

Developing 3-D shear wave models using a multi-objective joint inversion scheme

Lennox Thompson, Aaron A. Velasco, and Vladik Kreinovich

Abstract For this research, our main purpose is to obtain a better understanding of the Earth's tectonic processes in the Texas region, which requires us to analyze the Earth structure. We expand on a constrained optimization approach for a joint inversion least-squares (LSQ) algorithm to characterize a one-dimensional Earth's structure of Texas with the use of multiple geophysical data sets. We employed a joint inversion scheme using multiple geophysical datasets for the sole purpose of obtaining a three-dimensional velocity structure of Texas in order to identify an ancient rift system within Texas. In particular, we use data from the USArray, which is part of the EarthScope experiment, a 15-year program to place a dense network of permanent and portable seismographs across the continental United States. Utilizing the USArray data has provided us with the ability to image the crust and upper mantle structure of Texas. We simultaneously inverted multiple datasets from USArray data, to help us to better obtain an estimate of the true Earth Structure model. We prove through numerical and experimental testing that our Multi-Objective Optimization (MOP) scheme performs inversion in a more accurate, robust, and flexible manner than traditional inversion approaches.

Key words: Teleseismic, Receiver Functions, Seismographs, Body Waves, Multi-Objective Optimization, Primal-Dual Interior Point Method

Lennox Thompson

Dept of Geological Sciences, University of Texas at El Paso (UTEP), 500 W. University Ave, El Paso, TX 79968 e-mail: lethompson@miners.utep.edu

Aaron A. Velasco

Dept of Geological Sciences, University of Texas at El Paso (UTEP), 500 W. University Ave, El Paso, TX 79968 e-mail: aavelasco@utep.edu

Vladik Kreinovich

University of Texas at El Paso (UTEP), 500 W. University Ave, El Paso, TX 79968 e-mail: vladik@utep.edu

1 Introduction

For this research, we propose to combine multiple geophysical datasets for the purpose of assisting us in better determining physical properties of the Earth structure. By simultaneously inverting multiple datasets, we obtain a better estimate of the true Earth Structure. In general, there are two reasons why the estimated Earth Structure model differs from the true Earth structure. The first reason is the inherent non-uniqueness of the inverse problem that causes several (usually infinitely many) models to satisfy the data. The second reason is that real geophysical data is always affected by noise, which introduces error associated with the estimation of the Earth Structure model after inversion. By jointly inverting multiple geophysical data sets, we avoid the inherent non-uniqueness typical for the geophysical datasets (e.g., receiver functions, surface wave dispersion, teleseismic delay travel times, and gravity) individually [Vozoff et al., 1975]. For this research, we use receiver functions, surface wave dispersion measurements, travel times, and gravity data sets to characterize the crust and upper mantle of various region such as Texas.

In general, geophysical data sets such as receiver functions are suited to constrain the depth of discontinuities and are sensitive to relative changes in S-wave velocities in different layers. Surface waves measurements on the other hand, constrain the absolute shear velocities between discontinuities whereas receiver functions are unable to do that [Shen et al., 2013]. Seismic first-arrival travel times and gravity data are complementary to each other because one can recover the causative slowness and density distributions of the Earth structure. The complementary information provided by the following datasets, reduces the inherent ambiguity or non-uniqueness of performing inversion [e.g., Haber & Oldenburg 1997]. By jointly inverting seismic data along with gravity data, we will be able to overcome the difficulties of non-uniqueness and be able to facilitate the construction of the true Earth model.

When we process a single data set (e.g., Surface Wave Dispersion), we use the least squares method to find the best-fit model. For multiple data sets (e.g., Surface Wave Dispersion and Receiver Functions), if we knew the variance (uncertainty of data) of the different measurements of the multiple data sets, we could still be able to use the least squares approach to finding the model space. In practice, we only have an approximate knowledge of the variances. So, instead of producing a single model, we want to generate several models corresponding to different possible variances. Once several models corresponding to different possible variances are computed, we can then proceed to select the most geophysical meaningful model from the Pareto Front. The reason we will use an optimization technique is to find the best possible solution for nonlinear geophysics inverse problem. For example, in geophysics, most inverse problems require finding some minimization and that is why we will use an optimization technique called MultiObjective Optimization Problem (MOP). The MOP technique generates several possible models. This is what sets it apart from other various joint inversion techniques. We will be able to select the final solution from a population of alternative solutions from the model space. Such methods are described in [Sambridge 1999a,b; Kozlovskaya 2000].

There are two types of seismic waves that travel through the Earth: the body waves and the surface waves. Both types of waves give us different sensitivities and information about the Earth Structure, since they are sampling the interior and surface of the medium with different velocities and directions. The information collected from the body waves travels deeper into the Earth and translates into teleseismic P-wave receiver functions. In order to obtain information about the Earth surface, surface waves are analyzed, in our case, by means of surface waves dispersion. On one hand we have receiver functions, which resolve discontinuities (impedance contrasts) in seismic velocities, and provide good measurement of crustal thickness, without providing a good average of shear wave velocity. On the other hand, we have surface (Love and Rayleigh) waves whose energy is concentrated near the Earth's surface, and provide good average of absolute shear wave velocity, without a good shear-wave velocity contrasts in layered structures. Therefore these two data sets can be considered as complimentary and consistent, as long as we sample the same medium. Hence, we expect a mutually consistent estimate of the Earth's structure. Since both data sets are sensitive to shear wave velocity structure [Julia et al., 2000], we can assume a forward operator F depending nonlinearly on our model parameter $x \in \mathbb{R}^n$ that represents the different shear velocities of a half space with n horizontal layers (a standard way of modeling Earth's structure). In the next subsections we explain in more detail the nonlinear relationship with respect to shear wave velocities of this operator and the techniques used to compute each synthetic dataset.

1.1 Receiver Functions

A receiver function is simply a time series representation of the Earth's response relative to an incoming P-wave propagating near a recording station. Positive or negative spike amplitudes represent positive or negative seismic velocity contrast. A receiver function technique can model the structure of the earth by using seismograms from three component (vertical, north, and east) seismic stations from teleseismic earthquakes. The receiver function technique takes advantage of the fact that part of the energy of seismic P waves is converted into S waves at discontinuities along the ray path [Dzierma et al., 2011], and has been utilized in many studies [e.g., Wilson et al., 2005; Bailey et al., 2012; Wilson and Aster, 2005]. For data collection and processing, we use the Standing Order for Data (SOD) [Owens et al., 2004; Bailey et al., 2012] to request three component seismograms for P-wave arrivals and for events with a minimum magnitude 5.5, depth in the range of 1600 km, and an epicentral distance ranging from $30^\circ - 95^\circ$ [e.g., Bailey et al., 2012].

Receiver functions were first applied in the late 1970s at solitary stations to obtain local one-dimensional structural estimates [Langston, 1981]. Since then, there was an increase in the number of stations deployed seismic experiments. It is now possible to generate detailed two or three-dimensional images of structures, such as

the moho and upper mantle transition zone discontinuities near 410 km and 670 km depth [e.g., Wilson et al., 2003].

Receiver functions are derived using deconvolution, a mathematical method used to filter a signal and isolate the superimposed harmonic waves. Specially, receiver functions are calculated by deconvolving the vertical component of a seismogram from the radial component, resulting in the identification of converted phases where there is an impedance contrast (crustal-mantle boundary). The receiver functions are sensitive to V_p/V_s ratios and crustal thickness.

Deconvolution is again, a mathematical method used to filter signals and isolate harmonic waves. For example, when given a convolved signal $u(t) = s(t) * p(t) * i(t)$, the system should isolate the components $s(t)$, the source, $p(t)$, the earth structure, and $i(t)$, the instrumentation so that they can be analyzed individually from the seismogram $u(t)$.

Convolution of the source $S(t)$, propagation $P(t)$, instrument $I(t)$, will result in the seismogram denoted as in the time domain.

$$\begin{aligned} U_z(t) &= S(t) * P_z(t) * I(t) \\ U_r(t) &= S(t) * P_r(t) * I(t) \end{aligned} \quad (1)$$

Taking the Fourier transform [Spiegel, 1974] of the vertical $U_z(t)$ and radial $U_r(t)$ component of the seismogram will result in the vertical and radial components from time domain to frequency domain,

$$\begin{aligned} F[U_z(t)] &= U_z(\omega) = S(\omega) * P_z(\omega) * I(\omega) \\ F[U_r(t)] &= U_r(\omega) = S(\omega) * P_r(\omega) * I(\omega) \end{aligned} \quad (2)$$

By performing deconvolution (spectral division) of the vertical component from the radial component,

$$R(\omega) = \frac{P_z(\omega)}{P_r(\omega)} \quad (3)$$

then taking the inverse Fourier transform of $R(\omega)$, results in the approximation of a receiver function $r(t)$.

$$r(t) = F^{-1} \left[\frac{P_z(\omega)}{P_r(\omega)} \right] \quad (4)$$

1.2 Receiver function stacking

We used the receiver function stacking technique introduced by Zhu and Kanamori (2000), which estimates the crustal thickness and a V_p/V_s ratio based on the radial

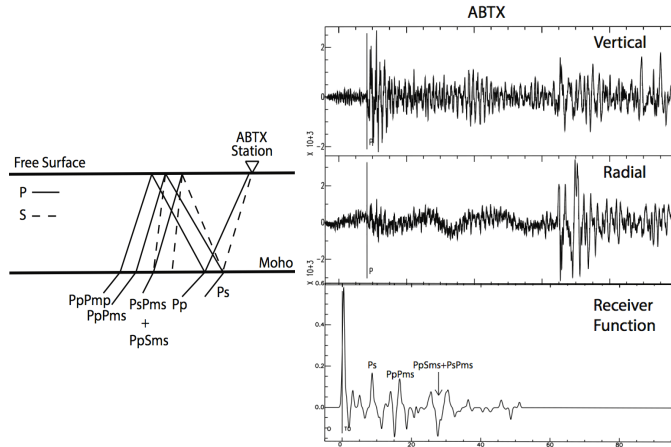


Fig. 1 (Left) Illustration of a simplified ray diagram, which identifies the P_s converted phases, which comprise the receiver function for a single layer. (Right) Vertical and radial seismograms and the corresponding receiver function resulting from the deconvolution of the vertical component from the radial component

receiver function. This technique is the standard approach used by EARS. Assuming that no lateral velocity heterogeneities exist, the time separation between the P_s converted wave and the direct P-wave obtained from receiver functions (t_{P_s}) can then be used to estimate crustal thickness (H), given the average crustal velocities V_p and a V_p/V_s ratio (κ), and the constant ray parameter p of the incident wave [e.g., Gurrola et al., 1995]. The trade-off between the thickness and the crustal velocities presents an ambiguity that can be reduced by using the later multiple phases $t_{p_p P_s}$ and $t_{p_s P_s} + p_p s_s$, which provide additional constraints to both V_p/V_s and the crustal thickness [e.g., Gurrola et al., 1995; Zhu and Kanamori, 2000]. Using and stacking multiple events helps to increase the signal-to-noise ratio (SNR), which may be caused by background noise, scattering from crustal heterogeneities, and P-to-S multiple conversions from other velocity discontinuities [Lodge and Helffrich, 2009]. The H - κ domain stacking weights each phase and plots the stacked phases as a gridded image $s(H, \kappa)$, which reaches a maximum when all three phases (t_{P_s} , $t_{p_p P_s}$, $t_{p_s P_s} + p_p s_s$) are stacked coherently with the correct H and κ [Zhu and Kanamori, 2000]. The main advantage of this grid-search based technique is that (1) large amounts of receiver functions can be processed without the need of picking P_s arrival times, and (2) the stacking results in an enhancement of the signal/noise ratio and a suppression of lateral variations in the vicinity of the recording station [Lodge and Helffrich, 2009]. We will use this technique to derive an average crustal model including H and V_p/V_s (κ). An example of this technique is shown in Figure 5 for one of the Earthscope USArray stations, 219A. The dark dot with the white circle around the dot represents the possible solution in H and V_p/V_s space (Figure 3).

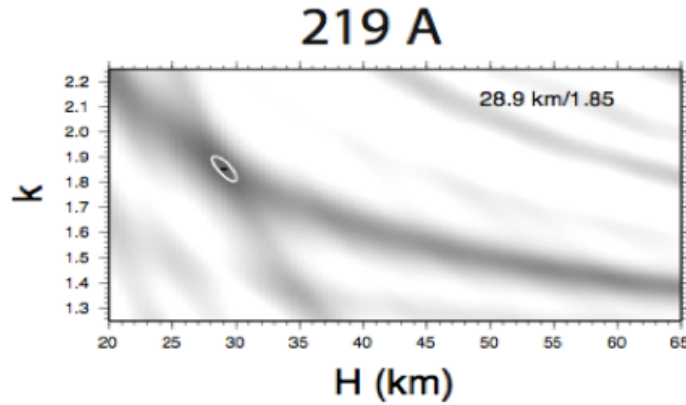


Fig. 2 This is a receiver function stack of station 219A, V_p/V_s vs H (km). The black dot with the white circle around it represents the preferred value. Note the multiple shaded regions might result in a poor choice of crustal thickness.

1.3 Surface Wave Dispersion

Surface waves in general differ from body waves in many respects they travel slower, lower frequencies, largest amplitudes, and their velocities are in fact dependent on frequency [Shearer, 2009]. The surface wave velocities vary with respect to depth being sampled by each period of the surface wave. The sampling by each period of the surface wave is known as dispersion [Sosa et al., 2013]. Valuable information can be inferred by measuring surface wave dispersion because it will allow you to be able to better understand the Earth's crustal and mantle velocity structure [Obrebski et al., 2010, Sosa, 2013]. In particular, Love and Rayleigh wave group dispersion observations generally account for average velocity structure as a function of depth [Julia et al., 2000; Maceira and Ammon 2009]. The dispersion curves for surface waves are extracted from station records of three component seismograms for different frequencies and distances, by using reduction algorithms that rely on spectral analysis techniques. The important fact here is that, based on Rayleigh's principle, surface wave velocities are more sensitive to S wave velocity, although they are also theoretically sensitive to P wave velocity and density. The Rayleigh's principle states that the phase velocity perturbation, denoted by $\frac{\partial c}{c}$ can be viewed as a function of $(K_\alpha, K_\beta, K_\rho)$, the sensitivity coefficients for P wave velocity, S wave velocity and density, respectively, i.e.

$$\frac{\partial c(T)}{c(T)} = \int \left(K_\alpha \frac{\partial \alpha(z)}{\alpha(z)} + K_\beta \frac{\partial \beta(z)}{\beta(z)} + K_\rho \frac{\partial \rho(z)}{\rho(z)} \right) \quad (5)$$

where T is the period and z is the depth. By investigating sensitivity function variation in depth, the relative contribution of each property to dispersion can be shown. This subject is beyond the scope of our work, thus we just mention here that such analysis allows geophysicists to show that the relative contribution of P wave velocity, and density to dispersion is smaller than the one for S wave velocity [Julia et al., 2000]. This is, surface wave dispersion is much more sensitive with respect to S wave velocity, and therefore we have established the dependence of this data set on shear wave velocity.

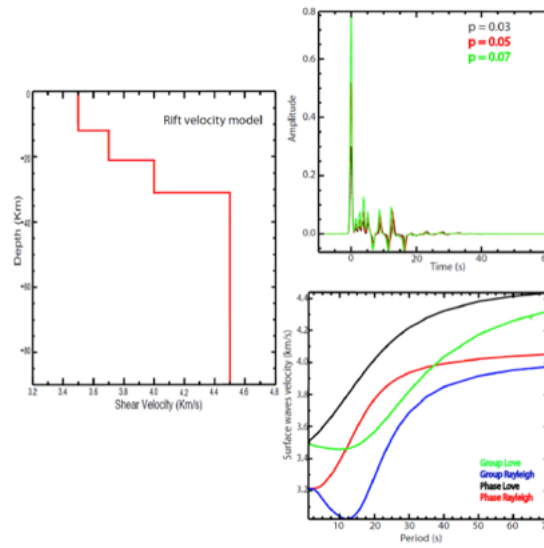


Fig. 3 (Left) S-wave synthetic velocity model (Rift) [Laske et al., 2000]. (Right: top) Receiver functions for three different ray parameters p , and (Right: bottom) surface wave dispersion curves (Love and Rayleigh) computed for the velocity model on the left.

1.4 Delay Travel Times

The travelttime T between a source and receiver along a ray L is given in integral form for a velocity field as

$$T = \int_L \frac{ds}{v(s)} \quad (6)$$

where s is the position vector in 2D or 3D media. Travel times are considered a nonlinear inverse problem given the relationship between the measured data (travel times) and the unknown model parameters (the velocity field). However, by transforming variables to use slowness, the reciprocal of velocity, instead of velocity as the unknown, a seemingly linear inversion problem is created:

$$\int_L \Delta u(s) ds = \Delta T = T_{obs} - T_{pred} \quad (7)$$

However, the ray is also dependent on the velocity (or slowness) model, thus making the inverse problem nonlinear regardless of what form of model variable or parameterization is used. If the medium is subdivided into blocks, the path length l_j in the j th block and can be discretize to

$$\Delta T = \sum_j l_j \Delta u_j \quad (8)$$

The model can be parameterized any number of ways using velocity or slowness, and cells, nodes, or splines, since the problems nonlinearity must be dealt with regardless of the parameterization. Most often a linearized gradient approach is applied in which a starting model is used and both the model and rays are updated over a series of iterations with the hope that there will be convergence to an acceptable model (the final model). The model is almost always discretized using cells, nodes, or other interpolating functions; in the latter two cases, the discrete model parameters are the coefficients of the interpolating functions. For the formulation of travel times for a tomography problem, the model is parameterized using constant-slowness cells, in which case the equation for the i th data becomes

$$\Delta T_i = \sum_j l_{ij} \Delta u_j \quad (9)$$

where l_{ij} is the length of the i th ray in the j th model cell and Δu_j is the slowness in the j th cell. In this case the path length of each ray in a block, l_{ij} is the partial derivative, $\partial T_i / \partial u_j$ of the travel time with respect to the slowness of that block [Stein and Wysession, 2006].

1.5 Gravity Anomalies

In geophysics gravity anomalies are generally defined as the difference between observed gravity field and the field of a reference model. Depending on the reference gravity model, two different types of anomaly variations are considered: gravity anomalies and gravity disturbances. The geodetic gravity anomaly is defined as the difference between gravity on the geoid and normal gravity on the reference ellipsoid [Heiskanen and Moritz, 1967]. On the other hand, the gravity disturbance is defined as the difference of the fields at the same point on the reference ellipsoid. It has been demonstrated that the gravity disturbances are more appropriate for geo-

physical purposes [e.g., Hackney and Featherstone, 2003]. In any case, its necessary to take into account the difference in the interpretation.

The observed gravity anomalies reflect the effect of density variations relative to the homogeneous reference model. Interpretation of the gravity anomalies implies an estimation of the density heterogeneities. The density model should reproduce the observed gravity field, taking into account that the observations may be affected by measurement errors. Density heterogeneity of the Earth, associated with thermal and compositional variations or with deflected boundaries separating layers of different density, is one of the main factors that control dynamic processes and deformations at both shallow and deep levels. Therefore, interpretation of the gravity anomalies or gravity modeling is one of the principal methods, which help to understand the nature and origin of the tectonic processes and the Earths dynamics.

2 Forward Problem

If we know the layered shear velocity distribution $x = (x_1, \dots, x_n)$ at n different horizontal layers, then we can evaluate the measured quantities $y = (x_1, \dots, x_m)$ (e.g., the travel times) by applying an appropriate nonlinear operator $F(x)$ that uses the velocities x to predict the Earth's response $y = F(x)$;

$$F(x) = (F_1(x), \dots, F_m(x)) \in \mathbb{R}^m, x = (x_1, \dots, x_n) \in \mathbb{R}^n (m \gg n) \quad (10)$$

The operator F relates the data space and the model space. In other words, if we know the velocity model x , then we can predict the Earths response based on the velocity model.

3 Inverse Problem

Given an observed data vector, $y \in \mathbb{R}^m$, we want to find the unknown model, x , such that $F(x)$ approximates as much as possible. For each specific type T of observations, this means that we are minimizing

$$\min_x \|F^T(x) - y^T\|^2 = \min_x (F_i^T(x) - y_i^T)^2 \quad (11)$$

to match measurements of different types, researchers traditionally use weighted non-linear least squares method (LSQ). For example, to simultaneously match the teleseismic receiver functions (RF), surface wave dispersion velocities (SW), travel times (TT), and gravity (GR), we minimize $\min_x J$, where

$$J = w_{RF}^2 \|F^{RF}(x) - y^{RF}\|^2 + w_{SW}^2 \|F^{SW}(x) - y^{SW}\|^2 + w_{TT}^2 \|F^{TT}(x) - y^{TT}\|^2 + w_{GR}^2 \|F^{GR}(x) - y^{GR}\|^2 \quad (12)$$

This minimization problem can be reformulated as

$$\min_x \|F(x) - y\|^2 \quad (13)$$

where

$$F(x) = W \begin{pmatrix} F^{SW}(x) \\ F^{RF}(x) \\ F^{TT}(x) \\ F^{GR}(x) \end{pmatrix} \in \mathbb{R}^m,$$

$$y = W \begin{pmatrix} y^{SW} \\ y^{RF} \\ y^{TT} \\ y^{GR} \end{pmatrix} \in \mathbb{R}^n$$

where

$$W = \text{diag}(w_i), w_i = \sqrt{\frac{\eta_1}{\sigma_i^2 p}}, i = 1, \dots, p, w_i = \sqrt{\frac{\eta_2}{\sigma_i^2 q}}, i = p+1, \dots, p+q, w_i = \sqrt{\frac{\eta_3}{\sigma_i^2 r}}$$

$$i = p+q+1, \dots, p+q+r, w_i = \sqrt{\frac{1 - \eta_1 - \eta_2 - \eta_3}{\sigma_i^2 s}}, i = p+q+r+1, \dots, m = p+q+r+s \quad (14)$$

with W a weighted diagonal matrix used to equalize the contribution of each dataset with respect to physical units and number of data points, $\eta_i \in [0, 1]$ are influence parameters that measures the reliability of each dataset used for the inversion, σ_i^2 is the approximate standard deviation of each point, and p, q, r and s are the number of RF, SW, TT, and GR observations [Sosa et al., 2013].

4 Need for multi-objective optimization

In practice, we do not know the exact values of the influence parameters. For different values of the influence parameters, we get, in general, different velocity distributions x ; some of these velocity models are geophysically meaningful, some are not (e.g., some models x predict higher velocities in the crust and lower velocities in the mantle contrary to geophysics).

Traditionally, researchers avoid non-physical non-smooth velocity models by adding

a regularization form $+\lambda \|Lx\|^2$ to the minimized function [Tikhonov, 1977]. The problem with this term is that it is not clear how to select λ , and different values of λ lead to different solutions; see, e.g., [Hansen 1987] and [Vogel, 2002].

In this work, instead of using regularization, we explicitly formulate constraints that need to be satisfied, for example, the desired smoothness can be described as a bound on $|x_i - x_j| \leq \Delta$ on the difference between velocities x_i and x_j at nearby locations. Then, we find the model x for which $J(x)$ is the smallest under these constraints. Additionally, we include bounds $a \leq x \leq b$ on the velocities at different depths. In geophysical applications, it is crucial to keep the physical parameters within appropriate bounds.

So, instead of selecting a single combination of influence parameters (and thus, of weights), we propose to use multi-objective optimization (MOP); namely, we generate all possible models x corresponding to different combinations of weights, and then we use one of the MOP criteria to select the most promising model [Sambridge 1999a,b; Kozlovskaya, 2000].

In this case, we want to minimize the four criteria $f_1(x) = \|F^{RF}(x) - y^{RF}\|^2$, $f_2(x) = \|F^{SW}(x) - y^{SW}\|^2$, $f_3(x) = \|F^{TT}(x) - y^{TT}\|^2$, $f_4(x) = \|F^{GR}(x) - y^{GR}\|^2$. First, we find the Pareto optimal set, P^* , i.e. the set of all feasible solutions x for which, there is no other feasible solution x' which is better with respect to all criteria $f_1(x') < f_1(x), \dots, f_k(x') < f_k(x)$.

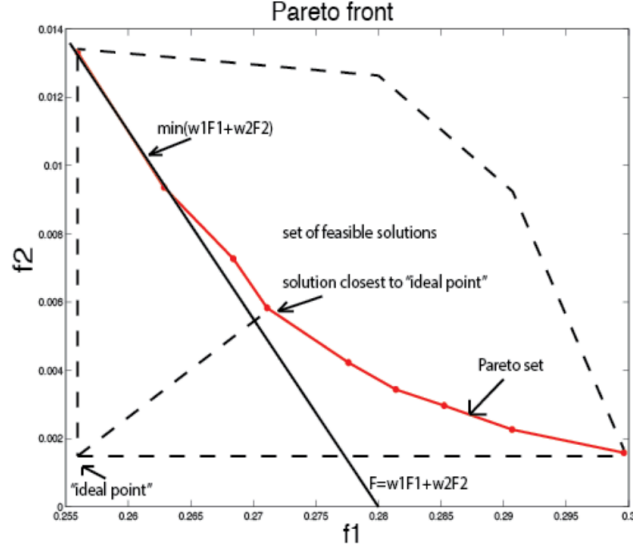


Fig. 4 Illustration of the solution set or Pareto front, which is, defined as the weights times the perspective objective functions.

Definition: (Pareto Optimal Set) For a given multi-objective problem, $F(x) = (f_1(x), \dots, f_k(x))$, the Pareto Optimal Set, $P^* \triangleq$, is defined as:

$$P \triangleq \{x \in \Omega \mid \neg \exists x' \in \Omega (F(x') \leq F(x))\} \quad (15)$$

It is known that elements of the Pareto set can be obtained by solving the one-objective (scalar) optimization problem.

$$\min_{x \in X} f(x) = \sum_{i=1}^k w_i f_i(x) \quad (16)$$

where $w = (w_1, \dots, w_k) \geq 0$ is the vector of weighting coefficients assigned prior to the solution of the problem. So, in our computations, we try all possible combinations of weights, and we find all solutions x corresponding to different combinations. For each criterion f_i , we then find the smallest value f_i^{min} and the largest value f_i^{max} . The smallest values form an ideal point $f^{min} = (f_1^{min}, \dots, f_k^{min})$. We then select a solution x which is the closest to this ideal point. Specifically, we normalize each differences $f_i(x) - f_i^{min}(x)$ to the interval (0,1) by dividing it by $f_i^{max}(x) - f_i^{min}(x)$, and then we minimize the corresponding normalized distance. In other words, we select a solution x for which the distance

$$d^2(f^{min}, f(x)) = \sum_{i=1}^k \left(\frac{f_i(x) - f_i^{min}(x)}{f_i^{max}(x) - f_i^{min}(x)} \right)^2 \quad (17)$$

5 Numerical Algorithm

First, we use a first order Taylor approximation of the operator F around some suitable model \bar{x}_k :

$$F(x) \cong F(\bar{x}_k) + F'(\bar{x}_k)\Delta x = F(\bar{x}_k) + F'(\bar{x}_k)(x - \bar{x}_k), \quad (18)$$

where $F'(\bar{x}_k)$ is the matrix formed by the partial derivatives of F . Therefore, we rewrite the problem (13) as

$$\begin{aligned} \min_x \quad & \left\| F'(\bar{x}_k) + r(\bar{x}_k) \right\|^2 \\ \text{s.t.} \quad & g(\bar{x}_k) \geq 0 \\ & g(\bar{x}_k) = \begin{pmatrix} \bar{x}_k - a \\ b - \bar{x}_k \end{pmatrix} \end{aligned} \quad (19)$$

where $r(\bar{x}_k) = F(\bar{x}_k) - y - F'(\bar{x}_k)\bar{x}_k$, and $g(x)$ is a vector of constraints, including constraints $x_i - a_i \geq 0$ and $b_i - x_i \geq 0$ that describe the bounds $a_i \leq x_i \leq b_i$ on

velocities x_i at different layers. Problem (20) can be solved using the Primal-Dual Interior-Point method [Sosa, 2012].

6 Primal Dual Interior-Point Method

To implement the Primal Dual Interior-Point method, we first rewrite our problem in a standard form as follows:

$$\begin{aligned} \min_x \quad & \frac{1}{2} \left\| F'(\bar{x}_k)x + r(\bar{x}_k) \right\|^2 \\ \text{s.t.} \quad & g(\bar{x}_k) - s = 0 \\ & s \geq 0 \end{aligned} \quad (20)$$

where $s \in \mathbb{R}^{2n}$ is a slack variable. Then we define the Lagrange function associated to problem (21) as:

$$l(\bar{x}_k, z, s, w) = \frac{1}{2} \left\| F'(\bar{x}_k)x + r(\bar{x}_k) \right\|^2 - (g(\bar{x}_k) - s)^T z - s^T w \quad (21)$$

with the Lagrangian multipliers $z, w \in \mathbb{R}^{2n}$, $(z, w) \geq 0$. For a given perturbation parameter $\mu > 0$, the perturbed Karush-Kuhn-Tucker (KKT) or necessary conditions are given by:

$$\hat{F}(\bar{x}_k, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T (F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) - s \\ z - w \\ SWe - \mu e \end{pmatrix} = 0 \quad (22)$$

where

$$\hat{F} : \mathbb{R}^{n+2n+2n} \rightarrow \mathbb{R}^{n+2n+2n} \quad S = \text{diag}(s_1, \dots, s_{2n}), \quad W = \text{diag}(w_1, \dots, w_{2n})$$

and $e = (1, \dots, 1) \in \mathbb{R}^{2n}$. It is easy to see that $z = w$, hence the perturbed KKT system (22) is rewritten as

$$\hat{F}(x, z, s, w) = \begin{pmatrix} F'(\bar{x}_k)^T (F'(\bar{x}_k)x + r(\bar{x}_k)) - \nabla g^T(\bar{x}_k)z \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{pmatrix} = 0 \quad (23)$$

thus the Jacobian associated to (23) is then computed as

$$F' \begin{pmatrix} x \\ z \\ s \end{pmatrix} = \begin{bmatrix} F'(\bar{x}_k)^T F'(\bar{x}_k) - \nabla g^T(\bar{x}_k) & 0_{n \times n} \\ \nabla g(\bar{x}_k) & 0_{n \times m} & -I_{m \times m} \\ 0_{m \times n} & S & Z \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \\ \Delta s \end{bmatrix} = - \begin{bmatrix} \nabla_x l(x, z, s) \\ g(\bar{x}_k) - s \\ SZe - \mu e \end{bmatrix} \quad (24)$$

System (24) can be reduced further by eliminating the third block of equations as follows. From the last block of equation in (24) we have

$$S\Delta z + Z\Delta s = -SZe + \mu e,$$

therefore

$$\begin{aligned} Z\Delta s &= -SZe + \mu e - S\Delta z \\ \Delta s &= -s + \mu Z^{-1}e - Z^{-1}S\Delta z, \end{aligned}$$

and then

$$\begin{aligned} \nabla g^T(\bar{x}_k)\Delta x - \Delta s &= \nabla g^T(\bar{x}_k)\Delta x + s - \mu Z^{-1}e + Z^{-1}e + Z^{-1}S\Delta z = -\nabla g^T(\bar{x}_k)x + s \\ \nabla g^T(\bar{x}_k)\Delta x + Z^{-1}S\Delta z &= \mu Z^{-1}e - g(\bar{x}_k) \end{aligned}$$

which allow us to write the reduced linear system

$$\begin{bmatrix} -F'(\bar{x}_k)^T F'(\bar{x}_k) & \nabla g^T(\bar{x}_k) \\ \nabla g(\bar{x}_k) & Z^{-1}S \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta z \end{bmatrix} = \begin{bmatrix} \nabla_x l(x, z, s) \\ Z^{-1}\mu e - g(\bar{x}_k) \end{bmatrix} \quad (25)$$

7 Conclusion

In summary, for this research project we propose to utilize the MOP technique to perform joint inversion of multiple data sets (Receiver functions and Surface Wave Dispersion). We will incorporate different weights in the MOP inversion scheme in order to map the Pareto Set (Solution Space) of receiver functions and surface wave dispersion measurements. We used the MOP technique to help characterize the crust and upper mantle of an ancient rift system in Texas using seismic data from USArray and Earthscope Network. We will extend the Primal Dual Interior Point Method (PDIP) and Truncated Singular Value Decomposition (TSVD) algorithm with the MOP in order to obtain high-resolution 3D imagery of Texas using teleseismic receiver functions, surface wave dispersion measurements, delay travel times, and gravity. We chose this optimization approach because we want to find the best possible solution for our nonlinear geophysics inverse problem. In geophysics, most inversion problems require finding some minimization. The optimization tech-

nique that we chose to solve our non-linear inverse problem requires the search of the global minimum and this technique will be able to define the entire solution based from using different weights to map the Pareto Set. From the Pareto Set, the MOP technique performs a direct search method that basically selects the final solution from a set of alternative solutions from the model space [Sambridge 1999a,b; Kozlovskaya 2000].

Acknowledgements We would like to take the time to thank the computational science, mathematical science, and computer science departments from University of Texas at El Paso (UTEP). We would also like to thank Ezer Patlan, Dr. Anibal Sosa, and Dr. Rodrigo Romero for all of their contributions to this work. This work was sponsored by the NSF CREST under Grant Cybershare HRD-0734825.

References

1. Bashir, L., S. S. Gao, K. H. Liu, and K. Mickus (2011), Crustal structure and evolution beneath the Colorado Plateau and the southern Basin and Range Province: Results from receiver function and gravity studies, *Geochem. Geophys. Geosyst.*, 12, Q06008, doi:10.1029/2011GC003563.
2. Bailey, I. W., M. S. M. S. Miller, K. Liu, and A. Levander (2012), Vs and density structure Beneath the Colorado Plateau constrained by gravity anomalies and joint inversions Of receiver function and phase velocity data, *J. Geophys. Res.*, 117, B02313, doi:10.1029/2011JB0085.
3. Chu, R., B. Schmandt, and D. V. Helmberger (2012), Upper mantle P velocity structure beneath the Midwestern United States derived from triplicated waveforms, *Geochem. Geophys. Geosyst.*, 13, Q0AK04, doi:10.1029/2011GC003818.
4. Dzierma, Y., Rabbel, W., Thorwart, M.M., Flueh, E.R., Mora, M.M. and Alvarado, G.E. (2011). The steeply subducting edge of the Cocos Ridge: Evidence from receiver functions beneath the northern Talamanca Range, south-central Costa Rica. *Geochemistry Geophysics Geosystems* 12: doi: 10.1029/2010GC003477. issn: 1525-2027.
5. Eaton, D. W., F. Darbyshire, R. L. Evans, H. Grutter, A. G. Jones, and X. Yuan (2009). The elusive lithosphere-asthenosphere boundary (LAB) beneath cratons. *Lithos* 109, 1-22; doi.org/10.1016/j.lithos.2008.05.009.
6. Fishcher, K. M., H. A. Ford, D. L. Abt, and C. A. Rychert (2010). The lithosphere-asthenosphere boundary. *Annual Review of Earth and Planetary Sciences* 38, 551-575; doi:10.1146/annurev-earth-040809-152438.
7. Gurrrola, H., Baker E. G., and Minster B. J., (1995), Simultaneous time-domain deconvolution with application to the computation of receiver functions, *J. geophys. Int.*, 120, 537-543.
8. Haber, E. & Oldenburg, D., 1997. Joint inversion: a structural approach, *Inverse Problems*, 13, 63-77.
9. Hackney, R.I., and W.E. Featherstone (2003). Geodetic versus geophysical perspectives of the gravity anomaly. *Geophysical Journal International*, 154(1), pp. 35-43.
10. Heiskanen, W. A. & Moritz, H., 1967. *Physical Geodesy*, W. H. Freeman and Company, San Francisco.
11. Julia, J., C. J. Ammon, R. Hermann, and M. Correig. Joint inversion of receiver function and surface wave dispersion observations. *Geophysics Int. J.* 142: 99-112, 2000.
12. Kozlovskaya, E., 2000. An algorithm of geophysical data inversion based on non-probabilistic Presentation of a-prior information and definition of Pareto-optimality, *Inverse Problems*, 16, 839-861.

13. Kumar, P., Yuan, X., Kind, R., and Mechie, J.: The lithosphere-asthenosphere boundary observed with USArray receiver functions, *Solid Earth*, 3, 149-159, doi:10.5194/se-3-149-2012, 2012.
14. Langston, C. A (1981), Evidence for the subducting lithosphere under southern Vancouver Island and western Oregon from teleseismic P wave conversions, *J. Geophys. Res.*, 86, 3857-3866.
15. Laske, G., G. Masters and C. Reif. Crust 2.0. The Current Limits of Resolution for Surface Wave Tomography in North America, *EOS Trans AGU*, 81, F897, 2000.<http://igpppublic.ucsd.edu/gabi/ftp/crust2/>
16. Lodge, A., and Helffrich G., (2009), Grid search inversion of teleseismic receiver functions, *Geophys. J. Int.* 178, 513-523.
17. Maceira, M. and Ammon, C.J., 2009. Joint inversion of surface wave velocity and gravity observations and its application to central Asian basins s-velocity structure. *J. Geophys. Res.*, 114, B02314, doi:10.1029/2007JB0005157.
18. Mickus, K., Stern, R.J., Keller, G.R. and Anthony, E. Y., 2009, Potential field evidence for volcanic rifted margin along the Texas Gulf Coast, *Geology*, Vol 37, No. 5, 387-390.
19. Mickus, K., and Keller, G. R., 1992, Lithospheric structure of the south-central U.S: *Geology*, v. 20, p. 335-338, doi: 10.1130/0091-763(1992)020;0335:LSOTSC₂.3.CO:2.
20. Mosher, S., 1998, Tectonic evolution of the Southern Laurentian Grenville orogenic belt: *Geological Society of America Bulletin*, v. 110, p. 1357-1375, doi:10.1130/0016-7606(1998)110;1357:TEOTSL₂.3.CO:2.
21. Moschetti, M. P., Ritzwoller, M. H., Lin, F. C., & Yang, Y. (2010b), Crustal shear wave velocity structure of the western United States inferred from ambient seismic noise and earthquake data, *Journal of Geophysical Research*, 115, B10, 1020, doi:10.1029/2010JB007448.
22. Obrebski, M., S. Kiselev, L. Vinnik, and J.-P. Montagner (2010), Anisotropic stratification beneath Africa from joint inversion of SKS and P receiver functions, *J. Geophys. Res.*, 115, B09313, doi:10.1029/2009JB006923.
23. Owens, T. J., H. P. Crotwell, C. Groves, and P. Oliver-Paul (2004). SOD: Standing Order for Data, *Seismological Research Letters* 75, 515520.
24. P.C. Hansen. The truncated SVD as a method for regularization. *BIT*, 27, 1987.
25. Raye U., Anthony E.Y., Stern R.J., Kimura J., Ren M., Qing C. and Tani K., (2011), Composition of the mantle lithosphere beneath south-central Laurentia: Evidence from peridotite xenoliths, Knippa, Texas, *Geosphere*, June 2011; v. 7; no. 3; p. 1-14.
26. Rowley, D.B., and Pindell, J.L., 1989, End Paleozoic/Early Mesozoic western Pangean reconstruction and its implications for the distribution of Precambrian and Paleozoic rocks around Meso-America: *Precambrian Research*, v. 42, p. 411444, doi: 10.1016/0301-9268 (89)90022-3.
27. Sambridge, M., 1999a. Geophysical inversion with a neighborhood algorithm I: searching a parameter space, *Geophys. J. Int.*, 138, 479-494.
28. Sambridge, M., 1999b. Geophysical inversion with a neighborhood algorithm II: appraising the ensemble, *Geophys. J. Int.*, 138, 727-746.
29. Sawyer, D.S., Buffler, R. T., and Pilger, R.H.P., Jr., 1991, The crust under the Gulf of Mexico Basin, in Salvador, A., ed., *The Gulf of Mexico Basin: Geological Society of America, The Geology of North America*, v. J. p. 53-72.
30. Shearer, P.M., *Introduction to Seismology*, Second Edition, Cambridge University Press, Cambridge, 2009.
31. Shen W., M. H. Ritzwoller, and V. Schulte-Pelkum (2013), A 3-D model of the crust and uppermost mantle beneath the Central and Western US by joint inversion of receiver functions and surface wave dispersion, *J. Geophys. Res. Solid Earth*, 118, doi:10.1029/2012JB009602.
32. Sosa, A., A. A. Velasco, L. Velasquez, M. Argaez, and R. Romero., 2013. Constrained Optimization framework for joint inversion of geophysical data sets, *Geophys. J. Int.*, 195, 197-211.
33. Spiegel, M. R. (1974), *Fourier Analysis with applications to boundary value problems*. Schaums Outlines. 1-187.

34. S. Stein and M. Wysession. *An Introduction to Seismology Earthquakes and Earth Structure*. Blackwell Publishing, 2006.
35. Tikhonov, A. N. & Arsenin, V.Y., 1977. *Solution of Ill-Posed Problems*, V.H. Winston & Sons, Washington, DC.
36. Vogel, C. R., *Computational Methods for Inverse Problems*. SIAM, 2002.
37. Vozoff, K. and D. L. B. Jupp, 1975: Joint inversion of geophysical data. *Geophys. J. Roy Astr. Soc.*, 42, 977-991.
38. Wilson, D. (2003), Imagining crust and upper mantle seismic structure in the southwestern United States using teleseismic receiver functions. *Leading Edge* 22, 232-237.
39. Wilson, D., and R. Aster (2005), Seismic imaging of the crust and upper mantle using Regularized joint receiver functions, frequency-wave number filtering, and Multimode Kirchhoff migration, *J. Geophys. Res.*, B05305, doi:10.1029/2004JB003430.
40. Wilson, D., R. Aster, J. Ni, S. Grand, M. West, W. Gao, W.S. Baldrige, and S. Semken (2005), Imaging the structure of the crust and upper mantle beneath the Great Plains, Rio Grande Rift, and Colorado Plateau using receiver functions, *J. Geophys. Res.*, 110, B05306, doi:10.1029/2004JB003492.
41. Zhu, L., Kanamori, H., (2000), Moho depth variation in southern California from teleseismic receiver functions, *J. Geophys. Res.*, 105, 2969-2980.