Efficient Implementation of Interior-Point Optimization Methods for Constrained Nonlinear Seismic Traveltime Tomography

*Donald C. Williams**

ABSTRACT

Traveltime tomography is a nonlinear ill-posed inverse problem. Effective solution methods require imposition of constraints on the estimated slowness model, beyond fit to traveltime data. These constraints may be imposed through penalty terms ("Tikhonov regularization") or explicitly, as equalities and inequalities to be satisfied by the slowness estimate. Interior point methods have generated a great deal of recent interest in the nonlinear programming community. These methods appear to be particularly promising for large scale inequality constrained optimization. This proposal presents evidence that the primal-dual interior point method performs effectively in application to inequality constrained traveltime tomography. It also suggests a program of further research on this topic.

1. INTRODUCTION

The aim of tomography is to generate a cross-sectional image (tomogram) of a region from the region's response to probing energy. The probing energy is typically applied and measured at the boundaries of the region to be imaged. The theory for using measured boundary data to infer internal properties of a region can be traced back to the work of Radon in the early 1900's (Radon 1917; cf., Deans 1983 for an English translation of Radon's paper). Over the last 30 years, however, tomographic techniques have become far-reaching and broadly applicable in many areas, which include astronomical imaging (e.g., Bracewell and Riddle 1967), diagnostic medicine (e.g., Cormack 1973, Scudder 1978, Kak 1979, Louis and Natterer 1983), and seismology (e.g., Aki et al. 1976, Dines and Lytle 1979, Bates et al. 1982). The focus of this thesis is the first-arrival seismic traveltime tomography problem, in which subsurface velocities are to be estimated from travel times of acoustic waves in the earth.

Several instances of successful acquisition, processing, and interpretation of crosswell seismic tomograms from hydrocarbon reservoirs exist in the literature. For example, single

*The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, Houston TX 77005-1892 USA, email: `donald@caam.rice.edu`

P-wave images have been used to interpret structure (Lo et al., 1990; Harris et al., 1990; Justice et al., 1992), lithology and porosity (Inderwiesen and Lo, 1990), and saturation (Lo et al., 1990). Researchers have found that without imposition of external constraints, the application of first-arrival seismic traveltime tomography as a high-resolution imaging tool for velocity estimation is fundamentally limited by recording aperture and the Fresnel zone of the direct wave (Williamson and Worthington, 1993; Rector and Washbourne, 1994). One such constraint is an inequality constraint on model slowness derived from Fermat's principle (Berryman, 1989; Berryman, 1990b; Berryman, 1991). We analyze the effectiveness and efficiency of interior-point optimization techniques for solving general constrained seismic traveltime tomography problems with both equality and inequality constraints.

For Ω (image region) an open, bounded, connected subset of $\mathbf{R}^2$ with a continuous acoustic wave speed velocity distribution, $c(\mathbf{x})$, and slowness (reciprocal wave speed), $s(\mathbf{x}) = 1/c(\mathbf{x})$, where $\mathbf{x} \in \Omega$; the *seismic traveltime tomography problem* seeks to estimate $s(\mathbf{x})$ from measured traveltime data, $t^d(\mathbf{x}_s, \mathbf{x}_r)$, where $\mathbf{x}_s, \mathbf{x}_r \in \Omega$ are known source and receiver positions. The measured traveltime data, $t^d(\mathbf{x}_s, \mathbf{x}_r)$, consist of measured times, from recordings of acoustic waves traveling through the image region between the source-receiver pairs.

Existing mathematical formulations for seismic traveltime tomography mostly cast the problem as an unconstrained optimization problem (Nolet, 1987; Berryman, 1990a). That is, given a vector of measured traveltime data, $\mathbf{t}^d$, the *seismic traveltime tomography problem* is: find $s(\mathbf{x})$ that solves

$$\min_{s(\mathbf{x}) \in A} \; E_p\left(\tau[s], \mathbf{t}^d\right) \tag{1}$$

where $E_p\left(\tau[s], \mathbf{t}^d\right)$ is a measure of the misfit (prediction error) between the computed traveltimes, $\tau[s](\mathbf{x}_s, \mathbf{x}_r)$, and the measured traveltime data. The nonlinear operator $\tau[s](\mathbf{x}_s, \mathbf{x}_r) : \mathcal{M} \rightarrow \mathcal{D}$ defines a traveltime functional that maps a certain Hilbert "model" space $\mathcal{M}$ into a Hilbert "data" space $\mathcal{D}$. The model space, $\mathcal{M}$, and data space, $\mathcal{D}$, are the slowness model space and traveltime data space. The set of admissible slowness models $A \subset \{s(\mathbf{x}) : s(\mathbf{x}) > 0\}$; the only physically meaningful models satisfy $s(\mathbf{x}) > 0$. However most formulations in the literature do not impose even this constraint explicitly. The computed traveltime functional, $\tau[s](\mathbf{x}_s, \mathbf{x}_r)$, is defined below in Sections (2.1), (2.2), and some of its useful properties for this work are given in Section (2.3). If E is the Euclidean norm distance between computed and measured data, we have the standard least-squares optimization problem for seismic traveltime tomography.

The mathematical formulation of seismic traveltime tomography as an optimization problem usually gives rise to ill-posed problems in the sense of Hadamard (cf., Hurt 1989, p.9). In fact, solutions of (1) are not unique even for complete and precise data (Coen, 1994; Bube and Meadows, 1996). See also, (Berryman, 1990a).

Remedies for the non-uniqueness of ill-posed seismic traveltime tomography problems include regularization, the addition of constraints, or some combination of both (Eriksson et al., 1995; Engl et al., 1996; Kaltenbacher, 1997). Some interesting constraints for the seismic traveltime tomography problem include: physical bounds derived from known subsurface stratification and material properties, smoothness measures, and Fermat feasibility constraints (Berryman, 1989).

The introduction of inequality constraints in seismic traveltime tomography begets a need for an effective and efficient optimization algorithm to solve the constrained problem. There exist at least four rather standard approaches in the literature for solving nonlinear constrained optimizations problems with inequality constraints. These approaches include: the method of squared slack variables (e.g., Tapia 1980), active set strategies (e.g., Gill et al. 1981), exterior-point methods (cf., Fiacco and McCormick 1990), and interior-point methods (e.g., El-Bakry, Tapia, Tsuchiya, and Zhang 1996; Wright 1997). Each of these approaches can, in general, be viewed as vehicles for obtaining a square nonlinear system that represents the original problem to varying degrees of accuracy (Tapia, 1997).

The method of squared slack variables has the disadvantages of increased dimension, increased degree of nonlinearity of inequality constraints which leads to loss of convexity. The use of active set strategies suffer from exponential combinatorial explosion in the number of possible active sets. Exterior-point method can fail to yield a feasible solution. Interior point methods are, however, fast when they work, and all iterates (approximations of the slowness model) generated by the algorithm are feasible. Thus algorithms based on interior-point techniques will always obtain a feasible model. This is important from the standpoint of obtaining a meaningful model, even if not optimal.

Within the family of interior-point methods, we choose to employ a primal-dual interior-point method since it tends to avoid the ill-conditioning problems of barrier methods as the iterates approach the solution (El-Bakry et al., 1996). Moreover, since the data sets for the constrained seismic traveltime tomography problems are often very large, the efficiency of the primal-dual method in handling a large number of inequality constraints is critical.

The proposal is organized as follows. Section 2 presents background material on Ray theory, Fermat's principle, and properties of the traveltime functional. Section 3 introduces the formulation of seismic traveltime tomography in terms of five different optimization problems: unconstrained least squares, regularized least squares, Fermat constrained least squares, Fermat constrained least squares with regularization, and Fermat constrained L_1 minimization. This section also presents the linearized seismic traveltime tomography problem. Section 4 presents the algorithms employed to solve the aforementioned optimization problems. Section 5 presents the discretization and preliminary results in terms of the linearized seismic traveltime tomography problem. A discussion of the preliminary results is given in section 6. The final section, 7, summarizes the proposed work of this thesis.

2. BACKGROUND

2.1. Ray Theory

Ray theory is applicable to the traveltime tomography problem under the assumption that the medium (image region) of interest doesn't vary rapidly with respect to the wavelength of seismic energy used to probe the medium (Wu and Aki, 1988). Thus, in this work, we assume that all seismic sources produce seismic wave energy with wavelengths much smaller than the size of the inhomogeneities encountered in the image region. Otherwise, other methods such as seismic diffraction tomography must be applied to solve the traveltime tomography problem (Aki and Richards, 1980, Section 13.3.5). The use of ray theory to model seismic wave energy has a well-known relation to using geometric optics to describe light wave propagation through lenses.

We are interested in first arrival traveltime of seismic wave energy. The theory of *viscosity solution* of Hamilton-Jacobi equations, developed by Crandall, Lions, and others (e.g., Lions 1982), gives sense to the first arrival traveltime, in a slowness (reciprocal velocity) field $s(\mathbf{x})$ defined on an open subset Ω of $\mathbf{R}^2$, as a possibly nondifferentiable solution of the eikonal equation

$$s(\mathbf{x}) = |\nabla\tau[s](\mathbf{x}_s, \mathbf{x})|. \tag{2}$$

Traveltime solutions, $\tau[s](\mathbf{x}_s, \mathbf{x})$, of the eikonal equation (2) represent the time taken by seismic wave energy propagating from a source point $\mathbf{x}_s = (x_s, z_s)$ to another point $\mathbf{x} \in \Omega$. The traveltime may be expressed as

$$\tau[s](\mathbf{x}_s, \mathbf{x}) = s(\mathbf{x}_s)|\mathbf{x} - \mathbf{x}_s| + O\left(|\mathbf{x} - \mathbf{x}_s|^2\right).$$

Wavefronts $\tau[s](\mathbf{x}_s, \mathbf{x}) = t$ are orthogonal trajectories of rays $\mathbf{x}(t)$. That is, $\mathbf{x}(t)$ is the integral curve of

$$\dot{\mathbf{x}}(t) = c^2(\mathbf{x})\nabla\tau(\mathbf{x}_s, \mathbf{x}) \tag{3}$$

where $c^2(\mathbf{x}) = s^{-2}(\mathbf{x})$ is the square of the seismic wave velocity. Substituting $\mathbf{p}(\mathbf{x}) = \nabla\tau[s](\mathbf{x}_s, \mathbf{x})$ in the ray equation (3) yields

$$\dot{\mathbf{x}}(t) = c^2(\mathbf{x})\mathbf{p}(\mathbf{x}).$$

The slowness vector $\mathbf{p}(t) \equiv \mathbf{p}(\mathbf{x}(t))$ also satisfies a differential equation obtained by differentiating $\mathbf{p}(\mathbf{x}(t))$ with respect to t. Then $\mathbf{x}(t)$, $\mathbf{p}(t)$ is a solution of Hamilton's equations,

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \nabla_{\mathbf{p}}H(\mathbf{x}, \mathbf{p}) \\ \dot{\mathbf{p}}(t) &= -\nabla_{\mathbf{x}}H(\mathbf{x}, \mathbf{p}). \end{aligned} \tag{4}$$

where the Hamiltonian function is defined as

$$\mathbf{H}(\mathbf{x}, \mathbf{p}) = \frac{1}{2}c^2(\mathbf{x})|\mathbf{p}|^2.$$

The first component of the solution of Hamilton's equations (4) is the trajectory of a ray. The trajectories of this system (4) for a point source at $\mathbf{x} = \mathbf{x}_s$ have initial conditions satisfying

$$\mathbf{x}(0) = \mathbf{x}_s$$
$$\|\mathbf{p}(0)\| = s(\mathbf{x}_s)$$

2.2. Fermat's Principle

The traveltime of a seismic wave is expressed as the integral of the slowness, $s(\mathbf{x})$, along a ray path connecting a given source and receiver pair. More generally, an expression for the traveltime along an arbitrary path, P, in a continuous slowness distribution, $s(\mathbf{x})$, is given by

$$\tau^P(s) = \int_p s(\mathbf{x})\, dl^P \tag{5}$$

where, dl defines the infinitesimal distance along the path, P. The physical principle of Fermat postulates that the correct ray path between two points is the one of least overall traveltime. Thus, according to Fermat's principle, the true traveltime from a source position, $\mathbf{x}_s$, to position, $\mathbf{x}$, in slowness medium $s(\mathbf{x})$ is given by

$$\tau(\mathbf{x}_s, \mathbf{x}) = \min_{P \in \overline{P}} \tau^P(s) \tag{6}$$

where, $\overline{P}$ defines the set of all continuous paths connecting $\mathbf{x}_s$ to $\mathbf{x}$. Thus, the true traveltime is an extremal of the traveltime functional, $\tau^P(s)$, or more specifically, a minimum. The true traveltime functional will hereafter be denoted by $\tau[s](\mathbf{x}_s, \mathbf{x}) = \tau(\mathbf{x}_s, \mathbf{x})$ to indicate the explicit dependence of the true traveltime on slowness. We compute this true traveltime for a given slowness medium, $s(\mathbf{x})$, by solving the eikonal equation, (2). Furthermore, the traveltime functional for an arbitrary path, P, will hereafter be denoted by $\tau^P[s](\mathbf{x}_s, \mathbf{x}) = \tau^P(s)$.

2.3. Properties of the Traveltime Functional

The interesting property of the traveltime functional associated with an arbitrary path, $\tau^P[s](\mathbf{x}_s, \mathbf{x})$, and the traveltime functional associated with the Fermat ray path, $\tau[s](\mathbf{x}_s, \mathbf{x})$, is that both are concave homogeneous functionals. It is obvious that

$$\tau^P[s](\mathbf{x}_s, \mathbf{x}) = \int_p s(\mathbf{x})\, dl^P$$

is a linear functional of the slowness model as integration is a linear functional of the integrand and concave as linear functionals are both convex and concave. Not so obvious is the fact that

$$\tau[s](\mathbf{x}_s, \mathbf{x}) = \min_{P \in \overline{P}} \tau^P[s](\mathbf{x}_s, \mathbf{x})$$

is a homogeneous concave functional. This result is known (Berryman, 1991), but included here for completeness in characterizing the feasibility region of the yet to be defined constrained optimization problem for nonlinear traveltime tomography. The convexity of the feasibility regions for the constrained problem will derive in part from the concavity of $\tau^P[s](\mathbf{x}_s, \mathbf{x})$, and $\tau[s](\mathbf{x}_s, \mathbf{x})$.

Proposition 1: *The traveltime functional for the Fermat ray path, $\tau[s](\mathbf{x}_s, \mathbf{x})$, is a homogeneous functional.*

Proof. Consider $\lambda > 0$. The traveltime functional with respect to γs is

$$\begin{aligned}\tau\left[\gamma s\right]\left(\mathbf{x}_s,\mathbf{x}\right) &= \min_P \tau^P \tau^P \left[\gamma s\right]\left(\mathbf{x}_s,\mathbf{x}\right) \\ &= \min_P \int_P \gamma s\left(\mathbf{x}\right) dl^P.\end{aligned}$$

From the linearity of τ^P, we have

$$\tau\left[\gamma s\right]\left(\mathbf{x}_s,\mathbf{x}\right) = \min_P \tau^P \left[\gamma s\right]\left(\mathbf{x}_s,\mathbf{x}\right) = \min_P \gamma\tau^P \left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) = \gamma \min_P \tau^P \left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) = \gamma\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right).$$

□

Proposition 2: *The traveltime functional for the Fermat ray path, $\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right)$, is a concave functional.*

Proof. Given slowness models s_1,s_2, and $\lambda \in [0,1]$. Let s be an affine combination of the given slowness models,

$$s = \lambda s_1 + \left(1-\lambda\right) s_2.$$

Let $P^*\left(s\right)$ be the Fermat ray path for slowness medium s. Thus

$$\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) = \tau^{P^*}\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right).$$

From the linearity of τ^P, we have

$$\begin{aligned}\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) &= \tau^{P^*}\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) \\ &= \tau^{P^*}\left[\lambda s_1 + \left(1-\lambda\right) s_2\right]\left(\mathbf{x}_s,\mathbf{x}\right) \\ &= \lambda\tau^{P^*}\left[s_1\right]\left(\mathbf{x}_s,\mathbf{x}\right) + \left(1-\lambda\right)\tau^{P^*}\left[s_2\right]\left(\mathbf{x}_s,\mathbf{x}\right)\end{aligned}$$

Since τ minimizes τ^P for any fixed model, it must be the case that

$$\tau^{P^*}\left[s_1\right]\left(\mathbf{x}_s,\mathbf{x}\right) \geq \tau\left[s_1\right]\left(\mathbf{x}_s,\mathbf{x}\right)$$

and similarly

$$\tau^{P^*}\left[s_2\right]\left(\mathbf{x}_s,\mathbf{x}\right) \geq \tau\left[s_2\right]\left(\mathbf{x}_s,\mathbf{x}\right)$$

Thus

$$\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right) \geq \lambda\tau\left[s_1\right]\left(\mathbf{x}_s,\mathbf{x}\right) + \left(1-\lambda\right)\tau\left[s_2\right]\left(\mathbf{x}_s,\mathbf{x}\right)$$

□

From Proposition (1) and (2), we have that the traveltime functional, $\tau\left[s\right]\left(\mathbf{x}_s,\mathbf{x}\right)$, is homogeneous and concave.

3. METHOD (SET UP AND QUESTIONS)

Seismic traveltime tomography can be subdivided into three main parts: first, a selection (possibly adaptive) of the parametrization and discretization used to represent the velocity model; second, the definition of an objective functional, which measures the prediction error between the computed traveltimes and observed traveltime data, with possible regularization and/or constraints; and, third, the optimization algorithm used to search the model space for the correct velocity model. In this section, we focus our attention on the second part of the seismic traveltime tomography problem, selection of the objective function, regularization, and constraints.

3.1. Formulations of Traveltime Tomography Problem

The optimization problem for nonlinear seismic traveltime tomography can be further subdivided into four main parts which include: (1) The definition of a measure (e.g., L_p-norm) for the prediction error (2) Regularization, and (3) Inequality constraints (e.g. physical bounds, Fermat feasibility). Additionally, as a part of the solution method, the penalty parameter which characterizes the trade-off between prediction error and regularization is selected.

3.1.1. Prediction Error

The primary data for the traveltime tomography problem are the measured traveltimes, $t^d(\mathbf{x}_s, \mathbf{x}_r)$, between source-receiver pairs located at position $(\mathbf{x}_s, \mathbf{x}_r)$ within the image region Ω, where $\mathbf{x}_s = (x_s, z_s)$, and $\mathbf{x}_r = (x_r, z_r)$. By selecting a measure,

$$E\left(\tau[s](\mathbf{x}_s, \mathbf{x}_r), t^d(\mathbf{x}_s, \mathbf{x}_r)\right),$$

for the prediction error between the computed traveltimes and the measured traveltime data, we can quantify how well the computed traveltimes, $\tau[s](\mathbf{x}_s, \mathbf{x}_r)$, approximate the measured traveltime data, $t^d(\mathbf{x}_s, \mathbf{x}_r)$. We define the L_p measure for the prediction error as,

$$E_p\left(\tau[s], \mathbf{t}^d\right) = \left\|\tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r)\right\|_p^p \tag{7}$$

where

$$\left\|\tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r)\right\|_p^p \equiv \sum_{s\in\mathcal{S},\, r\in\mathcal{R}} \left|\ \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r)\ \right|^p.$$

The sets $\mathcal{S}$ and $\mathcal{R}$ define indexed sets of sources and receivers. Using only the prediction error, (7), the *unconstrained optimization problem* for seismic traveltime tomography is expressed as

$$\min_s\ E_p\left(\tau[s], \mathbf{t}^d\right) \tag{8}$$

where $p = 2$ implies the standard least-squares (L_2-norm squared) problem. The unconstrained optimization problem, (8), is dependent on the measured traveltime data only. No *a priori* information about the problem is considered.

3.1.2. Regularization

A priori information about the structure of the slowness (velocity) model is typically available. It is natural to somehow introduce this information into the formulation of (8). In the case of an under-determined problem, there is more than one solution with zero prediction error. This implies that we must impose *a priori* information to remove the non-uniqueness. For example, we can seek a slowness model that is "closest" to a prescribed starting model, the "flattest" model which minimizes the first derivative of the model parameters, the "smoothest" model which minimizes the second derivative of the model parameters, or etc... (e.g., Phillips and Fehler 1991, Zelt and Barton 1998).

A common technique to overcome the ill-posedness of (8) is the regularization method of Tikhonov, which provides a one parameter family of problems. This regularization technique consist of solving an associated problem defined by a perturbation of (8),

$$\min_{\mathbf{s}} \ \Psi(s) = E_p\left(\tau[s], \mathbf{t}^d\right) + \rho^2 \|\mathbf{M}(\mathbf{s})\|_2^2 \tag{9}$$

where $\rho > 0$ is a Tikhonov regularization parameter (or penalty parameter) and $\mathbf{M} : \mathcal{M} \rightarrow \mathcal{A}$ is a bounded linear (model weighting) operator mapping the real Hilbert space $\mathcal{M}$ (slowness model space) into the real Hilbert space $\mathcal{A}$ (auxiliary data space). The operator $\mathbf{M}(\mathbf{s})$ typically represents a first difference operator for the flattest model, or a second difference operator for the smoothest model. The basic results of Tikhonov state that under certain weak conditions on $\tau[s]$, $\mathbf{M}$, s, and $\mathbf{t}^d$,

$$s(\mathbf{x}) = \arg\min_{\mathbf{s}} \ \Psi(s) = E_p\left(\tau[s], \mathbf{t}^d\right) + \rho^2 \|\mathbf{M}(\mathbf{s})\|_2^2$$

if and only if (1) there is an $\delta_\rho \geq 0$ so that

$$\begin{aligned} & s(\mathbf{x}) = \arg\min_{\mathbf{s}} \Psi(s) = E_p\left(\tau[s], \mathbf{t}^d\right) \\ \text{subject to} \quad & \\ & \|\mathbf{M}(\mathbf{s})\|_2^2 \leq \delta_\rho \end{aligned} \tag{10}$$

and (2) there is an $\varepsilon_\rho > 0$ so that

$$\begin{aligned} & s(\mathbf{x}) = \arg\min_{\mathbf{s} \in \mathbf{A}} \|\mathbf{M}(\mathbf{s})\|_2^2 \\ \text{subject to} \quad & \\ & E_p\left(\tau[s], \mathbf{t}^d\right) \leq \varepsilon_\rho \end{aligned} \tag{11}$$

(Tikhonov and Arsenin, 1977). The regularization parameter ρ can be viewed as an explicit "trade-off" parameter between prediction error and regularization in the objective of (9), however, the prediction error always has the same objective cost in the inequality constrained formulation of (10). Under certain conditions, (9), (10), and (11), suggest that the Tikhonov regularized unconstrained optimization problem, (9), can be thought of as a constrained optimization problem with inequality constraints. Furthermore, the standard regularized least-squares tomography problem results from letting $p = 2$ in (9). If we let $p = 2$ and $\rho = 0$, the regularized unconstrained least squares problem (9) reduces to the unconstrained least squares problem (8).

3.1.3. Inequality Constraints

We use inequality constraints $\mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l$ to represent physical bounds on the slowness model. These bounds derive from known subsurface stratification and material properties. The constraints are introduced into the constrained optimization problem below to establish a better representation of the physical problem.

Berryman (1990a), introduces additional inequality constraints derived from Fermat's principle. These inequality constraints eliminate trial slowness (velocity) models that produce any ray paths with traveltime smaller than the measured traveltime; these are infeasible models. The feasibility region for the Fermat constrained model will derive in part from Fermat principle (6).

From Fermat's principle we have that the true traveltime between a given source-receiver pair is a minimum of the functional given by (5). Thus, if we assume, at this point, that the traveltime data represent true traveltimes, then a slowness model $s(\mathbf{x})$ can only be correct if

$$\tau[s](\mathbf{x}_s, \mathbf{x}_r) \geq t^d(\mathbf{x}_s, \mathbf{x}_r). \tag{12}$$

All slowness models which satisfy (12) are said to be feasible. Therefore, (12) defines a feasibility region $B \equiv \left\{s(\mathbf{x}) : \tau[s](\mathbf{x}_s, \mathbf{x}_r) \geq t^d(\mathbf{x}_s, \mathbf{x}_r)\right\}$. Depending on the point of view taken, this constraint can be viewed as either a constraint on the computed traveltime or a constraint on the slowness model. One should observe that B is convex as it is the intersection of a family of half spaces.

It is realistic to assume that the measured traveltimes, $t^d(\mathbf{x}_s, \mathbf{x}_r)$, may contain measurement errors and so are not equal to the true traveltimes, $\tau^*(\mathbf{x}_s, \mathbf{x}_r)$. Thus, (12) should really be written

$$\tau[s](\mathbf{x}_s, \mathbf{x}_r) \geq \tau^*(\mathbf{x}_s, \mathbf{x}_r). \tag{13}$$

We assume that a bound on the error in t^d is known *a priori*:

$$\left|t^d(\mathbf{x}_s, \mathbf{x}_r) - \tau^*(\mathbf{x}_s, \mathbf{x}_r)\right| \leq \varepsilon \tag{14}$$

From (14), we deduce a lower bound on the true traveltime data in terms of the noise level ε and the measured data t^d,

$$\tau^*(\mathbf{x}_s, \mathbf{x}_r) \geq t^d(\mathbf{x}_s, \mathbf{x}_r) - \varepsilon.$$

Substituting this bound in the feasibility constraint (13) yields

$$\boxed{\tau[s](\mathbf{x}_s, \mathbf{x}_r) \geq t^d(\mathbf{x}_s, \mathbf{x}_r) - \varepsilon} \tag{15}$$

This relaxation of the Fermat feasibility constraint allows for over-estimated traveltime data due to noise. The inequality constraint given in (15) is used within the Fermat constrained optimization problems. The magnitude of ε in (15) is established with respect to *a priori* knowledge of measurement error in the traveltime data. The geophysicist's empirical knowledge of measurement error can be useful in establishing good estimates of

ε. To avoid overfitting of data, the optimization algorithms in this work limit the data fit to a tolerance on the order of ε. Using (9), (15), and the bound constraints, we define the following *general Fermat constrained optimization problem*,

$$\begin{array}{ll} & \min_{\mathbf{s}} \Psi(\mathbf{s}) = E_p\left(\tau[s], \mathbf{t}^d\right) + \rho^2 \|\mathbf{M}(\mathbf{s})\|_2^2 \\ \text{subject to} & \\ & \tau[s] \geq \mathbf{t}^d - \varepsilon \mathbf{e} \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{array} \tag{16}$$

where $\mathbf{e}$ is a vector of ones of appropriated dimension, and the bound constraints $\mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l$ reflect known information about subsurface stratification and physical bounds on wave speeds through materials. As in the unconstrained model, the computed traveltime $\tau[s](\mathbf{x}_s, \mathbf{x}_r)$ satisfy the eikonal equation (2).

3.2. Optimization Problems for Tomography

The following subsections contain descriptions of optimization problems used as mathematical formulations for the seismic traveltime tomography. In each case the computed traveltimes, $\tau[s](\mathbf{x}_s, \mathbf{x}_r)$, are determined by solving the eikonal equation, (2), with respect to a given slowness field, $s(\mathbf{x})$.

3.2.1. Least Squares Tomography

The optimization problem for unconstrained nonlinear least-squares tomography is obtained by letting $p = 2$ in (8), which yields,

$$\min_{\mathbf{s}} \ \Psi(s) = \frac{1}{2} \sum_{s,r} \left| \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) \right|^2. \tag{17}$$

Although other norm measures may be used for the prediction error (7), techniques that use the squared error, as in (17), are probably the most prevalent in the literature (Phillips and Fehler, 1991).

3.2.2. Regularized Least Squares Tomography

In a similar manner, the optimization problem for regularized unconstrained least squares tomography is obtained by letting $p = 2$ in (9), which yields,

$$\min_{\mathbf{s}} \ \Psi(s) = \frac{1}{2} \sum_{s,r} \left| \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) \right|^2 + \rho^2 \|\mathbf{M}(\mathbf{s})\|_2^2. \tag{18}$$

3.2.3. Least Squares Tomography with Fermat Constraints

The optimization problem for nonlinear least-squares tomography with Fermat feasibility constraints is obtained by letting $p = 2$ and $\rho = 0$ in (16), which yields,

$$\begin{array}{ll} & \min_{\mathbf{s}} \Psi(\mathbf{s}) = \frac{1}{2} \sum_{s,r} \left| \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) \right|^2 \\ \text{subject to} & \\ & \tau[s] \geq \mathbf{t}^d - \varepsilon \mathbf{e} \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{array} \tag{19}$$

The feasibility region for (19) is

$$\mathcal{F} = \left\{ s : \tau\left[s\right]\left(\mathbf{x}_s, \mathbf{x}_r\right) \geq t^d\left(\mathbf{x}_s, \mathbf{x}_r\right) - \varepsilon,\ s_u \geq s \geq s_l \right\}$$

which is convex.

3.2.4. Regularized Least Squares Tomography with Fermat Constraints

The optimization problem for regularized nonlinear least-squares tomography with Fermat feasibility constraints is obtained by letting $p = 2$ in (16), which yields,

$$\begin{array}{ll} & \min\limits_{\mathbf{s}} \Psi\left(\mathbf{s}\right) = \frac{1}{2} \sum\limits_{s,r} \left|\ \tau\left[s\right]\left(\mathbf{x}_s, \mathbf{x}_r\right) - t^d\left(\mathbf{x}_s, \mathbf{x}_r\right)\ \right|^2 + \rho^2 \left\|\mathbf{M}\left(\mathbf{s}\right)\right\|_2^2 \\ \text{subject to} & \\ & \tau\left[s\right] \geq \mathbf{t}^d - \varepsilon \mathbf{e} \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{array} \tag{20}$$

The feasibility region for (20) is also given by $\mathcal{F}$.

3.2.5. L_1 Tomography with Fermat Constraints

We can replace the L_2 measure for the prediction error between the computed traveltimes and traveltime data in (20) with an L_1 measure by letting $p = 1$ in (16), which yields,

$$\begin{array}{ll} & \min\limits_{\mathbf{s}} \Psi\left(\mathbf{s}\right) = \sum\limits_{s,r} \left|\ \tau\left[s\right]\left(\mathbf{x}_s, \mathbf{x}_r\right) - t^d\left(\mathbf{x}_s, \mathbf{x}_r\right)\ \right| + \rho^2 \left\|\mathbf{M}\left(\mathbf{s}\right)\right\|_2^2 \\ \text{subject to} & \\ & \tau\left[s\right] \geq \mathbf{t}^d - \varepsilon \mathbf{e} \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{array} \tag{21}$$

For exact traveltime data (i.e., $\varepsilon = 0$), and no regularization ($\rho = 0$), linearized L_1 tomography problem with Fermat constraints reduces to a linear programming problem. If regularization is employed, ($\rho > 0$), and traveltime data assumed exact (i.e., $\varepsilon = 0$) the linearized L_1 tomography problem with Fermat constraints reduces to a simple-quadratic programming problem.

3.3. Linearized Traveltime Tomography Problem

Consider the unconstrained nonlinear least-squares seismic traveltime tomography problem, (17). Least-squares tomography as presented above seeks to minimize a nonlinear functional

$$\Psi\left(\mathbf{s}\right) = \frac{1}{2} \sum_{i \in \mathcal{I}} \mid R_i\left(s\right) \mid^2 \tag{22}$$

where $\mathcal{I} \equiv \left\{\left(s, r\right)_i : s \in \mathcal{S},\ r \in \mathcal{R}\right\}$ is an indexed set of source-receiver pairs, $\mathcal{S}$ and $\mathcal{R}$ define indexed sets of sources and receivers. The ith traveltime residual is defined as

$$R_i\left(s\right) = \tau\left[\mathbf{s}\right]\left(\mathbf{x}_s, \mathbf{x}_r\right) - t^d\left(\mathbf{x}_s, \mathbf{x}_r\right)$$

with $i = r + (s-1) \cdot |\mathcal{R}|$.

An affine approximation (linearization) of the traveltime residual, $R(s)$, about a slowness model, s_c, is

$$\begin{aligned} \widehat{\Phi}_c(\mathbf{s}) &= \tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) + \mathbf{D}\tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r)\,\delta\mathbf{s} \\ &= \mathbf{D}\tau[\mathbf{s}]\,\delta\mathbf{s} + R(s) \end{aligned} \tag{23}$$

For each traveltime functional, $\tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r)$, if we consider the perturbation induced by a small change in the model, $\delta\mathbf{s}$,

$$\tau[\mathbf{s} + \delta\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r) = \tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r) + \mathbf{D}\tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r)\,\delta\mathbf{s} + O\left(\delta\mathbf{s}^2\right)$$

To first order in $\delta\mathbf{s}$, we have,

$$\mathbf{D}\tau[\mathbf{s}](\mathbf{x_s}, \mathbf{x_r})\,\delta\mathbf{s} = \tau[\mathbf{s} + \delta\mathbf{s}](\mathbf{x_s}, \mathbf{x_r}) - \tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r)$$

If $dl^*[\mathbf{s}]$ is the infinitesimal increment of the Fermat least-time ray along a given path for model $\mathbf{s}$, then

$$\tau[\mathbf{s} + \delta\mathbf{s}](\mathbf{x_s}, \mathbf{x_r}) - \tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r) = \int (\mathbf{s} + \delta\mathbf{s})\,dl^*[\mathbf{s} + \delta\mathbf{s}] - \int \mathbf{s}dl^*[\mathbf{s}]$$

Since the ray paths do not change to first order in $\delta\mathbf{s}$ near the one of least time,

$$\int \mathbf{s}dl^*[\mathbf{s} + \delta\mathbf{s}] = \int \mathbf{s}\left[dl^*[\mathbf{s}] + \delta dl^*\right] \simeq \int \mathbf{s}dl^*[\mathbf{s}]$$

Substituting this approximation into the previous equation yields,

$$\begin{aligned} \tau[\mathbf{s} + \delta\mathbf{s}](\mathbf{x_s}, \mathbf{x_r}) - \tau[\mathbf{s}](\mathbf{x}_s, \mathbf{x}_r) &\simeq \int \mathbf{s}dl^*[\mathbf{s}] + \int \delta\mathbf{s}dl^*[\mathbf{s} + \delta\mathbf{s}] - \int \mathbf{s}dl^*[\mathbf{s}] \\ &= \int \delta\mathbf{s}dl^*[\mathbf{s} + \delta\mathbf{s}] \\ &\simeq \int \delta\mathbf{s}dl^*[\mathbf{s}] \end{aligned} \tag{24}$$

The approximation in (24) results from Fermat's Principle and it follows that for each source-receiver pair,

$$\mathbf{D}\tau[\mathbf{s}](\mathbf{x_s}, \mathbf{x_r})\,\delta\mathbf{s} \simeq \int \delta\mathbf{s}dl^*[\mathbf{s}] \tag{25}$$

where the number of source-receiver pairs is $m = |\mathcal{S}| \times |\mathcal{R}|$. The slowness field, $\mathbf{s}(x, z) \in \Omega \subset \mathbf{R}^2$, is defined over a two dimensional image region and is typically discretized into a rectangular grid by specifying slowness values at the nodal points of a grid and choosing an interpolation scheme such as bilinears between the nodes or as a special case, using cells of constant slowness. With respect to the cell slowness model, let l_{ij} be the length of the i-th ray path through the j-th cell. Thus,

$$l_{ij} = \int_{P_i \cap cell_j} dl_i^*[\mathbf{s}]$$

For a model discretization with n cells of constant slowness, we write equation (25) as

$$\mathbf{D}\tau_i[\mathbf{s}]\left(\mathbf{x_s}, \mathbf{x_r}\right)\delta\mathbf{s} \simeq \int \delta \mathbf{s} dl_i^*\left[\mathbf{s}\right] = \sum_{j=1}^{n} l_{ij}\delta s_j, \qquad i = 1, \ldots, m$$

which can be express in matrix notation as,

$$\mathbf{D}\tau[\mathbf{s}]\left(\mathbf{x_s}, \mathbf{x_r}\right)\delta\mathbf{s} \simeq J\left(\mathbf{s}\right)\delta\mathbf{s}.$$

Substituting $\mathbf{D}\tau[\mathbf{s}]\left(\mathbf{x_s}, \mathbf{x_r}\right)\delta\mathbf{s}$ into the linearized traveltime tomography model, (23), yields,

$$\widehat{\Phi}_c\left(s\right) = R\left(s_c\right) + J\left(s_c\right)\left(s - s_c\right) \tag{26}$$

where the Jacobian matrix $J\left(s_c\right) \in \mathbf{R}^{m\times n}$ is the Jacobian of $R\left(s\right)$, $\widehat{\Phi}_c\left(s\right) : \mathbf{R}^n \to \mathbf{R}^m$. Typically, $m > n$. We can not in general expect to find a slowness model, s_+, such that $\widehat{\Phi}_c\left(s_+\right) = 0$ since this is an over-determined system of linear equations. However, a logical way to solve the nonlinear least-squares problem (22) using the linearization given in (26) is to choose the next approximation to the slowness model, as the solution of the linear least-squares problem

$$\min_{\delta s} \ \frac{1}{2}\left\|\widehat{\Phi}_c\left(s_c + \delta s\right)\right\|_2^2 \tag{27}$$

The updated approximation of the slowness model is taken as $s_+ = s_c + \delta s$. Iterates generated in this manner represent a Gauss-Newton algorithm (Dennis and Schnabel, 1996).

3.4. Special Properties of L_1 Tomography

3.4.1. Linear Programming Problem

For exact traveltime data (i.e., $\varepsilon = 0$) and no regularization ($\rho = 0$), the L_1 tomography problem with Fermat constraints, (21), reduces to

$$\begin{aligned} &\min_{\mathbf{s}} \ \Psi\left(\mathbf{s}\right) = \sum_{s,r} \left|\ \tau\left[s\right]\left(\mathbf{x}_s, \mathbf{x}_r\right) - t^d\left(\mathbf{x}_s, \mathbf{x}_r\right)\ \right| \\ \text{subject to}\quad & \\ &\tau\left[s\right] \geq \mathbf{t}^d \\ &\mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{aligned}$$

Replacing the nonlinear traveltime residual by the linearization given in (26) from Section 3.3 yields

$$\begin{aligned} &\min_{\delta\mathbf{s}} \ \Phi\left(\delta\mathbf{s}\right) = \sum_{i\in\mathcal{I}} \left|\ \left(J\left(\mathbf{s}\right)\delta\mathbf{s}\right)_i + R_i\left(\mathbf{s}\right)\ \right| \\ \text{subject to}\quad & \\ &J\left(\mathbf{s}\right)\delta\mathbf{s} + R\left(\mathbf{s}\right) \geq 0 \\ &\mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \end{aligned}$$

With the Fermat feasibility constraint enforced,

$$J(\mathbf{s})\,\delta\mathbf{s} + R(\mathbf{s}) \geq 0,$$

the objective function can be expressed as,

$$\begin{aligned}\Phi(\delta s) &= \sum_{i\in\mathcal{I}} (J(\mathbf{s})\,\delta\mathbf{s})_i + \sum_{i\in\mathcal{I}} R_i(\mathbf{s}) \\ &= e^T (J(\mathbf{s})\,\delta\mathbf{s}) + \|R(\mathbf{s})\|_1 .\end{aligned}$$

We now define a *coverage vector*, c, such that,

$$c^T = e^T J(\mathbf{s}).$$

The coverage vector is the column sums of the Jacobian (ray path matrix). Therefore, the element c_j is the total length of the ray path arcs passing through slowness cell-block, s_j. Utilizing the coverage vector, c, and neglecting the constant term in the objective function, we obtain the following linear programming problem,

$$\boxed{\begin{array}{ll} & \min\limits_{\delta s} \Phi(\delta\mathbf{s}) = c^T \delta\mathbf{s} \\ s.t. & \\ & J(\mathbf{s})\,\delta\mathbf{s} + R(\mathbf{s}) \geq 0 \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \\ where & \\ & \mathbf{s} = \mathbf{s}_0 + \delta\mathbf{s} \end{array}} \tag{28}$$

3.4.2. Quadratic Programming Problem

For exact traveltime data (i.e., $\varepsilon = 0$) and regularization $\rho > 0$, the L_1 tomography problem with Fermat constraints, (28), reduces to

$$\min_{\mathbf{s}} \Psi(\mathbf{s}) = \sum_{s,r} \left|\; \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) \;\right| + \rho^2 \|\mathbf{M}(\mathbf{s})\|_2^2$$

subject to

$$\begin{aligned}\tau[s] &\geq \mathbf{t}^d \\ \mathbf{s}_u \geq \mathbf{s} &\geq \mathbf{s}_l.\end{aligned}$$

In a manner similar to the derivation of the linear programming problem (28), we obtain the following simple quadratic programming problem

$$\boxed{\begin{array}{ll} & \min\limits_{\delta s} \Phi(\delta\mathbf{s}) = c^T \delta\mathbf{s} + \rho\delta\mathbf{s}^T Q \delta\mathbf{s} \\ s.t. & \\ & J(\mathbf{s})\,\delta\mathbf{s} + R(\mathbf{s}) \geq 0 \\ & \mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \\ where & \\ & \mathbf{s} = \mathbf{s}_0 + \delta\mathbf{s} \end{array}} \tag{29}$$

where $Q = \mathbf{I}$ for model weighting $\mathbf{M}(\mathbf{s}) \equiv \mathbf{I}\,\delta\mathbf{s}$.

4. ALGORITHMS

4.1. Unconstrained Gauss-Newton

Gauss-Newton is the approach typically taken to solve unconstrained nonlinear seismic traveltime tomography problems (Berryman, 1990a). This approach minimizes a linear model of the nonlinear objective, and takes successive steps to improve the model estimate as characterized by the following algorithmic framework,

Algorithm 1: (Gauss-Newton (GN))

1. Set an initial slowness model, $s = s_0$
2. (Forward problem). Compute traveltime, t_0, corresponding to the initial slowness, s.
3. Compute Jacobian, $J(s)$, corresponding to the initial slowness, s.
4. Compute traveltime residuals, $R(s)$.
5. While (norm of traveltime residual, $R(s)$, not sufficiently small) {

   Compute $J(s)$.

   (Linearized inversion step). Find the model correction δs as the solution to the linear least-squares problem

$$\min_{\delta s} |J(s)\,\delta s + R(s)|^2$$

   (Update model)

$$s \leftarrow s + \delta s$$

   Compute $R(s)$.

   }

The behavior of Gauss-Newton is dependent on the size of the residual at the solution. The linear least squares problem that is solved in the linearized inversion step derives from a linearization of the traveltime residual. Thus, second derivative information is not present in Gauss-Newton's method. The convergence or non-convergence of Gauss-Newton's method is dependent on whether the omitted second derivative information is significant. That is, the degree of nonlinearity of the problem. The method can fail to converge to a solution for problems with "large" residuals. (Dennis and Schnabel, 1996).

4.2. Constrained Gauss-Newton

The addition of constraints in the seismic traveltime tomography problem leads us to solve a constrained optimization problem. One approach to solving the constrained least squares optimization problem, 19, for seismic traveltime tomography is to use the GN algorithm with the constraints applied at the level of the subproblem (linearized inversion step). This approach leads to the constrained Gauss-Newton framework for nonlinear inversion,

Algorithm 2: (Constrained Gauss-Newton (CGN))

1. Set an initial slowness model, $s = s_0$
2. (Forward problem). Compute traveltime, t_0, corresponding to the initial slowness, s.
3. Compute Jacobian, $J(s)$, corresponding to the initial slowness, s.
4. Compute traveltime residuals, $R(s)$.
5. While (norm of traveltime residual, $R(s)$, not sufficiently small) {

   Compute $J(s)$.

   (Linearized inversion step). Find the model correction δs as the solution to the linear least-squares problem

$$\min_{\mathbf{s}\delta} \; |J(s)\,\delta s + R(s)|^2$$

subject to

$$J(s)\,\delta s + \tau[s] \geq \mathbf{t}^d - \varepsilon\mathbf{e}$$
$$\mathbf{s}_u \geq \mathbf{s} + \delta s \geq \mathbf{s}_l$$

   (Update model)

$$s \leftarrow s + \delta s$$

   Compute $R(s)$.

   }

The CGN algorithm derives directly form the GN algorithm by adding linearized Fermat constraints and bound constraints to the linearized inversion step of GN. Although the CGN algorithm solves a sequence of constrained quadratic subproblems, the sequence of quadratic subproblems are not representative of Sequential Quadratic Programming (SQP) in the standard mathematical optimization sense (Boggs, 1996). Defining the CGN algorithm for the general constrained nonlinear seismic traveltime tomography problem (19) with constraints applied in the linearized inversion step is very different than applying SQP or primal-dual interior point, where the constraints are addressed through the Lagrangian (Boggs, 1996; Wright, 1997).

4.3. Successive Quadratic Programming (SQP)

SQP in the standard mathematical optimization sense minimizes a quadratic model of the Lagrangian of the original constrained problem subject to linearized constraints. We will consider SQP with respect to the following three levels,

1. SQP with no second derivative information.
2. SQP with an approximation (BFGS) to second derivative information
3. SQP with analytic (Hessian) second derivative information.

For notational simplicity, we consider the constrained optimization problem given in (19) without the bound constraints,

$$\begin{array}{ll} & \min\limits_{\mathbf{s}} \Psi(\mathbf{s}) = \frac{1}{2}\sum\limits_{s,r} \left| \tau[s](\mathbf{x}_s, \mathbf{x}_r) - t^d(\mathbf{x}_s, \mathbf{x}_r) \right|^2 \\ \text{subject to} & \\ & \tau[s] \geq \mathbf{t}^d - \varepsilon \mathbf{e} \end{array} \tag{30}$$

The Lagrangian of (30) is

$$\begin{aligned} l(s,z) &= \frac{1}{2} R(s)^T R(s) - z(R(s) + \varepsilon) \\ &= \frac{1}{2}\left(\tau[s] - t^d\right)^T \left(\tau[s] - t^d\right) - z\left(\tau[s] - t^d + \varepsilon\right) \end{aligned}$$

where $z \equiv$ Lagrangian multipliers for the inequality constraint. The derivative of the Lagrangian is

$$\begin{aligned} \nabla l(s,z) &= J(s)^T R(s) - J(s)^T z \\ &= \nabla\tau[s]^T \left(\tau[s] - t^d\right) - \nabla\tau[s]^T z. \end{aligned}$$

where $J(s)$ is defined as the Jacobian of the traveltime residual, $R(s)$. Similarly, the second derivative of the Lagrangian is

$$\begin{aligned} \nabla^2 l(s,z) &= J(s)^T J(s) + \sum_{i=1}^{m} R_i(s) \nabla^2 R_i(s) - \sum_{i=1}^{m} z_i \nabla^2 R_i(s) \\ &= J(s)^T J(s) + \sum_{i=1}^{m} \left(R_i(s) - z_i\right) \nabla^2 R_i(s) \\ &= \nabla\tau[s]^T \nabla\tau[s] + \sum_{i=1}^{m} \left(\tau_i[s] - t_i^d - z_i\right) \nabla^2 \tau_i[s]. \end{aligned}$$

Thus, the SQP subproblem for (30) is

$$\begin{array}{ll} & \min\limits_{\delta s} [\nabla l(s,z)]^T \delta s + \frac{1}{2}\delta s^T \left(\nabla^2 l(s,z)\right) \delta s \\ \text{subject to} & \\ & J(s)\,\delta s + R(s) + \varepsilon \geq 0 \end{array} \tag{31}$$

or

$$\begin{array}{ll} & \min\limits_{\delta s} \left[\nabla\tau^T\left(\tau - t^d\right) - \nabla\tau^T z\right]^T \delta s \\ & \quad + \frac{1}{2}\delta s^T \left(\nabla\tau^T\nabla\tau + \sum_{i=1}^{m}\left(\tau_i - t_i^d - z_i\right)\nabla^2\tau_i\right)\delta s \\ \text{subject to} & \\ & \nabla\tau\,\delta s + \tau - t^d + \varepsilon \geq 0 \end{array}$$

The basic SQP algorithmic framework for (30) is thus given as,

Algorithm 3: (SQP)

1. Given initial slowness model, $s = s_0$, and a merit function ϕ.
2. (Forward problem). Compute traveltime, t_0, corresponding to the initial slowness, s.
3. Compute Jacobian, $J(s)$, corresponding to the initial slowness, s.
4. Compute Hessian, $B(s) = \nabla^2 l(s, z)$, or an approximation $B(s) \simeq \nabla^2 l(s, z)$.
5. Compute traveltime residuals, $R(s)$.
6. While (ϕ not sufficiently small) {

   Compute $J(s)$, $B(s)$

   Find model correction δs as the solution of (31).

   (Update model)

$$s \leftarrow s + \delta s$$

   Update $R(s)$, ϕ.

   }

In the SQP algorithm, the merit function ϕ represents a quantitative measure of the quality of the slowness model estimate (e.g., not necessarily the residual $R(s)$). Whether the analytic Hessian is used in the SQP algorithm or some approximation is dependent on the degree of nonlinearity and computational expense. We will investigate the degree of nonlinearity of the constrained traveltime tomography problem by evaluating the performance of CGN, SQP with no second derivative information, SQP with an approximation (BFGS) to second derivative information, and SQP with analytic (Hessian) second derivative information. For example, ignoring second derivative terms, we obtain the following SQP subproblem.

$$\boxed{\begin{array}{ll} & \min_{\delta s} \ \left[\nabla\tau^T\left(\tau - t^d\right) - \nabla\tau^T z\right]^T \delta s + \frac{1}{2}\delta s^T\left(\nabla\tau^T\nabla\tau\right)\delta s \\ \text{subject to} & \\ & \nabla\tau\delta s + \tau - t^d + \varepsilon \geq 0 \end{array}} \tag{32}$$

Although, the constraint equation for the CGN subproblem is the same as that given for the SQP subproblem, (32), the objective functional for the CGN subproblem differs from the SQP subproblem. The objective of the CGN subproblem is

$$\begin{aligned} \min_{\delta s} \frac{1}{2}\left\|R(s) + J(s)\,\delta s\right\|_2^2 &= \min_{\delta s} \frac{1}{2}[R(s)^T R(s) + R(s)^T J(s)\,\delta s \\ &\quad + \delta s^T J(s)^T R(s) + \delta s^T J(s)^T J(s)\,\delta s] \\ &= \min_{\delta s}\left[R(s)^T J(s)\,\delta s + \frac{1}{2}\delta s^T J(s)^T J(s)\,\delta s\right] \\ &= \min_{\delta s}\left[\left(\tau - t^d\right)^T \nabla\tau\delta s + \frac{1}{2}\delta s^T\nabla\tau^T\nabla\tau\delta s\right] \end{aligned}$$

Thus the CGN subproblem can be written as

$$\boxed{\begin{array}{ll} & \min\limits_{\delta s}\ \left[\nabla\tau^T\left(\tau - t^d\right)\right]^T\delta s + \frac{1}{2}\delta s\left(\nabla\tau^T\nabla\tau\right)\delta s \\ \text{subject to} & \\ & \nabla\tau\delta s + \tau - t^d + \varepsilon \geq 0 \end{array}} \tag{33}$$

In comparing (32) and (33), we see that the SQP formulation (32) incorporates the additional $-z^T\nabla\tau\delta s$ term in the objective. Thus, there exists Lagrange multiplier information in the SQP approach that is not present in adding constraints to the linearized inversion step of the unconstrained Gauss-Newton framework. Furthermore, using better approximation of the Hessian (second derivative information) allow the SQP algorithm to better represent the nonlinear nature of the objective of the fully nonlinear seismic traveltime tomography problem.

The subproblem of the CGN algorithm minimizes a quadratic model of the original objective function subject to a linear model of the constraints as opposed to minimizing a quadratic model of the Lagrangian subject to a linear model of the constraints, in the SQP algorithm. The problem with a method that solves

$$\begin{array}{ll} & \min\limits_{x} f(x) \\ \text{s.t.} & \\ & h(x) = 0 \\ & g(x) \geq 0 \end{array}$$

where f and g are nonlinear functions by minimizing a model of the function subject to linearized constraints is illustrated in the following example. Consider a subproblem that employs a local quadratic approximation to f at x_k. That is,

$$f(x) \simeq \nabla f^T\Delta x + \frac{1}{2}\Delta x^T B_k \Delta x$$

where B_k is taken as the Hessian. While this is a reasonable approximation to use if the constraints are linear, the presence of nonlinear constraints makes this choice inappropriate. For example, the problem

$$\begin{array}{ll} & \min\limits_{x}\ -x_1 - \frac{1}{2}(x_2)^2 \\ \text{s.t.} & \\ & (x_1)^2 + (x_2)^2 - 1 = 0 \end{array}$$

has a local solution $(1,0)$ satisfying standard Newton assumptions (Boggs, 1996), but at the point $(1+\varepsilon, 0)$ the approximating quadratic program is

$$\begin{array}{ll} & \min\limits_{\Delta x}\ -\Delta x_1 - \frac{1}{2}(\Delta x_2)^2 \\ \text{s.t.} & \\ & \Delta x_1 = \frac{1}{2}\varepsilon(2+\varepsilon)(1+\varepsilon)^{-1} \end{array}$$

which is unbounded regardless of how small ε is. To take nonlinearities in the constraints into account while maintaining the linearity of the constraints in the subproblem, the

SQP method uses a quadratic model of the Lagrangian function as the objective (e.g., subproblem (31)). The major theoretical ideas of nonlinear constrained optimization problems can be found in numerous sources. See, e.g., (Gill et al., 1981), or (Luenberger, 1989). Convergence rate theory can be found in (Ortega and Rheinboldt, 1970), and (Dennis and Schnabel, 1996). The effect nonlinearities will be addressed in solving the constrained seismic traveltime tomography problem in full generality.

4.4. Direct Nonlinear Interior Point

In this section, we introduce the primal dual interior-point algorithm in the context of the linearized traveltime tomography problem with Fermat constraints (19),

$$\begin{aligned} &\min_{\delta s} \Phi(\delta \mathbf{s}) = \tfrac{1}{2}\left\|J(s)\,\delta s - R(s)\right\|_2^2 \\ \text{s.t.}\quad & \\ &J(\mathbf{s})\,\delta\mathbf{s} + R(\mathbf{s}) \geq 0 \\ &\mathbf{s}_u \geq \mathbf{s} \geq \mathbf{s}_l \\ \text{where}\quad & \\ &s \equiv s_0 + \delta s \end{aligned}$$

To make the presentation of the primal dual interior-point algorithm lucid, the bound constraints on the slowness model are dropped and the linearized L_2-Fermat optimization problem is expressed as

$$\begin{aligned} &\min_{\delta s} \Phi(\delta s) = \tfrac{1}{2}\left\|J(s)\,\delta s - R(s)\right\|_2^2 \\ s.t.\quad & \\ &g(s) = J(s)\,\delta s - R(s) + \varepsilon \geq 0 \end{aligned}$$

Adding slack variables, y, yields

$$\begin{aligned} &\min_{\delta s} \Phi(\delta s) = \tfrac{1}{2}\left\|J(s)\,\delta s - R(s)\right\|_2^2 \\ s.t.\quad & \\ &g(s) - y = 0 \\ &\qquad\quad y \geq 0 \end{aligned} \tag{34}$$

The Lagrangian for (34) is given by,

$$L(s, y, \lambda, \eta) = \Phi(s) + \lambda^T\left[g(s) + y\right] - \eta^T y$$

The gradient of the Lagrangian with respect to primal variables (s, y) is given by,

$$\begin{aligned} \nabla_s L &= \nabla\Phi(s) + \nabla g(s)\,\lambda \\ \nabla_y L &= \lambda - \eta \end{aligned}$$

The corresponding Karush-Kuhn-Tucker (KKT), optimality, condition are

$$F(s, y, \lambda) = \begin{bmatrix} \nabla\Phi(s) + \nabla g(s)\,\lambda \\ g(s) - y \\ \Lambda Y e \end{bmatrix} = 0, \quad (y, \lambda) \geq 0 \tag{35}$$

and the Perturbed KKT conditions are

$$F_\mu(s, y, \lambda) = \begin{bmatrix} \nabla\Phi(s) + \nabla g(s)\lambda \\ g(s) - y \\ \Lambda Y e - \mu e \end{bmatrix} = 0, \quad (y, \lambda) \geq 0 \tag{36}$$

There is an inherent difficulty associated with applying a Newton's method or a modification thereof to the KKT conditions for a given problem (El-Bakry et al., 1996). This difficulty involves what is considered in the optimization community as "sticking to the walls". That is, if a variable becomes zero prematurely during an iteration of the algorithm, global convergence may be precluded. We avoid the problem of "sticking to the walls" by applying a Newton method to the perturbed-KKT conditions given in (36). The typical primal-dual step is a Newton step on the perturbed-KKT. For example,

$$\begin{bmatrix} \nabla^2_{ss}L(s,\lambda) & \nabla g(s) & 0 \\ \nabla g(s)^T & 0 & I \\ 0 & Y & \Lambda \end{bmatrix} \begin{bmatrix} \Delta s \\ \Delta\lambda \\ \Delta y \end{bmatrix} = - \begin{bmatrix} \nabla\Phi(s) + \nabla g(s)\lambda \\ g(s) - y \\ \Lambda Y e - \mu e \end{bmatrix}$$

where

$$\mu = \sigma\frac{\lambda^T y}{m}, \text{ defines a pertubation parameter}$$
$$\sigma \in [0,1] \text{ defines a centering parameter}$$

A modification of the primal-dual interior point method presented by El-Bakry, Tapia, Tsuchiya, and Zhang in (1992) has been implemented in MATLAB. The algorithm is utilized to obtain some initial experimental results for the linearized traveltime tomography problem. Although, not included in the simplified derivation above, the bounds on updated slowness, $\mathbf{s}^{(k+1)}$, is enforced in the numerical experiments. This updated slowness satisfies linearized Fermat constraints.

In the fully nonlinear problem the constraint region will be updated (or adjusted) with each model update. The updated slowness model may be infeasible with respect to the new linearized feasibility constraints. Thus it may be advantageous at each iteration of the interior point algorithm to "loosely" solve the optimization subproblems. That is, at each iteration, solve the perturbed-KKT system to within a specified tolerance with respect to a centrality measure (Argaez and Tapia, 1995). The tolerance will be loose when the algorithm is far from the solution, and as the algorithms converges the tolerance will tighten.

5. PRELIMINARY RESULTS

The aim of the numerical experiments in this section is to illustrate the viability of the ideas of this thesis in the context of linearized inversion. Synthetic seismic tomographic data sets for borehole-to-borehole arrays are computed and inverted. In all cases, inversion is performed on the linearized traveltime tomography problem. The linearized problem is an essential component of the techniques we will apply to solving the nonlinear traveltime tomography problem, as well as a useful tool in its own right. There exist obvious inefficiencies in the current numerical implementation which will be dealt with in the thesis.

For the numerical experimentation, we consider a two-dimensional source-receiver cross-hole configuration involving boreholes, Figure(1). The region to be imaged lies

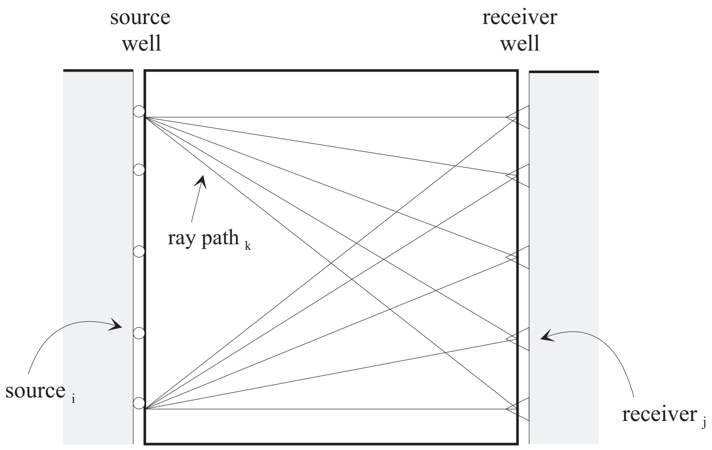

FIG. 1. Schematic of the cross-hole configuration for borehole tomography. The region to be imaged resides in the vertical plane containing the source and receiver wells.

in the vertical plane containing the boreholes.

5.1. Discretization

The image region is typically discretized into a rectangular grid either using cells of constant slowness or by specifying slowness values at the nodal points of a grid and choosing an interpolation scheme such as bilinears between the nodes. At present, the former discretization approach is employed to represent the slowness in $\mathbf{R}^2$ (cf., Figure(2)). The cell discretization defines a n-dimensional slowness model space for the problem, with s_i being the slowness of the i-th cell. The m-dimensional data space of the problem derives from the measured traveltime data between the m source-receiver pairs. We let, $\mathbf{t}^d$ be an m-dimensional vector of measured traveltime with $t^d(\mathbf{x}_s, \mathbf{x}_r)$ the measured traveltime between the source-receiver pairs.

Although, the chosen cross-hole configuration is geometrically simple, it allows us to investigate the behavior of algorithms used to solve the linearized inversion problem. The following considerations are made in the numerical experiments:

- Effects of data perturbation (i.e., $\pm$ random perturbation of traveltime data, t^d)
- Error in fitting traveltime data (norm of constraint error)
- Relative error in model (slowness)

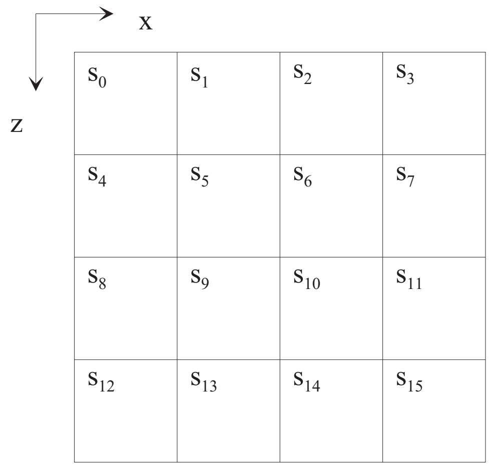

FIG. 2. Schematic of the cell discretization of the image space.

- Influence of Fermat feasibility constraint on model estimate

The numerical experiments are based on the optimization problems presented in Section 3.2. A linearization of each model is assumed. The problems include: unconstrained least-squares model (UC-LS), L_1-norm objective with Fermat feasibility constraints (l_1-Fermat), L_2-norm objective and Fermat feasibility constraints (l_2-Fermat), unconstrained least-squares with Tikhonov regularization (UC-LS-R), and L_2-norm objective with Fermat feasibility and Tikhonov regularization constraints (l_2-Fermat-R).

The unconstrained least-squares problem (UC-LS) and regularized unconstrained least squares problem (UC-LS-R) were solved directly via normal equations. The l_1-Fermat problem was solved as a linear programming problem with a linear interior point solver (LIPSOL). A primal-dual interior-point algorithm is employed to solve the convex-quadratic models l_2-Fermat and l_2-Fermat-R.

We considered a physical model with 15 sources and 15 receivers at know locations within a two dimensional rectangular region as illustrated in Figure(1). The slowness model is discretized into a 9×9 cell structure in a manner similar to Figure(2). It follows that the dimension of the data space is 225 and the dimension of the model space is 81. The ray path matrix, $J(s_e)$, is obtained with respect to a straight ray assumption, where s_e is the exact slowness for the region. The straight ray assumption is most appropriate for the constant slowness model that contains no anomalies. Although useful for models with low contrast, ray bending must be taken into account for high contrast media. Forward modelling is used to generate the synthetic traveltime data, t_*^d, for each specified media. The synthetic traveltime data, t_*^d, is obtain using Fermat's principle,

$$t_*^d = J(s_e)\, s_e$$

and taken to be exact. The measured traveltime, t^d, is given by

$$t^d = t_e^d \pm t_\varepsilon$$

	$\left\| \tau\left[s_f\right] - t^d \right\|_2$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	3.7533e-09	4.9795e-07	2.2144e-03	3.2194e-05	2.2171e-03
0.000285	1.0687e-04	2.2391e-04	2.2356e-03	1.1510e-04	2.2351e-03
0.002660	1.1495e-03	2.4193e-03	2.8907e-03	9.9991e-04	3.5309e-03
0.028466	1.1773e-02	2.5125e-02	2.4197e-02	1.0837e-02	2.3187e-02
0.299592	1.1074e-01	2.4165e-01	2.2732e-01	1.1410e-01	2.2695e-01
2.992190	1.2103e+00	2.7322e+00	1.9199e+00	1.1716e+00	2.3222e+00

Table 1. (Problem A). Illustrates error in the traveltime data fit for each optimization model considered as a function of noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e{-}6$.

where t_ε defines a percentage perturbation of the exact traveltime data. The measured traveltime, t^d, is exact for $t_\varepsilon = 0$. A constant initial slowness is assumed, $s_{init}(\mathbf{x}) \equiv 5.0\ \sec/km$, and traveltimes between the first source and the fifteen receivers range from 4.4 sec (horizontal) to 6.1 sec (diagonal) for this constant initial model.

5.2. Results from Numerical Experiments

We employ a primal-dual interior point algorithm to solve the following optimization problems for the linearized seismic traveltime tomography problem:

- Unconstrained Least Squares
- Least Squares w/ Fermat Constraints
- L_1Tomography w/ Fermat Constraints
- Regularized Unconstrained Least Squares
- Regularized Least Squares w/ Fermat Constraints

The numerical experiments are based on the true model being constant, constant with low slowness anomaly, and constant with high slowness anomaly.

5.2.1. Constant Model

Problem A: (Constant Slowness Model)

This problem considers a constant slowness model with the exact slowness taken to be, $s_e(\mathbf{x}) \equiv 4.4\ \sec/km$. For the optimization problem that use regularization $\rho = 1.0e-6$. The results for the traveltime data fit and model error are summarized in Table(1) and Table(2) respectively. The tables illustrate the effects of various levels of perturbation, in the measured traveltime data, on the optimization models considered.

	$\frac{\|s_f - s_e\|_2}{\|s_e\|_2}$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	1.3855e-04	8.8708e-03	2.390974e-003	7.8680e-06	3.8639e-05
0.000285	1.9705e+00*	1.0111e+00	9.350135e-003	3.6378e-05	4.4834e-05
0.002660	3.1560e+01*	2.6543e+00	2.055057e-001	2.8536e-04	2.7466e-04
0.028466	1.6982e+02*	2.7287e+00	8.842487e-002	3.4580e-03	4.3390e-03
0.299592	9.7152e+02*	2.0718e+00	1.174613e+000	2.4944e-02	2.8177e-02
2.992190	1.9050e+04*	8.9350e-01	1.919864e+000	2.9085e-01	2.9912e-01

Table 2. (Problem A). Relative error in the computed slowness model with respect to noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e-6$. The * indicates negative slowness values.

In Table(1), observe that all models considered give an accurate data fit when the traveltime data is exact. As the percent perturbation in the data increases, the accuracy of the data fit deteriorates in all cases. The l_1-Fermat and l_2-Fermat optimization models yield similar results for the traveltime data fit.

Table(2), indicate models with negative slowness values can be obtained when using UC-LS for the data perturbations considered. Adding regularization to the unconstrained problem, as in the UC-LS-R eliminates the negative slowness values which appear in the solution of the UC-LS problem. Table(2) also illustrates that the deterioration in the relative error of the slowness model with respect to data perturbation is more rapid for UC-LS and l_1-Fermat, in comparison with the l_2-Fermat model. The results for the regularized models, UC-LS-R and l_2-Fermat-R are very similar. Since the exact slowness model is constant for Problem A, the results are only illustrated for UC-LS-R and l_2-Fermat-R optimization with 2.9% data perturbations. These results are illustrated in Figure(3) and Figure(4) respectively.

The effects of regularization in UC-LS-R and the l_2-Fermat-R on data fit and accuracy of slowness model estimate in illustrated in Figure(5). The results in Figure(5) are given with respect to the traveltime data perturbation fixed at approximately 2.9%.

5.2.2. Low Contrast Anomaly

Problem B: (Low contrast anomaly)

This problem considers a constant slowness model with a low slowness (high velocity) anomaly near the top of the image region. The anomaly has constant slowness with 10% contrast from the background. The exact slowness is taken to be constant, $s_e(\mathbf{x}) \equiv 4.4$ sec$/km$, with a low slowness, $s_e(\mathbf{x}) \equiv 4.0$ sec$/km$, anomaly near the top of the model. For the optimization problems that use regularization $\rho = 1.0e-6$.

The fit of the traveltime data and model error for Problem B are summarized in Table(3) and Table(4) respectively. The tables illustrate the effects of various levels of perturbation, in the measured traveltime data, on the optimization models considered.

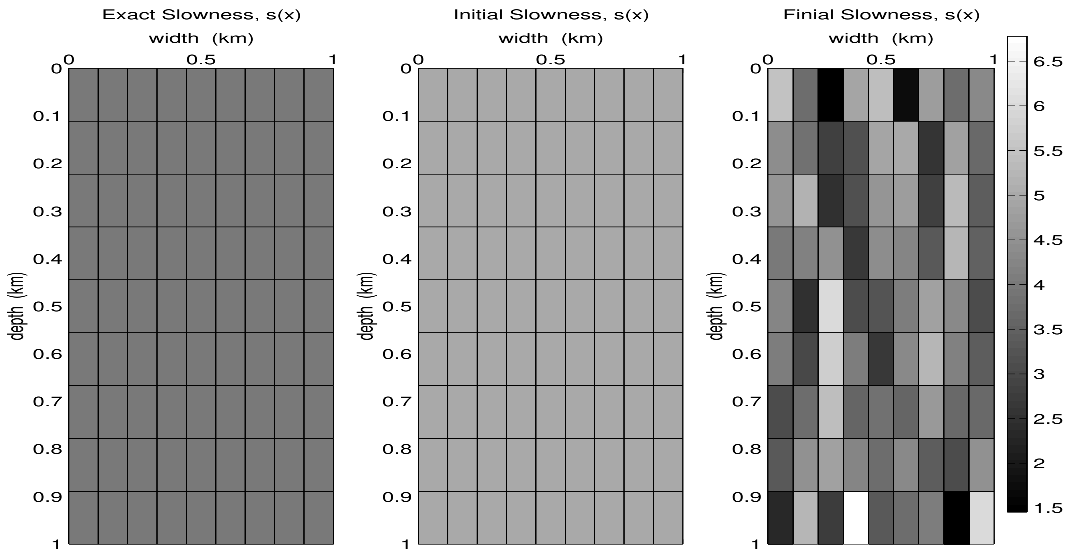

FIG. 3. (Problem A). Illustration of exact, initial, and final slowness distributions for the UC-LS-R optimization model with 2.9% perturbation error in traveltime data. The regularization parameter, $\rho = 1.0e-6$.

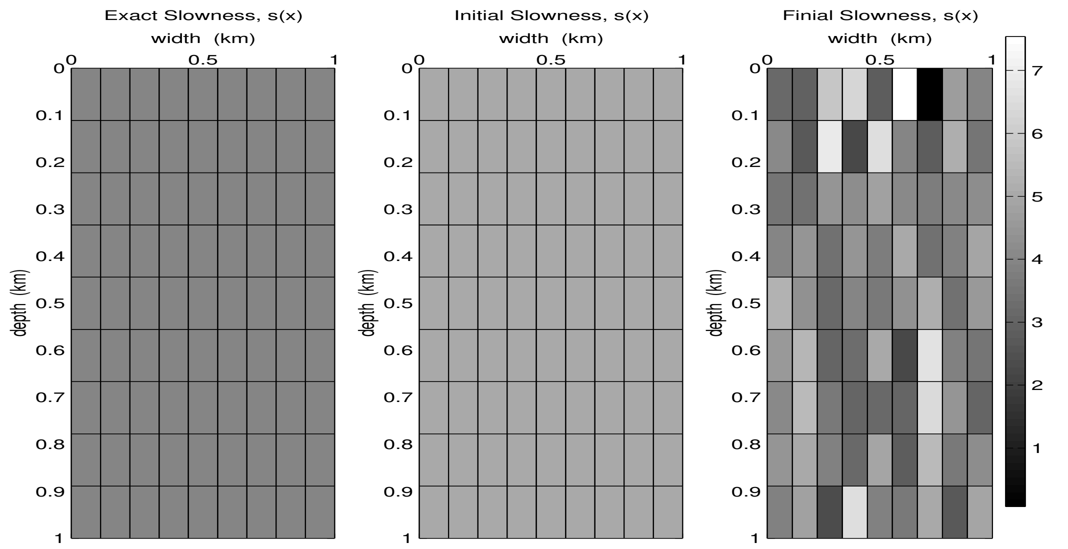

FIG. 4. (Problem A). Illustration of exact, initial, and final slowness distributions for the l_2-Fermat-R optimization model with 2.9% perturbation error in traveltime data. The regularization parameter, $\rho = 1.0e-6$.

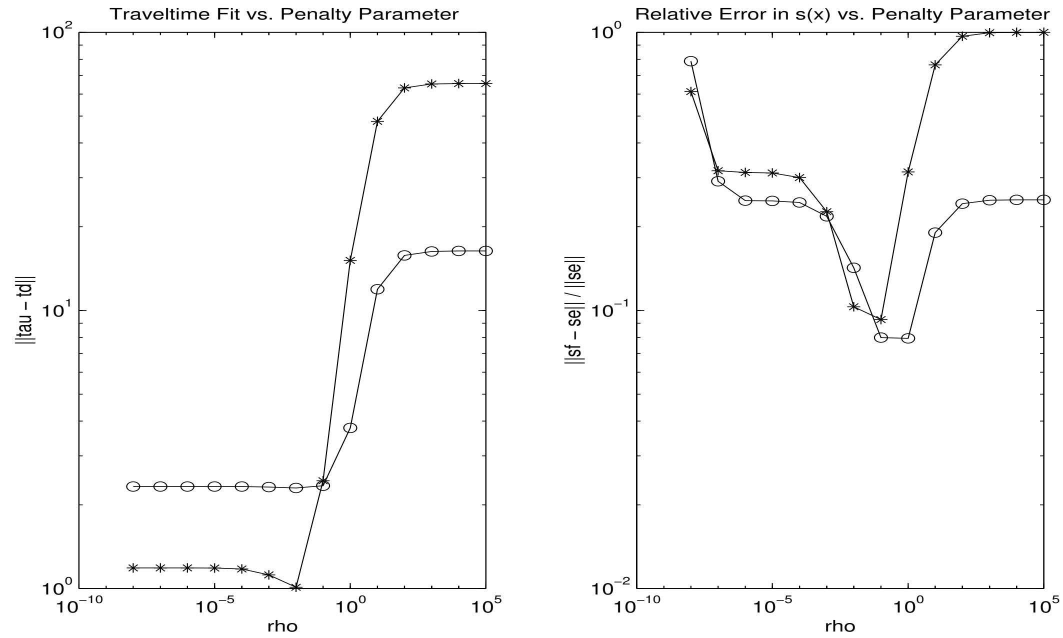

FIG. 5. (Problem A). The UC-LS-R, "$*$", and the l_2-Fermat-R, "o", optimization models are plotted as a function of the regularization parameter, ρ.

	$\left\|\tau^*(s_f) - t^d\right\|_2$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	4.0222e-09	9.7098e-07	2.0889e-03	3.4704e-05	2.0908e-03
0.000285	1.3167e-04	2.4615e-04	2.0937e-03	1.0822e-04	2.0839e-03
0.002660	1.1556e-03	2.6282e-03	3.3668e-03	1.1186e-03	3.3903e-03
0.028466	1.2408e-02	2.5668e-02	2.4918e-02	1.2213e-02	2.5869e-02
0.299592	1.1582e-01	2.8131e-01	2.5014e-01	1.1781e-01	2.6538e-01
2.992190	1.2437e+00	2.7912e+00	2.6034e+00	1.2561e+00	2.7505e+00

Table 3. (Problem B). Illustrates error in the traveltime data fit for each optimization model considered as a function of noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e{-}6$.

	$\frac{\|s_f - s_e\|_2}{\|s_e\|_2}$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	1.3668e-04	1.4140e-02	3.0479e-03	7.8433e-06	3.3736e-05
0.000285	2.4805e+00*	1.0592e+00	2.1126e-02	2.7620e-05	4.8512e-05
0.002660	1.5142e+01*	2.0141e+00	1.4041e-01	3.2714e-04	2.4861e-04
0.028466	1.9756e+02*	2.5059e+00	1.3958e-01	2.2527e-03	2.1856e-03
0.299592	9.1551e+02*	2.2965e+00	5.9058e-01	2.8964e-02	4.1320e-02
2.992190	2.2077e+04*	9.5754e-01	2.0232e+00	3.2741e-01*	3.3297e-01

Table 4. (Problem B). Relative error in the computed slowness model with respect to noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e - 6$. The * indicates negative slowness values.

In Table(3), observe that, as the percent perturbation in the data space increases, the accuracy of the data fit deteriorates in all cases. For the low contrast problem, the l_2-Fermat problem obtains a more accurate model estimate than the l_1-Fermat model (see, e.g. Table(4)). The fact that models with negative slowness values can be obtained when using UC-LS or UC-LS-R in the presence of noisy data is indicated in Table(4). Table(4) also illustrates that the deterioration in the accuracy of the slowness model estimate with respect to data perturbation is more rapid for the UC-LS in comparison to the l_2-Fermat model. The results for the regularized models, UC-LS-R and l_2-Fermat-R are very similar.

I would like to note that, as one might expect for a least-squares data fitting problems, the data is fit with approximately half of the un-imposed Fermat feasibility constraints are slightly violated by UC-LS. However, for this problem, the constraint violations are on the order of the data perturbations. The l_2-Fermat optimization problem with no noise in the measured traveltime data is used to illustrate the slowness model in Figure(6). Figure(7) illustrates the ghosts which manifest as stripes in the numerical solution of cross-well tomography as a result of the non-uniqueness of the solution (cf., Berryman 1990b). Although the l_2-Fermat optimization model gives improved results over the UC-LS and l_1-Fermat models, striping still effects the model. Figure(7) show the l_2-Fermat optimization model with 0.029% percent traveltime data perturbation.

5.2.3. High Contrast Anomaly

Problem C: (High contrast anomaly)

Consider a constant slowness model with a low slowness (high velocity) anomaly near the top of the image region. The anomaly has constant slowness with 80% contrast from the background. The exact slowness is taken to be constant, $s_e(\mathbf{x}) \equiv 4.4$ sec $/km$, with a low slowness, $s_e(\mathbf{x}) \equiv .88$ sec $/km$, anomaly near the top of the model. For the optimization problems that use regularization $\rho = 1.0e - 6$.

The results for the traveltime data fit and error in the recovered slowness model for Problem C are summarized in Table(5) and Table(6) respectively. The tables illustrate

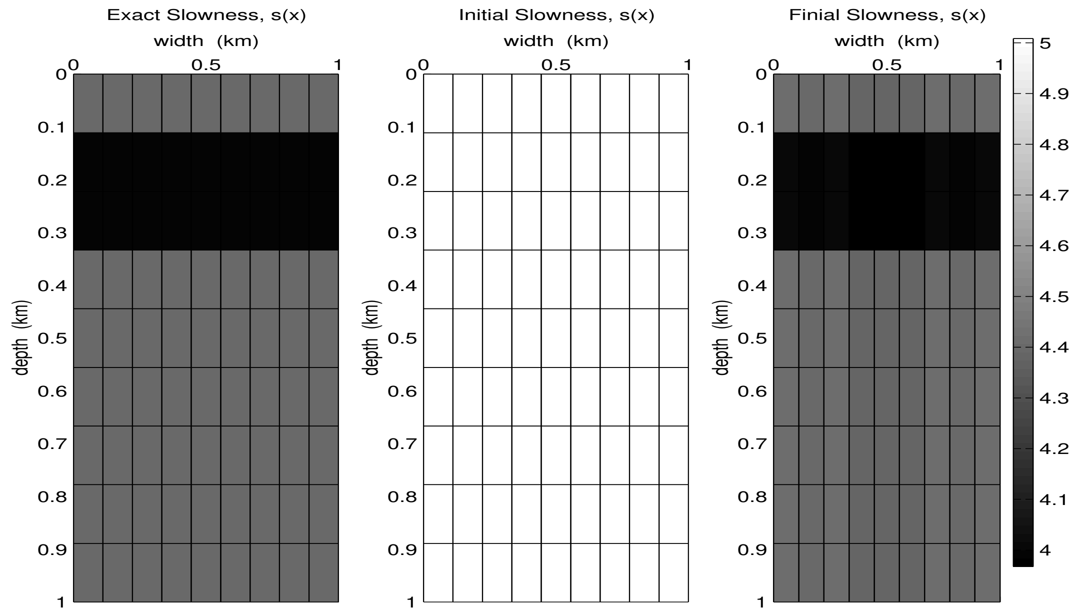

FIG. 6. (Problem B). Illustration of exact, initial, and final slowness distributions for the l_2-Fermat optimization model with no noise. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness anomaly, $s_e(\mathbf{x}) = 4.0\ s/km$, in the upper part of the region. The starting model is constant, $s_i(\mathbf{x}) = 5\ s/km$.

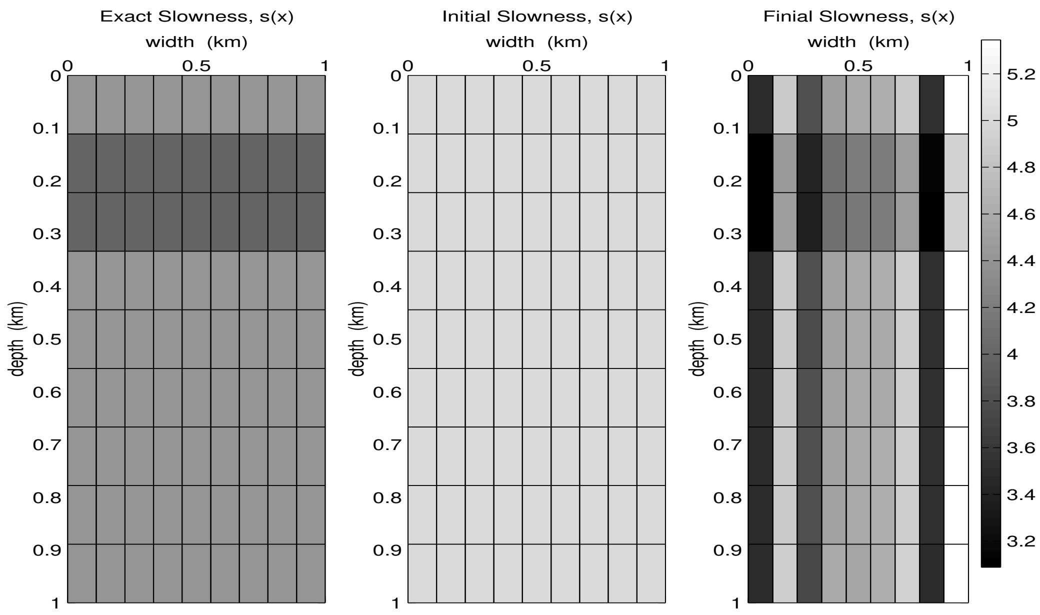

FIG. 7. (Problem B). Illustration of exact, initial, and final slowness distributions for the l_2-Fermat optimization model with 0.029% noise in the traveltime data. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness, anomaly, $s_e(\mathbf{x}) = 4.0\ s/km$, in the upper part of the region. The starting model is constant $s_i(\mathbf{x}) = 5\ s/km$.

	$\left\Vert \tau^*(s_f) - t^d \right\Vert_2$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	3.3683e-009	1.3383e-005	2.4565e-003	3.2844e-003	2.4602e-003
0.000285	1.0082e-004	2.1494e-004	2.4486e-003	3.2919e-003	2.4635e-003
0.002660	1.0714e-003	2.4687e-003	3.3266e-003	3.5255e-003	3.4464e-003
0.028466	9.6079e-003	2.2834e-002	2.3233e-002	1.0756e-002	2.0364e-002
0.299592	1.1324e-001	2.2300e-001	2.2281e-001	9.6065e-002	2.1508e-001
2.992190	1.1232e+000	2.4474e+000	2.0942e+000	1.0609e+000	2.0826e+000

Table 5. (Problem C). Illustrates error in the traveltime data fit for each optimization model considered as a function of noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e{-}6$.

	$\dfrac{\left\Vert s_f - s_e \right\Vert_2}{\left\Vert s_e \right\Vert_2}$				
% pertubation	UC-LS	l_1-Fermat	l_2-Fermat	UC-LS-R	l_2-Fermat-R
0	1.2294e-04	1.7416e-02	3.6312e-04	7.2907e-04	4.4051e-05
0.000285	1.5891e+00*	4.0678e-01	1.0366e-03	7.3350e-04	5.6800e-05
0.002660	1.5893e+01*	6.1951e-01	1.5282e-02	7.6527e-04	2.6459e-04
0.028466	1.6558e+02*	5.5758e-01	2.7485e-02	3.5662e-03	2.9876e-03
0.299592	1.5346e+03*	5.6351e-01	1.7534e-01	3.1090e-02	3.1282e-02
2.992190	1.8725e+04*	5.9411e-01	4.0712e-01	3.2065e-01*	3.8663e-01

Table 6. (Problem C). Relative error in the computed slowness model with respect to noise (percent perturbations of measured traveltime, t^d, from true traveltime data, t^*). For the optimization models with regularization, $\rho = 1e - 6$. The * indicates negative slowness values. The exact slowness model is $s_e\,(bfx) \equiv 4.4 \sec/km$, with a low slowness anomaly, $s_e\,(bfx) \equiv .88 \sec/km$, near the top of the model.

the effects of various levels of perturbation in the measured traveltime data on the optimization models considered.

In Table(5), observe that as the percent perturbation in the data space increases, the accuracy of the traveltime data fit deteriorates in all cases. The table also show that the traveltime data fit is within 2% for all models when the data perturbation is less that 0.0284%. There exist slight violations of the Fermat feasibility for the UCLS model. The magnitude of the feasibility violations range from $1.0e{-}4$ for 0.000285% data perturbation to 0.2 for 2.9% data perturbation.

Table(6), illustrates the fact that models with negative slowness values can be obtained when using UC-LS or UC-LS-R. The largest negative slowness values are approximately $1.2e4$ and $4.02e-1$ for UC-LS and UC-LS-R respectively. Table(6) also illustrates that the deterioration in the relative error in the slowness model estimate with respect to data perturbation is more rapid for the UC-LS model than either the l_1-Fermat or l_2-Fermat models. In comparison to the low contrast example in Problem B, the l_1-Fermat model yields better results for the high contrast model. The slowness models obtain with the l_2-

Fermat optimization problem are still slightly better than those obtained by the l_1-Fermat optimization problem for the high contrast slowness model.

With regards to image reconstruction, the resolution of the solution (i.e., the anomaly) for the UC-LS optimization model is severely degraded by striping for extremely small data perturbations (e.g., 0.000285%), and unrecognizable for data perturbations of approximately 0.00266%. In contrast, although stripes do begin to appear in the l_2-Fermat optimization model for approximately 0.29% data perturbations, the anomaly is still clearly recognizable for 2.9% data perturbation as illustrated in Figure(8). For comparison, the UC-LS optimization model on Problem C with 2.9% data perturbation is given in Figure(9). Figure(8) and Figure(9) illustrate the definite advantage of the model with Fermat feasibility constraints over the unconstrained problem UC-LS. In resolving high contrast anomalies, the results for the regularized models, UC-LS-R and l_2-Fermat-R are very similar. This similarity is illustrated in Figure(10) and Figure(11).

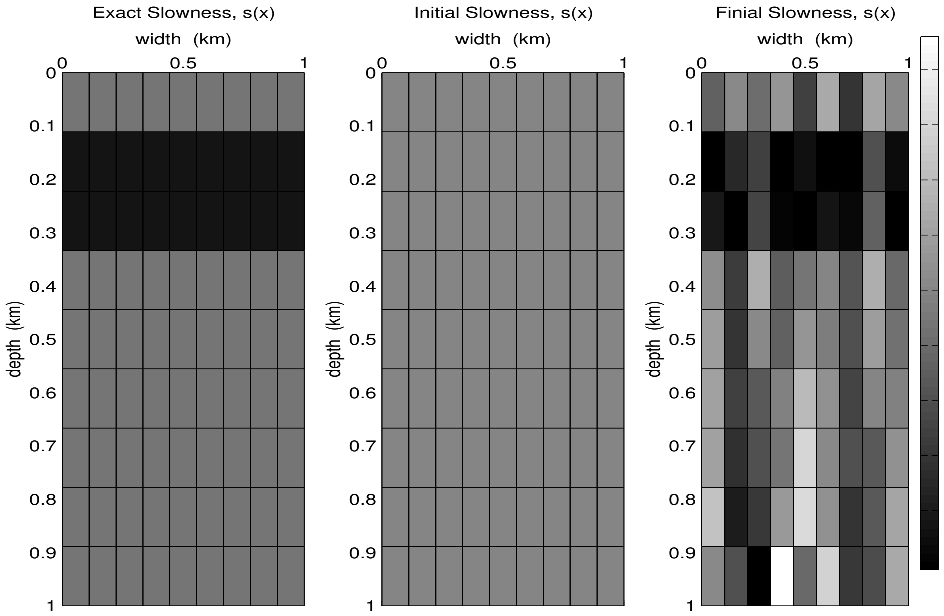

FIG. 8. (Problem C). Illustration of exact, initial, and final slowness distributions for the l_2-Fermat optimization model with 2.9% noise in the traveltime data. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness, anomaly, $s_e(\mathbf{x}) = 0.88\ s/km$, in the upper part of the region. The starting model is constant $s_i(\mathbf{x}) = 5\ s/km$.

6. DISCUSSION

The preliminary experiments of this proposal illustrates the characteristics of the linearized seismic traveltime tomography problem with respect to unconstrained and constrained optimization problems. The constrained problems utilize Fermat feasibility constraints and simple bound constraints on the model space to restrict the algorithm to a well-defined convex feasible region. The experimentation of this proposal was limited to

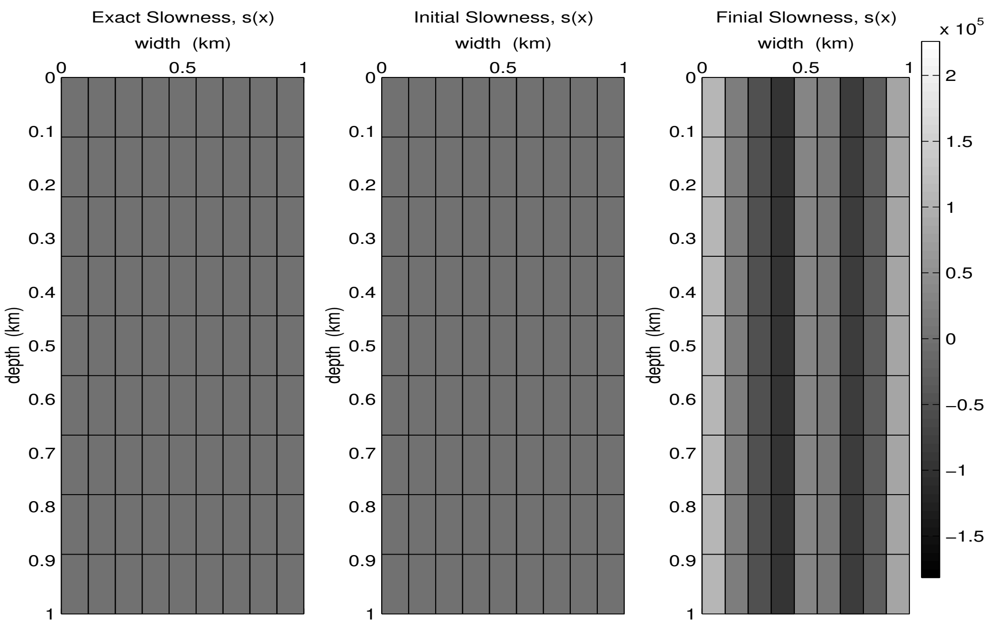

FIG. 9. (Problem C). Illustration of exact, initial, and final slowness distributions for the UC-LS optimization model with 2.9% noise in the traveltime data. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness, anomaly, $s_e(\mathbf{x}) = 0.88\ s/km$, in the upper part of the region. The starting model is constant $s_i(\mathbf{x}) = 5\ s/km$.

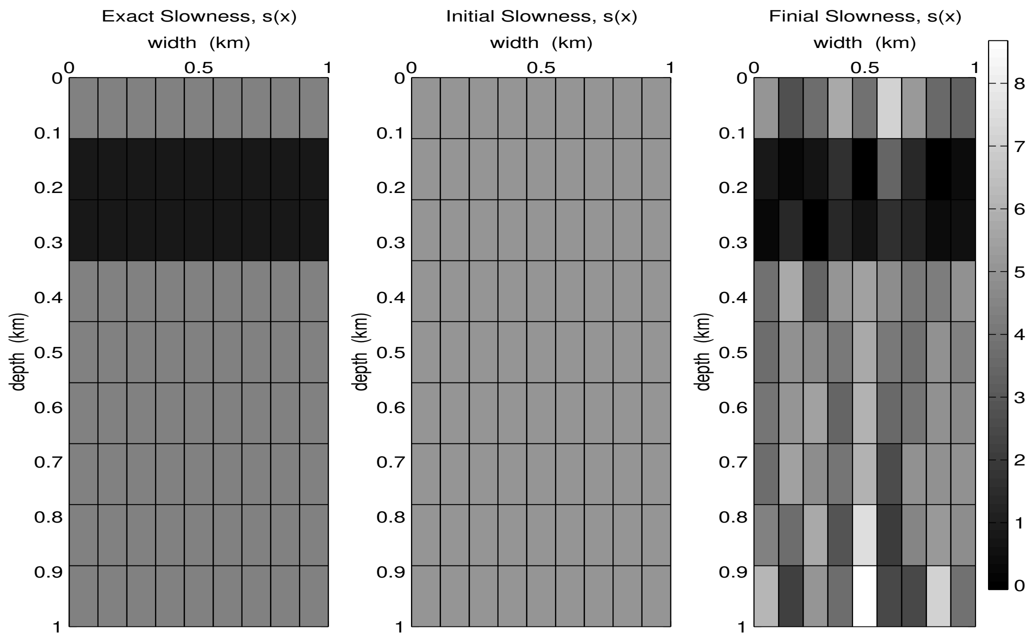

FIG. 10. (Problem C). Illustration of exact, initial, and final slowness distributions for the l_2-Fermat-R optimization model with 2.9% noise in the traveltime data. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness, anomaly, $s_e(\mathbf{x}) = 0.88\ s/km$, in the upper part of the region. The starting model is constant $s_i(\mathbf{x}) = 5\ s/km$.

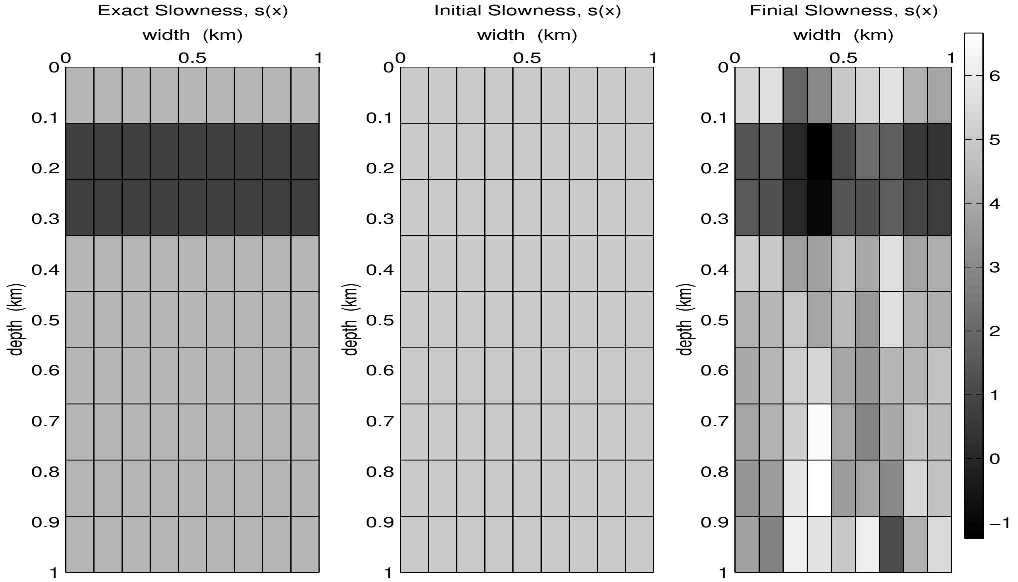

FIG. 11. (Problem C). Illustration of exact, initial, and final slowness distributions for the UC-LS-R optimization model with 2.9% noise in the traveltime data. The exact slowness distribution is $s_e(\mathbf{x}) = 4.4\ s/km$ with a low slowness, anomaly, $s_e(\mathbf{x}) = 0.88\ s/km$, in the upper part of the region. The starting model is constant $s_i(\mathbf{x}) = 5\ s/km$.

the study of the linearized portion of the inversion problem and thus ray bending was not taken into account in the results. The assumption of straight rays for the linearized problem limits our investigation of the full benefit of Fermat feasibility constraints within a constrained problems for the fully nonlinear traveltime tomography problem.

The unconstrained least-squares optimization model (UC-LS) was solved directly via normal equations. As expected, this model performed very poorly in the presence of noisy data. That is, oscillations in the solution, which varied wildly over positive and negative values, caused infeasible models to be generated. The primary infeasibility being that the recovered slowness model contained negative values within the model for problems with perturbations in the synthetic traveltime data.

The l_1-Fermat optimization model was solved using LIPSOL, which is a linear interior point solver (Zhang, 1996). This approach, as well as the other, has the potential of yielding accurate solutions for the linearized traveltime tomography problem when exact synthetic data is used. However, although this model yields better results than the UC-LS model, the approach appears less accurate than the l_2-Fermat method. Further study with the fully nonlinear problem is needed to establish the effectiveness of the methods.

The UC-LS-R and l_2-Fermat-R converged to the same solution within an acceptable tolerance for various starting models. To improve the l_1-Fermat optimization model, one could add upperbounds to the step, or apply the same Tikhonov regularization which was applied to UC-LS and l_2-Fermat optimization models. Employing Tikhonov regularization in the l_1-Fermat optimization problem should improve its performance at the expense of converting the linear programming problem to a simple quadratic programming problem as mentioned earlier. A modification of the l_1-Fermat method to accommodate regularization is currently being considered. This approach may have some added benefit when considerations are made concerning efficiency or possibly a relaxed optimization model for large scale problems.

A nonlinear interior-point solver based on a modification of the algorithm presented in (El-Bakry et al., 1996) was implemented and used to solve the l_2-Fermat optimization models. Although, the l_2-Fermat optimization model performed better than the UC-LS and l_1-Fermat methods in recovering the desired slowness models, each of the methods suffered from rapid deterioration in accuracy as the noise levels increased. To improve the performance of UC-LS and l_2-Fermat, minimum change models were consider. To obtain the minimum change models, Tikhonov regularization was employed as a part of UC-LS and l_2-Fermat. The Tihkonov regularization scheme used here limits the resolution of the model by limiting the change of the model from an initial model. For the linearized tomography problem, Tikhonov regularization is a means of minimizing model variations, thus minimizing the influence of ghost in the solution of the linearized traveltime tomography problem. The performance of the optimization problem on Example Problem C reveal that model fluctuations (e.g., ghost) are controlled, to a certain degree, by the regularization method. Preliminary results show that the regularized optimization problems UC-LS-R and l_2-Fermat-R performed the best in terms of traveltime data misfit and relative error from the true slowness model.

The present experimentation illustrate a definite benefit in using model regularization for both the unconstrained least-squares model and the l_2-Fermat model, however, no clear benefit of the Fermat constraints is established. I attribute this to the fact that the reported experiments only deal with the linearized problem. The linearized problem assumes straight rays, thus the data fit is dependent only on changes in the slowness model space. In the presence of model anomalies, and as the contrast of these anomalies increasingly differ from a constant background model, the straight ray assumption can not be substantiated and ray bending must be included as part of the model correction. This leads us to make further research consideration in terms of the fully nonlinear problem. It is my contention that, the merits of using feasibility constraints over solving an unconstrained model will become apparent in terms of fully nonlinear problem.

Two things which I expect to become apparent with the fully nonlinear model. First, the feasibility boundary, defined by the Fermat constraints, is bounded away from the origin in the model (slowness) space. Thus, very slow slowness models (or very fast wave speeds) are discriminated against. This gives a lowerbound on the slowness models (or upperbound on wave speeds). At the other extreme, if a region is slow (large components of slowness or slow wave speed) compared to the surrounding regions, then once that region becomes "quite slow" compared to the surrounding regions, then the region will effectively drop out of the problem due to the fact that the rays will tend to find faster paths around the slow region. This yields an upperbound on the slowness models (or lowerbound on wave speeds). This upperbound on the slowness models is established by the Fermat constraints and is not a part of the unconstrained model. Secondly, far from the solution, the feasibility constraints will restrict the algorithm to the space of feasible models, thus adding to the stability of the algorithm.

7. PROPOSED WORK

There are several issues involved in the formulation and solution of the traveltime tomography problems. I propose to conduct my thesis work in the area of seismic traveltime tomography, addressing the following issues:

- Formulation of a constrained optimization model for traveltime tomography
- Effective application of primal-dual interior point methods, and
- Finite-difference calculation of traveltimes and traveltime gradients.

Object-oriented software design and development, and automatic differentiation techniques will also be considered.

The problem formulation involves the addition of inequality constraints derived from a priori knowledge of the image region and Fermat feasibility constraints within the traditional unconstrained traveltime tomography model, as discussed in Section(3.1.3). The solution of the resulting constrained optimization problem is addressed using a primal-dual interior point optimization methodology, which requires the computation of both traveltime fields and traveltime gradients. There exist several method in the literature for calculating traveltimes directly on a regular grid. For example, Reshef and Kosloff (1986) used a finite-difference approximation to the eikonal equation, in calculating depth gradient of traveltimes, which they integrated with a Runge-Kutta method. Vidale (1988) also approximated the eikonal equation via finite-differences, but his solution was obtained directly for traveltime using a planar or circular wavefront extrapolation. Although useful, these methods suffer from some disadvantages. Both the method of Reshef and Kosloff and the method of Vidale (as originally published) fail to compute the traveltimes of first arrivals in complex velocity structures. Furthermore, Vidale's method requires an ordering of the points on the computational front which complicates the algorithm in regards to computational time and space complexities. I will employ a high-order Essentially NonOscillatory (ENO) finite-difference scheme for the calculation of traveltimes (cf. (Osher and Shu, 1991)).

The adjoint state approach, commonly used in other optimization and optimal control problems, will be utilized for the computation of the traveltime gradient. I propose to explore convergence and consistency issues related to employing the adjoint state method in this application. As shown by Sei and Symes (1995), the consistency of a numerical scheme with a continuous equation does not imply the consistency of the adjoint scheme. This yields the interesting fact, that the adjoint difference scheme is either of lower order than the forward scheme or not adjoint to (discrete) forward scheme. In choosing the former of these two consequences, there should exist a projection onto the smooth velocity space that will bring the accuracy of the adjoint scheme back up. This conjecture will be established in the thesis. Furthermore, the viability of automatic differentiation tools, such and ADIFOR (Automatic Differentiation In FORtran), will be investigated for the computation of the adjoint state.

Preliminary experimentation, Section(5), suggest that interior point methods are viable for addressing the constrained seismic traveltime tomography problem. In this thesis, I shall further investigate the structure of the constrained model and regularization techniques. Although preliminary results were obtained using MATLAB implementations of algorithms, all numerical algorithms for the thesis will be developed using the Hilbert Class Library (HCL). This library is a collection of C++ classes designed for implementing numerical optimization algorithms in the context of Hilbert spaces (Gockenbach and Symes, 1996). It is my contention that the use of the object-oriented programming paradigm provided by HCL will allow for maximum code reuse with respect to model variations in our work. Eventually, a general HCL based primal-dual interior-point algorithm class will be developed which should prove useful in examining the effects of different constrained models for seismic traveltime tomography and other applications.

REFERENCES

Aki, K., Christoffersson, A., and Husebye, E. S. (1976). Determination of the three-dimensional seismic structures of the lithosphere. *J. Geophys. Res.*, 82:277–296.

Aki, K. and Richards, P. (1980). *Quantitative Seismology: Theory and Methods*, volume II. W. H. Freeman & Co., San Francisco.

Argaez, M. and Tapia, R. A. (1995). On the global convergence of a modified augmented lagrangian linesearch interior point newton method for nonlinear programming. Technical Report TR95-38, Rice University, Department of Computational and Applied Mathematics, Houston, TX.

Bates, C. C., Gaskell, T., and Price, R. (1982). *Geophysics in the Affairs of Man.* Pergamon Press, Oxford.

Berryman, J. G. (1989). Fermat principle and nonlinear traveltime tomography. *Physical Review Letters*, 62(25):2953–2956.

Berryman, J. G. (1990a). Lecture notes on nonlinear inversion and tomography. Technical Report UCRL-LR-105358, Lawrence Livermore National Laboratory.

Berryman, J. G. (1990b). Stable interative reconstruction algorithm for nonlinear traveltime tomography. *Inverse Problems*, 6(1):21–42.

Berryman, J. G. (1991). Convexity properties of inverse problems with variational constraints. *Journal of the Franklin Institute*, 328(1):1–13.

Boggs, P. T. (1996). Sequential quadratic programming. *Acta Numerica*, pages 1–52.

Bracewell, R. N. and Riddle, A. C. (1967). Inversion of fan beam scans in radio astronomy. *Astrophs. J.*, 150:427–434.

Bube, K. P. and Meadows, M. A. (1996). On the null space in linearized anisotropic crosswell traveltime tomography. *66th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 2027–2030.

Coen, S. (1994). White holes in the fresnel zone causing ambiguous first arrival time tomography interpretation. *Physical Review Letters*, 73(24):3219–3222.

Cormack, A. M. (1973). Reconstruction of densities from their projections with applications in radiological physics. *Phys. Med. Biol.*, 18:195–207.

Dennis, J. E. and Schnabel, R. B. (1996). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics (SIAM): Classics in Applied Mathematics, Philadelphia, PA. Originally published by Prentice-Hall, Inc., Englewood Cliffs,NJ, 1983.

Dynes, K. A. and Lyte, R. J. (1979). Computerized geophysical tomography. *Proc. IEEE*, 67:1065–1073.

El-Bakry, A. S., Tapia, R. A., Tsuchiya, T., and Zhang, Y. (1996). On the formulation and theory of the newton interior-point method for nonlinear programming. *Journal of Optimization Theory and Applications*, 89(3):507–541. Paper derives from Rice University Department of Computational and Applied Mathematics Technical Report TR92-40 (1992).

Engl, H. W., Hanke, M., and Neubauer, A. (1996). *Regularization of Inverse Problems*. Dordrecht, Kluwer.

Eriksson, J., Lindström, P., and Wedin, P.-Å. (1995). A new regularization method for rank-deficient nonlinear least squares problems. Technical report, Umeå University, Department of Computing Science, Umeå, Sweden.

Fiacco, A. V. and McCormick, G. P. (1990). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Society for Industrial and Applied Mathematics (SIAM): Classics in Applied Mathematics, Philadelphia, Pennsylvania. Originally published by John Wiley and Sons, Inc., New York-London-Sydney, 1968.

Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical Optimization*. Academic Press Inc., New York.

Gockenbach, M. S. and Symes, W. W. (1996). Hilbert class library: A library of abstract C++ classes for optimization and inversion. *Computers and Mathematics with Applications*, 32(6):1–13.

Harris, J. M., Tan, H., Lines, L., Pearson, C., Treitel, S., Mavko, G., Moos, D., and Nolen-Hoeksma, R. N. (1990). Cross-well tomographic images of geological structures in Gulf Coast sediments. *60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 37–40.

Hurt, N. E. (1989). *Phase Retrieval and Zero Crossings: Mathematical Methods in Image Reconstruction*. Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Inderwiesen, P. I. and Lo, T. (1990). Cross-hole seismic tomographic imaging of reservoir inhomogeneities in the Midway Sunset Field, California. *60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 22–25.

Justice, J. H., Vassiliou, A. A., Mathisen, M. E., Singh, S., Cunningham, P. S., and Hutt, P. R. (1992). Acoustin tomography in reservoir surveillance. *Reservoir Geophysics: Soc. Expl. Geophys., Investigations in Geophysics*, 7:321–334.

Kak, A. C. (1979). Computerized tomography with x-ray, emission, and ultrasound sources. *Proc. IEEE*, 67:1245–1272.

Kaltenbacher, B. (1997). Some newton-type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13:729–753.

Lions, P. L. (1982). *Generalized Solutions of Hamilton-Jacobi Equations*, volume 69 of *Research Notes on Mathematics*. Pitman Publishing Inc, 1020 Plain Street, Marshfield, Massachusetts.

Lo, T., Inderwiesen, P. L., Howlett, D. L., Melton, D. R., Livingston, N. D., Paulsson, B. N. P., and Fairborn, J. W. (1990). McKittrick cross-well seismology project: Part II. Tomographic processing and interpretation. *60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 30–33.

Louis, A. K. and Natterer, F. (1983). Mathematical problems of computerized tomography. *Proc. IEEE*, 71:379–389.

Luenberger, D. G. (1989). *Linear and Nonlinear Programming*. Addison-Wesley Publishing Co., Reading, Massachusetts.

Nolet, G. (1987). *Seismic Tomography: With Applications in Global Seismology and Exploration Geophysics*. Reidel, Dordrecht, Holland.

Ortega, J. M. and Rheinboldt, W. C. (1970). *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, Inc, San Diego, CA.

Osher, S. and Shu, C. (1991). Higher-order essentially NonOscillatory schemes for hamilton-jacobi equations. *SIAM J. Numer. Anal.*, 28:907–922.

Phillips, W. S. and Fehler, M. C. (1991). Traveltime tomography: A comparison of popular methods. *Geophysics*, 56(10):1639–1649.

Radon, J. (1917). Uber die Bestimmung von Funktionen durch ihre integralwerte langs gewisser mannigfaltigkeiten. *Math. Phys. K.*, 69:262.

Rector, J. W. and Washbourne, J. (1994). Characterization of resolution and uniqueness in cross-well direct arrival traveltime tomography using the Fourier projection slice theorem. *Geophysics*, 59:1056–1064.

Reshef, M. and Kosloff, D. (1986). Migration of common shot gathers. *Geophysics*, 51:324–331.

Scudder, H. J. (1978). Introduction to computer aided tomography. *Proc. IEEE*, 66:628–637.

Sei, A. and Symes, W. W. (1995). A note on consistency and adjointness for numerical schemes. In *The Rice Inversion Project: Annual Report*, Rice University, Houston, TX.

Symes, W. W. (1995). A slowness matching finite difference method for traveltimes beyond transmission caustics. Technical report, The Rice Inversion Project (TRIP) Annual Report, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005 USA.

Tapia, R. A. (1980). On the role of slack variables in quasi-newton methods for constrained optimization. In Dixon, L. C. W. and Szegö, G. P., editors, *Numerical Optimization of Dynamic Systems*, pages 235–246. North Holland, Amsterdam, Holland.

Tapia, R. A. (1997). On the fundamental role of interior-point methodology in constrained optimization. Technical Report TR97-09, Rice University, Department of Computational and Applied Mathematics, MS-134, Houston, TX 77005.

Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solution of Ill-Posed Problems.* Winston-Wiley, New York.

Vidale, J. (1988). Finite-difference calculation of traveltimes. *Bulletin of the Seismological Society of America*, 78:2062–2076.

Williamson, P. R. and Worthington, M. H. (1993). Resolution limits in ray tomography due to wave behavior. *Geophysics*, 58:727–736.

Wright, S. J. (1997). *Primal-Dual Interior-Point Methods.* Society for Industrial and Applied Mathematics, Philadelphia.

Wu, R. and Aki, K. (1988). Introduction: Seismic wave scattering in three-dimensionally heterogeneous earth. *PAGEOPH*, 128:1–6.

Zelt, C. A. and Barton, P. J. (1998). 3d seismic refraction tomography: A comparison of two methods applied to data from the faeroe basin. *Journal of Geophysical Research*, 103:7187–7210.

Zhang, Y. (1996). Solving large-scale linear programs by interior-point methods under the MATLAB environment. Technical Report TR96-01, Department of Mathematics and Statistics, University of Maryland Baltimore County. (to appear in Optimization Methods and Software).