

A convex inversion framework for identifying parameters in saddle point problems with applications to inverse incompressible elasticity[†]

Baasansuren Jadamba¹, Akhtar A Khan¹,
Michael Richards² and Miguel Sama^{3,4} 

¹ Center for Applied and Computational Mathematics, School of Mathematical Sciences, Rochester Institute of Technology, 85 Lomb Memorial Drive, Rochester, New York, 14623, United States of America

² Department of Biomedical Engineering Kate Gleason College of Engineering Rochester Institute of Technology, Rochester, New York, 14623, United States of America

³ Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Calle Juan del Rosal, 12, 28040 Madrid, Spain

E-mail: msama@ind.uned.es

Received 26 July 2019, revised 6 February 2020

Accepted for publication 30 March 2020

Published 30 June 2020



Abstract

This work investigates the elasticity imaging inverse problem of tumor identification in a fully incompressible medium through a family of inverse problems in a nearly incompressible medium. We develop an inversion framework for saddle point problems that goes far beyond the elasticity imaging inverse problem and applies to a wide variety of inverse problems. We introduce a family of convex optimization problems with regularized saddle point problems as the constraint and prove its convergence. We discretize the inverse problem by using the finite element approach and prove the convergence of the discrete problems. We offer formulas for the gradient and the Hessian computation. The outcome of detailed numerical computations, carried out using the tissue phantom data, shows the efficacy of the developed framework.

Keywords: saddle point problems, parameter identification, inf-sup/Babuska-Breezi condition, elliptic regularization, modified output least-squares, finite element discretization, elasticity imaging

(Some figures may appear in colour only in the online journal)

[†] Dedicated to the blessed memory of Prof. Paul Wenger.

⁴ Author to whom any correspondence should be addressed.

1. Introduction

Let V and Q be real Hilbert spaces with V^* and Q^* as their dual spaces, B be a real Banach space, and A be a nonempty, closed, convex, and bounded subset of B . In the applications that we have in mind, the typical choices are $B := L^\infty(\Omega)$ and $A := \{\ell \in B \mid 0 < a_0 \leq \ell(x) \leq a_1 < \infty\}$, where a_0 and a_1 are constants. Let $a : B \times V \times V \rightarrow \mathbb{R}$ be a trilinear form which is symmetric in the last two arguments, that is, $a(\cdot, u, v) = a(\cdot, v, u)$, for every $u, v \in V$, $b : V \times Q \rightarrow \mathbb{R}$ be a bilinear form, and $f \in V^*$, and $g \in Q^*$ be given. For a fixed $\ell \in B$, the trilinear form a and the bilinear form b are associated to linear maps $\mathbb{A}_\ell \in \mathcal{L}(V, V^*)$ and $\mathbb{B} \in \mathcal{L}(V, Q^*)$ by the relationships $\langle \mathbb{A}_\ell u, v \rangle = a(\ell, u, v)$ for all $v \in V$, and $\langle \mathbb{B}u, q \rangle = b(u, q)$ for all $q \in Q$.

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

$$a(\ell, u, v) + b(v, p) = f(v), \quad \text{for every } v \in V, \quad (1.1a)$$

$$b(u, q) = g(q), \quad \text{for every } q \in Q. \quad (1.1b)$$

Saddle point problem (1.1), which constitutes the direct problem in this work, provides a convenient mathematical framework for studying numerous applied models and has been studied in great detail, from a theoretical as well as a numerical viewpoint. Our focus, however, is on the inverse problem of identifying the parameter ℓ from a measurement of the solution of (1.1). To discuss the solvability of (1.1), we set $L := \{u \in V \mid b(u, q) = 0, \text{ for all } q \in Q\}$, and formulate the conditions:

$$|a(\ell, u, v)| \leq \varsigma_0 \|\ell\|_B \|u\|_V \|v\|_V, \quad \text{for all } \ell \in B, u, v \in V, \varsigma_0 > 0, \quad (1.2a)$$

$$a(\ell, v, v) \geq \varsigma_1 \|v\|_V^2, \quad \text{for all } \ell \in A, v \in L, \varsigma_1 > 0, \quad (1.2b)$$

$$|b(v, q)| \leq \kappa_0 \|v\|_V \|q\|_Q, \quad \text{for all } v \in V, q \in Q, \kappa_0 > 0, \quad (1.2c)$$

$$\inf_{p \in Q} \sup_{u \in V} \frac{b(u, q)}{\|u\|_V \|q\|_Q} \geq \kappa_1, \quad \kappa_1 > 0. \quad (1.2d)$$

It is known that under conditions (1.2c) and (1.2d), for a fixed $\ell \in A$, saddle point problem (1.1) has a unique solution (u, p) which depends continuously on the data (f, g) , see [1].

We note that for saddle point problems emerging from applied models, continuity conditions (1.2a) and (1.2c) are often easy to verify and the coercivity condition (1.2b) also holds in many applications. Condition (1.2d), commonly termed as the Babuska–Brezzi or the inf-sup condition, is a natural substitute for the coercivity condition and plays an essential role in the theoretical as well as the numerical treatment of saddle point problems. In finite element discretization of (1.1), a discrete analog of (1.2d), which is obtained by replacing V and Q , by finite-dimensional subspaces V_h and Q_h , is used. The discrete inf-sup condition imposes compatibility restrictions on the choices of finite-dimensional subspaces of V and Q . Furthermore, there are situations when either the inf-sup condition cannot be verified or is not completely satisfactory from an analysis viewpoint. Interesting works on the use of (1.2d), its variants, limitations, and extensions can be found in [2–6], and the cited references. We also refer to the recent work [7] where a saddle point framework was employed to solve an inverse problem.

System (1.1) is related to the following linearly constrained optimization problem:

$$\min_{u \in V} \mathbb{J}(u) := \frac{1}{2}a(\ell, u, u) - f(u), \quad \text{subject to } b(u, q) = g(q), \quad \text{for every } q \in Q. \quad (1.3)$$

Indeed, if we define the Lagrangian $L : V \times Q \mapsto \mathbb{R}$ by

$$L(u, p) := \frac{1}{2}a(\ell, u, u) - f(u) + b(u, p) - g(p),$$

then under (1.2a)–(1.2d), $L(u, p)$ has a unique saddle point which solves (1.1); see Braess ([8], p 132).

The use of Lagrange multipliers, although it avoids the explicit constraints, poses additional challenges. In particular, increases the dimension from u to (u, p) , and the discrete optimality system is not positive definite, and hence, one has to resort to specialized numerical techniques.

A commonly used technique in optimization is to consider the regularized Lagrangian given by

$$L_\varepsilon(u, p) := \frac{1}{2}a(\ell, u, u) - f(u) + b(u, p) - g(p) - \frac{\varepsilon}{2}\|p\|^2, \quad \varepsilon > 0,$$

which is also an unconstrained optimization problem, however, quadratic in both u and p .

Another commonly adopted viewpoint is to consider the following penalized version of (1.3):

$$\min_{u \in V} P_\varepsilon(u) := \frac{1}{2}a(\ell, u, u) - f(u) - \frac{1}{2\varepsilon}\|Bu - g\|^2, \quad \varepsilon > 0.$$

These perturbation techniques have led to investigating the regularized saddle point problem: given $\varepsilon > 0$ and a coercive and continuous bilinear form $c : Q \times Q \rightarrow \mathbb{R}$, find $(u_\varepsilon, p_\varepsilon) \in V \times Q$ such that

$$a(\ell, u_\varepsilon, v) + b(v, p_\varepsilon) = f(v), \quad \text{for every } v \in V, \quad (1.4a)$$

$$b(u_\varepsilon, q) - \varepsilon c(p_\varepsilon, q) = g(q), \quad \text{for every } q \in Q. \quad (1.4b)$$

Regularized saddle point problem (1.4) system has some advantages in comparison to (1.1) and has been explored extensively, see [1, 3, 9]. A typical result is that under conditions (1.2a)–(1.2d), the sequence of the regularized solutions $\{(u_\varepsilon, p_\varepsilon)\}$ converges to the solution of (1.1), as $\varepsilon \rightarrow 0$. Variants of these conditions have been used to give error estimates for the regularized solutions and to gauge their impact on the finite-dimensional discretization. The regularized saddle point problem has also been studied to mitigate the role of the inf-sup condition; see [5].

This work is partly motivated by the elasticity imaging inverse problem of identifying the cancerous tumor, which generalizes the practice of palpation by making use of varying elastic properties of healthy and cancerous tissue to locate tumors. The idea is to apply a relatively small external quasi-static compression force to the tissue, and then measure the tissue's axial displacement field either directly or indirectly through the comparison of the undeformed and deformed images (see [10–22]). A tumor is then located by the inverse problem of determining the tissue's elastic features from this measurement. The following system of partial

differential equations which describes the response of an isotropic elastic object to known body forces and boundary traction is central to the elasticity imaging inverse problem:

$$-\nabla \cdot \sigma = f \quad \text{in } \Omega, \quad (1.5a)$$

$$\sigma = 2\mu\epsilon(u) + \lambda \operatorname{div} u I, \quad (1.5b)$$

$$u = 0 \quad \text{on } \Gamma_1, \quad (1.5c)$$

$$\sigma n = h \quad \text{on } \Gamma_2. \quad (1.5d)$$

Here the domain Ω is a subset of $\mathbb{R}^2$ with $\partial\Omega = \Gamma_1 \cup \Gamma_2$ as its boundary, the vector-valued function $u = u(x)$ is the displacement of the elastic body, 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 (1.5b) is obtained under the condition that the elastic body is isotropic, and the displacement is sufficiently small so that a linear relationship remains valid. Here μ and λ are the Lamé parameters that quantify the elastic properties of the object.

In this study, we will treat the elastic object as incompressible. The incompressibility can be understood from the relationship $\lambda := \frac{2\nu\mu}{1-2\nu}$, where ν is the Poisson's ratio. If $\nu \approx 0.5$, λ is large, and the elastic object is called nearly incompressible. If $\nu \rightarrow \frac{1}{2}$, the elastic object is called fully incompressible. In this case, the relationship (1.5a) is no more defined, and an alternative, the Lagrangian formulation, is obtained using the explicit incompressibility constraint, see Hughes [23].

Taking $Q = L^2(\Omega)$, and $V = \{v = (v_1, v_2) \in H^1(\Omega) \times H^1(\Omega) : v = 0 \text{ on } \Gamma_1\}$, the variational formulation of (1.5) in the incompressible case reads: find $(u, p) \in V \times Q$ such that

$$\int_{\Omega} 2\mu\epsilon(u) \cdot \epsilon(v) + \int_{\Omega} p(\operatorname{div} v) = \int_{\Omega} f v + \int_{\Gamma_2} v h, \quad \text{for every } v \in V, \quad (1.6a)$$

$$\int_{\Omega} (\operatorname{div} u) q = 0, \quad \text{for every } q \in Q, \quad (1.6b)$$

which corresponds to the saddle point problem (1.1) by taking

$$a(\mu, u, v) = \int_{\Omega} 2\mu\epsilon(u) \cdot \epsilon(v), \quad b(u, q) = \int_{\Omega} p \operatorname{div} u,$$

where $\mu = \mu(x)$ is the sought parameter. We note that inhomogeneous boundary conditions can be incorporated by using the natural data shifting technique.

Note that if the body is nearly incompressible ($\lambda \gg \mu$), then the mixed formulation reads

$$\int_{\Omega} 2\mu\epsilon(u) \cdot \epsilon(v) + \int_{\Omega} p(\operatorname{div} v) = \int_{\Omega} f v + \int_{\Gamma_2} v h, \quad \text{for every } v \in V, \quad (1.7a)$$

$$\int_{\Omega} (\operatorname{div} u) q - \int_{\Omega} \frac{1}{\lambda} p q = 0, \quad \text{for every } q \in Q, \quad (1.7b)$$

which corresponds to the regularized saddle point problem (1.4) with $\varepsilon := \frac{1}{\lambda}$ and $c(p, q) := \int_{\Omega} p q$.

We study the inverse problem by formulating an identification problem in an abstract saddle point problem. This generality significantly enhances the reachability of the results beyond the elasticity imaging inverse problem. We explore the inverse problem as an optimization

problem by using a new modified output least-squares functional (MOLS). The key idea is to consider a family of optimization problems given through convex MOLS functionals with a regularized saddle point problem as the constraint. We prove that under suitable conditions, the optimization problems with the regularized constraints converge to the optimization problem with the original saddle point problem as the constraint. We also study the differentiability properties of the regularized parameter-to-solution map and give necessary and sufficient optimality conditions. It turns out that the derivative of the MOLS objective only includes the trilinear form a . Consequently, as a byproduct of regularizing the saddle point problem, one can eliminate the variable p from the system, which significantly reduces the computational cost. This feature of the developed framework is particularly relevant for the elasticity imaging inverse problems as there as the clinical/experimental data is only available for u but not for p . We also discretize the direct problem and the inverse problem by using the finite element approach and provide the convergence of the discrete optimization problems to the continuous one. We provide discrete formulas for the gradient and the Hessian computation. We offer detailed numerical examples to show the efficacy of the developed framework. To fully understand the usefulness of the developed regularization framework, we consider an analytic example and observe the role of the diminishing regularization parameter. Obtained numerical results are in full compliance with the theoretical results on the regularized solutions. We also give detailed numerical experimentation on a phantom data mimicking the behavior of the tissue closely. The numerical results on the real data are quite encouraging, and the inclusion in the phantom is identified satisfactorily. We note that the regularized saddle point problem corresponds to the mixed variational formulation corresponding to the nearly incompressible elasticity system. Therefore, the developed framework has an appealing physical interpretation of identifying the tumor in the incompressible medium as a limit of a family of identification problems in a nearly incompressible medium.

The contents of this paper are organized into six sections. Section 2 contains the problem formulation, regularization process, convergence analysis, and optimality conditions. Section 3 presents the discretization framework and the convergence of the discrete problems to the continuous one. Section 4 gives the discrete formulas, and the numerical results are provided in section 5. The paper concludes with some remarks on future work.

2. Identification by the modified output least squares

The nonlinear inverse problem of parameter identification is ill-posed in most usual spaces, and some regularization is necessary for a stable identification process, see [24–37]. For regularization, we assume that the set A of feasible coefficients also belongs to a Hilbert space H that is compactly embedded in the space B . An example is $B = L^\infty(\Omega)$ and $H = H^2(\Omega)$, for $\Omega \subset \mathbb{R}^2$. The compactness assumption is to facilitate the abstract framework; however, for specific problems, it is often sufficient for the identification of smooth coefficients to regularize via the $H^1(\Omega)$ -norm.

For any $\ell \in A$, denoting by $(u(\ell), p(\ell))$, a solution of saddle point problem (1.1) and by $(\bar{z}, \hat{z}) \in V \times Q$, the measured data, we define the modified output least-squares (MOLS) functional by

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

and consider the following regularized MOLS-based constrained optimization problem:

$$\min_{\ell \in A} J_\kappa(\ell) := \frac{1}{2} a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, p(\ell) - \hat{z}) + \kappa \|\ell\|_H^2, \quad (2.1)$$

where $\kappa \geq 0$ is the regularization parameter and $\|\cdot\|_H^2$ is the quadratic regularizer.

To highlight some of the features of the MOLS functional, we recall the commonly used output least-squares functional

$$\hat{J}(\ell) := \frac{1}{2} \|u(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2} \|p(\ell) - \hat{z}\|_{\mathbb{Q}}^2, \quad (2.2)$$

where $(u(\ell), p(\ell))$ is a solution of (1.1) for $\ell \in A$ and $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$ is the given data where $\mathbb{V}$ and $\mathbb{Q}$ are the suitable data spaces. The key idea behind the above OLS approach is to minimize the gap between the computed solutions $(u(\ell), p(\ell))$ of (1.1) and the measured data $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$. Even for simpler identification problems, the OLS functional is nonconvex, in general.

Remark 2.1. Note that the OLS functional is defined using the norm of the data spaces $\mathbb{V}$ and $\mathbb{Q}$ which are typically larger than the spaces of the computed solutions V and Q . On the other hand, the MOLS functional requires that the data belongs to the spaces V and Q . In other words, the MOLS functional imposes regularity restrictions on the data than the OLS objective.

As a prelude to the forthcoming discussion, we begin with the following existence result:

Theorem 2.1. Assume that for each $\ell \in A$, saddle point problem (1.1) is uniquely solvable and the set of all solutions $\{(u(\ell), p(\ell)), \ell \in A\}$ is a bounded set, the trilinear form a satisfies the continuity condition (1.2a), and the bilinear form b satisfies the continuity condition (1.2c). Furthermore, either $\kappa > 0$ or the set A is bounded in H . Then optimization problem (2.1) has a solution.

Proof. Since for each $\ell \in A$, saddle point problem (1.1) has the solution $(u(\ell), p(\ell))$, optimization problem (2.1) is well-defined. Note that for each $\ell \in A$, $J_\kappa(\ell)$ is bounded from below, and hence there is a minimizing sequence $\{\ell_n\}$ in A such that $J_\kappa(\ell_n) \rightarrow \inf\{J_\kappa(\ell), \ell \in A\}$, as $n \rightarrow \infty$. If $\kappa = 0$, then the minimizing sequence $\{\ell_n\}$ is bounded in H by the assumption. On the other hand, if $\kappa > 0$, then the minimizing sequence $\{\ell_n\}$ is bounded by the definition of the MOLS objective. Therefore, by the compact embedding of H into B , there is a subsequence converging in $\|\cdot\|_B$. By keeping the same notations for subsequences as well, let $\{\ell_n\}$ be the subsequence converging to some $\bar{\ell} \in A$. Let $\{(u_n, p_n)\}$ be the corresponding sequence of solutions of (1.1), which by assumption remains bounded, and hence 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 $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$. By the definition of (ℓ_n, u_n, p_n) , we have

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

and after a rearrangement, we obtain

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

which, due to (1.2c) and (1.2a), when passed to the limit $n \rightarrow \infty$, yields

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

proving that $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$.

We now claim that for $\ell_n \rightarrow \bar{\ell}$, $u_n := u_n(\ell_n) \rightharpoonup u(\bar{\ell})$, and $p_n := p_n(\ell_n) \rightharpoonup p(\bar{\ell})$, we have

$$\frac{1}{2}a(\ell_n, u_n - \bar{z}, u_n - \bar{z}) + b(u_n - \bar{z}, p_n - \hat{z}) \mapsto \frac{1}{2}a(\bar{\ell}, \bar{u}(\bar{\ell}) - \bar{z}, \bar{u}(\bar{\ell}) - \bar{z}) + b(\bar{u}(\bar{\ell}) - \bar{z}, p(\bar{\ell}) - \hat{z}). \quad (2.3)$$

We set $v = u_n - \bar{z}$ and $q = p_n - \hat{z}$ in the saddle point problem defining (ℓ_n, u_n, p_n) and obtain

$$\begin{aligned} a(\ell_n, u_n - \bar{z}, u_n - \bar{z}) + 2b(u_n - \bar{z}, p_n - \hat{z}) &= -a(\ell_n, \bar{z}, u_n - \bar{z}) + f(u_n - \bar{z}) \\ &\quad + g(p_n - \hat{z}) - b(u_n - \bar{z}, \hat{z}) - b(\bar{z}, p_n - \hat{z}), \end{aligned}$$

and since the right-hand side of the above identity converges to $a(\bar{\ell}, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + 2b(\bar{u} - \bar{z}, \bar{p} - \hat{z})$, we obtain (2.3). Using (2.3) and the weak lower semicontinuity of any norm, we conclude that $\bar{\ell} \in A$ is a minimizer. The proof is complete. $\square$

Remark 2.2. We note that the commonly used functionals, such as the output least-squares or the equation error functional (see [17]) are bounded below by zero because they are defined by using a norm. The MOLS objective for the scalar PDEs is also bounded below by zero because the energy norm defines it, see [32]. On the other hand, the MOLS objective for saddle point problem (1.1) emerges from adding the two equations (1.1a) and (1.1b) and is not necessarily bounded below by zero. However, due to the continuity assumption (2.8), which involves $\|\ell\|_B$, and the boundedness of the set A of feasible parameters in B , the MOLS functional remains bounded below, but in general not by zero. We also note that if we subtract the two equations (1.1a) and (1.1b), and then define an objective functional, the resulting function would be bounded below by zero.

We have the following corollary:

Corollary 2.3. Assume that the conditions given in (1.2) hold. Furthermore, either $\kappa > 0$ or the set A is bounded in H . Then optimization problem (2.1) has a solution.

Proof. Since under (1.2), for each $\ell \in A$, saddle point problem (1.1) is uniquely solvable and the set of all solutions is a bounded set, the proof follows from theorem 2.1. $\square$

We will approximate a solution of (2.1) by a family of regularized MOLS-based optimization problems where the constraint is a regularized saddle point problem. The regularization of the saddle point problem allows us to consider data that is contaminated by some noise in the sense described below. Let ε , δ , and ν be positive reals and $f_\nu \in V^*$, $g_\nu \in Q^*$, and $(\bar{z}_\delta, \hat{z}_\delta) \in V \times Q$ be the noisy data satisfying the following conditions:

$$\max \{ \|f_\nu - f\|_{V^*}, \|g_\nu - g\|_{Q^*} \} \leq \nu, \quad (2.4)$$

$$\max \{ \|\bar{z}_\delta - \bar{z}\|_V, \|\hat{z}_\delta - \hat{z}\|_Q \} \leq \delta. \quad (2.5)$$

$$\left\{ \varepsilon, \nu, \delta, \frac{\delta}{\varepsilon}, \frac{\nu}{\varepsilon} \right\} \rightarrow 0. \quad (2.6)$$

We consider the following family of regularized saddle point problems: given the regularization parameter $\varepsilon > 0$ and $\ell \in A$, find $(u_\sigma, p_\sigma) \equiv (u_\sigma(\ell), p_\sigma(\ell)) \in V \times Q$ such that

$$a(\ell, u_\sigma, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma) = f_\nu(v), \quad \text{for every } v \in V, \quad (2.7a)$$

$$b(u_\sigma, q) - \varepsilon c(p_\sigma, q) = g_\nu(q), \quad \text{for every } q \in Q, \quad (2.7b)$$

where $s : V \times V \mapsto \mathbb{R}$ and $c : Q \times Q \mapsto \mathbb{R}$ are continuous and coercive bilinear forms and $\sigma \equiv (\varepsilon, \nu)$ represents the dependence on the noise in the data and the regularization parameter.

In the following, we only assume that the bilinear form b is continuous and the trilinear form a is continuous and positive, that is,

$$a(\ell, v, v) \geq 0, \quad \text{for all } \ell \in A, v \in V. \quad (2.8)$$

Then, by the Lax–Milgram lemma, for a fixed σ , and for an arbitrary but fixed $\ell \in A$, regularized saddle point problem (2.7) has a unique solution $(u_\sigma(\ell), p_\sigma(\ell))$.

Our objective is to approximate (2.1) by the following family of regularized MOLS-based optimization problems: given $\tau \equiv (\kappa, \varepsilon, \nu, \delta)$, find $\ell_\tau \in A$ by solving

$$\begin{aligned} \min_{\ell \in A} J_\tau(\ell) := & \frac{1}{2} a(\ell, u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta) + b(u_\sigma(\ell) - \bar{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta) \\ & + \frac{\varepsilon}{2} s(u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta) - \frac{\varepsilon}{2} c(p_\sigma(\ell) - \hat{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta) + \kappa \|\ell\|_H^2, \end{aligned} \quad (2.9)$$

where $(u_\sigma(\ell), p_\sigma(\ell))$ is the unique solution of the regularized saddle point problem (2.7).

The following is the main convergence result:

Theorem 2.2. *Assume that for each $\ell \in A$, saddle point problem (1.1) is uniquely solvable, the set of all solutions $\{(u(\ell), p(\ell)), \ell \in A\}$ is a bounded set, the trilinear form a satisfies the continuity condition (1.2a) and the positivity (2.8), the bilinear form b satisfies the continuity condition (1.2c), and the set A is bounded in H . Then, for each parameter τ , regularized optimization problem (2.9) has a solution ℓ_τ . Moreover, there is a subsequence $\{\ell_\tau\}$ converging in $\|\cdot\|_B$ to a solution of (2.1) (with $\kappa = 0$), as $\tau \rightarrow 0$ in the sense of (2.6), that is, as $\{\varepsilon, \nu, \delta, \frac{\delta}{\varepsilon}, \frac{\nu}{\varepsilon}\} \rightarrow 0$.*

Proof. By the arguments used in the proof of theorem 2.1, it follows that for a fixed $\tau \equiv (\kappa, \varepsilon, \nu, \delta)$, (2.9) has a solution. In fact, it suffices to notice that for any fixed τ , for all $\ell \in A$, the regularized saddle point problem (2.7) is uniquely solvable and the solutions are bounded (dependent on τ).

We claim that the sequence $\{\ell_\tau\}$ of solutions of (2.9) and the regularized solutions $\{(u_\sigma(\ell_\tau), p_\sigma(\ell_\tau))\}$ of (2.7) are uniformly bounded. Indeed the sequence $\{\ell_\tau\}$ is bounded since A is bounded in H .

For the boundedness of the sequences of regularized solutions $\{(u_\sigma, p_\sigma)\}$, we note that

$$a(\ell_\tau, u_\sigma, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma) = f_\nu(v), \quad \text{for every } v \in V,$$

$$b(u_\sigma, q) - \varepsilon c(p_\sigma, q) = g_\nu(q), \quad \text{for every } q \in Q.$$

We will use the assumption that for every $\ell \in A$, saddle point problem (1.1) is uniquely solvable. For $\ell_\tau \in A$, let $(\tilde{u}_\tau, \tilde{p}_\tau)$ be the solution of (1.1). Since the set of solutions is bounded by assumption, the sequence $\{(\tilde{u}_\tau, \tilde{p}_\tau)\}$ is bounded. Moreover, we have

$$\begin{aligned} a(\ell_\tau, \tilde{u}_\tau, v) + b(v, \tilde{p}_\tau) &= f(v), \quad \text{for every } v \in V, \\ b(\tilde{u}_\tau, q) &= g(q), \quad \text{for every } q \in Q. \end{aligned}$$

We combine the above two saddle point problems and rearrange them to obtain

$$\begin{aligned} a(\ell_\tau, u_\sigma - \tilde{u}_\tau, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma - \tilde{p}_\tau) &= (f_\nu - f)(v), \quad \text{for every } v \in V, \\ b(u_\sigma - \tilde{u}_\tau, q) - \varepsilon c(p_\sigma, q) &= (g_\nu - g)(q), \quad \text{for every } q \in Q. \end{aligned}$$

We set $v = u_\sigma - \tilde{u}_\tau$, $q = p_\sigma - \tilde{p}_\tau$ in the above system, combine the resulting equations, and using $a(\ell_\tau, \tilde{u}_\tau - u_\sigma, \tilde{u}_\tau - u_\sigma) \geq 0$, obtain (with m_1, m_2, L_1, L_2 as modulus of coercivity, continuity of s, c)

$$\begin{aligned} \varepsilon m_1 \|u_\sigma\|_V^2 + \varepsilon m_2 \|p_\sigma\|_Q^2 &\leq \varepsilon s(u_\sigma, \tilde{u}_\tau) + \varepsilon c(p_\sigma, \tilde{p}_\tau) + (f_\nu - f)(u_\sigma - \tilde{u}_\tau) + (g - g_\nu)(p_\sigma - \tilde{p}_\tau) \\ &\leq \varepsilon L_1 \|u_\sigma\|_V \|\tilde{u}_\tau\|_V + \varepsilon L_2 \|p_\sigma\|_Q \|\tilde{p}_\tau\|_Q + \nu [\|p_\sigma - \tilde{p}_\tau\|_Q + \|\tilde{u}_\tau - u_\sigma\|_V], \end{aligned}$$

which, by taking $m = \min(m_1, m_2)$, further results in

$$\|u_\sigma\|_V^2 + \|p_\sigma\|_Q^2 \leq \|u_\sigma\|_V \left[\frac{L_1}{m} \|\tilde{u}_\tau\|_V + \frac{\nu}{\varepsilon} \right] + \|p_\sigma\|_Q \left[\frac{L_2}{m} \|\tilde{p}_\tau\|_Q + \frac{\nu}{\varepsilon} \right] + \frac{\nu}{\varepsilon} \|\tilde{u}_\tau\|_V + \frac{\nu}{\varepsilon} \|\tilde{p}_\tau\|_Q, \quad (2.10)$$

and consequently, we have

$$\|u_\sigma\|_V^2 + \|p_\sigma\|_Q^2 \leq K_1 \|u_\sigma\|_V + K_2 \|p_\sigma\|_Q + K_3,$$

where

$$\begin{aligned} K_1 &:= \max_\sigma \left\{ \frac{L_1}{m} \|\tilde{u}_\tau\|_V + \frac{\nu}{\varepsilon} \right\}, \\ K_2 &:= \max_\sigma \left\{ \frac{L_2}{m} \|\tilde{p}_\tau\|_Q + \frac{\nu}{\varepsilon} \right\}, \\ K_3 &:= \max_\sigma \left\{ \frac{\nu}{\varepsilon} \|\tilde{u}_\tau\|_V + \frac{\nu}{\varepsilon} \|\tilde{p}_\tau\|_Q \right\}, \end{aligned}$$

are positive constants. We therefore conclude that the sequence $\{(u_\sigma, p_\sigma)\}$ is bounded.

Since the regularized solutions are bounded, we can extract a subsequence $\{(u_\sigma, p_\sigma)\}$ converging weakly to some $(\bar{u}, \bar{p})$. Recalling that $\ell_\tau \rightarrow \bar{\ell}$, we claim that $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$.

Since ℓ_τ is a solution of (2.9), we have

$$\begin{aligned} a(\ell_\tau, u_\sigma, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma) &= f_\nu(v), \quad \text{for every } v \in V, \\ b(u_\sigma, q) - \varepsilon c(p_\sigma, q) &= g_\nu(q), \quad \text{for every } q \in Q, \end{aligned}$$

or equivalently,

$$\begin{aligned} a(\ell_\tau - \bar{\ell}, u_\sigma, v) + a(\bar{\ell}, u_\sigma - \bar{u}, v) + a(\bar{\ell}, \bar{u}, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma) &= f_\nu(v), \quad \text{for every } v \in V, \\ b(u_\sigma, q) - \varepsilon c(p_\sigma, q) &= g_\nu(q), \quad \text{for every } q \in Q, \end{aligned}$$

which due to the imposed conditions, when passed to the limit $\sigma \rightarrow 0$, implies that

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

confirming that $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$.

We will next show that for $\ell_\tau \rightarrow \bar{\ell}$, $u_\sigma := u_\sigma(\ell_\tau) \rightarrow u(\bar{\ell})$, and $p_\sigma := p_\sigma(\ell_\tau) \rightarrow p(\bar{\ell})$, we have

$$\begin{aligned} \frac{1}{2}a(\ell_\tau, u_\sigma(\ell_\tau) - \bar{z}_\delta, u_\sigma(\ell_\tau) - \bar{z}_\delta) + b(u_\sigma(\ell_\tau) - \bar{z}_\delta, p_\sigma(\ell_\tau) - \hat{z}_\delta) + \frac{1}{2}\varepsilon s(u_\sigma(\ell_\tau) - \bar{z}_\delta, u_\sigma(\ell_\tau) - \bar{z}_\delta) \\ - \frac{\varepsilon}{2}c(p_\sigma(\ell_\tau) - \hat{z}_\delta, p_\sigma(\ell_\tau) - \hat{z}_\delta) \\ \mapsto \frac{1}{2}a(\bar{\ell}, \bar{u}(\bar{\ell}) - \bar{z}, \bar{u}(\bar{\ell}) - \bar{z}) + b(\bar{u}(\bar{\ell}) - \bar{z}, p(\bar{\ell}) - \hat{z}). \end{aligned} \quad (2.11)$$

For this, we note that for every $(v, q) \in V \times Q$, we have

$$\begin{aligned} a(\ell_\tau, u_\sigma, v) + \varepsilon s(u_\sigma, v) + b(v, p_\sigma) &= f_\nu(v), \quad \text{for every } v \in V, \\ b(u_\sigma, q) - \varepsilon c(p_\sigma, q) &= g_\nu(q), \quad \text{for every } q \in Q, \end{aligned}$$

which for the choice $v = u_\sigma - \bar{z}_\delta$, and $q = p_\sigma - \hat{z}_\delta$, leads to the following identity:

$$\begin{aligned} a(\ell_\tau, u_\sigma - \bar{z}_\delta, u_\sigma - \bar{z}_\delta) + 2b(u_\sigma - \bar{z}_\delta, p_\sigma - \hat{z}_\delta) + \varepsilon s(u_\sigma - \bar{z}_\delta, u_\sigma - \bar{z}_\delta) - \varepsilon c(p_\sigma - \hat{z}_\delta, p_\sigma - \hat{z}_\delta) \\ = -a(\ell_\tau, \bar{z}_\delta, u_\sigma - \bar{z}_\delta) + f(u_\sigma - \bar{z}_\delta) + g(p_\sigma - \hat{z}_\delta) - b(u_\sigma - \bar{z}_\delta, \hat{z}_\delta) + f_\nu(u_\sigma - \bar{z}_\delta) \\ - f(u_\sigma - \bar{z}_\delta) - b(\bar{z}_\delta, p_\sigma - \hat{z}_\delta) + g_\nu(p_\sigma - \hat{z}_\delta) - g(p_\sigma - \hat{z}_\delta), \end{aligned}$$

and because the right-hand side of the above identity converges to

$$a(\bar{\ell}, -\bar{z}, \bar{u} - \bar{z}) + b(\bar{u} - \bar{z}, -\hat{z}) + b(-\bar{z}, \bar{p} - \hat{z}) + f(\bar{u} - \bar{z}) + g(\bar{p} - \hat{z}),$$

which, due to the fact that $(\bar{u}, \bar{p})$ is a solution of (1.1) equals to $a(\bar{\ell}, \bar{u} - \bar{z}, \bar{u} - \bar{z}) + 2b(\bar{u} - \bar{z}, \bar{p} - \hat{z})$, results in the desired convergence (2.11).

Let $\hat{\ell}$ be a solution of (2.1) (with $\kappa = 0$), and $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ be the corresponding solution of saddle point problem (1.1). We consider the regularized solutions $\{(u_\sigma(\hat{\ell}), p_\sigma(\hat{\ell}))\}$ for the coefficient $\hat{\ell}$. That is,

$$a(\hat{\ell}, u_\sigma(\hat{\ell}), v) + \varepsilon s(u_\sigma(\hat{\ell}), v) + b(v, p_\sigma(\hat{\ell})) = f_\nu(v), \quad \text{for every } v \in V, \quad (2.12a)$$

$$b(u_\sigma(\hat{\ell}), q) - \varepsilon c(p_\sigma(\hat{\ell}), q) = g_\nu(q), \quad \text{for every } q \in Q. \quad (2.12b)$$

By the arguments used above, we can show that $\{(u_\sigma(\hat{\ell}), p_\sigma(\hat{\ell}))\}$ converges weakly to $(\hat{u}, \hat{p})$. In fact, the convergence is strong as we show next. We rearrange the above system as follows:

$$\begin{aligned} \varepsilon m_1 \|u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell})\|_V^2 + m_2 \varepsilon \|p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell})\|_Q^2 \\ = \varepsilon s(u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell}), u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell})) + \varepsilon c(p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell}), p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell})) \\ \leq (f_\nu - f)(u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell})) - (g_\nu - g)(p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell})) \\ + \varepsilon \langle \hat{u}(\hat{\ell}), u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell}) \rangle_V + \varepsilon \langle \hat{p}(\hat{\ell}), p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell}) \rangle_Q, \end{aligned}$$

which after a simple calculation implies that

$$\lim_{n \rightarrow \infty} \|u_\sigma(\hat{\ell}) - \hat{u}(\hat{\ell})\|_V^2 + \lim_{n \rightarrow \infty} \|p_\sigma(\hat{\ell}) - \hat{p}(\hat{\ell})\|_Q^2 \leq 0,$$

and consequently the strong convergence of $\{(u_\sigma(\hat{\ell}), p_\sigma(\hat{\ell}))\}$ to $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ follows.

Summarizing the above observations, we have

$$\begin{aligned} J(\bar{\ell}) &= \frac{1}{2} a(\bar{\ell}, \bar{u}(\bar{\ell}) - \bar{z}, \bar{u}(\bar{\ell}) - \bar{z}) + b(\bar{u}(\bar{\ell}) - \bar{z}, \bar{p}(\bar{\ell}) - \bar{z}) \\ &\leq \lim_{n \rightarrow \infty} \left\{ \frac{1}{2} a(\ell_\tau, u_\sigma(\ell_\tau) - \bar{z}_\delta, u_\sigma(\ell_\tau) - \bar{z}_\delta) + b(u_\sigma(\ell_\tau) - \bar{z}_\delta, p_\sigma(\ell_\tau) - \bar{z}_\delta) \right. \\ &\quad \left. + \frac{1}{2} \varepsilon s(u_\sigma(\ell_\tau) - \bar{z}_\delta, u_\sigma(\ell_\tau) - \bar{z}_\delta) - \frac{\varepsilon}{2} c(p_\sigma(\ell_\tau) - \bar{z}_\delta, p_\sigma(\ell_\tau) - \bar{z}_\delta) \right\} + \liminf_{n \rightarrow \infty} \kappa \|\ell_\tau\|_H^2 \\ &\leq \lim_{n \rightarrow \infty} \left\{ \frac{1}{2} a(\hat{\ell}, u_\sigma(\hat{\ell}) - \bar{z}_\delta, u_\sigma(\hat{\ell}) - \bar{z}_\delta) + b(u_\sigma(\hat{\ell}) - \bar{z}_\delta, p_\sigma(\hat{\ell}) - \bar{z}_\delta) \right. \\ &\quad \left. + \frac{1}{2} \varepsilon s(u_\sigma(\hat{\ell}) - \bar{z}_\delta, u_\sigma(\hat{\ell}) - \bar{z}_\delta) - \frac{\varepsilon}{2} c(p_\sigma(\hat{\ell}) - \bar{z}_\delta, p_\sigma(\hat{\ell}) - \bar{z}_\delta) + \kappa \|\hat{\ell}\|_H^2 \right\} \\ &\leq \frac{1}{2} a(\hat{\ell}, \hat{u}(\hat{\ell}) - \bar{z}, \hat{u}(\hat{\ell}) - \bar{z}) + b(\hat{u}(\hat{\ell}) - \bar{z}, \hat{p}(\hat{\ell}) - \bar{z}), \end{aligned}$$

and since $\hat{\ell} \in A$ was chosen arbitrarily, we have established that $\bar{\ell} \in A$ is a minimizer of (2.1). $\square$

To derive optimality conditions for (2.1), the following result that sheds some light on the smoothness of the regularized parameter-to-solution map will play a key role

Theorem 2.3. *For a fixed σ , let ℓ be in the interior of A which we assume to be nonempty. The first-order derivative $(Du_\sigma(\ell)\delta\ell, Dp_\sigma(\ell)\delta\ell)$ of the regularized parameter-to-solution map $\ell \rightarrow (u_\sigma(\ell), p_\sigma(\ell))$ at ℓ in the direction $\delta\ell \in B$ is the unique solution of the regularized saddle point problem:*

$$a(\ell, Du_\sigma(\ell)\delta\ell, v) + \varepsilon s(Du_\sigma(\ell)\delta\ell, v) + b(v, Dp_\sigma(\ell)\delta\ell) = -a(\delta\ell, u_\sigma(\ell), v), \quad \text{for every } v \in V, \quad (2.13a)$$

$$b(Du_\sigma(\ell)\delta\ell, q) - \varepsilon c(Dp_\sigma(\ell)\delta\ell, q) = 0, \quad \text{for every } q \in Q. \quad (2.13b)$$

Moreover, the second-order derivative $(D^2u_\sigma(\ell)(\delta\ell_1, \delta\ell_2), D^2p_\sigma(\ell)(\delta\ell_1, \delta\ell_2))$ of $(u_\sigma(\ell), p_\sigma(\ell))$ at ℓ in the direction $(\delta\ell_1, \delta\ell_2) \in B \times B$ is the unique solution of the regularized saddle point problem:

$$\begin{aligned} a(\ell, D^2u_\sigma(\ell)(\delta\ell_1, \delta\ell_2), v) + \varepsilon s(D^2u_\sigma(\ell)(\delta\ell_1, \delta\ell_2), v) + b(v, D^2p_\sigma(\ell)(\delta\ell_1, \delta\ell_2)) \\ = -a(\delta\ell_2, Du_\sigma(\ell)\delta\ell_1, v) - a(\delta\ell_1, Du_\sigma(\ell)\delta\ell_2, v), \quad \text{for every } v \in V, \end{aligned} \quad (2.14a)$$

$$b(D^2u_\sigma(\ell)(\delta\ell_1, \delta\ell_2), q) - \varepsilon c(D^2p_\sigma(\ell)(\delta\ell_1, \delta\ell_2), q) = 0, \quad \text{for every } q \in Q \quad (2.14b)$$

Proof. The proof follows by similar arguments that were used in [32]. $\square$

We again consider the following MOLS objective with perturbed data:

$$J_\varepsilon(\ell) := \frac{1}{2}a(\ell, u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta) + b(u_\sigma(\ell) - \bar{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta) \\ + \frac{\varepsilon}{2}s(u_\sigma(\ell) - \bar{z}_\delta, u_\sigma - \bar{z}_\delta) - \frac{\varepsilon}{2}c(p_\sigma(\ell) - \hat{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta), \quad (2.15)$$

where $(u_\sigma(\ell), p_\sigma(\ell))$ is the unique solution of regularized saddle point problem (2.7).

We have the following result: $\square$

Theorem 2.4. *For each fixed σ , the modified output least-squares functional (2.15) is convex in A .*

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

$$DJ_\varepsilon(\ell)(\hat{\ell}) = \frac{1}{2}a(\hat{\ell}, u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta) + a(\ell, Du_\sigma(\ell)(\hat{\ell}), u_\sigma(\ell) - \bar{z}_\delta) \\ + \varepsilon s(Du_\sigma(\ell)(\hat{\ell}), u_\sigma - \bar{z}_\delta) - \varepsilon c(Dp_\sigma(\ell)(\hat{\ell}), p_\sigma(\ell) - \hat{z}_\delta) \\ + b(Du_\sigma(\ell)(\hat{\ell}), p_\sigma(\ell) - \hat{z}_\delta) + b(u_\sigma(\ell) - \bar{z}_\delta, Dp_\sigma(\ell)(\hat{\ell})). \quad (2.16)$$

Moreover, by using (2.13), we obtain

$$a(\ell, Du_\sigma(\ell)(\hat{\ell}), u_\sigma(\ell) - \bar{z}_\delta) + \varepsilon s(Du_\sigma(\ell)(\hat{\ell}), u_\sigma(\ell) - \bar{z}_\delta) + b(u_\sigma(\ell) - \bar{z}_\delta, Dp_\sigma(\ell)(\hat{\ell})) \\ = -a(\hat{\ell}, u_\sigma(\ell), u_\sigma(\ell) - \bar{z}_\delta), \\ b(Du_\sigma(\ell)(\hat{\ell}), p_\sigma(\ell) - \hat{z}_\delta) - \varepsilon c(Dp_\sigma(\ell)(\hat{\ell}), p_\sigma(\ell) - \hat{z}_\delta) = 0,$$

which, when substituted in (2.16), yields

$$DJ_\varepsilon(\ell)(\hat{\ell}) = \frac{1}{2}a(\hat{\ell}, u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \hat{z}_\delta) - a(\hat{\ell}, u_\sigma(\ell), u_\sigma(\ell) - \bar{z}_\delta) \\ = -\frac{1}{2}a(\hat{\ell}, u_\sigma(\ell) + \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta).$$

Furthermore,

$$D^2J_\varepsilon(\ell)(\hat{\ell}, \hat{\ell}) = -\frac{1}{2}a(\hat{\ell}, u_\sigma(\ell) - \bar{z}_\delta, Du_\sigma(\ell)(\hat{\ell})) - \frac{1}{2}a(\hat{\ell}, u_\sigma(\ell) + \bar{z}_\delta, Du_\sigma(\ell)(\hat{\ell})) \\ = -a(\hat{\ell}, u_\sigma(\ell), Du_\sigma(\ell)(\hat{\ell})) \\ = a(\ell, Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) + \varepsilon s(Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) \\ + b(Dp_\sigma(\ell)(\hat{\ell}), Dp_\sigma(\ell)(\hat{\ell})) \\ = a(\ell, Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) + \varepsilon s(Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) \\ + \varepsilon c(Dp_\sigma(\ell)(\hat{\ell}), Dp_\sigma(\ell)(\hat{\ell})),$$

where we used the following consequence of (2.13):

$$\begin{aligned}
a(\ell, Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) + \varepsilon s(Du_\sigma(\ell)(\hat{\ell}), Du_\sigma(\ell)(\hat{\ell})) + b(Dp_\sigma(\ell)(\hat{\ell}), Dp_\sigma(\ell)(\hat{\ell})) \\
= -a(\hat{\ell}, u_\sigma(\ell), Du_\sigma(\ell)(\hat{\ell})), \\
b(Du_\sigma(\ell)(\hat{\ell}), Dp_\sigma(\ell)(\hat{\ell})) - \varepsilon s(Dp_\sigma(\ell)(\hat{\ell}), Dp_\sigma(\ell)(\hat{\ell})) = 0.
\end{aligned}$$

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

$$D^2 J_\varepsilon(\ell)(\hat{\ell}, \hat{\ell}) \geq \varepsilon m_1 \|Du_\sigma(\ell)(\hat{\ell})\|_V^2 + \varepsilon m_2 \|Dp_\sigma(\ell)(\hat{\ell})\|_Q^2, \quad (2.17)$$

and consequently J_ε is a smooth and convex functional. $\square$

In the following optimality condition, for simplicity we assume that the parameter-to-solution map is defined on a larger set that contains the closed and convex set A in its interior.

Theorem 2.5. *Under the setting of theorem 2.3, a necessary and sufficient optimality condition for an element $\ell_\tau \in A$ to be a solution of regularized optimization problem (2.9) is the variational inequality*

$$-\frac{1}{2}a(\ell - \ell_\tau, u_\sigma + \bar{z}_\delta, u_\sigma - \bar{z}_\delta) + 2\kappa \langle \ell_\tau, \ell - \ell_\tau \rangle_H \geq 0, \quad \text{for every } \ell \in A, \quad (2.18)$$

where $(u_\sigma, p_\sigma) = (u_\sigma(\ell_\tau), p_\sigma(\ell_\tau))$ is the solution of regularized saddle point problem (2.7). If $u_\sigma(\ell_\tau) \rightarrow u(\bar{\ell})$, then the following variational inequality holds:

$$-\frac{1}{2}a(\ell - \bar{\ell}, u(\bar{\ell}) + \bar{z}, u(\bar{\ell}) - \bar{z}) \geq 0, \quad \text{for every } \ell \in A. \quad (2.19)$$

Proof. We notice that due to the convexity of the MOLS functional, a necessary and sufficient optimality condition for ℓ_τ to be a solution of (2.9) is the variational inequality

$$DJ_\varepsilon(\ell_\tau)(\ell - \ell_\tau) + 2\kappa \langle \ell_\tau, \ell - \ell_\tau \rangle_H \geq 0, \quad \text{for every } \ell \in A, \quad (2.20)$$

where J_ε is given in (2.15), that is,

$$\begin{aligned}
J_\varepsilon(\ell) := & \frac{1}{2}a(\ell, u_\sigma(\ell) - \bar{z}_\delta, u_\sigma(\ell) - \bar{z}_\delta) + b(u_\sigma(\ell) - \bar{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta) \\
& + \frac{\varepsilon}{2} \langle u_\sigma - \bar{z}_\delta, u_\sigma - \bar{z}_\delta \rangle_V - \frac{\varepsilon}{2} \langle p_\sigma(\ell) - \hat{z}_\delta, p_\sigma(\ell) - \hat{z}_\delta \rangle_Q.
\end{aligned}$$

Since for any direction $\delta\ell$, we have

$$DJ_\varepsilon(\ell)(\delta\ell) = -\frac{1}{2}a(\delta\ell, u_\sigma + \bar{z}_\delta, u_\sigma - \bar{z}_\delta).$$

We at once obtain (2.18). The variational inequality (2.19) follows by passing the limit in variational inequality (2.20). The proof is complete. $\square$

Remark 2.4. If a is coercive, then s is not needed, and $u_\sigma(\ell_\tau) \rightarrow u(\bar{\ell})$.

3. Finite-dimensional approximation

We will now use a finite-element-based discretization scheme for the inverse problem. We assume that there is a parameter h converging to 0, and families $\{V_h\}$, $\{Q_h\}$, and $\{B_h\}$, of finite-dimensional subspaces of V , Q , and B , respectively. We set $A_h = B_h \cap A$ and assume that $\cap_h A_h \neq \emptyset$. We define a projection operator $P_h = (\bar{P}_h, \hat{P}_h)$ with

$$\|\bar{P}_h v - v\|_V \rightarrow 0, \text{ for every } v \in V, \quad (3.1a)$$

$$\|\hat{P}_h q - q\|_Q \rightarrow 0, \text{ for every } q \in Q. \quad (3.1b)$$

Furthermore, to have an analogous approximation property for the parameter space, we assume that for any $\ell \in A$, there exists a sequence $\{\hat{\ell}_h\} \subset A_h$ such that $\hat{\ell}_h \rightarrow \ell$ in B and $\|\hat{\ell}_h\|_H \rightarrow \|\ell\|_H$.

For simplicity in presentation, we incorporate the data contamination by assuming that κ , ε , δ , and ν are positive parameters with $f_\nu \in V_h^*$, $g_\nu \in Q_h^*$, and $(\bar{z}_\delta, \hat{z}_\delta) \in V_h \times Q_h$ such that

$$\max \{\|f_\nu - f\|_{V^*}, \|g_\nu - g\|_{Q^*}\} \leq \nu, \quad (3.2)$$

$$\max \{\|\bar{z}_\delta - \bar{z}\|_V, \|\hat{z}_\delta - \hat{z}\|_Q\} \leq \delta. \quad (3.3)$$

Moreover, we assume that

$$\left\{ \kappa, \varepsilon, \nu, \delta, \frac{\delta}{\varepsilon}, \frac{\nu}{\varepsilon} \right\} \rightarrow 0. \quad (3.4)$$

Given $\ell_h \in A_h$, the discrete saddle point problem seeks $(u_h, p_h) \in V_h \times Q_h$ such that

$$a(\ell_h, u_h, v_h) + b(v_h, p_h) = f(v_h), \text{ for every } v_h \in V_h, \quad (3.5a)$$

$$b(u_h, q_h) = g(q_h), \text{ for every } q_h \in Q_h, \quad (3.5b)$$

and the discrete regularized saddle point problem seeks $(u_\sigma^h, p_\sigma^h) \in V_h \times Q_h$ with

$$a(\ell_h, u_\sigma^h, v_h) + \varepsilon s(u_\sigma^h, v_h) + b(v_h, p_\sigma^h) = f_\nu(v_h), \text{ for every } v_h \in V_h, \quad (3.6a)$$

$$b(u_\sigma^h, q) - \varepsilon c(p_\sigma^h, q) = g_\nu(q_h), \text{ for every } q_h \in Q_h, \quad (3.6b)$$

where $\sigma \equiv (\varepsilon, \nu)$ represents the influence of the involved parameters.

We also consider discrete MOLS-based minimization problem: find $\ell_\tau^h \in A_h$ by solving

$$\begin{aligned} \min_{\ell_h \in A_h} J_\kappa^h(\ell_h) &= \frac{1}{2} a(\ell_h, u_\sigma^h - \bar{z}_\delta, u_\sigma^h - \bar{z}_\delta) + \frac{\varepsilon}{2} s(u_\sigma^h - \bar{z}_\delta, u_\sigma^h - \bar{z}_\delta) + b(u_\sigma^h - \bar{z}_\delta, p_\sigma^h - \hat{z}_\delta) \\ &\quad - \frac{\varepsilon}{2} c(p_\sigma^h - \hat{z}_\delta, p_\sigma^h - \hat{z}_\delta) + \kappa \|\ell_h\|_H^2, \end{aligned} \quad (3.7)$$

where (u_σ^h, p_σ^h) is the unique solution of (3.6) and $\tau \equiv (\kappa, \varepsilon, \nu, \delta)$.

The following result ensures the convergence of the discrete optimization problem:

Theorem 3.1. Besides the hypotheses of theorem 2.2, assume that for each $\ell_h \in A_h$, (3.5) is uniquely solvable and the set of all solutions is a bounded set. Then, for each (h, τ) , (3.7) has

a minimizer $\ell_\tau^h \in A_h$. Moreover, there is a subsequence $\{\ell_\tau^h\}$ that converges to a solution $\bar{\ell}$ of the optimization problem (2.1).

Proof. For fixed (h, σ) , (3.6) is uniquely solvable, and the minimization problem (3.7) has a solution $\ell_\tau^h \in A_h$ by arguments analogous to those given in theorem 2.2. Since $A_h \subset A$, and since the set A is bounded in H by assumption, the sequences $\{\ell_\tau^h\}$ is bounded in $\|\cdot\|_H$, and by the compact embedding of H into B , there is a subsequence, still denoted by $\{\ell_\tau^h\}$, that converges to some $\bar{\ell} \in A$. Let (u_σ^h, p_σ^h) be the unique solution of the discrete regularized saddle point problem (3.6) that corresponds to coefficient ℓ_τ^h . The boundedness of the sequence $\{(u_\sigma^h, p_\sigma^h)\} \subset V_h \times Q_h$ can be proved exactly as in the previous section. Let $\{(\bar{u}, \bar{p})\}$ be the subsequence converging weakly to some $(\bar{u}, \bar{p}) \in V \times Q$. We will show that $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$, the unique solution of (1.1).

Let $(v, q) \in V \times Q$ be arbitrary. We take $(v_h, q_h) := (\bar{P}_h v, \hat{P}_h q) \in V_h \times Q_h$ in the discrete regularized saddle point problem (3.6), and after rearranging the resulting equations, obtain

$$\begin{aligned} a(\bar{\ell}, u_\sigma^h, v) + \varepsilon_n s(u_h, v_h) + b(v, p_\sigma^h) + a(\ell_\tau^h, u_\sigma^h, v_h - v) + b(v_h - v, p_\sigma^h) + a(\ell_\tau^h - \bar{\ell}, u_\sigma^h, v) &= f_v(v_h), \\ b(u_\sigma^h, q) + b(u_\sigma^h, q_h - q) - \varepsilon c(p_\sigma^h, q_h) &= g_v(q_h). \end{aligned}$$

We pass the above system to the limit $h \rightarrow 0$, $\sigma \rightarrow 0$, and using (3.1), deduce that

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

and consequently $(\bar{u}, \bar{p}) = (u(\bar{\ell}), p(\bar{\ell}))$ is the unique solution of saddle point problem (1.1).

Furthermore, following the argument used in the proof of (2.11), we can show that if $\{\ell_h\} \subset A_h$ converges to some $\ell \in A$ as $h \rightarrow 0$ in B , then

$$\begin{aligned} \lim_{h \rightarrow 0} & \left[\frac{1}{2} a(\ell_h, u_\sigma^h(\ell_h) - \bar{z}_\delta, u_\sigma^h(\ell_h) - \bar{z}_\delta) + \frac{\varepsilon}{2} s(u_\sigma^h(\ell_h) - \bar{z}_\delta, u_\sigma^h(\ell_h) - \bar{z}_\delta) \right. \\ & \left. + b(u_\sigma^h(\ell_h) - \bar{z}_\delta, p_\sigma^h(\ell_h) - \hat{z}_\delta) - \frac{\varepsilon}{2} c(p_\sigma^h(\ell_h) - \hat{z}_\delta, p_\sigma^h(\ell_h) - \hat{z}_\delta) \right] \\ &= \frac{1}{2} a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, p(\ell) - \hat{z}). \end{aligned}$$

Let $\ell \in A$ be arbitrary. By the assumption, there exists a sequence $\{\hat{\ell}_h\} \subset A_h$ such that $\hat{\ell}_h \rightarrow \ell$ in B and $\|\hat{\ell}_h\|_H \rightarrow \|\ell\|_H$. This fact when combined with the above observations, yields

$$\begin{aligned}
J(\bar{\ell}) &= \frac{1}{2}a(\bar{\ell}, u(\bar{\ell}) - \bar{z}, u(\bar{\ell}) - \bar{z}) + b(u(\bar{\ell}) - \bar{z}, p(\bar{\ell}) - \hat{z}) \\
&\leq \lim_{h \rightarrow 0} \left\{ \frac{1}{2}a(\ell_\tau^h, u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta, u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta) + \frac{\varepsilon}{2}s(u_\sigma^h(\ell_\tau^h) - \hat{z}_\delta, u_\sigma^h(\ell_\tau^h) - \hat{z}_\delta) \right. \\
&\quad \left. + b(u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta, p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta) - \frac{\varepsilon}{2}c(p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta, p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta) \right\} + \liminf_{h \rightarrow 0} \kappa \|\ell_\tau^h\|_H^2 \\
&\leq \liminf_{h \rightarrow 0} \left\{ \frac{1}{2}a(\ell_\tau^h, u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta, u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta) + \frac{\varepsilon}{2}s(u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta, u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta) \right. \\
&\quad \left. + b(u_\sigma^h(\ell_\tau^h) - \bar{z}_\delta, p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta) - \frac{\varepsilon}{2}c(p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta, p_\sigma^h(\ell_\tau^h) - \hat{z}_\delta) + \kappa \|\ell_\tau^h\|_H^2 \right\} \\
&\leq \liminf_{h \rightarrow 0} \left\{ \frac{1}{2}a(\hat{\ell}_h, u_\sigma^h(\hat{\ell}_h) - \bar{z}_h, u_\sigma^h(\hat{\ell}_h) - \bar{z}_h) + \frac{\varepsilon}{2}s(u_\sigma^h(\hat{\ell}_h) - \bar{z}_h, u_\sigma^h(\hat{\ell}_h) - \bar{z}_h) \right. \\
&\quad \left. + b(u_\sigma^h(\hat{\ell}_h) - \bar{z}_h, p_\sigma^h(\hat{\ell}_h) - \hat{z}_h) - \frac{\varepsilon}{2}c(p_\sigma^h(\hat{\ell}_h) - \hat{z}_h, p_\sigma^h(\hat{\ell}_h) - \hat{z}_h) + \kappa \|\hat{\ell}_h\|_H^2 \right\} \\
&= \frac{1}{2}a(\ell, u(\ell) - \bar{z}, u(\ell) - \bar{z}) + b(u(\ell) - \bar{z}, p(\ell) - \hat{z}) = J(\ell),
\end{aligned}$$

where we also used the fact that ℓ_τ^h is a minimizer. Since the element $\ell \in A$ was chosen arbitrarily, we obtain that $\bar{\ell}$ is a solution of the continuous optimization problem. This completes the proof. $\square$

Remark 3.1. If a is coercive and b satisfies discrete inf-sup condition, then for each ℓ , discrete saddle point problem (3.5) is uniquely solvable and the set of all solutions is bounded, uniformly in h .

4. Computational framework

We shall now collect discrete formulas for the regularized saddle point problem and the inverse problem. For this, we begin with a triangulation $\mathcal{T}_h$ on the domain Ω . Assume that L_h is the space of all continuous piecewise polynomials of degree d_ℓ relative to $\mathcal{T}_h$, U_h is the space of all continuous piecewise polynomials of degree d_u relative to $\mathcal{T}_h$, and Q_h is the space of all continuous piecewise polynomials of degree d_q relative to $\mathcal{T}_h$. We specify the 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 $L_i = \ell(x_i)$, $i = 1, 2, \dots, m$, 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 $\ell = \sum_{i=1}^m L_i \varphi_i$. Analogously, $u \in U_h$ 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, however, the nodes will be different if $d_\ell \neq d_u \neq d_q$.

Let $S : \mathbb{R}^m \rightarrow \mathbb{R}^{n+k}$ be the finite element solution map that related to any $\ell \in A_h$, the unique solution $(u_\sigma^h, p_\sigma^h) \in U_h \times Q_h$ of (3.6). Then $S_{sn}(L) = U_\sigma^h$, where U_σ^h is given by

$$K_\sigma^h(L)U_\sigma^h = F_\sigma^h, \quad (4.1)$$

with the stiffness matrix $K_\sigma^h(L) \in \mathbb{R}^{(n+k) \times (n+k)}$ and the load vector $F_\sigma^h \in \mathbb{R}^{n+k}$ are given by

$$K_\sigma(L) = \begin{bmatrix} \widehat{K}(L) + \varepsilon S & B^\top \\ B & -\varepsilon C \end{bmatrix}$$

with

$$\begin{aligned} \widehat{K}(L)_{i,j} &= a(\ell, \psi_j, \psi_i), & i, j &= 1, 2, \dots, n, \\ S &= s(\psi_j, \psi_i), & i, j &= 1, 2, \dots, n, \\ B_{i,j} &= b(\psi_j, \chi_i), & i &= 1, 2, \dots, k, \quad j = 1, 2, \dots, n, \\ C_{i,j} &= c(\chi_j, \chi_i), & i, j &= 1, 2, \dots, k, \\ (F_\sigma^h)_i &= f_\nu(\psi_i), & i &= 1, 2, \dots, n, \\ (F_\sigma^h)_j &= g_\nu(\chi_i), & j &= n+1, n+2, \dots, n+k. \end{aligned}$$

In the following, for notational simplicity, in vectors and matrices, we drop the dependence on the parameters. A straightforward calculation leads to the discrete MOLS:

$$\begin{aligned} J(L) &= \frac{1}{2} \begin{bmatrix} U(L) - \bar{Z} \\ P(L) - \widehat{Z} \end{bmatrix}^\top \begin{bmatrix} \widehat{K}(L) + \varepsilon S & B^\top \\ B & -\varepsilon C \end{bmatrix} \begin{bmatrix} U(L) - \bar{Z} \\ P(L) - \widehat{Z} \end{bmatrix} \\ &= \frac{1}{2} (U(L) - \bar{Z})^\top \left(\widehat{K}(L) + \varepsilon S \right) (U(L) - \bar{Z}) + (U(L) - \bar{Z})^\top B^\top (P(L) - \widehat{Z}) \\ &\quad - \frac{\varepsilon}{2} (P(L) - \widehat{Z})^\top C (P(L) - \widehat{Z}). \end{aligned}$$

Denoting by $\mathbb{A}$ an adjoint stiffness matrix defined by the condition that $\widehat{K}(L)V = \mathbb{A}(V)L$, for every $L \in \mathbb{R}^m$, and for every $V \in \mathbb{R}^n$, the gradient and the Hessian are given by

$$\begin{aligned} \nabla J(L) &= -\frac{1}{2} \mathbb{A}(U(L))^\top U(L) + \frac{1}{2} \mathbb{A}(\bar{Z})^\top \bar{Z}, \\ \nabla^2 J(L) &= \nabla U(L)^\top \mathbb{A}(\nabla U(L)) + \varepsilon \nabla U(L)^\top S \nabla U(L) + \varepsilon \nabla P(L)^\top C \nabla P(L). \end{aligned}$$

One of the main advantages of $\mathbb{A}$ is the computational flexibility that it provides by exploiting the relationship between the bases of u and the bases for ℓ in the trilinear form a , which leads to explicit formulas for the discrete MOLS functional and its gradient.

5. Computational experimentation for the elasticity imaging inverse problem

We now present numerical results for the incompressible system (1.6). Since, for this case, a is elliptic, we do not regularize by using the bilinear form s . However, a natural choice for s is $s(u, v) = \langle u, v \rangle_V$, the inner product in the space V . For the bilinear form c , we use $c(p, q) := \int_\Omega p q$, the inner product in the space Q . In all the numerical experiments, we choose compatible elements for which the inf-sup condition holds.

5.1. An analytical example

We will first present the outcome of numerical experimentation carried out on the incompressible system (1.6). Let $\Omega = (0, 1) \times (0, 1)$ be the domain. The sought parameter is $\mu(x, y) = 1 + x^2 y + y x$, and the load function is

Table 1. Identification error $\frac{\|\mu^h - \mu\|_{L^2(\Omega)}}{\|\mu\|_{L^2(\Omega)}}$ (CPU time) for $\kappa = 1 \times 10^{-5}$, $h = 0.0707107$.

ε	$\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ (non Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (non Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (Hessian)
1×10^{-7}	8.406×10^{-3} (50.2 s)	6.688×10^{-3} (160.8 s)	7.378×10^{-3} (1814 s)
1×10^{-8}	8.686×10^{-3} (47.9 s)	7.157×10^{-3} (168.3 s)	7.378×10^{-3} (1762 s)
1×10^{-9}	9.110×10^{-3} (57.4 s)	6.947×10^{-3} (153.3 s)	7.378×10^{-3} (1760 s)
1×10^{-10}	9.802×10^{-3} (47.6 s)	7.487×10^{-3} (179.4 s)	7.378×10^{-3} (1798 s)

Table 2. Regularized discretization error for $\varepsilon = 1 \times 10^{-10}$ ($\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ elements).

h	$\frac{\ u_1^{h,\varepsilon} - \bar{u}_1\ _{L^2(\Omega)}}{\ \bar{u}_1\ _{L^2(\Omega)}}$	$\frac{\ u_2^{h,\varepsilon} - \bar{u}_2\ _{L^2(\Omega)}}{\ \bar{u}_2\ _{L^2(\Omega)}}$	$\ P^{h,\varepsilon}\ _{L^2(\Omega)}$
$\sqrt{2}/10$	8.674×10^{-4}	8.741×10^{-4}	1.042×10^{-3}
$\sqrt{2}/12$	6.021×10^{-4}	6.067×10^{-4}	7.140×10^{-4}
$\sqrt{2}/14$	4.422×10^{-4}	4.455×10^{-4}	5.196×10^{-4}
$\sqrt{2}/16$	3.384×10^{-4}	3.410×10^{-4}	3.950×10^{-4}
$\sqrt{2}/18$	2.673×10^{-4}	2.693×10^{-4}	3.105×10^{-4}
$\sqrt{2}/20$	2.165×10^{-4}	2.181×10^{-4}	2.505×10^{-4}

$$f(x, y) = (2x^3 + 4x^2y + 4xy - 2x + 2, 6x^2y + 4xy^2 + 2y^2 - 2y + 2).$$

On Γ_1 (bottom and right boundaries), we have Dirichlet boundary conditions. On the bottom boundary, $y = 0$, we consider $g(x, y) = (0, x(1 - x))$, and the right part, $x = 1$, we consider $g(x, y) = (y(1 - y), 0)$. We consider Neumann conditions Γ_2 in the top and left part. In the top part, $y = 1$, we consider $h(x, y) = ((-x^2y + xy + 1)(2x + 2y - 2), 0)$, and in the left boundary, $x = 0$, we consider $h(x, y) = (0, (x^2y + xy + 1)(2x + 2y - 2))$. The corresponding displacement vector is $\bar{u}(x, y) = (y(1 - y), x(1 - x))$.

In the numerical experiments, we use the finite element discretization scheme outlined in section 4, and for simplicity, we consider no data contamination. For discretization, we use the finite element library FreeFem++ [38]. To solve the optimization problem, we use the IPOPT library integrated with FreeFem++. IPOPT is a software library for large-scale nonlinear constrained optimization problems using a primal-dual interior-point method, see [39]. IPOPT uses the first-order or second-order derivatives depending on whether the Hessian matrix is provided. If the Hessian is not supplied, then it is approximated by using the BFGS-quasi-Newton methods. IPOPT allows us to incorporate the bounds on the constraints, and in all the experiments, we set the lower bound $l_b = 1$ and the upper bound $u_b = 3.5$. In the present context, this means that the discrete analog of the set A of the feasible parameters is a box constrained set with the lower bound l_b and the upper bound u_b . In all the numerical experiments, we take $H = H^1(\Omega)$ as regularization space.

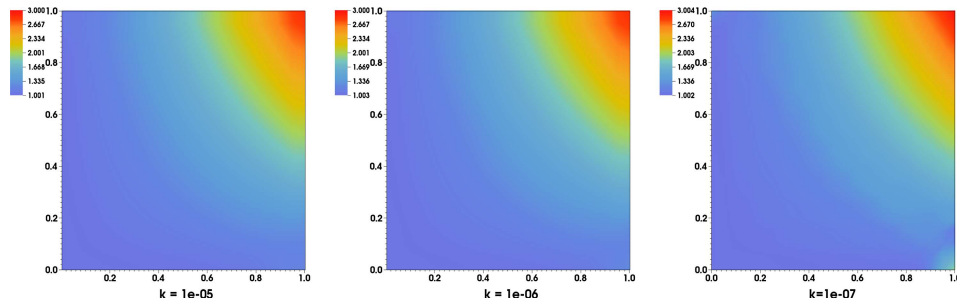
For discretization, we use the triangular elements. Moreover, we consider $\mathbb{P}_1$ elements for the pressure term p and the parameter μ , while for the displacement vector $u = (u_1, u_2)$, we have two choices, namely, $\mathbb{P}_2$ and $\mathbb{P}_1$ elements. We consider three possible scenarios. In the first case, we consider $\mathbb{P}_2$ elements for displacement and $\mathbb{P}_1$ for pressure term and use the Hessian-based methods. In the second scenario, we consider the same $\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ discretization scheme, but we do not provide the Hessian. In the third scenario, we consider $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ and Hessian is given.

Table 3. Regularized discretization error for $\varepsilon = 1 \times 10^{-10}$ ($\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ elements).

h	$\frac{\ u_1^{h,\varepsilon} - \bar{u}_1\ _{L^2(\Omega)}}{\ \bar{u}_1\ _{L^2(\Omega)}}$	$\frac{\ \bar{u}_2^{h,\varepsilon} - \bar{u}_2\ _{L^2(\Omega)}}{\ \bar{u}_2\ _{L^2(\Omega)}}$	$\ p^{h,\varepsilon}\ _{L^2(\Omega)}$
$\sqrt{2}/10$	3.871×10^{-4}	7.560×10^{-4}	3.795×10^{-3}
$\sqrt{2}/12$	2.704×10^{-4}	5.304×10^{-4}	2.597×10^{-3}
$\sqrt{2}/14$	1.995×10^{-4}	3.924×10^{-4}	1.888×10^{-3}
$\sqrt{2}/16$	1.532×10^{-4}	3.020×10^{-4}	1.434×10^{-3}
$\sqrt{2}/18$	1.214×10^{-4}	2.396×10^{-4}	1.127×10^{-3}
$\sqrt{2}/20$	9.851×10^{-5}	1.947×10^{-4}	9.088×10^{-4}

Table 4. Identification error $\frac{\|\mu^h - \mu\|_{L^2(\Omega)}}{\|\mu\|_{L^2(\Omega)}}$ (CPU time) for $\kappa = 1 \times 10^{-6}$, $h = 0.0707107$.

ε	$\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ (non Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (non Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (Hessian)
1×10^{-7}	7.542×10^{-3} (105.1 s)	4.131×10^{-3} (320.7 s)	5.560×10^{-3} (1805.1 s)
1×10^{-8}	9.668×10^{-3} (90.4 s)	5.186×10^{-2} (370.6 s)	5.560×10^{-3} (1759.1 s)
1×10^{-9}	1.113×10^{-2} (75.6 s)	4.629×10^{-3} (301.9 s)	5.560×10^{-3} (1779.5 s)
1×10^{-10}	1.114×10^{-2} (97.6 s)	3.092×10^{-3} (334.0 s)	5.560×10^{-3} (1805.8 s)

**Figure 1.** Scenario $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$. Estimated parameter for different κ , $\varepsilon = 1 \times 10^{-10}$, $h = \frac{\sqrt{20}}{20}$.

In the second scenario, we studied the influence of the regularization parameters κ and ε on the identification error for a fixed discretization parameter $h = \frac{\sqrt{2}}{20}$, where h is defined as the length of the longest edge of all elements in the mesh. For this, we measure the identification error by the quantity

$$\frac{\|\mu^h - \mu\|_{L^2(\Omega)}}{\|\mu\|_{L^2(\Omega)}},$$

where μ is the interpolated exact parameter and μ^h is the computed solution.

We performed several groups of numerical experiments. Firstly, we studied the influence of ε for the discrete saddle point problem. Tables 2 and 3 show that the stability of discretization error. For $\varepsilon \in \{1 \times 10^{-7}, 1 \times 10^{-8}, 1 \times 10^{-9}, 1 \times 10^{-10}\}$, we obtained similar results. For κ , the most stable option appears to be $\kappa = 1 \times 10^{-5}$, which gives excellent reconstruction for all the

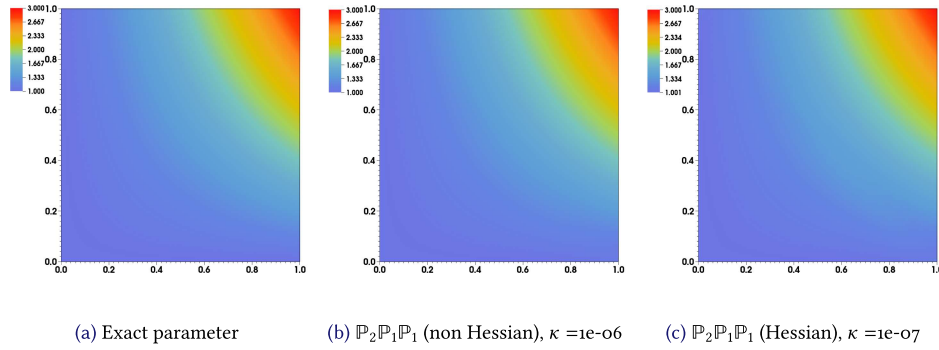


Figure 2. Exact and estimated parameter for $\varepsilon = 1.0 \times 10^{-10}$, $h = \frac{\sqrt{20}}{20}$.

Table 5. Identification error $\frac{\|\mu^h - \mu\|_{L^2(\Omega)}}{\|\mu\|_{L^2(\Omega)}}$ (CPU time) for $\kappa = 1 \times 10^{-7}$, $h = \frac{\sqrt{2}}{20}$.

ε	$\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ (non-Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (non-Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (Hessian)
1×10^{-7}	4.254×10^{-2} (163.3 s)	2.212×10^{-2} (418.2 s)	7.283×10^{-3} (1763.4 s)
1×10^{-8}	3.445×10^{-2} (144.5 s)	2.776×10^{-2} (424.4 s)	7.283×10^{-3} (1754.0 s)
1×10^{-9}	7.748×10^{-2} (113.2 s)	3.252×10^{-2} (409.9 s)	7.283×10^{-3} (1765.0 s)
1×10^{-10}	2.823×10^{-2} (148.8 s)	2.275×10^{-2} (302.6 s)	7.283×10^{-3} (1821.6 s)

Table 6. Identification error $\frac{\|\mu^h - \mu\|_{L^2(\Omega)}}{\|\mu\|_{L^2(\Omega)}}$ for $\kappa = 1 \times 10^{-5}$, $\varepsilon = 1 \times 10^{-10}$.

h	$\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ (non-Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (non-Hessian)	$\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ (Hessian)
$\sqrt{2}/10$	1.525×10^{-2}	5.766×10^{-3}	6.200×10^{-3}
$\sqrt{2}/12$	1.251×10^{-2}	6.259×10^{-3}	6.485×10^{-3}
$\sqrt{2}/14$	1.010×10^{-2}	6.211×10^{-3}	7.145×10^{-3}
$\sqrt{2}/16$	1.110×10^{-2}	7.066×10^{-3}	7.511×10^{-3}
$\sqrt{2}/18$	8.741×10^{-3}	6.684×10^{-3}	7.521×10^{-3}
$\sqrt{2}/20$	9.802×10^{-3}	7.487×10^{-3}	7.378×10^{-3}

chosen values of ε (table 1). For smaller values of κ , we noticed some instability in the reconstruction, particularly for the non-Hessian-based schemes. For $\kappa = 1 \times 10^{-6}$ (see table 4), the discretization scheme $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ gives a less satisfactory reconstruction (see figure 1), but still a bit better than the other two scenarios, see figure 2. For $\kappa = 1 \times 10^{-7}$, the Hessian-based identification scheme gives excellent reconstruction, while $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ scenarios give acceptable reconstruction. On the other hand, $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$, discretization is by far the fastest method, especially if we compare with the Hessian-based schemes which are computationally expensive. In conclusion, $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ gives a much faster method and excellent reconstruction for specific a regularization parameter, while still acceptable reconstruction for other values of the regularization parameter. For $\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$, with the use of Hessian, the reconstruction is very good, but the computational cost is high (table 5). The choice $\mathbb{P}_2\mathbb{P}_1\mathbb{P}_1$ resides in between the two. The same behavior can be observed, when studying the effects of the discretization parameter for fixed regularization parameters $\kappa = 1 \times 10^{-5}$, $\varepsilon = 1 \times 10^{-10}$; see table 6.

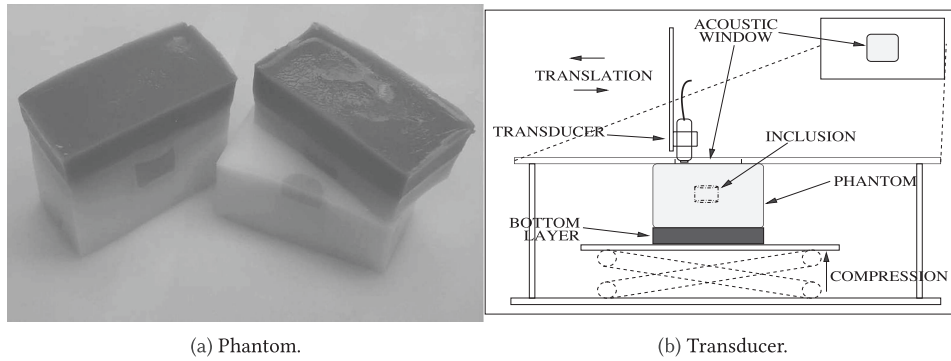


Figure 3. Elastography experiment.

5.2. 3D reconstruction using a tissue phantom data

We will now test the developed framework on a 3D-reconstruction of the elasticity modulus μ using a tissue phantom data. The phantom used in this work was created using gelatin, with silica added for acoustic scatter, to mimic elastic properties of soft tissue (see figure 3(a), which is taken from [40]). A complete description of the phantom construction and experimental imaging setup can be found in Richards, *et al* [40]. The whole phantom was cuboid in shape ($60 \times 60 \times 50$ mm in width, length, and height, respectively) with an 8% by mass background gelatin concentration and a centrally located, stiffer, cylindrical inclusion of 12% gelatin concentration (4.80 mm in diameter and 5 mm in height). This led to an approximate inclusion to background contrast of 1.89 ± 0.11 as measured by an independent mechanical test [40]. A bottom layer (approximately 10 mm of additional height) was also added to the phantom with a layer to background contrast of 2.01 ± 0.22 as measured by an independent mechanical test [40]. Ultrasound (US) image sequences were collected using an Analogic AN2300 (Analogic Corp., 8 Centennial Drive, Peabody, MA 01960) with a 40 Hz linear array transducer. Three dimensional, static images were acquired by scanning the US transducer at a fixed rate while triggering the two-dimensional US frame acquisitions at a fixed elevational distance of 0.14 mm (see figure 3(b), which is taken from [40]). Two 3D images were acquired, a pre-deformation 3D image at approximately no compression, and a second, post-deformation image, after ≈ 1 –2 strain was applied to the phantom in the axial image direction. The scanned volume measured approximately $27.44 \text{ mm} \times 55.62 \text{ mm} \times 27.44 \text{ mm}$ in the lateral (x), axial (y), and elevational (z) directions, respectively. The full 3D displacement vector field was measured from the static images using an image registration based, 3D displacement estimator described in [40].

We consider a 3D discretization by following the same scheme as used in the previous example. We also take into account the practical aspects given in [40], see also [41]. For our computations, we considered the $\mathbb{P}_1\mathbb{P}_1\mathbb{P}_1$ discretization scheme and used a projected data on a mesh of size $20 \times 30 \times 20$ (full data available to us corresponds to a mesh of size $41 \times 61 \times 41$). Another important aspect, when dealing with phantom data, is the choice of boundary conditions. Following [40, 41], we considered Dirichlet boundary conditions in the top and bottom parts of the boundary. We set the vertical direction/component of the remaining boundary conditions to be Dirichlet and allow lateral components to be traction free. We use $l_b = 1$ and $u_b = 3.5$ as the lower and the upper bound. Since we are dealing with a discontinuous parameter, we have chosen the standard TV regularizer (see [40, 42, 43]).

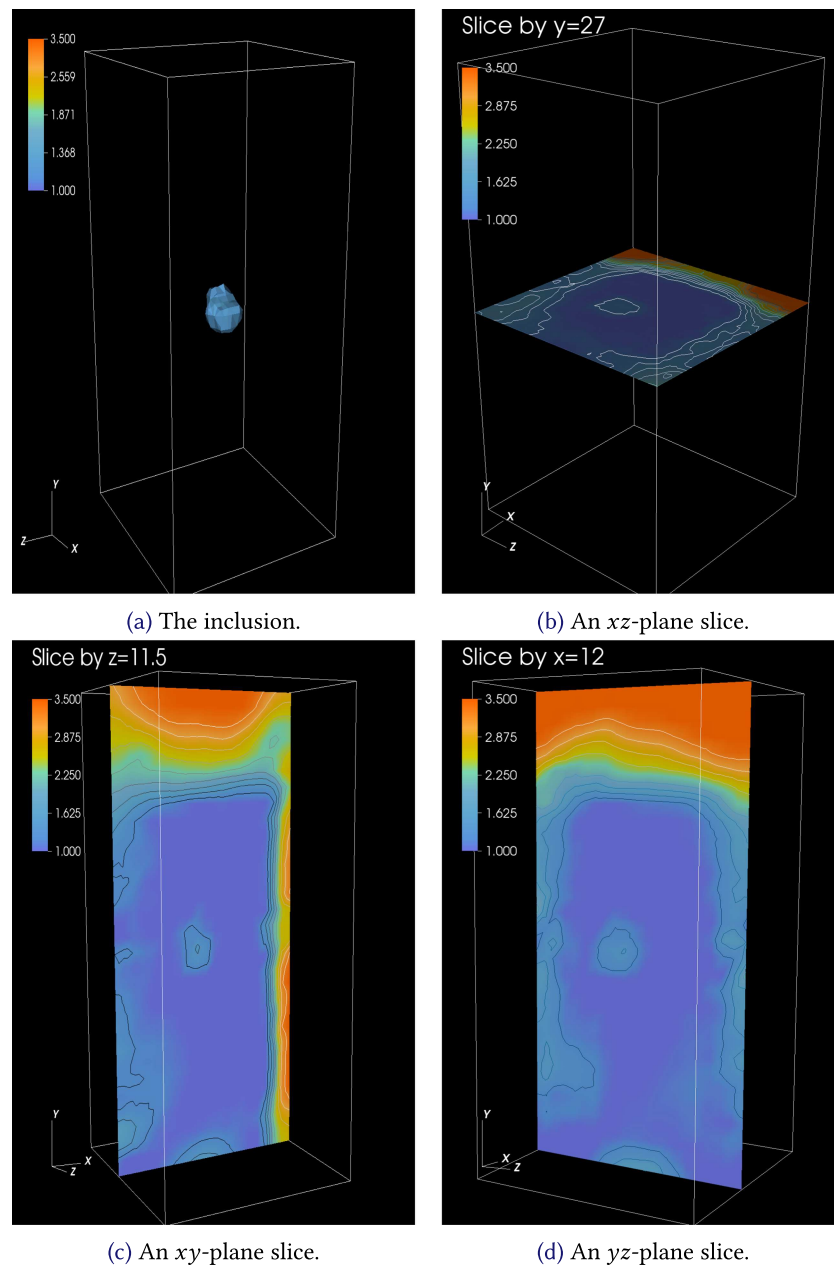


Figure 4. A 3D visualization of the inclusion slices.

In the 3D plots shown in figure 4, the inclusion is highlighted by using a color map corresponding to the identified parameter μ . Considering the use of a projected data into a coarse mesh, the developed scheme identifies the inclusion quite satisfactorily. The artifacts in the boundaries and the stand-off layer correspond to the stiffer bottom part. The artifacts are more pronounced in the lateral part than it is in the upper/bottom part. The background contrast is

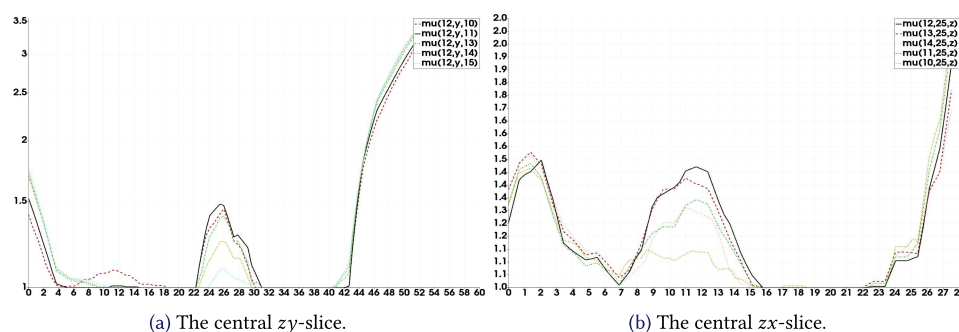


Figure 5. Five lines of parameter reconstruction for two central slices.

approximately 1, corresponding to the lower bound, whereas the inclusion contrast is between 1.3 and 1.5, where the stiffer bottom part corresponds to the stand-off layer. Figure 5 shows parameter μ reconstruction along five lines for two central slices (by $x = 12$ and $y = 25$).

6. Concluding remarks

Saddle point problems have been regularized extensively for stable approximation schemes. In this work, we successfully used this philosophy to devise a stable procedure of the inverse problem of identifying a variable parameter in general saddle point problems. We would also like to note that although the MOLS-type functionals have been used extensively for inverse problems, this is the first instance of their use on a real-world problem with real data. Our numerical experiments provide very encouraging results. Furthermore, it appears that some data smoothing could be used to mitigate the artifacts in the reconstruction procedure. Tools such as Sobolev gradients or stochastic gradients seem to be adequate for it. Finally, in the elasticity imaging inverse problem, the boundary conditions are not known. Since the boundary conditions have a direct impact on the overall success of the identification process, it is of interest to develop a framework that also identifies boundary conditions along with the elasticity modulus.

Acknowledgments

We are grateful to the reviewers for their incredibly detailed and insightful reports that helped us significantly improve the results as well as the presentation. The research of Baasansuren Jadamba and Akhtar Khan is supported by the National Science Foundation under Award No. 1720067. Miguel Sama is 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) and Grant No. 2020-MAT11 (ETSI Industriales, UNED).

ORCID iDs

Miguel Sama  <https://orcid.org/0000-0002-9698-0714>

References

- [1] Girault V and Raviart P-A 1979 *Finite Element Approximation of the Navier–Stokes Equations* (Berlin: Springer)
- [2] Barbosa H and Hughes T 1992 Circumventing the Babuška–Brezzi condition in mixed finite element approximations of elliptic variational inequalities *Comput. Methods Appl. Mech. Eng.* **97** 193–210
- [3] Bochev P and Lehoucq R 2006 Regularization and stabilization of discrete saddle-point variational problems *Electron. Trans. Numer. Anal.* **22** 97–113
- [4] Bochev P, Gunzburger M and Lehoucq R 2007 On stabilized finite element methods for the Stokes problem in the small time step limit *Int. J. Numer. Methods Fluids* **53** 573–97
- [5] Ito K, Kunisch K and Peichl G 2002 On the regularization and approximation of saddle point problems without inf-sup condition *Comput. Appl. Math.* **21** 245–74
- [6] Pestana J and Wathen A J 2015 Natural preconditioning and iterative methods for saddle point systems *SIAM Rev.* **57** 71–91
- [7] Beigl A, Scherzer O, Sogn J and Zulehner W 2020 Preconditioning inverse problems for hyperbolic equations with applications to photoacoustic tomography *Inverse Problems* **36** 014002
- [8] Braess D 2007 *Finite Elements* (Cambridge: Cambridge University Press)
- [9] Brezzi F and Fortin M 1991 *Mixed and Hybrid Finite Element Methods* (New York, NY: Springer)
- [10] Albocher U, Barbone P, Oberai A and Harari I 2014 Uniqueness of inverse problems of isotropic incompressible three-dimensional elasticity *J. Mech. Phys. Solids* **73** 55–68
- [11] Ammari H, Garapon P, Kang H and Lee H 2008 A method of biological tissues elasticity reconstruction using magnetic resonance elastography measurements *Q. Appl. Math.* **66** 139–75
- [12] Banks H T and Luke N S 2008 Modelling of propagating shear waves in biotissue employing an internal variable approach to dissipation *Commun. Comput. Phys.* **3** 603–40
- [13] Cahill N C, Jadamba B, Khan A A, Sama M and Winkler B 2013 A first-order adjoint and a second-order hybrid method for an energy output least squares elastography inverse problem of identifying tumor location *Bound. Value Problems* **263** 1–14
- [14] Crossen E, Gockenbach M, Jadamba B, Khan A A and Winkler B 2014 An equation error approach for the elasticity imaging inverse problem for predicting tumor location *Comput. Math. Appl.* **67** 122–35
- [15] Doyley M M 2012 Model-based elastography: a survey of approaches to the inverse elasticity problem *Phys. Med. Biol.* **57** R35
- [16] Doyley M M, Jadamba B, Khan A A, Sama M and Winkler B 2014 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** 984–1017
- [17] Gockenbach M, Jadamba B, Khan A A, Tammer C and Winkler B 2015 Proximal methods for the elastography inverse problem of tumor identification using an equation error approach *Advances in Variational and Hemivariational Inequalities* (Cham: Springer) 173–97
- [18] Hubmer S, Sherina E, Neubauer A and Scherzer O 2018 Lamé parameter estimation from static displacement field measurements in the framework of nonlinear inverse problems *SIAM J. Imaging Sci.* **11** 1268–93
- [19] Jadamba B, Khan A A, Oberai A and Sama M 2017 First-order and second-order adjoint methods for parameter identification problems with an application to the elasticity imaging inverse problem *Inverse Problems Sci. Eng.* **25** 1768–87
- [20] Ji L and McLaughlin J 2004 Recovery of the Lamé parameter μ in biological tissues *Inverse Problems* **20** 1–24
- [21] McLaughlin J R and Yoon J R 2004 Unique identifiability of elastic parameters from time-dependent interior displacement measurement *Inverse Problems* **20** 25–45
- [22] Widlak T and Scherzer O 2015 Stability in the linearized problem of quantitative elastography *Inverse Problems* **31** 035005
- [23] Hughes T 1987 *The Finite Element Method* (Englewood Cliffs, NJ: Prentice Hall)
- [24] Ammari H, Garapon P and Jouve F 2010 Separation of scales in elasticity imaging: a numerical study *J. Comput. Math.* **28** 354–70
- [25] Boehm C and Ulbrich M 2015 A semismooth Newton-CG method for constrained parameter identification in seismic tomography *SIAM J. Sci. Comput.* **37** S334–64

- [26] Boiger R and Kaltenbacher B 2016 An online parameter identification method for time dependent partial differential equations *Inverse Problems* **32** 045006
- [27] Clason C 2012 L^∞ fitting for inverse problems with uniform noise *Inverse Problems* **28** 104007
- [28] Evstigneev R O, Medvedik M Y and Smirnov Y G 2016 Inverse problem of determining parameters of inhomogeneity of a body from acoustic field measurements *Comput. Math. Math. Phys.* **56** 483–90
- [29] Gholami A, Mang A and Biros G 2016 An inverse problem formulation for parameter estimation of a reactiondiffusion model of low grade gliomas *J. Math. Biol.* **72** 409–33
- [30] Guchhait S and Banerjee B 2015 Constitutive error based material parameter estimation procedure for hyperelastic material *Comput. Methods Appl. Mech. Eng.* **297** 455–75
- [31] Hager W, Ngo C, Yashtini M and Zhang H-C 2015 An alternating direction approximate Newton algorithm for illconditioned inverse problems with application to parallel MRI *J. Oper. Res. Soc. China* **3** 139–62
- [32] Jadamba B, Khan A A, Sama M and Tammer C 2017 On convex modified output least-squares for elliptic inverse problems: stability, regularization, applications, and numerics *Optimization* **66** 983–1012
- [33] Kindermann S, Mutumbu L D and Resmerita E 2014 A numerical study of heuristic parameter choice rules for total variation regularization *J. Inverse Ill-Posed Problems* **22** 63–94
- [34] Kuchment P and Steinhauer D 2015 Stabilizing inverse problems by internal data. II: non-local internal data and generic linearized uniqueness *Anal. Math. Phys.* **5** 391–425
- [35] Liu T 2016 A wavelet multiscale-homotopy method for the parameter identification problem of partial differential equations *Comput. Math. Appl.* **71** 1519–23
- [36] Manservigi S and Gunzburger M 2000 A variational inequality formulation of an inverse elasticity problem *Appl. Numer. Math.* **34** 99–126
- [37] Neubauer A, Hein T, Hofmann B, Kindermann S and Tautenhahn U 2010 Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces *Appl. Anal.* **89** 1729–43
- [38] Hecht F 2012 New development in FreeFem++ *J. Numer. Math.* **20** 251–66
- [39] Wachter A and Biegler L T 2006 On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming *Math. Program.* **106** 25–57
- [40] Richards M, Barbone P and Oberai A 2009 Quantitative three-dimensional elasticity imaging from quasi-static deformation: a phantom study *Phys. Med. Biol.* **54** 757
- [41] Richards M 2007 Quantitative three dimensional elasticity imaging *PhD Thesis* Boston University
- [42] Jadamba B, Khan A A, Rus G, Sama M and Winkler B 2014 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** 1486–510
- [43] Nashed N Z and Scherzer O 1998 Least squares and bounded variation regularization with nondifferentiable functionals *Numer. Funct. Anal. Optim.* **19** 873–901