INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **19** (2003) 297–313 PII: S0266-5611(03)54272-1

Solution of inverse problems in elasticity imaging using the adjoint method

Assad A Oberai[1], **Nachiket H Gokhale**[1] **and Gonzalo R Feijóo**[2]

[1] Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, MA 02215, USA
[2] Sandia National Laboratories, MS 9405, Livermore, CA 94551, USA

Received 1 October 2002, in final form 10 December 2002
Published 7 February 2003
Online at stacks.iop.org/IP/19/297

Abstract

We consider the problem of determining the shear modulus of a linear-elastic, incompressible medium given boundary data and one component of the displacement field in the entire domain. The problem is derived from applications in quantitative elasticity imaging. We pose the problem as one of minimizing a functional and consider the use of gradient-based algorithms to solve it. In order to calculate the gradient efficiently we develop an algorithm based on the *adjoint elasticity operator*. The main cost associated with this algorithm is equivalent to solving two forward problems, independent of the number of optimization variables. We present numerical examples that demonstrate the effectiveness of the proposed approach.

1. Introduction

Elasticity imaging is a relatively new and promising technique in medical imaging. It relies on using the difference in elastic modulus of tissues to distinguish them. The basic steps are

(1) subject the specimen to a deformation,
(2) measure the displacement field in the entire domain and
(3) compute the elastic modulus (usually the shear modulus) by solving an inverse problem.

The displacement field may be determined using either ultrasound (see [1] for an example) or nuclear magnetic resonance (NMR) (see [2] for an example). In the case of ultrasound, a speckle image of the specimen in the undeformed state is recorded. Thereafter, the specimen is deformed and another speckle image is recorded. These images are registered to yield the displacement field. Typically, the speckle images, and hence the displacement field, are obtained under static conditions, and effort is made to keep the deformations small. As a result, the displacements are governed by the equations of equilibrium of an incompressible, linear-elastic solid undergoing small, *quasi-static* deformation. On the other hand, when using NMR, displacements are calculated from the phase of the measured magnetic field and are required to be time harmonic. Thus the governing equations are time harmonic. In this study,

 © 2003 IOP Publishing Ltd Printed in the UK

we consider the case of quasi-static deformations, although the algorithms developed may be easily extended to the time-harmonic case.

Once the displacement field is determined, the inverse elasticity problem is solved in one of two ways. The first is the so-called direct method (see [3–7] for applications). This method is based on interpreting the equations of equilibrium as a single, linear, hyperbolic partial differential equation for the shear modulus in which strains appear as known 'material parameters'. This equation is solved with appropriate boundary data to reconstruct the shear modulus. While the direct approach is appealing in that the inverse problem is reduced to a forward problem, it suffers from the following drawbacks:

(1) calculation of the strain field requires differentiating a noisy, measured displacement field and the process of numerical differentiation degrades the accuracy of the method and
(2) the boundary data required to render the equation for the shear modulus well posed are often not known.

The second method for solving the inverse elasticity problem is based on recasting the problem as a non-linear optimization problem (see [8–10] for applications). In this approach, a distribution of shear modulus that minimizes the difference between the measured and predicted displacement fields is sought. The predicted displacement field is required to satisfy the appropriate elasticity equations. In relation to the direct method this approach is more robust, and it addresses the drawbacks described in the previous paragraph. However, it tends to be computationally expensive.

In this study we develop and implement efficient algorithms for solving the non-linear optimization problem resulting from the inverse elasticity problem. In particular, we consider the problem of finding that distribution of shear modulus which minimizes the L_2 norm of the difference between the measured and predicted displacement fields, subject to the constraint that the predicted displacement field satisfy the equations of equilibrium for an incompressible, linear-elastic solid undergoing small, quasi-static deformations. In discretizing this problem, we represent the shear modulus using N piece-wise linear finite-element shape functions whose nodal values are the parameters of the inverse problem. We consider the use of gradient-based algorithms and show that a straightforward calculation of the gradient requires N solves of the forward elasticity problem. This cost is computationally prohibitive for typical values of N ($\geqslant 10^3$). To circumvent this difficulty, we develop and implement a new algorithm based on the adjoint elasticity operator which requires only two solves (independent of N) to compute the gradient. Similar ideas involving the use of adjoint equations in solving optimization problems have been considered in areas such as impedance tomography [11], electromagnetic imaging [12], acoustic shape identification [13] and optimal shape design in aerodynamics [14].

The format of the paper is as follows. In section 2, we describe the forward elasticity problem. In section 3, we define the inverse elasticity problem and consider its solution using gradient based methods. In section 4, we present a straightforward method for computing the gradient, and in section 5 we present a more efficient approach based on the adjoint elasticity operator. In section 6, we present numerical examples, and close with concluding remarks in section 7.

2. Forward elasticity problem

In this section we define the equations that the measured and predicted displacement fields are assumed to satisfy. These describe the equilibrium of an isotropic, linear-elastic solid undergoing small, quasi-static deformations. We present the strong formulation first and

then an equivalent weak or variational formulation. The weak formulation is subsequently approximated using Galerkin's method.

2.1. Strong formulation

The strong form of the forward problem is given as follows: given

(i) the Lamé parameters, $\lambda(\boldsymbol{x})$ and $\mu(\boldsymbol{x})$, in the entire spatial domain Ω,
(ii) the traction vector $\boldsymbol{h}(\boldsymbol{x})$ on a part of the boundary denoted by Γ_h and
(iii) the prescribed displacement vector $\boldsymbol{r}(\boldsymbol{x})$ on a part of the boundary denoted by Γ_r,

find the displacement and pressure fields denoted by $\boldsymbol{u}(\boldsymbol{x})$ and $p(\boldsymbol{x})$, respectively, that satisfy the equations of equilibrium for a linear, incompressible material given by

$$\nabla \cdot (-p\mathbf{1} + \mu(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathrm{T}})) = \mathbf{0}, \qquad \text{in } \Omega \tag{1}$$

$$\nabla \cdot \boldsymbol{u} + p/\lambda = 0, \qquad \text{in } \Omega \tag{2}$$

$$(-p\mathbf{1} + \mu(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathrm{T}}))\boldsymbol{n} = \boldsymbol{h}, \qquad \text{on } \Gamma_h \tag{3}$$

$$\boldsymbol{u} = \boldsymbol{r}, \qquad \text{on } \Gamma_r. \tag{4}$$

In the equations above, $\Omega \subset \mathcal{R}^s$, represents the interior of a body whose boundary is $\partial\Omega = \overline{\Gamma_h \cup \Gamma_r}$, $\boldsymbol{n}$ is the unit outward normal on $\partial\Omega$ and the superscript T denotes the transpose of a tensor. Equation (1) represents the balance of linear momentum within an elastic solid. The term that appears within the divergence operator is the stress tensor. For an isotropic material, Lamé's parameters are related to the more familiar Young modulus (E) and Poisson ratio (ν) via $\mu = E/(2(1+\nu))$ and $\lambda = \nu E/((1+\nu)(1-2\nu))$. Note that μ is also referred to as the shear modulus. Equation (2) is a convenient way of modelling materials that are nearly incompressible. In this equation for a finite value of pressure p, the limit $\lambda \to +\infty$ implies $\nabla \cdot \boldsymbol{u} = 0$, which is the incompressibility condition. In our formulation we approximate this limit replacing $\lambda(\boldsymbol{x})$ by $\lambda(\boldsymbol{x}) = \beta\mu(\boldsymbol{x})$ with β chosen to be a large constant (in our numerical calculations $\beta = 10^6$). Equation (3) specifies the tractions on a part of the domain denoted by Γ_h, and equation (4) specifies the displacement on the other part Γ_r. Note that for a well posed forward problem the union of Γ_h and Γ_r must be the entire boundary $\partial\Omega$.

2.2. Weak formulation

An equivalent weak formulation of (1)–(4) for the incompressible limit is the following: find $\boldsymbol{U} = (\boldsymbol{u}, p) \in \mathcal{S}$, such that

$$A(\boldsymbol{W}, \boldsymbol{U}; \mu) = (\boldsymbol{w}, \boldsymbol{h})_{\Gamma_h}, \qquad \forall \boldsymbol{W} = (\boldsymbol{w}, q) \in \mathcal{V} \tag{5}$$

where $A(\cdot, \cdot; \mu)$ and $(\cdot, \cdot)_{\Gamma_h}$ are bilinear forms defined as

$$A(\boldsymbol{W}, \boldsymbol{U}; \mu) \equiv \int_\Omega \left(\frac{\mu}{2}(\nabla \boldsymbol{w} + \nabla \boldsymbol{w}^{\mathrm{T}}) : (\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathrm{T}}) - \frac{qp}{\beta\mu} - (\nabla \cdot \boldsymbol{w})p - q(\nabla \cdot \boldsymbol{u}) \right) \mathrm{d}\Omega, \tag{6}$$

$$(\boldsymbol{w}, \boldsymbol{h})_{\Gamma_h} \equiv \int_{\Gamma_h} \boldsymbol{w} \cdot \boldsymbol{h} \, \mathrm{d}\Gamma. \tag{7}$$

In (6) the symbol : is used to represent the inner product between two tensors. That is, for two tensors, $\bar{\boldsymbol{T}}$ and $\hat{\boldsymbol{T}}$, $\bar{\boldsymbol{T}} : \hat{\boldsymbol{T}} = \bar{T}_{ij}\hat{T}_{ij}$, where the sum on i and j is implied. The weighting function space $\mathcal{V}$ and the trial solution space $\mathcal{S}$ are defined as

$$\mathcal{V} = \{(\boldsymbol{w}, q) | w_i \in H^1(\Omega), \boldsymbol{w} = \mathbf{0} \text{ on } \Gamma_r, q \in L_2(\Omega)\} \tag{8}$$

$$\mathcal{S} = \{(\boldsymbol{u}, p) | u_i \in H^1(\Omega), \boldsymbol{u} = \boldsymbol{r} \text{ on } \Gamma_r, p \in L_2(\Omega)\}. \tag{9}$$

Note that the bilinear form $A(\cdot, \cdot; \mu)$ depends on the shear modulus μ.

The weak form of the problem may be derived from the strong form by

(1) multiplying (1) and (2) with the weighting functions $\boldsymbol{w}$ and q respectively,
(2) integrating the sum over the entire domain and
(3) performing integration by parts and then making use of the boundary conditions.

It may also be obtained directly by evaluating the conditions for minimizing the total potential energy. In this case, the potential energy is the total elastic strain energy minus the work done by the external body forces and boundary tractions. For the derivation of the weak formulation by minimizing the potential energy the reader is referred to material under the theorem of virtual work in [15]. For the equivalence of strong and weak formulations the reader is referred to [16].

2.3. Galerkin's approximation

In order to solve (5) numerically we introduce finite-dimensional subspaces associated with the weighting and trial solution function spaces. These are denoted by $\boldsymbol{V}^h \subset \mathcal{V}$ and $\boldsymbol{S}^h \subset \mathcal{S}$ respectively. Throughout this manuscript a superscript h denotes the finite-dimensional (numerical) counterpart of a continuous quantity. To obtain the Galerkin approximation to the weak formulation we introduce a function $\boldsymbol{R}^h = (\boldsymbol{r}^h, 0) \in \mathcal{S}^h$, where $\boldsymbol{r}^h$ is a specified function in Ω that approximates the prescribed Dirichlet data $\boldsymbol{r}$ on Γ_r. Thereafter we set $\boldsymbol{U}^h = \boldsymbol{V}^h + \boldsymbol{R}^h$ and derive an equation to determine $\boldsymbol{V}^h \in \mathcal{V}^h$. This leads to the Galerkin approximation: find $\boldsymbol{V}^h = (\boldsymbol{v}^h, p^h) \in \mathcal{V}^h$, such that

$$A(\boldsymbol{W}^h, \boldsymbol{V}^h; \mu^h) = (\boldsymbol{w}^h, \boldsymbol{h})_{\Gamma_h} - A(\boldsymbol{W}^h, \boldsymbol{R}^h; \mu^h), \qquad \forall \boldsymbol{W}^h = (\boldsymbol{w}^h, q^h) \in \mathcal{V}^h; \tag{10}$$

note that the superscript h is used to denote finite-dimensional approximation of quantities that appear in the weak formulation. The split $\boldsymbol{U}^h = \boldsymbol{V}^h + \boldsymbol{R}^h$ is typical of a finite-element solution to the weak form. In this case the function $\boldsymbol{R}^h$ is a known function that is constructed from finite-element shape functions with nodes on the Dirichlet boundary. The contribution from $\boldsymbol{R}^h$ appears in the right-hand side of the algebraic equations resulting from the finite-element discretization.

As part of the proposed algorithm to solve the inverse problem, we solve (10) using the finite-element method. We represent $\boldsymbol{u}^h$, $\boldsymbol{w}^h$ and μ^h as a linear combination of continuous, piece-wise bilinear finite-element shape functions. For example, μ^h is given by

$$\mu^h = \sum_{i=1}^{N} \mu_i \phi_i(\boldsymbol{x}), \tag{11}$$

where N is the number of nodes in the finite-element model and ϕ_i represents a typical bilinear finite-element shape function. We have chosen the same shape functions to represent displacements and shear modulus in order to simplify the programming requirements. In future work other choices will be considered. We represent p^h using discontinuous, piece-wise constant shape functions. This choice allows for the pressure degrees of freedom to be statically condensed from the element matrix, thereby greatly simplifying the overall solution process (see [16] for a discussion).

3. Inverse elasticity problem

In the forward problem described in the previous section we were given the material properties and boundary data and were asked to calculate the displacement field. In the inverse problem the situation is reversed. We are now given the displacement field (or a related measurement)

and the boundary data and are asked to calculate the material properties. While this situation is typical of inverse problems in other fields such as ultrasound tomography and seismic imaging, the inverse elasticity problem is distinct in that the measured data is known on a significant subset of the domain, and not on a lower-dimensional manifold.

The inverse problem is defined as follows: given boundary data $\boldsymbol{r}$ and $\boldsymbol{h}$, and the measurement $\boldsymbol{T}(\boldsymbol{u}^m)$ in Ω, find the distribution of shear modulus $\mu(\boldsymbol{x})$ and a displacement field $\boldsymbol{u}$, such that (1)–(4) hold, and $\boldsymbol{T}(\boldsymbol{u}^m) - \boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{0}$. Here $\boldsymbol{T} : \mathbb{R}^s \to \mathbb{R}^s$ is a second-order tensor with Cartesian components T_{ij}, and $\boldsymbol{u}^m$ represents the measured displacement field. This problem may be posed as a minimization problem as follows: find $\mu(\boldsymbol{x})$ such that the functional

$$\pi(\mu) = \frac{1}{2}\|\boldsymbol{T}(\boldsymbol{u}) - \boldsymbol{T}(\boldsymbol{u}^m)\|_\Omega^2 + \frac{\alpha}{2}\|\mu\|_b^2 \tag{12}$$

is minimized subject to the constraint that $\boldsymbol{u}$ satisfies the forward elasticity problem. In the above equation $\|\cdot\|_\Omega^2 = (\cdot,\cdot)$ represents the L_2 norm in Ω, and $\|\cdot\|_b$ represents an appropriate norm chosen for regularizing the solution. We assume that the norm is derived from an appropriate inner product, that is $\|\cdot\|_b^2 = (\cdot,\cdot)_b$. The parameter α is the Tikhonov parameter and is chosen according to the theory of residues due to Morozov (see [17, 18] for example).

Remark.

(i) The tensor $\boldsymbol{T}$ is introduced to account for the fact that the measurement may be a linear function of the displacement field. For example, in the case when ultrasound is used to measure the displacement field, the resolution in the direction perpendicular to the axis of the transducer is poor, and only one component of the displacement field is measured. Thus

$$[\boldsymbol{T}] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{13}$$

where it is assumed that the transducer is aligned along the x_2-direction.

(ii) The solution to the inverse problem defined above is not unique. To render it unique either the value of the shear modulus must be known on a part of the boundary, or more than one linearly independent, measured displacement field must be specified. The reader is referred to [19, 20] for a detailed discussion. In the examples considered, we have ensured uniqueness by specifying the shear modulus on a part of the boundary.

(iii) In practice when solving the minimization problem, $\pi(\mu)$ is replaced by $\pi(\mu^h)$, which takes into account the finite-dimensional approximation of all the fields appearing in π. Thus the discretized minimization problem reads as follows: find $\mu^h(\boldsymbol{x})$ such that the functional

$$\pi(\mu^h) = \frac{1}{2}\|\boldsymbol{T}(\boldsymbol{u}^h) - \boldsymbol{T}(\boldsymbol{u}^m)\|_\Omega^2 + \frac{\alpha}{2}\|\mu^h\|_b^2 \tag{14}$$

is minimized subject to the constraint that $\boldsymbol{u}^h$ is the numerical solution (i.e. the solution of (10)). Note that the first term in this expression now evaluates the difference between the predicted *numerical solution* and the measured data.

We consider the solution of this problem using a class of optimization algorithms that require the value of the functional and its derivative (also called the gradient) at each iteration. Several quasi-Newton algorithms such as steepest descent, BFGS (Broyden–Fletcher–Goldfarb–Shanno) and DFP (Davidson–Fletcher–Powell) fall under this category (see [21] for a description of these algorithms).

4. Straightforward calculation of the gradient

We consider the straightforward approach to evaluating the gradient in the continuous and discrete cases. In the latter case, we estimate the cost of this computation.

4.1. Continuous case

We define the differential of a functional in the usual way. That is, for a functional $\rho(v)$ which depends on v, the differential in the direction δv is given by

$$\delta\rho = D_v\rho \cdot \delta v = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\rho(v + \epsilon\delta v)\bigg|_{\epsilon=0}, \tag{15}$$

where $D_v\rho$ is the derivative or the gradient of the functional. Using this definition in (12) we have

$$\delta\pi = D_\mu\pi \cdot \delta\mu = (\boldsymbol{T}(\delta\boldsymbol{u}), \boldsymbol{T}(\boldsymbol{u} - \boldsymbol{u}^m)) + \alpha(\delta\mu, \mu)_b, \tag{16}$$

where the differential $\delta\boldsymbol{u}$ represents a change in the displacement $\boldsymbol{u}$ corresponding to a change in shear modulus from μ to $\mu + \delta\mu$. It is evaluated by differentiating (5). This yields the following: find $\delta\boldsymbol{U} \in \mathcal{V}$ such that

$$A(\boldsymbol{W}, \delta\boldsymbol{U}; \mu) = -D_\mu A(\boldsymbol{W}, \boldsymbol{U}; \mu) \cdot \delta\mu, \qquad \forall \boldsymbol{W} \in \mathcal{V}, \tag{17}$$

where $\delta\boldsymbol{U} = (\delta\boldsymbol{u}, \delta p)$, $\boldsymbol{U}$ is the solution to (5), and the right-hand side is given by

$$D_\mu A(\boldsymbol{W}, \boldsymbol{U}; \mu) \cdot \delta\mu = \int_\Omega \delta\mu \left(\frac{1}{2}(\nabla\boldsymbol{w} + \nabla\boldsymbol{w}^{\mathrm{T}}) : (\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^{\mathrm{T}}) + \frac{qp}{\beta\mu^2} \right) \mathrm{d}\Omega. \tag{18}$$

In deriving (17), we have made use of the linearity of the form $A(\boldsymbol{W}, \boldsymbol{U}; \mu)$ on $\boldsymbol{U}$, i.e. $D_{\boldsymbol{U}} A(\cdot, \boldsymbol{U}; \cdot) \cdot \delta\boldsymbol{U} = A(\cdot, \delta\boldsymbol{U}; \cdot)$.

4.2. Discrete case

In the discrete case, μ is approximated by μ^h and $\pi(\mu)$ by $\pi(\mu^h)$. Thus the appropriate differential is given by

$$\delta\pi = D_{\mu^h}\pi \cdot \delta\mu^h = (\boldsymbol{T}(\delta\boldsymbol{u}^h), \boldsymbol{T}(\boldsymbol{u}^h - \boldsymbol{u}^m)) + \alpha(\delta\mu^h, \mu^h)_b, \tag{19}$$

where the differential $\delta\boldsymbol{u}^h$ is evaluated by differentiating (10). This yields

$$A(\boldsymbol{W}^h, \delta\boldsymbol{V}^h; \mu^h) = -D_\mu A(\boldsymbol{W}^h, \boldsymbol{U}^h; \mu^h) \cdot \delta\mu^h, \qquad \forall \boldsymbol{W}^h \in \mathcal{V}^h, \tag{20}$$

where $\delta\boldsymbol{V}^h = (\delta\boldsymbol{u}^h, \delta p^h)$, and $\boldsymbol{U}^h$ is the solution to (10).

4.3. Gradient vector

Let μ^h be represented by (11). Then it is clear that the scalars $\mu_i, i = 1, \ldots, N$ are the parameters of the inverse problem, and that any variation in μ^h is represented by

$$\delta\mu^h = \sum_{i=1}^{N} \delta\mu_i \phi_i(\boldsymbol{x}). \tag{21}$$

Using (11) in (19) and (20) we have

$$\delta\pi = \boldsymbol{G} \cdot \delta\boldsymbol{\mu}, \tag{22}$$

where $\boldsymbol{G} \in \mathbb{R}^N$ is the gradient vector whose components are given by

$$G_i = D_{\mu^h}\pi \cdot \phi_i = (\boldsymbol{T}(\delta\boldsymbol{u}^h), \boldsymbol{T}(\boldsymbol{u}^h - \boldsymbol{u}^m)) + \alpha(\phi_i, \mu^h)_b. \tag{23}$$

In the above equation, for each i, $\delta \boldsymbol{u}^h$ is obtained by solving

$$A(\boldsymbol{W}^h, \delta \boldsymbol{V}^h; \mu^h) = -D_\mu A(\boldsymbol{W}^h, \boldsymbol{V}^h; \mu^h) \cdot \phi_i, \qquad \boldsymbol{W}^h \in \mathcal{V}^h, \tag{24}$$

where $\delta \boldsymbol{V}^h = (\delta \boldsymbol{u}^h, \delta p^h)$, and $\boldsymbol{V}^h$ is the solution to (10). Using (23), the following algorithm may be used to calculate the gradient vector $\boldsymbol{G}$.

Algorithm 4.1.

(i) Solve (10) to evaluate $\boldsymbol{V}^h$.
(ii) For $i = 1, \ldots, N$
  (a) solve (24) to evaluate $\delta \boldsymbol{V}^h$;
  (b) calculate G_i using (23).

The above algorithm requires $N + 1$ solves (one in step (i) and N in step (ii)) of the size of a forward elasticity problem. For typical values of $N \geqslant 10^3$, this represents a large computational cost.

5. Calculation of the gradient using the adjoint equations

In this section we propose an efficient algorithm for calculating the gradient vector. This approach is based on using the adjoint of the bilinear form $A(\cdot, \cdot; \mu)$.

5.1. The continuous case

We begin by introducing the Lagrangian L,

$$L(\boldsymbol{U}, \boldsymbol{W}, \mu) = \frac{1}{2}\|\boldsymbol{T}(\boldsymbol{u}) - \boldsymbol{T}(\boldsymbol{u}^m)\|_\Omega^2 + \frac{\alpha}{2}\|\mu\|_b^2 + A(\boldsymbol{W}, \boldsymbol{U}; \mu) - (\boldsymbol{w}, \boldsymbol{h})_{\Gamma_h}, \tag{25}$$

where $\boldsymbol{W} = (\boldsymbol{w}, q)$ plays the role of a Lagrange multiplier. The differential of L is given by

$$\delta L = D_U L \cdot \delta \boldsymbol{U} + D_{\boldsymbol{W}} L \cdot \delta \boldsymbol{W} + D_\mu L \cdot \delta \mu. \tag{26}$$

In the above equation, setting $D_{\boldsymbol{W}} L \cdot \delta \boldsymbol{W} = 0, \forall \delta \boldsymbol{W} \in \mathcal{V}$, yields the following equation for $\boldsymbol{U}$:

$$A(\delta \boldsymbol{W}, \boldsymbol{U}; \mu) = (\delta \boldsymbol{w}, \boldsymbol{h})_{\Gamma_h}, \qquad \forall \delta \boldsymbol{W} \in \mathcal{V}, \tag{27}$$

which implies that $\boldsymbol{U}$ satisfies the original elasticity problem (5). For this value of $\boldsymbol{U}$, from (12) and (25) we conclude that $L(\boldsymbol{U}, \boldsymbol{W}, \cdot) = \pi(\cdot), \forall \boldsymbol{W}$, and consequently for a given $\delta \mu$

$$\delta L = \delta \pi. \tag{28}$$

Thus, we may use the expression for δL to evaluate $\delta \pi$. The expression for δL can be simplified considerably if $D_U L \cdot \delta \boldsymbol{U} = 0, \forall \delta \boldsymbol{U} \in \mathcal{V}$. This condition yields the following equation for $\boldsymbol{W}$:

$$A(\boldsymbol{W}, \delta \boldsymbol{U}; \mu) = -(\boldsymbol{T}(\delta \boldsymbol{u}), \boldsymbol{T}(\boldsymbol{u} - \boldsymbol{u}^m)), \qquad \forall \delta \boldsymbol{U} \in \mathcal{V}. \tag{29}$$

This may be written as

$$A^*(\delta \boldsymbol{U}, \boldsymbol{W}; \mu) = -(\boldsymbol{T}(\delta \boldsymbol{u}), \boldsymbol{T}(\boldsymbol{u} - \boldsymbol{u}^m)), \qquad \forall \delta \boldsymbol{U} \in \mathcal{V}, \tag{30}$$

where $A^*(\cdot, \cdot; \mu)$ is the *adjoint* of the bilinear form $A(\cdot, \cdot; \mu)$. Using (6) it is easily verified that $A(\cdot, \cdot; \mu)$ is self-adjoint, that is $A^*(\boldsymbol{V}_1, \boldsymbol{V}_2, \mu) = A(\boldsymbol{V}_1, \boldsymbol{V}_2, \mu) \quad \forall \boldsymbol{V}_1, \boldsymbol{V}_2 \in \mathcal{V}$, and hence the above equation reduces to

$$A(\delta \boldsymbol{U}, \boldsymbol{W}; \mu) = -(\boldsymbol{T}(\delta \boldsymbol{u}), \boldsymbol{T}(\boldsymbol{u} - \boldsymbol{u}^m)), \qquad \forall \delta \boldsymbol{U} \in \mathcal{V}. \tag{31}$$

With $\boldsymbol{U}$ and $\boldsymbol{W}$ given by (27) and (31) respectively, from (26) and (28) we have

$$\delta \pi = \delta L = D_\mu A(\boldsymbol{W}, \boldsymbol{U}; \mu) \cdot \delta \mu + \alpha(\mu, \delta \mu)_b. \tag{32}$$

Remark.

(i) The equation for determining $\boldsymbol{U}$, that is (27), is the same as (5), the forward elasticity problem. This equation is often referred to as the primal problem.

(ii) The equation for the Lagrange multiplier $\boldsymbol{W}$, that is (31), is of the same type as the primal problem and is driven by the difference in the predicted and measured data. This equation is referred to as the dual problem. The strong form associated with the problem may be obtained by performing integration by parts on (31) and is given by

$$\nabla \cdot (-q\mathbf{1} + \mu(\nabla \boldsymbol{w} + \nabla \boldsymbol{w}^{\mathrm{T}})) = -\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T}(\boldsymbol{u} - \boldsymbol{u}^m), \qquad \text{in } \Omega \tag{33}$$

$$\nabla \cdot \boldsymbol{w} + q/\lambda = 0, \qquad \text{in } \Omega \tag{34}$$

$$(-q\mathbf{1} + \mu(\nabla \boldsymbol{w} + \nabla \boldsymbol{w}^{\mathrm{T}}))\boldsymbol{n} = \mathbf{0}, \qquad \text{on } \Gamma_h \tag{35}$$

$$\boldsymbol{w} = \mathbf{0}, \qquad \text{on } \Gamma_r. \tag{36}$$

Thus the dual displacement $\boldsymbol{w}$ and the dual pressure field q are determined by solving the original elasticity equations forced by the difference between the predicted and measured displacements (see (33), (34)). The boundary conditions for these fields (35) and (36) are the homogeneous counterparts of the boundary conditions for the primal field.

5.2. Discrete case

Repeating the derivation in the previous section with $\pi(\mu)$ replaced by $\pi(\mu^h)$, which is the discrete approximation to $\pi(\mu)$, we have

$$\delta\pi = D_\mu A(\boldsymbol{W}^h, \boldsymbol{U}^h; \mu^h) \cdot \delta\mu^h + \alpha(\mu^h, \delta\mu^h)_b, \tag{37}$$

where $\boldsymbol{U}^h = \boldsymbol{V}^h + \boldsymbol{R}^h$, $\boldsymbol{V}^h$ satisfies

$$A(\delta\boldsymbol{W}^h, \boldsymbol{V}^h; \mu^h) = (\delta\boldsymbol{w}^h, \boldsymbol{h})_{\Gamma_h} - A(\delta\boldsymbol{W}^h, \boldsymbol{R}^h; \mu^h), \qquad \forall \delta\boldsymbol{W}^h \in \mathcal{V}^h, \tag{38}$$

and $\boldsymbol{W}^h$ satisfies

$$A(\delta\boldsymbol{U}^h, \boldsymbol{W}^h; \mu^h) = -(\boldsymbol{T}(\delta\boldsymbol{u}^h), \boldsymbol{T}(\boldsymbol{u}^h - \boldsymbol{u}^m)), \qquad \forall \delta\boldsymbol{U}^h \in \mathcal{V}^h. \tag{39}$$

5.3. Gradient vector

Once again the gradient may be explicitly expressed as a finite-dimensional vector. This is accomplished by utilizing (21) in (37)–(39) to arrive at

$$\delta\pi = \boldsymbol{G} \cdot \delta\boldsymbol{\mu}. \tag{40}$$

The components of the gradient vector are given by

$$G_i = A(\boldsymbol{W}^h, \boldsymbol{U}^h; \phi_i) + \alpha(\mu^h, \phi_i)_b, \tag{41}$$

where $\boldsymbol{U}^h$ is given by (38) and $\boldsymbol{W}^h$ is given by (39). Note that (38) is an equation for calculating $\boldsymbol{V}^h$, and that $\boldsymbol{U}^h$ may be obtained from $\boldsymbol{V}^h$ by the relation $\boldsymbol{U}^h = \boldsymbol{V}^h + \boldsymbol{R}^h$. Thus, using (41), the following algorithm may be used to compute $\boldsymbol{G}$.

Algorithm 5.1.

(i) Solve (38) to evaluate $\boldsymbol{U}^h$.

(ii) Solve (39) to evaluate $\boldsymbol{W}^h$.

(iii) Use $\boldsymbol{U}^h$ and $\boldsymbol{W}^h$ in (41) to compute $\boldsymbol{G}$.

Remarks.

(i) Algorithm 5.1 involves two solves (one in step (i) and one in step (ii)) of the forward elasticity problem to compute $\boldsymbol{G}$. This in contrast to the $N + 1$ solves required in algorithm 4.1.

(ii) Previous studies aimed at solving the inverse elasticity problem by treating it as an optimization problem have utilized a regularized version of the Gauss–Newton algorithm (see [8–10] for example). This algorithm involves evaluation of the so-called Jacobian matrix, a calculation that is embedded in step (ii) of algorithm 4.1. The cost associated with this algorithm also scales as $N + 1$ forward solves per iteration.

(iii) Interesting connections can be drawn between the algorithms proposed in [11, 12] and our approach. For the case of a single measurement, the algorithms proposed in these studies are equivalent to using steepest descent instead of a quasi-Newton algorithm with algorithm 5.1 to calculate the gradient.

6. Numerical examples

In this section we present numerical examples that demonstrate the effectiveness of the proposed approach. In all examples the gradient is computed using algorithm 5.1. To ensure that the shear modulus remains bounded away from zero in the entire domain, a dummy variable ξ is used in the optimization algorithm. The shear modulus μ is related to ξ via

$$\mu = 1 + \xi^2. \tag{42}$$

The expression $\mu = 1 + |\xi|$ has also been tried and found to yield similar results. However we prefer to use (42) since the derivative of $\mu = 1 + |\xi|$ is not defined at $\xi = 0$.

6.1. Problem description

A schematic of the example problem is shown in figure 1. The domain $\Omega = (0, L) \times (0, L)$, where $L = 10$ units. On the $x_2 = 0$ edge, $u_2 = 0.1$ and $h_1 = 0$; on the $x_2 = 10$ edge, $u_2 = 0$ and $h_1 = 0$, and on the $x_1 = 0$, 10 edges, $h_1 = 0$ and $h_2 = 0$, where u_i and h_i are components of the displacement and traction vectors $\boldsymbol{u}$ and $\boldsymbol{h}$, respectively. To prevent rigid translation of the body in the x_1-direction the centre point of the domain is constrained not to move in the x_1-direction. It is assumed that only the x_2-component of displacement is measured. Thus the only non-zero entry of the tensor $\boldsymbol{T}$ is $T_{22} = 1$. The shear modulus is assumed to be constant on the $x_2 = 0$ edge. The uniqueness of this problem for the incompressible case has been proven in [19].

For the inverse problem, we generate the 'measured' displacement field by solving the forward elasticity problem with the target material field, and then adding white Gaussian noise to the displacements. We consider three levels of noise corresponding to $\Delta = 3$, 1 and 0.3%, where Δ is given by

$$\Delta = \|u_2^m - \bar{\boldsymbol{u}} \cdot \boldsymbol{e}_2\|_\Omega / \|u_2^m\|_\Omega. \tag{43}$$

In the above equation $\bar{\boldsymbol{u}}$ is the displacement field with no noise. Note that values of $\Delta = 3$, 1 and 0.3% correspond approximately to a signal to noise ratios (SNR) of 30, 40 and 50 dB, respectively. It is believed that an SNR of 40 dB is typical for systems that use ultrasound to measure displacements (see [10] for example).

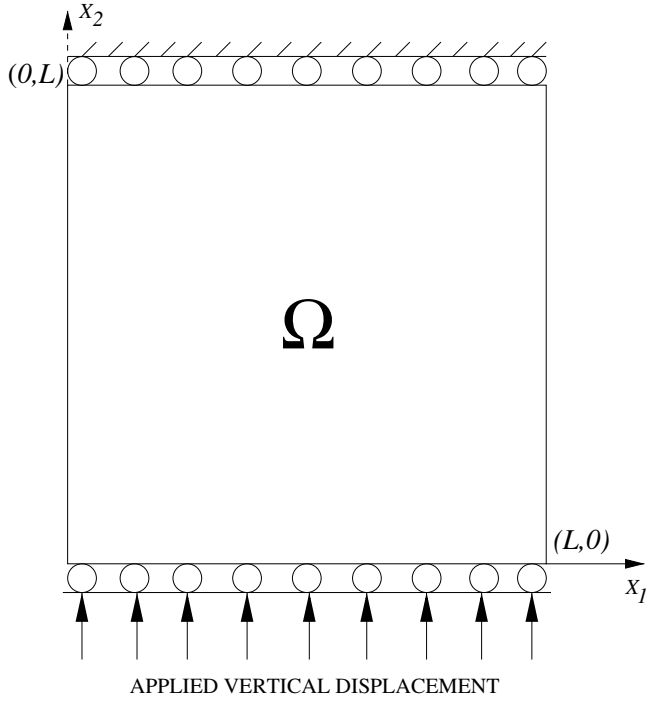


Figure 1. Schematic of the example problem.

Table 1. Choice of the regularization parameter.

Noise level, Δ	3×10^{-2}	1×10^{-2}	3×10^{-3}
Regularization parameter, α	2.5×10^{-6}	2.5×10^{-7}	2.5×10^{-8}

6.2. Regularization parameter

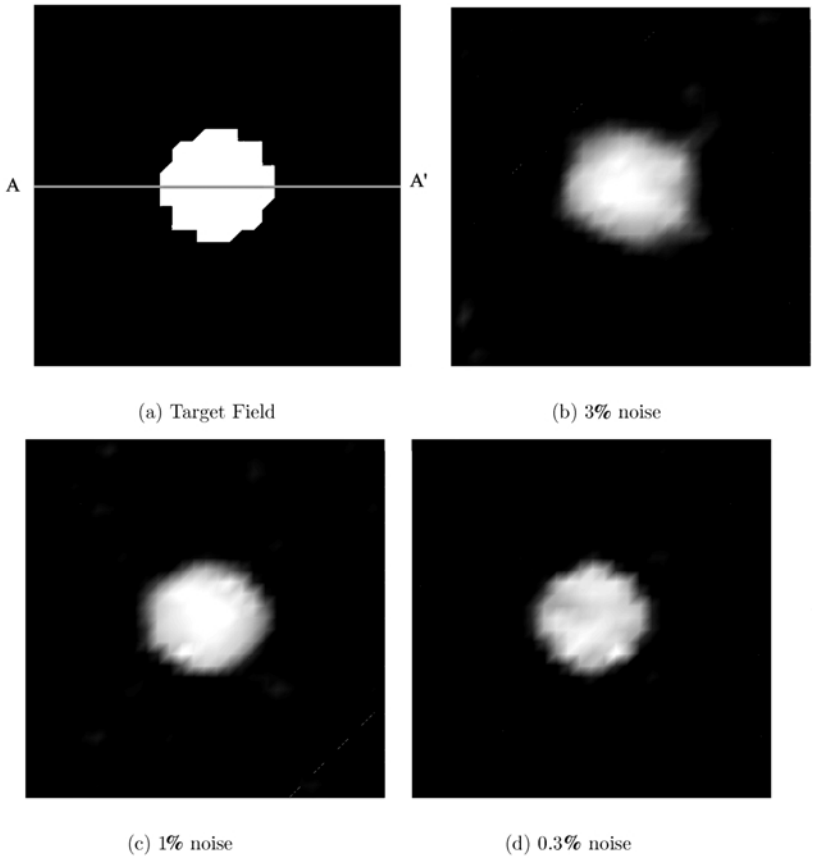
We choose the norm for the regularization term to be the L_2 norm on Ω. That is, $\|\cdot\|_b = \|\cdot\|_\Omega$. In choosing the regularization parameter α, we assume that the measured solution u_2^m has a known level of noise Δ. According to the theory of residues due to Morozov, α should be chosen to be the largest real number that allows

$$\|u_2^m - \boldsymbol{u}^h \cdot \boldsymbol{e}_2\|_\Omega = C\|u_2^m - \bar{\boldsymbol{u}} \cdot \boldsymbol{e}_2\|_\Omega = C\Delta\|u_2^m\|_\Omega, \tag{44}$$

for the optimal solution. In the equation above, C is a number that is of order 10^0, and is greater than unity. In practice, we have found that a value of $C \approx 1/\sqrt{2}$ provides the best results. Values much larger than this over-regularize the solution. Based on this criterion the values of α used for three different noise levels are shown in table 1. From this table, it is clear that, within the range considered, $\alpha \sim \Delta^2$.

6.3. Reconstructions

We perform all reconstructions using the limited memory BFGS algorithm [22]. The computer program that implements this algorithm was downloaded from http://www-fp.mcs.anl.gov/otc/Tools/LBFGS-B. In BFGS, as in any quasi-Newton method, within each iteration several calls may be made to the routine that calculates the functional and the gradient.



(a) Target Field (b) 3% noise

(c) 1% noise (d) 0.3% noise

Figure 2. Shear modulus reconstructions for a specimen with a single inclusion on a coarse mesh.

Each such call is termed a sub-iteration. Thus 25 sub-iterations correspond to 25 evaluations of the functional and the gradient, and in our case 25 solves of the primal and dual problems. In all cases we terminate the optimization algorithm at the smallest iteration number n for which $(\pi(\mu^n) - \pi(\mu^{n-5}))/\pi(\mu^{n-5}) < 0.01$. The total number of sub-iterations is labelled n_s. The system of linear equations resulting from the finite-element discretization of the primal and dual problems is solved using conjugate gradient iterations with an incomplete LU preconditioner (see [23] for example). The subroutines to perform these solves are part of the PETSC library [24]. All calculations are performed on a Linux workstation with a 1.7 GHz Pentium processor.

6.3.1. Single inclusion on a coarse mesh. First we consider the problem of a single inclusion in a uniform, soft matrix. The contrast, defined as the ratio of the modulus of the inclusion to the surrounding medium, is five. At 0.05% nominal strain, this ratio is typical of ductal carcinoma *in situ* surrounded by glandular tissue in the human breast (see [25] for example). The problem is solved on a 30×30 quadrilateral finite-element mesh. This leads to 931 parameters in the inverse problem. Their break-up is as follows. There are 961 nodes in the domain, out of which 31 were on the $x_2 = 0$ edge. It is assumed that the distribution of the shear modulus on this edge is known (uniform), however its value is not. Thus the 31 nodal values on the $x_2 = 0$ edge are represented by a single parameter, leading to $961 - 31 + 1 = 931$ total parameters.

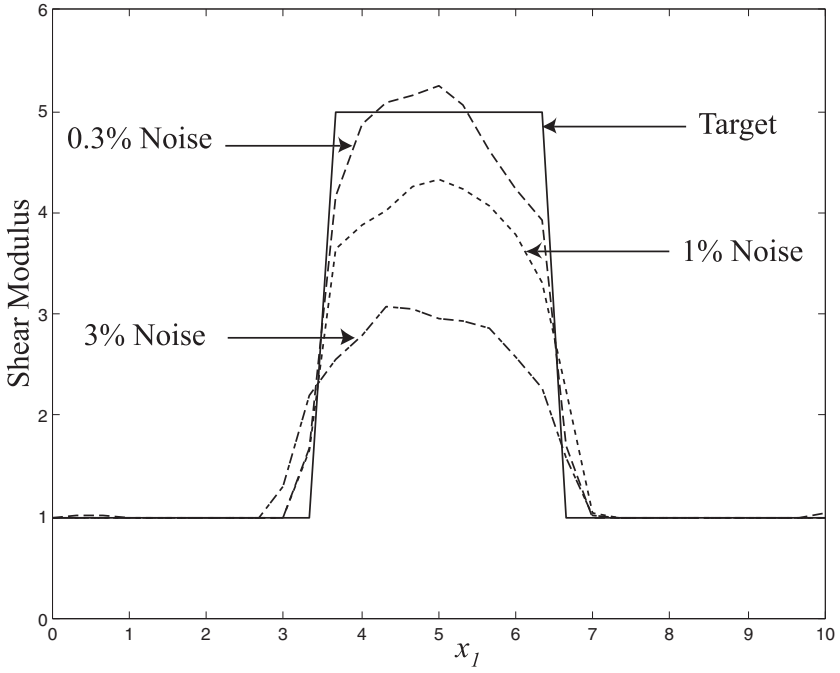


Figure 3. Shear modulus along A–A′.

Table 2. Output data for reconstructions.

Reconstruction	Sub-iterations (n_s)	Wall clock time (s)	Error (ϵ) (%)
30^2 mesh, $\Delta = 3\%$	17	14	39
30^2 mesh, $\Delta = 1\%$	24	20	23
30^2 mesh, $\Delta = 0.3\%$	55	46	13
100^2 mesh, $\Delta = 3\%$	20	261	40
100^2 mesh, $\Delta = 1\%$	32	418	23
100^2 mesh, $\Delta = 0.3\%$	41	535	14
150^2 mesh, $\Delta = 3\%$	22	882	40
150^2 mesh, $\Delta = 1\%$	28	1122	26
150^2 mesh, $\Delta = 0.3\%$	58	2324	15

In figure 2, the target distribution and three reconstructions of μ corresponding to different levels of noise are shown. To highlight the contrast, different grey scale maps are used within each plot. We observe that the shape of the inclusion is recovered accurately in all cases. In figure 3, the shear modulus along the line A–A′ is compared. In this figure, differences among the three reconstructions are apparent. With 3% noise the shear modulus of the inclusion is underestimated by about 40%, with 1% noise this value is 20%, and with 0.3% noise it is smaller still. These observations are quantified by calculating the L_2 error in the reconstructions, namely $\epsilon = \|\mu_{\text{target}} - \mu_{\text{predicted}}\|_\Omega / \|\mu_{\text{target}}\|_\Omega$. The values of ϵ, listed in table 2, are consistent with the observations made for figure 3.

The reduced accuracy of the reconstructed contrast ratio in the high-noise-level case can be ascribed to the higher value of regularization parameter used in that case. A higher value of α enhances the role of the regularizing term leading to solutions with lower L_2 norm of the

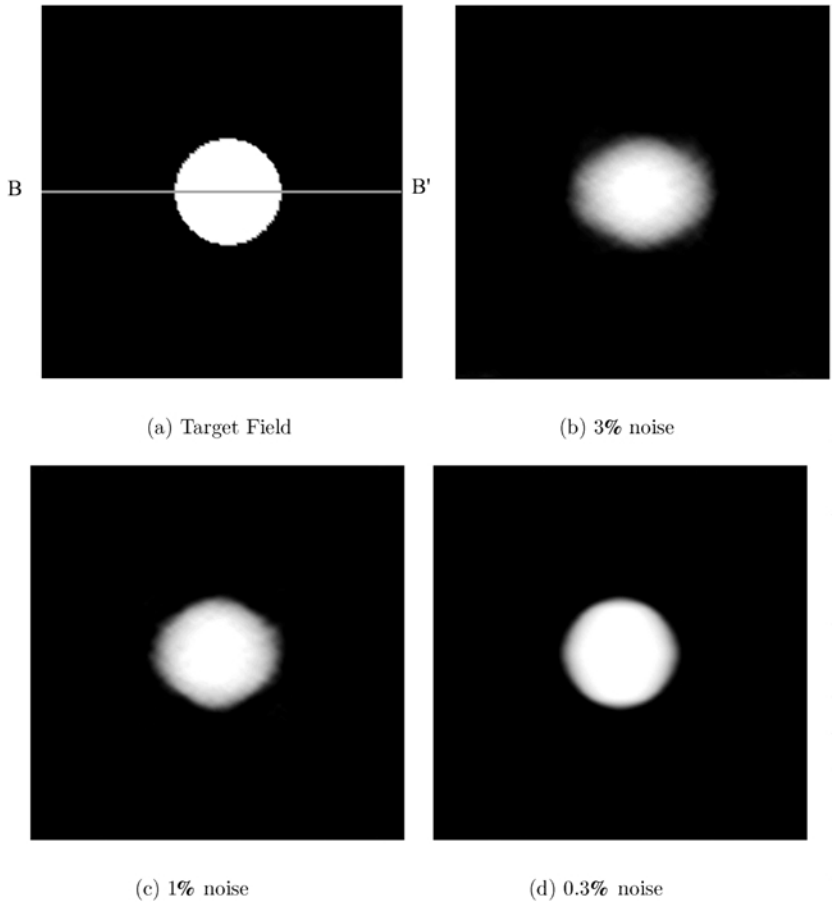


Figure 4. Shear modulus reconstructions for a specimen with a single inclusion on a fine mesh.

reconstructed field. Regularization terms based on other norms will have different effects. For instance, for the H^1 norm, we would expect reconstructions with smeared-out boundaries, but more accurate contrast ratios. We are currently exploring possibilities for what might be an optimal norm for regularizing the inverse elasticity problem.

6.3.2. Single inclusion with a fine mesh. We consider the same inclusion on a finer mesh of 150×150 quadrilateral finite elements. This represents a relatively large inverse problem with 22 651 parameters.

In figure 4, the target distribution and three reconstructions of μ with different levels of noise are shown. As for the previous case, different grey scale maps are used for each plot. From the figure it is clear that for all levels of noise the shape of the inclusion is recovered accurately, though for the 3.0 and 1.0% cases the circular inclusion is somewhat elongated in the x_1-direction. This artefact is absent from the reconstruction with 0.1% noise. In figure 5, the shear modulus along the line B–B′ is shown. Trends observed for the coarse mesh are seen to persist. These results are validated by calculating ϵ, the L_2 error in the reconstructions (see table 2).

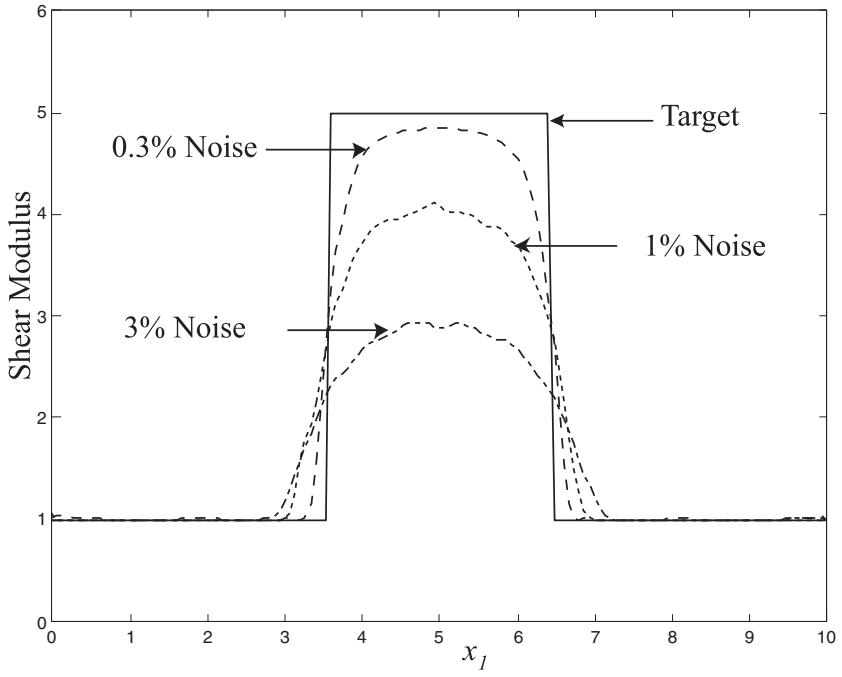


Figure 5. Shear modulus along line B–B′.

6.3.3. Multiple inclusions. Finally, we considerer two inclusions, each of different size and contrast. This problem is relevant to the imaging of human breast where fibroadenoma has been found to have a smaller contrast when compared with ductal carcinoma *in situ* (see [25]). We solve this problem on a 100×100 mesh with a total of 10 101 parameters.

The target shear modulus and three reconstructions are shown in figure 6. As before, different grey scale maps are used for each plot. From this figure it is clear that the location of both inclusions is recovered accurately for all cases; however, their shapes are recovered most accurately for data with least noise. In addition, an artefact in the form of a strip of material with lower contrast is seen running through the smaller of the two inclusions for cases with lower noise. It is found that this streak does not vanish even when the convergence criterion is tightened. Tests are being performed to estimate the minimum level of noise in the measured data required to remove this artefact. In figure 7, the shear modulus along the line C–C′ is plotted. In all cases the relative magnitude of the shear modulus of the two inclusions is recovered fairly accurately; however, the contrast with respect to the surrounding medium is most accurate for the reconstruction with least noise. The L_2 error in reconstruction is listed in table 2, and is consistent with the problems with single inclusions. In addition, from data presented in this table we draw the following conclusions.

(i) For a given level of noise in measured data the error in reconstruction is independent of the mesh size for the problem.
(ii) The number of sub-iterations to convergence increases with decreasing noise level. This is because the functional drops by a greater factor for cases with less noise and hence the optimization algorithm does more work in these cases.
(iii) For all cases, the overall time taken to solve the problem is reasonable. For $\Delta = 1\%$, which corresponds to the level of noise in a typical measurement, this value is about half

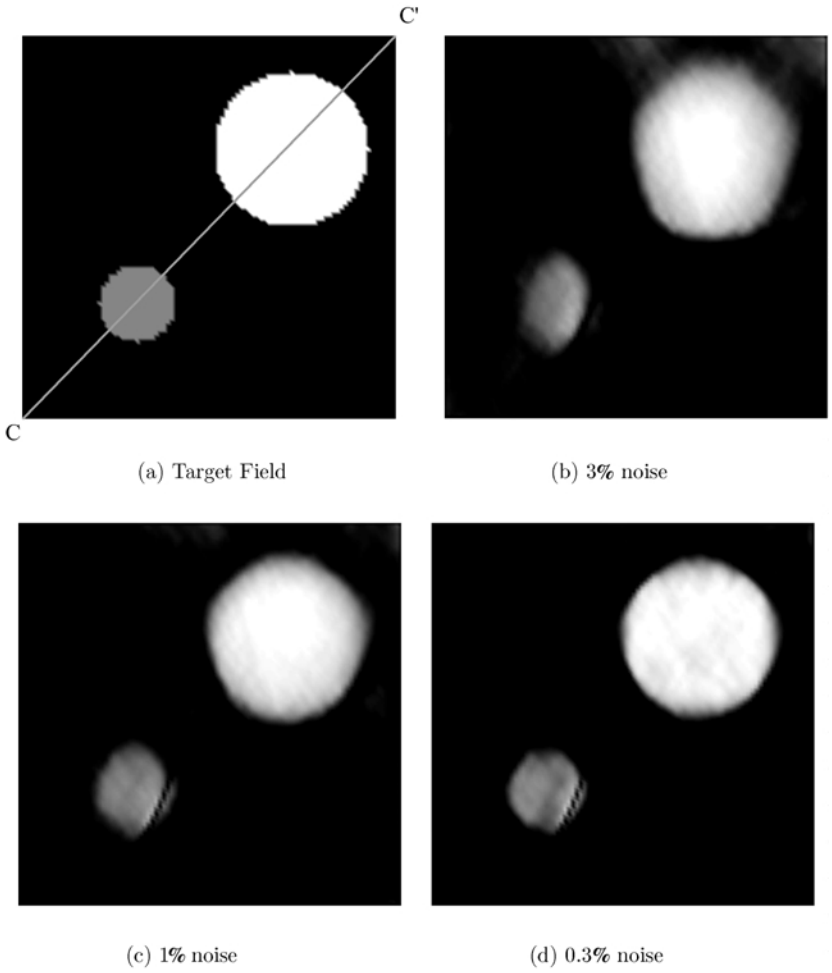


Figure 6. Shear modulus reconstructions for specimen with two inclusions.

a minute for the 30^2 mesh, 7 min for the 100^2 mesh and about 20 min for the 150^2 mesh. These results are obtained on a single 1.7 GHz Pentium processor.

7. Conclusions

We have developed and implemented an efficient formulation to solve the inverse elasticity problem. The main features of this approach are the use of a gradient based algorithm and efficient computation of the gradient using the adjoint equations. Using this formulation we have solved model problems with noise and studied the effect of regularizing terms. We have found the proposed algorithm is fast and yields accurate results for typical values of noise (SNR $\approx$ 40 dB) attained in experiments. Future work includes the application of this methodology to problems with multiple displacement measurements, time-harmonic fields and three-dimensional reconstructions.

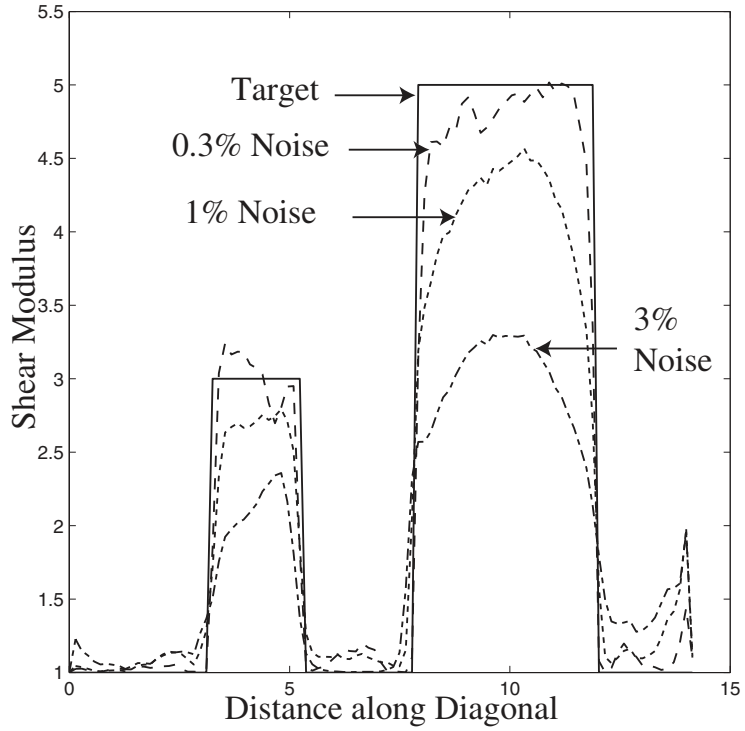


Figure 7. Shear modulus along C–C′.

Acknowledgments

The authors would like to acknowledge P E Barbone for many helpful and stimulating discussions on quantitative elasticity imaging and A Dwitama for help in writing the finite-element code used to solve the forward problem. The third author of this paper is supported by Sandia, a multiprogram laboratory operated by the Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-ACO4-94AL85000.

References

[1] Ophir J, Cespedes I, Ponnekanti H, Yazdi Y and Li X 1991 Elastography—a quantitative method for imaging the elasticity of biological tissues *Ultrason. Imaging* **13** 111–34

[2] Muthupillai R, Lomas D J, Rossman P J, Greenleaf J F, Manduca A and Ehman R L 1995 Magnetic resonance elastography by direct visualization of propagating acoustic strain waves *Science* **269** 1854–7

[3] Raghavan K R and Yagle A E 1994 Forward and inverse problems in elasticity imaging of soft tissues *IEEE Trans. Nucl. Sci.* **41** 1639–48

[4] Skovoroda A R, Emelianov S Y and O'Donnell M 1995 Tissue elasticity reconstruction based on ultrasonic displacement and strain images *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42** 747–65

[5] Sumi C, Suzuki A and Nakayama K 1995 Estimation of shear modulus distribution in soft tissue from strain distribution *IEEE Trans. Biomed. Eng.* **42** 193–202

[6] Plewes D B, Bishop J, Samani A and Sciaretta J 2000 Visualization and quantification of breast cancer biomechanical properties with magnetic resonance elastography *Phys. Med. Biol.* **45** 1591–610

[7] Bishop J, Samani A, Sciaretta J and Plewes D B 2000 Two dimensional MR elastography with linear inversion reconstruction: methodology and noise analysis *Phys. Med. Biol.* **45** 2081–91

[8] Kallel F and Bertrand M 1996 Tissue elasticity reconstruction using linear perturbation method *IEEE Trans. Med. Imaging* **15** 299–313
[9] Van Houten E E W, Paulsen K D, Miga M I, Kennedy F E and Weaver J B 1999 An overlapping subzone technique for MR-based elastic property reconstruction *Magn. Reson. Med.* **99** 779–86
[10] Doyley M M, Meaney P M and Bamber J C 2000 Evaluation of an iterative reconstruction method for quantitative elasticity *Phys. Med. Biol.* **45** 1521–40
[11] Natterer F and Wubbeling F 1995 A propagation–backpropagation method for ultrasound tomography *Inverse Problems* **11** 1225–32
[12] Dorn O, Bertete-Aguirre H, Berryman J G and Papanicolaou G C 1999 A nonlinear inversion method for 3D electromagnetic imaging using adjoint fields *Inverse Problems* **15** 1523–58
[13] Feijóo G R, Malhotra M, Oberai A A and Pinsky P M 2001 Shape sensitivity calculations for exterior acoustics problems *Eng. Comput.* **18** 376–91
[14] Jameson A 1988 Aerodynamic design via control theory *J. Sci. Comput.* **3** 233–60
[15] Gurtin M E 1981 *An Introduction to Continuum Mechanics* (New York: Academic)
[16] Hughes T J R 2000 *The Finite Element Method—Linear Static and Dynamic Finite Element Analysis* (Mineola, NY: Dover)
[17] Colton D and Kress R 1998 *Inverse Acoustic and Electromagnetic Scattering Theory* 2nd edn (Berlin: Springer)
[18] Isakov V 1998 *Inverse Problems for Partial Differential Equations* 1st edn (New York: Springer)
[19] Barbone P E and Bamber J C 2002 Quantitative elasticity imaging: what can and cannot be inferred from strain images *Phys. Med. Biol.* **47** 2147–64
[20] Barbone P E and Gokhale N H, Elastic modulus imaging: ill-posed and well-posed incompressible elastic inverse problems, in preparation
[21] Gill P E, Murray W and Wright M H 2000 *Practical Optimization* 12th edn (London: Academic)
[22] Zhu C, Byrd R H and Nocedal J 1997 L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization *ACM Trans. Math. Softw.* **23** 550–60
[23] Saad Y 1995 *Iterative Methods for Sparse Linear Systems* 1st edn (Boston, MA: PWS)
[24] Balay S, Gropp W D, McInnes L C and Smith B F 2002 PETSc users manual *Argonne National Laboratory Technical Report* ANL-95/11—Revision 2.1.3
[25] Wellman P, Howe R H, Dalton E and Kern K A 1999 Breast tissue stiffness in compression is correlated to histological diagnosis *Harvard BioRobotics Laboratory Technical Report*