell et al.*, 1996; *Zeng and Anderson*, 1996; *Ji et al.*, 2002].

[7] Although various numerical methods have been developed for the simulation of wave propagation in complex media, the synthetic Green's functions used in previous finite fault inversions have been calculated from 1-D (flat-layered) homogeneous velocity models that did not account the more realistic heterogeneous structure. Because the seismogram is a convolution of the source and path effects, the path effects that are unaccounted for can be mapped into the source parameters. While 1-D approximations can be justified in certain areas, the rupture process determined from such inversions may be significantly influenced by the omission of three-dimensional (3-D) structural effects [*Hartzell*, 1989]. Given the recent advent of efficient 3-D numerical wave propagation methods and improved knowledge about the 3-D structure model in some active seismic zones [e.g. *Brocher et al.*, 1997; *Hauksson and Haase*, 1997; *Magistrale et al.*, 2000], *Liu and Archuleta* [1999] used 3-D Green's functions to invert for the kinematic parameters on a finite fault. *Graves and Wald* [2001] and *Wald and Graves* [2001] inverted synthetic data to examine source parameters resulting from both 3-D and 1-D Green's functions. They found that inaccurate Green's functions, either 3-D or 1-D, produce incomplete source descriptions. As noted by *Hartzell* [1989] this is the problem of resolution. We cannot know how different the approximate model is from a true model, and thus we cannot "make an accurate determination of resolution of the source" [*Hartzell*, 1989]. He further added "inversion theory and source modeling have advanced to the point where the Green's functions are the weakest link in the analysis." Using synthetics, *Graves and Wald* [2001] similarly showed that because a 3-D Green's function is used the resolution is not necessarily improved. *Ramos-Martínez and McMechan* [2001] inverted full waveforms for point source descriptions of Northridge aftershocks. They found that that using the 3-D structure of *Magistrale et al.* [1998] reduced the residuals and significantly improved the uniqueness of the source compared to a 1-D model. They also conclude that there is still room for lots of improvement in the 3-D structural model that will lead to better results [e.g., *Hartzell*, 1989; *Graves and Wald*, 2001]. Given that the Earth structure will never be perfectly known, we are left to use structural models that approximate the actual earth as well as it is known.

[8] While determining an accurate structural approximation to the earth remains a fundamental problem that is beyond the intent of this paper, even the choice of the subfault size can have a significant effect on a finite-fault

inversion [*Hartzell and Langer*, 1993; *Das and Suhadolc*, 1996]. In most finite fault inversions, a source parameter—such as slip amplitude, rake, or rise time—is almost always taken to be a constant within a subfault. This may be the primary reason that solutions are dependent on the subfault size. What is often left unstated is that the subfault is often subdivided by factors of 10 or more in order to compute the synthetics that are compared with the data. If so, the spatial variation of slip amplitude and rise time within a subfault cannot be neglected in the inversion. As shown later (equation (1)) a synthetic seismogram is represented as the summation of many point sources. To simulate accurately the ground motion in the opposite direction of the rupture, the distance between two point sources should be less than the rupture velocity divided by eight times the maximum frequency of interest [e.g., *Spudich and Archuleta*, 1987]. In a forward simulation (described later) we show that the interval between two adjacent point sources should be less than one-tenth the shortest wavelength. This limitation on the spacing of point source comes from the fact that the spatial variation of source parameters, especially the rupture time and rise time, will generate high frequency synthetic ground motion. In a finite fault inversion, the dimension of subfault is often more than half of the minimum wave length of interest. Thus one can expect that variation of source parameters within a subfault will affect the synthetic ground motions; these in turn will affect the resulting determination of kinematic parameters. In fact, most finite fault inversion methods approximate the variation of rupture time by summing the Green's function with a time delay that is calculated using a constant rupture velocity across a subfault.

[9] Another factor to be considered is that the source parameters of a particular subfault affect only the seismogram radiated from this subfault. Thus parameters among neighboring subfaults may have no correlation. To reduce this incoherency one can smooth or constrain the variation of parameters between neighboring subfaults [e.g., *Olson and Apsel*, 1982; *Hartzell and Heaton*, 1983; *Wald et al.*, 1991; *Sekiguchi et al.*, 2000]. The smoothing of source parameters enforces a kind of spatial correlation among source parameters during the inversion process. The difficulty of implementing a smoothing criterion is determining the relative weight between the waveform fit and the smoothing criterion, which typically varies from problem to problem [*Sekiguchi et al.*, 2000] and is often subjective in evaluating the tradeoff between resolution and the fit between synthetic and recorded waveforms [*Graves and Wald*, 2001]. On the other hand, the requirement of smoothing in a finite fault inversion implies that the spatial incoherence of source parameters tends to reduce the stability of the inversion. Consequently, it is necessary to find an algorithm that mitigates the effect of the subfault size on the solution.

[10] The basic nature of the kinematic parameters and their relationship to the data presents its own problem. The two slip components (one in the dip direction, another for the strike direction) are linearly related to the ground motion. They can be conveniently determined using linear least squares. However as mentioned earlier the other source parameters, such as rupture time and rise time that are essential for fully describing the kinematics of the rupture, are nonlinearly related to the data [*Archuleta*, 1984]. Although a linearized, iterative technique can be used to solve the problem [*Tarantola and Valette*, 1982], the solution may depend on the starting model [*Cotton and Campillo*, 1995]. This feature of parameterization compels us to look for nonlinear inversion methods that can be applied to determining the kinematic parameters of the rupture [*Liu et al.*, 1995; *Hartzell and Liu*, 1996; *Hartzell et al.*, 1996; *Zeng and Anderson*, 1996; *Ji et al.*, 2002].

[11] In this paper we present a new procedure to invert data for the kinematic parameters that reflect an earthquake rupture on a finite fault. First, we describe how we consider the effects of complex geology on finite inversions by using Green's functions calculated from a 3-D velocity structure. Second, we introduce a bilinear interpolation technique to represent variation of source parameters within a subfault area. Third, using this bilinear fault parameterization, we develop a global inversion method for efficiently determining source parameters from inversion of ground motion data. Finally, to demonstrate the efficiency of bilinear interpolation technique and to study the effect of velocity structure on the inversion solution, we reanalyze the 1989 **M** 6.9 Loma Prieta earthquake.

2. Calculation of 3-D Green's Functions

[12] The 3-D seismic wave propagation problem is currently solved by discrete numerical methods such as the finite-difference, finite-element, spectral-element and pseudospectral methods. In general there are two ways to calculate the 3-D Green's functions. One can apply an impulsive slip (or equivalent body force) at a point on the fault surface and the Green's functions are obtained at various observers for each applied force. If a Green's function is required at *NP* points over a finite fault, the total number of computations will be $2 \times NP$, regardless the number of observers. (The factor of 2 arises because each point on the fault needs a computation for each of the two slip components.) Alternatively, one can use reciprocity where a single body force is applied at the observer location and a Green's function (actually a traction) is evaluated everywhere on the fault [*Day*, 1977; *Archuleta and Day*, 1980; *Spudich and Archuleta*, 1987]. In this case three numerical calculations must be performed for each observer location (one for each component of ground motion). This approach requires $3 \times NS$ computations to obtain all of the Green's functions where *NS* is the number of stations.

[13] Given that the value of *NP* is much greater than *NS*, we choose the second method to calculate the Green's functions. Thus using reciprocity we can reduce the computational burden dramatically [*Spudich and Archuleta*, 1987]. Moreover, the displacement $u_m(\omega)$ at a station due to the slip discontinuity across fault plane Σ can be simply represented as

$$u_m(\omega) = \int\int_\Sigma \Delta u_s(\mathbf{x},\omega) \cdot \left[\cos(\lambda(\mathbf{x}))T_1^m(\mathbf{x},\omega) + \sin(\lambda(\mathbf{x}))T_2^m(\mathbf{x},\omega)\right]d\Sigma, \tag{1}$$

where $\Delta u_s(\mathbf{x}, \omega)$ is the slip discontinuity (dislocation) of point $\mathbf{x} = (x, y)$ on Σ; λ is the rake angle of the dislocation;

Table 1. Velocity Model

V_P, km/s	V_Sm km/s	Density g/cm^3	Q_P	Q_S	Thickness, km
4.0	2.0	2.6	180	100	1.0
6.0	3.464	2.7	250	150	...

(T_1^m, T_2^m) is the traction vector on Σ caused by a point force applied in the mth direction at the station [*Spudich and Archuleta*, 1987]; the subscripts 1 and 2 indicate strike and up-dip direction, respectively. Using the staggered-grid velocity-stress finite difference scheme this form is optimum for the Green's function calculations because we directly obtain the traction with the same accuracy as the particle velocity.

[14] We use a 3-D viscoelastic finite difference (FD) algorithm of *Liu and Archuleta* [1999]) to calculate the traction Green's functions in equation (1). The FD algorithm is accurate to fourth order in space and second order in time. It employs an explicit time-stepping, staggered-grid, velocity-stress formulation; it can simulate constant or frequency dependent Q. We first compute the six-components of stress at points on the fault caused by a body-force pulse inserted at a recording station, generally on the free surface. These stresses are then projected to the fault surface to obtain the two components of traction on the fault.

[15] To demonstrate the accuracy of reciprocal Green's functions, we compare the synthetic seismograms calculated using reciprocity of the representation theorem and the FD technique with seismograms obtained using a frequency wave number method. We use a point source double-couple with a strike of 30°, dip of 60°, rake of 50°, and at a depth of 1 km. The moment-rate function is $S(t) = M_0 t/T^2 \exp(-t/T)$, where $T = 0.1$ and $M_0 = 10^{18}$ Nm. The velocity structure model is laterally homogeneous (Table 1). The observation point is located at the surface at a distance of 10 km due east of the epicenter. In the FD calculation, the grid spacing is 0.1 km and the time step is 0.008 sec. With this grid spacing, the model has 6 grids per minimum shear wavelength at the upper frequency limit of 3.0 Hz. The velocity seismograms are low-pass filtered with a corner frequency of 3.0 Hz. The seismograms from the two methods are compared in Figure 1. The agreement between the waveforms is very good; there is a small difference in a later arriving wave generated by the artificial boundary in the FD simulation.

3. Finite Fault Parameterization

[16] Similar to a conventional finite fault inversion, our fault parameterization also requires discretizing a seismic fault plane into finite number of subfaults. However, our subfaults are quadrilateral elements, which are not necessarily rectangular and equal in area. That is, the shape and area of each quadrilateral can differ from others. We assign the unknown source parameters to the nodes (or corners) of the subfaults. We do not assume that the source parameters within a subfault are constant. The source parameters within each subfault are calculated by bilinear interpolating the four nodal quantities of the subfault.

[17] Inspired by the works of *Cotton and Campillo* [1995], *Hartzell et al.* [1996], and *Ji et al.* [2002], the temporal variation of slip discontinuity on the fault is represented by a functional form. We assume that the temporal behavior of a slip rate function is given by

$$\frac{\partial S(t,r,p)}{\partial t} = C\left(\frac{t}{r}\right)^p\left(1-\frac{t}{r}\right)^{5-p}, \quad 0 < t < r, \tag{2}$$

where r is rise time or duration of slip; C is a normalizing constant such that $S(t) = 1$ when $t \geq r$; and the exponent p controls the shape of slip rate function. The exponent p can vary from 1 to 4. In the particular case $p = 1$ this slip rate function has an amplitude spectrum that is practically identical to the Aki-Brune ω-squared model [*Aki*, 1967; *Brune*, 1970]. Using a slip rate function that requires only one or two parameters makes the inversion more stable as compared to the multiple time window technique that is frequently used [e.g., *Hartzell and Heaton*, 1983; *Wald and Heaton* [1994]; *Sekiguchi et al.*, 1996].

[18] Considering the limited amount of the information that can be extracted from ground motions we restrict the inversion to solve for a few spatial and temporal source parameters at all nodes: slip amplitude D, rake angle λ, rupture velocity c, rise time r, and/or the exponential p. Here the rupture velocity of a node is defined as the average speed of rupture from the hypocenter to this node. The local rupture velocity at a node may be significantly greater or less than the corresponding average velocity.

[19] By knowing the nodal values of source parameters, values of source parameters at any point within a subfault is calculated using bilinear interpolation. Without loss of generality, we use a subfault with an index e (left side, Figure 2) to explain the interpolation of a source parameter m^e. The subfault is a quadrilateral element defined by the location of its four nodal points (x_i^e, y_i^e). The values of a source parameter m^e at node i is m_i^e, where $i = 1, \ldots, 4$. If the quadrilateral element is rectangular, the bilinear interpolating can be directly applied to calculate the value of m^e at any

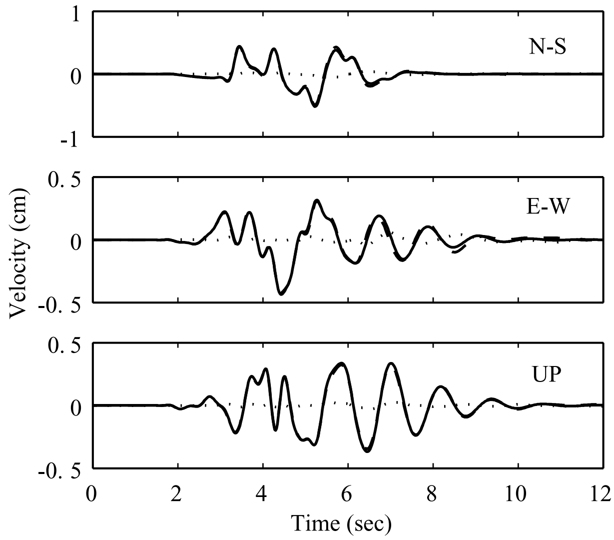

Figure 1. Comparison of three-component synthetic velocity seismograms calculated with frequency-wave number technique (solid line) and 3-D finite difference technique (dashed line). The difference between the waveforms from the two techniques (dotted line) is very small.

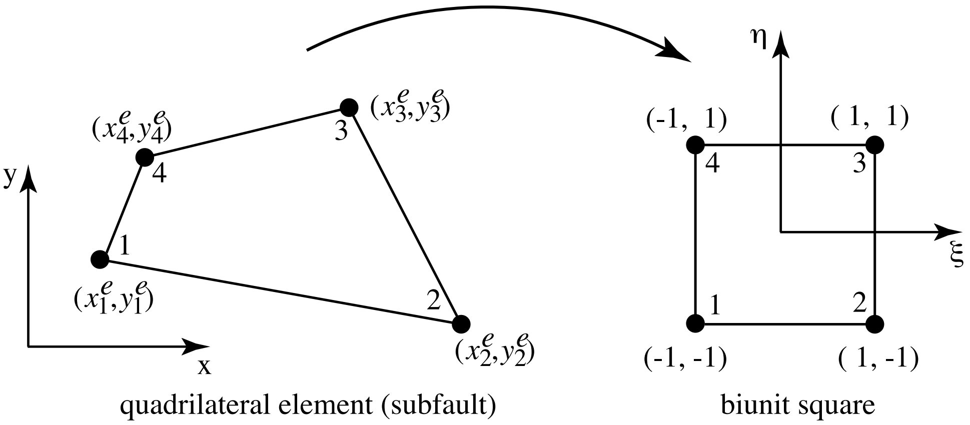

Figure 2. A change of coordinates that maps the given quadrilateral element (subfault) into the bi-unit square. The solid circles represent the nodes of subfault. The number near a node denotes local node ordering.

point (x^e, y^e) within the subfault. In the general case, we borrow the concept of isoparametric elements used in the finite element method [*Hughes*, 1987] to map a given quadrilateral into a bi-unit square shown at the right of Figure 2. The point (x^e, y^e) is first related to a point (ξ, η) in the bi-unit square as follows:

$$\begin{aligned} x^e(\xi,\eta) &= \sum_{i=1}^{4} N_i(\xi,\eta) x_i^e \\ y^e(\xi,\eta) &= \sum_{i=1}^{4} N_i(\xi,\eta) y_i^e \end{aligned}, \tag{3}$$

where N_i are the shape functions, and has the form

$$\begin{aligned} N_1(\xi,\eta) &= \frac{1}{4}(1-\xi)(1-\eta) \quad N_2(\xi,\eta) = \frac{1}{4}(1+\xi)(1-\eta) \\ N_3(\xi,\eta) &= \frac{1}{4}(1+\xi)(1+\eta) \quad N_4(\xi,\eta) = \frac{1}{4}(1-\xi)(1+\eta) \end{aligned}. \tag{4}$$

The value of m^e at point (x^e, y^e) is then calculated by

$$m^e(x^e, y^e) = \sum_{i=1}^{4} N_i(\xi,\eta) m_i^e. \tag{5}$$

For a point (ξ, η) in the bi-unit square, the coordinate (x^e, y^e) of the corresponding point in subfault e is determined by equation (3). Values of the source parameters at point (x^e, y^e) are obtained through bilinear interpolation according to equation (5).

[20] Following the basic assumption in seismology that a large earthquake can be simulated by a distributed set of point dislocations that cover the fault, we assign N_ξ points of equal spacing along the ξ direction, and N_η points along η. Consequently every subfault contains $N_\xi N_\eta$ interior point sources. Synthetics are computed for every interior point and summed to produce a seismogram from each subfault. Thus equation (1) can be approximated as

$$\begin{aligned} u(\omega) \approx \frac{1}{N_\xi N_\eta} & \sum_{e=1}^{N_e} \sum_{j=1}^{N_\xi} \sum_{k=1}^{N_\eta} D_{jk}^e S\left(r_{jk}^e, p_{jk}^e, \omega\right) \exp\left(-i\omega\tau_{jk}^e\right) \\ & \cdot \left[\cos\left(\lambda_{jk}^e\right) T_1\left(x_{jk}^e, y_{jk}^e, \omega\right) + \sin\left(\lambda_{jk}^e\right) T_2\left(x_{jk}^e, y_{jk}^e, \omega\right)\right], \\ & \cdot J^e\left(\xi_j, \eta_k\right) \end{aligned} \tag{6}$$

where the coordinate (ξ_j, η_k) of a point in a bi-unit square is

$$\xi_j = \frac{2j-1}{N_\xi} - 1, \quad \eta_j = \frac{2k-1}{N_\eta} - 1 \tag{7}$$

and the value of the Jacobian determinant $J^e(\xi_j, \eta_k)$ is given explicitly by

$$J^e(\xi,\eta) = \det \begin{bmatrix} \sum_{i=1}^{4} \frac{\partial N_i}{\partial \xi} x_i^e & \sum_{i=1}^{4} \frac{\partial N_i}{\partial \xi} y_i^e \\ \sum_{i=1}^{4} \frac{\partial N_i}{\partial \eta} x_i^e & \sum_{i=1}^{4} \frac{\partial N_i}{\partial \eta} y_i^e \end{bmatrix}. \tag{8}$$

N_e is number of subfaults; D_{jk}^e is the dislocation amplitude at the point (x_{jk}^e, y_{jk}^e) within a subfault e, λ_{jk}^e is the rake, τ_{jk}^e is the rupture time (the distance between this point and hypocenter divided by the rupture velocity c_{jk}^e), r_{jk}^e is the rise time and p_{jk}^e is the exponent that determines the shape of the slip function S. The source parameters are evaluated using equation (5). Normally, the Green's functions at point (x_{jk}^e, y_{jk}^e) are also calculated through bilinear interpolation. The accuracy of equation (6) is highly dependent upon frequency and upon the observer's position with respect to the rupture propagation direction [*Spudich and Archuleta*, 1987]. Our numerical tests demonstrate that the interval between two adjacent point sources should be less than one-tenth of the shortest wavelength.

[21] Although the spatial distributions of source parameters are not smooth at the interface of two adjacent subfaults, they are continuous within the whole fault. The seismograms radiated from a subfault depend on the source parameters at four nodes rather than just one. Moreover, a change in the parameters at a node also affects the synthetics resulting from the adjacent subfaults because a node normally connects four subfaults. Thus the bilinear interpolation enforces a spatial correlation among source parameters in our finite fault inversion. The efficiency of this fault parameterization is demonstrated in the following section by applying it to analysis of the Loma Prieta earthquake.

4. Inversion Procedure

[22] Source parameters at the nodes of subfaults are determined by minimizing a misfit or objective function

that compares synthetic time histories with the data. In the following, we first discuss the choice of a suitable objective function and then describe an inversion method that minimizes the objective function by finding the optimal set of source parameters.

[23] There is a wide range of objective functions that have been used in the process of comparing the synthetics with the data [e.g., *Hartzell et al.*, 1991, 1996; *Zeng and Anderson*, 1996]. Minimizing the L_2 norm is the most popular approach used in inverse problems because it leads to the easiest computations—a least squares problem. The basic problem is that least squares solutions are not robust, i.e., solutions are very sensitive to a small number of large errors in the data set [*Tarantola*, 1987]. Following *Sen and Stoffa* [1991] and *Hartzell et al.* [1996], we define our objective function e as

$$E(\mathbf{M}) = \sum_{1}^{N_d} W_d \left(1 - \frac{\sum_{\omega_b}^{\omega_e} \left(u_o(\omega)u_s^*(\omega) + u_o^*(\omega)u_s(\omega)\right)}{\sum_{\omega_b}^{\omega_e} u_o(\omega)u_o^*(\omega) + \sum_{\omega_b}^{\omega_e} u_s(\omega)u_s^*(\omega)} \right) + W_c(\text{constraints}), \tag{9a}$$

where $\mathbf{M}$ is the vector of model parameters, N_d is the number of data records, W_d is a weight reflecting the quality of the data record, (ω_b, ω_e) is frequency band in which we compare synthetics $u_s(\omega)$ with observed data $u_o(\omega)$, and the asterisk indicates complex conjugate. Two types of constraints are chosen: one that minimizes the difference between slip on adjacent subfaults; and one that minimizes the total moment [*Hartzell et al.*, 1996]. Using additional constraints increases the stability of inversion but reduces the resolution. The weight W_c is used to adjust the trade-off between fitting the data and satisfying the constraints. This value is obtained by trial-and-error to ensure that the fits to data are not strongly degraded. The first part of this objective function is equivalent to the L_2 norm divided by the sum of squares of the data and the synthetics. It is a tradeoff between cross-correlation and L_2 norm. Besides considering signal shape as cross-correlation does, this objective function also uses the amplitude information of signal like least squares. However, this objective function is less sensitive to the amplitudes than the standard least squares.

[24] Because it is very fast to compute the convolution of source functions and Green's functions synthetics in frequency domain, it is more efficient to compare the synthetic data with observations using equation (9a). There are situations where it is more desirable to make the comparison in the time domain. One good reason is that we can easily limit the time window used in inversion to the direct arriving waves in order to mitigate the stronger influence of the uncertainties in the velocity structure on later arrivals. In the time domain the objective function given in equation (9a) has the form:

$$E(\mathbf{M}) = \sum_{1}^{N_d} W_d \left(1 - \frac{2\sum_{t_b}^{t_e} \left(u_o(t)u_s(t)\right)}{\sum_{t_b}^{t_e} u_o^2(t) + \sum_{t_b}^{t_e} u_s^2(t)} \right) + W_c(\text{constraints}), \tag{9b}$$

where (t_b, t_e) is the time window for the inversion. The synthetic time history $u_s(t)$ is calculated from the inverse Fourier transform of $u_s(\omega)$.

```
Start with a random source model
Loop over number of iterations/temperature (k=1, …, K)
    Loop over number of nodes (n=1, …, N)
        Loop over number of perturbations (p=1, …, P)
            generate source parameters for node n, and
            evaluate the objective function
        End loop
        Select a model to update current model
    End loop
End loop
```

Figure 3. A pseudo Fortran code for the developed simulated annealing algorithm.

[25] *Hartzell et al.* [1996] applied the hybrid global search algorithm of *Liu et al.* [1995] to invert for the two slip amplitudes, rupture time, and rise time. This algorithm perturbs all of the source parameters at the same time. On the basis of our numerical tests, we find that the efficiency of this algorithm decreases dramatically when a fault has more than 100 subfaults. *Ji et al.* [2002] took another approach. They chose a particular kind of the simulated annealing method (SA), called the heat-bath algorithm [*Rothman*, 1986], to search for the best model. This algorithm acts by perturbing the model parameters one by one. As indicated by *Ji et al.* [2002], synthetic seismograms from only one subfault need to be updated at each perturbation. Thus, the calculation of objective function can be very fast. Another advantage of this method is that it is well suited for problems with a large number of free parameters [*Sen and Stoffa*, 1995].

[26] Using our fault parameterization, we have developed a simulated annealing algorithm by combining the algorithm of *Liu et al.* [1995] with the heat-bath algorithm used by *Ji et al.* [2002]. In order to describe the algorithm, we assume that the source model vector $\mathbf{M}$ consists of N nodes, and each node has I model parameters, thus $\mathbf{M}$ = ($m_{1,1},\ldots,m_{I,1},m_{1,2},\ldots,m_{I,2},\ldots,m_{1,N},\ldots,m_{I,N}$).

[27] Following the analogy with annealing in a thermodynamic system, this algorithm is initiated at a suitably high temperature and then allowed to cool slowly according to an annealing schedule. At each temperature, the algorithm is designed to search the entire model space node-by-node. This search procedure is similar to the heat-bath algorithm. However, this method simultaneously perturbs all of the model parameters at one node. The range of the perturbations is proportional to the temperature. After perturbing the current model parameters at a node has generated a prespecified number of new models, one of the new models is selected to replace the current model, and the search procedure moves to the next node. A schematic diagram of the algorithm is given in Figure 3. The details of each step of the algorithm are described below.

[28] The algorithm starts with a random model whose parameters are initialized by

$$m_{i,n} = m_{i,n}^{\min} + \xi_{i,n}\left(m_{i,n}^{\max} - m_{i,n}^{\min}\right), \tag{10}$$

where $m_{i,n}$ is the ith source parameter of current model at node n; $m_{i,n}^{\min}$ and $m_{i,n}^{\max}$ are predetermined minimum and maximum constrains on $m_{i,n}$. A random number $\xi_{i,n}$ is uniformly distributed in the interval (0, 1). This initialization allows the model parameters to range uniformly over the preset model space we wish consider. Ideally, we would like to initiate the algorithm at a high initial temperature T_0, and then cool very slowly until the model converges to the global minimum. However, this process may require enormous computing time. In practice we choose an exponential cooling schedule suggested by *Rothman* [1986],

$$T_k = T_0 \eta^k, \tag{11}$$

where T_k is the temperature at iteration k; η is cooling rate. Iterations continue until some predetermined minimum value of objective function E is reached or until a preset number of iterations are completed. In practice, we do not know a priori the minimum value of the objective function; we use a preset iteration number K as convergence criterion of the algorithm. The choice of initial temperature and cooling rate, as well as the number of iterations, is problem dependent and is discussed below.

[29] At each iteration or temperature T_k the search process examines the subfault nodes sequentially. When a node (say node n) is visited, the current model parameters at this node are perturbed repeatedly to obtain a number of new models $\mathbf{M}_n^p$, which are the same as the current model $\mathbf{M}$ except that at node n the new model parameters are given by

$$m_{n,i}^p = m_{i,n} + y_k^p v_i^p \left(m_{n,i}^{\max} - m_{n,i}^{\min} \right); \quad i = 1, \cdots, I, \tag{12}$$

where

$$y_k^p = T_k \tan(\pi(\alpha^p - 0.5))$$
$$v_i^p = \frac{\beta_i^p}{\left(\sum_{j=1}^{M} \left(\beta_j^p \right)^2 \right)^{1/2}} \quad . \tag{13}$$

The superscript p refers to perturbation number; $m_{i,n}$ are the current model parameters; α^p and β_i^p in equation (13) are random numbers distributed uniformly in the interval (0, 1) and $(-1, 1)$, respectively. The vector $(v_1^p, \ldots, v_I^p)$ determines the perturbation direction. The perturbation distance y_k^p depends on the temperature T_k and follows a one-dimensional Cauchy distribution. Equation (12) is very similar to the perturbation form used by *Lane* [1992] and *Liu et al.* [1995]. It yields a perturbation to the current model (at one node) that is uniformly distributed in direction and Cauchy distributed in distance. In addition, the flat tail of Cauchy distribution allows the inversion algorithm to escape relatively easily from local minima in the search for the desired global minimum. To limit the search to the predetermined model space, the new model parameter $m_{i,n}^p$ of $\mathbf{M}_n^p$ must lie on the interval $(m_{i,n}^{\min}, m_{i,n}^{\max})$. If not, a new random variable α^p is generated for this parameter until the constraint is satisfied.

[30] After the perturbations to the current model parameters at node n are repeated P times, we get P new models $\mathbf{M}_n^p$ and corresponding objective functions $\mathrm{E}(\mathbf{M}_n^p)$. One of the new models is then selected to update the current model. To do this, first we calculate the change in objective function Δe_n^p from the current model $\mathbf{M}$ to the new model $\mathbf{M}_n^p$

$$\Delta E_n^p = E\left(\mathbf{M}_n^p\right) - E(\mathbf{M}), \tag{14}$$

and we calculate the average value $\overline{\Delta E_n}$,

$$\overline{\Delta E_n} = \sum_{p=1}^{P} \mathrm{abs}\left(\Delta E_n^p\right). \tag{15}$$

Next we evaluate following cumulative probability C_p of new model $\mathbf{M}_n^p$

$$C_p = \frac{\sum_{j=1}^{p} \exp\left(-\frac{E\left(\mathbf{M}_n^j\right)}{\Delta E_a}\right)}{\sum_{j=1}^{P} \exp\left(-\frac{E\left(\mathbf{M}_n^j\right)}{\Delta E_a}\right)} = \frac{\sum_{j=1}^{p} \exp\left(-\frac{\Delta E_n^j}{\Delta E_a}\right)}{\sum_{j=1}^{P} \exp\left(-\frac{\Delta E_n^j}{\Delta E_a}\right)}, \tag{16}$$

where ΔE_a is the average of $\overline{\Delta E_n}$ over all the nodes during previous iteration. Finally, we draw a random number γ from uniform distribution between 0 and 1. At the point j where

$$C_{j-1} \leq \gamma < C_j, \tag{17}$$

a new model $\mathbf{M}_n^j$ is picked to replace the current model. Thus $\mathbf{M} = \mathbf{M}_n^j$, and $\mathrm{E}(\mathbf{M}) = \mathrm{E}(\mathbf{M}_n^p)$. Here we assume that $C_0 = 0$. In a complete iteration the current model is updated N times because each node is examined once. At high temperature, the strong perturbations result in a large value of ΔE_a and therefore, it is nearly random when selecting a new model according (equations (16) and (17)). As the temperature is lowered the model with the smallest value of the objective function has a higher probability of being selected.

[31] We have described in detail each element of the algorithm. However, to implement the algorithm, there are still several free parameters that need to be chosen properly: the initial temperature T_0, cooling rate η, the number of perturbations P, and the number of iterations K. Often physical insight and/or a significant amount of trial-and-error runs are required to choose the values of these parameters. Normally, the number of perturbations should be at least twenty times larger than the number of source parameters at a node, i.e., $P \approx 20I$. For most cases that we have explored, teleseismic and near-source inversions, an initial temperature in the range 0.1–0.2 is sufficient. We find that it is more efficient to preset a final temperature T_f from which a cooling rate can be determined using equation (10):

$$\eta = \left(\frac{T_f}{T_0}\right)^{\frac{1}{K}},$$

where the final temperature T_f is in the range 0.001–0.005.

[32] Our inversion procedure is one kind of global optimization algorithm. Unlike a local inversion method, such as linearized iterative least squares, it has the ability to escape from local minima to find a global minimum.

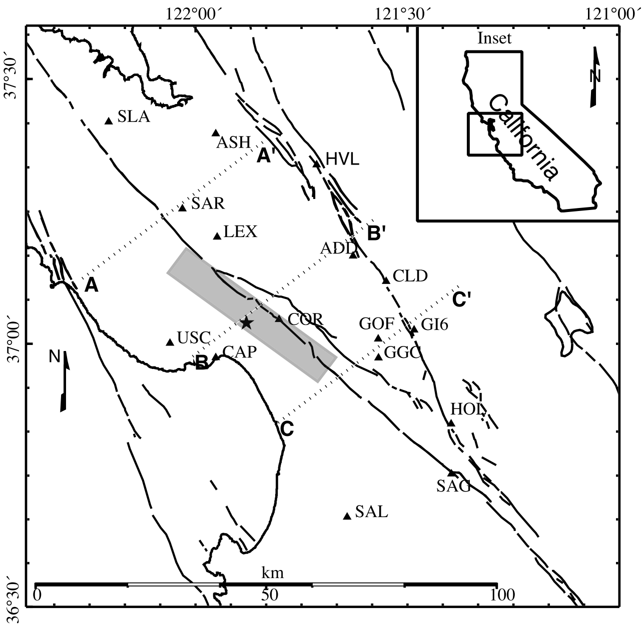

Figure 4. Map of the local region surrounding the Loma Prieta earthquake. Inset shows the local area with respect to the state of California. Fault projection to the Earth's surface is shown as the shaded area. A star denotes epicenter location. Stations used to invert for the kinematic parameters of faulting are shown as triangles. The locations of three cross sections A-A′, B-B′ and C-C′ (Figure 5) are indicated by dashed lines.

However, the cost for this desirable feature is that the forward problem—which consists mainly of computing synthetic seismograms—must be calculated a large number of times. For example, in one iteration of the algorithm the objective function is evaluated $N \times P$ times. Fortunately, the computation of synthetic seismograms can be very fast in our inversion procedure. As we pointed out earlier, when a node is examined, only the current source parameters of this node are perturbed P times. This node connects two to four subfaults whose synthetics need to be recomputed each time. The seismograms from the other subfaults are not changed and can be reused during the perturbations. Therefore, only a few computations are required to get the synthetic seismograms from each newly generated model.

5. Application of the Inversion Procedure to the Loma Prieta Earthquake

[33] The 1989 Loma Prieta **M** 6.9 earthquake occurred in a region that was fairly well instrumented with strong-motion accelerographs. Because strong-motion instruments record data in the near-source region where Green's functions from different parts of the fault can be very different, these data have the potential to resolve detailed aspects of the rupture process. In addition, a large number of teleseismic stations recorded this earthquake. The abundance of high-quality data provided one of the best opportunities to study the rupture process of Loma Prieta earthquake. *Beroza* [1995] compared four finite fault source models of Loma Prieta earthquake [*Beroza*, 1991; *Hartzell et al.*, 1991; *Steidl et al.*, 1991; *Wald et al.*, 1991]. Each of the studies finds low slip in the hypocentral region and generally a bimodal distribution of slip with respect to the hypocenter. The seismic moment ranges from 2.3×10^{19} to 3.5×10^{19} N-m, spanning approximately the same range determined teleseismically at longer periods. *Beroza* [1995] also pointed out the differences among the four models: the location of the high-slip areas and, especially the rake differ among the four models. He attributed the substantial bias in the rake to the inadequacy of the synthetic Green's functions due to lateral variations in the velocity structure, and the differences in the slip distributions to different data sets and frequency bands used for the inversions. We suspect that the different velocity structures used in these four studies may be another reason for the differences [*Stidham et al.*, 1999]. To address this issue and to demonstrate the applicability of the new inversion procedure, we re-analyze the data from the Loma Prieta earthquake.

[34] Because the FD technique requires a large computational effort, we selected only 16 stations (Figure 4 and Table 2) in the near-source region of the earthquake. To

Table 2. Stations Used to Invert for Kinematic Parameters

Station	Latitude, °N	Longitude °W	Epicentral Distance, km	Owner	TTimes-OT[a]
COR Coralitos	37.046	121.803	5	CDMG	5.2
CAP Capitola	36.974	121.952	10	CDMG	NAT
UCS U.C. Santa Cruz	37.011	122.060	15	CDMG	NAT
GOF Gilroy-Old Firehouse	37.009	121.569	30	CDMG	8.2
GGC Gilroy-Gavilan College	36.973	121.568	30	CDMG	NAT
GI6 Gilroy Array #6	37.026	121.484	35	CDMG	10.8
SAR Saratoga	37.255	122.031	30	CDMG	NAT
HOL Hollister (South and Pine St.)	36.848	121.397	50	CDMG	12.3
ASH Agnews State Hospital	37.397	121.952	45	CDMG	9.0
ADD Anderson Dam-Downstream	37.166	121.628	30	USGS	7.8
SLA Stanford-SLAC	37.419	122.205	50	USGA	NAT
LEX Lexington Dam-Abutment	37.202	121.949	20	CDMG	5.9
CLD Coyote Lake Dam-Downstream	37.118	121.550	30	CDMG	9.3
HVL Halls Valley-Grant Park	37.338	121.714	35	CDMG	9.7
SAG SAGO South-Hollister	36.758	121.396	50	CDMG	15.0
SAL Salinas	36.671	121.642	45	CDMG	10.7

[a]Definitions are as follows: OT, origin time at 00 04 15.21 GMT 18 October 1989; Ttime, trigger time of the instrument; NAT, no absolute time.

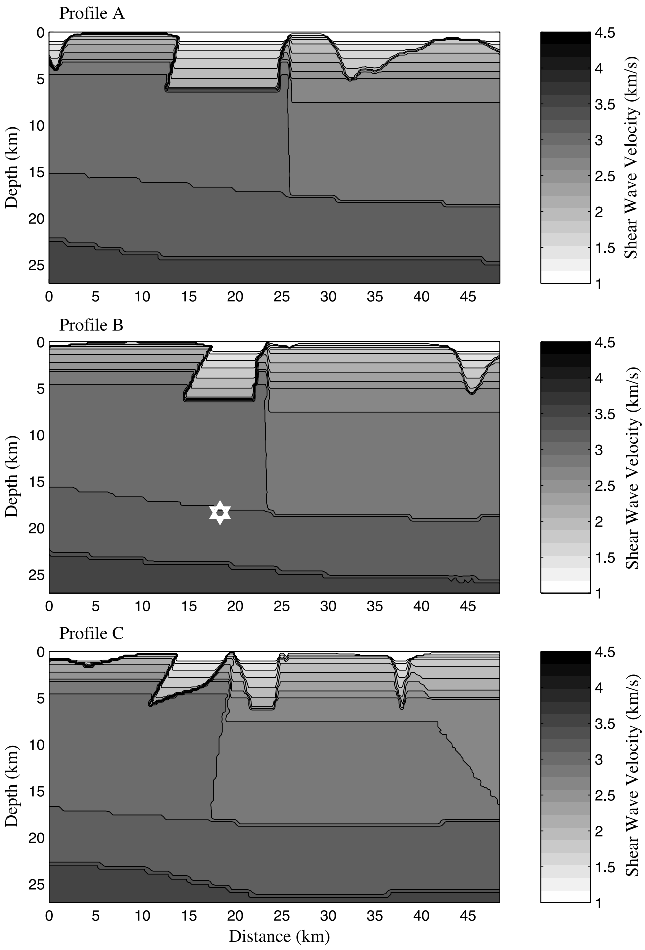

Figure 5. Cross sections of shear wave velocity for three profiles indicated on Figure 4. In Profile B-B′, the hypocenter is shown as a white star. One can see the difference in velocity across the San Andreas to a depth of about 20 km in each profile. The local basins penetrate to depths about 5 km. Velocity structure is the latest USGS model (T. M. Brocher, personal communication). See color version of this figure at back of this issue.

examine the effect of the structure we invert these data using both 1-D and 3-D Green's functions. The 3-D Green's functions are calculated numerically using the 3-D velocity and *Q* structure provided by T. M. Brocher (1999 written communication). To illustrate the 3-D variability of the velocity structure we show three cross-sections in Figure 5. These Green's functions are computed using a fourth-order staggered grid FD code that includes new absorbing boundary conditions and coarse-grained attenuation [*Liu and Archuleta*, 1999]). We limit the maximum frequency to 1.0 Hz and the minimum S wave velocity to 1.0 km/s at the Earth's surface. Parameters of the 3-D FD calculation are given in Table 3. The 1-D Green's functions are calculated from the velocity model used by *Wald et al.* [1991]. Because there is no *Q* value in this 1-D model, we assume $Q_P = 2Q_S$ and $Q_S = 0.1\ V_S$, where the units of V_S are m/s. Both the observed data (particle velocities) and Green's functions are band passed in the frequency range 0.05–1.0 Hz. Some of the strong motion records do not have trigger time information. For theses stations the synthetic shear wave from hypocenter is aligned with the first impulsive S wave in the data.

[35] We chose the same geometry for the fault as *Wald et al.* [1991]—strike of 128°, dip of 70° to the southwest. The fault measures 41.25 km in length and extends from a depth of 1.5 km to 20.3 km, giving a down-dip width of 20 km. The hypocenter is at 37.04°N, 121.88°W, with a depth of 18 km. For simplicity the fault area is discretized into 15 rectangular elements along strike and 8 elements down-dip for a total of 120 subfaults of equal area (dimensions 2.75 km by 2.5 km). For each subfault we have 16 equally spaced Green's functions. These Green's functions are calculated by linearly interpolating from the eight nearest points for which the Green's functions were computed. These are further bilinearly interpolated so that each subfault is represented by 144 Green's functions. The corresponding source parameters are also calculated through bilinear interpolation. Synthetic seismograms are computed for 144 points for each subfault and summed over all subfaults to produce a synthetic time history.

[36] Each node has four model parameters: slip amplitude, rake angle, rise time, and rupture velocity. We fix the exponent *p* (equations (2) and (6)) at 1.5. Prior to inverting the data we set bounds for the model parameters: slip amplitude—0.0 to 7.0 m; rake angle—80° to 170°; rise-times—0.6 to 6.0 sec; rupture time is bounded by the time for the rupture to reach the node from the hypocenter with the rupture velocity bound between 2.2 km/s to 3.5 km/s. In this inversion, we set the W_c be zero, i.e., no smoothing or minimum moment constraints are applied. The inversion starts at a temperature of 0.2 and stops at a temperature of 0.005 with a maximum of 1500 iterations. For each node the perturbation number is 70.

[37] To ensure that the inversion results are comparable, both 1-D and 3-D inversions use exactly the same procedure, data, source parameterization and constraint conditions, except for small differences in timing shifts (~0.5 s) between the 1-D and 3-D Green's functions to account for the difference between the1-D and 3-D velocity structures. A two-second time shift is used to account for the triggering of the Loma Prieta earthquake by a small event that did not trigger the strong motion accelerographs [*Wald et al.*, 1991].

[38] The seismic moment from 1-D and 3-D inversion are 3.9×10^{19} and 4.3×10^{19} N-m, respectively. The larger moment from 3-D inversion can be attributed to larger attenuation in the 3-D structure. A comparison of the observed and synthetic ground velocities is given in

Table 3. Parameters Used in the Finite Difference Calculation of the 3-D Green's Functions

Parameter	Value
V_s (min)	1000 m/s
V_p (min)	1700 km/s
Q_s (min)	20
Q_p (min)	50
Density (min)	1600 kg/m^3
Grid spacing	180 m
Time step	0.02 s
Time window	20 s
Total time	35 s
$f_{\max}$ (Hz)	1.0

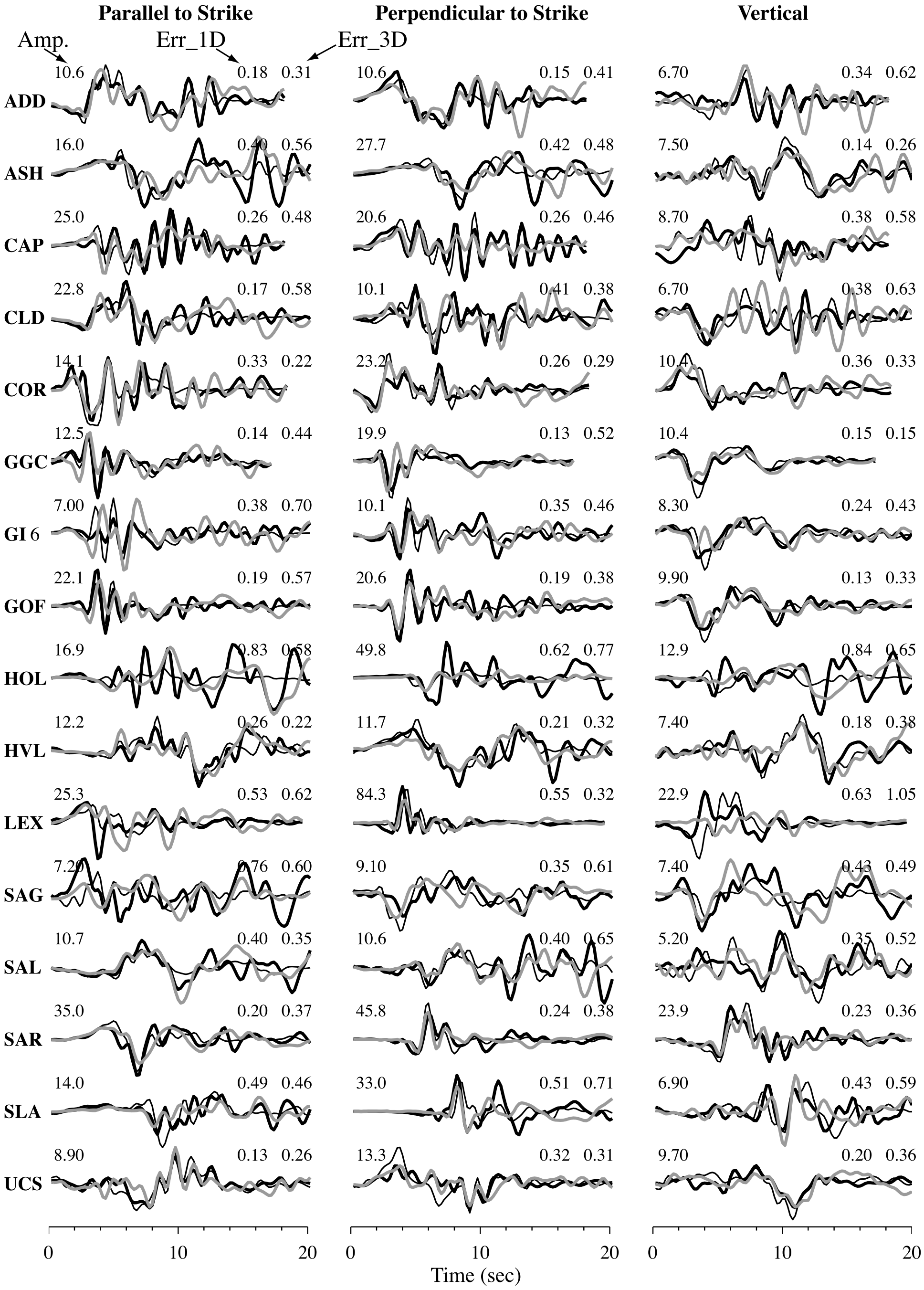

Figure 6. Comparison of observed velocity records (black thick lines) with the synthetics (gray thick lines) for the 3-D inversion and the synthetics (black thin lines) for 1-D inversion. Station names are indicated at the left. Peak amplitude of observation (cm/s) is given for each trace. Data and synthetics are plotted on the same vertical scales.

Figure 6. For each station we show the misfit between the data and the synthetic, both 1-D and 3-D, based on the objective function in equation (9a) (ignoring the additional term related to constraints). The synthetic seismograms from the 1-D inversion fit the data a slightly better than that from 3-D inversion at most stations. We think that the Q values given for the 3-D model are too low. In terms of total CPU time recomputing the 3-D Green's functions is no trivial matter. Nonetheless, it appears that the fit to the data could be improved with a higher Q in 3-D velocity structure for the near-surface material. It is clear that there are local effects that are not captured by even a 3-D model. There are a few stations where the misfit is consistently large for all three components, such as Hollister (HOL). As noted by *Wald et al.* [1991] the amplitude of the velocity on the perpendicular component is anomalous for the assumed fault geometry. They suggest additional faulting on the San Andreas southeast of the Loma Prieta fault zone. Lexington Dam (LEX) is well fit for its largest signal, perpendicular to strike, but substantially misfit on the

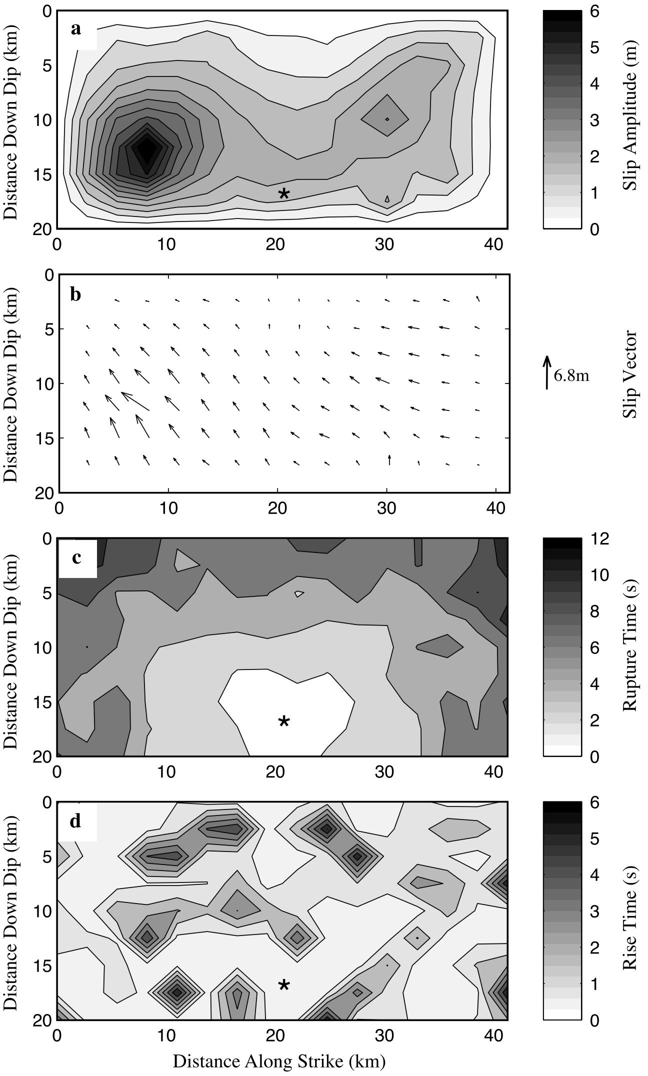

Figure 7. Plots of (a) slip amplitude at 0.5-m contour interval, (b) rake, (c) rupture time at 2-s intervals, and (d) risetime at 0.8-s contours for the inversion using 1-D Green's functions.

vertical and parallel components. The station SAR is about 5 km farther along strike from LEX, but all three components are fairly well matched by the synthetics. It may be that with the LEX station being sited on the abutment of the dam affects the ground motion. Neither a 1-D or 3-D velocity structure accounts for all the variation seen in the data.

[39] The spatial distributions of slip amplitude, rake, rupture time and rise time determined by 1-D and 3-D inversions are shown in Figures 7 and 8, respectively. Both inversions find a bimodal distribution of the slip with respect to the center of the fault. There is very little slip directly above the hypocenter. The two source models also have very similar distribution of rake angle: the asperity southeast of the hypocenter slips with a rake angle of ~135°; the northwest asperity has a rake angle of ~155°. In addition to the similarities, large differences in the two models can also be observed. In the 3-D model, the southeast asperity has larger slip amplitudes of ~4.0 m compared to the slip amplitudes ~3 m in the asperity northwest of the hypocenter. The 1-D model has almost a reversed picture. The maximum slip amplitudes in southeast asperity is only ~2.5 m, which is less than one half of the slip amplitudes ~5.5 m in the northwest asperity. The large differences in slip distributions between 1-D and 3-D source models can be attributed to lateral variations in velocity structure.

[40] The time when a point starts to rupture is equal to the distance between a point on the fault and the hypocenter divided by the rupture velocity (found through inversion) for this point. Contours of rupture times describe the position of the rupture front on the fault (Figures 7 and 8). In general the 3-D source model has a more heterogeneous rupture velocity than that found for the 1-D model, although both have a similar rupture velocity of 2.8 km/s. In general, rise times cannot be well resolved by these inversions because of the limited frequency band. Here we only considered the rise times in the regions where the slip

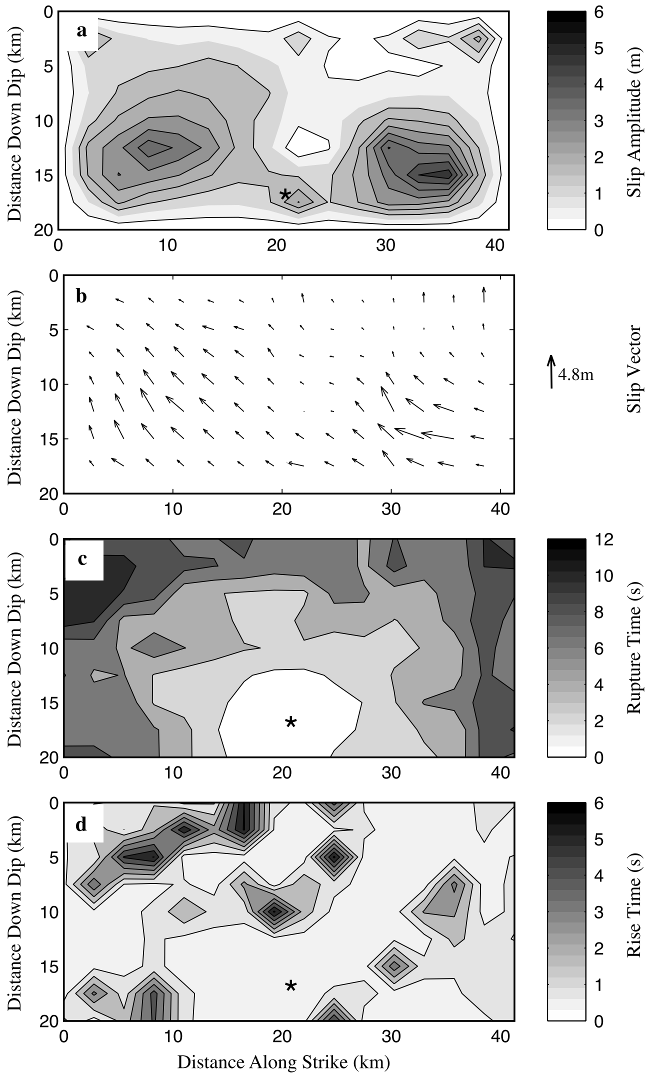

Figure 8. Plots of (a) slip amplitude at 0.5-m contour interval, (b) rake, (c) rupture time at 2-s intervals, and (d) risetime at 0.8-s contours for the inversion using 3-D Green's functions.

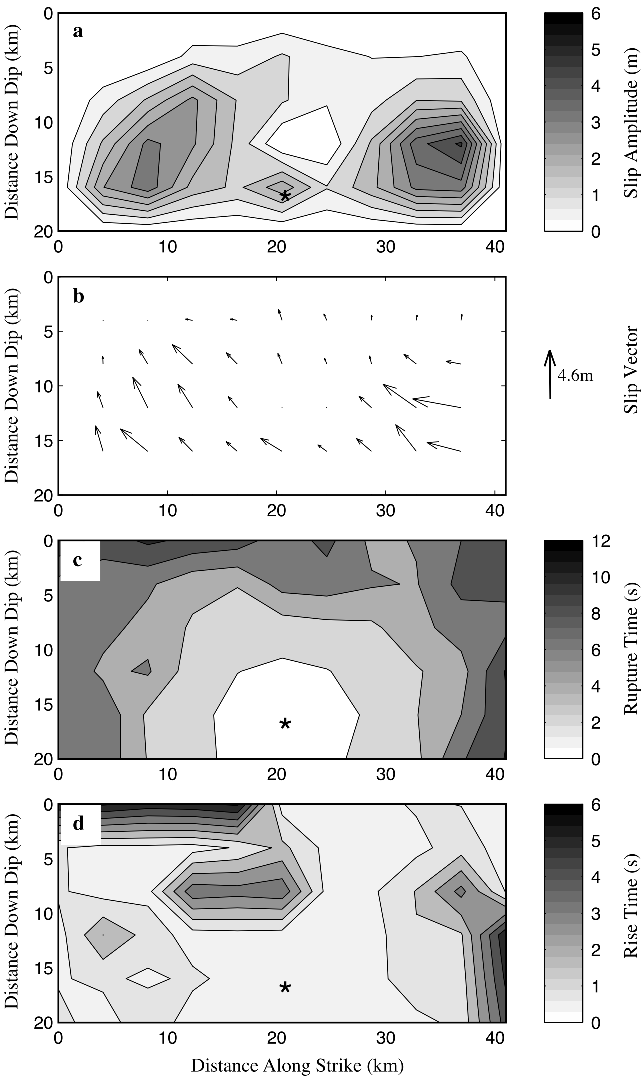

Figure 9. Plots of (a) slip amplitude at 0.5-m contour interval, (b) rake, (c) rupture time at 2-s intervals, and (d) risetime at 0.8-s contours for alternative 3-D inversion using 50 subfaults instead of 120.

amplitude is large: we find about 1.0 s rise time in the 3-D model and around 2.0 s in the 1-D model.

[41] In addition to comparisons with the 1-D inversion, we also ran another 3-D inversion to evaluate the performance of bilinear interpolation. In this inversion we use 50 subfaults (10×5) instead of 120. As shown in Figure 9, the 50-subfault model is very similar to the 120-subfault model (Figure 8). The 50-subfault model has smoother distributions of source parameters as a result of its larger subfault size. This comparison indicates that the bilinear interpolation is working to reduce the effect of subfault size on the inversion outcome.

6. Discussion

[42] While there has always been a strong effort to determine Earth structure at every scale, the existing models used in most inversions have significant uncertainty. Most inversions that rely on comparing waveforms automatically limit the maximum frequency that will even be attempted. Even at frequencies less than 1.0 Hz the Green's functions calculated from models cannot represent exactly wave propagation in the real Earth structure. Uncertainty in structure model will significantly affect the synthetic seismograms, in both phase and amplitude that in turn lead to the variation in the inverted source model. The basic issue is that the Earth structure is not well known but is almost certainly heterogeneous in all three dimensions [*Brocher et al.*, 1997; *Hauksson and Haase*, 1997; *Magistrale et al.*, 2000]. *Graves and Wald* [2001] examined the effect of 3-D versus 1-D Green's functions by computing synthetic seismograms from a hypothetical earthquake. They found that an inaccurate velocity structure could strongly bias the inverted slip distribution even when the rupture velocity, rise time and rake angle are fixed. *Ramos-Martínez and McMechan* [2001] found that the 3-D structural model for San Fernando was significantly better for determining a point source seismic moment, rake and dip than a 1-D model. Although the 3-D structure will not be perfectly known, we cannot assume that a 1-D approximation to the Earth structure is a better model. Regardless of the structural model being used, inaccuracies in the Green's functions will be mapped into the source parameters.

[43] In our inversion of data from the Loma Prieta earthquake, both 1-D and 3-D inversions used exactly same procedure, data, source parameterization and constraints, such that the difference in the two source models results only from the assumed structure models. While the overall pattern of the source model from the 1-D Green's function inversion is close to that found from 3-D inversion, the details of the two models are quite different (Figures 7 and 8). Regions with maximum slip amplitudes (as well as the rake) differ between the two models. We can also observe that the synthetic seismograms calculated separately from the two models fit the data equally well, although 1-D synthetics fit the data better at many stations. These results demonstrate that the global inversion method will find an optimal solution for a given set of Green's functions. However, the differences between the 1-D and 3-D solutions clearly demonstrate that some elements of source complexity result from inexact knowledge of the Greens functions. It can be expected that these effects will become more evident as the size of the fault and rupture duration increase. Considering the importance of Green's functions for source inversion and our limited knowledge of detailed velocity structure, seismic waveform inversions for large earthquakes should be based on the available structure models (including 3-D models) for the region of interest. Although we do not know which of these source models is more accurate in describing the source process, quantifying the difference among these models provides an estimation of the uncertainty in the solution obtained by inverting the data.

[44] The available information that can be used in the finite fault inversion is not sufficient to determine a continuous distribution of source process of large earthquake. Thus finite fault inversions usually require parameterization of the faulting process by dividing the finite fault into a grid of subfaults. However there is no criterion to decide how large a subfault should be. Normally a subfault is large enough that the variations of source parameters within the

subfault cannot be neglected. We apply the bilinear interpolation technique to describe these variations and find that it works well. Based on this, it is natural to infer that a higher order interpolation, such as a bi-cubic spline interpolation, may work even better. However, if a higher order interpolation is adopted, synthetic seismic waveforms from a subfault are no longer linearly related to the slip amplitudes of each corresponding node. The higher order interpolation will also force a stronger spatial correlation among the source parameters. Thus, determining even the slip distribution becomes a problem that cannot be solved with a linear least squares method, the favorite method used for inverting for the spatial distribution of slip on a finite fault. In addition, the non-linear relationship between the slip and ground motion and the enhanced stronger correlation of the source parameters makes the inversion process much more difficult. Our numerical tests show that higher order interpolation makes the inversion converge much more slowly; bi-cubic spline interpolation requires more than five times the CPU time than the bilinear interpolation for the same problem.

[45] *Ide and Takeo* [1997] proposed a different approach to representing the spatial distribution of slip on the fault plane. The slip distribution is expanded as linear combination of 2-D linear B-spline base functions. The expansion coefficients are defined by the slip amplitudes on knots distributed uniformly along the strike and dip directions, respectively. While our subfaults are rectangular elements of equal area and the locations of the subfault nodes coincide with those of knots of the 2-D linear B-spline, the two approaches give same representation of slip distribution. However using bilinear quadrilateral elements is more flexible for modeling the complex rupture process. For example, we can assign the distribution of nodes so that the spacing between nodes is smaller in the zones where the source parameters have the largest variations. This will allow us to improve the resolution of a finite fault model without increasing the number of free parameters. This advantage will be explored in our future research.

[46] The Loma Prieta inversions were done with exactly the same method so that one can compare the results based only on the differences resulting from 1-D versus 3-D Green's functions. Of course, it is not surprising that there are differences; the overall misfit is, in fact, less for the 1-D inversion. While we are in no position to judge which kinematic model more accurately reflects the earthquake, we have to accept that the velocity structure is three-dimensional [*Brocher et al.*, 1997; *Stidham et al.*, 1999]. Our preference is the inversion results using the 3-D structure with the caveat that we think that the attenuation given for the model was too large based on comparing the synthetics with the data. The 1-D inversion results can be compared with previous 1-D inversions for which 3-D structure models were not available [*Steidl et al.*, 1991; *Wald et al.*, 1991; *Beroza*, 1991, 1995]. Our 1-D model reflects both differences between our global nonlinear inversion and linear methods as well as the choice of the stations used in the inversion [*Wald et al.*, 1991; *Beroza*, 1995]. *Beroza* [1995] contrasted the various 1-D models; we want to comment only on how our 1-D model differs or is similar to that found by others.

[47] The distribution of slip (Figure 7) is more asymmetric comparing the fault southeast of the hypocenter to that northwest of the hypocenter than found in previous inversions. While it may look like the results from *Wald et al.* [1991], it is significantly different in that most of the slip northwest of the hypocenter is dominated by dip-slip, whereas *Wald et al.* [1991] had found primarily strike-slip northwest of the hypocenter using only strong-motion data. This distribution of slip is very different from *Beroza* [1995], who found the largest slip southeast of the hypocenter dominated by strike-slip motion. All of the models based on strong-motion data resulted in areas of maximum slip that are deeper than 10 km. Also, we find the regions of maximum slip are concentrated 10 km or more along strike from the hypocenter as did *Steidl et al.* [1991]. Models by *Wald et al.* [1991] and *Beroza* [1991, 1995] generally put the concentrations of slip within 10 km (measured along strike) of the hypocenter. In our 1-D model the rake angle dip slip motion dominates northwest of the hypocenter while to the southeast there are nearly equal parts of strike-slip and dip-slip. This is similar to *Beroza* [1991, 1995], *Steidl et al.* [1991] and the combined teleseismic and strong motion model of *Wald et al.* [1991]. Of all the models, our rake is similar to the combined teleseismic and strong motion model of *Wald et al.* [1991] and to the model of *Steidl et al.* [1991]. Our 1-D model is least like that of *Hartzell et al.* [1991], who inverted for slip on the fault using only teleseismic data. Their model is dominated by strike-slip motion southeast of the hypocenter and is concentrated at much shallower depths. All of the models do have one common feature: the rise times are generally less than 1.0 s.

[48] The primary differences between our 3-D and 1-D models are (1) that the 3-D model distribution of slip is more symmetric with nearly equal amplitudes northwest and southeast of the hypocenter with slightly more amplitude to the southeast and (2) the rake angle shows that in the southeast there is significantly more dip-slip motion. The maximum amplitude is 4.8 m for the 3-D inversion as compared to 6.8 for the 1-D inversion. As in the 1-D inversion there is almost no slip in the region directly above the hypocenter giving a bimodal distribution (Figure 8). The rise times for the 3-D inversion are still short, approximately 1.0 s or less in the regions of maximum slip. The rupture velocity is different from the 1-D in detail, but the general pattern is symmetric with respect to the hypocenter; both have an average rupture velocity of 2.8 km/s.

7. Conclusions

[49] We have developed a procedure to perform finite fault inversions. The problem of inverting seismic data for the spatial distribution of slip amplitude, rupture velocity and rise time on a fault is nonlinear. The global inversion method presented here efficiently solves this problem. Our approach requires only the value of the objective function, not its derivative, and uses random processes to search the model space to find the optimal solution. Throughout the process the nonzero probability of strong perturbations to a current model allow the method to escape from local minima. Our application of this approach shows that the

final solution is almost independent of the choice of the starting model. *Cotton and Campillo* [1995] using a linearized, iterative technique to solve a similar nonlinear problem also found a slight dependence of the solution on the starting model. Another advantage of the global inversion method is that many physical constraints about the model parameters are easily incorporated into the inversion. We find that it is necessary to consider the variations of source parameters within a subfault. The bilinear interpolation provides an efficient way to account for these variations. It mitigates the dependence of finite fault inversion on the choice of subfault size.

[50] Using identical numerical parameters and method we inverted near-source data from the Loma Prieta earthquake using both 1-D and 3-D Green's functions. Both Green's functions result in source models that produce synthetics that match the data with almost the same misfit in the objective function with the 1-D model producing a slightly smaller misfit. In practice, we do not know which Earth model more accurately represents the real structure, but this region has 3-D structure that penetrates to depth including variation of the velocity on different sides of the San Andreas (Figure 5 [*Brocher et al.*, 1997; *Stidham et al.*, 1999]). A prudent approach is to use each Earth model that is available to deduce the range of possible faulting models and look for the elements that are common to the different Earth models [*Hartzell*, 1989; *Graves and Wald*, 2001]. Here we find the common elements are that slip northwest of the hypocenter is deep and dominated by dip-slip; slip southeast of the hypocenter has a strong component of strike-slip; the rise time on the fault is less than 1.0 s; the rupture velocity is about 2.8 km/s.

[51] **Acknowledgments.** We thank Steve Hartzell and Chen Ji for their advice during the course of this work. We also thank Dave Wald and Robert Graves and the Associate Editor for their thoughtful reviews. The research was funded by the National Science Foundation under contract EAR-0073899 and by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. This is ICS contribution 0639 and SCEC contribution 755.

References

Aki, K. (1967), Scaling law of seismic spectrum, *J. Geophys. Res.*, *72*, 1231–1271.

Archuleta, R. J. (1984), A faulting model for the 1979 Imperial Valley earthquake, *J. Geophys. Res.*, *89*, 4559–4585.

Archuleta, R. J., and S. M. Day (1980), Dynamic rupture in a layered medium: The 1966 Parkfield earthquake, *Bull. Seismol. Soc. Am.*, *70*, 671–689.

Beroza, G. C. (1991), Near-source modeling of the Loma Prieta earthquake: Evidence for heterogeneous slip and implications for earthquake hazard, *Bull. Seismol. Soc. Am.*, *81*, 1603–1621.

Beroza, G. C. (1995), Rupture history of the earthquake from high-frequency strong-motion data, in *The Loma Prieta, California Earthquake of October 17, 1989 -Main-Shock Characteristics*, *U.S. Geol. Surv. Prof. Pap.*, *P 1550-A*, 9–32.

Bouchon, M. (1997), The state of stress on some faults of the San Andreas system as inferred from near-field strong motion data, *J. Geophys. Res.*, *102*, 11,731–11,744.

Brocher, T. M., et al. (1997), A crustal-scale 3-D seismic velocity model for the San Francisco Bay area, California (abstract), *Eos Trans. AGU*, *78*(46), Fall. Meet. Suppl., F435.

Brune, J. N. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, *75*, 4997–5009.

Cohee, B. P., and G. C. Beroza (1994), Slip distribution of the 1992 Landers earthquake and its implications for earthquake source mechanism, *Bull. Seismol. Soc. Am.*, *84*, 692–712.

Cotton, F., and M. Campillo (1995), Frequency domain inversion of strong motions: Application to the 1992 Landers earthquake, *J. Geophys. Res.*, *100*, 3961–3975.

Das, S., and P. Suhadolc (1996), On the inversion problem for earthquake rupture: The Haskell-type source model, *J. Geophys. Res.*, *101*, 5725–5738.

Day, S. M. (1977), Finite element analysis of seismic scattering problems, Ph.D. dissertation, 149 pp., Univ. of Calif., San Diego.

Day, S. M., Y. Guang, and D. J. Wald (1998), Dynamic stress changes during earthquake rupture, *Bull. Seismol. Soc. Am.*, *88*, 515–522.

Favreau, P., and R. J. Archuleta (2003), Direct seismic energy modeling and application to the 1979 Imperial Valley earthquake, *Geophys. Res. Lett.*, *30*(5), 1198, doi:10.1029/2002GL015968.

Graves, R. W., and D. J. Wald (2001), Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions: 1. Strong motions, *J. Geophys. Res.*, *106*, 8745–8766.

Hartzell, S. H. (1989), Comparison of seismic waveform inversion results for the rupture history of a finite fault: Application to the 1986 North Palm Springs, California, earthquake, *J. Geophys. Res.*, *94*, 7515–7534.

Hartzell, S. H., and T. H. Heaton (1983), Inversion of strong ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California, earthquake, *Bull. Seismol. Soc. Am.*, *73*, 1553–1583.

Hartzell, S. H., and D. V. Helmberger (1982), Strong-motion modeling of Imperial Valley earthquake of 1979, *Bull. Seismol. Soc. Am.*, *72*, 571–596.

Hartzell, S. H., and C. J. Langer (1993), Important of model parameterization in finite fault inversions: Application to the 1974 Mw 8.0 Peru earthquake, *J. Geophys. Res.*, *98*, 2123–2134.

Hartzell, S. H., and P. Liu (1996), Calculation of earthquake rupture histories using a hybrid global search algorithm: Application to the 1992 Landers, California, earthquake, *Phys. Earth Planet. Inter.*, *95*, 79–99.

Hartzell, S. H., G. S. Stewart, and C. Mendoza (1991), L_1 and L_2 norms in a teleseismic waveform inversion for the slip history of the Loam Prieta, California, earthquake, *Bull. Seismol. Soc. Am.*, *81*, 1518–1539.

Hartzell, S. H., P. Liu, and C. Mendoza (1996), The 1994 Northridge, California, earthquake: Investigation of rupture velocity, rise time, and high-frequency radiation, *J. Geophys. Res.*, *101*, 20,091–20,108.

Haskell, N. A. (1964), Total energy and energy spectra density of elastic wave from propagating fault, *Bull. Seismol. Soc. Am.*, *54*, 1811–1841.

Hauksson, E., and J. S. Haase (1997), Three-dimensional V_P and V_P/V_S velocity models of the Los Angeles basin and central Transverse Ranges, California, *J. Geophys. Res.*, *102*, 5423–5453.

Hughes, T. J. R. (1987), *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis*, pp. 109–119, Prentice-Hall, Old Tappan, N. J.

Ide, S., and M. Takeo (1997), Determination of constitutive relations of fault slip based on seismic wave analysis, *J. Geophys. Res.*, *102*, 27,379–27,391.

Ji, C., D. J. Wald, and D. V. Helmberger (2002), Source description of the 1999 Hector Mine, California, earthquake, part I: Wavelet domain inversion theory and resolution analysis, *Bull. Seismol. Soc. Am.*, *92*, 1192–1207.

Lane, F. (1992), Estimation of kinematic rupture parameters from historical seismograms: An application of simulated annealing to a nonlinear optimization problem, Ph.D. thesis, Colo. Sch. of Mines, Golden, Colo.

Liu, P., and R. J. Archuleta (1999), Source parameter inversion using 3-D Green's functions: Application to the 1996 Northridge, California earthquake (abstract), *Eos Trans. AGU*, *80*(46), Fall Meet. Suppl., F709.

Liu, P., S. H. Hartzell, and W. J. Stephenson (1995), Nonlinear multiparameter inversion using a hybrid global search algorithm: Applications in reflection seismology, *Geophys. J. Int.*, *122*, 991–1000.

Magistrale, H., R. W. Graves, and R. Clayton (1998), A standard three dimensional seismic velocity model for southern California reference, version 1 (abstract), *Eos Trans. AGU*, *79*(45), Fall Meet. Suppl., F605.

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

Mai, P., and G. C. Beroza (2002), A spatial random-field model to characterize complexity in earthquake slip, *J. Geophys. Res.*, *107*(B11), 2308, doi:10.1029/2001JB000588.

Mikumo, T., and T. Miyatake (1995), Heterogeneous distribution of dynamic stress strop and relative fault strength recovered from the results of waveform inversion: The 1984 Morgan Hill, California, earthquake, *Bull. Seismol. Soc. Am.*, *85*, 178–193.

Nielsen, S., and K. Olsen (2000), Constraints on stress and friction from dynamic rupture models of the 1994 Northridge, California, earthquake, *Pure Appl. Geophys.*, *157*, 2029–2046.

Olsen, K., R. Madariaga, and R. J. Archuleta (1997), Three-dimensional dynamic simulation of the 1992 Landers earthquake, *Science*, *278*, 834–838.
Olson, A. H., and J. G. Anderson (1988), Implications of frequency domain inversion of earthquake ground motions for resolving the space-time dependence of slip on an extended fault, *Geophys. J. R. Astron. Soc.*, *94*, 443–455.
Olson, A. H., and R. J. Apsel (1982), Finite fault and inversion theory with applications to 1979 Imperial Valley earthquake, *Bull. Seismol. Soc. Am.*, *72*, 1969–2001.
Peyrat, S., K. Olsen, and R. Madariaga (2001), Dynamic modeling of the 1992 Landers earthquake, *J. Geophys. Res.*, *106*, 26,467–26,482.
Ramos-Martínez, J., and G. A. McMechan (2001), Source parameter estimation by full waveform inversion in 3-D heterogeneous, viscoelastic anisotropic media, *Bull. Seismol. Soc. Am.*, *85*, 276–291.
Rothman, D. (1986), Automatic estimation of large residual statics corrections, *Geophysics*, *51*, 332–346.
Sekiguchi, H., K. Irikura, T. Iwata, Y. Kachi, and M. Hoshiba (1996), Minute locating of faulting beneath Kobe and waveform inversion of source process during the 1995 Hyogo-ken Nanbu, Japan, earthquake using strong motion records, *J. Phys. Earth*, *44*, 473–487.
Sekiguchi, H., K. Irikura, and T. Iwata (2000), Fault geometry at the rupture termination of the 1995 Hyogo-ken Nanbu, earthquake, *Bull. Seismol. Soc. Am.*, *90*, 117–134.
Sen, M., and P. L. Stoffa (1991), Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics*, *56*, 1624–1638.
Sen, M., and P. L. Stoffa (1995), *Global Optimization Methods in Geophysical Inversion*, Elsevier Sci., New York.
Somerville, P. G., K. Irikura, R. Graves, S. Sawata, D. J. Wald, N. Abrahamson, Y. Iwasaki, T. Kagawa, N. Smith, and A. Kowada (1999), Characterizing crustal earthquake slip model for the prediction of strong ground motion, *Seismol. Res. Lett.*, *70*, 59–80.
Spudich, P. (1980), The Dehoop-Knopoff representation theorem as a linear inversion problem, *Geophys. Res. Lett.*, *9*, 717–720.
Spudich, P., and R. J. Archuleta (1987), Techniques for earthquake ground-motion calculation with applications to source parameterization of finite faults, in *Seismic Strong Motion Synthetics*, edited by B. A. Bolt, pp. 205–265, Academic, San Diego, Calif.
Steidl, J. H., R. J. Archuleta, and S. H. Hartzell (1991), Rupture history of the 1989 Loma Prieta, California, earthquake, *Bull. Seismol. Soc. Am.*, *81*, 1573–1602.
Stidham, C., M. Antolik, D. Dreger, S. Larson, and B. Romanowicz (1999), Three-dimensional structure influences on the strong-motion wavefield for the 1989 Loma Prieta earthquake, *Bull. Seismol. Soc. Am.*, *89*, 1184–1202.
Tarantola, A. (1987), *Inverse Problem Theory*, Elsevier Sci., New York.
Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problem solved using the least squares criterion, *Rev. Geophys.*, *20*, 219–232.
Wald, D. J., and R. W. Graves (2001), Resolution analysis of finite fault source inversion using one- and three-dimensional Green's functions: 2. Combining seismic and geodetic data, *J. Geophys. Res.*, *106*, 8767–8788.
Wald, D. J., and T. H. Heaton (1994), Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake, *Bull. Seismol. Soc. Am.*, *84*, 668–691.
Wald, D. J., D. V. Helmberger, and T. H. Heaton (1991), Rupture model of the 1989 Loma Prieta earthquake from the inversion of strong-motion and broadband teleseismic data, *Bull. Seismol. Soc. Am.*, *81*, 1540–1572.
Zeng, Y., and J. G. Anderson (1996), A composite source model of the 1994 Northridge earthquake using genetic algorithms, *Bull. Seismol. Soc. Am.*, *86*, S71–S83.

R. J. Archuleta and P. Liu, Institute for Crustal Studies, University of California, Santa Barbara, CA 93106, USA. (pcliu@crustal.ucsb.edu)

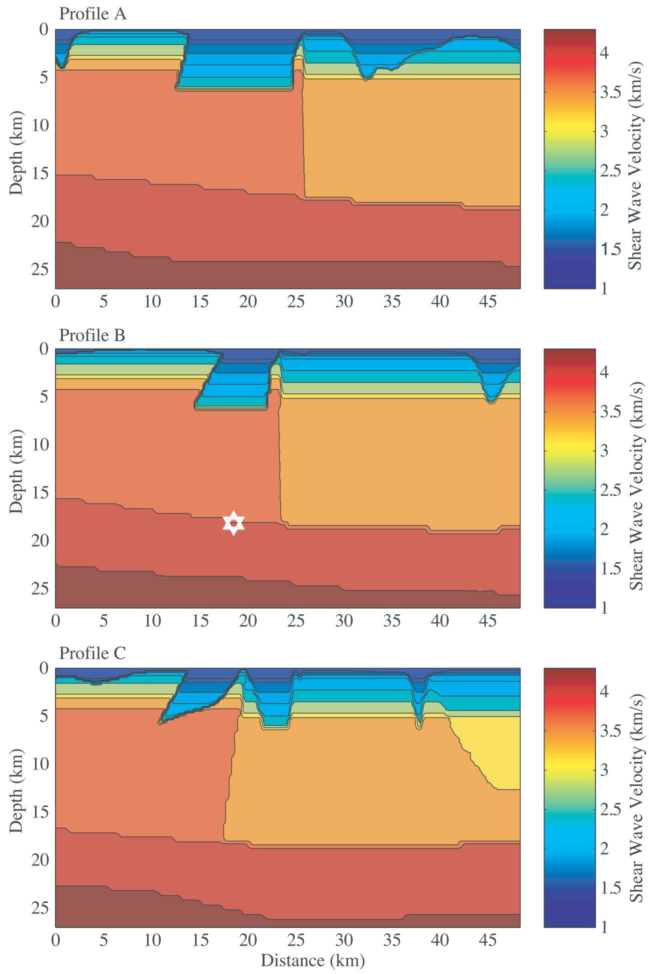

Figure 5. Cross sections of shear wave velocity for three profiles indicated on Figure 4. In Profile B-B′, the hypocenter is shown as a white star. One can see the difference in velocity across the San Andreas to a depth of about 20 km in each profile. The local basins penetrate to depths about 5 km. Velocity structure is the latest USGS model (T. M. Brocher, personal communication).