

RESEARCH

Open Access

A first-order adjoint and a second-order hybrid method for an energy output least-squares elastography inverse problem of identifying tumor location

Nathan D Cahill¹, Baasansuren Jadamba¹, Akhtar A Khan^{1*}, Miguel Sama² and Brian C Winkler¹

*Correspondence: aaksma@rit.edu
¹Center for Applied and Computational Mathematics, School of Mathematical Sciences, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623, USA
Full list of author information is available at the end of the article

Abstract

In this paper we investigate the elastography inverse problem of identifying cancerous tumors within the human body. From a mathematical standpoint, the elastography inverse problem consists of identifying the variable Lamé parameter μ in a system of linear elasticity where the underlying object exhibits nearly incompressible behavior. This problem is subsequently posed as an optimization problem using an energy output least-squares (EOLS) functional, but the nonlinearity that arises makes the computation of the EOLS functional's derivatives challenging. We employ an adjoint method for the computation of the gradient, something shown to be an efficient method in recent studies, and also give a parallelizable hybrid method for the computation of the EOLS functional's second derivative. Detailed discrete formulas and nontrivial computational examples are provided to show the feasibility of both the adjoint and hybrid approaches. Furthermore, all results are given in the framework of a general saddle point problem allowing easy adaptation to numerous other inverse problems.

MSC: 35R30; 65N30

Keywords: inverse problems; parameter identification; nearly incompressible elasticity; tumor identification; energy output least squares; output least squares; regularization; mixed finite element method; saddle point problems

1 Introduction

Consider the following system of partial differential equations describing the response of an isotropic elastic object to certain body forces and traction applied to its boundary:

$$-\nabla \cdot \sigma = f \quad \text{in } \Omega, \quad (1a)$$

$$\sigma = 2\mu\epsilon(u) + \lambda \operatorname{div} uI, \quad (1b)$$

$$u = g \quad \text{on } \Gamma_1, \quad (1c)$$

$$\sigma n = h \quad \text{on } \Gamma_2. \quad (1d)$$

Here the domain Ω is a subset of \mathbb{R}^2 or \mathbb{R}^3 and $\partial\Omega = \Gamma_1 \cup \Gamma_2$ is its boundary. In (1a)-(1d), the vector-valued function $u = u(x)$ represents the displacement of the elastic object, f is

the applied body force, n is the unit outward normal, and

$$\epsilon(u) = \frac{1}{2}(\nabla u + \nabla u^T)$$

is the linearized strain tensor. The resulting stress tensor σ in the stress-strain law (1b) is obtained under the assumption that the elastic object is isotropic and the displacement is small enough so that a linear relationship holds. The Lamé parameters μ and λ quantify the elastic properties of the material. (In the following, for simplicity we set $g = 0$.)

In this work our objective is to investigate the elastography (also known as elasticity imaging) inverse problem of locating cancerous tumors within the human body. This inverse problem consists of identifying the variable parameter μ in (1a)-(1d) from a measurement of the displacement field u . Conversely, the direct problem for (1a)-(1d) is to find the displacement u when function h , the variable coefficients μ and λ , and the body force f are all known. The underlying idea is that differences in molecular makeup as well as microscopic and macroscopic structure result in significant differences in the stiffness of living soft tissue (see [1]). Moreover, changes in tissue stiffness generally correlate with changes in pathological state, with many cancers appearing as hard nodules within the surrounding softer tissue. In a clinical setting, measurements of displacement in human tissue can be obtained using ultrasound and this can then serve as data in the context of the elastography inverse problem. By solving this inverse problem and recovering μ , tumor locations can be identified using the marked differences in elastic properties between the healthy and unhealthy tissue. Additionally, we note that in the elastography inverse problem the human body is treated as a nearly incompressible object where the parameter λ is significantly large and hence only the parameter μ is sought.

Although numerous authors have contributed to using the elasticity properties of soft tissue as a tool to differentiate between normal and cancerous tissue, Raghavan and Yagle [2] were among the first authors to realize that this study can be best done in an inverse problem framework using measured strains and the equations of equilibrium to recover elasticity (*cf.* (1a)-(1d)). Since then, many studies have been devoted to investigating various aspects of the elastography inverse problem and the interested reader is referred to [3–8] and the cited reference therein. Additionally, a detailed account of the recent developments in elastography inverse problem can be found in the survey article by Doyley [1]. See also [9–21] and the cited references therein for more details.

One of the main technical challenges in the study of this inverse problem stems from the fact that the human body is treated as a nearly incompressible object. That is, the elasticity modulus λ is significantly large (and particularly $\lambda \gg \mu$), rendering classical finite element methods ineffective due to the so-called locking effect. In the literature, several approaches have been proposed to overcome the locking effect, and in this work we employ the mixed finite elements strategy.

In the following, we provide the necessary details for the transformation of system (1a)-(1d) into a saddle point problem to which the mixed finite element approach can be applied.

We begin by recalling that the dot product of two tensors A and B can be denoted by $A \cdot B$. That is, for 2×2 tensors A and B , we have

$$A \cdot B = A_{11}B_{11} + A_{12}B_{12} + A_{21}B_{21} + A_{22}B_{22}.$$

Given a sufficiently smooth domain $\Omega \subset \mathbb{R}^2$, the L_2 -norm of a tensor-valued function $A = A(x)$ is given by

$$\|A\|_{L^2}^2 = \|A\|_{L^2(\Omega)}^2 = \int_{\Omega} A \cdot A = \int_{\Omega} (A_{11}^2 + A_{12}^2 + A_{21}^2 + A_{22}^2).$$

On the other hand, for a vector-valued function $u(x) = (u_1(x), u_2(x))^T$, the L_2 -norm is given by

$$\|u\|_{L^2}^2 = \|u\|_{L^2(\Omega)}^2 = \int_{\Omega} (u_1^2 + u_2^2),$$

whereas the H^1 -norm by

$$\|u\|_{H^1}^2 = \|u\|_{H^1(\Omega)}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2.$$

In the following discussion, for the sake of simplicity, in (1a)-(1d) we set $g = 0$. For this case, the space of test functions, denoted by \widehat{V} , is given by

$$\widehat{V} = \{\bar{v} \in H^1(\Omega) \times H^1(\Omega) : \bar{v} = 0 \text{ on } \Gamma_1\}.$$

By using Green's identity and boundary conditions (1c) and (1d), we obtain the following weak form of elasticity system (1a)-(1d): Find $\bar{u} \in \widehat{V}$ such that

$$\int_{\Omega} 2\mu \epsilon(\bar{u}) \cdot \epsilon(\bar{v}) + \int_{\Omega} \lambda (\operatorname{div} \bar{u})(\operatorname{div} \bar{v}) = \int_{\Omega} f \bar{v} + \int_{\Gamma_2} \bar{v} h \quad \text{for every } \bar{v} \in \widehat{V}. \quad (2)$$

The mixed finite elements strategy, which, in the present context, consists of introducing a pressure term $p \in Q = L^2(\Omega)$, is as follows:

$$\operatorname{div} \bar{u} = \frac{p}{\lambda}. \quad (3)$$

As $\lambda \rightarrow \infty$, (3) yields the incompressibility limit

$$\operatorname{div} \bar{u} = 0.$$

The weak formulation of (3) reads

$$\int_{\Omega} (\operatorname{div} \bar{u}) q - \int_{\Omega} \frac{1}{\lambda} p q = 0 \quad \text{for every } q \in Q. \quad (4)$$

By using relation (3), the weak form (2) can be expressed as follows: Find $\bar{u} \in \widehat{V}$ such that

$$\int_{\Omega} 2\mu \epsilon(\bar{u}) \cdot \epsilon(\bar{v}) + \int_{\Omega} p (\operatorname{div} \bar{v}) = \int_{\Omega} f \bar{v} + \int_{\Gamma_2} \bar{v} h \quad \text{for every } \bar{v} \in \widehat{V}, \quad (5)$$

where the pressure p is also an unknown.

Therefore, the problem of finding $\bar{u} \in \widehat{V}$, satisfying (2), has now been converted into the saddle point problem of finding $(\bar{u}, p) \in \widehat{V} \times Q$ such that

$$\int_{\Omega} 2\mu \epsilon(\bar{u}) \cdot \epsilon(\bar{v}) + \int_{\Omega} p(\operatorname{div} \bar{v}) = \int_{\Omega} f \bar{v} + \int_{\Gamma_2} \bar{v} h \quad \text{for every } \bar{v} \in \widehat{V}, \quad (6a)$$

$$\int_{\Omega} (\operatorname{div} \bar{u}) q - \int_{\Omega} \frac{1}{\lambda} p q = 0 \quad \text{for every } q \in Q, \quad (6b)$$

where $Q = L^2(\Omega)$ and $\widehat{V} = \{\bar{v} \in H^1(\Omega) \times H^1(\Omega) : \bar{v} = 0 \text{ on } \Gamma_1\}$.

For the saddle point formulation, the Babuška-Brezzi condition provides guidance in the choice of finite element spaces necessary for a stable numerical approximation (see [22]).

The primary objective of this work is to develop an efficient computational framework for the elastography inverse problem. For this we employ an adjoint approach for the derivative computation of a recently proposed energy output least-squares (EOLS) functional [23]. We recall that Oberai *et al.* [24] used the adjoint approach to compute efficiently the gradient of the output least-squares functional. Inspired by Tortorelli and Michaleris [25], we also devise a hybrid method for an efficient computation of the second-order derivative of the EOLS functional. In this direction, we would also like to draw attention to an interesting paper by Cioacaa, Alexea, and Sandua [26] where a second-order adjoint method is studied. All the results and formulas given are for a general saddle point problem and hence can easily be adapted to a wide range of inverse problems for variational problems (see [27]). In the derivation of the adjoint formulas, we do not include the regularization functional while considering the EOLS functional. However, we use a smooth regularizer for the identification of a smooth parameter and a BV regularizer for the identification of discontinuous coefficients.

2 Optimization approach for inverse problems in saddle point problems

Let \widehat{V} and Q be real Hilbert spaces, let B be a real Banach space, and let A be a nonempty, closed, and convex subset of B . Here B is the coefficient/parameter space and A is the set of all admissible coefficients. Let $a : B \times \widehat{V} \times \widehat{V} \rightarrow \mathbb{R}$ be a trilinear map which we assume to be symmetric with respect to the second and third arguments. That is, for every $\ell \in B$ and for all $\bar{u}, \bar{v} \in \widehat{V}$, we have $a(\ell, \bar{u}, \bar{v}) = a(\ell, \bar{v}, \bar{u})$. Let $b : \widehat{V} \times Q \rightarrow \mathbb{R}$ be a bilinear form, let $c : Q \times Q \rightarrow \mathbb{R}$ be a symmetric bilinear form, and let $m : \widehat{V} \rightarrow \mathbb{R}$ be a linear and continuous map. We assume that there are positive constants $\kappa_1, \kappa_2, \varsigma_1, \varsigma_2$, and κ_0 such that the following inequalities hold:

$$a(\ell, \bar{v}, \bar{v}) \geq \kappa_1 \|\bar{v}\|^2 \quad \text{for every } \bar{v} \in \widehat{V}, \text{ for every } \ell \in A, \quad (7a)$$

$$a(\ell, \bar{u}, \bar{v}) \leq \kappa_2 \|\ell\| \|\bar{u}\| \|\bar{v}\| \quad \text{for every } \bar{u}, \bar{v} \in \widehat{V}, \text{ for every } \ell \in A, \quad (7b)$$

$$c(q, q) \geq \varsigma_1 \|q\|^2 \quad \text{for every } q \in Q, \quad (7c)$$

$$|c(p, q)| \leq \varsigma_2 \|p\| \|q\| \quad \text{for every } p, q \in Q, \quad (7d)$$

$$|b(\bar{v}, q)| \leq \kappa_0 \|\bar{v}\| \|q\| \quad \text{for every } \bar{v} \in \widehat{V}, \text{ for every } q \in Q. \quad (7e)$$

Remark 2.1 We remark that for the subsequent development of our approach, it suffices to assume that A is a closed and convex set of admissible parameters. Most commonly, it is chosen as the set of box-constraints. In some works, the space in which A resides is

required to be compactly embedded in the solution space (see [28–31]). In our discrete examples, we have used linear elements to approximate the imposed box-constraints.

We consider the following saddle point problem: Given $\ell \in A$, find $(\bar{u}, p) \in \widehat{V} \times Q$ such that

$$a(\ell, \bar{u}, \bar{v}) + b(\bar{v}, p) = m(\bar{v}) \quad \text{for every } \bar{v} \in \widehat{V}, \quad (8a)$$

$$b(\bar{u}, q) - c(p, q) = 0 \quad \text{for every } q \in Q. \quad (8b)$$

Given all the data, the direct problem in this setting is to find (\bar{u}, p) . However, our focus is on the inverse problem of finding a parameter $\ell \in A$ that makes (8a)-(8b) true for a measurement (\bar{z}, \hat{z}) of (\bar{u}, p) .

Evidently, saddle point problem (6a)-(6b) connected to the elastography inverse problem of identifying a variable parameter μ in the system of incompressible linear elasticity can be deduced by setting:

$$a(\mu, \bar{u}, \bar{v}) = \int_{\Omega} 2\mu \epsilon(\bar{u}) \cdot \epsilon(\bar{v}), \quad (9a)$$

$$b(\bar{u}, q) = \int_{\Omega} (\operatorname{div} \bar{u}) q, \quad (9b)$$

$$c(p, q) = \int_{\Omega} \frac{1}{\lambda} p q, \quad (9c)$$

$$m(\bar{v}) = \int_{\Omega} f \bar{v} + \int_{\Gamma_2} \bar{v} h. \quad (9d)$$

A common approach to solve inverse problems of parameter identification in PDEs is to minimize the output least-squares functional, which, in the present context, can be defined by

$$J_{\text{OLS}}(\ell) = \frac{1}{2} \|u(\ell) - z\|_V^2 = \frac{1}{2} \|\bar{u}(\ell) - \bar{z}\|_{\widehat{V}}^2 + \frac{1}{2} \|p(\ell) - \hat{z}\|_Q^2, \quad (10)$$

where $V = \widehat{V} \times Q$, $z = (\bar{z}, \hat{z}) \in V$ is the measured data, and $u(\ell) = (\bar{u}(\ell), p(\ell)) \in V$ is the solution of (8a)-(8b) corresponding to ℓ .

The output least-squares solution to the inverse problem of identifying ℓ is the one that solves the following optimization problem: Find $\bar{\ell} \in A$ such that

$$J_{\text{OLS}}(\bar{\ell}) \leq J_{\text{OLS}}(\ell) \quad \text{for every } \ell \in A.$$

Recently, in [23], the following objective functional was proposed to solve the inverse problem of identifying the variable parameter $\ell \in A$ in saddle point problem (8a)-(8b):

$$J(\ell) = \frac{1}{2} a(\ell, \bar{u}(\ell) - \bar{z}, \bar{u}(\ell) - \bar{z}) + \frac{1}{2} c(p(\ell) - \hat{z}, p(\ell) - \hat{z}), \quad (11)$$

where $z = (\bar{z}, \hat{z})$ is the measured data and $u(\ell) = (\bar{u}(\ell), p(\ell))$ is the solution of (8a)-(8b) corresponding to ℓ .

Clearly, to solve an optimization problem with the above objective functional, we need to compute its derivative which, in turn, requires us to compute the derivative of the solution map. It is well known that one of the most challenging aspects in the study of inverse problems is in finding an efficient computation of the derivative of the solution map. We will now develop an adjoint method for the computation of the first derivative of the EOLS functional and then a new hybrid method for the computation of the functional's second derivative.

For every $\ell \in A$, the map $\ell \rightarrow S(\ell) = (\bar{u}(\ell), p(\ell))$ is well defined and single-valued. The following result for the differentiability of S , which was announced in [23] without a proof, will be needed.

Theorem 2.1 *For each ℓ in the interior of A , $u = u(\ell) = (\bar{u}(\ell), p(\ell))$ is infinitely differentiable at ℓ .*

1. *Given u , the first derivative $\delta u = (\delta \bar{u}, \delta p) = (D\bar{u}(\ell)\delta\ell, Dp(\ell)\delta\ell)$ is the unique solution of the saddle point problem:*

$$a(\ell, \delta \bar{u}, \bar{v}) + b(\bar{v}, \delta p) = -a(\delta \ell, \bar{u}, \bar{v}) \quad \text{for every } \bar{v} \in \widehat{V}, \quad (12a)$$

$$b(\delta \bar{u}, q) - c(\delta p, q) = 0 \quad \text{for every } q \in Q. \quad (12b)$$

2. *The second-order derivative*

$$\delta^2 u = (\delta^2 \bar{u}, \delta^2 p) = (D^2 \bar{u}(\ell)(\delta \ell_1, \delta \ell_2), D^2 p(\ell)(\delta \ell_1, \delta \ell_2))$$

is the unique solution of the saddle point problem

$$a(\ell, \delta^2 \bar{u}, \bar{v}) + b(\bar{v}, \delta^2 p) = -a(\delta \ell_2, D\bar{u}(\ell)\delta \ell_1, \bar{v}) - a(\delta \ell_1, D\bar{u}(\ell)\delta \ell_2, \bar{v}) \quad \text{for every } \bar{v} \in \widehat{V}, \quad (13a)$$

$$b(\delta^2 \bar{u}, q) - c(\delta^2 p, q) = 0 \quad \text{for every } q \in Q. \quad (13b)$$

Proof We define a map $G : A \times \widehat{V} \rightarrow \widehat{V}^* \times Q^*$ by $G(\ell, (\bar{u}, p)) = (a(\ell, \bar{u}) + b(p) - m, b(\bar{u}) - c(p))$, where \widehat{V}^* and Q^* are the duals of \widehat{V} and Q , and $a(\ell, \bar{u})$, $b(p)$, and $c(p)$ are the associated dual elements given by the Riesz theorem. Then saddle point problem (8a)-(8b) is equivalent to the following implicit equation:

$$G(\ell, u) = (0_{\widehat{V}^*}, 0_{Q^*}). \quad (14)$$

The differentiability of $u = u(\ell)$ follows from the implicit function theorem. In fact, the map G is infinitely differentiable and the partial derivative with respect to variable $u = (\bar{u}, p)$ is given by

$$D_u G(\ell, (\bar{u}, p))(\delta \bar{u}, \delta p) \equiv (a(\ell, \delta \bar{u}) + b(\delta p), b(\delta \bar{u}) - c(\delta p)) \quad \text{for every } \delta \bar{u} \in \widehat{V}.$$

By [22, Proposition 4], the map $D_u G(\ell, (\bar{u}, p)) : A \times \widehat{V} \rightarrow \widehat{V}^* \times Q^*$ is an isomorphism. Therefore, using the implicit function theorem, the map $u = u(\ell)$ is infinitely differentiable at any ℓ in the interior of A .

We now compute the first and second derivatives of the coefficient-to-solution map. By using equation (8a), for any $\hat{\ell} \in A$ and for any sufficiently small $t \in \mathbb{R}_+$, we have

$$\begin{aligned} a(\ell + t\hat{\ell}, \bar{u}(\ell + t\hat{\ell}), \bar{v}) + b(\bar{v}, p(\ell + t\hat{\ell})) &= m(\bar{v}), \\ a(\ell, \bar{u}(\ell), \bar{v}) + b(\bar{v}, p(\ell)) &= m(\bar{v}), \end{aligned}$$

and by manipulating the terms in these two equations, we obtain

$$a\left(\ell + t\hat{\ell}, \frac{\bar{u}(\ell + t\hat{\ell}) - \bar{u}(\ell)}{t}, \bar{v}\right) + a(\hat{\ell}, \bar{u}(\ell), \bar{v}) + b\left(\bar{v}, \frac{p(\ell + t\hat{\ell}) - p(\ell)}{t}\right) = 0,$$

which, by passing the above equation to the limit when $t \rightarrow 0^+$, yields (12a)

$$a(\ell, D\bar{u}(\ell)(\hat{\ell}), \bar{v}) + b(\bar{v}, Dp(\ell)(\hat{\ell})) = -a(\hat{\ell}, \bar{u}(\ell), \bar{v}).$$

Analogously, using equation (8b), for any $\hat{\ell} \in A$ and for every sufficiently small $t \in \mathbb{R}_+$, we have

$$\begin{aligned} b(u(\ell + t\hat{\ell}), q) - c(p(\ell + t\hat{\ell}), q) &= 0, \\ b(u(\ell), q) - c(p(\ell), q) &= 0, \end{aligned}$$

and by manipulating the above two equations, we obtain

$$b\left(\frac{u(\ell + t\hat{\ell}) - u(\ell)}{t}, q\right) - c\left(\frac{p(\ell + t\hat{\ell}) - p(\ell)}{t}, q\right) = 0,$$

which, by passing the above equation to limit $t \rightarrow 0^+$, gives

$$b(D\bar{u}(\ell)(\hat{\ell}), q) - c(Dp(\ell)(\hat{\ell}), q) = 0,$$

which is (12b). Consequently, (12a) and (12b) characterize the first derivative.

The same arguments can be used to compute the form of the second derivative. From (12a), for any $\hat{\ell}_1, \hat{\ell}_2 \in A$ and for any sufficiently small $t \in \mathbb{R}_+$, we have

$$\begin{aligned} a(\ell + t\hat{\ell}_1, D\bar{u}(\ell + t\hat{\ell}_1)(\hat{\ell}_2), \bar{v}) + b(\bar{v}, Dp(\ell + t\hat{\ell}_1)(\hat{\ell}_2)) &= -a(\hat{\ell}_2, \bar{u}(\ell + t\hat{\ell}_1), \bar{v}), \\ a(\ell, D\bar{u}(\ell)(\hat{\ell}_2), \bar{v}) + b(\bar{v}, Dp(\ell)(\hat{\ell}_2)) &= -a(\hat{\ell}_2, \bar{u}(\ell), \bar{v}), \end{aligned}$$

and by rearranging the above set of equations, we obtain

$$\begin{aligned} a(\hat{\ell}_1, D\bar{u}(\ell + t\hat{\ell}_1)(\hat{\ell}_2), \bar{v}) + a\left(\ell, \frac{D\bar{u}(\ell + t\hat{\ell}_1)(\hat{\ell}_2) - D\bar{u}(\ell)(\hat{\ell}_2)}{t}, \bar{v}\right) \\ + b\left(\bar{v}, \frac{Dp(\ell + t\hat{\ell}_1)(\hat{\ell}_2) - Dp(\ell)(\hat{\ell}_2)}{t}\right) \\ = -a\left(\hat{\ell}_2, \frac{\bar{u}(\ell + t\hat{\ell}_1) - \bar{u}(\ell)}{t}, \bar{v}\right). \end{aligned}$$

Since the solution map $u = u(\ell)$ is twice Fréchet differentiable, by passing to the limit $t \rightarrow 0^+$, we get

$$\begin{aligned} a(\hat{\ell}_1, D\bar{u}(\ell)(\hat{\ell}_2), \bar{v}) + a(\ell, D^2\bar{u}(\hat{\ell}_1, \hat{\ell}_2), \bar{v}) + b(\bar{v}, D^2p(\hat{\ell}_1, \hat{\ell}_2)) \\ = -a(\hat{\ell}_2, D\bar{u}(\ell)(\hat{\ell}_1), \bar{v}), \end{aligned} \quad (15)$$

which, after a rearrangement of terms, yields (13a)

$$a(\ell, D^2\bar{u}(\hat{\ell}_1, \hat{\ell}_2), \bar{v}) + b(\bar{v}, D^2p(\hat{\ell}_1, \hat{\ell}_2)) = -a(\hat{\ell}_2, D\bar{u}(\ell)(\hat{\ell}_1), \bar{v}) - a(\hat{\ell}_1, D\bar{u}(\ell)(\hat{\ell}_2), \bar{v}).$$

From (12b), for any $\hat{\ell}_1, \hat{\ell}_2 \in A$ and any sufficiently small $t \in \mathbb{R}_+$, we have

$$\begin{aligned} b(D\bar{u}(\ell + \hat{\ell}_1)(\hat{\ell}_2), q) - c(Dp(\ell + \hat{\ell}_1)(\hat{\ell}_2), q) &= 0, \\ b(D\bar{u}(\ell)(\hat{\ell}_2), q) - c(Dp(\ell)(\hat{\ell}_2), q) &= 0, \end{aligned}$$

and by rearranging the above two equations, we get

$$b\left(\frac{D\bar{u}(\ell + t\hat{\ell}_1)(\hat{\ell}_2) - D\bar{u}(\ell)(\hat{\ell}_2)}{t}, q\right) - c\left(\frac{Dp(\ell + \hat{\ell}_1)(\hat{\ell}_2) - Dp(\ell)(\hat{\ell}_2)}{t}, q\right) = 0.$$

By passing to limit $t \rightarrow 0^+$, we finally deduce

$$b(D^2\bar{u}(\hat{\ell}_1, \hat{\ell}_2), q) - c(D^2p(\hat{\ell}_1, \hat{\ell}_2), q) = 0,$$

which in conjunction with (13b) forms the corresponding saddle point whose unique solution characterizes the second derivative $D^2u(\hat{\ell}_1, \hat{\ell}_2) = (D^2\bar{u}(\hat{\ell}_1, \hat{\ell}_2), D^2p(\hat{\ell}_1, \hat{\ell}_2))$. \square

3 An adjoint and a hybrid method for the energy output least squares

The developed adjoint method for the EOLS functional,

$$J(\ell) = \frac{1}{2}a(\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + \frac{1}{2}c(p - \hat{z}, p - \hat{z}),$$

is based on the key observation that the underlying saddle point problem can equivalently be posed as a variational equation of finding $u = (\bar{u}, p) \in V = \hat{V} \times Q$ such that

$$T(\ell, u, v) = m(\bar{v}) \quad \text{for every } v = (\bar{v}, q) \in V, \quad (16)$$

where

$$T(\ell, u, v) = a(\ell, \bar{u}, \bar{v}) + b(\bar{v}, p) + b(\bar{u}, q) - c(p, q). \quad (17)$$

By a direct computation, we have

$$\begin{aligned} DT(\ell, u, v)(\delta\ell) &= a(\delta\ell, \bar{u}, \bar{v}) + a(\ell, \delta\bar{u}, \bar{v}) \\ &\quad + b(\bar{v}, \delta p) + b(\delta\bar{u}, q) - c(\delta p, q) \\ &= a(\delta\ell, \bar{u}, \bar{v}) + t(\ell, \delta u, v). \end{aligned} \quad (18)$$

We define

$$J(\ell, v) = J(\ell) + T(\ell, u, v) - m(\bar{v}) \quad \text{for every } v \in V,$$

and, by using (16), notice that

$$J(a, v) = J(a) \quad \text{for every } v \in V.$$

Therefore, for any 'test function' $v = (\bar{v}, q) \in V$, we have

$$DJ(\ell)(\delta\ell) = D_\ell J(\ell, v)(\delta\ell) \quad \text{for every } \delta\ell \in A, \quad (19)$$

where D_ℓ stands for the partial derivative with respect to ℓ .

The key idea behind the adjoint method is to choose a particular v to avoid the computation of δu . By a direct computation and taking into account (18), we obtain

$$\begin{aligned} D_\ell J(\ell, v)(\delta\ell) &= \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\ell, \delta\bar{u}, \bar{u} - \bar{z}) + c(\delta p, p - \hat{z}) \\ &\quad + a(\delta\ell, \bar{u}, \bar{v}) + T(\ell, \delta u, v). \end{aligned} \quad (20)$$

Now, let $w = w(\ell)$ be the unique solution of the saddle point problem

$$a(\ell, \bar{w}, \bar{v}) + b(\bar{v}, p_w) = -a(\ell, \bar{u} - \bar{z}, \bar{v}) - b(\bar{v}, p - \hat{z}) \quad \text{for every } \bar{v} \in \widehat{V}, \quad (21a)$$

$$b(\bar{w}, q) - c(p_w, q) = 0 \quad \text{for every } q \in Q, \quad (21b)$$

which exists, by standard arguments, since the above problem is just (8a)-(8b) with $m(\cdot) = -a(\ell, \bar{u}, \cdot) - b(\cdot, p)$.

By setting $v = w$ in (20), we obtain

$$\begin{aligned} D_\ell J(\ell, w)(\delta\ell) &= \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\ell, \delta\bar{u}, \bar{u} - \bar{z}) + c(\delta p, p - \hat{z}) \\ &\quad + a(\delta\ell, \bar{u}, \bar{w}) + T(\ell, \delta u, w) \\ &= \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\ell, \delta\bar{u}, \bar{u} - \bar{z}) + c(\delta p, p - \hat{z}) \\ &\quad + a(\delta\ell, \bar{u}, \bar{w}) - a(\ell, \bar{u} - \bar{z}, \delta\bar{u}) - b(\delta\bar{u}, p - \hat{z}) \\ &= \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\delta\ell, \bar{u}, \bar{w}) + a(\ell, \delta\bar{u}, \bar{u}) + c(\delta p, p) \\ &\quad - b(\delta\bar{u}, p - \hat{z}) + a(\delta\ell, \bar{u}, \bar{w}) - a(\ell, \delta\bar{u}, \bar{u} - \bar{z}) \\ &= \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\delta\ell, \bar{u}, \bar{w}) \\ &\quad + c(\delta p, p - \hat{z}) - b(\delta\bar{u}, p - \hat{z}), \end{aligned}$$

where we have used the symmetry of the trilinear form T , $a(\ell, \cdot, \cdot)$, and (21a)-(21b). Since $c(\delta p, p - \hat{z}) = b(\delta\bar{u}, p - \hat{z})$, we obtain

$$D_\ell J(\ell, w)(\delta\ell) = \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\delta\ell, \bar{u}, \bar{w}).$$

Therefore, using (19), we have

$$DJ(\ell)(\delta\ell) = \frac{1}{2}a(\delta\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\delta\ell, \bar{u}, \bar{w}). \quad (22)$$

Summarizing, we have the following scheme to compute the derivative $DJ(\ell)$ given $a \in A$:

1. Compute u by (16).
2. Compute w by (21a)-(21b).
3. Compute $DJ(\ell)$ by (22).

Let us now develop the hybrid method for the computation of the second-order derivative. In the hybrid method proposed below, the derivative δu is computed directly while the computation of the second derivative $\delta^2 u$ is avoided by using an adjoint method. We will follow the same general scheme that was used above, but here we will use derivative formula (12a)-(12b).

Let $\delta\ell_2 \in A$ be a fixed direction. Then, for any $v = (\bar{v}, q) \in V$, we define

$$\begin{aligned} H(\ell, v) &= DJ(\ell)(\delta\ell_2) + a(\ell, D\bar{u}(\ell)(\delta\ell_2), \bar{v}) + b(\bar{v}, Dp(\ell)(\delta\ell_2)) \\ &\quad + b(D\bar{u}(\ell)(\delta\ell_2), q) - c(Dp(\ell)(\delta\ell_2), q) + a(\delta\ell_2, \bar{u}, \bar{v}) \\ &= \frac{1}{2}a(\delta\ell_2, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\ell, D\bar{u}(\ell)(\delta\ell_2), \bar{u} - \bar{z}) \\ &\quad + c(Dp(\ell)(\delta\ell_2), p - \hat{z}) + a(\ell, D\bar{u}(\ell)(\delta\ell_2), \bar{v}) + b(\bar{v}, Dp(\ell)(\delta\ell_2)) \\ &\quad + b(D\bar{u}(\ell)(\delta\ell_2), q) - c(Dp(\ell)(\delta\ell_2), q) + a(\delta\ell_2, \bar{u}, \bar{v}). \end{aligned}$$

By the construction of H , for every $v \in V$, we have

$$\frac{\partial H}{\partial \ell}(\ell, v)(\delta\ell_1) = D^2J(\ell)(\delta\ell_1, \delta\ell_2) \quad \text{for every } \delta\ell_1 \in A. \quad (23)$$

By a simple calculation, we have

$$\begin{aligned} \frac{\partial H}{\partial \ell}(\ell, v)(\delta\ell_1) &= a(\delta\ell_2, D\bar{u}(\ell)(\delta\ell_1), \bar{u} - \bar{z}) + a(\delta\ell_1, D\bar{u}(\ell)(\delta\ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D^2\bar{u}(\ell)(\delta\ell_1, \delta\ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D\bar{u}(\ell)(\delta\ell_2), D\bar{u}(\ell)(\delta\ell_1)) \\ &\quad + c(D^2p(\ell)(\delta\ell_1, \delta\ell_2), p - \hat{z}) + c(Dp(\ell)(\delta\ell_2), Dp(\ell)(\delta\ell_1)) \\ &\quad + a(\delta\ell_1, D\bar{u}(\ell)(\delta\ell_2), \bar{v}) + a(\ell, D^2\bar{u}(\ell)(\delta\ell_1, \delta\ell_2), \bar{v}) \\ &\quad + b(\bar{v}, D^2p(\ell)(\delta\ell_1, \delta\ell_2)) + b(D^2\bar{u}(\ell)(\delta\ell_1, \delta\ell_2), q) \\ &\quad - c(D^2p(\ell)(\delta\ell_1, \delta\ell_2), q) + a(\delta\ell_2, D\bar{u}(\ell)(\delta\ell_1), \bar{v}). \end{aligned} \quad (24)$$

Let $w(\ell) = (\bar{w}(\ell), p_w(\ell))$ be the unique solution of the saddle point problem (cf. (12a)-(12b)):

$$a(\ell, \bar{w}, \bar{v}) + b(\bar{v}, p_w) = a(\ell, \bar{v}, \bar{z} - \bar{u}) + b(\bar{v}, \hat{z} - p) \quad \text{for every } \bar{v} \in \widehat{V}, \quad (25a)$$

$$b(\bar{w}, q) - c(p_w, q) = 0 \quad \text{for every } q \in Q. \quad (25b)$$

By setting $v = w$ in (24), we have

$$\begin{aligned} \frac{\partial H}{\partial \ell}(\ell, w)(\delta \ell_1) &= a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{u} - \bar{z}) + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), \bar{u} - \bar{z}) + a(\ell, D\bar{u}(\ell)(\delta \ell_2), D\bar{u}(\ell)(\delta \ell_1)) \\ &\quad + c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), p - \hat{z}) + c(Dp(\ell)(\delta \ell_2), Dp(\ell)(\delta \ell_1)) \\ &\quad + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{w}) + a(\ell, D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), \bar{w}) \\ &\quad + b(\bar{w}, D^2p(\ell)(\delta \ell_1, \delta \ell_2)) + b(D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), q) \\ &\quad - c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), q) + a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{w}) \\ &= a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{u} - \bar{z}) + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D\bar{u}(\ell)(\delta \ell_2), D\bar{u}(\ell)(\delta \ell_1)) + c(Dp(\ell)(\delta \ell_2), Dp(\ell)(\delta \ell_1)) \\ &\quad + a(\ell, D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), \bar{u} - \bar{z}) + c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), p - \hat{z}) \\ &\quad + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{w}) + a(\ell, D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), \bar{w}) \\ &\quad + b(\bar{w}, D^2p(\ell)(\delta \ell_1, \delta \ell_2)) + b(D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), q) \\ &\quad - c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), q) + a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{w}) \\ &= a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{u} - \bar{z}) + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D\bar{u}(\ell)(\delta \ell_2), D\bar{u}(\ell)(\delta \ell_1)) + c(Dp(\ell)(\delta \ell_2), Dp(\ell)(\delta \ell_1)) \\ &\quad + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{w}) + a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{w}) \\ &\quad + c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), p - \hat{z}) + b(D^2\bar{u}(\ell)(\delta \ell_1, \delta \ell_2), \hat{z} - p). \end{aligned}$$

Recall that by derivative formula (13a)-(13b), we have

$$c(D^2p(\ell)(\delta \ell_1, \delta \ell_2), p - \hat{z}) = b(D^2p(\ell)(\delta \ell_1, \delta \ell_2), p - \hat{z}),$$

which implies

$$\begin{aligned} \frac{\partial H}{\partial \ell}(\ell, w)(\delta \ell_1) &= a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{u} - \bar{z}) + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D\bar{u}(\ell)(\delta \ell_2), D\bar{u}(\ell)(\delta \ell_1)) + c(Dp(\ell)(\delta \ell_2), Dp(\ell)(\delta \ell_1)) \\ &\quad + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{w}) + a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{w}). \end{aligned}$$

Consequently, from (23), we get

$$\begin{aligned} D^2J(\ell)(\delta \ell_1, \delta \ell_2) &= a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{u} - \bar{z}) + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{u} - \bar{z}) \\ &\quad + a(\ell, D\bar{u}(\ell)(\delta \ell_2), D\bar{u}(\ell)(\delta \ell_1)) + c(Dp(\ell)(\delta \ell_2), Dp(\ell)(\delta \ell_1)) \\ &\quad + a(\delta \ell_1, D\bar{u}(\ell)(\delta \ell_2), \bar{w}) + a(\delta \ell_2, D\bar{u}(\ell)(\delta \ell_1), \bar{w}) \end{aligned}$$

and, in particular,

$$D^2J(\ell)(\delta \ell, \delta \ell) = 2a(\delta \ell, \delta \bar{u}, \bar{u} - \bar{z}) + a(\ell, \delta \bar{u}, \delta \bar{u}) + c(\delta p, \delta p) + 2a(\delta \ell, \delta \bar{w}, \bar{w}). \quad (26)$$

Summarizing, we propose the following scheme to compute the derivative $D^2J(\ell)(\delta\ell, \delta\ell)$ given $\ell \in A$, $\delta\ell \in A$.

1. Compute $u(\ell) = (\bar{u}(\ell), p)$ by (16).
2. Compute $\delta u = (\delta\bar{u}, \delta p)$ by (12a)-(12b).
3. Compute $w(\ell) = (\bar{w}(\ell), q(\ell))$ by (25a)-(25b).
4. Compute $D^2J(\ell)(\delta\ell, \delta\ell)$ by (26).

4 Discretization formulas for the adjoint and the hybrid method

In this section, we collect discrete formulae for saddle point problem (8a)-(8b) and the associated inverse problem. We begin, therefore, with a triangulation \mathcal{T}_h on Ω , L_h is the space of all piecewise continuous polynomials of degree d_ℓ relative to \mathcal{T}_h , U_h is the space of all piecewise continuous polynomials of degree d_u relative to \mathcal{T}_h , and Q_h is the space of all piecewise continuous polynomials of degree d_q relative to \mathcal{T}_h .

In order to represent the discrete saddle point problem in a computable form, we proceed as follows. We represent bases for L_h , U_h , and Q_h by $\{\varphi_1, \varphi_2, \dots, \varphi_m\}$, $\{\psi_1, \psi_2, \dots, \psi_n\}$, and $\{\chi_1, \chi_2, \dots, \chi_k\}$, respectively. The space L_h is then isomorphic to \mathbb{R}^m and for any $\ell \in L_h$, we define $L \in \mathbb{R}^m$ by $L_i = \ell(x_i)$ for $i = 1, 2, \dots, m$, where the nodal basis $\{\varphi_1, \varphi_2, \dots, \varphi_m\}$ corresponds to the nodes $\{x_1, x_2, \dots, x_m\}$. Conversely, each $L \in \mathbb{R}^m$ corresponds to $\ell \in L_h$ defined by $\ell = \sum_{i=1}^m L_i \varphi_i$. Similarly, $u \in U_h$ will correspond to $U \in \mathbb{R}^n$, where $\bar{U}_i = u(y_i)$, $i = 1, 2, \dots, n$, and $u = \sum_{i=1}^n \bar{U}_i \psi_i$, where y_1, y_2, \dots, y_n are the nodes of the mesh defining U_h . Finally, $q \in Q_h$ will correspond to $Q \in \mathbb{R}^k$, where $Q_i = q(z_i)$, $i = 1, 2, \dots, k$, and $q = \sum_{i=1}^k Q_i \chi_i$, where z_1, z_2, \dots, z_k are the nodes of the mesh defining Q_h . (The spaces A_h , U_h , and Q_h are defined relative to the same elements, but the nodes will be different if $d_\ell \neq d_u \neq d_q$.)

Recall that the discrete saddle point problem seeks, for each ℓ_h , the unique $(\bar{u}_h, p_h) \in V_h \times Q_h$ with

$$a(\ell_h, \bar{u}_h, \bar{v}) + b(\bar{v}, p_h) = m(\bar{v}) \quad \text{for every } \bar{v} \in U_h, \quad (27a)$$

$$b(\bar{u}_h, q) - c(p_h, q) = 0 \quad \text{for every } q \in Q_h. \quad (27b)$$

We define $S: \mathbb{R}^m \rightarrow \mathbb{R}^{n+k}$ to be the finite element solution operator that assigns to each coefficient $\ell_h \in A_h$ the unique approximate solution $u_h = (\bar{u}_h, p_h) \in U_h \times Q_h$. Then $S(L) = U$, where U is defined by

$$K(L)U = F, \quad (28)$$

and where the stiffness matrix $K(L) \in \mathbb{R}^{(n+k) \times (n+k)}$ and the load vector $F \in \mathbb{R}^{n+k}$ are given by

$$K(L) = \begin{bmatrix} \hat{K}(L) & B^T \\ B & -C \end{bmatrix}$$

with

$$\hat{K}(L)_{ij} = a(\ell, \psi_j, \psi_i), \quad i, j = 1, 2, \dots, n,$$

$$B_{ij} = b(\psi_j, \chi_i), \quad i = 1, 2, \dots, k, n = 1, 2, \dots, n,$$

$$\begin{aligned}C_{ij} &= c(\chi_j, \chi_i), \quad i, j = 1, 2, \dots, k, \\F_i &= m(\psi_i), \quad i = 1, 2, \dots, n, \\F_j &= 0, \quad j = n+1, n+2, \dots, n+k.\end{aligned}$$

For future reference, it will be useful to note that

$$\widehat{K}(L)_{ij} = T_{ijk} L_k,$$

where the summation convention is used and T is the tensor defined by

$$T_{ijk} = a(\varphi_k, \psi_i, \psi_j) \quad \text{for every } i, j = 1, \dots, n, k = 1, \dots, m.$$

Let us now compute the discrete analogue of energy least-squares objective functional. By using the above notations, the discrete form of

$$J(L) = \frac{1}{2} a(\ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + \frac{1}{2} c(p - \hat{z}, p - \hat{z})$$

is given by

$$J(L) = \frac{1}{2} (\bar{U}(L) - \bar{Z})^T \widehat{K}(L) (\bar{U}(L) - \bar{Z}) + \frac{1}{2} (P(L) - \widehat{Z})^T C (P(L) - \widehat{Z}).$$

In order to get an operative expression for the gradient, we need to consider the so-called adjoint stiffness matrix \mathbb{A} defined by the following condition:

$$\widehat{K}(L) \bar{V} = \mathbb{A}(\bar{V}) L \quad \text{for every } L \in \mathbb{R}^m, \text{ for every } \bar{V} \in \mathbb{R}^n. \quad (29)$$

4.1 Computation of the gradient by using the adjoint method

Using the above notation, we have the following discrete adjoint method for the computation of gradient of $J(\cdot)$.

1. We compute $U = \begin{bmatrix} \bar{U}(L) \\ P(L) \end{bmatrix}$ by solving the linear system

$$\begin{bmatrix} \widehat{K}(L) & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} \bar{U}(L) \\ P(L) \end{bmatrix} = \begin{bmatrix} F \\ 0 \end{bmatrix}. \quad (30)$$

2. We compute $W = \begin{bmatrix} \bar{W}(L) \\ P_W(L) \end{bmatrix}$ by solving the linear system

$$\begin{bmatrix} \widehat{K}(L) & B^T \\ B & -C \end{bmatrix} \begin{bmatrix} \bar{W}(L) \\ P_W(L) \end{bmatrix} = \begin{bmatrix} -\widehat{K}(L)(\bar{U} - \bar{Z}) - B^T(P - \widehat{P}) \\ 0 \end{bmatrix}. \quad (31)$$

3. The gradient $\nabla J(L)$ can be calculated by using the adjoint stiffness matrix. From (22), we have

$$DJ(\ell)(\delta \ell) = \frac{1}{2} a(\delta \ell, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + a(\delta \ell, \bar{u}, \bar{w}), \quad (32)$$

a direct discretization gives the following:

$$\begin{aligned}\nabla J(L)(\delta L) &= \frac{1}{2}(\bar{U} - \bar{Z})^T \widehat{K}(\delta L)(\bar{U} - \bar{Z}) + \bar{U}^T \widehat{K}(\delta L) \bar{W} \\ &= \frac{1}{2} \bar{U}^T \mathbb{A}(\bar{U} - \bar{Z}) \delta L + \bar{U}^T \mathbb{A}(\bar{W}(L)) \delta L,\end{aligned}$$

and therefore the gradient $\nabla J(L)$ is given by

$$\nabla J(L) = \frac{1}{2}(\bar{U} - \bar{Z})^T \mathbb{A}(\bar{U} - \bar{Z}) + \bar{U}^T \mathbb{A}(\bar{W}(L)). \quad (33)$$

4.2 Computation of the Hessian by using a hybrid method

Recall that we have established the following:

$$D_{\ell\ell}^2 J(\ell)(\delta \ell, \delta \ell) = 2a(\delta \ell, \delta \bar{u}, \bar{u} - \bar{z}) + a(\ell, \delta \bar{u}, \delta \bar{u}) + c(\delta p, \delta p) + 2a(\delta \ell, \delta \bar{u}, \bar{w}). \quad (34)$$

By the standard discretization scheme, we have

1. $a(\delta L, \delta \bar{U}, \bar{U} - \bar{Z}) = \delta L^T \nabla \bar{U}^T \widehat{K}(\delta L)(\bar{U} - \bar{Z}) = \delta L^T \nabla \bar{U}^T \mathbb{A}(\bar{U} - \bar{Z}) \delta L,$
2. $a(L, \delta \bar{U}, \delta \bar{U}) = \delta L^T \nabla \bar{U}^T \widehat{K}(L) \nabla \bar{U} \delta L = \delta L^T \nabla \bar{U}^T \widehat{K}(L) \nabla \bar{U} \delta L,$
3. $c(\delta P, \delta P) = \delta L^T \nabla P^T C \nabla P \delta L,$
4. $a(\delta \ell, \delta \bar{u}, \bar{w}) = \delta L^T \nabla \bar{U}^T \widehat{K}(\delta L) \bar{W} = \delta L^T \nabla \bar{U}^T \mathbb{A}(\bar{W}) \delta L.$

Consequently, we have the following explicit formula for the Hessian:

$$\nabla^2 J(L) = 2^T \nabla \bar{U}^T \mathbb{A}(\bar{U} - \bar{Z}) + \nabla \bar{U}^T \widehat{K}(L) \nabla \bar{U} + \nabla P^T C \nabla P + 2 \nabla \bar{U}^T \mathbb{A}(\bar{W}). \quad (35)$$

Summarizing, we have the following scheme for the computation of the second derivative of the EOLS:

1. Compute $U = (\bar{U}, P)$ by solving linear system (30).
2. Compute $W = (\bar{W}, P)$ by solving linear system (31).
3. Compute $\nabla U = (\nabla \bar{U}, \nabla P)$ by solving m linear systems.
4. Compute $\nabla^2 J(L)$ by using formula (35).

We note that to compute the Hessian using the hybrid method requires the solution of $m + 2$ linear systems.

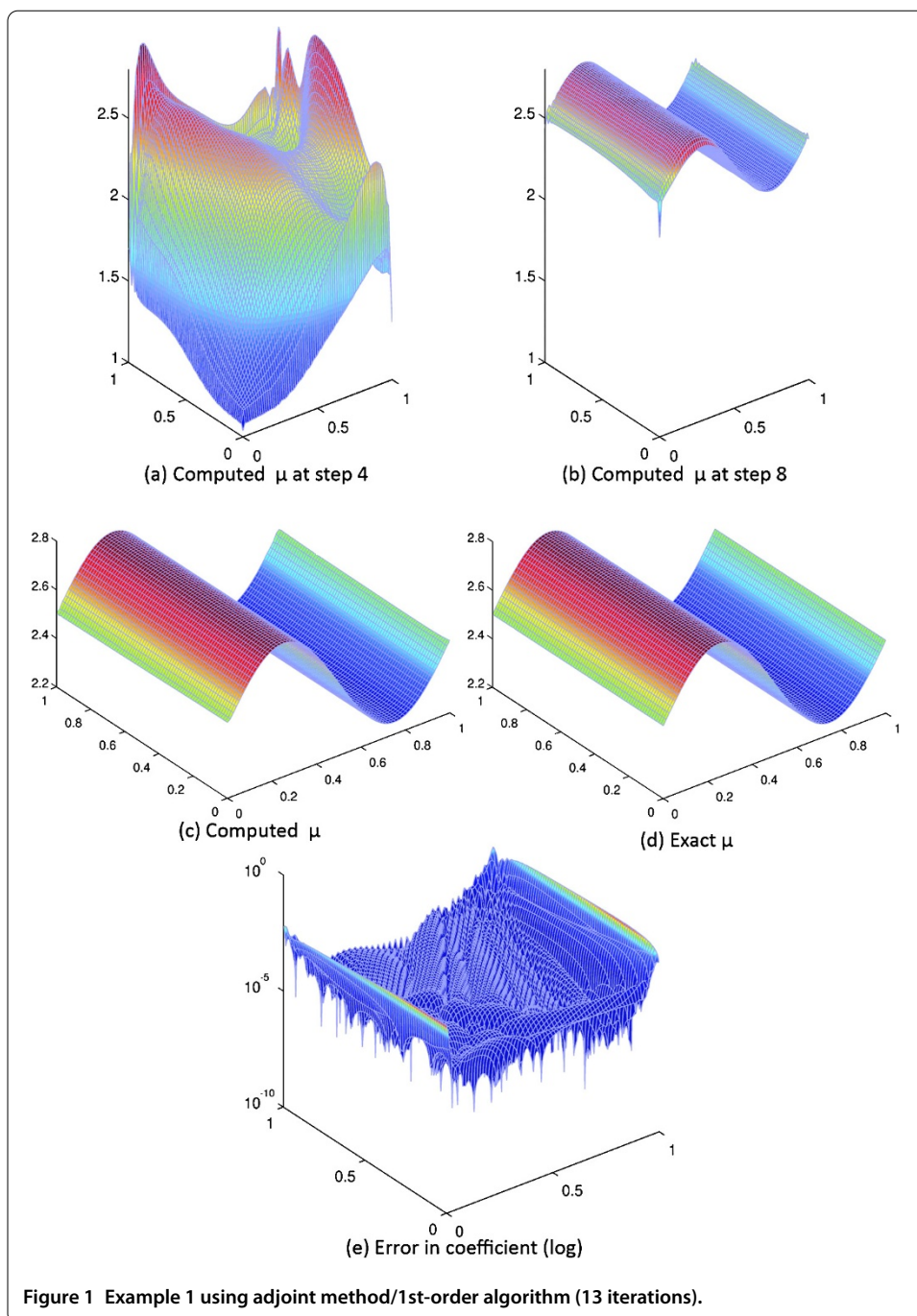
5 Numerical experiments

We consider here two representative examples of elastography inverse problems for the recovery of a variable μ on a two-dimensional isotropic domain $\Omega = (0, 1) \times (0, 1)$ with boundary $\partial\Omega = \Gamma_1 \cup \Gamma_2$. In the first example, a smooth coefficient is recovered using both the adjoint and hybrid gradient calculation methods. For the second example, we examine the recovery of a discontinuous coefficient using the adjoint method.

All examples are solved on a 75×75 quadrangular mesh with 5,476 quadrangles and 16,576 total degrees of freedom. Example 1 uses a smooth Tikhonov-type regularization method, whereas the discontinuities in Example 2 necessitate the use of a BV-regularization scheme (see [23] for a more thorough discussion of regularization).

5.1 Example 1

In this example we consider the recovery of a smooth coefficient in which the left and right domain boundaries (Γ_1) are fixed with static condition $g(x, y)$ and the top and bottom



boundaries have Neumann condition $h(x, y)$. The functions defining the coefficient, load, and boundary conditions are as follows:

$$\mu(x, y) = 2.5 + \frac{1}{4} \sin(2\pi x), \quad f(x, y) = \begin{bmatrix} 2.3 + \frac{1}{10}x \\ 2.3 + \frac{1}{10}y \end{bmatrix},$$

$$g(x, y) = \frac{1}{100} \begin{bmatrix} x \\ y^2 \end{bmatrix} \quad \text{on } \Gamma_1, \quad h(x, y) = \frac{1}{2} \begin{bmatrix} 1 + 2x^2 \\ 1 + 2y^2 \end{bmatrix} \quad \text{on } \Gamma_2.$$

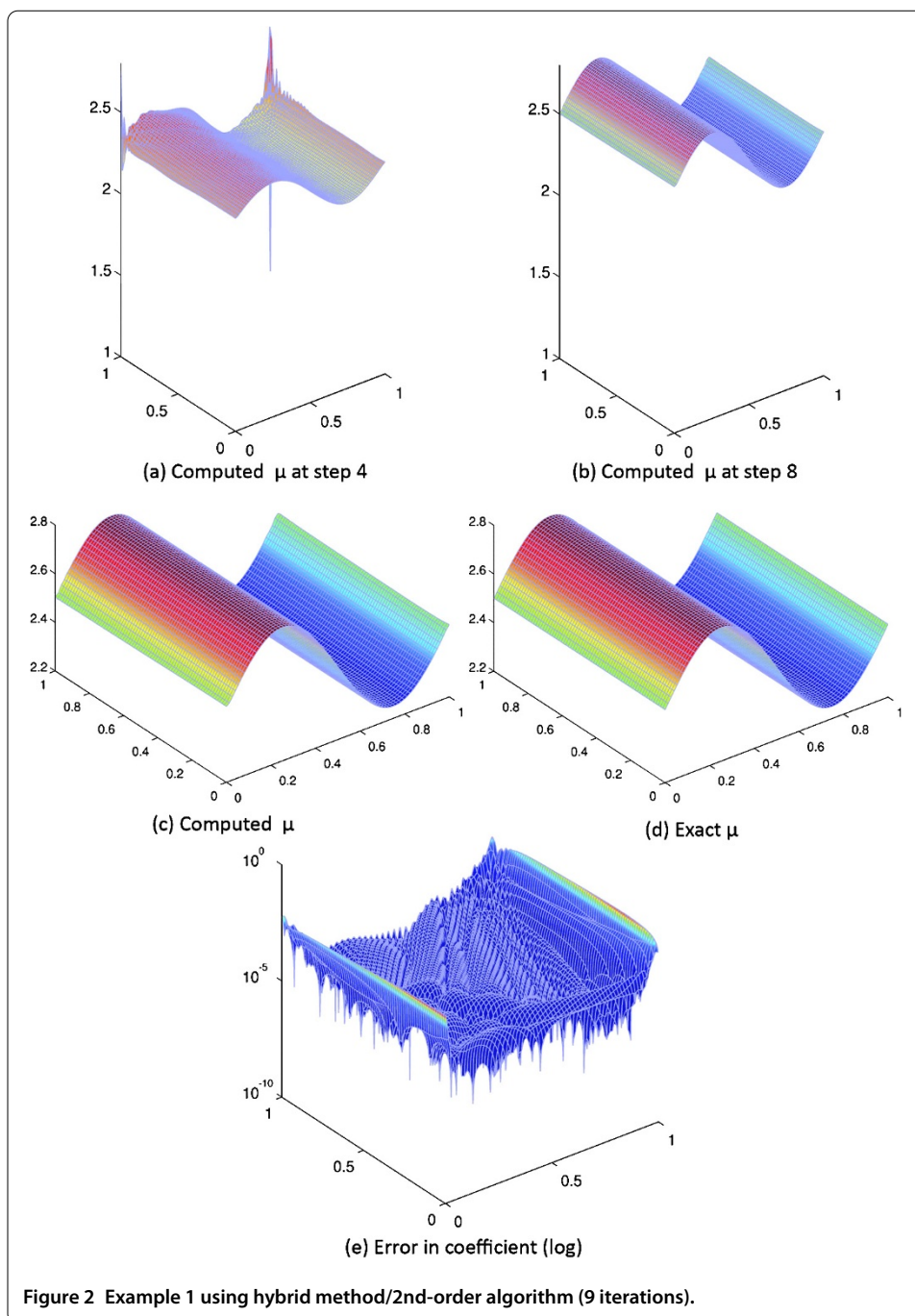


Figure 2 Example 1 using hybrid method/2nd-order algorithm (9 iterations).

For this example, the underlying optimization problem was solved using both a first-order Newton-CG-Trust Region algorithm as well as a second-order quasi-Newton method, using the adjoint and hybrid gradient calculations outlined in the preceding sections, respectively. Comparatively, the hybrid method converges faster to the solution in only 9 algorithm iterations compared to 13 iterations for the adjoint method when both are started from the same initial point and under the same stopping criteria ($\nabla J < 10^{-10}$). This can be seen qualitatively in Figures 1 and 2 through the comparison of the computed μ at selected intermediary algorithm steps (subfigures (a) and (b)).

5.2 Example 2

For the discontinuous example, the top of the region is taken as Γ_1 and fixed with (constant) Dirichlet condition $g(x, y)$. The remaining edges of the region are taken as Γ_2 with Neumann condition $h(x, y)$. The functions defining the coefficient, load, and boundary conditions are as follows:

$$\mu(x, y) = \begin{cases} 0.3 & \text{for } (x, y) \in R_1, \\ 0.5 & \text{for } (x, y) \in R_2, \\ 0.2 & \text{otherwise,} \end{cases} \quad f(x, y) = \begin{bmatrix} 2.0 + \frac{x}{10} \\ 2.0 + \frac{y}{10} \end{bmatrix},$$

$$g(x, y) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{on } \Gamma_1, \quad h(x, y) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{on } \Gamma_2,$$

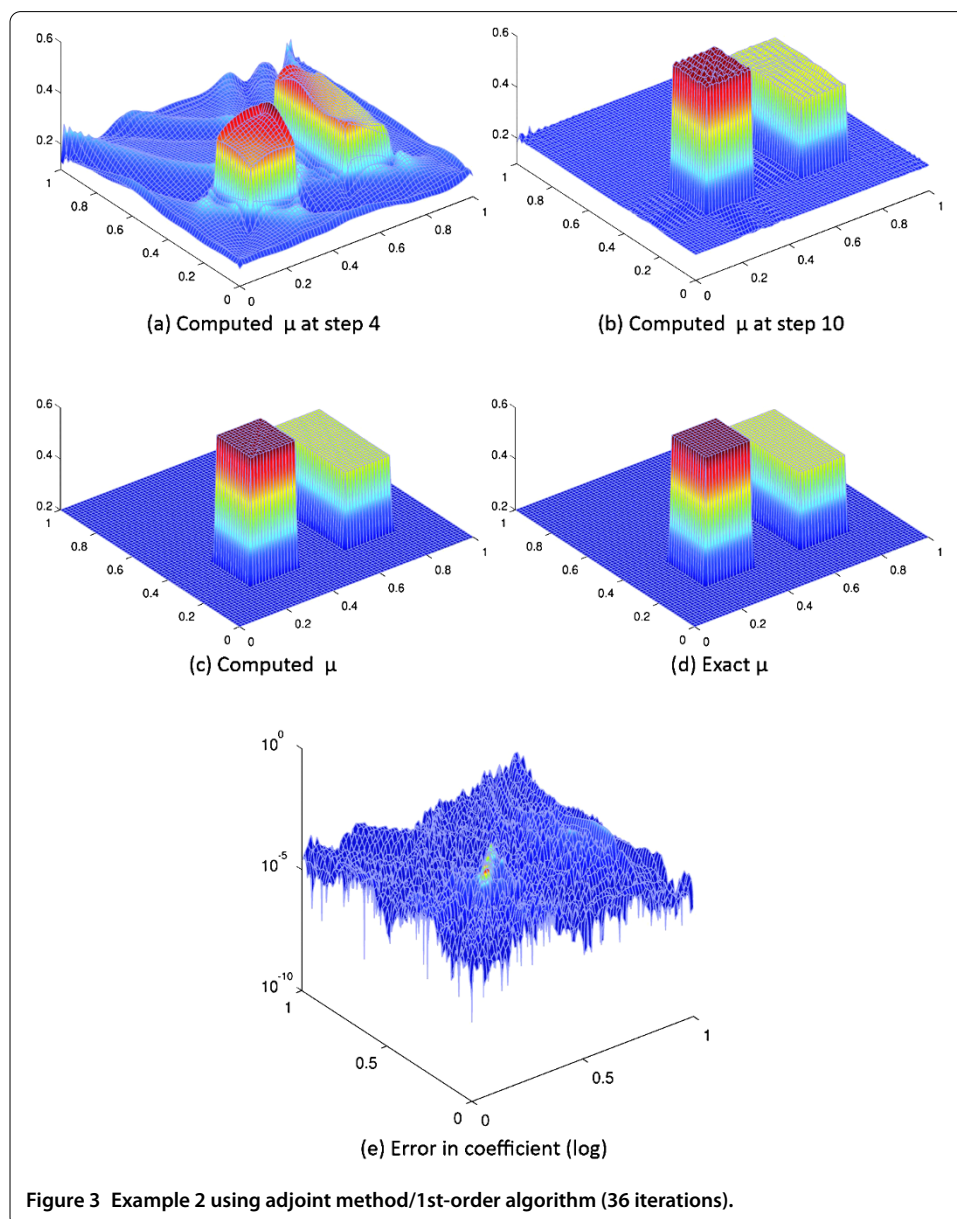


Figure 3 Example 2 using adjoint method/1st-order algorithm (36 iterations).

where $R_1 = \{(x, y) : 0.6 \leq x \leq 0.8, 0.2 \leq y \leq 0.6\}$ and $R_2 = \{(x, y) : 0.2 \leq x \leq 0.4, 0.2 \leq y \leq 0.4\}$.

6 Concluding remarks

In this work we have presented a detailed application of the adjoint method for efficiently computing the gradient of the energy output least-squares functional as well as a hybrid method for calculating the functional's second derivative. We have also provided two numerical examples of elastography inverse problems to demonstrate the overall feasibility of implementation and establish the relative effectiveness of these methods when coupled with the appropriate first-order and second-order optimization algorithms. See Figure 3.

One issue not addressed in depth was the comparative performance of these methods, measured both against existing schemes and against one other. In short, we note that the hybrid method requires the solution of $m + 2$ linear systems with m scaling along with the size of the mesh. However, the m systems remain entirely independent, allowing for the parallelization of parts of the computation and thus granting significant performance gains and potential advantages over other strategies. In a future work, we look to extend our study here into just such a thorough analysis and carefully consider the performance of the adjoint and hybrid derivative computation methods.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

This research was carried out during Prof. Miguel Sama's visit at RIT and all the work was done at that time in a collaborative manner. All authors read and approved the final manuscript.

Author details

¹Center for Applied and Computational Mathematics, School of Mathematical Sciences, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, NY 14623, USA. ²Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Calle Juan del Rosal, 12, Madrid, 28040, Spain.

Acknowledgements

The work of AA Khan is supported by RIT's COS D-RIG Acceleration Research Funding Program 2012-2013 and a grant from the Simons Foundation (#210443 to Akhtar Khan). The work of M Sama is partially supported by Ministerio de Ciencia (Spain), project (MTM2012-30942).

Received: 14 August 2013 Accepted: 5 November 2013 Published: 02 Dec 2013

References

- Dooley, MM: Model-based elastography: a survey of approaches to the inverse elasticity problem. *Phys. Med. Biol.* **57**, R35 (2012). doi:10.1088/0031-9155/57/3/R35
- Raghavan, KR, Yagle, AE: Forward and inverse problems in elasticity imaging of soft tissues. *IEEE Trans. Nucl. Sci.* **41**, 1639-1648 (1994)
- Aguilo, MA, Aquino, W, Brigham, JC, Fatemi, M: An inverse problem approach for elasticity imaging through vibroacoustics. *IEEE Trans. Med. Imaging* **29**, 1012-1021 (2010)
- Ammari, H, Garapon, P, Jouve, F: Separation of scales in elasticity imaging: a numerical study. *J. Comput. Math.* **28**, 354-370 (2010)
- Arnold, A, Reichling, S, Bruhns, O, Mosler, J: Efficient computation of the elastography inverse problem by combining variational mesh adaption and clustering technique. *Phys. Med. Biol.* **55**, 2035-2056 (2010)
- Beretta, E, Bonnetier, E, Francini, E, Mazzucato, A: Small volume asymptotics for anisotropic elastic inclusions. *Inverse Probl. Imaging* **6**, 1-23 (2012)
- Ji, L, McLaughlin, J: Recovery of Lamé parameter μ in biological tissues. *Inverse Probl.* **20**, 1-24 (2004)
- Kallel, F, Bertrand, M: Tissue elasticity reconstruction using linear perturbation method. *IEEE Trans. Med. Imaging* **15**, 299-313 (1996)
- Barbone, PE, Bamber, JC: Quantitative elasticity imaging: what can and cannot be inferred from strain images. *Phys. Med. Biol.* **47**, 2147-2164 (2002)
- Barbone, PE, Gokhale, NH: Elastic modulus imaging: on the uniqueness and nonuniqueness of the elastography inverse problem in two dimensions. *Inverse Probl.* **20**, 283-296 (2004)
- Braess, D: *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, 3rd edn. Cambridge University Press, Cambridge (2007)

12. Chan, TF, Tai, XC: Identification of discontinuous coefficients in elliptic problems using total variation regularization. *SIAM J. Sci. Comput.* **25**, 881-904 (2003)
13. Gockenbach, MS, Khan, AA: Identification of Lamé parameters in linear elasticity: a fixed point approach. *J. Ind. Manag. Optim.* **1**, 487-497 (2005)
14. Gockenbach, MS, Jadamba, B, Khan, AA: Numerical estimation of discontinuous coefficients by the method of equation error. *Int. J. Math. Comput. Sci.* **1**, 343-359 (2006)
15. Gockenbach, MS, Jadamba, B, Khan, AA: Equation error approach for elliptic inverse problems with an application to the identification of Lamé parameters. *Inverse Probl. Sci. Eng.* **16**, 349-367 (2008)
16. Harrigan, T, Konofagou, EE: Estimation of material elastic moduli in elastography: a local method, and an investigation of Poisson ratio sensitivity. *J. Biomech.* **37**, 1215-1221 (2004)
17. Jadamba, B, Khan, AA, Raciti, F: On the inverse problem of identifying Lamé coefficients in linear elasticity. *Comput. Math. Appl.* **56**, 431-443 (2008)
18. Jadamba, B, Khan, AA, Sama, M: Inverse problems of parameter identification in partial differential equations. In: *Mathematics in Science and Technology*, pp. 228-258. World Scientific, Hackensack (2011)
19. Konofagou, E, Harrigan, T, Ophir, J, Krouskop, T: Poroelastography: estimation and imaging of the poroelastic properties of tissues. In: *IEEE Proceedings of the Symposium in Ultrasonics, Ferroelectrics and Frequency Control*, Lake Tahoe, NV, pp. 1627-1630 (1999)
20. McLaughlin, J, Yoon, JR: Unique identifiability of elastic parameters from time-dependent interior displacement measurement. *Inverse Probl.* **20**, 25-45 (2004)
21. Mehrabian, H, Campbell, G, Samani, A: A constrained reconstruction technique of hyperelasticity parameters for breast cancer assessment. *Phys. Med. Biol.* **53**, 7489-7508 (2012)
22. Brezzi, F, Fortin, M: *Mixed and Hybrid Finite Element Methods*. Springer, New York (1991)
23. Doyley, MM, Jadamba, B, Khan, AA, Sama, M, Winkler, B: A new energy inversion for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location (2013, submitted)
24. Oberai, AA, Gokhale, NH, Feijóo, GR: Solution of inverse problems in elasticity imaging using the adjoint method. *Inverse Probl.* **19**, 297-313 (2003)
25. Tortorelli, DA, Michaleris, P: Design sensitivity analysis: overview and review. *Inverse Probl. Eng.* **1**, 71-105 (1994)
26. Cioaca, A, Alexea, M, Sandua, A: Second-order adjoints for solving PDE-constrained optimization problems. *Optim. Methods Softw.* **27**, 625-653 (2012)
27. Goeleven, D, Motreanu, D: *Variational and Hemivariational Inequalities - Theory, Methods and Applications*, vol. 2. Springer, Berlin (2003)
28. Bush, N, Jadamba, B, Khan, AA, Raciti, F: Identification of a parameter in fourth-order partial differential equations by an equation error approach (2014, to appear)
29. Crossen, E, Gockenbach, MS, Jadamba, B, Khan, AA, Winkler, B: An equation error approach for the elasticity imaging inverse problem for predicting tumor location. *Comput. Math. Appl.* (2013, to appear)
30. Gockenbach, MS, Khan, AA: An abstract framework for elliptic inverse problems. Part 1: an output least-squares approach. *Math. Mech. Solids* **12**, 259-276 (2007)
31. Gockenbach, MS, Khan, AA: An abstract framework for elliptic inverse problems. Part 2: an augmented Lagrangian approach. *Math. Mech. Solids* **14**, 517-539 (2009)

10.1186/1687-2770-2013-263

Cite this article as: Cahill et al.: A first-order adjoint and a second-order hybrid method for an energy output least-squares elastography inverse problem of identifying tumor location. *Boundary Value Problems* 2013, **2013**:263

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com