J. Nonlinear Var. Anal. 4 (2020), No. 2, pp. 301-318
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.4.2020.2.10

THREE OPTIMIZATION FORMULATIONS FOR AN INVERSE PROBLEM IN SADDLE POINT PROBLEMS WITH APPLICATIONS TO ELASTICITY IMAGING OF LOCATING TUMOR IN INCOMPRESSIBLE MEDIUM

O. BABANIYI[1], B. JADAMBA[1,*], A. A. KHAN[1], M. RICHARDS[2], M. SAMA[3], C. TAMMER[4]

[1]*School of Mathematical Sciences, Rochester Institute of Technology, Rochester, New York, 14623, USA*
[2]*Department of Biomedical Engineering, Kate Gleason College of Engineering, Rochester Institute of Technology, Rochester, New York, 14623, USA*
[3]*Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Calle Juan del Rosal, 12, 28040 Madrid, Spain*
[4]*Institute of Mathematics, Martin-Luther-University of Halle-Wittenberg, Theodor-Lieser-Str. 5, D-06120 Halle-Saale, Germany*

Dedicated to our dear friend Juan Enrique Martínez Legaz on the occasion of his 70th birthday

Abstract. This work focuses on identifying a distributed parameter in a saddle point problem with application to the elasticity imaging inverse problem. We examine three optimization formulations for the inverse problem, namely, the output least-squares (OLS), the modified output least-squares (MOLS), and the energy output least-squares (EOLS). The OLS functional and the EOLS functional are, in general, nonconvex; however, we show that the MOLS functional is convex. We provide existence results for optimization problems involving the regularized variants of the OLS, the EOLS, and the MOLS functional. We give first-order and second-order adjoint methods in the continuous setting to compute the first-order and the second-order derivative of the OLS/EOLS functionals. The derivative of the MOLS objective does not involve the derivative of the solution map and hence does not require the adjoint approach. We provide numerical experimentation on tissue phantom data.

Keywords. Saddle point problems; Parameter identification: Regularization; Output least-squares; Elasticity imaging.

1. INTRODUCTION

Soft tissue cancers are among the deadliest forms of the disease, and early detection is crucial for a successful treatment. A change in soft tissue's stiffness is attributed to tissue health, making palpation an ancient practice of detecting tumors, which is limited to finding hard nodules near the skin's surface. Elasticity imaging inverse problem is a quantitative substitute for palpation and uses the difference in elastic properties of healthy and diseased tissue to locate

*Corresponding author.
E-mail addresses: obsma@rit.edu (O. Babaniyi), bxjsma@rit.edu (B. Jadamba), aaksma@rit.edu (A. A. Khan), msrbme@rit.edu (M. Richards), msama@ind.uned.es (M. Sama), christiane.tammer@mathematik.uni-halle.de (C. Tammer).

Received October 21, 2019; Accepted July 10, 2020.

©2020 Journal of Nonlinear and Variational Analysis

tumors. The following isotropic linear elasticity system provides the necessary mathematical tools for the elasticity imaging inverse problem:

$$-\nabla\cdot\sigma = f \text{ in } \Omega, \tag{1.1a}$$

$$\sigma = 2\mu\varepsilon(u) + \lambda \operatorname{div} u I, \tag{1.1b}$$

$$u = g \text{ on } \Gamma_1, \tag{1.1c}$$

$$\sigma n = h \text{ on } \Gamma_2. \tag{1.1d}$$

Here the domain Ω is a sufficiently smooth subset of $\mathbb{R}^2$ or $\mathbb{R}^3$ with boundary $\partial\Omega = \Gamma_1 \cup \Gamma_2$. The vector-valued function $u = u(x)$ is the displacement of the elastic object, f is the applied body force, n is the unit outward normal, and $\varepsilon(u) = \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}})$ is the linearized strain tensor. The (spatially varying) Lamé parameters μ and λ quantify the elastic properties of the material.

The elasticity imaging inverse problem locates likely tumors by identifying the elasticity parameters from a clinical measurement of the displacement u. For measuring the displacement, one applies a sufficiently small external quasistatic compression force to the tissue and then measures the tissue's axial displacement field either directly or through the comparison of an undeformed and deformed image.

The primary goal of this study is to develop an abstract framework for the elasticity imaging inverse problem by treating the underlying elastic object as incompressible. The incompressibility can be understood using the relationship $\lambda := \frac{2\nu\mu}{1-2\nu}$, where ν is the Poisson's ratio. If $\nu \approx 0.5$, then λ is large, and the elastic object is called nearly incompressible. If $\nu \to \frac{1}{2}$, the elastic object is called fully incompressible. For fully incompressible case, (1.1) is not valid, and an alternative formulation is derived via an explicit incompressibility constraint, see [18].

Taking $Q = L^2(\Omega)$, $V = \{v = (v_1, v_2) \in H^1(\Omega) \times H^1(\Omega) : v = 0 \text{ on } \Gamma_1\}$, we recall that if the body is nearly incompressible ($\lambda \gg \mu$), then the variational form of (1.1) ($g = 0$ for simplicity) reads: Find $(u, p) \in V \times Q$ such that reads

$$\int_\Omega 2\mu\varepsilon(u)\cdot\varepsilon(v) + \int_\Omega p(\operatorname{div} v) = \int_\Omega fv + \int_{\Gamma_2} vh, \quad \text{for every } v \in V, \tag{1.2a}$$

$$\int_\Omega (\operatorname{div} u)q - \int_\Omega \frac{1}{\lambda} pq = 0, \quad \text{for every } q \in Q. \tag{1.2b}$$

Although (1.1) is not valid for the fully incompressible case ($\lambda \to \infty$), an alternate penalty formulation (see [18]) leads to the variational problem of finding $(u, p) \in V \times Q$ such that

$$\int_\Omega 2\mu\varepsilon(u)\cdot\varepsilon(v) + \int_\Omega p(\operatorname{div} v) = \int_\Omega fv + \int_{\Gamma_2} vh, \quad \text{for every } v \in V, \tag{1.3a}$$

$$\int_\Omega (\operatorname{div} u)q = 0, \quad \text{for every } q \in Q. \tag{1.3b}$$

The elasticity imaging inverse problems is often studied in the nearly incompressible medium under the simplifying assumption that λ is a large constant. However, without such modeling assumption, in (1.2), both μ and λ are space dependent. This formulation, yet, results in non-trivial technical obstacles when computing the derivatives of the solution map. Although (1.3) also has drawbacks, it seems more suitable for developing a rigorous mathematical framework for the elasticity imaging inverse problem. Some results related to the elasticity imaging inverse problem can be found in [1, 2, 3, 4, 7, 8, 16, 17, 21, 24, 28] and the references cited therein.

1.1. **Main contributions.** Although our motivation stems from the elasticity imaging inverse problem for a fully incompressible elastic medium, we conduct this study in the context of a general saddle point problem. Therefore, the given results are not limited to this specific application and can be applied to numerous other models that lead to saddle point problems.

The fundamental approach and the main contributions of this work are as follows:

(1) Inverse problems are most commonly studied using optimization tools that offer a convenient way of incorporating regularization. Following this trend, we examine three optimization formulations to investigate the inverse problem, namely, the output least-squares (OLS), the modified output least-squares (MOLS), and the energy output least-squares (EOLS). The OLS formulation's underlying principle is to minimize the gap between the computed solution and the measured data. In contrast, the philosophy of the MOLS/EOLS objectives is to find an optimal solution by minimizing the energy associated with the saddle point problem. The OLS functional and the EOLS functional are, in general, nonconvex; however, we show that the MOLS functional is convex.
(2) We give existence results for optimization problems involving the regularized variants of the OLS, the EOLS, and the MOLS functional. We use a quadratic regularization to counter the adverse effects of the ill-posedness of the inverse problem at hand.
(3) Besides the lack of convexity of the OLS/EOLS functionals, another drawback is the need to compute the derivative of the solution map for evaluating the derivative of the OLS/EOLS functionals. Adjoint methods provide efficient schemes to circumvent this difficulty. We give first-order and second-order adjoint methods in the continuous setting to compute the first-order and the second-order derivative of the OLS/EOLS functionals. The derivative of the MOLS objective does not involve the derivative of the solution map and hence does not require the adjoint approach.
(4) We provide numerical experimentation on tissue phantom data.

We organize the contents of this paper into six sections. Section 2 introduces the inverse problem and explores the smoothness of the solution map. Section 3 investigates the inverse problem by using the OLS approach, the MOLS approach, and the EOLS approach. Sections 4 and 5 are devoted to the first-order and the second-order adjoint approaches for the OLS and the EOLS objectives. In Section 6, we report the outcome of some experiments.

2. Inverse problem formulation

First off, we formulate the saddle point problem, which is the direct problem in this study, and then introduce the inverse problem of parameter identification. In this work, the parameter space B is a real Banach space which contains the nonempty, closed, convex, and bounded set A of admissible parameters. We pose the saddle point problem in real Hilbert spaces V and Q whose topological duals are V^* and Q^*. We take the measured data in real Hilbert spaces $\mathbb{V}$ and $\mathbb{Q}$ and assume that $V \times Q$ is continuously embedded in $\mathbb{V} \times \mathbb{Q}$. We specify the strong convergence by $\to$, the weak convergence by $\rightharpoonup$, and denote the norm of a normed space N by $\|\cdot\|_N$. Assume that $a : B \times V \times V \to \mathbb{R}$ is a trilinear form which is symmetric in the last two arguments, that is, $a(\cdot, u, v) = a(\cdot, v, u)$, for every $u, v \in V$, $b : V \times Q \to \mathbb{R}$ is a bilinear map,

$f \in V^*$, and $g \in Q^*$. Assume that there are positive constants α, β, γ, and τ such that

$$|a(\ell,u,v)| \le \beta \|\ell\|_B \|u\|_V \|v\|_V, \quad \text{for all } \ell \in B,\ u,v \in V, \tag{2.1a}$$

$$|b(v,q)| \le \gamma \|v\|_V \|q\|_Q, \quad \text{for all } v \in V,\ q \in Q, \tag{2.1b}$$

$$a(\ell,v,v) \ge \alpha \|v\|_V^2, \quad \text{for all } \ell \in A,\ v \in V, \tag{2.1c}$$

$$\inf_{q\in Q} \sup_{u\in V} \frac{b(u,q)}{\|u\|_V \|q\|_Q} \ge \tau. \tag{2.1d}$$

Throughout this paper, we will work under the above setting and assumption (2.1).

Consider the saddle point problem: Given $\ell \in A$, find $(u(\ell), p(\ell)) := (u,p) \in V \times Q$ such that

$$a(\ell,u,v) + b(v,p) = f(v), \quad \text{for every } v \in V, \tag{2.2a}$$

$$b(u,q) = g(q), \quad \text{for every } q \in Q. \tag{2.2b}$$

Our objective is to study the inverse problem of identifying a parameter $\ell \in A$ such that the solution $(u(\ell), p(\ell))$ is closest to the given data $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$. Under the set of assumptions (2.1), for $\ell \in A$, the parameter-to-solution map $\ell \mapsto (u(\ell), p(\ell))$ is well-defined and single-valued. More precisely, we have the following well-known result:

Lemma 2.1. *For any $\ell \in A$, the saddle point problem (2.2) has a unique solution $(u(\ell), p(\ell))$. Furthermore, there is a positive constant c such that the following estimate holds:*

$$\|u\|_V + \|p\|_Q \le c\left(\|f\|_{V^*} + \|g\|_{Q^*}\right). \tag{2.3}$$

Proof. See Girault and Raviart [11, Theorem 4.1]. □

Remark 2.1. *Lemma 2.1 holds under a weaker assumption that (2.1a) is valid not on the entire space V but only on the closed subspace $M := \{u \in V \mid b(u,q) = 0, \text{ for all } q \in Q\}$.*

The following result establishes the Lipschitz continuity of the map $\ell \mapsto (u(\ell), p(\ell))$:

Theorem 2.1. *For $\ell_1, \ell_2 \in A$, let $(u(\ell_1), p(\ell_1))$ and $(u(\ell_2), p(\ell_2))$ be the unique solutions of the saddle point problem (2.2). Then, for a constant $L > 0$, the following estimate holds:*

$$\|u(\ell_1) - u(\ell_2)\|_V + \|p(\ell_1) - p(\ell_2)\|_Q \le L\|\ell_1 - \ell_2\|_B. \tag{2.4}$$

Proof. By the definitions of $(u(\ell_1), p(\ell_1))$ and $(u(\ell_2), p(\ell_2))$, we have

$$a(\ell_1, u(\ell_1), v) + b(v, p(\ell_1)) = f(v), \quad \text{for every } v \in V, \tag{2.5a}$$

$$b(u(\ell_1), q) = g(v), \quad \text{for every } q \in Q, \tag{2.5b}$$

and

$$a(\ell_2, u(\ell_2), v) + b(v, p(\ell_2)) = f(v), \quad \text{for every } v \in V, \tag{2.6a}$$

$$b(u(\ell_2), q) = g(v), \quad \text{for every } q \in Q. \tag{2.6b}$$

We set $v = u(\ell_1) - u(\ell_2)$ and $q = p(\ell_1) - p(\ell_2)$ in (2.5) and (2.6), and rearrange them to obtain

$$\begin{aligned} a(\ell_1, u(\ell_1) - u(\ell_2), u(\ell_1) - u(\ell_2)) &+ b(u(\ell_1) - u(\ell_2), p(\ell_1) - p(\ell_2)) \\ &= a(\ell_2 - \ell_1, u(\ell_2), u(\ell_1) - u(\ell_2)), \\ b(u(\ell_1) - u(\ell_2), p(\ell_1) - p(\ell_2)) &= 0, \end{aligned}$$

which, by using (2.1) and the estimate (2.3), implies that

$$\begin{aligned}\alpha\|u(\ell_1)-u(\ell_2)\|_V^2 &\le a(\ell_1,u(\ell_1)-u(\ell_2),u(\ell_1)-u(\ell_2))\\ &= a(\ell_2-\ell_1,u(\ell_2),u(\ell_1)-u(\ell_2))\\ &\le \beta\|\ell_1-\ell_2\|_B\|u(\ell_2)\|_V\|u(\ell_1)-u(\ell_2))\|_V\\ &\le c\beta\left(\|f\|_{V^*}+\|g\|_{Q^*}\right)\|\ell_1-\ell_2\|_B\|u(\ell_1)-u(\ell_2))\|_V,\end{aligned}$$

and consequently

$$\|u(\ell_1)-u(\ell_2)\|_V \le c_1\|\ell_1-\ell_2\|_B, \tag{2.7}$$

where $c_1 := c\beta\alpha^{-1}\left(\|f\|_{V^*}+\|g\|_{Q^*}\right)$.

Now, using (2.5) and (2.6), we also have

$$a(\ell_1,u(\ell_1)-u(\ell_2),v)+b(v,p(\ell_1)-p(\ell_2)) = a(\ell_2-\ell_1,u(\ell_2),v),$$

which, due to the inf-sup condition (cf. (2.1d)), yields,

$$\begin{aligned}\tau\|p(\ell_1)-p(\ell_2)\|_Q &\le \sup_{v\in V}\frac{b(v,p(\ell_1)-p(\ell_2))}{\|v\|_V}\\ &\le \beta c_1\|\ell_1-\ell_2\|_B+\beta\|\ell_1\|_B\|u(\ell_1)-u(\ell_2)\|_V \le c_2\|\ell_1-\ell_2\|_B,\end{aligned}$$

where c_2 is a positive constant. Consequently,

$$\|p(\ell_1)-p(\ell_2)\|_V \le c_2\tau^{-1}\|\ell_1-\ell_2\|_B. \tag{2.8}$$

The desired estimate follows by combining (2.7) and (2.8). The proof is complete. □

The following result gives the smoothness properties of the parameter-to-solution map.

Theorem 2.2. *For each ℓ in the interior of A, the map $\ell \to (u(\ell),p(\ell))$ is infinitely differentiable at ℓ. The first-order derivative $(\delta u,\delta p) = (Du(\ell)\delta\ell, Dp(\ell)\delta\ell)$ is the unique solution of the saddle point problem:*

$$a(\ell,\delta u,v)+b(v,\delta p) = -a(\delta\ell,u,v), \text{ for every } v\in V \tag{2.9a}$$

$$b(\delta u,q) = 0, \text{ for every } q\in Q. \tag{2.9b}$$

The second-order derivative $(\delta^2 u,\delta^2 p) = (D^2u(\ell)(\delta\ell_1,\delta\ell_2), D^2p(\ell)(\delta\ell_1,\delta\ell_2))$ is the unique solution of the following saddle point problem so that, for every $v\in V$ and $q\in Q$,

$$a(\ell,\delta^2u,v)+b(v,\delta^2p) = -a(\delta\ell_2,Du(\ell)\delta\ell_1,v)-a(\delta\ell_1,Du(\ell)\delta\ell_2,v), \tag{2.10a}$$

$$b(\delta^2u,q) = 0. \tag{2.10b}$$

Proof. We begin by introducing the following notations:

$$\delta w := u(\ell+\delta\ell)-u(\ell),$$
$$\delta r := p(\ell+\delta\ell)-p(\ell).$$

Then, we have

$$a(\ell+\delta\ell,u+\delta w,v)+b(v,p+\delta r) = f(v), \quad \text{for every } v\in V, \tag{2.11a}$$

$$b(u+\delta w,q) = g(q), \quad \text{for every } q\in Q, \tag{2.11b}$$

which, when coupled with (2.2), yields

$$a(\ell+\delta\ell,\delta w,v)+b(v,\delta r)=-a(\delta\ell,u,v),\quad \text{for every } v\in V, \tag{2.12a}$$

$$b(\delta w,q)=0,\quad \text{for every } q\in Q, \tag{2.12b}$$

Combining (2.12) and (2.9) we have

$$a(\ell,\delta w-\delta u,v)+b(v,\delta r-\delta p)=-a(\delta\ell,\delta w,v),\quad \text{for every } v\in V, \tag{2.13a}$$

$$b(\delta w-\delta u,q)=0,\quad \text{for every } q\in Q. \tag{2.13b}$$

We take $(v,q)=(\delta w-\delta u,\delta r-\delta p)$ in the above system and obtain

$$a(\ell,\delta w-\delta u,\delta w-\delta u)+b(\delta w-\delta u,\delta r-\delta p)=-a(\delta\ell,\delta w,\delta w-\delta u), \tag{2.14a}$$

$$b(\delta w-\delta u,\delta r-\delta p)=0, \tag{2.14b}$$

and it follows that

$$\|\delta w-\delta u\|_V\le\frac{\beta}{\alpha}\|\delta\ell\|_B\|\delta w\|_V\le c\|\delta\ell\|_B.$$

Hence

$$\frac{\|u(\ell+\delta\ell)-u(\ell)-\delta u\|_V}{\|\delta\ell\|_B}=\frac{\|\delta w-\delta u\|_V}{\|\delta\ell\|_B}\le c\|\delta\ell\|_B=O(\|\delta\ell\|_B),$$

which proves that $Du(\ell)(\delta\ell)=\delta u$.

Now using, (2.13b), we have

$$b(v,\delta r-\delta p)=-a(\ell,\delta w-\delta u,v)-a(\delta\ell,\delta w,v),\quad \text{for every } v\in V,$$

and hence

$$\begin{aligned}\tau\|\delta r-\delta p\|_Q&\le\sup_{v\in V}\frac{b(v,\delta r-\delta p)}{\|v\|_V}\\&\le\beta\|\ell\|\|\delta w-\delta u\|_V+\beta\|\delta\ell\|_B\|\delta w\|_V\\&\le O(\|\delta\ell\|_V^2).\end{aligned}$$

As a consequence, $Dp(\ell)(\delta\ell)=\delta p$. The proof for the second-order and higher-order derivatives follows by combining the above arguments and the arguments given in [14]. □

3. Optimization formulations of the inverse problem

It is a well-known fact that the nonlinear inverse problem of parameter identification is severely ill-posed, and some regularization must be incorporated for a stable inversion. The right choice of the regularization space is crucial and depends on the regularity of the coefficient. A regularization by the aid of the square of a suitable norm, the so-called quadratic regularizer, has been a common choice for smooth coefficients. On the other hand, for rapidly varying or even discontinuous coefficients, the total variation regularization, an example of the nonsmooth-nonquadratic regularization, has rendered promising results. For simplicity, in this work, we resort to the quadratic regularization. For this, we assume that the set A of feasible coefficients belongs to a Hilbert space H that is compactly embedded in the space B. A typical example is $B=L^\infty(\Omega)$ and $H=H^2(\Omega)$, for a suitable domain Ω.

In this work, for the given $\ell \in A$, we will study the following three objective functions:

$$J_{\boldsymbol{O}}(\ell) = \frac{1}{2}\|u(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p(\ell) - \hat{z}\|_{\mathbb{Q}}^2, \tag{3.1}$$

$$J_{\boldsymbol{M}}(\ell) = \frac{1}{2}a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, p(\ell) - \hat{z}), \tag{3.2}$$

$$J_{\boldsymbol{E}}(\ell) = \frac{1}{2}a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}). \tag{3.3}$$

Here $(u(\ell), p(\ell))$ is a solution of saddle point problem (2.2) for $\ell \in A$, $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$ is the given data. In the following, we will consider the regularized analogues of the above functionals.

The functional $J_{\boldsymbol{O}}$ is the most commonly used output least-squares (OLS) objective functional, which aims to minimize the gap between the computed solutions $(u(\ell), p(\ell))$ of (2.2) and the measured data $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$. It turns out that the OLS objective is nonconvex, and hence it is only suitable to investigate local minimizers. Inspired by the shortcomings of the OLS, an abstract convex variant of the OLS was proposed in [14] by minimizing the energy of a general variational problem. However, since saddle point problems involve a coupled system of equations, the two ways of combining them to obtain a single variational problem and then minimizing the associated energy, give rise to two different functionals; namely, the MOLS $J_{\boldsymbol{M}}$ and the EOLS $J_{\boldsymbol{E}}$. Interestingly, the MOLS objective and the EOLS objective have two different features; the MOLS functional preserves the convexity of the original functional (that is, the MOLS proposed in [14]) but loses positivity, whereas the EOLS objective retains positivity, but it is nonconvex, in general.

Note that the MOLS/EOLS functionals require that measured data $(\bar{z}, \hat{z}) \in V \times Q$, which is a stronger regularity assumption in comparison to the OLS formulation of the previous section which only requires that $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$. We note that in many applications, such as the elasticity imaging inverse problem, we only have the displacement component data. In this sense, the EOLS has the slight advantage of not involving the pressure variable data. On the other hand, the OLS functional can be considered without the pressure term. We note that studies related to the MOLS/EOLS functional and its extensions can be found in [5, 9, 10, 12, 13, 15, 19, 20, 25].

3.1. **Inverse problem by the regularized output least squares approach.** We now introduce the following regularized output least-squares (OLS) based optimization problem:

$$\min_{\ell \in A} J_{\boldsymbol{O}}^{\kappa}(\ell) := \frac{1}{2}\|u(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p(\ell) - \hat{z}\|_{\mathbb{Q}}^2 + \kappa\|\ell - \hat{\ell}\|_H^2, \tag{3.4}$$

where $(u(\ell), p(\ell))$ is a solution of (2.2) for $\ell \in A$, $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$ is the data, $\kappa > 0$ is a regularization parameter, $\|\cdot\|_H^2$ is the regularizer, and $\hat{\ell}$ is a-priori estimate of the sought parameter.

We first prove the solvability of the original OLS problem.

Theorem 3.1. *For each fixed $\kappa > 0$, the optimization problem (3.4) has a solution.*

Proof. For each parameter $\ell \in A$, due to Lemma 2.1, saddle point problem (2.2) has a unique solution $(u(\ell), p(\ell))$, and consequently the OLS-based optimization problem (3.4) is well-defined. Since, for each $\ell \in A$, the functional $J_{\boldsymbol{O}}^{\kappa}(\ell)$ is bounded from below, there exists a minimizing sequence $\{\ell_n\}$ in A such that $\lim_{n\to\infty} J_{\boldsymbol{O}}^{\kappa}(\ell_n) = \inf\{J_{\boldsymbol{O}}^{\kappa}(\ell),\ \ell \in A\}$. By the aid of the regularization, the minimizing sequence $\{\ell_n\}$ is bounded in H. By using the compact embedding of H into B, there exists a subsequence, which converges strongly in $\|\cdot\|_B$. By keeping the same

notation for subsequences as well, let $\{\ell_n\}$ be the subsequence, which converges in $\|\cdot\|_L$ to some $\bar{\ell} \in A$. Let $(u_n, p_n) := (u(\ell_n), p(\ell_n))$ be the unique solution of (2.2) that corresponds to ℓ_n. It can be shown that the sequence $\{(u_n, p_n)\}$ remains bounded, and therefore it possesses a weakly convergent subsequence. Let $\{(u_n, p_n)\}$ be the subsequence which converges weakly to some $(\bar{u}, \bar{p}) \in V \times Q$. We claim that $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$. Note that the definition of (ℓ_n, u_n, p_n) implies

$$\begin{aligned} a(\ell_n, u_n, v) + b(v, p_n) &= f(v), \text{ for every } v \in V, \\ b(u_n, q) &= g(q), \text{ for every } q \in Q. \end{aligned}$$

We rearrange the above saddle point problem to obtain

$$\begin{aligned} a(\ell_n - \bar{\ell}, u_n, v) + a(\bar{\ell}, u_n - \bar{u}, v) + a(\bar{\ell}, \bar{u}, v) + b(v, p_n) &= f(v), \text{ for every } v \in V, \\ b(u_n, q) &= g(q), \text{ for every } q \in Q, \end{aligned}$$

which when passed to the limit $n \to \infty$, due to the properties of a and b, implies that

$$\begin{aligned} a(\bar{\ell}, \bar{u}, v) + b(v, \bar{p}) &= f(v), \text{ for every } v \in V, \\ b(\bar{u}, q) &= g(q), \text{ for every } q \in Q, \end{aligned}$$

prove the claim. The optimality of $\bar{\ell}$ now is a consequence of the weak-lower-semicontinuity of any norm. The proof is complete. □

3.2. **Inverse problem by the regularized modified output least squares approach.** We shall now focus on the following regularized MOLS-based constrained optimization problem

$$\min_{\ell \in A} J_{\boldsymbol{M}}^{\kappa}(\ell) := \frac{1}{2} a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, p(\ell) - \hat{z}) + \kappa \|\ell - \hat{\ell}\|_H^2, \tag{3.5}$$

where $(u(\ell), p(\ell))$ is a solution of (2.2) for $\ell \in A$, $(\bar{z}, \hat{z}) \in V \times Q$ is the data, $\kappa > 0$ is a regularization parameter, $\|\cdot\|_H^2$ is the regularizer and $\hat{\ell}$ is a-priori estimate of the sought parameter.

We shall first give the following useful result.

Theorem 3.2. *The modified output least-squares functional* (3.2) *is convex in* A.

Proof. Let us first compute the derivatives of $J_{\boldsymbol{M}}$. For any direction $\hat{\ell} \in B$, we have

$$\begin{aligned} DJ_{\boldsymbol{M}}(\ell)(\hat{\ell}) = \frac{1}{2} a(\hat{\ell}, u(\ell) - \bar{z}, u(\ell) - \bar{z}) &+ a(\ell, Du(\ell)(\hat{\ell}), u(\ell) - \bar{z}) \\ &+ b(Du(\ell)(\hat{\ell}), p(\ell) - \hat{z}) + b(u(\ell) - \bar{z}, Dp(\ell)(\hat{\ell})). \end{aligned} \tag{3.6}$$

Moreover, by using the derivative characterization (2.9), we obtain

$$\begin{aligned} a(\ell, Du(\ell)(\hat{\ell}), u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, Dp(\ell)(\hat{\ell})) &= -a(\hat{\ell}, u(\ell), u(\ell) - \bar{z}), \\ b(Du(\ell)(\hat{\ell}), p(\ell) - \hat{z}) &= 0, \end{aligned}$$

and hence (3.6) yields

$$\begin{aligned} DJ_{\boldsymbol{M}}(\ell)(\hat{\ell}) &= \frac{1}{2} a(\hat{\ell}, u(\ell) - \bar{z}, u(\ell) - \hat{z}) - a(\hat{\ell}, u(\ell), u(\ell) - \bar{z}) \\ &= -\frac{1}{2} a(\hat{\ell}, u(\ell) + \bar{z}, u(\ell) - \bar{z}). \end{aligned}$$

Furthermore,

$$\begin{aligned} D^2 J_{\boldsymbol{M}}(\ell)(\hat{\ell},\hat{\ell}) &= -\frac{1}{2}a(\hat{\ell},u(\ell)-\bar{z},Du(\ell)(\hat{\ell})) - \frac{1}{2}a(\hat{\ell},u(\ell)+\bar{z},Du(\ell)(\hat{\ell})) \\ &= -a(\hat{\ell},u(\ell),Du(\ell)(\hat{\ell})) \\ &= a(\ell,Du(\ell)\delta\ell,Du(\ell)(\hat{\ell})) + b(Du(\ell)(\hat{\ell}),Dp(\ell)(\hat{\ell})) \\ &= a(\ell,Du(\ell)(\hat{\ell}),Du(\ell)(\hat{\ell})), \end{aligned}$$

where we used the following consequence of the derivative formula:

$$\begin{aligned} a(\ell,Du(\ell)(\hat{\ell}),Du(\ell)(\hat{\ell})) + b(Du(\ell)(\hat{\ell}),Dp(\ell)(\hat{\ell})) &= -a(\hat{\ell},u(\ell),Du(\ell)(\hat{\ell})), \\ b(Du(\ell)(\hat{\ell}),Dp(\ell)(\hat{\ell})) &= 0. \end{aligned}$$

We notice, in particular, that the following inequality holds for all ℓ in the interior of A

$$D^2 J_{\boldsymbol{M}}(\ell)(\hat{\ell},\hat{\ell}) \geq \alpha \|Du(\ell)(\hat{\ell})\|_V^2, \tag{3.7}$$

and consequently the convexity of $J_{\boldsymbol{M}}$ follows. □

We have the following existence result.

Theorem 3.3. *For each fixed $\kappa > 0$, the optimization problem (3.5) has a solution.*

Proof. Since the proof is quite similar to the proof of Theorem 3.1, we only point out the differences in the two proofs. First of all, we note that $J_{\boldsymbol{M}}^{\kappa}$ is not positive. However, due to the boundedness of the solution of saddle point problem (2.2), it remains bounded from below. Furthermore, we can show that $J_{\boldsymbol{M}}(\ell_n) \to J_{\boldsymbol{M}}(\bar{\ell})$ for any sequence $\{\ell_n\}$ converging to $\bar{\ell}$ in $\|\cdot\|_B$. Because of these observations, the proof follows at once. □

3.3. **Inverse problem by the regularized energy output least squares approach.** We now introduce the following regularized EOLS-based optimization problem:

$$\min_{\ell \in A} J_E^{\kappa}(\ell) := \frac{1}{2}a(\ell,u(\ell)-\bar{z},u(\ell)-\bar{z}) + \kappa\|\ell-\hat{\ell}\|_H^2, \tag{3.8}$$

where $(u(\ell),p(\ell))$ is a solution of (2.2) for $\ell \in A$, $\bar{z} \in V$ is the given data, $\kappa > 0$ is a regularization parameter, $\|\cdot\|_H^2$ is the regularizer and $\hat{\ell}$ is a-priori known estimate of the sought parameter.

We have the following existence result:

Theorem 3.4. *For each fixed $\kappa > 0$, the optimization problem (3.8) has a solution.*

Proof. The proof is similar to the proof of Theorem 3.3 and hence omitted. □

4. Evaluation of First-order and Second-order Derivatives of the OLS

One of the shortcomings of an OLS-based approach is the evaluation of the derivatives of the OLS functional, which demands computationally expensive evaluation of the solution map. Adjoint methods render a framework for computing the derivatives of the OLS functional, and have been explored intensively. Recent developments in adjoint methods can be found in [6, 22, 23, 25]. Furthermore, in [19], first-order and second-order adjoint methods were applied to nearly incompressible elasticity imaging.

4.1. **Evaluation of the first-order derivative.** The OLS objective, without the regularizer, with the saddle point problem as the constraint, reads

$$\min_{\ell\in A} J_{\boldsymbol{O}}(\ell) := \frac{1}{2}\|u(\ell)-\bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p(\ell)-\hat{z}\|_{\mathbb{Q}}^2, \tag{4.1}$$

where $(u(\ell), p(\ell))$ is the unique solution of the saddle point problem:

$$a(\ell, u(\ell), v) + b(v, p(\ell)) = f(v), \quad \text{for each } v \in V, \tag{4.2a}$$

$$b(u(\ell), q) = g(q), \quad \text{for each } q \in Q. \tag{4.2b}$$

The derivative of $J_{\boldsymbol{O}}$ at $\ell \in A$ in a direction $\delta\ell$ is given by

$$DJ_{\boldsymbol{O}}(\ell)(\delta\ell) := \langle Du(\ell)(\delta\ell), u(\ell)-\bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell), p(\ell)-\hat{z}\rangle_{\mathbb{Q}},$$

where $(Du(\ell)(\delta\ell), Dp(\ell)(\delta\ell))$ is the derivative of the coefficient-to-solution map $\ell \mapsto (u,p)$ computed at ℓ in the direction $\delta\ell$.

For an arbitrary $(v,q) \in V \times Q$, we define the map $L : B \times V \times Q \to \mathbb{R}$ by

$$L(\ell, v, q) = J_{\boldsymbol{O}}(\ell) + a(\ell, u(\ell), v) + b(v, p(\ell)) + b(u(\ell), q) - f(v) - g(v).$$

Since $(u(\ell), p(\ell))$ is the unique solution of (4.2), the following identity holds

$$L(\ell, v, q) = J_{\boldsymbol{O}}(\ell), \text{ for every } \ (v,q) \in V \times Q,$$

and consequently, for every $(v,q) \in V \times Q$, the following identity holds for any direction $\delta\ell$:

$$\partial_\ell L(\ell, v, q)\,(\delta\ell) = DJ_{\boldsymbol{O}}(\ell)\,(\delta\ell)\,. \tag{4.3}$$

The primary obstacle in computing the derivative of the OLS objective is the involvement of the solution map. The adjoint method chooses the test function (v,q) to avoid its direct computation as we shall see shortly. First, we note that

$$\begin{aligned}\partial_\ell L(\ell, v, q)\,(\delta\ell) = {} & \langle Du(\ell)(\delta\ell), u(\ell)-\bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell), p(\ell)-\hat{z}\rangle_{\mathbb{Q}} + a(\delta\ell, u(\ell), v) \\ & + a(\ell, Du(\ell)(\delta\ell), v) + b(v, Dp(\ell)(\delta\ell)) + b(Du(\ell)(\delta\ell), q).\end{aligned} \tag{4.4}$$

For $\ell \in A$, let $(w(\ell), t(\ell))$ be the unique solution of the adjoint problem

$$a(\ell, w(\ell), v) + b(v, t(\ell)) = \langle \bar{z} - u(\ell), v\rangle\,, \quad \text{for every } v \in V, \tag{4.5a}$$

$$b(w(\ell), q) = \langle \hat{z} - p(\ell), q\rangle\,, \quad \text{for every } q \in Q, \tag{4.5b}$$

where $(u(\ell), p(\ell))$ solves (4.2) for the given ℓ and $(\bar{z}, \hat{z})$ is the given data.

We set $(v,q) = (w(\ell), t(\ell))$ in (4.4) and after a simplification obtain

$$\begin{aligned}
&\partial_\ell L(\ell, w(\ell), t(\ell))\,(\delta\ell) \\
&= \langle Du(\ell)(\delta\ell), u(\ell)-\bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell), p(\ell)-\hat{z}\rangle_{\mathbb{Q}} + a(\delta\ell, u(\ell), w(\ell)) \\
&+ a(\ell, Du(\ell)(\delta\ell), w(\ell)) + b(w(\ell), Dp(\ell)(\delta\ell)) + b(Du(\ell)(\delta\ell), t(\ell)) \\
&= \langle Du(\ell)(\delta\ell), u(\ell)-\bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell), p(\ell)-\hat{z}\rangle_{\mathbb{Q}} + a(\delta\ell, u(\ell), w(\ell)) \\
&+ a(\ell, w(\ell), Du(\ell)(\delta\ell)) + b(w(\ell), Dp(\ell)(\delta\ell)) + b(Du(\ell)(\delta\ell), t(\ell)) \\
&= \langle Du(\ell)(\delta\ell), u(\ell)-\bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell), p(\ell)-\hat{z}\rangle_{\mathbb{Q}} + a(\delta\ell, u(\ell), w(\ell)) \\
&+ \langle \bar{z} - u(\ell), Du(\ell)(\delta\ell)\rangle_{\mathbb{V}} + \langle \hat{z} - p(\ell), Dp(\ell)(\delta\ell)\rangle_{\mathbb{Q}} \\
&= a(\delta\ell, u(\ell), w(\ell)),
\end{aligned}$$

which at once gives the formula for the first-order derivative of $J_{\boldsymbol{O}}$:

$$DJ_{\boldsymbol{O}}(\ell)\,(\delta\ell) = a(\delta\ell, u(\ell), w(\ell)). \tag{4.6}$$

In summary, we deduce the following scheme to compute $DJ_{\boldsymbol{O}}(\ell)\,(\delta\ell)$:

(1) Compute $(u(\ell), p(\ell))$ by solving the regularized saddle point problem (4.2).
(2) Compute $(w(\ell), t(\ell))$ by solving the regularized adjoint problem (4.5).
(3) Compute $DJ_{\boldsymbol{O}}(\ell)\,(\delta\ell)$ by using (4.6).

4.2. **Evaluation of the second-order derivative.** We now provide a second-order adjoint method for the evaluation of the second-order derivative of the OLS objective. The objective of the second-order adjoint approach is to give a formula for the second-order derivative that does not involve the second-order derivative of the regularized parameter-to-solution map u. The key idea of the second-order method is to compute the derivative δu directly by using its characterization while the computation of the second-order derivative is avoided by the strategy of the adjoint method.

We recall the derivative characterization of the regularized parameter-to-solution map:

$$a(\ell, Du(\ell)\delta\ell, v) + b(v, Dp(\ell)\delta\ell) = -a(\delta\ell, u(\ell), v), \quad \text{for all } v \in V, \tag{4.7a}$$

$$b(Du(\ell)\delta\ell, q) = 0, \quad \text{for all } q \in Q. \tag{4.7b}$$

For any $(v,q) \in V \times Q$, $(u(\ell), p(\ell)) \in V \times Q$, and for a fixed direction $\delta\ell_2$, we define

$$\begin{aligned}\mathscr{L}(\ell, v, q) &:= DJ_{\boldsymbol{O}}(\ell)(\delta\ell_2) + a(\ell_2, Du(\ell)\delta\ell_2, v) + b(v, Dp(\ell)\delta\ell_2) + a(\delta\ell_2, u(\ell), v) \qquad (4.8)\\ &\quad + b(Du(\ell)\delta\ell_2, q)\\ &= \langle Du(\ell)(\delta\ell_2), u(\ell) - \bar{z}\rangle_{\mathbb{V}} + \langle Dp(\ell)(\delta\ell_2), p(\ell) - \hat{z}\rangle_{\mathbb{Q}} + a(\ell, Du(\ell)\delta\ell_2, v)\\ &\quad + b(v, Dp(\ell)\delta\ell_2) + a(\delta\ell_2, u(\ell), v) + b(Du(\ell)\delta\ell_2, q). \qquad (4.9)\end{aligned}$$

Using the definition of $\mathscr{L}$, for every $(v,q) \in V \times Q$, and for any direction $\delta\ell_1$, we have

$$\partial_\ell \mathscr{L}(\ell, v, q)(\delta\ell_1) = D^2 J_{\boldsymbol{O}}(\ell)(\delta\ell_1, \delta\ell_2), \tag{4.10}$$

By computing the right-hand side of the above identity and using (4.9), we obtain

$$\begin{aligned}&\partial_\ell \mathscr{L}(\ell, v, q)(\delta\ell_1)\\ &= \left\langle D^2 u(\ell)(\delta\ell_1, \delta\ell_2), u - \bar{z}\right\rangle + \langle Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle\\ &+ \left\langle D^2 p(\ell)(\delta\ell_1, \delta\ell_2), p - \hat{z}\right\rangle + \langle Dp(\ell)(\delta\ell_2), Dp(\ell)(\delta\ell_1)\rangle\\ &+ a(\delta\ell_2, Du(\ell)(\delta\ell_1), v) + a(\delta\ell_1, Du(\ell)(\delta\ell_2), v) + a(\ell, D^2 u(\ell)(\delta\ell_1, \delta\ell_2), v)\\ &+ b(v, D^2 p(\ell)(\delta\ell_1, \delta\ell_2)) + b(D^2 u(\ell)(\delta\ell_1, \delta\ell_2), q).\end{aligned}$$

By setting $(v,q) = (w(\ell),t(\ell))$; the solution of (4.5), we obtain

$$\begin{aligned}
&\partial_\ell \mathscr{L}(\ell,v,q)(\delta\ell_1)\\
&= \left\langle D^2u(\ell)(\delta\ell_1,\delta\ell_2), u-\bar{z}\right\rangle + \langle Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle\\
&+ \left\langle D^2p(\ell)(\delta\ell_1,\delta\ell_2), p-\hat{z}\right\rangle + \langle Dp(\ell)(\delta\ell_2), Dp(\ell)(\delta\ell_1)\rangle + a(\delta\ell_2, Du(\ell)(\delta\ell_1), w(\ell))\\
&+ a(\delta\ell_1, Du(\ell)(\delta\ell_2), w(\ell)) + a(\ell, D^2u(\ell)(\delta\ell_1,\delta\ell_2), w(\ell))\\
&+ b(w(\ell), D^2p(\ell)(\delta\ell_1,\delta\ell_2)) + b(D^2u(\ell)(\delta\ell_1,\delta\ell_2), t(\ell))\\
&= \left\langle D^2u(\ell)(\delta\ell_1,\delta\ell_2), u-\bar{z}\right\rangle + \langle Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle + \left\langle D^2p(\ell)(\delta\ell_1,\delta\ell_2), p-\hat{z}\right\rangle\\
&+ \langle Dp(\ell)(\delta\ell_2), Dp(\ell)(\delta\ell_1)\rangle + a(\delta\ell_1, Du(\ell)(\delta\ell_2), w) + a(\delta\ell_2, D\bar{u}(\ell)(\delta\ell_1), w)\\
&+ \left\langle \bar{z}-u, D^2u(\ell)(\delta\ell_1,\delta\ell_2)\right\rangle + \left\langle \hat{z}-p, D^2p(\ell)(\delta\ell_1,\delta\ell_2)\right\rangle\\
&= \langle Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle + \langle Dp(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle + a(\delta\ell_1, Du(\ell)(\delta\ell_2), w)\\
&+ a(\delta\ell_2, Du(\ell)(\delta\ell_1), w).
\end{aligned}$$

Therefore, using (4.10), we get the following formula for the second-order derivative of the OLS that has no explicit involvement of the second-order derivatives of the solution map:

$$\begin{aligned}
D^2J_{\boldsymbol{O}}(\ell)(\delta\ell_1,\delta\ell_2) = &\langle Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle + \langle Dp(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)\rangle\\
&+ a(\delta\ell_1, Du(\ell)(\delta\ell_2), w) + a(\delta\ell_2, Du(\ell)(\delta\ell_1), w).
\end{aligned}$$

In particular, we have

$$D^2J_{\boldsymbol{O}}(\ell)(\delta\ell,\delta\ell) = \langle \delta u, \delta u\rangle + \langle \delta p, \delta p\rangle + 2a(\delta\ell, \delta u, w). \tag{4.11}$$

In summarizing, we obtain the following scheme to compute $D^2J_{\boldsymbol{O}}(\ell)(\delta\ell,\delta\ell)$:

(1) Compute $(u(\ell), p(\ell))$ by solving the saddle point problem (2.2).
(2) Compute $(\delta u, \delta p)$ by solving the regularized saddle point problem (4.7).
(3) Compute $(w(\ell), t(\ell))$ by solving the regularized adjoint problem (4.5).
(4) Compute $D^2J_{\boldsymbol{O}}(\ell)(\delta\ell,\delta\ell)$ by (4.11).

We note that the second-order adjoint approach given above is based on evaluating the second-order derivative of regularized OLS by a direct computation of its first-order derivative. However, if we employ the first-order derivative formula of the OLS obtained from the first-order adjoint approach, we devise a different second-order adjoint approach.

5. Evaluation of First-order and Second-order Derivatives of the EOLS

We will now develop adjoint methods for the EOLS functional given by

$$J_{\boldsymbol{E}}(\ell) := \frac{1}{2}a(\ell, u-\bar{z}, u-\bar{z}).$$

As before, we define the Lagrangian as follows

$$L_{\boldsymbol{E}}(\ell,v,q) := J_{\boldsymbol{E}}(\ell) + a(\ell,u,v) + b(v,p) + b(u,q) - f(v) - g(q), \quad \text{for every } (v,q) \in V\times Q,$$

and notice that $L_{\boldsymbol{E}}(\ell,v,q) = J_{\boldsymbol{E}}(\ell)$, for every $(v,q) \in V\times Q$, and consequently, for every $(v,q) \in V\times Q$, and for every feasible direction $\delta\ell$, we have $DJ_{\boldsymbol{E}}(\ell)(\delta\ell) = D_\ell J(\ell,v,q)(\delta\ell)$. Furthermore,

we have

$$D_\ell J(\ell,v,q)(\delta\ell) = \frac{1}{2}a(\delta\ell,u-\bar{z},u-\bar{z}) + a(\ell,\delta u,u-\bar{z}) + a(\delta\ell,u,v) + a(\ell,\delta u,v) + b(v,\delta p) + b(\delta u,q). \tag{5.1}$$

Let (w,p_w) be the unique solution of the following adjoint saddle point problem:

$$a(\ell,w,v) + b(v,p_w) = -a(\ell,u-\bar{z},v), \quad \text{for every } v \in V, \tag{5.2a}$$

$$b(w,q) = 0, \quad \text{for every } q \in Q. \tag{5.2b}$$

By setting $(v,q) = (w,p_w)$ in (5.1) and using (5.2), we obtain

$$\begin{aligned} D_\ell J(\ell,w,p_w)(\delta\ell) &= \frac{1}{2}a(\delta\ell,u-\bar{z},u-\bar{z}) + a(\ell,\delta u,u-\bar{z}) + a(\delta\ell,u,w) + a(\ell,\delta u,w) \\ &\quad + b(w,\delta p) + b(\delta u,p_w) \\ &= \frac{1}{2}a(\delta\ell,u-\bar{z},u-\bar{z}) + a(\delta\ell,u,w). \end{aligned} \tag{5.3}$$

Summarizing, the following scheme computes the derivative $DJ(\ell)(\delta\ell)$:

(1) Compute (u,p) by solving saddle point problem (2.2).
(2) Compute (w,p_w) by solving adjoint saddle point problem (5.2).
(3) Compute $DJ_{\boldsymbol{E}}(\ell)(\delta\ell)$ by (5.3).

5.1. **Second-order adjoint method for the EOLS.** We shall now develop a second-order adjoint method for the computation of the second-order derivative. As in the case of the OLS, the derivative δu is computed directly while the computation of the second derivative $\delta^2 u$ is avoided by using the adjoint method philosophy.

Let $\delta\ell_2$ be a fixed feasible direction. Then, for any $(v,q) \in V$,

$$\begin{aligned} \mathscr{L}_{\boldsymbol{E}}(\ell,v,q) &:= DJ_{\boldsymbol{E}}(\ell)(\delta\ell_2) + a(\ell,Du(\ell)(\delta\ell_2),v) + b(v,Dp(\ell)(\delta\ell_2)) \\ &\quad + b(Du(\ell)(\delta\ell_2),q) + a(\delta\ell_2,u,v) \\ &= \frac{1}{2}a(\delta\ell_2,u-\bar{z},u-\bar{z}) + a(\ell,Du(\ell)(\delta\ell_2),u-\bar{z}) + a(\ell,Du(\ell)(\delta\ell_2),v) \\ &\quad + b(v,Dp(\ell)(\delta\ell_2)) + b(Du(\ell)(\delta\ell_2),q) + a(\delta\ell_2,u,v). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} \partial_\ell \mathscr{L}_{\boldsymbol{E}}(\ell,v,q)(\delta\ell_1) &= a(\delta\ell_2,Du(\ell)(\delta\ell_1),u-\bar{z}) + a(\delta\ell_1,Du(\ell)(\delta\ell_2),u-\bar{z}) \\ &\quad + a(\ell,D^2u(\ell)(\delta\ell_1,\delta\ell_2),u-\bar{z}) + a(\ell,Du(\ell)(\delta\ell_2),Du(\ell)(\delta\ell_1)) \\ &\quad + a(\delta\ell_1,Du(\ell)(\delta\ell_2),v) + a(\ell,D^2u(\ell)(\delta\ell_1,\delta\ell_2),v) + b(v,D^2p(\ell)(\delta\ell_1,\delta\ell_2)) \\ &\quad + b(D^2u(\ell)(\delta\ell_1,\delta\ell_2),q) + a(\delta\ell_2,Du(\ell)(\delta\ell_1),v). \end{aligned} \tag{5.4}$$

Let (w, p_w) be the unique solution of (5.2). By setting $(v,q) = (w, p_w)$ in (5.4), we obtain

$$\begin{aligned}
\partial_\ell \mathscr{L}_E(\ell, w, p_w)(\delta\ell) &= a(\delta\ell_2, Du(\ell)(\delta\ell_1), u-\bar{z}) + a(\delta\ell_1, Du(\ell)(\delta\ell_2), u-\bar{z}) \\
&+ a(\ell, D^2u(\ell)(\delta\ell_1,\delta\ell_2), u-\bar{z}) + a(\ell, Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)) \\
&+ a(\delta\ell_1, Du(\ell)(\delta\ell_2), w) + a(\ell, D^2u(\ell)(\delta\ell_1,\delta\ell_2), w) \\
&+ b(w, D^2p(\ell)(\delta\ell_1,\delta\ell_2)) + b(D^2u(\ell)(\delta\ell_1,\delta\ell_2), p_w) + a(\delta\ell_2, Du(\ell)(\delta\ell_1), w) \\
&= a(\delta\ell_2, Du(\ell)(\delta\ell_1), u-\bar{z}) + a(\delta\ell_1, Du(\ell)(\delta\ell_2), u-\bar{z}) + a(\ell, Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)) \\
&+ a(\ell, D^2u(\ell)(\delta\ell_1,\delta\ell_2), u-\bar{z}) + a(\delta\ell_1, Du(\ell)(\delta\ell_2), w) + a(\ell, D^2u(\ell)(\delta\ell_1,\delta\ell_2), w) \\
&+ b(w, D^2p(\ell)(\delta\ell_1,\delta\ell_2)) + b(D^2u(\ell)(\delta\ell_1,\delta\ell_2), p_w) + a(\delta\ell_2, Du(\ell)(\delta\ell_1), w) \\
&= a(\delta\ell_2, Du(\ell)(\delta\ell_1), u-\bar{z}) + a(\delta\ell_1, Du(\ell)(\delta\ell_2), u-\bar{z}) + a(\ell, Du(\ell)(\delta\ell_2), Du(\ell)(\delta\ell_1)) \\
&+ a(\delta\ell_1, Du(\ell)(\delta\ell_2), w) + a(\delta\ell_2, Du(\ell)(\delta\ell_1), w) \\
&= D^2J_E(\ell)(\delta\ell_1,\delta\ell_2).
\end{aligned}$$

In particular, we have

$$D^2J(\ell)(\delta\ell,\delta\ell) = 2a(\delta\ell, \delta u, \bar{u}-\bar{z}) + a(\ell, \delta u, \delta u) + 2a(\delta\ell, \delta u, w). \tag{5.5}$$

Summarizing, the following scheme computes the derivative $D^2J(\ell)(\delta\ell,\delta\ell)$:

(1) Compute $(u(\ell), p)$ by solving saddle point problem (2.2).
(2) Compute $(\delta u, \delta p)$ by solving derivative characterization (2.9).
(3) Compute (w, p_w) adjoint saddle point problem (5.2).
(4) Compute $D^2J(\ell)(\delta\ell,\delta\ell)$ by (5.5).

6. Numerical Experiments

In this paper, we acquired simulated displacement data using a finite element model of a tissue like phantom designed to simulate the homogenous soft tissue of an organ surrounding a stiff tumor. The simulated phantom and experimental conditions are shown in Figure 1. The model simulated an incompressible ($\nu = 0.495$), plane strain phantom with a 1 cm calibration pad situated in the bottom of the imaged domain. The stand-off pad had a shear modulus value 3 times that of the background material. This phantom was designed with a cylindrical inclusion of 1.5 cm of diameter and a shear modulus contrast of 2.5 relative to the background modulus. The center of the inclusion is situated 3.25 cm from the bottom of the phantom and centered laterally. The overall size of the phantom was 8 cm tall and 10 cm wide. In the forward simulation, a 3% axial compressive strain with no slip was applied to the top boundary of the phantom. The bottom of the phantom was fixed and the lateral walls of the phantom were traction free. Two dimensional displacement vector data was extracted starting at the bottom of the phantom, as if imaged from underneath, with a resolution in both the axial (y) and lateral (x) directions of approximately 0.12 cm. The forward simulation was done using ANSYS.

Following [26, 27], we consider Dirichlet boundary conditions in the top and bottom boundaries of the imaged domian. We set the vertical direction/component of the remaining boundary conditions to be Dirichlet and allow the lateral components to be traction free. We solve the problem with IPOPT and take $l_b = 0.5$, $u_b = 5$ as the lower and the upper bound for the parameters. Since we are dealing with a discontinuous parameter, we have chosen the standard TV regularizer, see [20].

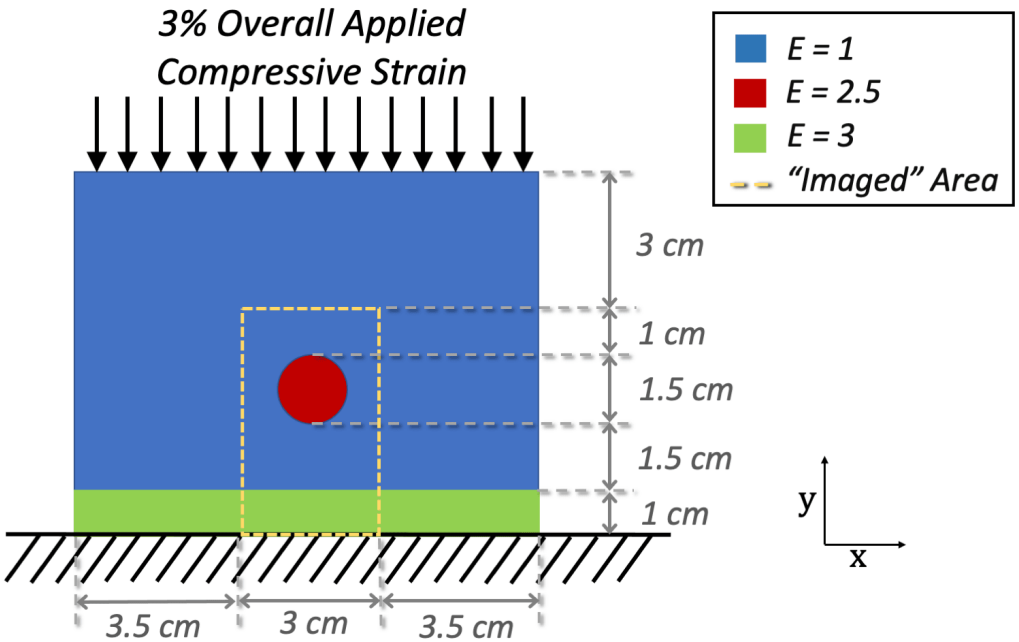
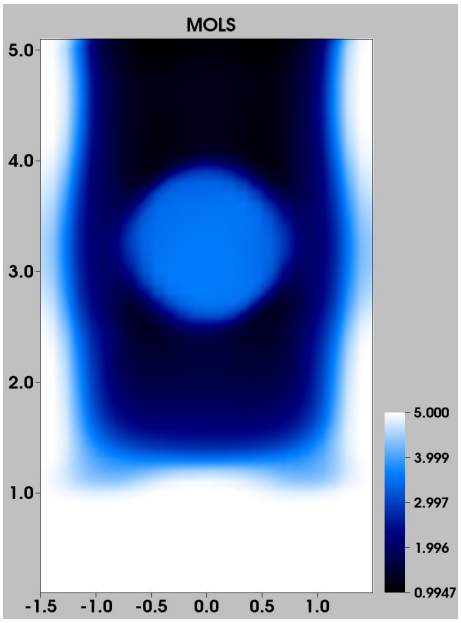
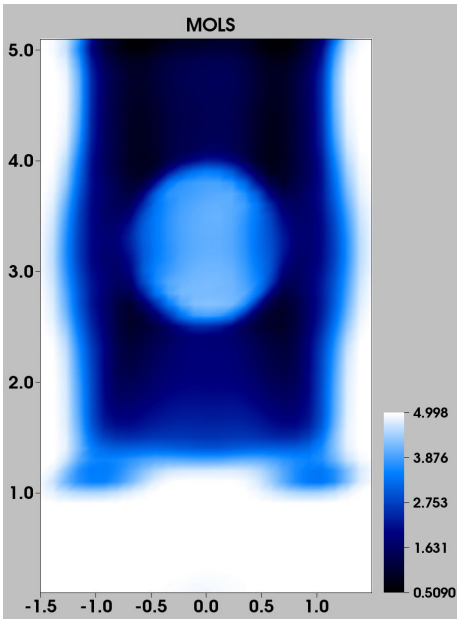


FIGURE 1. Schematic diagram of the finite element, forward simulation setup. An overall 3% compressive strain was applied at the top of a 2D plane strain phantom (blue) with a 1.5 cm diameter, cylindrical inclusion (red) and a 1 cm calibration pad (green). Data was interpolated from the finite element mesh at regularly sampled points (0.12 cm spacing in x and y) within the "imaged" region (yellow). The relative shear modulus of the background, inclusion and standoff pad are 1, 2.5 and 3, respectively.

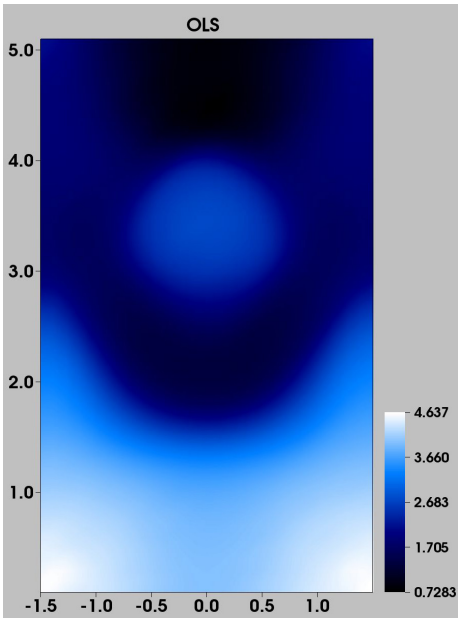
Numerical results are presented in Figure 2 and Figure 3 for different choices of regularization parameters. The selection of the regularization parameter is purely heuristic; we considered four parameters $\kappa \in \{1e-04, 1e-05, 1e-06, 1e-07\}$ and choose the one producing better results. Here we provide the comparison only between the MOLS and the OLS objectives. Both approaches identify the location of the standoff pad and inclusion correctly. Furthermore, both methods can recover the circular shape of the inclusion. However, the estimation obtained by MOLS seems superior, while MOLS shows certain artifacts close to the boundary. On the contrary, for the chosen regularization parameters, OLS is better for determining the contrast of inclusion concerning the standoff pad and the background. However, MOLS still gives a reasonable measure.



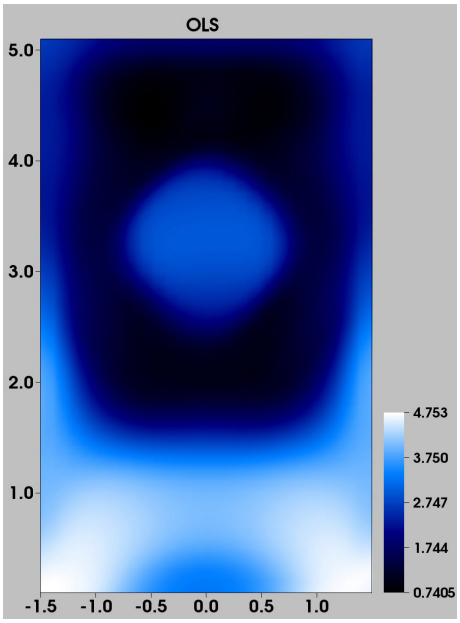
(A) (MOLS) $\kappa = 1e-05$



(B) (MOLS) $\kappa = 1e-06$

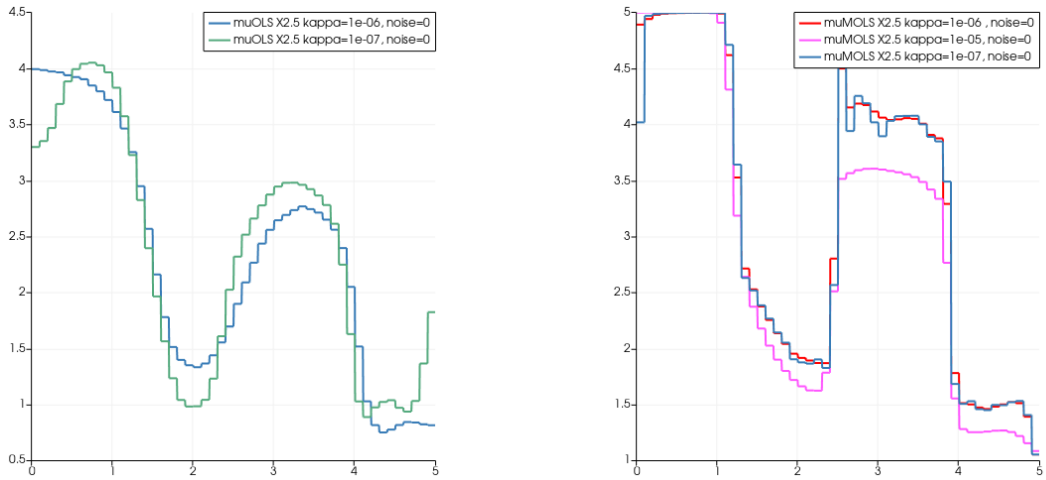


(C) (OLS) $\kappa = 1e-06$



(D) (OLS) $\kappa = 1e-07$

FIGURE 2. Estimated inclusion for the two approaches.



(A) (MOLS) (B) (OLS)

FIGURE 3. Vertical cut ($x = 0$) for different regularization parameters.

Acknowledgements

The research of Baasansuren Jadamba and Akhtar Khan was supported by the National Science Foundation under Award No. 1720067. Miguel Sama was supported by Ministerio de Ciencia, Innovación y Universidades (MCIU), Agencia Estatal de Investigación (AEI) (Spain) and Fondo Europeo de Desarrollo Regional (FEDER) under project PGC2018-096899-B-I00 (MCIU/AEI/FEDER, UE).

REFERENCES

[1] U. Albocher, P. E. Barbone, A. A. Oberai, I. Harari, Uniqueness of inverse problems of isotropic incompressible three-dimensional elasticity, J. Mech. Phys. Solids 73 (2014) 55-68.

[2] H. Ammari, P. Garapon, F. Jouve, Separation of scales in elasticity imaging: a numerical study, J. Comput. Math. 28 (2010) 354-370.

[3] H. Ammari, P. Garapon, H. Kang, H. Lee, A method of biological tissues elasticity reconstruction using magnetic resonance elastography measurements, Quart. Appl. Math. 66 (2008) 139-175.

[4] N. Cahill, B. Jadamba, A. A. Khan, M. Sama, B. Winkler, 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 Probl. 2013 (2013), 263.

[5] M. Cho, B. Jadamba, R. Kahler, A. A. Khan, M. Sama, First-order and second-order adjoint methods for the inverse problem of identifying nonlinear parameters in pdes, Industrial Mathematics and Complex systems, Springer, pp. 1–16, 2017.

[6] A. Cioaca, M. Alexe, A. Sandu, Second-order adjoints for solving PDE-constrained optimization problems, Optim. Methods Softw. 27 (2012) 625-653.

[7] E. Crossen, M. S. Gockenbach, B. Jadamba, A. A. Khan, B. Winkler, An equation error approach for the elasticity imaging inverse problem for predicting tumor location, Comput. Math. Appl. 67 (2014) 122-135.

[8] M. M. Doyley, Model-based elastography: a survey of approaches to the inverse elasticity problem, Phys. Med. Biol. 57 (2012), R35-R73.

[9] M. M. Doyley, B. Jadamba, A. A. Khan, M. Sama, B. Winkler, A new energy inversion for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location, Numer. Funct. Anal. Optim. 35 (2014), 984–1017.

[10] A. Gibali, B. Jadamba, A. A. Khan, F. Raciti, B. Winkler, Gradient and extragradient methods for the elasticity imaging inverse problem using an equation error formulation: a comparative numerical study, in: Nonlinear analysis and optimization, vol. 659 of Contemp. Math., Amer. Math. Soc., Providence, RI, pp. 65–89, 2016.

[11] V. Girault, P.-A. Raviart, Finite element approximation of the Navier-Stokes equations, vol. 749 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 1979.

[12] M. S. Gockenbach, B. Jadamba, A. A. Khan, Numerical estimation of discontinuous coefficients by the method of equation error, Int. J. Math. Comput. Sci. 1 (2006), 343-359.

[13] M. S. Gockenbach, A. A. Khan, Identification of Lamé parameters in linear elasticity: a fixed point approach, J. Ind. Manag. Optim. 1 (2005), 487-497.

[14] M. S. Gockenbach, A. A. Khan, An abstract framework for elliptic inverse problems: Part 1. an output least-squares approach, Math. Mech. Solids 12 (2007), 259–276.

[15] M. S. Gockenbach, A. A. Khan, An abstract framework for elliptic inverse problems. II. An augmented Lagrangian approach, Math. Mech. Solids 14 (2009), 517-539.

[16] S. Guchhait, B. Banerjee, Constitutive error based material parameter estimation procedure for hyperelastic material, Comput. Methods Appl. Mech. Engrg. 297 (2015), 455–475.

[17] S. Hubmer, E. Sherina, A. Neubauer, O. Scherzer, Lamé parameter estimation from static displacement field measurements in the framework of nonlinear inverse problems, SIAM J. Imaging Sci. 11 (2018) 1268-1293.

[18] T.J.R. Hughes, The finite element method, Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.

[19] B. Jadamba, A. A. Khan, A. Oberai, M. Sama, First-order and second-order adjoint methods for parameter identification problems with an application to the elasticity imaging inverse problem, Inverse Probl. Sci. Eng. (2017), 1–20.

[20] B. Jadamba, A. A. Khan, G. Rus, M. Sama, B. Winkler, A new convex inversion framework for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location, SIAM J. Appl. Math. 74 (2014), 1486-1510.

[21] L. Ji, J. McLaughlin, Recovery of the Lamé parameter μ in biological tissues, Inverse Probl. 20 (2004) 1-24.

[22] D. A. Knopoff, D. R. Fernández, G. A. Torres, C. V. Turner, Adjoint method for a tumor growth PDE-constrained optimization problem, Comput. Math. Appl. 66 (2013) 1104-1119.

[23] D. Kourounis, L.J. Durlofsky, J.D. Jansen, K. Aziz, Adjoint formulation and constraint handling for gradient-based optimization of compositional reservoir flow, Comput. Geosci. 18 (2014) 117-137.

[24] J. R. McLaughlin, J.-R. Yoon, Unique identifiability of elastic parameters from time-dependent interior displacement measurement, Inverse Probl. 20 (2004) 25-45.

[25] A. A. Oberai, N. H. Gokhale, G. R. Feijóo, Solution of inverse problems in elasticity imaging using the adjoint method, Inverse Probl. 19 (2003), 297-313.

[26] M. S. Richards, Quantitative three dimensional elasticity imaging, Ph.D. thesis, Boston University, 2007.

[27] M. S. Richards, P. E. Barbone, A. A. Oberai, Quantitative three-dimensional elasticity imaging from quasistatic deformation: a phantom study, Physics in Medicine & Biology 54 (2009) 757.

[28] T. Widlak, O. Scherzer, Stability in the linearized problem of quantitative elastography, Inverse Probl. 31 (2015), 035005.