Robust Ultrasound Travel-time Tomography Using the Bent Ray Model

Ali Hormati, Ivana Jovanović, Olivier Roy, and Martin Vetterli

Audiovisual Communications Laboratory,
Ecole Polytechnique Fédérale de Lausanne,
CH-1015 Switzerland

ABSTRACT

We present a bent ray reconstruction algorithm for an ultrasound tomography (UT) scanner designed for breast screening. The scanner consists of a circular array of transmitters and receivers which encloses the object to be imaged. By solving a nonlinear system of equations, the reconstruction algorithm estimates the sound speed of the object using the set of travel-time measurements. The main difficulty in this inverse problem is to ensure the convergence and robustness to noise. In this paper, we propose a gradient method to find a solution for which the corresponding travel-times are closest to the measured travel-times in the least squares sense. To this end, first the gradient of the cost function is derived using Fermat's Principle. Then, the iterative nonlinear conjugate gradient algorithm solves the minimization problem. This is combined with the backtracking line search method to efficiently find the step size in each iteration. This approach is guaranteed to converge to a local minimum of the cost function where the convergence point depends on the initial guess. Moreover, the method has the potential to easily incorporate regularity constraints such as sparsity as a priori information on the model. The method is tested both numerically and using in vivo data obtained from a UT scanner. The results confirm the stability and robustness of our approach for breast screening applications.

Keywords: ultrasound tomography, ray-based reconstruction, Snell's law, inverse problems, Gauss-Newton, nonlinear conjugate gradient, regularization, sparsity

1. INTRODUCTION

Acoustic tomography aims at recovering the parameters of an unknown medium by studying the characteristics of sound propagated through the medium. This requires first an accurate forward model which describes well the underlying physical system and second, measurements of good quality. Then, the unknown model is estimated by solving an inverse problem which looks for a model that describes well the measurements. The accuracy of the forward model, the fidelity of the measurements and the choice of the inverse method have direct influence on the quality of the estimated model.

The propagation of acoustic energy in inhomogeneous medium is well modeled by the linear wave equation which is a second order partial differential equation in space and time[1]

$$\nabla^2 p(\boldsymbol{r}, t) - \frac{1}{c^2(\boldsymbol{r})} \frac{\partial^2 p(\boldsymbol{r}, t)}{\partial t^2} = s(\boldsymbol{r}, t).$$

Note that although the pressure $p(\boldsymbol{r}, t)$ is linear with respect to the input signal $s(\boldsymbol{r}, t)$, it is a nonlinear function of the unknown sound speed model $c(\boldsymbol{r})$. Knowing the source signal $s(\boldsymbol{r}, t)$, the inverse problem finds a model $c(\boldsymbol{r})$ which describes well the measured pressure $p(\boldsymbol{r}, t)|_\Omega$ recorded on a known boundary Ω. Although the inverse problem based on the full wave equation is computationally expensive, it makes use of the full information content of the recorded waveforms.

Further author information: (Send correspondence to Ali Hormati)
Ali Hormati: email: ali.hormati@usense.org, phone: +41 21 693 7663
Ivana Jovanović: email: ivana.jovanovic@usense.org, phone: +41 21 693 1271
Olivier Roy: email: olivier.roy@usense.org
Martin Vetterli: email: martin.vetterli@usense.org, phone: +41 21 693 5698

Instead of modeling the forward problem using the wave equation, travel-time tomography employs the principles of geometrical acoustics to estimate the sound speed distribution. It is based on the fact that acoustic energy travels along the lines perpendicular to the equal-phase wavefronts which is a valid assumption at high frequencies. This allows one to find the ultrasound ray paths using Fermat's principle.[2] According to this principle, the ray path between two points is the path of stationary time and can be found using Snell's law. Then, the inverse problem consists of reconstructing the sound speed distribution $c(\boldsymbol{r})$ from the travel-time data estimated from the recorded waveforms. In contrast to the problems with straight-line propagations such as X-ray tomography,[3,4] the ultrasound propagation paths are not straight in an inhomogeneous medium and depend on the sound speed distribution. Therefore, travel-times are a nonlinear function of the unknown sound speed values. However, travel-time tomography is much less computationally expensive than full wave inversion and yields a very good estimate of the sound speed distribution.

Travel-time acoustic tomography has its first roots in seismology where it was used to determine the velocity distribution of the earth between two boreholes.[5,6] Currently, a number of studies have shown that ultrasound tomography has also the potential in detecting and diagnosing breast cancer.[7–9] The results indicate that sound speed and attenuation, the two parameters measured in ultrasound tomography, can help distinguish benign from malignant tissues. The experimental results of this paper are mainly based on the scanner setup of Duric et al.[10] They have developed the Computed Ultrasound Risk Evaluation (CURE) prototype which is an operator independent breast imaging device and aims at differentiation of different masses in the breast tissue. Although the system can use both ultrasound reflection and transmission, we only focus on the transmission part of the system.

In this work, we present techniques to solve the inverse problem in ultrasound travel-time tomography for breast screening. Since the nonlinearity of the problem originates from the dependence of the acoustic paths on the sound speed distribution, a common iterative approach for the inversion is the Gauss-Newton method where one iteratively finds the acoustic paths for the current model estimate and solves the resulting linear system for an updated model. After discussing the drawbacks of this method, we show that the nonlinear conjugate gradient approach works well for the inversion phase. To this end, using the Fermat's principle, we first compute the gradient of the nonlinear sum of squares cost function with respect to the model. Then, we employ the nonlinear conjugate gradient algorithm combined with the backtracking line search method to efficiently solve the least squares minimization problem. Moreover, this method has the potential to easily incorporate regularity constraints such as sparsity[11,12] as a priori information on the model to improve the robustness of the solution to the noise in the measurements. Our experimental results show the effectiveness of our approach to improve the reconstruction quality in terms of estimation accuracy and robustness.

The organization of the paper is as follows. In Section 2, we explain the problem formulation and the setup for the screening procedure. In Section 3, we first present the Gauss-Newton approach and discuss its benefits and drawbacks. Then, we propose the nonlinear conjugate gradient inversion method to minimize the sum of squares cost function followed by its sparsity regularized form. This section is followed by experimental results, both in vivo and numerical which confirms the effectiveness of the proposed algorithm. Finally, we conclude the paper in Section 5.

2. SETUP

The measurement setup consists of a set of ultrasound transmitters and receivers on a 2D ring that encloses the region of interest, as shown in Figure 1. This is the base for the CURE prototype designed at the Karmanos Cancer Institute, Detroit.[10] There are 256 water-coupled ultrasound transducers situated around the ring with an operation frequency of 1.5 MHz. The ring has 20 cm diameter and encompasses the breast to be imaged. Since the sound speeds of water and breast tissue are close, the breast is suspended in water to ensure good acoustic coupling. Each transmitter sends an ultrasound signal that is then received by all the receivers. The ring moves vertically to scan the whole breast slice by slice. One complete scan leads to approximately 45-75 slices and takes about 1 minute to image the entire breast volume. The received signals are sampled at the rate of 6.25 MHz and transferred to a work station for reconstruction.

The first step in the bent ray reconstruction procedure is to estimate the time of flights from the recorded waveforms. Then, the estimated time of flights are used in an inverse problem formulation to reconstruct the
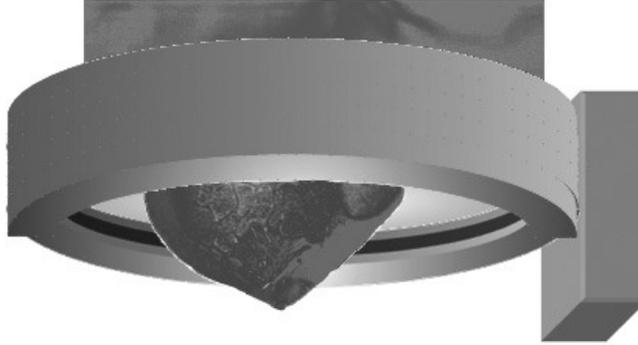
Figure 1. The transducer ring surrounds the tissue to be imaged and both are immersed in water. The initial position of the ring is at the chest wall. The ring is translated vertically to image the whole breast.

unknown sound speed distribution. In this work, we focus on the inverse problem assuming that the time of flights are estimated using an appropriate method [*]. In the following, we first describe the widely used Gauss-Newton method for the inversion phase. Then, after a discussion on the drawbacks of this method, we present the nonlinear conjugate gradient approach as our method of choice for reconstruction.

3. SOLVING THE INVERSE PROBLEM

3.1 Problem Formulation

The time of flight from a transmitter to a receiver is equal to the integral of the slowness (reciprocal of the sound speed) along the propagation path

$$d = \int_\Gamma \frac{1}{u(\boldsymbol{r})} ds, \tag{1}$$

where d is the time of flight, Γ is the acoustic propagation path, $u(\boldsymbol{r})$ is the sound speed at position $\boldsymbol{r}$ and ds is the infinitesimal length along the propagation path. Note that the acoustic path Γ depends on the sound speed distribution $u(\boldsymbol{r})$ and follows the Snell's law of propagation. Therefore, there is a a nonlinear relationship between the time of flights and the sound speed values. In our setup, we discretize equation (1) by dividing the sound speed field into a total of N tiles, each one representing a constant unknown sound speed.

Based on (1), the discretized ray acoustics inverse problem is given by

$$\boldsymbol{d} = \boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m}, \tag{2}$$

where $\boldsymbol{m}$ is the unknown slowness distribution model of size N, $\boldsymbol{G}(\boldsymbol{m})$ is the ray-length matrix of size $M \times N$ and $\boldsymbol{d}$ is a vector of size M which denotes the travel-times estimated from the recorded waveforms. The goal of the inverse problem is to find a slowness distribution $\hat{\boldsymbol{m}}$ which can well describe the travel-times through the nonlinear equation (2).

To quantify the performance of the reconstruction algorithm, we use the least squares criterion, i.e., we try to find a slowness model $\hat{\boldsymbol{m}}$ which minimizes the cost function

$$C(\boldsymbol{m}) \triangleq \|\boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2. \tag{3}$$

This form of the cost function is well justified when there is no a priori information on the model and the errors in the data have the same variance and are independent[†].

On the way to solve this nonlinear problem, we first analyze the Gauss-Newton method, a widely used iterative approach to solve (2).

---

[*]One such algorithm is the cross-correlation method.

[†]We will see how we can incorporate a priori information on the model later in this section.

3.2 Gauss-Newton Approach

Gauss-Newton is an iterative approach to solve nonlinear least squares problems. For the bent-ray tomographic inverse problem in (2), this approach consists of iteratively performing the following steps:

1. Find the ray paths for the current slowness estimate and build the corresponding ray-length matrix.

2. Solve the resulting linear system of equations (2) and update the slowness distribution.

Consider step n of the iterative Gauss-Newton approach with the current slowness distribution estimate $\boldsymbol{m}_n$. First of all, we should linearize the residual $\boldsymbol{r}(\boldsymbol{m}) = \boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}$ around the current slowness estimate $\boldsymbol{m}_n$. This linearization step makes use of the Jacobian matrix

$$\boldsymbol{J}_r[i,j] = \frac{\partial \boldsymbol{r}(i)}{\partial \boldsymbol{m}(j)} \quad i = 1,\dots,M \quad j = 1,\dots,N,$$

where M is the number of travel-time measurements and N is the size of the slowness model. Using the Fermat's principle, it is shown that the Jacobian of the residuals at the current estimate $\boldsymbol{m}_n$ is the corresponding ray-length matrix,[13] i.e.,

$$\boldsymbol{J}_r(\boldsymbol{m}_n) = \boldsymbol{G}(\boldsymbol{m}_n). \tag{4}$$

Therefore, $\boldsymbol{r}(\boldsymbol{m})$ is linearized around $\boldsymbol{m}_n$ as

$$\begin{aligned}
\boldsymbol{r}(\boldsymbol{m}_{n+1}) &= \boldsymbol{G}(\boldsymbol{m}_{n+1}) \cdot \boldsymbol{m}_{n+1} - \boldsymbol{d} \\
&\approx \boldsymbol{r}(\boldsymbol{m}_n) + \boldsymbol{J}_r(\boldsymbol{m}_n) \cdot (\boldsymbol{m}_{n+1} - \boldsymbol{m}_n) \\
&= \boldsymbol{G}(\boldsymbol{m}_n) \cdot \boldsymbol{m}_n - \boldsymbol{d} + \boldsymbol{G}(\boldsymbol{m}_n) \cdot (\boldsymbol{m}_{n+1} - \boldsymbol{m}_n) \\
&= \boldsymbol{G}(\boldsymbol{m}_n) \cdot \boldsymbol{m}_{n+1} - \boldsymbol{d}.
\end{aligned}$$

Thus, in order to find the model estimate at iteration $n+1$ of the Gauss-Newton method, we should solve the following linear problem at iteration n

$$\boldsymbol{m}_{n+1} = \underset{\boldsymbol{m}}{\operatorname{argmin}} \, \|\boldsymbol{G}(\boldsymbol{m}_n) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2. \tag{5}$$

Assuming the system of equations is overdetermined ($M \geq N$), the solution to (5) is given by

$$\boldsymbol{m}_{n+1} = \left[\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{m}_n)\right]^{-1} \boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{d}. \tag{6}$$

Equivalently, by substituting $\boldsymbol{d} = \boldsymbol{G}(\boldsymbol{m}_n)\boldsymbol{m}_n - r(\boldsymbol{m}_n)$ in (6), the Gauss-Newton update equation can be written as

$$\begin{aligned}
\boldsymbol{m}_{n+1} &= \boldsymbol{m}_n + \triangle\boldsymbol{m}_n \tag{7} \\
&= \boldsymbol{m}_n - \left[\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{m}_n)\right]^{-1} \boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{r}(\boldsymbol{m}_n).
\end{aligned}$$

This setups the Gauss-Newton algorithm for iteration $n+1$. In practice, equation (6) is solved using iterative methods like gradient descent or conjugate gradient algorithms.

Although this method is quite straightforward, it suffers from drawbacks which prevents it from being effectively used in practice. First of all, for the bent-ray tomography problem, the local linearity assumption is not valid for large perturbations of the model and therefore, there is no guarantee for the convergence of the algorithm even to a local minimum of the initial sum of squares cost function. The second drawback is related to the solution of the normal equations (6). Because of the inhomogeneous coverage of the rays, the ray-length matrix is very sparse and not well-conditioned in general. Taking into account that the condition number of $\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{m}_n)$ is square of the condition number of $\boldsymbol{G}(\boldsymbol{m}_n)$, any attempt in finding the exact solution of (6)

results in a solution $\boldsymbol{m}_{n+1}$ very sensitive to the errors in the travel-times data $\boldsymbol{d}$. Finally, this method requires one to keep the ray-length matrix $\boldsymbol{G}(\boldsymbol{m}_n)$ in memory which needs special strategies for saving sparse matrices and will directly affect the memory requirement and the speed of the inversion process.

In the following section, we propose the nonlinear conjugate gradient method to solve (2) in the least squares sense and see how it overcomes the problems inherent in the Gauss-Newton approach.

3.3 Nonlinear Conjugate Gradient Approach

In this section we present the method of nonlinear conjugate gradient to minimize the sum of squares cost function in the bent-ray tomography inverse problem. In each iteration of the method, we calculate the gradient of the cost function and also an appropriate step size. We calculate the step size using the back-tracking line search method. The step size will approximately lead us to the minimum of the cost function along the descent direction and eventually provides us with the model update for the next iteration. The conjugate gradient method can follow narrow (ill-conditioned) valleys where the steepest descent method slows down and follows a criss-cross pattern. As we will see, the conjugate gradient approach overcomes the drawbacks of the Gauss-Newton method and can be implemented efficiently in practice. The same approach has been used in[14] to solve the inverse problem in magnetic resonance imaging.

To derive the nonlinear conjugate gradient method, the first step is to find the gradient of the cost function with respect to the sound speed model. To this end, we write the cost function as an inner product

$$C(\boldsymbol{m}) = \|\boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2$$
$$= \langle \boldsymbol{r}(\boldsymbol{m}), \boldsymbol{r}(\boldsymbol{m}) \rangle,$$

where $\boldsymbol{r}(\boldsymbol{m}) = \boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}$ is the vector of residuals. The gradient vector of the cost function with respect to the model is then calculated as

$$\nabla C(\boldsymbol{m}) \overset{(a)}{=} 2 \langle \frac{\partial \boldsymbol{r}(\boldsymbol{m})}{\partial \boldsymbol{m}}, \boldsymbol{r}(\boldsymbol{m}) \rangle$$
$$= 2 \langle \boldsymbol{J}_r(\boldsymbol{m}), \boldsymbol{r}(\boldsymbol{m}) \rangle$$
$$\overset{(b)}{=} 2 \langle \boldsymbol{G}(\boldsymbol{m}), \boldsymbol{r}(\boldsymbol{m}) \rangle$$
$$= 2 \, \boldsymbol{G}(\boldsymbol{m})^{\mathrm{T}} \cdot \boldsymbol{r}(\boldsymbol{m}),$$

where in (a) we used the linearity of the inner product and the assumption of working in a real Hilbert space and in (b), we used (4). Using the gradient information, the decent direction $\triangle \boldsymbol{m}_n$ at iteration n of the conjugate gradient method is calculated as

$$\triangle \boldsymbol{m}_n = -\nabla C(\boldsymbol{m}_n) + \gamma_n \triangle \boldsymbol{m}_{n-1},$$

where we initialize $\triangle \boldsymbol{m}_0 = -\nabla C(\boldsymbol{m}_0)$ and $\boldsymbol{m}_0$ is the initial guess for the sound speed distribution. The value of γ_n is given by the Fletcher-Reeves formula[15]

$$\gamma_n = \frac{\|\nabla C(\boldsymbol{m}_n)\|_2^2}{\|\nabla C(\boldsymbol{m}_{n-1})\|_2^2}.$$

The update of the model at iteration n for the nonlinear conjugate gradient approach is then given by

$$\boldsymbol{m}_{n+1} = \boldsymbol{m}_n + \mu_n \triangle \boldsymbol{m}_n, \tag{8}$$

where μ_n is a small real number denoting the step size in the descent direction at iteration n and is calculated using a line search method. In an exact line search algorithm, the step size is chosen to minimize the cost function along the descent direction $\triangle \boldsymbol{m}_n$

$$\mu_n = \underset{t>0}{\operatorname{argmin}} \, C(\boldsymbol{m}_n + t\triangle \boldsymbol{m}_n).$$

However, finding the exact minimum may need many evaluations of the cost function and can be too complex for practical purposes. Thus, in practice, the step size is chosen to approximately minimize the cost function along

---
Algorithm 1 Back-tracking Line Search Algorithm[16]

1: Let $\triangle\boldsymbol{m}$ be a descent direction at point $\boldsymbol{m}$ and $0 < \alpha < 0.5$, $0 < \beta < 1$.
2: $t := 1$.
3: **while** $C(\boldsymbol{m} + t\triangle\boldsymbol{m}) > C(\boldsymbol{m}) + \alpha t\nabla C(\boldsymbol{m})^{\mathrm{T}}\triangle\boldsymbol{m}$ **do**
4: $\quad t = \beta t$.
5: **end while**
---

the descent direction or just decrease it enough. In our problem, we use the simple and efficient *back-tracking line search* method to find the appropriate step size.[16] This line search algorithm depends on two constants $0 < \alpha < 0.5$ and $0 < \beta < 1$. It starts with the unit step size $t = 1$ and decreases it by a factor β until the stopping condition

$$C(\boldsymbol{m} + t\triangle\boldsymbol{m}) \leq C(\boldsymbol{m}) + \alpha t\nabla C(\boldsymbol{m})^{\mathrm{T}}\triangle\boldsymbol{m},$$

is satisfied.

In general, the complexity of the inverse problem is proportional to the number of forward models it needs to run. In each iteration of the conjugate gradient algorithm, we need to evaluate the gradient $\nabla C(\boldsymbol{m}_n) = 2\,\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}} \cdot \boldsymbol{r}(\boldsymbol{m}_n)$ which requires the computation of the ray-length matrix and the current residual. This step needs one evaluation of the forward model. Next, each iteration of the back-tracking line search algorithm requires one evaluation of the cost function which again needs one forward model evaluation. In total, each iteration of the nonlinear conjugate gradient algorithm requires at least one forward model evaluation. Our experiments show that the minimization algorithm is fast enough to be considered a method of choice in practice.

We can now compare the nonlinear conjugate gradient method with the Gauss-Newton approach. The update equation (7) for the Gauss-Newton method can be rewritten as

$$\boldsymbol{m}_{n+1} = \boldsymbol{m}_n - \frac{1}{2}\left[\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{m}_n)\right]^{-1}\nabla C(\boldsymbol{m}).$$

Comparing the above equation with the update equation of the conjugate gradient algorithm in (8), we notice that the inverse operator $\left[\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{m}_n)\right]^{-1}$ is replaced by the real valued step size μ_n. This indicates that the conjugate gradient method does not suffer from the ill-conditioning of the ray-length matrix in his updates. Moreover, the steepest descent nature of the conjugate gradient algorithm will lead the inversion to converge to a local minimum of the cost function. This indicates that if the initial guess of the model is good enough, we can hope to converge close to the true solution whereas, we do not know the convergence behavior of the Gauss-Newton method. Finally, in contrast to the Gauss-Newton method, we do not need to store the ray-length matrix. This happens because each ray contributes independently to the gradient and the cost function evaluations. Therefore, we can reduce the memory requirements and increase the speed of the inversion process considerably. Similar techniques to solve the ray-based seismic tomography inverse problem along with experimental results can be found in.[17]

3.4 Regularization Using Sparsity

Inverse problems often lead to solutions which are unstable under data perturbations. This is the case for our inverse problem in ray acoustics since the system is nonlinear and the ray-length matrix $\boldsymbol{G}(\boldsymbol{m})$ is sparse and has a high condition number. Regularization methods are special numerical techniques to cope with the instability of the results by using a priori information on the solution.

Under the gradient descent formulation, it is straightforward to add a priori information on the model in order to regularize the solution of the inverse problem. This involves adding the regularization terms to the sum of squares cost function and minimize the new cost function with the nonlinear conjugate gradient algorithm. In the following, we focus on the sparsity prior and how we can incorporate it in the inversion process.

The a priori information that we consider is based on the concept of sparse signal representation. Namely, it assumes that the slowness distribution can be represented as a linear combination of a set of functions (e.g.,

---

Algorithm 2 Nonlinear Conjugate Gradient Algorithm with Sparsity Regularization

---

1: Set ϵ to be a small constant.
2: Take $\boldsymbol{m}_0$ as the initial guess for the model.
3: $n := 0$.
4: **repeat**
5:     Calculate the gradient direction $\nabla C(\boldsymbol{m}_n) = 2\,\boldsymbol{G}(\boldsymbol{m}_n)^{\mathrm{T}} \cdot \boldsymbol{r}(\boldsymbol{m}_n) + \lambda \boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{m}_n$.
6:     Set $\gamma_n = \frac{\|\nabla C(\boldsymbol{m}_n)\|_2^2}{\|\nabla C(\boldsymbol{m}_{n-1})\|_2^2}$ $n \geq 1$, $\gamma_0 = 0$.
7:     Update the conjugate direction $\triangle \boldsymbol{m}_n = -\nabla C(\boldsymbol{m}_n) + \gamma_n \triangle \boldsymbol{m}_{n-1}$.
8:     Perform back-tracking line search in Algorithm 1 and find the step size μ_n.
9:     Update the model estimate $\boldsymbol{m}_{n+1} = \boldsymbol{m}_n + \mu_n \triangle \boldsymbol{m}_n$.
10:    $n = n + 1$.
11: **until** $\frac{\|\mu_n \triangle \boldsymbol{m}_n\|}{\|\boldsymbol{m}_n\|} < \epsilon$

---

vectors in an orthonormal basis or frame) where many of the coefficients are zero.[11, 12, 18] The sparsity basis Ψ can be either an existing one, for example a 2D wavelet basis, or it can be a learned dictionary from a set of similar images. In this work, we assume that the sparsity basis is an orthonormal system or tight frame. The new optimization problem can be represented as

$$\min \ \|\boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2 \ \text{subject to} \ \ \|\Psi^{\mathrm{T}}\boldsymbol{m}\|_0 \leq K,$$

where the $\|\boldsymbol{x}\|_0$ gives the number of non-zeros in vector $\boldsymbol{x}$. This minimization problem can also be represented in Lagrangian form

$$\min \ \|\boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2 + \lambda \|\Psi^{\mathrm{T}}\boldsymbol{m}\|_0. \tag{9}$$

where λ determines the tradeoff between the data fidelity term and the sparsity constraint. Solving (9) is known to be NP-hard in general.[12] There are two main approaches to approximately solve the problem: greedy optimization strategies[11] and relaxation of the cost function. In this work, we follow the relaxation strategy which is more suited to large scale optimization problems.

It is well known that the minimization of the ℓ_1 norm of the transform coefficients promotes sparsity.[19] Therefore, as in Basis Pursuit De-noising,[20] we relax the ℓ_0 term in (9) and replace it with the convex ℓ_1 penalty. This leads to the new optimization problem

$$C(\boldsymbol{m}) = \|\boldsymbol{G}(\boldsymbol{m}) \cdot \boldsymbol{m} - \boldsymbol{d}\|_2^2 + \lambda \|\Psi^{\mathrm{T}}\boldsymbol{m}\|_1.$$

In order to minimize the new cost function using the nonlinear conjugate gradient algorithm, we need the gradient of the new regularization term. However, the absolute value function is not differentiable at the origin. To overcome this issue, we follow the approach in[14] and approximate the absolute value by

$$|x| \approx \sqrt{x^2 + \epsilon},$$

where ϵ is a smoothing parameter. Under this approximation, we get $\frac{d|x|}{dx} \approx \frac{x}{\sqrt{x^2+\epsilon}}$. Therefore, we have for the gradient of the ℓ_1 norm

$$\nabla[\Psi^{\mathrm{T}}\boldsymbol{m}] \approx \Psi\boldsymbol{\Sigma}^{-1}\Psi^{\mathrm{T}}\boldsymbol{m},$$

where the diagonal matrix W has diagonal elements $\boldsymbol{\Sigma}(i,i) = \sqrt{|[\Psi^{\mathrm{T}}\boldsymbol{m}](i)|^2 + \epsilon}$. The nonlinear conjugate gradient algorithm with the sparsity prior is shown in Algorithm 2.

4. EXPERIMENTAL RESULTS

In this section, we present numerical results to demonstrate the effectiveness of the proposed inversion algorithm for the breast screening application based on the CURE setup.[10] The results are divided in two parts. In the first part, we numerically simulate the travel-time measurements in a synthetic phantom by directly solving
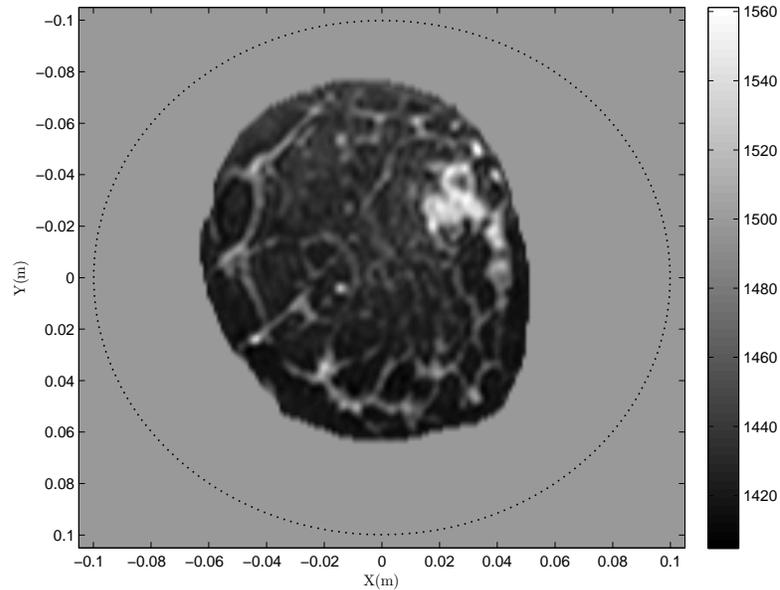
Figure 2. Synthetic sound speed distribution to evaluate the performance of the inversion algorithm. This distribution is synthesized from an MRI slice of a breast with cancer. The dotted ring shows the positions of the 256 water-coupled ultrasound transducers in the CURE imaging setup.

the eikonal equation [‡]. Therefore, we bypass the travel-time estimation phase and focus on solving the inverse problem which is the main topic of this work. This setup let us evaluate the reconstruction quality by looking at the root mean square error (rmse) between the true and reconstructed sound speed distributions. In the second part, we apply the reconstruction algorithm to a set of *in vivo* travel-times which are estimated from the received signals measured by the CURE scanner.

In each iteration of the reconstruction algorithm, we should find the current gradient vector and also the current residual. Since the gradient vector depends on the ray-length matrix, we need to find the travel-time path from each source to each receiver. In order to find the rays, we run the eikonal solver for each source and then trace back the rays from each receiver to the corresponding source by following the opposite direction of the gradient on the travel-time field. Moreover, the current residual is easily computed using the travel-time measurements and the values of the travel-time field for the current sound speed distribution estimate at the position of receivers.

The synthetic phantom along with the ring of ultrasound transducers is shown in Figure 2. This sound speed distribution is synthesized from an MRI slice of a breast with a big cancer tumor indicated by the part with the high sound speed values. In the forward modeling, we get the measurements data by solving the eikonal equation for each source and calculate the corresponding travel-times for the receivers. This process provides us with a total of 32640 travel-time measurements which are the inputs to the reconstruction algorithm.

To solve the inverse problem, we divide the region of interest into 200×200 equal-sized tiles. Each tile represents a constant sound speed value. We use the conjugate gradient algorithm with back-tracking line search method with the initial guess set to a constant water sound speed of 1500 m/s. The rmse for this constant water sound speed is 41 m/s. Figure 3 shows the reconstructions and the corresponding root mean square errors and cost functions for the non-regularized (only sum of squares) and regularized reconstructions. For the regularized reconstruction, we use the Daubechies D2 wavelet as the sparsifying transform. The results show that in our application, using wavelet regularization improves the reconstruction quality both visually and also in the root mean square sense.

[‡]We used the FDTIMES package[21] written by Pascal Podvin from Centre de Géosciences/Géophysique - MINES ParisTech.

(a) Least Squares

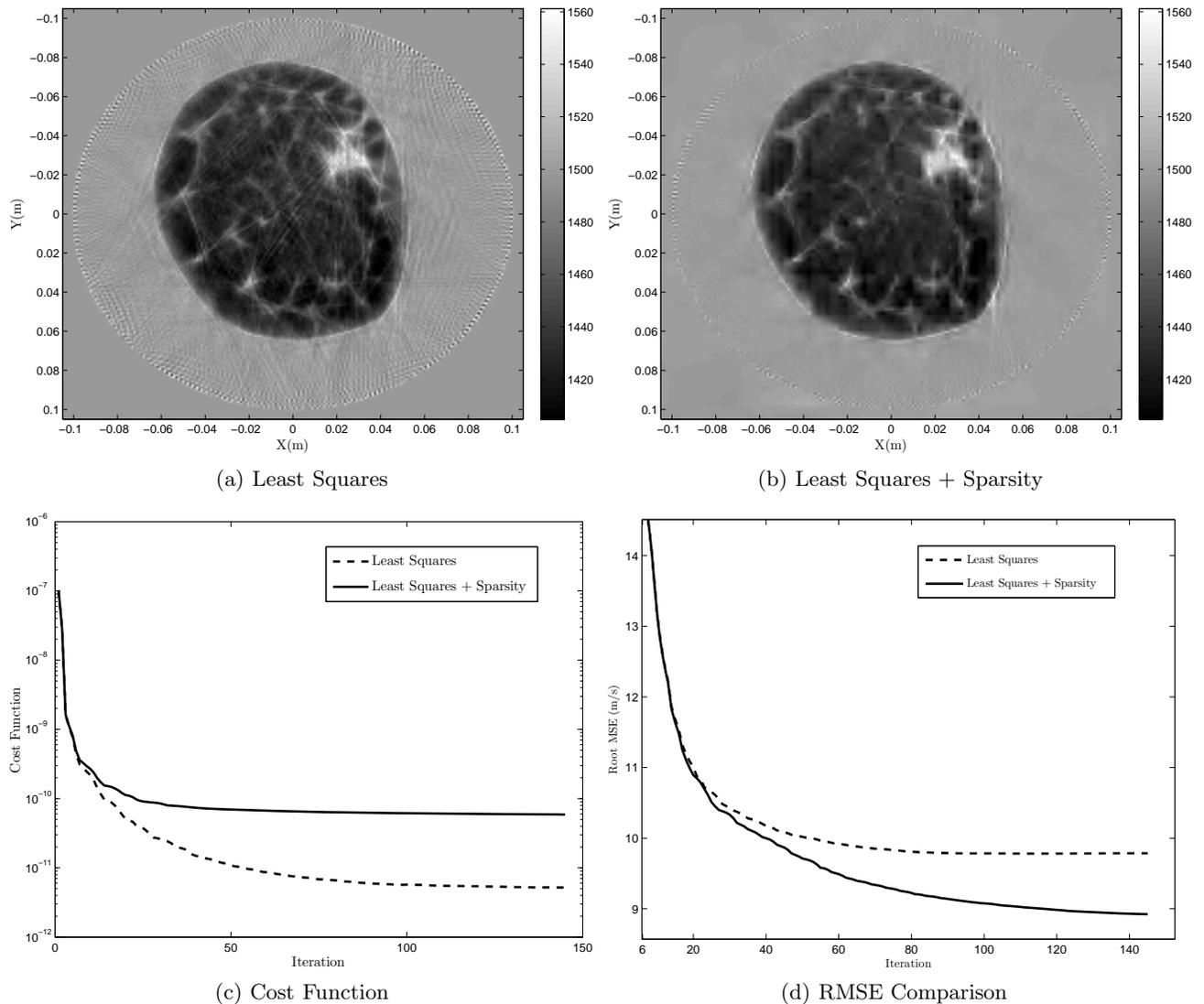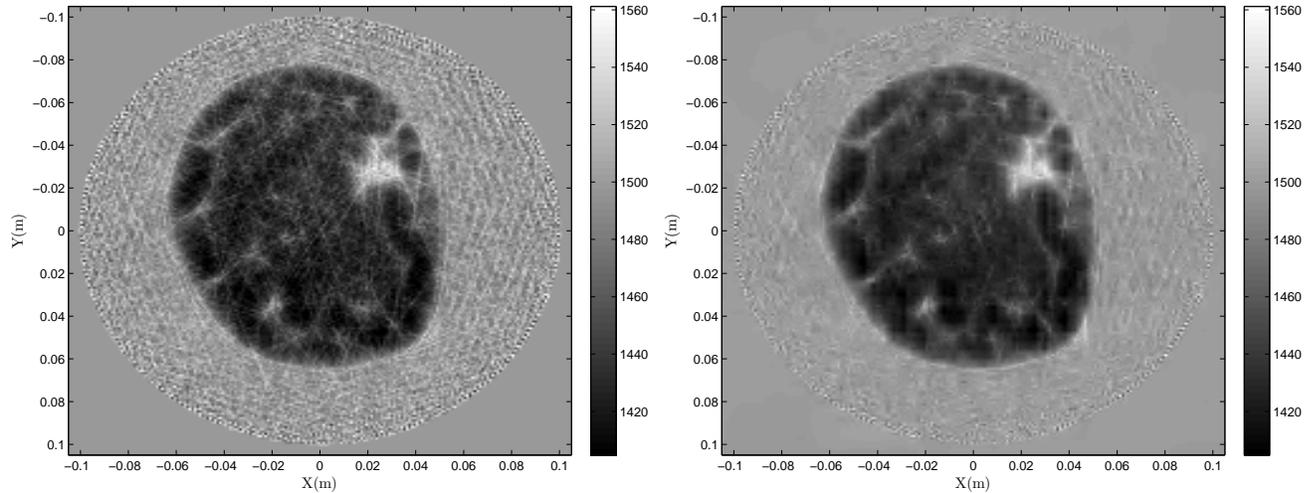(b) Least Squares + Sparsity

(c) Cost Function

(d) RMSE Comparison

Figure 3. Noiseless Setting: The reconstructed sound speed distributions for the regularized and non-regularized formulation of the inverse problem using the conjugate gradient algorithm with back-tracking line search. The sparsifying transform is the Daubechies D2 wavelet. The rmse for the initial guess (constant water sound speed) is 41 m/s.

To evaluate the robustness of our algorithm to noise, we add uniformly distributed noise in the range $[-.16, .16]$ microseconds to the time of flight measurements. Note that for the sampling frequency of 6.25 MHz in the CURE setup, the sampling interval is equal to 1.6 microseconds. Therefore, this amount of noise is equal to at most one sample of uncertainty in the time of flight estimation procedure. The reconstruction results in both non-regularized (only sum of squares) and regularized reconstructions along with corresponding root mean square errors and cost functions are shown in Figure 4. Although the sound speed values in the non-regularized reconstruction goes as high as 1600 m/s, for the sake of visual comparison, we modify the gray scales of all reconstructions to match the scale of Figure 2. Note that in this noisy setup, the wavelet domain sparsity shows more improvement in the root mean square sense compared to the noiseless setting. Moreover, as shown in Figure 4(d), the rmse in the non-regularized scenario increases from iteration 25 although its cost function is still decreasing. This can be explained as overfitting the reconstruction to the noisy input data. This phenomenon does not exist in the regularized reconstruction.

The last experiment is a reconstruction with in vivo data directly from the CURE prototype. As already
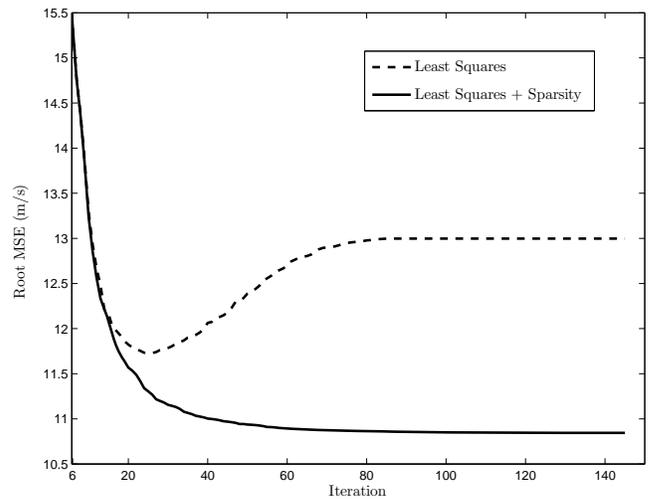
(a) Least Squares

(b) Least Squares + Sparsity
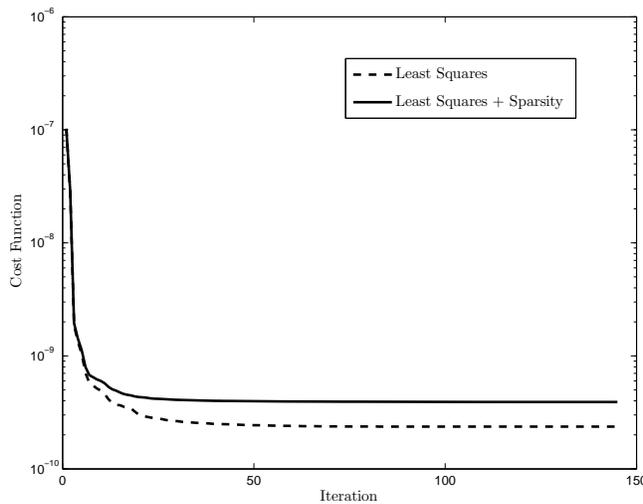
(c) Cost Function

(d) RMSE Comparison

Figure 4. Noisy Setting: The reconstructed sound speed distributions for the regularized and non-regularized formulation of the inverse problem using the conjugate gradient algorithm with back-tracking line search. The sparsifying transform is the Daubechies D2 wavelet. Note that overfitting in the non-regularized setting results in the increase in rmse from iteration 25.

explained in Section 2, ultrasound signals with center frequency of 1.5 MHz is sent by each transducer. The received signals at the receivers are sampled at the rate of 6.25 MHz and sent to a computer for processing. We first estimate the travel times from the received waveforms. This data along with other necessary information are sent to the reconstruction algorithm, which generates the reconstructed slice shown in Figure 5. The recovery algorithm is the nonlinear conjugate gradient method with sparsity regularization. To calculate the gradient, we find the rays using a posteriori ray tracing on the travel time field obtained by solving the eikonal equation.

5. CONCLUSIONS

We investigated the task of reconstructing sound speed distribution of a human breast tissue from the travel time data obtained with a set of transducers placed around the region of interest. Using the bent ray model for sound propagation, we studied the nonlinear formulation which relates the travel time data to the unknown sound speeds. The inverse problem was defined as to minimize the sum of squared error between the modeled
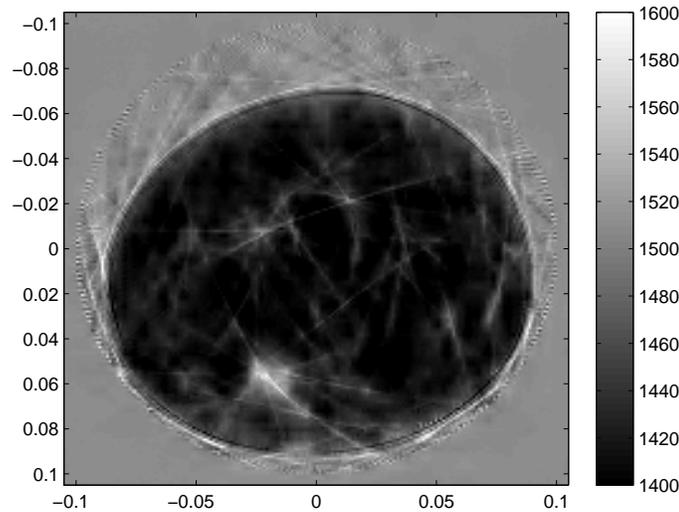
Figure 5. Sound speed reconstruction using the in vivo data from the CURE prototype. The reconstruction is with the nonlinear conjugate gradient method with sparsity constraint in the wavelet domain.

data and the measurements. After discussing the Gauss-Newton method and its properties, we proposed to use the nonlinear conjugate gradient method with back-tracking line search to effectively minimize the cost function. Then, we added regularity constraints in the form of sparsity to the sum of squares term and modified the algorithm to effectively minimize the new cost function. Our experimental results showed that our algorithm converges to a very good estimate of the unknown sound speed distribution and is also robust to the noise in the measurements.

REFERENCES

[1] Goodman, J. W., [*Introduction to Fourier Optics*], San Francisco, CA: McGraw-Hill (1968).

[2] Schuster, A., [*An Introduction to the Theory of Optics*], London: Edward Arnold (1904).

[3] Tanabe, K., "Projection method for solving a singular system of linear equations and its applications," *Num. Math.* **17**, 203–214 (Jun. 1971).

[4] Louis, A. K. and Natterer, F., "Mathematical problems of computerized tomography," **71**, 379–389 (Mar. 1983).

[5] Bois, P., Porte, M. L., Lavergne, M., and Thomas, G., "Essai de détermination automatique des vitesses sismique par mesure entre puits," *Geophys. Prospect.* **19**, 42–83 (1971).

[6] Dines, K. A. and Lytle, R. J., "Computerized geophysical tomography," **67**, 1065– 1073 (Jul. 1979).

[7] Greenleaf, J. F., Johnson, A., Bahn, R. C., and Rajagopalan, B., "Quantitative cross-sectional imaging of ultrasound parameters," in [*1977 Ultrasonics Symposium Proc., IEEE Cat. 77CH1264-1SU*], 989–995 (1977).

[8] Carson, P. L., Meyer, C. R., Scherzinger, A. L., and Oughton, T. V., "Breast imaging in coronal planes with simultaneous pulse echo and transmission ultrasound," *Science 214* **11413** (1981).

[9] Andre, M. P., Janee, H. S., Martin, P. J., Otto, G. P., Spivey, B. A., and Palmer, D. A., "High-speed data acquisition in a diffraction tomography system employing large-scale toroidal arrays," *Int. J. Imaging Syst. Technol.* **8**, 137–147 (1997).

[10] Duric, N., Littrup, P., Poulo, L., Babkin, A., Holsapple, E., Rama, O., and Glide, C., "Detection of breast cancer with ultrasound tomography: First results with the computed ultrasound risk evaluation (CURE) prototype," *Medical Physics* **2**(34), 773–785 (2007).

[11] Mallat, S. G. and Zhang, Z., "Matching pursuits with time-frequency dictionaries," **41**, 3397–3415 (Dec. 1993).

[12] Natarajan, B. K., "Sparse approximate solutions to linear systems," *SIAM J. Comput.* **24**, 227–234 (Apr. 1995).

[13] Tarantola, A., [*Inverse Problem Theory and Methods for Model Parameter Estimation*], Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2004).

[14] Lustig, M., Donoho, D. L., and Pauly, J. M., "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine* **58**(6), 1182–1195 (2007).

[15] Fletcher, R. and Reeves, C., "Function minimization by conjugate gradients," *SIAM J. Comput.* **7**, 149–154 (1964).

[16] Boyd, S. and Vandenberghe, L., [*Convex Optimization*], Cambridge University Press (Mar. 2004).

[17] Taillandier, C., *Formulation de la tomographie des temps de première arrivèe par une méthode de gradient : un pas vers une tomographie interactive*, PhD thesis, Ecole Nationale Superieure des Mines de Paris, France (Dec. 2008).

[18] Jovanović, I., Hormati, A., Sbaiz, L., and Vetterli, M., "Efficient and stable acoustic tomography using sparse reconstruction methods," *19th International Congress on Acoustics* (Sep. 2007).

[19] Donoho, D. L., "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Comm. Pure Appl. Math.* **59**, 797–829 (Jun. 2006).

[20] Chen, S. S., Donoho, D. L., and Saunders, M. A., "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing* **20**, 33–61 (1998).

[21] Podvin, P. and Lecomte, I., "Finite difference computation of traveltimes in very contrasted velocity model: a massively parallel approach and its associated tools," *Geophys. J. Intern.* **105**, 271–284 (1991).