A Fast and Global Two Point Low Storage Optimization Technique for Tracing Rays in 2D and 3D Isotropic Media

Debora Cores * **Glenn M. Fung** † **Reinaldo J. Michelena** ‡

October 1999

Abstract

We present the problem of tracing rays in 2D and 3D heterogeneous isotropic media as a set of optimization problems. Each optimization problem is obtained by applying Fermat's principle to an approximation of the travel time equation from a fixed source to a fixed receiver. We assume a piecewise linear raypath which simplifies the computations of the problem, in the same way Mao and Stuart suggested in a very recent paper. Here, instead, the reflectors geometry and the velocity function are computed by using nonuniformly biharmonic splines. This biharmonic spline is easy to apply to interpolation problems in three and more dimensions and it also has minimum curvature. On the other hand, to solve the optimization problem we use the Global Spectral Gradient method. This recent developed optimization scheme is a low storage optimization technique that requires very few floating point operations. It only requires the gradient of the travel time function, and it does not require a close initial raypath. These three properties of the optimization method and the assumption of piecewise linear rays make this ray tracing scheme a very fast and effective method when estimating velocities via tomography. Besides, in a 2D or 3D stratified homogeneous medium, we show that this ray tracing approach converges to the global minimum, if the optimization problem has a solution. We present some numerical results that indicate that this scheme outperforms some classical optimization methods recently used for solving the same problem.

*Scientific Computing and Statistics Department, Universidad Simón Bolívar, Apartado 89000, Caracas 1080-A, Venezuela.(e-mail:cores@cesma.usb.ve, fax # 9063232)

†Scientific Computing and Statistics Department, Universidad Simón Bolívar, Apartado 89000, Caracas 1080-A, Venezuela. (e-mail: gfung@cs.wisc.edu, fax # 58-2-9063232)

‡PDVSA-Intevep (Research Center of the Venezuelan Oil Company), Apartado 76343, Caracas 1070-A, Venezuela. (e-mail: michelen@intevep.pdv.com, fax# 58-2-9087078)

Keywords
Ray tracing, unconstrained optimization, forward modeling, isotropic media.

1 Introduction

The problem of estimating reflector depth and seismic velocity via tomography is a very crucial step when migrating seismic data in complex media, since horizontal or lateral displacements of the structures, and distorsions on the final seismic images highly depend on the velocity of the model. Moreover, the travel time tomography problem is a nonlinear inverse problem, which is usually solved by Gauss-Newton methods. See for example [4], [5], [8] and [2]. Each iteration of the Gauss-Newton method requires the raypaths and the travel time associated to each raypath. This implies that each Gauss-Newton iteration is very time consuming, even with recent computer technologies. So, a computational efficient and low time consuming algorithm for tracing rays in 2D and 3D media is a very important problem in seismic exploration.

Traditionally the problem of tracing seismic rays have been posed as a system of first order differential equations or a two point boundary value problem, as for example in Jackson [11], Jacob [12], Wesson [24], [25], and Pereyra et al. [18]. Moreover, Julian and Gubbins [13] present a comparison between both formulations and they conclude that the problem of tracing seismic rays is computationally faster if the two point boundary value problem is solved. In the last decade, some researchers have been solving the ray tracing problem as an optimization problem by using Fermat's principle which consists in finding the minimum travel time raypath between a source and a receiver. Different optimization techniques have been used for solving the ray tracing problem. Specifically, Um and Thurber [23] present an iterative optimization two point bending method which is based on Fermat's principle and uses a three point perturbation technique. Convergence of this method has not been established. At the same time, Prothero et al. [19] propose a two point optimization technique that consists in finding the minimum travel time ray arc connecting the source and the receiver by using a family of perturbations of arcs and the well-known Nelder and Mead direct search method, [17]. The advantage of using the Nelder and Mead optimization scheme is that no derivatives are required, however the convergence of the method is extremely slow. More recently, Mao and Stuart [15], present a two point ray tracing based on Fermat's principle assuming piecewise linear raypath. To solve the problem they use Newton's method which is a very attractive method since it is local and q-quadratic convergent. However, each iteration requires to solve a linear system of equations and to evaluate the gradient vector and the Hessian matrix. Newton's method requires to be embedded in a globalization strategy, since it is only locally convergent so, convergence to a maximum or stationary point could occur and even worse, the method could diverge. In their paper, the authors propose a strategy to find a close initial raypath to the solution raypath by a shooting scheme utilizing the Snell's law since they use a nonglobalized Newton's method and therefore, the choice of the initial guess is relevant in the convergence of the method.

In this work, we propose to obtain the minimum travel time raypath from a fixed source to a fixed receiver, and its associated travel time, as the solution of a nonlinear optimization problem by using Fermat's principle as Mao and Stuart suggest in [15]. The Spectral Gradient method (SG) developed by Raydan [20], [21] is used to solve each optimization problem. This global optimization method is a low storage technique that requires very few floating point operations, it only requires gradient evaluations, and it does not require a close raypath to the solution raypath since it is a global scheme. These properties of the SG method together with the use of piecewise linear raypath proposed by Mao and Stuart [15] makes the proposed ray tracing method a low time consuming and efficient method especially when the size of the problem to solve is large, which is the case in the reflection travel time tomography problem and in some other geophysical applications.

Besides, the reflectors are expressed as straight lines (2D) or planes (3D) for stratified or dipped earth models, otherwise, in the case of more complex data, the biharmonic spline proposed by Sandwell [22] is used for interpolating the velocity field and depth reflectors. This biharmonic spline is a simple algorithm for finding the minimum "approximated" curvature surface that passes through a set of nonuniformly spaced data set. The algorithm is based on the Green function of the biharmonic operator. For more details, see [22]. The major advantage of this technique is that slope measurements can be used as data. This feature is important for some remote sensing applications where slopes are measured more accurately than heights. Another advantage of this technique is that it is easily applied to interpolation problems in three or more dimensions.

The SG method proposed by Raydan [20], [21] has been used successfully in some large real applications. For example, it has been used in chemistry to determine molecular structures from nuclear magnetic resonance data, Glunt et al. [9]; in statistics and cartography to solve multidimensional scaling problems, Luengo et al. citeRayFran:96; in the solution of partial differential equations, where the dimensions of the linear systems are very large, Molina and Raydan [16]; and very recently in seismic exploration to estimate velocities by reflection tomography, which brings enormous computational advantages, Castillo et al. [3].

This work is presented in the following manner. In Section 2 we pose the ray tracing problem as a nonlinear optimization problem and we also present the model parametrization. A description of the SG method when it is used for solving the ray tracing problem is presented in Section 3. In Section 4 some convergence properties of the SG method and of the proposed optimization problem are presented. In section 5 we present some numerical results and we also illustrate the computational advantages of the method when compared to other optimization techniques. Finally, in Section 6 we present some conclusions.

2 Optimization Problem

In this section we formulate the problem of tracing rays from a fixed source to a fixed receiver as the solution of an optimization problem.

A ray traveling from a source $X_s \in \mathbb{R}^3$ to a receiver $X_r \in \mathbb{R}^3$ and crossing different reflectors of the subsurface satisfies the Fermat's principle, i.e., the ray follows the trajectory of minimum travel time between X_s and X_r. For that reason, for each pair of source-receptor (X_s, X_r) crossing n layers, where each layer is defined by the region delimited between two consecutive interfaces of the subsurface, the problem is a nonlinear unconstrained optimization problem

$$Minimize \quad T_{X_s}^{X_r}(x,y,z) = \int_{X_s}^{X_r} \frac{dl}{v(x,y,z)}, \tag{1}$$

where $v(x,y,z)$ is the velocity of the medium and, dl is the differential length of the ray.

It is important to mention that the proposed scheme consists in solving m nonlinear optimization problems of the form (1) where, m is the total number of rays. In this work we assume that the segment of a ray between two consecutive interfaces, say the $(i-1)$th and ith interfaces, is a straight line segment l_i, as Mao and Stuart proposed in [15] (See Figure 1).

$$l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}, \tag{2}$$

where (x_i, y_i, z_i) are coordinates of the ray at the interfaces, $i = 1, \ldots, 2n+1$. See Fugure 1.

Under the assumption that the segment of a ray between any two consecutive interfaces is a straight line segment we have that the trajectory of a ray for a pair of source-receptor (X_s, X_r) is a piecewise linear path crossing $n+1$ interfaces. So, the optimization problem (1) can be written as follows:

$$\begin{array}{ll} Minimize & T_{X_s}^{X_r}(X,Y,Z) = \sum_{i=2}^{2n+1} \frac{l_i(X,Y,Z)}{v_i(X,Y,Z)}, \\ (X,Y,Z) \in \mathbb{R}^{3\times(2n-1)} & \end{array} \tag{3}$$

where

$$\begin{array}{rcl} X_s &=& (x_1, y_1, z_1)^T, \\ X_r &=& (x_{2n+1}, y_{2n+1}, z_{2n+1})^T, \\ X &=& (x_2, \ldots, x_{2n})^T, \\ Y &=& (y_2, \ldots, y_{2n})^T, \\ Z &=& (z_2, \ldots, z_{2n})^T, \end{array} \tag{4}$$

v_i is the velocity in the midpoint of the raypath segment delimited by the $(i-1)$th and ith interfaces, and it satisfies $v_i = v_{3+2n-i}$ for $i = 2, \ldots, n+1$ in the case of non-converted waves; otherwise, the down going velocities v_i, $i = 2, \ldots, n+1$, are different from the up going velocities v_i, $i = n+2, \ldots, 2n+1$ (see Figure 1).

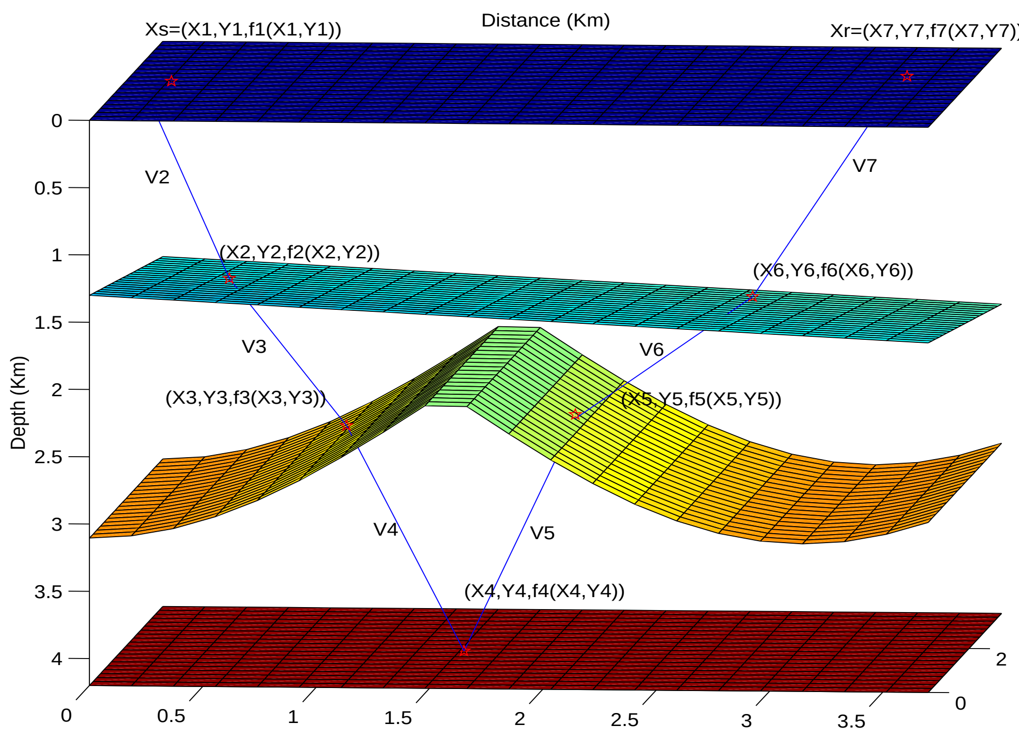

Figure 1: Model Parametrization

The z_i coordinates of the ray trajectory, for $i = 1, \ldots, 2n+1$ in equation (2) and equation (4) depend on the ray trajectory coordinates (x_i, y_i) in the following way

$$z_i = f_i(x_i, y_i), \qquad \text{for } \; i = 1, \ldots, 2n+1, \tag{5}$$

where the function f_i defines the ith interface, $i = 1, \ldots, n+1$ and, for $i = n+2, \ldots, 2n+1$, $\quad f_i = f_{2n-i+2}$. Then, it is clear that the $(n+1)$th interface, f_{n+1}, corresponds to the reflecting interface (see Figure 1). Moreover, the interface functions f_i and the velocities v_i are obtained by biharmonic splines, Sandwell [22] . Therefore, the optimization problem (3) can be expressed as a function of $(X, Y) \in \mathbb{R}^{4n-2}$ as follows

$$Minimize \qquad T_{X_s}^{X_r}(X, Y) = \sum_{i=2}^{2n+1} \frac{l_i(X,Y)}{v_i(X,Y)}. \tag{6}$$

It is well-known that a necessary condition for a ray trajectory (X, Y) to be a minimizer of problem (6) is that the gradient of the travel time functional, $\nabla T_{X_s}^{X_r}(X, Y)$, is equal to zero. Here, the gradient vector is given by

$$\nabla T_{X_s}^{X_r}(X,Y) := \left(\ldots, \frac{\partial T_{X_s}^{X_r}(X,Y)}{\partial x_i}, \ldots, \frac{\partial T_{X_s}^{X_r}(X,Y)}{\partial y_i}, \ldots\right)^T, \quad i = 2, \ldots, 2n \tag{7}$$

where the partial derivatives are:

$$\begin{aligned}
\frac{\partial T_{X_s}^{X_r}(X,Y)}{\partial x_i} &= \frac{\frac{\partial l_i(X,Y)}{\partial x_i} v_i(X,Y) - \frac{\partial v_i(X,Y)}{\partial x_i} l_i(X,Y)}{v_i(X,Y)^2} + \\
&\quad \frac{\frac{\partial l_{i+1}(X,Y)}{\partial x_i} v_{i+1}(X,Y) - \frac{\partial v_{i+1}(X,Y)}{\partial x_i} l_{i+1}(X,Y)}{v_{i+1}(X,Y)^2}, \qquad (8) \\
\frac{\partial T_{X_s}^{X_r}(X,Y)}{\partial y_i} &= \frac{\frac{\partial l_i(X,Y)}{\partial y_i} v_i(X,Y) - \frac{\partial v_i(X,Y)}{\partial y_i} l_i(X,Y)}{v_i(X,Y)^2} + \\
&\quad \frac{\frac{\partial l_{i+1}(X,Y)}{\partial y_i} v_{i+1}(X,Y) - \frac{\partial v_{i+1}(X,Y)}{\partial y_i} l_{i+1}(X,Y)}{v_{i+1}(X,Y)^2}, \qquad (9)
\end{aligned}$$

for $i = 2, \ldots, 2n$.

Now, for the general 3D ray tracing we write the Snell's law as a nonlinear system of equations as follows:

$$\frac{\sin(\alpha_i)}{v_i} = \frac{\sin(\alpha_{i+1})}{v_{i+1}}, \quad i = 2, \ldots, 2n \tag{10}$$

where for each i, α_i is the angle of incidence and α_{i+1} is the refraction angle at the ith interface. It is clear that the number of Snell's nonlinear equations is $2n - 1$, one for each point (x_i, y_i, z_i) in the trajectory of the ray except for the source and the receiver. Using the definition of the sine and cosine of an angle we have that

$$\begin{aligned}
x_i - x_{i-1} &= l_i \sin(\alpha_i) \cos(\phi_i), \\
y_i - y_{i-1} &= l_i \sin(\alpha_i) \sin(\phi_i), \\
z_i - z_{i-1} &= l_i \cos(\alpha_i),
\end{aligned}$$

where α_i is the angle that the segment between $(x_{i-1}, y_{i-1}, z_{i-1})$ and (x_i, y_i, z_i) forms with the vertical axes and the angle ϕ_i is the corresponding azimuthal angle, as Figure 2 illustrates. Moreover, these angles satisfy,

$$\begin{aligned}
\sin(\phi_i) &= \frac{y_i - y_{i-1}}{\hat{l}_i}, \\
\cos(\phi_i) &= \frac{x_i - x_{i-1}}{\hat{l}_i}, \\
\hat{l}_i &= l_i \sin(\alpha_i) = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}. \qquad (11)
\end{aligned}$$

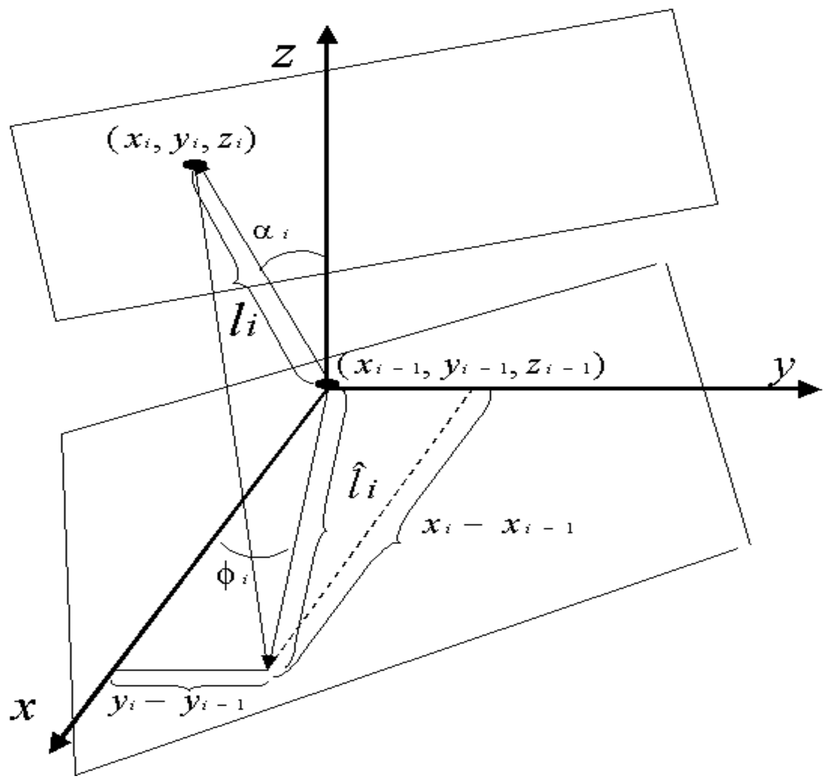

Figure 2: Snell's Law Parametrization

Therefore, from these equations we can write the Snell's equations as a nonlinear system of equations of the form,

$$F(X,Y) = 0, \tag{12}$$

where,

$$F(X,Y) = \frac{\hat{l}_i}{l_i v_i} - \frac{\hat{l}_{i+1}}{l_{i+1} v_{i+1}}, \quad , i = 2, \ldots, 2n. \tag{13}$$

Since in a 2D homogeneous medium, $\frac{\hat{l}_i}{l_i} = \frac{\partial l_i}{\partial x_i}$ and $\frac{\partial v_i}{\partial x_i} = 0$, we can establish from the nonlinear system of equations (12) and (13) the following result.

Theorem 2.1:
In a 2D homogeneus isotropic medium solving the Snell's nonlinear system (12) and (13) is equivalent to forcing the gradient of the travel time function given by (7), (8) and (9) to be equal to zero.

Therefore, the first order necessary conditions for the travel time function given by (6) are identical to Snell's nonlinear system of equations (12). Thus, Newton's method applied to (6) will generate identical iterates to the ones generated by the Newton's method applied to the Snell's nonlinear system of equations (12). Moreover, in a 2D homogeneous medium finding a ray path using Snell's equations to obtain an initial iterate to solve problem (6) by an optimization scheme seems to be redundant. Therefore, an optimization technique for solving problem (6) that does not require an initial raypath close to the true one to converge and, that requires very inexpensive computations, will avoid the redundancy on the use of different approaches for solving the problem and also will guarantee fast convergence. In the 3D homogeneous medium we can only establish that the first order necessary conditions associated to the travel time function (6) imply that the Snell's nonlinear system of equations are satisfied.

Based on the conclusions made on the above paragraph, we propose to solve the unconstrained nonlinear optimization problem (6) with the global SG method [20], [21]. This low storage optimization technique has been used recently in inversion tomography [3] obtaining many computational advantages. This low storage technique presents four advantages: requires low computational storage and few floating point operations when compared with other optimization techniques, second order information (Hessian of the travel time function) is not needed and, it is a global technique, in the sense that it converges from any initial ray trajectory (X_0, Y_0), which is in sharp contrast with the method proposed by Mao and Stuart in [15]. A brief but consice description of the SG method when applied to problem (6) can be found in the next section.

3 Global Spectral Gradient Method

In this section we briefly discuss the Spectral Gradient (SG) Method proposed by Raydan (see [20] and [21]) for solving problem (6).

Assume the travel time function $T_{X_s}^{X_r} : \mathbb{R}^{4n-2} \to \mathbb{R}$ is continuously differentiable in $\mathbb{R}^{4n-2}$ and let

$$(X_k, Y_k) = (x_2^k, x_3^k, \ldots, x_{2n}^k, y_2^k, y_3^k, \ldots, y_{2n}^k)^T \tag{14}$$

be the kth iterate, then the SG method is given by the following iterative process

$$(X_{k+1}, Y_{k+1}) = (X_k, Y_k) - \lambda_k G_k, \tag{15}$$

where G_k is the gradient of $T_{X_s}^{X_r}$ evaluated at (X_k, Y_k), i.e., $G_k := \nabla T_{X_s}^{X_r}(X_k, Y_k)$ and, λ_k is the steplength.

Notice that the search direction in (15) is the negative gradient direction, as in the case of the steepest descent method. The SG is not a descent method since the objective function does not decrease at every iteration. Unfortunately, forcing decrease at every iteration will reduce the SG method to the steepest descent method which is known for being slow. Therefore, the steplength λ_k is chosen by the nonmonotone line search of Grippo, Lampariello and Lucidi [10] which imposes a much weaker condition of decrease on the objective function that makes the SG method much faster than any globalization of steepest descent method. For a given $\lambda = \frac{1}{\alpha_k}$ at iteration kth with,

$$\alpha_k = -\frac{G_{k-1}^T \Lambda_{k-1}}{\lambda_{k-1} G_{k-1}^T G_{k-1}},$$

where $\Lambda_{k-1} = G_k - G_{k-1}$, the λ_k is obtained by the nonmonotone line search which consists on verifying the following weak condition, at each iteration,

$$T_{X_s}^{X_r}(X_{k+1}, Y_{k+1}) \leq \max_{0 \leq j \leq \min(k,M)} T_{X_s}^{X_r}(X_{k-j}, Y_{k-j}) - \gamma \lambda G_k^T ((X_{k+1}, Y_{k+1})^T - (X_k, Y_k)^T), \tag{16}$$

where M is a non-negative integer and γ is a small real number (for details see Luengo at al [14]).

This nonmonotone line search has been recently incorporated to many optimization algorithms. Condition (16) forces the objective function to decrease after M iterations, which guarantees the fast global convergence of the method, i.e, fast convergence of the SG method from any initial guess (X_0, Y_0).

It is important to stress out that the numerical results used to compare the behavior of the SG method with recent extensions of the conjugate gradient method for the non-quadratic case, indicate that the SG method allows, in many cases, a significant reduction of the number of line search and also the number of gradient evaluations [21]. Moreover, the SG method does not require the evaluation of the Hessian matrix and also does not require to solve a linear system of equations at each iteration, as the Newton's method. The low computational work of the SG method together with the properties mentioned above significantly reduce the computational cost and the CPU time of tracing rays in a 3D medium. Next, we present the SG algorithm when applying to trace rays in a 3D medium.

Spectral Gradient Method (SG)

Given $(X_0, Y_0), \alpha_0, integer M \geq 0$, $\gamma \in (0,1)$, $\delta > 0$,

$0 < \sigma_1 < \sigma_2 < 1$, $0 < \epsilon < 1$. Set $k = 0$

Step 1: If $\|G_k\|$ is sufficiently small then stop

Step 2: If $\alpha_k \leq \epsilon$ or $\alpha_k \geq 1/\epsilon$ then set $\alpha_k = \delta$

Step 3: Let $\lambda = 1/\alpha_k$

Step 4: (Nonmonotone Line Search)

If $T_{X_s}^{X_r}((X_k, Y_k) - \lambda G_k) \leq \max_{0 \leq j \leq \min(k,M)} T_{X_s}^{X_r}(X_{k-j}, Y_{k-j}) - \gamma \lambda G_k^T G_k$

then set $\lambda_k = \lambda$, $(X_{k+1}, Y_{k+1}) = (X_k, Y_k) - \lambda_k G_k$, and

$Z_{k+1} = (f_2(x_2^{k+1}, y_2^{k+1}), f_3(x_3^{k+1}, y_3^{k+1}), \ldots, f_{2n}(x_{2n}^{k+1}, y_{2n}^{k+1}))$ and go to Step 6

Step 5: Choose $\sigma \in [\sigma_1, \sigma_2]$, set $\lambda = \sigma\lambda$, and go to step 4

Step 6: Set $\alpha_{k+1} = -(\frac{G_k^T \Lambda_k}{\lambda_k G_k^T G_k})$, $k = k+1$, and go to step 1.

Observe that the choice of the steplength α_k is not the classical choice of the steepest descent method. This choice of the steplength is related to the eigenvalues of the Hessian matrix at the minimizer, which makes the method faster than the steepest descent method (see [9]).

4 Convergence Properties

From the theoretical and practical point of view it is interesting to know when the solution of the optimization problem (6), obtained iteratively by an optimization scheme could reach the global minimum. In the case of stratified and dipped earth models (i.e., $z_i = f_i(x_i, y_i) = a_i x_i + b_i y_i + c_i$) with constant velocity between layers we can establish that the function $T_{X_s}^{X_r}$ given in problem (6) is convex, and therefore any local minimum of problem (6) is in fact a global minimum.

Theorem 4.1:
Let $T_{X_s}^{X_r}$ be the function defined in problem (6). If we have stratified and dipped earth models ($f_i = a_i x_i + b_i y_i + c_i$) with constant velocity between layers then $T_{X_s}^{X_r}$ is convex.

Proof. Let $(X, Y) = (x_2, x_3, \ldots, x_{2n}, y_2, y_3, \ldots, y_{2n})^T \in \mathbb{R}^{4n-2}$. Then the travel time function given by the equation (6) can be written as

$$T_{X_s}^{X_r}(X, Y) = \sum_{i=2}^{2n+1} \frac{\|B_i(X, Y)\|_2}{v_i} = \sum_{i=2}^{2n+1} \frac{\|R_i(X) + S_i(Y) + P_i(X, Y)\|_2}{v_i} \tag{17}$$

where

$$\begin{aligned} R_i(X) &= (x_i - x_{i-1}, 0, 0) \in \mathbb{R}^3, \\ S_i(X) &= (0, y_i - y_{i-1}, 0) \in \mathbb{R}^3, \\ P_i(X, Y) &= (0, 0, f_i(x_i, y_i) - f_{i-1}(x_{i-1}, y_{i-1})) \in \mathbb{R}^3. \end{aligned}$$

Now let (X, Y) and (W, U) be any two vectors in $\mathbb{R}^{4n-2}$. Then for all $\alpha \in [0, 1]$ we have that

$$T(\alpha(X, Y) + (1-\alpha)(W, U)) = \sum_{i=2}^{2n+1} \frac{\|B_i(\alpha X + (1-\alpha)W, \alpha Y + (1-\alpha)U)\|_2}{v_i}. \tag{18}$$

On the other hand, it is straightforward to verify that, for $i = 2, \ldots, 2n+1$

$$R_i(\alpha X + (1-\alpha)W) = \alpha R_i(X) + (1-\alpha)R_i(W) \quad \text{and,} \tag{19}$$
$$S_i(\alpha Y + (1-\alpha)U) = \alpha S_i(Y) + (1-\alpha)S_i(U). \tag{20}$$

Moreover, if $z_i = f_i(x_i, y_i) = a_i x_i + b_i y_i + c_i$, for some $a_i, b_i, c_i \in \mathbb{R}$, $i = 2, \ldots, 2n+1$, where, $f_i = f_{2n-i+2}$ for $i = n+2, \ldots, 2n+1$ and the velocities are constants between layers (i.e., it is a stratified model), it is trivial to show that

$$P_i(\alpha X + (1-\alpha)W, \alpha Y + (1-\alpha)U) = \alpha P_i(X,Y) + (1-\alpha)P_i(W,U) \qquad (21)$$

The result follows directly from (19), (20), (21) and the fact that the norm satisfies the triangular inequality.

The previous result gives us a strong property of the travel time function (6) for stratified and dipped models. In this case, no matter how far we are from the solution any globalized optimization scheme will converge to the global minimum of the problem, if it exists. In particular, the global SG method described in Section 4, will converge fast to the global minimum but with inexpensive computations that is, requiring few line search and so requiring few gradient evaluations. On the other hand, the computational storage is also very low, 3 vectors and, the number of floating point operations per iteration is much smaller than the Newton's method (order 3(2n-1) in 2D and 3(4n-2) in 3D plus the order of evaluating the gradient vector). Newton's method requires to be embedded in a globalization strategy to guarantee convergence. Since it is a monotone optimization scheme, that is the objective function decrease at each iteration, it requires a more exhaustive line search and therefore the computational work and storage per iteration is higher than the required by the global SG method.

5 Numerical Results

The purpose of this work is to present a ray tracing algorithm that has low computational costs and therefore the CPU time is relative small compared with the complexity of solving the problem. Thus, we illustrate the computational work of the scheme by utilizing the algorithm for tracing rays in a 2D and 3D synthetic five layer model of complex stratigraphy for P-S converted waves. Based on the CPU time, total number of iterations, total number of line searches, total floating point operations and computational storage, required for the algorithm in each of the models, we establish its computational advantages. Moreover, we also report if the method converges for all the pairs of sources and receptors. The numerical solutions for the tested problems were obtained in a SUN station ULTRA 1, with 64 Mg Ram, using fortran 77. Also, applications of the method in 2D real models are presented in this section.

Example 1:
First we present a 2D synthetic model that consists of 157 geophones (17 m. apart) and 3 sources (17 m. apart also). The model has 5 layers of complex stratigraphy

where P-S converted waves velocities are considered, as Figure 3 shows. Figure 4 illustrates the rays for the P-S converted waves of the first shot. The CPU time for obtaining all the rays in each shot is approximately of 1 second. Moreover, the algorithm converges for all pairs of sources and receptors and for different initial raypaths, not necessarily close to the solution. Since the model satisfies the conditions of Theorem 4.1, the obtained rays are the global minimum for each pair source-receptor in the model. The total number of iterations for tracing 785 rays, i.e., solving 785 optimization problems, is 17.461 which required the total of 1.671 line searches. Furtheremore, in the 2D case, the number of floating point operations, per iteration, is order $3(2n-1)$ plus the order of evaluating the gradient vector, where n is the number of layers that each ray crosses. Then the average number of total floating point operations, without considering the complexity of the gradient, is 1650 operations. We also observed that the globalization strategy was activated very few times, so the number of gradient evaluations is only 10% more than the number of function evaluations.

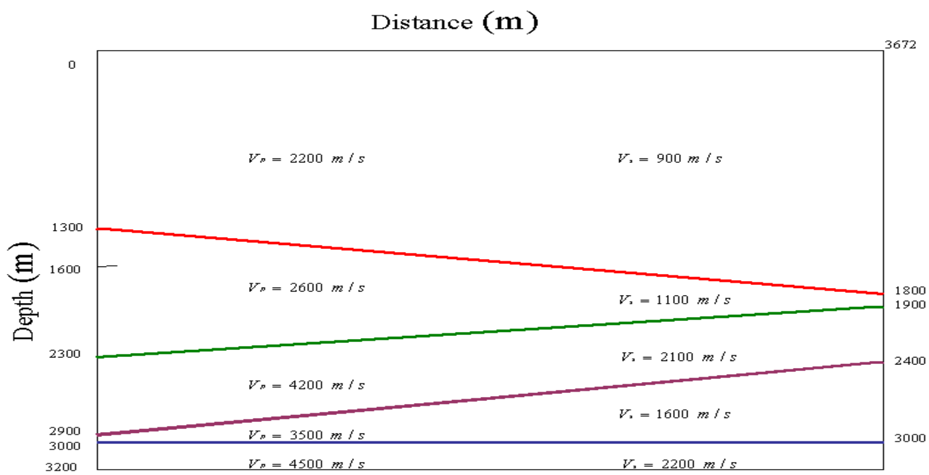

Figure 3: 2D Model

2D Synthetic Application:

This ray tracing algorithm was also used recently by D'Agosto and Michelena in [7] to estimate velocities of compressional and shear waves from P-S converted waves. In this work the authors considered a model of 10 horizontal layers of equal thickness and equal to 500 m. The recording geometry consisted of 100 geophones (15 m. apart) and 3 sources (15 m. apart also). The minimum offset was 30 m. The use of different initial P-S converted waves velocities allowed them to establish some interesting convergenece relations between P-wave velocities, S-wave velocities and layer thicknesses. For more details see [7].

Example 2:
Now we consider a 3D extension of the 2D model previously presented where the sources have the same x and z coordinates and zero y coordinates. The receivers have also the same x and z coordinates but randomly generated y coordinates. The P-S converted waves velocities are given in Figure 5. In this example the CPU time for obtaining all the rays in each shot is approximately 3 seconds. Convergence to the global minimum for each pair of source-receptor is attained by starting from different initial raypaths, since conditions of Theorem 4.1 are satisfied. Figure 6 shows the optimal rays obtained for the first shot. The total number of iterations for tracing 785 rays is 29.136 which required the total of 3.684 line searches. As before the number of gradient evaluations is approximately 12% more than the number of function evaluations. In a 3D model the number of floating point operations, per iteration, is order $3(4n-2)$ plus the order of the gradient evaluation. So, the average number of total floating point operations in this case is 5550 operations.

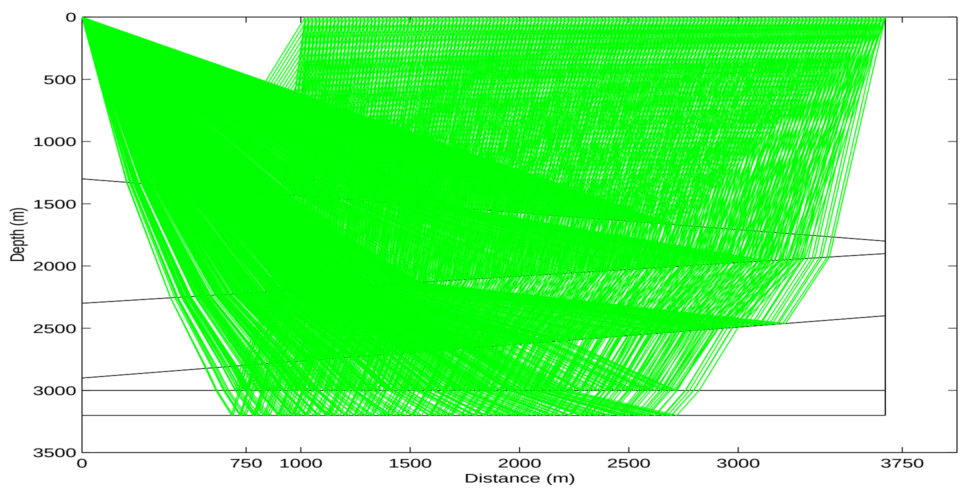

Figure 4: 2D Ray Tracing for P-S Converted Waves

2D Real Application:
This new approach for tracing rays was used again, very recently, by D'Agosto et al [6] for tomography and pre-stack depth migration of P-S converted waves. In this particular work, the authors inverted and migrated P-S field data over the Manporal field, in the south-west of Venezuela, where the geological setting is composed mainly of nearly flaying sediments which dip towards the NE approximately 4 degress. The target zone is the member "0" of the Escandalosa formation, 25 m thick, fracture limestone located at a depth of approximately 3000 m. The portion of the data used to test the algorithms consisted of 124 shot gathers with 216 channels each. The maximum offset for each shot was 3672 m. The distance between sources was 51 m. For more details about the data acquisition, processing and interpretation see [1]. The simplicity, low computational costs and the effectiveness of the ray tracing

algorithm for P-S converted waves allow the authors to conclude that the use of the ray tracing for converted waves in the tomography scheme together with the use of pre-stack migration seems to be the way to image converted waves in structurally complex areas, in areas where P and P-S resolutions are very different that velocity models may vary a lot one from another.

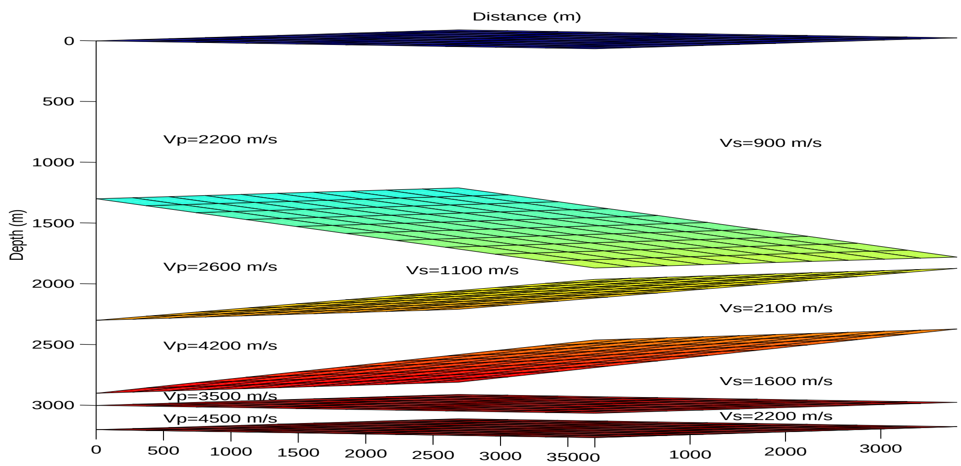

Figure 5: 3D Model

We can extract from the numerical results in the Examples that the average number of of iterations, per each optimization problem, is 22 iterations in Example 1 and 37 iterations in the Example 2. On the other hand, the number of iterations required in Example 2 is 1.67 times the number of iterations used in Example 1. However, the CPU time in Example 2 increases by three the CPU time used in Example 1. This increase in the CPU time come from the fact that the average number of total floating point operations in Example 2 is 3.36 times the number of floating point operations in Example 1. These increases are obvious consequences of the increase in the size of a 3D problem and therefore, in the increase of the complexity order of a 3D problem.

On the other hand, we obtain convergence for all pairs of sources and receptors from any initial guess by activating the globalization strategy very few times, since it is not an exhaustive line search. This property of the nonmonotone line search also reduce the number of gradient evaluations and as a consequence the order and CPU time of the algorithm is reduced considerably, which is not the case if an exhaustive line search is considered. Moreover, the value of M in the SG method for these particular examples is 8. However, we try different values of M and the results obtained present the same behavior. The tolerance in Step 1 of the algorithm is 10^{-5}.

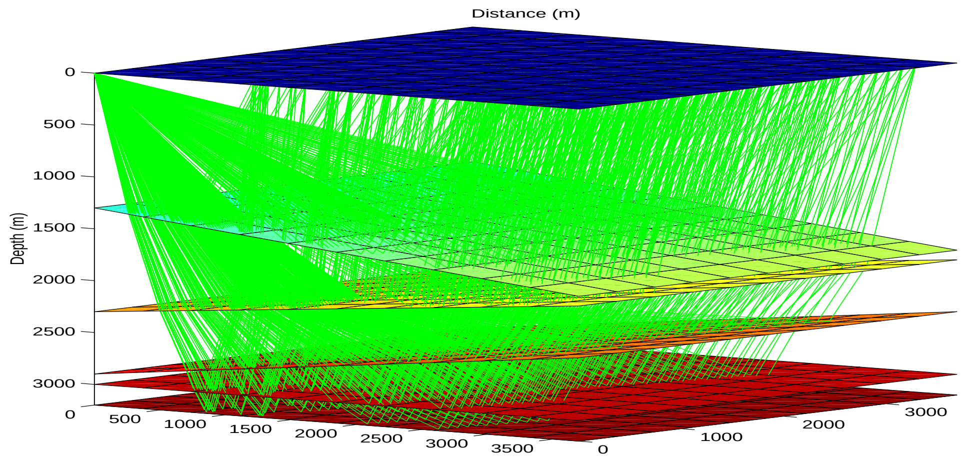

Figure 6: 3D Ray Tracing for P-S Converted Waves

6 Conclusions

We have presented a new and fast low cost optimization technique for tracing rays in 2D and 3D media that converges to a global minimum for homogeneous stratified and dipped earth models, if the problem has a solution. In more complex models the algorithm converges to a minimum but there is not guarantee that the minimum is global. Also, the algorithm converges from any initial guess.

We can see from the numerical results that the CPU time required for computing the raypaths for a shot is relatively small, compared to the complexity of the problem. This low computational work is a consequence of the fact that the SG method requires only gradient evaluations, few floating point operations, and does not require an exhaustive line search to guarantee convergence. Even though Newton's method is locally q-quadratic, it requires an exhaustive line search to guarantee convergence from any initial iterate, so additional gradient evaluations are used, and therefore the number of floating point operations increases. Besides, Newton's method needs to evaluate the Hessian matrix and also to solve a nonlinear system of equations, at each iteration. Thus, the complexity order of Newton's method depends strongly on the solution of the linear system and the evaluation of the Hessian matrix.

The low computational work and storage, per iteration, of the SG method, even when it could take more iterations than the globalized Newton's (GN) method, makes the SG method faster and preferable in most of the cases, since only few gradient evaluations and few floating point operations are required, and in some others very competitive than the GN method.

Finally, the application of the SG method on a synthetic and real examples shows an excellent performance of the algorithm both in accuracy and efficiency.

7 Acknowledgments

We thank Marcos Raydan for making available his Global Spectral Gradient Code.

References

[1] E. ATA ana R.J. MICHELENA. Mapping distribution of fractures in a reservoir with P-S waves. *The Leading Edge*, 14:664–673, 1995.

[2] T. N. BISHOP, K.P. BUBE, R.T. CUTLER, R.T. LAGAN, P.L. LOVE, J.R. RESNICK, R.T. SHUEY, D. A. SPINDLER, and H.W. WYLD. Tomografhic determination of velocity and depth in laterally varing media. *Geophysics*, 50(6):903–923, 1985.

[3] Z. CASTILLO, D. CORES, and M. RAYDAN. Low-cost optimization techniques for the nonlinear seismic reflection tomography problem, 1999. submittted to Optimization and Engineering.

[4] S. K. L. CHIU, E. R. KANASEWICH, and S. PHADKE. Three dimensional determination of structure and velocity by seismic tomography. *Geophysics*, 51(8):1559–1571, 1986.

[5] S. K. L. CHIU and R. R. STEWART. Tomographic determination of three-dimensional seismic velocity structure using well logs, vertical seismic profiles, and surface seismic data. *Geophysics*, 52(8):1085–1098, 1987.

[6] C. D'AGOSTO, M.S. DONATI, and R.J. MICHELENA. Tomography + pre-stack depth migration of P-S converted waves. 68th Annual International Meeting of the Soc. Expl. Geophys., pp. 1361-1364, 1998.

[7] C. D'AGOSTO and R.J. MICHELENA. Tomographic estimation of compressional and shear wave velocities from P-S converted waves. 67th Annual International Meeting of the Soc. Expl. Geophys., pp. 1885-1888, 1997.

[8] V. FARRA and R. MADARIAGA. Non-linear reflection tomography. *Geophysical journal*, 95:135–147, 1988.

[9] W. GLUNT, L. HAYDEN, and M. RAYDAN. Molecular conformations from distance matrices. *J. of Computational Chemestry*, 15:114–120, 1993.

[10] L. GRIPPO, F. LAMPARIELLO, and S. LUDICI. A nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.*, 23:707–716, 1986.

[11] P. L. JACKSON. *Digital simulation os seismic waves*. PhD thesis, University of Michigan, Michigan, 1970.

[12] H. JACOB. Three dimensional seismic ray tracing in a laterally heterogeneous spherical earth. *J. Geophys. Res.*, 75:6685–6689, 1970.

[13] B.R. JULIAN and D. GUBBINS. Three dimensional seismic ray tracing. *J. Geophys.*, 43:95–114, 1977.

[14] F. LUENGO, M. RAYDAN, W. GLUNT, and T. L. HAYDEN. Preconditioned spectral gradient method. *submitted to SIAM Journal on Scientific Computing*, 1998. Technical Report 96-12, Computing Science Department, Universidad Central de Venezuela.

[15] W. MAO and G. W. STUART. Rapid multi-wave-type ray tracing in complex 2d and 3d isotropic media. *Geophysics*, 62(1):298–308, 1997.

[16] B. MOLINA and M. RAYDAN. Precondioned Barzilai-Borwein method for the numerical solution of partial differential equations. *Numerical Algorithms*, 13:45–60, 1996.

[17] A. NELDER and R. MEAD. A simplex method for function minimization. *Computer journal*, 7, 1965.

[18] V. PEREYRA, W. H. LEE, and H.B. KELLER. Solving two-point seismic-ray tracing problems in a heterogeneous medium. *Bulletin of the Seismological Society of America*, 70(1):79–99, 1980.

[19] W. PROTHERO, W. J. TAYLOR, and J. A. EICKEMEYER. A fast two-point three-dimensional ray tracing algorithm using a simple search method. *Bulletin of the Seismological Society of America*, 78:1190–119, 1988.

[20] M. RAYDAN. On the Barzilai and Borwein choice of the steplength for the gradient method. *IMA J. Numer. Anal.*, 13:321–326, 1993.

[21] M. RAYDAN. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. *SIAM J. Optim.*, 7(1):26–33, 1997.

[22] D. T. SANDWELL. Biharmonic spline interpolation os geos-3 and seasat altimeter data. *Geophys. Research Letters*, 14(2):139–142, 1987.

[23] J. UM and C. THURBER. A fast algorithm for two-point seismic ray tracing. *Bulletin of the Seismological Society of America*, 77(3):972–986, 1987.

[24] R. L. WESSON. *Seismic ray computation in laterally inhomogeneous crustal models*. PhD thesis, Stanford University, California, 1970.

[25] R. L. WESSON. Traveltime inversion for laterally inhomogeneous crustal velocity models. *Bulletin of the Seismological Society of America*, 61:729–746, 1971.