Optimization

A Journal of Mathematical Programming and Operations Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gopt20

Analyzing the role of the Inf-Sup condition for parameter identification in saddle point problems with application in elasticity imaging

Baasansuren Jadamba , Akhtar A. Khan , Michael Richards , Miguel Sama & Christiane Tammer

OPTIMIZATION
https://doi.org/10.1080/02331934.2020.1789128

Taylor & Francis
Taylor & Francis Group

Analyzing the role of the Inf-Sup condition for parameter identification in saddle point problems with application in elasticity imaging

Baasansuren Jadamba[a], Akhtar A. Khan[a], Michael Richards[b], Miguel Sama[c] and Christiane Tammer[d]

[a]Center for Applied and Computational Mathematics, School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, USA; [b]Department of Biomedical Engineering, Kate Gleason College of Engineering, Rochester Institute of Technology, Rochester, NY, USA; [c]Departamento de Matemática Aplicada, Universidad Nacional de Educación a Distancia, Madrid, Spain; [d]Institute of Mathematics, Martin-Luther-University of Halle-Wittenberg, Halle-Saale, Germany

ABSTRACT

We study the inverse problem of parameter identification in general saddle point problems. For saddle point problems, the use of elliptic regularization is an essential component. Saddle point problems, after discretization, lead to a non-invertible system, whereas the regularized saddle point problems result in an invertible system. Regularization methods, in the context of saddle point problems, have also been used to mitigate the role of the Inf-Sup condition, synonymously, also called the Babuska-Brezzi condition. This work aims to analyze the impact of regularizing the saddle point problem on the inverse problem. We investigate the inverse problem by using the output least-squares objective. To exploit the use of regularization fully, we work under the assumption that the solution map is nonempty. We regularize the saddle point problem and consider a family of optimization problems using the output least-squares objective for the regularized saddle point problem where some noise contaminates the whole data set. We give a complete convergence analysis showing that the optimization problems, given for the regularized output least-squares, approximate the original problem suitably. We also provide the first-order and the second-order adjoint method for the computation of the first-order and the second-order derivatives of the output least-squares objective. We present some heuristic numerical results. In the context of the elasticity imaging inverse problem, we conduct detailed numerical experiments on synthetic data (to study the role of the regularization parameter) as well as on phantom data.

ARTICLE HISTORY
Received 10 August 2019
Accepted 24 June 2020

KEYWORDS
Saddle point problems; parameter identification; elliptic regularization; output least-squares; elasticity imaging

2000 MATHEMATICS SUBJECT CLASSIFICATIONS
35R30; 49N45; 65J20; 65J22; 65M30

CONTACT Akhtar A. Khan aaksma@rit.edu

1. Introduction

We study the inverse problem of identifying a parameter in saddle point problems, which provide a unified framework for analyzing numerous applied models such as elasticity, Stokes equations, fourth-order boundary-value problems, the pure Neumann boundary-value problem, and many others. In recent years, the subject of inverse problems emerged as one of the most vibrant and expanding areas of research because of its ever-growing inventory of applications to biomedical sciences, finance, engineering, social sciences, and other related disciplines. For an overview of the recent developments in the dynamic field of inverse problems, we refer the interested reader to [1–22].

1.1. Problem formulation

We begin by introducing the saddle point problem (the direct problem) and the associated inverse problem of parameter identification. We denote the parameter space by B, which is a real Banach space. We denote the set of admissible parameters by A, which is a nonempty, closed, convex, and bounded subset of B. We pose the saddle point problem in real Hilbert spaces V and Q whose duals are denoted by V^* and Q^*. We take the measured data in real Hilbert spaces $\mathbb{V}$ and $\mathbb{Q}$ such that $V \times Q$ continuously embeds in $\mathbb{V} \times \mathbb{Q}$. We denote the norm of a normed space N by $\|\cdot\|_N$ and specify the strong convergence by $\to$ and the weak convergence by $\rightharpoonup$. 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 form, $f \in V^*$, and $g \in Q^*$. Assume that a and b are continuous, that is, there are constants $\kappa_0 > 0$ and $\varsigma_0 > 0$ such that

$$|a(\ell, u, v)| \leq \varsigma_0 \|\ell\|_B \|u\|_V \|v\|_V, \quad \text{for all } \ell \in B,\ u, v \in V, \tag{1}$$

$$\left|b(v, q)\right| \leq \kappa_0 \|v\|_V \|q\|_Q, \quad \text{for all } v \in V,\ q \in Q. \tag{2}$$

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

$$a(\ell, u, v) + b(v, p) = f(v), \quad \text{for every } v \in V, \tag{3a}$$

$$b(u, q) = g(q), \quad \text{for every } q \in Q. \tag{3b}$$

For a fixed $\ell \in B$, the trilinear form a and the bilinear form b are related to the 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 $q \in Q$. Furthermore, let $\mathbb{B}^* \in \mathcal{L}(Q, V^*)$ be the dual map of $\mathbb{B}$, that is,

$$\langle \mathbb{B}^* q, v\rangle = \langle q, \mathbb{B}v\rangle, \quad \text{for all } v \in V,\ \text{for all } q \in Q.$$

Then, (3a) can be equivalently written as the problem of finding $(u, p) \in V \times Q$ such that

$$\mathbb{A}_\ell u + \mathbb{B}^* p = f, \tag{4a}$$

$$\mathbb{B} u = g. \tag{4b}$$

1.2. Motivation and objectives

Our objective is to study the inverse problem of identifying a parameter $\ell \in A$ such that the corresponding solution $(u(\ell), p(\ell))$ is closest, in some norm, to the given data $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$. Inverse problems are most commonly studied using optimization tools which offer a convenient way of incorporating regularization. Adhering to this trend, we pose this nonlinear inverse problem as an output least-squares (OLS) based constrained optimization problem that minimizes the gap between the computed solution and the measured data. Find $\ell \in A$ by solving the following optimization problem

$$\min_{\ell \in A} J(\ell) := \frac{1}{2}\|u(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p(\ell) - \hat{z}\|_{\mathbb{Q}}^2. \tag{5}$$

Here, for each $\ell \in A$, $(u(\ell), p(\ell))$ is a solution of saddle point problem (3a), that is, (3a) is an explicit constraint to the above optimization problem; the set A being an implicit constraint.

For (5) to be meaningful, the parameter-to-solution map $\ell \mapsto (u(\ell), p(\ell))$ needs to be well-defined for each $\ell \in A$. The following well-known result is useful in this regard (see [23]):

Theorem 1.1: *Besides (1) and (2), assume that there are constants $\kappa_1 > 0$ and $\varsigma_1 > 0$ such that*

$$a(\ell, v, v) \geq \varsigma_1 \|v\|_V^2, \quad \textit{for all } \ell \in A,\ v \in L, \tag{6}$$

$$\sup_{u \in V} \frac{b(u, q)}{\|u\|_V} \geq \kappa_1 \|q\|_Q, \quad \textit{for all } q \in Q, \tag{7}$$

where $L := \{u \in V \mid b(u, q) = 0,\ \text{for all } q \in Q\}$. Then saddle point problem (3a) has a unique solution (u, p). That is, for $\ell \in A$, the map $\ell \mapsto (u(\ell), p(\ell))$ is a well-defined and single-valued.

Saddle point problem (3a) is associated to the linearly constrained optimization problem:

$$\min_{u \in V} \mathbb{J}(u) := \frac{1}{2} a(\ell, u, u) - f(u), \quad \text{subject to } b(u, q) = g(q), \quad \text{for every } q \in Q. \tag{8}$$

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

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

then under the hypotheses of Theorem 1.1, $L(u,p)$ has a unique saddle point which solves (3a), see [24, p. 132]. Here p is the Lagrange multiplier.

Saddle point problem (3a) leads to a non-invertible system after discretization, requiring specialized solution strategies. A commonly used technique is to regularize (3a), an instant advantage being that the regularized saddle point problem leads to an invertible system. Let $c: Q \times Q \to \mathbb{R}$ be a continuous and elliptic bilinear form. That is, there are constants $\omega_0 > 0$ and $\omega_1 > 0$ with

$$c(q,q) \geq \omega_0 \|q\|_Q^2, \quad \text{for all } q \in Q, \tag{9a}$$

$$c(p,q) \leq \omega_1 \|p\|_Q \|q\|_Q, \quad \text{for all } p,q \in Q. \tag{9b}$$

Given $\varepsilon > 0$, the regularized saddle point problem seeks $(u_\varepsilon, p_\varepsilon) := (u_\varepsilon(\ell), p_\varepsilon(\ell)) \in V \times Q$ with

$$a(\ell, u_\varepsilon, v) + b(v, p_\varepsilon) = f(v), \quad \text{for every } v \in V, \tag{10a}$$

$$b(u_\varepsilon, q) - \varepsilon c(p_\varepsilon, q) = g(q), \quad \text{for every } q \in Q. \tag{10b}$$

The regularized saddle point problem (10a) possesses some crucial computational advantages over (3a), and it has been studied extensively, see [23,25,26].

Associating c with the linear map $\mathbb{C} \in \mathcal{L}(Q, Q^*)$ by the relationship $\langle \mathbb{C}p, q\rangle = c(p,q)$ for every $p,q \in Q$, we can write (10a) into an equivalent problem of finding $(u_\varepsilon, p_\varepsilon) \in V \times Q$ such that

$$\mathbb{A}_\ell u_\varepsilon + \mathbb{B}^* p_\varepsilon = f, \tag{11a}$$

$$\mathbb{B} u_\varepsilon - \varepsilon \mathbb{C} p_\varepsilon = g. \tag{11b}$$

A fundamental result, connecting (3a) and (10a), asserts that under hypotheses (1), (2), (6), and (7), both (3a) and (10a) are uniquely solvable, and the regularized solutions $\{(u_\varepsilon, p_\varepsilon)\}$ converge to the unique solution of (3a), as $\varepsilon \to 0$. Some strengthening of ellipticity of a has been used to provide error estimates in terms of the regularization parameter ε, which also provide useful insight into the discrete counterparts of (3a) and (10a). To be precise, we have the following:

Theorem 1.2: *Besides (1), (2), and (7), assume that there is a constant $\omega > 0$ such that for each $\ell \in A$, we have*

$$a(\ell, v, v) + \langle \mathbb{C}^{-1}\mathbb{B}v, \mathbb{B}v\rangle \geq \omega \|v\|^2, \quad \textit{for every } v \in V. \tag{12}$$

Then, both (3a) and (10a) have unique solutions (u,p) and $(u_\varepsilon, p_\varepsilon)$. Moreover, there is a constant $K > 0$ such that the following error bound holds for each $\varepsilon (\leq 1)$:

$$\|u_\varepsilon - u\|_V + \|p_\varepsilon - p\|_Q \leq K\varepsilon(\|f\|_{V^*} + \|g\|_{Q^*}). \tag{13}$$

If $v \in L = \ker(\mathbb{B})$, then (12) recovers ellipticity (6). Furthermore, under the hypotheses of Theorem 1.1, the following estimate can be proved (see [27]): There is a constant $K > 0$ such that

$$\|u_\varepsilon - u\|_V + \|p_\varepsilon - p\|_Q \leq K\sqrt{\varepsilon}. \tag{14}$$

Condition (7), commonly termed as the Babuska-Brezzi or the Inf-Sup condition, is a natural substitute for the ellipticity condition and plays an essential role in the theoretical as well as the numerical treatment of saddle point problems, see [25,28–30]. In finite element discretization of (3a) for computing a solution, a discrete analog of (7) is used, which imposes stringent compatibility restrictions on the choices of the finite-dimensional subspaces of V and Q. Furthermore, there are explicit examples when either the Inf-Sup condition cannot be verified or is not entirely satisfactory from an analysis viewpoint. A classic example of the former is the Stokes problem for which the use of the liner elements for velocity and the pressure fails to satisfy the Inf-Sup condition; see [26, Section IV.2]. Pertaining to the latter, the following example, taken from [31], shows a scenario where the Inf-Sup condition is not optimal:

Example 1.3: Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with sufficiently smooth boundary, and Γ_0 be a closed, simple, and smooth curve in Ω. We consider the following saddle point problem:

$$\int_\Omega \left(\nabla u \cdot \nabla v - fv\right) \mathrm{d}x + \int_{\Gamma_0} pv \, \mathrm{d}s = 0, \quad \text{for all } v \in H_0^1(\Omega) \tag{15a}$$

$$u = 0, \quad \text{on } \Gamma_0. \tag{15b}$$

The above saddle point problem, which fits (3a) with $V = H_0^1(\Omega)$, $Q^* = H^{1/2}(\Gamma_0)$, and $\mathbb{B}$ the trace operator, is a necessary optimality condition of the following optimization problem that has relevance to shape optimization problems:

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \left(\frac{1}{2}|\nabla u|^2 - uf\right) \mathrm{d}x, \quad \text{subject to } u = 0 \text{ on } \Gamma_0. \tag{16}$$

The Inf-Sup condition holds and (15a) is well-posed in $H_0^1(\Omega) \times H^{-1/2}(\Gamma_0)$. The discrete variant of the Inf-Sup also condition holds in suitable finite-dimensional subspaces of V and Q^*, and the convergence to the solution in $H_0^1(\Omega) \times H^{-1/2}(\Gamma_0)$ can be proved. Since, for $f \in L^2(\Omega)$, the solution u of (15a) exhibits $H^2(\Omega)$ regularity in all subdomains not containing Γ_0, u is piecewise-H^2, implying that $p \in H^{1/2}(\Gamma_0)$. Thus, the convergence for p is preferred to be in a topology finer than that of $H^{-1/2}(\Gamma_0)$, such as $L^2(\Gamma_0)$ or $H^{1/2}(\Gamma_0)$. However, the Inf-Sup condition does not hold for $L^2(\Gamma_0)$ as the trace operator from $H_0^1(\Omega)$ onto $L^2(\Gamma_0)$ is not surjective. In conclusion, the convergence analysis cannot be obtained in $L^2(\Gamma_0)$ by the aid of the Inf-Sup condition, which only holds in $H^{-1/2}(\Gamma_0)$.

Remark 1.4: The unique solvability of (10a) holds without (7), and hence the regularization approach can well be used for ill-posed saddle point problems. Ito, Kunisch, and Peichil [31] used the regularization approach to circumvent the absence of the Inf-Sup condition. They posed the saddle point problem in Hilbert spaces X and Z^*, and developed an abstract framework using three Gelfand triples $X \hookrightarrow H \hookrightarrow X^*$, $Z \hookrightarrow Y \hookrightarrow Z^*$, and $W \hookrightarrow Y \hookrightarrow W^*$. The first Gelfand triple was used the describe the map defining the abstract variant of (16), the second Gelfand triple to prove the existence of the Lagrange multiplier in Z^*, and the third Gelfand triple for analyzing the additional regularity of the multiplier. In the context of Example 1.3, $W = Z = H^{1/2}(\Gamma_0)$, $Y = L^2(\Gamma_0)$, and $X = H_0^1(\Omega)$. Saddle point problems (3a) and (15a) are related by the choices $V = X$ and $Q = Z^*$. Assuming that a multiplier resides in a smaller space, in particular, taking the considered saddle point problem in $X \times W$, the authors proved the convergence of the regularized solutions in the norm of $X \times W$. Although we refrain from incorporating the functional framework of [31] in this work, we will present numerical results for the inverse problem of identifying a constant function in (15a). For more details on Example 1.3, see [32,33].

1.3. Main contributions

The primary objective of this work is to assimilate, for the first time, the usefulness of the regularized saddle point problems into the OLS formulation of the inverse problem of parameter identification in saddle point problems. The OLS functional, although the most commonly used optimization formulation for inverse problems, is typically nonconvex, and takes a large number of iterations to converge to a (local) minimum. Since after discretization (3a) leads to a non-invertible system, one has to reply on specific solution strategies that directly impact the inverse problem as the underlying system needs to be solved hundreds of times. Since regularization is a well-established technique for solving saddle point problems, replacing the constraint (3a) in the optimization problem (5) by its regularized variant (10a) has evident advantages. The approach consists of first showing that for each $\varepsilon > 0$, there is an optimal parameter ℓ_ε, and the goal then is to study the convergence of the sequence $\{\ell_\varepsilon\}$. In a nutshell, one of the main contributions of this work is an extension of Theorem 1.2 for the study of inverse problems. We show in Theroem 2.5 that under the continuity of a and b and the ellipticity of a and the Inf-Sup condition, the sequence $\{\ell_\varepsilon\}$ of optimal parameters converges to a minimizer of (5), as $\varepsilon \to 0$.

Note that under the ellipticity and the Inf-Sup condition, both (3a) and (10a) are well-posed. Hence, under these conditions, the regularization process provides a system with better properties to approximate the saddle point problem. On the other hand, for more general variational and quasi-variational inequalities, it is known that the regularization approach is quite efficient even for non-coercive problems. The key component of such studies is that if the solution set of an ill-posed variational problem is nonempty, then the regularized solutions converge to a solution. Inspired by such studies, we develop an abstract framework that does not explicitly rely on the ellipticity of a or the Inf-Sup condition. These conditions, however, play a crucial role, as our general assumptions can be verified when the Inf-Sup condition and the ellipticity hold.

We now provide an outline of the main contribution of this work:

(i) We work under the assumption that the trilinear form a is positive and continuous, the bilinear map b is continuous, and the saddle point problem (3a) is solvable for every parameter $\ell \in A$. Under these assumptions, the parameter-to-solution map is set-valued, in general. At this juncture, we note that under the ellipticity, the Inf-Sup condition is a necessary and sufficient condition for the well-posedness of (3a). However, this result has no direct impact on the present study, as we only work under the assumption that a is merely positive, and hence (3a) is ill-posed.

(ii) There are numerous obstacles associated with a satisfactory theoretical as well as numerical treatment of optimization problems that involve a set-valued parameter-to-solution map. The regularization process circumvents this difficulty and results in a single-valued (regularized) parameter-to-solution map. We prove the derivative characterization for this single-valued map. We consider a variant of the OLS (cf. (5)) for which the constraint is the regularized saddle point problem. We prove the solvability of the optimization problem and show that the OLS-based optimization problem, with the regularized saddle point problem as the constraint, approximates the original OLS-based optimization problem when the regularization parameter diminishes. This result is valid for the general case when the original saddle point problem has a set-valued solution map. However, the imposed conditions simplify significantly if the original saddle point problem is uniquely solvable. This happens when the inf-sup condition holds, and a is elliptic.

(iii) One of the significant drawbacks of the OLS formulation is the need to compute the derivative of the parameter-to-solution map in the computation of the derivative of the OLS objective. The so-called 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 functional.

(iv) We provide the outcome of numerical experimentation for the saddle point problem (15a). This example reflects the impact of the regularization approach when the desired Inf-Sup condition does not hold. We also provide detailed numerical experimentation for an analytic example related to the elasticity imaging inverse problem to show the role of the regularization parameter. Finally, we test the applicability of the developed framework for the elasticity imaging inverse problem using the phantom tissue data.

We organize the contents of this paper into five sections. Section 2 embeds the regularization process into the OLS formulation and provides the convergence analysis for the regularized solutions. Section 3 is devoted to the first-order, and the second-order adjoint approaches. In Section 4, we report the outcome of some preliminary numerical experiments. The paper concludes with some remarks.

2. An optimization framework for the inverse problem

In the following, besides (1) and (2), we assume that a is positive, that is,

$$a(\ell, v, v) \geq 0, \quad \text{for all } \ell \in A,\ v \in V. \tag{17}$$

Moreover, we assume that saddle point problem (3a) is solvable. For a given parameter $\ell \in A$, by $\mathcal{U}(\ell)$ we denote the set of all solutions of saddle point problem (3a). We begin with the following:

Lemma 2.1: *For any $\ell \in A$, the solution set $\mathcal{U}(\ell)$ of saddle point problem (3a) is closed and convex.*

Proof: The proof follows at once from the definition of the set-valued map $\mathcal{U} : A \rightrightarrows V \times Q$. ■

2.1. The output least-squares formulation

We define the set-valued output least-squares map $J : A \rightrightarrows \mathbb{R}$ that relates to each $\ell \in A$, the set

$$J(\ell) := \left\{ \tfrac{1}{2}\|u(\ell) - \bar{z}\|^2_{\mathbb{V}} + \tfrac{1}{2}\|p(\ell) - \hat{z}\|^2_{\mathbb{Q}} \mid (p(\ell), u(\ell)) \in \mathcal{U}(\ell) \right\},$$

where $(u(\ell), p(\ell))$ is a solution of (3a) for $\ell \in A$ and $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$ is the given data.

We now introduce the following output least-squares (OLS) based optimization problem:

$$\min_{\ell \in A} J(\ell). \tag{18}$$

An element $\bar{\ell} \in A$ is called a minimizer of (18), if there exists $(u(\bar{\ell}), p(\bar{\ell})) \in \mathcal{U}(\bar{\ell})$ with

$$\tfrac{1}{2}\|u(\bar{\ell}) - \bar{z}\|^2_{\mathbb{V}} + \tfrac{1}{2}\|p(\bar{\ell}) - \hat{z}\|^2_{\mathbb{Q}} \leq \tfrac{1}{2}\|u(\ell) - \bar{z}\|^2_{\mathbb{V}} + \tfrac{1}{2}\|p(\ell) - \hat{z}\|^2_{\mathbb{Q}}, \tag{19}$$

for every $(u(\ell), p(\ell)) \in \mathcal{U}(\ell)$, for every $\ell \in A$. To emphasize the role of $(u(\bar{\ell}), p(\bar{\ell}))$, we often say that $(\bar{\ell}, u(\bar{\ell}), p(\bar{\ell})) \in \operatorname{graph}(\mathcal{U})$ is a minimizer of (18).

We will approximate a solution of (18) by a family of regularized OLS-based optimization problems where the data set of the saddle point problem is contaminated by some noise in the sense described below. Let $\{\varepsilon_n\}$, $\{\delta_n\}$, and $\{\nu_n\}$ be sequences of positive reals. For each $n \in \mathbb{N}$, let $f_{\nu_n} \in V^*$, $g_{\nu_n} \in Q^*$, and $(\bar{z}_{\delta_n}, \hat{z}_{\delta_n}) \in \mathbb{V} \times \mathbb{Q}$ be the noisy data satisfying the following inequalities:

$$\max\left\{\|f_{\nu_n} - f\|_{V^*}, \|g_{\nu_n} - g\|_{Q^*}\right\} \le \nu_n, \tag{20}$$

$$\max\left\{\|\bar{z}_{\delta_n} - \bar{z}\|_{\mathbb{V}}, \|\hat{z}_{\delta_n} - \hat{z}\|_{\mathbb{Q}}\right\} \le \delta_n. \tag{21}$$

Moreover, as $n \to \infty$, the sequences $\{\varepsilon_n\}$, $\{\delta_n\}$, and $\{\nu_n\}$ satisfy

$$\left\{\varepsilon_n, \nu_n, \delta_n, \frac{\delta_n}{\varepsilon_n}, \frac{\nu_n}{\varepsilon_n}\right\} \to 0. \tag{22}$$

We consider the following family of regularized saddle point problems: For $n \in \mathbb{N}$, given the regularization parameter $\varepsilon_n > 0$ and $\ell \in A$, find $(u_n, p_n) \equiv (u_n(\ell), p_n(\ell)) \in V \times Q$ such that

$$a(\ell, u_n, v) + \varepsilon_n \langle u_n - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n) = f_{\nu_n}(v), \quad \text{for every } v \in V, \tag{23a}$$

$$b(u_n, q) - \varepsilon_n \langle p_n - \hat{z}_{\delta_n}, q\rangle_Q = g_{\nu_n}(q), \quad \text{for every } q \in Q. \tag{23b}$$

As a direct consequence of the Lax-Milgram lemma, for fixed $n \in \mathbb{N}$ and $\ell \in A$, regularized saddle point problem (23a) has a unique solution $(u_n(\ell), p_n(\ell))$. Therefore, the regularized parameter-to-solution map $\ell \to (u_n(\ell), p_n(\ell))$ is well-defined and single-valued. The following result sheds some light on the smoothness of the regularized parameter-to-solution map:

Theorem 2.2: *For $n \in \mathbb{N}$ fixed, let ℓ be in the interior of A which we assume to be nonempty. The first-order derivative $(Du_n(\ell)\delta\ell, Dp_n(\ell)\delta\ell)$ of the regularized parameter-to-solution map $\ell \to (u_n(\ell), p_n(\ell))$ at ℓ in the direction $\delta\ell \in B$ is the unique solution of the regularized saddle point problem*:

$$\begin{aligned} a(\ell, Du_n(\ell)\delta\ell, v) + \varepsilon_n \langle Du_n(\ell)\delta\ell, v\rangle_V + b(v, Dp_n(\ell)\delta\ell) \\ = -a(\delta\ell, u_n(\ell), v), \quad \textit{for every } v \in V, \end{aligned} \tag{24a}$$

$$b(Du_n(\ell)\delta\ell, q) - \varepsilon_n \langle Dp_n(\ell)\delta\ell, q\rangle_Q = 0, \quad \textit{for every } q \in Q. \tag{24b}$$

Moreover, the second-order derivative $(D^2u_n(\ell)(\delta\ell_1, \delta\ell_2), D^2p_n(\ell)(\delta\ell_1, \delta\ell_2))$ of $(u_n(\ell), p_n(\ell))$ 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:

$$\begin{aligned} a(\ell, D^2u_n(\ell)(\delta\ell_1, \delta\ell_2), v) + \varepsilon_n \langle D^2u_n(\ell)(\delta\ell_1, \delta\ell_2), v\rangle_V + b(v, D^2p_n(\ell)(\delta\ell_1, \delta\ell_2)) \\ = -a(\delta\ell_2, Du_n(\ell)\delta\ell_1, v) - a(\delta\ell_1, Du_n(\ell)\delta\ell_2, v), \quad \textit{for every } v \in V, \end{aligned} \tag{25a}$$

$$b(D^2u_n(\ell)(\delta\ell_1, \delta\ell_2), q) - \varepsilon_n \langle D^2p_n(\ell)(\delta\ell_1, \delta\ell_2), q\rangle_Q = 0, \quad \textit{for every } q \in Q. \tag{25b}$$

Proof: The proof follows by similar arguments that were used in [34]. ■

2.2. The regularized output least-squares formulation

The nonlinear inverse problem of parameter identification is known to be ill-posed, and regularization is necessary for a stable identification process. The choice of the regularization space is quite vital and depends on the nature of the sought 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 discontinuous or rapidly varying coefficients, a total-variation semi-norm has been used extensively in recent years. The success of the total variation regularization comes at the cost that the optimization is done in a non-reflexive Banach space, and the regularizer is nonsmooth. Hence, the computations rely on some smoothing.

In the following, we describe two sets of assumptions for the regularizer; the first one subsumes the total-variation regularization, and the second recovers the quadratic regularizer. The nonsmooth regularization framework imposes the following conditions on the regularizer.

(H1 (a)) The parameter space B, which is a Banach space, is continuously embedded in a Banach space L. There is another Banach space $\widehat{B}$ that is compactly embedded in L. The set A of admissible parameters is a subset of $B \cap \widehat{B}$, closed and bounded in B and also closed in L.

(H1 (b)) $R : \widehat{B} \to \mathbb{R}$ is positive, convex, and lower-semicontinuous in $\|\cdot\|_L$ such that

$$R(\ell) \geq \tau_1 \|\ell\|_{\widehat{B}} - \tau_2, \quad \text{for every } \ell \in A, \text{ for some } \tau_1 > 0, \tau_2 > 0. \tag{26}$$

(H1 (c)) For any $\{\ell_n\} \subset B$ with $\ell_n \to \ell$ in L, any bounded $\{u_n\} \subset V$, and fixed $v \in V$, we have

$$a(\ell_n - \ell, u_n, v) \to 0. \tag{27}$$

The following assumption is for the quadratic regularizer:

(H2) The set A belongs to a Hilbert space H that is compactly embedded in the space B.

An example for (H2) is $B = L^\infty(\Omega)$ and $H = H^2(\Omega)$, for a suitable domain Ω. The properties (H1) are inspired by the use of total variation regularization in the identification of discontinuous coefficients (see [35]). Recall that the total variation of $f \in L^1(\Omega)$ reads

$$\mathrm{TV}(f) = \sup\left\{\int_\Omega f\,(\nabla\cdot g) : g \in \left(C_0^1(\Omega)\right)^N,\ |g(x)| \le 1 \text{ for all } x \in \Omega\right\},$$

where $|\cdot|$ is the Euclidean norm. Clearly, if $f \in W^{1,1}(\Omega)$, then $\mathrm{TV}(f) = \int_\Omega |\nabla f|$.

If $f \in L^1(\Omega)$ satisfies $\mathrm{TV}(f) < \infty$, then f is said to have bounded variation, and $\mathrm{BV}(\Omega)$ is defined by $\mathrm{BV}(\Omega) = \{f \in L^1(\Omega) : \mathrm{TV}(f) < \infty\}$ with norm $\|f\|_{\mathrm{BV}(\Omega)} = \|f\|_{L^1(\Omega)} + \mathrm{TV}(f)$. The functional $\mathrm{TV}(\cdot)$ is a seminorm on $\mathrm{BV}(\Omega)$ and is often called the BV-seminorm.

We set $B = L^\infty(\Omega)$, $L = L^1(\Omega)$, $\widehat{B} = \mathrm{BV}(\Omega)$, and $R(\ell) = TV(\ell)$, and define

$$A = \{\ell \in L^\infty \,|\, 0 < c_1 \le \ell(x) \le c_2, \text{ a.e. in } \Omega,\ \mathrm{TV}(\ell) \le c_3 < \infty\}, \tag{28}$$

where c_1, c_2, and c_3 are positive constants. Clearly, A is bounded in $\|\cdot\|_{\widehat{B}}$ and compact in L. It is known that $L^\infty(\Omega)$ is continuously embedded in $L^1(\Omega)$, $\mathrm{BV}(\Omega)$ is compactly embedded in $L^1(\Omega)$, and $TV(\cdot)$ is convex and lower-semicontinuous in $L^1(\Omega)$-norm. Thus, (i) and (ii) hold.

In the following, for simplicity, we use $R(\ell) := \|\cdot\|_H^2$ as the regularizer. Minor changes in the proof allow incorporating the nonsmooth regularization framework given in (H1).

We shall now approximate (18) by the following family of regularized OLS-based optimization problems: For $n \in \mathbb{N}$, find $\ell_n \in A$ by solving

$$\min_{\ell \in A} J_{\kappa_n}(\ell) := \frac{1}{2}\|u_n(\ell) - \bar{z}_{\delta_n}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p_n(\ell) - \hat{z}_{\delta_n}\|_{\mathbb{Q}}^2 + \kappa_n R(\ell), \tag{29}$$

where $\kappa_n > 0$ is the regularization parameter with $\kappa_n \to 0$, and $(u_n(\ell), p_n(\ell))$ is the unique solution of the regularized saddle point problem (23a), that is,

$$a(\ell, u_n, v) + \varepsilon_n \langle u_n - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n) = f_{\nu_n}(v), \quad \text{for every } v \in V, \tag{30a}$$

$$b(u_n, q) - \varepsilon_n \langle p_n - \hat{z}_{\delta_n}, q\rangle_Q = g_{\nu_n}(q), \quad \text{for every } q \in Q. \tag{30b}$$

The following result proves that the OLS-based regularized optimization problem (29) approximates the original OLS-based optimization problem (18). To obtain the necessary optimality condition, we assume that the admissible set is in the interior of A, however, with an abuse of notation, we still denote it by A.

Theorem 2.3: *Assume that for each $\ell \in A, \mathcal{U}(\ell)$ is nonempty, the sets A and $\mathcal{U}(A)$ are bounded, and $V = \mathbb{V}$ and $Q = \mathbb{Q}$. Then, the optimization problem (18) has a solution, and for each $n \in \mathbb{N}$, the regularized optimization problem (29) has a solution ℓ_n. Moreover, there is a subsequence $\{\ell_n\}$ converging in $\|\cdot\|_B$ to a solution* *of* (18). *Finally, for any solution* ℓ_n *of* (29), *there is a unique* $(w_n, t_n) \in V \times Q$ *such that*

$$a(\ell_n, w_n, v) + \varepsilon_n \langle w_n, v\rangle_V + b(v, t_n) = \left\langle \bar{z}_{\delta_n} - u_n(\ell_n), v\right\rangle_{\mathbb{V}}, \quad \textit{for all } v \in V, \tag{31a}$$

$$b(w_n, q) - \varepsilon_n \langle t_n, q\rangle_Q = \left\langle \hat{z}_{\delta_n} - p_n(\ell_n), q\right\rangle_{\mathbb{Q}}, \quad \textit{for all } q \in Q, \tag{31b}$$

$$a(\ell - \ell_n, u_n(\ell_n), w_n) \geq \kappa_n(R(\ell_n) - R(\ell)), \quad \textit{for every } \ell \in A. \tag{31c}$$

Proof: We will divide the proof into several parts as follows:

(1) The OLS-based optimization problem (18) has a solution. By assumption, for each parameter $\ell \in A$, the solution set $\mathcal{U}(\ell)$ is nonempty and consequently the OLS-based optimization problem (18) is well-defined. Since for each $\ell \in A$, $J(\ell)$ is bounded from below, there is a minimizing sequence $\{\ell_n\}$ in A such that

$$\lim_{n\to\infty} J(\ell_n) = \inf\{J(\ell),\ \ell \in A\}.$$

By assumption, the set A is bounded in H, and hence 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 $\|\cdot\|_B$ to some $\bar{\ell} \in A$. Let $(u_n, p_n) \in \mathcal{U}(\ell_n)$ be arbitrarily chosen. Since by assumption the set $\mathcal{U}(A)$ is bounded, the sequence $\{(u_n, p_n)\}$ remains bounded, and therefore it possesses a weakly convergent subsequence. Let $\{(u_n, p_n)\}$ be the subsequence converging weakly to some $(\bar{u}, \bar{p}) \in V \times Q$. We claim that $(\bar{u}, \bar{p}) \in \mathcal{U}(\bar{\ell})$. Note that the definition of (ℓ_n, u_n, p_n) implies that

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

We rearrange the above saddle point problem to obtain

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

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

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

and hence $(\bar{u}, \bar{p}) \in \mathcal{U}(\bar{\ell})$. The optimality of $\bar{\ell}$ is a consequence of the weak lower-semicontinuity of a norm.

(2) For every $n \in \mathbb{N}$, the OLS-based regularized problem (29) has a solution ℓ_n. The arguments used in the first part of this proof ensure that for any fixed $n \in \mathbb{N}$, the regularized optimization problem (29) has a solution ℓ_n. In fact, we need to notice that for any fixed $n \in \mathbb{N}$, the regularized saddle point problem (23a) is uniquely solvable and the solution is bounded.

(3) The sequence $\{\ell_n\}$ of solutions of (29) and the regularized solutions $\{(u_n(\ell_n), p_n(\ell_n))\}$ of (30a) are bounded. Indeed the sequence $\{\ell_n\}$ is bounded by the assumption that A is bounded in H.

To prove that the sequence $\{(u_n, p_n)\}$, where $(u_n, p_n) \equiv (u_n(\ell_n), p_n(\ell_n))$, is bounded, we note that

$$\begin{aligned} a(\ell_n, u_n, v) + \varepsilon_n \langle u_n - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n) &= f_{v_n}(v), \quad \text{for every } v \in V, \\ b(u_n, q) - \varepsilon_n \langle p_n - \hat{z}_{\delta_n}, q\rangle_Q &= g_{v_n}(q), \quad \text{for every } q \in Q. \end{aligned}$$

We will prove that the sequence $\{(u_n, p_n)\}$ is bounded by using the assumption that for every $\ell \in A$, the solution set $\mathcal{U}(\ell)$ of the original saddle point problem (3a) is nonempty. For any $n \in \mathbb{N}$, and for $\ell_n \in A$, we choose an element $(\tilde{u}_n, \tilde{p}_n) \in \mathcal{U}(\ell_n)$ arbitrarily. Furthermore, since $\mathcal{U}(A)$ is bounded by assumption, the sequence $\{(\tilde{u}_n, \tilde{p}_n)\}$ is bounded. Moreover, we have

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

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

$$\begin{aligned} a(\ell_n, u_n - \tilde{u}_n, v) + \varepsilon_n \langle u_n, v\rangle_V + b(v, p_n - \tilde{p}_n) &= (f_{v_n} - f)(v) + \varepsilon_n \langle \bar{z}_{\delta_n}, v\rangle_V, \\ b(u_n - \tilde{u}_n, q) - \varepsilon_n \langle p_n, q\rangle_Q &= (g_{v_n} - g)(q) - \varepsilon_n \langle \hat{z}_{\delta_n}, q\rangle_Q, \end{aligned}$$

for each $v \in V$ and each $q \in Q$. We set $v = u_n - \tilde{u}_n$ and $q = p_n - \tilde{p}_n$ in the above system, combine the resulting equations, and use the fact that $a(\ell_n, \tilde{u}_n - u_n, \tilde{u}_n - u_n) \geq 0$ to obtain

$$\begin{aligned} &\varepsilon_n \|u_n\|_V^2 + \varepsilon_n \|p_n\|_Q^2 \\ &\quad \leq \varepsilon_n \langle u_n, \tilde{u}_n\rangle_V + \varepsilon_n \langle p_n, \tilde{p}_n\rangle_Q + (f_{v_n} - f)(u_n - \tilde{u}_n) + (g - g_{v_n})(p_n - \tilde{p}_n) \\ &\qquad + \varepsilon_n \langle \bar{z}_{\delta_n}, u_n - \tilde{u}_n\rangle_V + \varepsilon_n \langle \hat{z}_{\delta_n}, p_n - \tilde{p}_n\rangle_Q \\ &\quad \leq \varepsilon_n \|u_n\|_V \|\tilde{u}_n\|_V + \varepsilon_n \|p_n\|_Q \|\tilde{p}_n\|_Q + (v_n + \varepsilon_n \|\hat{z}_{\delta_n}\|_Q) \left[\|p_n\|_Q + \|\tilde{p}_n\|_Q\right] \\ &\qquad + (v_n + \varepsilon_n \|\bar{z}_{\delta_n}\|_V) \left[\|\tilde{u}_n\|_V + \|u_n\|_V\right], \end{aligned}$$

which further results in

$$\begin{aligned} \|u_n\|_V^2 + \|p_n\|_Q^2 \leq\; & \|u_n\|_V \left[\|\tilde{u}_n\|_V + v_n \varepsilon_n^{-1} + \|\bar{z}_{\delta_n}\|_V\right] \\ & + \|p_n\|_Q \left[\|\tilde{p}_n\|_Q + v_n \varepsilon_n^{-1} + \|\hat{z}_{\delta_n}\|_Q\right] \\ & + (v_n + \varepsilon_n \|\bar{z}_{\delta_n}\|_V) \|\tilde{u}_n\|_V + (v_n + \varepsilon_n \|\hat{z}_{\delta_n}\|_Q) \|\tilde{p}_n\|_Q, \end{aligned}$$

and consequently, we have

$$\|u_n\|_V^2 + \|p_n\|_Q^2 \leq c_1\|u_n\|_V + c_2\|p_n\|_Q + c_3,$$

where

$$\begin{aligned}
c_1 &:= \max\left\{\|\tilde{u}_n\|_V + \nu_n\varepsilon_n^{-1} + \|\bar{z}_{\delta_n}\|_V\right\},\\
c_2 &:= \max\left\{\|\tilde{p}_n\|_Q + \nu_n\varepsilon_n^{-1} + \|\hat{z}_{\delta_n}\|_Q\right\},\\
c_3 &:= \max\left\{(\nu_n + \varepsilon_n\|\bar{z}_{\delta_n}\|_V)\|\tilde{u}_n\|_V + (\nu_n + \varepsilon_n\|\hat{z}_{\delta_n}\|_Q)\|\tilde{p}_n\|_Q\right\},
\end{aligned}$$

are positive constants. We therefore deduce that $\{(u_n, p_n)\}$ is uniformly bounded.

(4) The sequence $\{(\ell_n, u_n(\ell_n), p_n(\ell_n))\}$ has a strong-weak limit point $(\bar{\ell}, \bar{u}, \bar{p})$ with $(\bar{u}, \bar{p}) \in \mathcal{U}(\bar{\ell})$. Since $\{\ell_n\}$ is bounded in H, due to the compact imbedding of H into B, there is a strongly convergent subsequence. Let $\{\ell_n\}$ be a subsequence that converges strongly to some $\bar{\ell} \in A$ in $\|\cdot\|_B$. Furthermore, since the spaces V and Q are reflexive, the sequence $\{(u_n, p_n)\}$ also has a weakly convergent subsequence. Using the same notation for subsequences, let $\{(u_n, p_n)\}$ be a subsequence converging weakly to some $(\bar{u}, \bar{p})$. We claim that $(\bar{u}, \bar{p}) \in \mathcal{U}(\bar{\ell})$. Since ℓ_n is a solution of (29), we have

$$\begin{aligned}
a(\ell_n, u_n, v) + \varepsilon_n\langle u_n - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n) &= f_{\nu_n}(v), \quad \text{for every } v \in V,\\
b(u_n, q) - \varepsilon_n\langle p_n - \hat{z}_{\delta_n}, q\rangle_Q &= g_{\nu_n}(q), \quad \text{for every } q \in Q,
\end{aligned}$$

or equivalently,

$$\begin{aligned}
&a(\ell_n - \bar{\ell}, u_n, v) + a(\bar{\ell}, u_n - \bar{u}, v) + a(\bar{\ell}, \bar{u}, v) + \varepsilon_n\langle u_n - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n)\\
&\quad = f_{\nu_n}(v), \quad \text{for every } v \in V,\\
&sb(u_n, q) - \varepsilon_n\langle p_n - \hat{z}_{\delta_n}, q\rangle_Q = g_{\nu_n}(q), \quad \text{for every } q \in Q,
\end{aligned}$$

which due to the imposed conditions, when passed to the limit $n \to \infty$, 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 $(\bar{u}, \bar{p}) \in \mathcal{U}(\bar{\ell})$.

(5) For $\hat{\ell} \in A$, the sequence of the regularized solutions $\{(u_n(\hat{\ell}), p_n(\hat{\ell})\}$ converges to $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$ which is the unique solution of the following

variational inequality:

$$\langle \hat{u}(\hat{\ell}) - \bar{z}, v - \hat{u}(\hat{\ell})\rangle_V + \langle \hat{p}(\hat{\ell}) - \hat{z}, q - \hat{p}(\hat{\ell})\rangle_Q \geq 0, \quad \text{for every } (v,q) \in \mathcal{U}(\hat{\ell}). \tag{32}$$

By the definition of $(u_n(\hat{\ell}), p_n(\hat{\ell}))$, we have

$$a(\hat{\ell}, u_n(\hat{\ell}), v) + \varepsilon_n \langle u_n(\hat{\ell}) - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n(\hat{\ell})) = f_{\nu_n}(v), \quad \text{for every } v \in V, \tag{33a}$$

$$b(u_n(\hat{\ell}), q) - \varepsilon_n \langle p_n(\hat{\ell}) - \hat{z}_{\delta_n}, q\rangle_Q = g_{\nu_n}(q), \quad \text{for every } q \in Q. \tag{33b}$$

By the arguments used in the earlier in this proof, it can be shown that $\{(u_n(\hat{\ell}), p_n(\hat{\ell}))\}$ is bounded and there is a subsequence converging weakly to some $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$.

The solution set $\mathcal{U}(\hat{\ell})$ is closed and convex, and consequently we can consider the following variational inequality: Find $(\tilde{u}(\hat{\ell}), \tilde{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$ such that

$$\langle \tilde{u}(\hat{\ell}) - \bar{z}, v - \tilde{u}(\hat{\ell})\rangle + \langle \tilde{p}(\hat{\ell}) - \hat{z}, q - \tilde{p}(\hat{\ell})\rangle \geq 0, \quad \text{for every } (v,q) \in \mathcal{U}(\hat{\ell}). \tag{34}$$

Variational inequality (34) has a unique solution $(\tilde{u}(\hat{\ell}), \tilde{p}(\hat{\ell}))$. We set $(v,q) = (\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ in (34) to get

$$\langle \tilde{u}(\hat{\ell}) - \bar{z}, \hat{u}(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle + \langle \tilde{p}(\hat{\ell}) - \hat{z}, \hat{p}(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle \geq 0. \tag{35}$$

Since $(\tilde{u}(\hat{\ell}), \tilde{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$, we also have

$$a(\hat{\ell}, \tilde{u}(\hat{\ell}), v) + b(v, \tilde{p}(\hat{\ell})) = f(v), \quad \text{for every } v \in V, \tag{36a}$$

$$b(\tilde{u}(\hat{\ell}), q) = g(q), \quad \text{for every } q \in Q. \tag{36b}$$

We combine (33a) and (36a) to deduce that for every $v \in V$ and every $q \in Q$, we have

$$\begin{aligned} a(\hat{\ell}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell}), v) + \varepsilon_n \langle u_n(\hat{\ell}) - \bar{z}_{\delta_n}, v\rangle_V + b(v, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})) &= (f_{\nu_n} - f)(v), \\ b(u_n(\hat{\ell}) - \tilde{u}(\hat{\ell}), q) - \varepsilon_n \langle p_n(\hat{\ell}) - \hat{z}_{\delta_n}, q\rangle_Q &= (g_{\nu_n} - g)(q). \end{aligned}$$

We set $v = u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})$ and $q = p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})$ in the above system to obtain

$$\begin{aligned} & a(\hat{\ell}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell}), u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})) + \varepsilon_n \langle u_n(\hat{\ell}), u_n(\hat{\ell}) \qquad (14) \\ & \quad - \tilde{u}(\hat{\ell})\rangle_V + \varepsilon_n \langle p_n(\hat{\ell}), p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q \\ & \quad = (f_{\nu_n} - f)(u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})) - (g_{\nu_n} - g)(p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})) \\ & \qquad + \varepsilon_n \langle \bar{z}_{\delta_n} - \bar{z}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V + \varepsilon_n \langle \bar{z}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V \\ & \qquad + \varepsilon_n \langle \hat{z}_{\delta_n} - \hat{z}, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q + \varepsilon_n \langle \hat{z}, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q, \end{aligned} \tag{37}$$

and by using the fact that $a(\hat{\ell}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell}), u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})) \geq 0$, further deduce that

$$\begin{aligned}
&\left[\nu_n \varepsilon_n^{-1} + \delta_n\right] \|u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\|_V + \left[\nu_n \varepsilon_n^{-1} + \delta_n\right] \|p_n(\hat{\ell}) \\
&\quad - \tilde{p}(\hat{\ell})\|_Q + \langle \bar{z}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V \\
&\quad + \langle \hat{z}, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q + \langle u_n(\hat{\ell}), \tilde{u}(\hat{\ell})\rangle_V + \langle p_n(\hat{\ell}), \tilde{p}(\hat{\ell})\rangle_Q \geq \langle u_n(\hat{\ell}), u_n(\hat{\ell})\rangle_V \\
&\quad + \langle p_n(\hat{\ell}), p_n(\hat{\ell})\rangle. \qquad (38)
\end{aligned}$$

Since any norm is weakly lower-semicontinuous, we have

$$\|\hat{u}(\hat{\ell})\|_V^2 + \|\hat{p}(\hat{\ell})\|_Q^2 \leq \liminf_{n\to\infty} \|u_n(\hat{\ell})\|_V^2 + \liminf_{n\to\infty} \|p_n(\hat{\ell})\|_Q^2,$$

and, consequently by using (38), we obtain

$$\langle \hat{u}(\hat{\ell}) - \bar{z}, \tilde{u}(\hat{\ell}) - \hat{u}(\hat{\ell})\rangle_V + \langle \hat{p}(\hat{\ell}) - \hat{z}, \tilde{p}(\hat{\ell}) - \hat{p}(\hat{\ell})\rangle_Q \geq 0, \qquad (39)$$

which, when combined with (35), implies that

$$\begin{aligned}
0 &\geq \langle \hat{u}(\hat{\ell}) - \tilde{u}(\hat{\ell}), \hat{u}(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V + \langle \hat{p}(\hat{\ell}) - \tilde{p}(\hat{\ell}), \hat{p}(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q \\
&= \|\hat{u}(\hat{\ell}) - \tilde{u}(\hat{\ell})\|_V^2 + \|\hat{p}(\hat{\ell}) - \tilde{p}(\hat{\ell})\|_Q^2,
\end{aligned}$$

and hence $\hat{u}(\hat{\ell}) = \tilde{u}(\hat{\ell})$ and $\hat{p}(\hat{\ell}) = \tilde{p}(\hat{\ell})$. Since $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ is unique, the whole sequence $\{u_n(\hat{\ell}), p_n(\hat{\ell})\}$ converges weakly to $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$. The prove that the convergence is strong, we rewrite (14) as follows

$$\begin{aligned}
&\varepsilon_n \|u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\|_V^2 + \varepsilon_n \|p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\|_Q^2 \\
&\quad = \varepsilon_n \langle u_n(\hat{\ell}) - \tilde{u}(\hat{\ell}), u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V + \varepsilon_n \langle p_n(\hat{\ell}) - \tilde{p}(\hat{\ell}), p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q \\
&\quad \leq (f_{\nu_n} - f)(u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})) - (g_{\nu_n} - g)(p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})) + \varepsilon_n \langle \bar{z}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V \\
&\qquad + \varepsilon_n \langle \bar{z}_{\delta_n} - \bar{z}, u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V + \varepsilon_n \langle \hat{z}, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q \\
&\qquad + \varepsilon_n \langle \hat{z}_{\delta_n} - \hat{z}, p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q + \varepsilon_n \langle \tilde{u}(\hat{\ell}), u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\rangle_V \\
&\quad + \varepsilon_n \langle \tilde{p}(\hat{\ell}), p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\rangle_Q,
\end{aligned}$$

which after a simple calculation implies that

$$\lim_{n\to\infty} \|u_n(\hat{\ell}) - \tilde{u}(\hat{\ell})\|_V^2 + \lim_{n\to\infty} \|p_n(\hat{\ell}) - \tilde{p}(\hat{\ell})\|_Q^2 \leq 0,$$

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

(6) For $\hat{\ell} \in A$, the regularized solutions $\{(u_n(\hat{\ell}), p_n(\hat{\ell}))\}$ converge to $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$ with

$$\begin{aligned}
&\|\hat{u}(\hat{\ell}) - \bar{z}\|_V^2 + \|\hat{p}(\hat{\ell}) - \hat{z}\|_Q^2 \\
&\quad \leq \|\breve{u}(\hat{\ell}) - \bar{z}\|_V^2 + \|\breve{p}(\hat{\ell}) - \hat{z}\|_Q^2, \quad \text{for all } (\breve{u}(\hat{\ell}), \breve{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell}).
\end{aligned} \tag{40}$$

As shown in the previous part, the sequence $\{(u_n(\hat{\ell}), p_n(\hat{\ell}))\}$ converges strongly to $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$ where $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ is the unique solution of (34). That is, for an arbitrary $(\breve{u}(\hat{\ell}), \breve{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$, we have

$$\langle \hat{u}(\hat{\ell}) - \bar{z}, \breve{u}(\hat{\ell}) - \hat{u}(\hat{\ell}) \rangle_V + \langle \hat{p}(\hat{\ell}) - \hat{z}, \breve{p}(\hat{\ell}) - \hat{p}(\hat{\ell}) \rangle_Q \geq 0$$

which can be rearranged as

$$\begin{aligned}
&\|\hat{u}(\hat{\ell}) - \bar{z}\|_V^2 + \|\hat{p}(\hat{\ell}) - \hat{z}\|_Q^2 \\
&\quad \leq \|\hat{u}(\hat{\ell}) - \bar{z}\|_V \|\breve{u}(\hat{\ell}) - \bar{z}\|_V + \|\hat{p}(\hat{\ell}) - \hat{z}\|_Q \|\breve{p}(\hat{\ell}) - \hat{z}\|_Q \\
&\quad \leq \frac{1}{2}\|\hat{u}(\hat{\ell}) - \bar{z}\|_V^2 + \frac{1}{2}\|\breve{u}(\hat{\ell}) - \bar{z}\|_V^2 + \frac{1}{2}\|\hat{p}(\hat{\ell}) - \hat{z}\|_Q^2 + \frac{1}{2}\|\breve{p}(\hat{\ell}) - \hat{z}\|_Q^2
\end{aligned}$$

and consequently

$$\|\hat{u}(\hat{\ell}) - \bar{z}\|_V^2 + \|\hat{p}(\hat{\ell}) - \hat{z}\|_Q^2 \leq \|\breve{u}(\hat{\ell}) - \bar{z}\|_V^2 + \|\breve{p}(\hat{\ell}) - \hat{z}\|_Q^2,$$

which implies that $(\hat{u}(\hat{\ell}), \hat{p}(\hat{\ell}))$ is the closest element to $(\bar{z}, \hat{z})$ among all $(\breve{u}(\hat{\ell}), \breve{p}(\hat{\ell})) \in \mathcal{U}(\hat{\ell})$.

(7) The element $\bar{\ell}$ is a minimizer of the OLS-based optimization problem (18). We recall that $\{\ell_n\}$ is the sequence of the regularized solutions of (29) which converges in $\|\cdot\|_B$ to $\bar{\ell} \in A$. For any fixed $n \in \mathbb{N}$, the optimality of $\ell_n \in A$ for (29) means that for each $\ell \in A$, we have

$$\begin{aligned}
J_{\kappa_n}(\ell_n) &:= \tfrac{1}{2}\|u_n(\ell_n) - \bar{z}_{\delta_n}\|_V^2 + \tfrac{1}{2}\|p_n(\ell_n) - \hat{z}_{\delta_n}\|_Q^2 + \kappa_n R(\ell_n) \\
&\leq \tfrac{1}{2}\|u_n(\ell) - \bar{z}_{\delta_n}\|_V^2 + \tfrac{1}{2}\|p_n(\ell) - \hat{z}_{\delta_n}\|_Q^2 + \kappa_n R(\ell),
\end{aligned} \tag{41}$$

where $(u_n(\ell), p_n(\ell))$, is the solution of regularized saddle point problem (23a) for the parameter ℓ.

Let $\hat{\ell}$ be a solution of (18), and let $(\hat{u}, \hat{p})$ be the corresponding solution of the saddle point problem. We will use $\hat{\ell}$ to generate a well-behaved feasible point for (41). Indeed, we take the sequence $\{(u_n(\hat{\ell}), p_n(\hat{\ell}))\}$ of the solutions of the regularized saddle point problem corresponding to the fixed coefficient $\hat{\ell}$, which evidently renders a feasible point for (41). Moreover, the limit of $\{(u_n(\hat{\ell}), p_n(\hat{\ell}))\}$ is characterized by (40).

Therefore,

$$\begin{aligned}
J(\bar{\ell}) &= \frac{1}{2}\|\bar{u}(\bar{\ell}) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|\bar{p}(\bar{\ell}) - \hat{z}\|_{\mathbb{Q}}^2 \\
&\le \liminf_{n\to\infty}\left\{\frac{1}{2}\|u_n(\bar{\ell}_n) - \bar{z}_{\delta_n}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p_n(\bar{\ell}_n) - \hat{z}_{\delta_n}\|_{\mathbb{Q}}^2 + \kappa_n R(\bar{\ell}_n)\right\}, \\
&\le \limsup_{n\to\infty}\frac{1}{2}\|u_n(\hat{\ell}) - \bar{z}\|_{\mathbb{V}}^2 + \limsup_{n\to\infty}\frac{1}{2}\|p_n(\hat{\ell}) - \hat{z}\|_{\mathbb{Q}}^2 \\
&= \frac{1}{2}\|\hat{u}(\hat{\ell}) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|\hat{p}(\hat{\ell}) - \hat{z}\|_{\mathbb{Q}}^2 \\
&\le \frac{1}{2}\|\breve{u}(\breve{\ell}) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|\breve{p}(\breve{\ell}) - \hat{z}\|_{\mathbb{Q}}^2,
\end{aligned}$$

for every $(\breve{u}(\breve{\ell}), \breve{p}(\breve{\ell})) \in \mathcal{U}(\breve{\ell})$ and every $\breve{\ell} \in A$. Consequently, $\bar{\ell} \in A$ is a minimizer of (18).

(8) The conditions (31a) are valid. We begin by noting that a necessary optimality condition for ℓ_n to be a minimizer of (29) is the following variational inequality

$$DJ_n(\ell_n)(\ell - \ell_n) \ge \kappa_n(R(\ell_n) - R(\ell)), \quad \text{for every } \ell \in A, \tag{42}$$

where

$$\begin{aligned}
J_n(\ell_n) &= \frac{1}{2}\|u_n(\ell_n) - \bar{z}_{\delta_n}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p_n(\ell_n) - \hat{z}_{\delta_n}\|_{\mathbb{Q}}^2, \\
DJ_n(\ell)(\delta\ell) &= \langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\rangle_{\mathbb{V}} + \langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\rangle_{\mathbb{Q}}.
\end{aligned}$$

For $n \in \mathbb{N}$, we consider the adjoint saddle point problem of finding $(w_n, t_n) \in V \times Q$ such that

$$a(\ell_n, w_n, v) + \varepsilon_n\langle w_n, v\rangle_V + b(v, t_n) = \left\langle \bar{z}_{\delta_n} - u_n(\ell_n), v\right\rangle_{\mathbb{V}}, \quad \text{for all } v \in V, \tag{43a}$$

$$b(w_n, q) - \varepsilon_n\langle t_n, q\rangle_Q = \left\langle \hat{z}_{\delta_n} - p_n(\ell_n), q\right\rangle_{\mathbb{Q}}, \quad \text{for all } q \in Q. \tag{43b}$$

Clearly (43) has a unique solution $(w_n, t_n) \in V \times Q$. We substitute $v = Du_n(\ell_n)(\ell - \ell_n)$ and $q = Dp_n(\ell_n)(\ell - \ell_n)$ in the above identities and combine them to obtain

$$\begin{aligned}
&\langle Du_n(\ell_n)(\ell - \ell_n), u_n(\ell_n) - \bar{z}_{\delta_n}\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell_n)(\ell - \ell_n), p_n(\ell_n) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} \\
&\quad = -a(\ell_n, w_n, Du_n(\ell_n)(\ell - \ell_n)) - \varepsilon_n\langle w_n, Du_n(\ell_n)(\ell - \ell_n)\rangle_V \\
&\qquad - b(Du_n(\ell_n)(\ell - \ell_n), t_n) - b(w_n, Dp_n(\ell_n)(\ell - \ell_n))
\end{aligned}$$

$$
\begin{aligned}
&+ \varepsilon_n \langle t_n, Dp_n(\ell_n)(\ell - \ell_n)\rangle_Q \\
&= -a(\ell_n, Du_n(\ell_n)(\ell - \ell_n), w_n) - \varepsilon_n \langle Du_n(\ell_n)(\ell - \ell_n), w_n\rangle_V \\
&\quad - b(Du_n(\ell_n)(\ell - \ell_n), t_n) - b(w_n, Dp_n(\ell_n)(\ell - \ell_n)) \\
&+ \varepsilon_n \langle Dp_n(\ell_n)(\ell - \ell_n), t_n\rangle_Q \\
&= a(\ell - \ell_n, u_n(\ell_n), w_n),
\end{aligned}
$$

where we used the following identity which follows from (24a) by taking $v = w_n$ and $q = t_n$

$$
\begin{aligned}
&- a(\ell, Du_n(\ell)(\ell - \ell_n), w_n) - \varepsilon_n \langle Du_n(\ell)(\ell - \ell_n), w_n\rangle_V - b(w_n, Dp_n(\ell)(\ell - \ell_n)) \\
&\quad - b(Du_n(\ell)(\ell - \ell_n), t_n) + \varepsilon_n \langle Dp_n(\ell)(\ell - \ell_n), t_n\rangle_Q = a(\ell - \ell_n, u_n(\ell), w_n),
\end{aligned}
$$

and (31c) follows by using the above expression in (42). The proof is complete. ■

Remark 2.4: The data in the regularized saddle point problem steers the regularized solutions towards the solution of (3a) that is closest to $(\bar{z}, \hat{z})$. However, for the following regularized problem

$$
a(\ell, u_n(\ell), v) + \varepsilon_n \langle u_n(\ell), v\rangle_V + b(v, p_n(\ell)) = f_{v_n}(v), \quad \text{for every } v \in V, \tag{44a}
$$

$$
b(u_n, q) - \varepsilon_n \langle p_n(\ell), q\rangle_Q = g_{v_n}(q), \quad \text{for every } q \in Q, \tag{44b}
$$

the regularized solutions converge to a minimum norm solution of (3a). We also emphasize that the imposed conditions don't ensure that the sequence of the adjoint solutions (w_n, t_n) is uniformly bounded. However, if the sequence of (w_n, t_n) is bounded, then by passing (31a) to limit, we shall derive optimality conditions for (18).

We will now give a particular case of Theorem 2.3 under the standard assumptions that the trilinear form a is continuous and elliptic (on V) and the bilinear form b is continuous and satisfies the Inf-Sup condition. Then, we can regularize (3a) by using the continuous, and elliptic bilinear form $c : Q \times Q \mapsto \mathbb{R}$. In the following, for simplicity, we don't involve data perturbation.

We consider the following OLS based optimization problem of finding $\ell \in A$ by solving

$$
\min_{\ell \in A} J(\ell) := \tfrac{1}{2}\|u(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \tfrac{1}{2}\|p(\ell) - \hat{z}\|_{\mathbb{Q}}^2 + \kappa \|\ell\|_H^2, \tag{45}
$$

where, for $\ell \in A$, $(u(\ell), p(\ell))$ is the unique solution of (3a), $(\bar{z}, \hat{z}) \in \mathbb{V} \times \mathbb{Q}$ is the given data, $\kappa > 0$ is the regularization parameter, and $\|\cdot\|_H^2$ is the quadratic regularizer.

We will approximate (45) by the following family of regularized OLS-based optimization problems: For $n \in \mathbb{N}$, find $\ell_n \in A$ by solving

$$\min_{\ell \in A} J_\kappa(\ell) := \frac{1}{2}\|u_n(\ell) - \bar{z}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p_n(\ell) - \hat{z}\|_{\mathbb{Q}}^2 + \kappa\|\ell\|_H^2, \tag{46}$$

where $\kappa > 0$ is the regularization parameter, and $(u_n(\ell), p_n(\ell))$ is the unique solution of the following saddle point problem:

$$a(\ell, u_n, v) + b(v, p_n) = f(v), \quad \text{for every } v \in V, \tag{47a}$$

$$b(u_n, q) - \varepsilon_n c(p_n, q) = g(q), \quad \text{for every } q \in Q. \tag{47b}$$

The following particular case of Theorem 2.3 shows that (46) approximates (45):

Theorem 2.5: *Assume that* (1), (2), (6) (*on* V), *and* (7) *hold. Then, the optimization problem* (45) *has a solution, and for each* $n \in \mathbb{N}$, *the regularized optimization problem* (46) *has a solution* ℓ_n. *Moreover, there is a subsequence* $\{\ell_n\}$ *converging in* $\|\cdot\|_B$ *to a solution of* (45).

3. Evaluation of the first-order and the second-order derivatives

One of the significant drawbacks of employing an OLS-based approach is the computation of the derivatives of the OLS functional, which involve computationally expensive evaluation of the solution map. Adjoint methods provide a computationally efficient framework for computing the derivatives of the OLS functional and have been explored intensively. Recent developments in adjoint methods can be found in [36–40]. Furthermore, in [41], the first-order and the second-order adjoint methods were applied to nearly incompressible elasticity imaging.

In the following, we give a quick derivation of the first-order and the second-order derivative formulas for the OLS objective by using the adjoint method. We will use the discrete variants of these formulas to compute the gradient and Hessian of the OLS objective.

3.1. Evaluation of the first-order derivative

We recall that the OLS objective (without the regularizer), with the regularized saddle point problem as the constraint, reads

$$\min_{\ell \in A} J_n(\ell) := \frac{1}{2}\|u_n(\ell) - \bar{z}_{\delta_n}\|_{\mathbb{V}}^2 + \frac{1}{2}\|p_n(\ell) - \hat{z}_{\delta_n}\|_{\mathbb{Q}}^2, \tag{48}$$

where $(u_n(\ell), p_n(\ell))$ is the unique solution of regularized saddle point problem (23a), that is,

$$a(\ell, u_n(\ell), v) + \varepsilon_n \langle u_n(\ell), v\rangle_V + b(v, p_n(\ell))$$
$$= f_{v_n}(v) + \varepsilon_n \langle \bar{z}_{\delta_n}, v\rangle_V, \quad \text{for each } v \in V, \tag{49a}$$
$$b(u_n(\ell), q) - \varepsilon_n \langle p_n(\ell), q\rangle = g_{v_n}(q) - \varepsilon_n \langle \hat{z}_{\delta_n}, q\rangle_Q, \quad \text{for each } q \in Q. \tag{49b}$$

As seen earlier, the derivative of J_n at $\ell \in A$ in a direction $\delta\ell$ is given by

$$DJ_n(\ell)(\delta\ell) = \left\langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}}.$$

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

$$L_n(\ell, v, q) = J_n(\ell) + a(\ell, u_n(\ell), v) + \varepsilon_n \langle u_n(\ell), v\rangle_V + b(v, p_n(\ell)) + b(u_n(\ell), q)$$
$$- \varepsilon_n \langle p_n(\ell), q\rangle_Q - f_{v_n}(v) - \varepsilon_n \langle \bar{z}_{\delta_n}, v\rangle_V - g_{v_n}(q) + \varepsilon_n \langle \hat{z}_{\delta_n}, q\rangle_Q.$$

Since $(u_n(\ell), p_n(\ell))$ is the unique solution of (49a), the following identity holds

$$L_n(\ell, v, q) = J_n(\ell), \quad \text{for every } (v, q) \in V \times Q,$$

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

$$\partial_\ell L_n(\ell, v, q)\,(\delta\ell) = DJ_n(\ell)\,(\delta\ell)\,. \tag{50}$$

The adjoint method chooses the test function (v, q) cleverly to avoid the direct computation of the solution map as we shall see shortly. First, we note that

$$\partial_\ell L_n(\ell, v, q)\,(\delta\ell)$$
$$= \left\langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} + a(\delta\ell, u_n(\ell), v)$$
$$+ a(\ell, Du_n(\ell)(\delta\ell), v) + \varepsilon_n \langle Du_n(\ell)(\delta\ell), v\rangle_V + b(v, Dp_n(\ell)(\delta\ell))$$
$$+ b(Du_n(\ell)(\delta\ell), q) - \varepsilon_n \langle Dp_n(\ell)(\delta\ell), q\rangle_Q. \tag{51}$$

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

$$a(\ell, w_n(\ell), v) + \varepsilon_n \langle w_n(\ell), v\rangle_Q + b(v, t_n(\ell)) = \left\langle \bar{z}_{\delta_n} - u_n(\ell), v\right\rangle_{\mathbb{V}}, \text{ for every } v \in V, \tag{52a}$$

$$b(w_n(\ell), q) - \varepsilon_n \langle p_n(\ell), q\rangle_Q = \left\langle \hat{z}_{\delta_n} - p_n(\ell), q\right\rangle_{\mathbb{Q}}, \quad \text{for every } q \in Q, \tag{52b}$$

where $(u_n(\ell), p_n(\ell))$ solves (49a) for the given ℓ and $(\bar{z}_{\delta_n}, \hat{z}_{\delta_n})$ is the given data.

We set $(v, q) = (w_n(\ell), t_n(\ell))$ in (51) and after a simplification obtain

$$
\begin{aligned}
\partial_\ell L_n(\ell, w_n(\ell), t_n(\ell))\,(\delta\ell) &= \left\langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} \\
&+ \left\langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} \\
&\quad + a(\delta\ell, u_n(\ell), w_n(\ell)) + a(\ell, Du_n(\ell)(\delta\ell), w_n(\ell)) + \varepsilon_n \langle Du_n(\ell)(\delta\ell), w_n(\ell)\rangle_V \\
&\quad + b(w_n(\ell), Dp_n(\ell)(\delta\ell)) + b(Du_n(\ell)(\delta\ell), t_n(\ell)) - \varepsilon_n \langle Dp_n(\ell)(\delta\ell), t_n(\ell)\rangle_Q \\
&= \left\langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} \\
&\quad + a(\delta\ell, u_n(\ell), w_n(\ell)) + a(\ell, w_n(\ell), Du_n(\ell)(\delta\ell)) + \varepsilon_n \langle w_n(\ell), Du_n(\ell)(\delta\ell)\rangle_V \\
&\quad + b(w_n(\ell), Dp_n(\ell)(\delta\ell)) + b(Du_n(\ell)(\delta\ell), t_n(\ell)) - \varepsilon_n \langle t_n(\ell), Dp_n(\ell)(\delta\ell)\rangle_Q \\
&= \left\langle Du_n(\ell)(\delta\ell), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} \\
&\quad + a(\delta\ell, u_n(\ell), w_n(\ell)) + \left\langle \bar{z}_{\delta_n} - u_n(\ell), Du_n(\ell)(\delta\ell)\right\rangle_{\mathbb{V}} \\
&\quad + \left\langle \hat{z}_{\delta_n} - p_n(\ell), Dp_n(\ell)(\delta\ell)\right\rangle_{\mathbb{Q}} \\
&= a(\delta\ell, u_n(\ell), w_n(\ell)),
\end{aligned}
$$

which at once gives the formula for the first-order derivative of J_n:

$$
DJ_n(\ell)\,(\delta\ell) = a(\delta\ell, u_n(\ell), w_n(\ell)). \tag{53}
$$

In summary, we deduce the following scheme to compute $DJ_n(\ell)(\delta\ell)$:

(1) Compute $(u_n(\ell), p_n(\ell))$ by solving the regularized saddle point problem (49a).
(2) Compute $(w_n(\ell), t_n(\ell))$ by solving the regularized adjoint problem (52a).
(3) Compute $DJ_n(\ell)(\delta\ell)$ by using (53).

3.2. *Evaluation of the second-order derivatives*

We now provide a second-order adjoint method for the evaluation of the second-order derivative of the OLS objective. The second-order adjoint approach yields a formula for the second-order derivative that does not involve the second-order derivative of the regularized parameter-to-solution map. The key idea is to compute the derivative directly via its variational characterization and avoid the computation of the second-order derivative by the adjoint philosophy.

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

$$
\begin{aligned}
& a(\ell, Du_n(\ell)\delta\ell, v) + \varepsilon_n \langle Du_n(\ell)\delta\ell, v\rangle_V + b(v, Dp_n(\ell)\delta\ell) && \text{(20a)} \\
& \quad = -a(\delta\ell, u_n(\ell), v), \quad \text{for all } v \in V, && \text{(54a)} \\
& b(Du_n(\ell)\delta\ell, q) - \varepsilon_n \langle Dp_n(\ell)\delta\ell, q\rangle_Q = 0, \quad \text{for all } q \in Q. && \text{(54b)}
\end{aligned}
$$

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

$$
\begin{aligned}
\mathcal{L}_n(\ell, v, q) &:= DJ_n(\ell)(\delta\ell_2) + a(\ell_2, Du_n(\ell)\delta\ell_2, v) + \varepsilon_n \langle Du_n(\ell)\delta\ell_2, v\rangle_V \qquad (55)\\
&\quad + b(v, Dp_n(\ell)\delta\ell_2) + a(\delta\ell_2, u_n(\ell), v)\\
&\quad + b(Du_n(\ell)\delta\ell_2, q) - \varepsilon_n \langle Dp_n(\ell)\delta\ell_2, q\rangle_Q\\
&= \left\langle Du_n(\ell)(\delta\ell_2), u_n(\ell) - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell_2), p_n(\ell) - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}}\\
&\quad + a(\ell, Du_n(\ell)\delta\ell_2, v) + \varepsilon_n \langle Du_n(\ell)\delta\ell_2, v\rangle_V + b(v, Dp_n(\ell)\delta\ell_2)\\
&\quad + a(\delta\ell_2, u_n(\ell), v) + b(Du_n(\ell)\delta\ell_2, q) - \varepsilon_n \langle Dp_n(\ell)\delta\ell_2, q\rangle_Q. \qquad (56)
\end{aligned}
$$

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

$$
\partial_\ell \mathcal{L}_n(\ell, v, q)(\delta\ell_1) = D^2 J_n(\ell)(\delta\ell_1, \delta\ell_2). \qquad (57)
$$

By computing the right-hand side of the above identity and using (56), we obtain

$$
\begin{aligned}
\partial_\ell \mathcal{L}_n(\ell, v, q)(\delta\ell_1) &= \left\langle D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), u_n - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}}\\
&+ \langle Du_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\rangle_{\mathbb{V}}\\
&+ \left\langle D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2), p_n - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} + \left\langle Dp_n(\ell)(\delta\ell_2), Dp_n(\ell)(\delta\ell_1)\right\rangle_{\mathbb{Q}}\\
&+ a(\delta\ell_2, Du_n(\ell)(\delta\ell_1), v)\\
&+ a(\delta\ell_1, Du_n(\ell)(\delta\ell_2), v) + a(\ell, D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), v)\\
&+ \varepsilon_n \langle D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), v\rangle_V\\
&+ b(v, D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2)) + b(D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), q)\\
&- \varepsilon_n \langle D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2), q\rangle_{\mathbb{Q}}.
\end{aligned}
$$

By setting $(v,q) = (w_n(\ell), t_n(\ell))$; the solution of (52a), we obtain

$$
\begin{aligned}
\partial_\ell \mathcal{L}_n(\ell, v, q)(\delta\ell_1) &= \left\langle D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), u_n - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}}\\
&+ \langle Du_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\rangle_{\mathbb{V}}\\
&\quad + \left\langle D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2), p_n - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} + \left\langle Dp_n(\ell)(\delta\ell_2), Dp_n(\ell)(\delta\ell_1)\right\rangle_{\mathbb{Q}}\\
&+ a(\delta\ell_2, Du_n(\ell)(\delta\ell_1), w_n(\ell))\\
&\quad + a(\delta\ell_1, Du_n(\ell)(\delta\ell_2), w_n(\ell)) + a(\ell, D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), w_n(\ell))\\
&+ \varepsilon_n \langle D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), w_n(\ell)\rangle_{\mathbb{V}}
\end{aligned}
$$

$$
\begin{aligned}
&+ b(w_n(\ell), D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2)) + b(D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), t_n(\ell)) \\
&- \varepsilon_n \langle D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2), t_n(\ell)\rangle_Q \\
&= \left\langle D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2), u_n - \bar{z}_{\delta_n}\right\rangle_{\mathbb{V}} + \langle Du_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\rangle_{\mathbb{V}} \\
&+ \left\langle D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2), p_n - \hat{z}_{\delta_n}\right\rangle_{\mathbb{Q}} \\
&\quad + \left\langle Dp_n(\ell)(\delta\ell_2), Dp_n(\ell)(\delta\ell_1)\right\rangle_{\mathbb{Q}} \\
&+ a(\delta\ell_1, Du_n(\ell)(\delta\ell_2), w_n) + a(\delta\ell_2, D\bar{u}(\ell)(\delta\ell_1), w_n) \\
&\quad + \left\langle \bar{z}_{\delta_n} - u_n, D^2 u_n(\ell)(\delta\ell_1, \delta\ell_2)\right\rangle + \left\langle \hat{z}_{\delta_n} - p_n, D^2 p_n(\ell)(\delta\ell_1, \delta\ell_2)\right\rangle \\
&= \langle Du_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\right\rangle_{\mathbb{Q}} \\
&+ a(\delta\ell_1, Du_n(\ell)(\delta\ell_2), w_n) \\
&\quad + a(\delta\ell_2, Du_n(\ell)(\delta\ell_1), w_n).
\end{aligned}
$$

Therefore, using (57) 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:

$$
\begin{aligned}
&D^2 J_n(\ell)(\delta\ell_1, \delta\ell_2) \\
&\quad = \langle Du_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\rangle_{\mathbb{V}} + \left\langle Dp_n(\ell)(\delta\ell_2), Du_n(\ell)(\delta\ell_1)\right\rangle_{\mathbb{Q}} \\
&\qquad + a(\delta\ell_1, Du_n(\ell)(\delta\ell_2), w_n) + a(\delta\ell_2, Du_n(\ell)(\delta\ell_1), w_n).
\end{aligned}
$$

In particular, we have

$$
D^2 J_n(\ell)(\delta\ell, \delta\ell) = \langle \delta u_n, \delta u_n\rangle_{\mathbb{V}} + \left\langle \delta p_n, \delta p_n\right\rangle_{\mathbb{Q}} + 2a(\delta\ell, \delta u_n, w_n). \tag{58}
$$

Summarizing, we obtain the following scheme to compute $D^2 J_n(\ell)(\delta\ell, \delta\ell)$:

(1) Compute $(u_n(\ell), p_n(\ell))$ by solving the regularized saddle point problem (49a).
(2) Compute $(\delta u_n, \delta p_n)$ by solving the regularized saddle point problem (20a).
(3) Compute $(w_n(\ell), t_n(\ell))$ by by solving the regularized adjoint problem (52a).
(4) Compute $D^2 J_n(\ell)(\delta\ell, \delta\ell)$ by (58).

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, employing the first-order derivative formula of the OLS obtained from the first-order adjoint approach results in an entirely different second-order adjoint approach.

4. An application to the elasticity imaging inverse problem

4.1. An analytical example

The following system describes the response of an isotropic elastic object to known body forces and boundary traction is the mathematical basis for the elasticity imaging inverse problem:

$$-\nabla \cdot \sigma = f \text{ in } \Omega, \tag{59a}$$

$$\sigma = 2\mu\epsilon(u) + \lambda \operatorname{div} u\, I, \tag{59b}$$

$$u = g \text{ on } \Gamma_1, \tag{59c}$$

$$\sigma n = h \text{ on } \Gamma_2. \tag{59d}$$

Here the domain Ω as a subset of $\mathbb{R}^2$ and $\partial\Omega = \Gamma_1 \cup \Gamma_2$ 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 $\epsilon(u) = \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}})$ is the linearized strain tensor. The resulting stress tensor σ in the stress-strain law (59b) 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 which quantify the elastic properties of the object.

In the numerical experiments, we will focus on the linear incompressible elasticity model. We recall that if the Poisson's ratio $\nu \approx 0.5$, then due to the relationship $\lambda := 2\nu\mu/(1-2\nu)$, λ is large, and the elastic object is termed nearly incompressible. On the other hand, if $\nu \to \frac{1}{2}$, the elastic object is said to fully incompressible. For incompressible materials, the relationship (59a) is not valid, and an alternative formulation is derived involving the incompressibility constraint.

By setting $Q = L^2(\Omega)$, and $\widehat{V} = \{v = (v_1, v_2) \in H^1(\Omega) \times H^1(\Omega) : \bar{v} = 0$ on $\Gamma_1\}$, the variational formulation of (59a) in the incompressible case reads: Find $(u, p) \in V \times Q$ such that

$$\int_\Omega 2\mu\epsilon(u) \cdot \epsilon(v) + \int_\Omega p(\operatorname{div} v) = \int_\Omega fv + \int_{\Gamma_2} vh, \quad \text{for every } \bar{v} \in \widehat{V}, \tag{60a}$$

$$\int_\Omega (\operatorname{div} u)q = 0, \quad \text{for every } q \in Q, \tag{60b}$$

which corresponds to the saddle point problem (3a) by taking

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

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

Let $\Omega = (0,1) \times (0,1)$ be the domain. The sought parameter is $\mu(x,y) = 1 + x^2y + yx$, and the chosen load function is

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

On the bottom and right boundaries (Γ_1), we use Dirichlet boundary conditions. On the bottom boundary, $y = 0$, we have $g(x,y) = (0, x(1-x))$, and on the right boundary, $x = 1$, we have $g(x,y) = (y(1-y), 0)$. On the top and the left boundaries we impose Neumann boundary conditions. On the top boundary, $y = 1$, we have $h(x,y) = ((-(x^2y + xy + 1)(2x + 2y - 2), 0))$, and on the left boundary, $x = 0$, we have $h(x,y) = (0, (x^2y + xy + 1)(2x + 2y - 2))$. This displacement vector in this case is $\bar{u}(x,y) = (y(1-y), x(1-x))$. For simplicity, we consider no data contamination.

For the finite element discretization, we use a simple regular triangulation $\mathcal{T}^h$ (h is the diameter). For the discretization spaces, we consider $\mathbb{P}_1$ elements for all variables, that is, the pressure term p, the parameter μ, and both components of the displacement $u = (u_1, u_2)$.

For discretization, we use the finite element library FreeFem++ [42]. We solve the optimization problem by using the IPOPT optimization library integrated with FreeFem++. We recall that IPOPT is a software library for large scale nonlinear constrained optimization, which implements a primal-dual interior-point method (see [43]). We approximate the Hessian by a BFGS update quasi-Newton method. We recall that IPOPT permits box constraints, and as the lower bound, we set $l_b(x) = 1$ and as the upper $u_b(x) = 3.5$. In all numerical experiments, we take $H = H^1(\Omega)$ as the regularization space.

Numerical results are summarized in Tables 1–3. Table 3 shows the stability of discretization error and Table 1 shows the effect of the regularization parameters κ and ε in the error for a fixed value $\frac{\sqrt{2}}{20}$ of h. Table 2 shows the effect of κ and h for a fixed value 1e−10 of ε. The identification error is measured 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. Given the collected data, the most stable option appears to be $\kappa \in \{1\text{e}{-04}, 1\text{e}{-05}\}$, which gave excellent reconstructions of the parameter for all chosen values of ε. For smaller values of κ (1e−06, 1e−07), we observe a decline in the quality of the reconstructions (see also Figure 1).

Remark 4.1: For a manageable dimension for the optimization problem, instead of the compatible $\mathbb{P}_2\mathbb{P}_1$ elements, we used $\mathbb{P}_1\mathbb{P}_1$ elements. However, the identification is still of satisfactory quality.

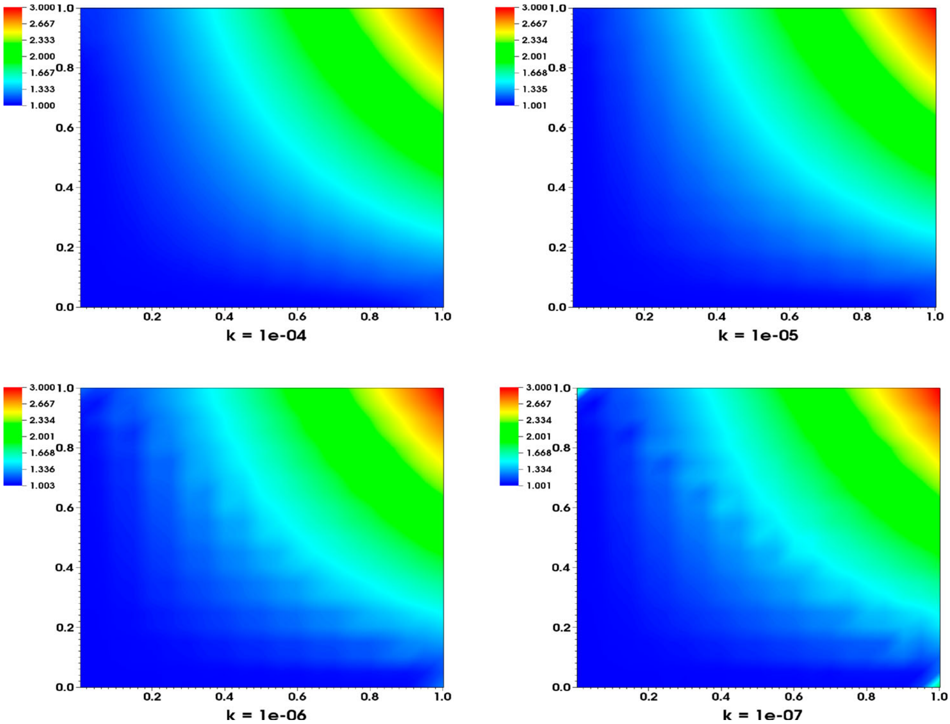


Figure 1. Identified parameter μ for different choices of regularization parameter κ.

Table 1. Error $\|\mu^h - \mu\|_{L^2(\Omega)}/\|\mu\|_{L^2(\Omega)}$ for $h = \frac{\sqrt{2}}{20}$, and various values of κ and ε.

$\varepsilon\backslash\kappa$	κ = 1e−04	κ = 1e−05	κ = 1e−06	κ = 1e−07
1e−05	6.083e−03	2.615e−03	6.297e−03	1.436e−02
1e−06	6.054e−03	2.630e−03	6.784e−03	1.567e−02
1e−07	6.147e−03	3.104e−03	7.265e−03	1.209e−02
1e−10	6.092e−03	3.646e−03	8.576e−03	1.448e−02

Table 2. Error $\|\mu^h - \mu\|_{L^2(\Omega)}/\|\mu\|_{L^2(\Omega)}$ for ε = 1e−10, and various values of h and κ.

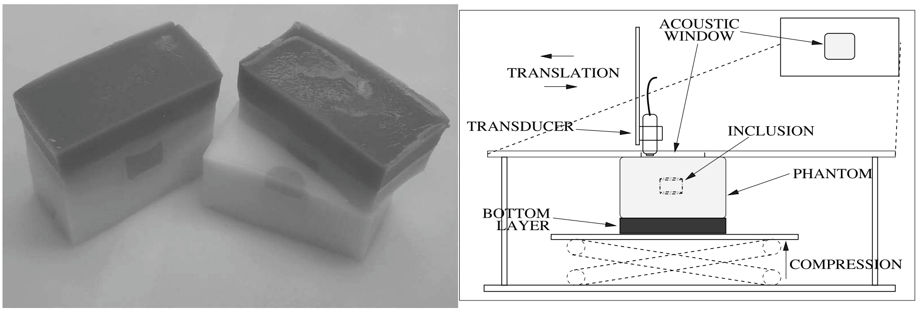
$h\backslash\kappa$	κ = 1e−04	κ = 1e−05	κ = 1e−06	κ = 1e−07
$\sqrt{2}/10$	8.593e−03	6.865e−03	8.854e−03	2.640e−02
$\sqrt{2}/12$	7.040e−03	5.374e−03	9.103e−03	2.260e−02
$\sqrt{2}/14$	6.344e−03	4.209e−03	6.668e−03	1.616e−02
$\sqrt{2}/16$	6.065e−03	3.500e−03	5.507e−03	1.728e−02
$\sqrt{2}/18$	6.037e−03	3.210e−03	8.758e−03	1.536e−02
$\sqrt{2}/20$	6.092e−03	3.646e−03	8.576e−03	1.448e−02

4.2. 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 was created using gelatin, with silica added for acoustic scatter, to mimic elastic properties of soft tissue (see Figure 2(a), which is taken from [44]). A complete description of

Table 3. Regularized discretization error for $\varepsilon = 1e{-}10$.

h	$\frac{\Vert u_1^{h,\varepsilon} - \bar{u}_1 \Vert_{L^2(\Omega)}}{\Vert \bar{u}_1 \Vert_{L^2(\Omega)}}$	$\frac{\Vert \bar{u}_2^{h,\varepsilon} - \bar{u}_1 \Vert_{L^2(\Omega)}}{\Vert \bar{u}_2 \Vert_{L^2(\Omega)}}$	$\Vert p^{h,\varepsilon} \Vert_{L^2(\Omega)}$
$\sqrt{2}/10$	3.871e−04	7.560e−04	3.795e−03
$\sqrt{2}/12$	2.704e−04	5.304e−04	2.597e−03
$\sqrt{2}/14$	1.995e−04	3.924e−04	1.888e−03
$\sqrt{2}/16$	1.532e−04	3.020e−04	1.434e−03
$\sqrt{2}/18$	1.214e−04	2.396e−04	1.127e−03
$\sqrt{2}/20$	9.851e−05	1.947e−04	9.088e−04

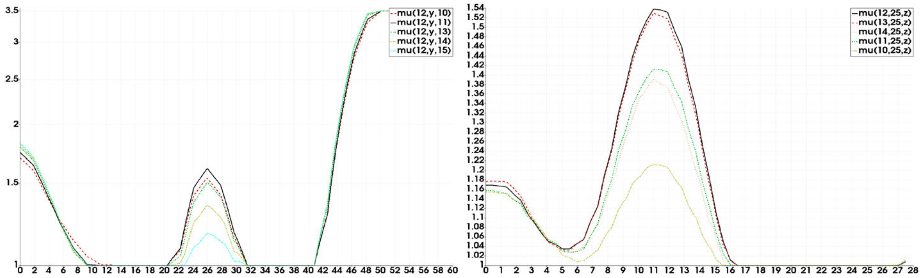


(a) Phantom. (b) Transducer.

Figure 2. (a) Phantom and (b) Transducer.

the phantom construction and experimental imaging setup is available in [44]. The whole phantom was cuboid in shape (60 × 60 × 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). It led to an approximate inclusion to background contrast of 1.89 ± 0.11 as measured by an independent mechanical test [44]. A bottom layer (approximately 10 mm of additional height) was also added to the phantom. Ultrasound (US) image sequences were collected using an Analogic AN2300 (Analogic Corp., 8 Centennial Drive, Peabody, MA 01960) with a 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 2(b)). 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 mm × 55.62 mm × 27.44 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 [44].

We use a discretization scheme described in the previous (analytical) example. We also take into account the practical aspects given in [44], see also [45]. In the simulation, we used a data projected on a mesh of size 30 × 30 × 30 (full data



(a) The central zy-slice (b) The central zx-slice

Figure 3. Five lines of parameter reconstruction for two central slices. (a) The central *zy*-slice and (b) The central *zx*-slice.

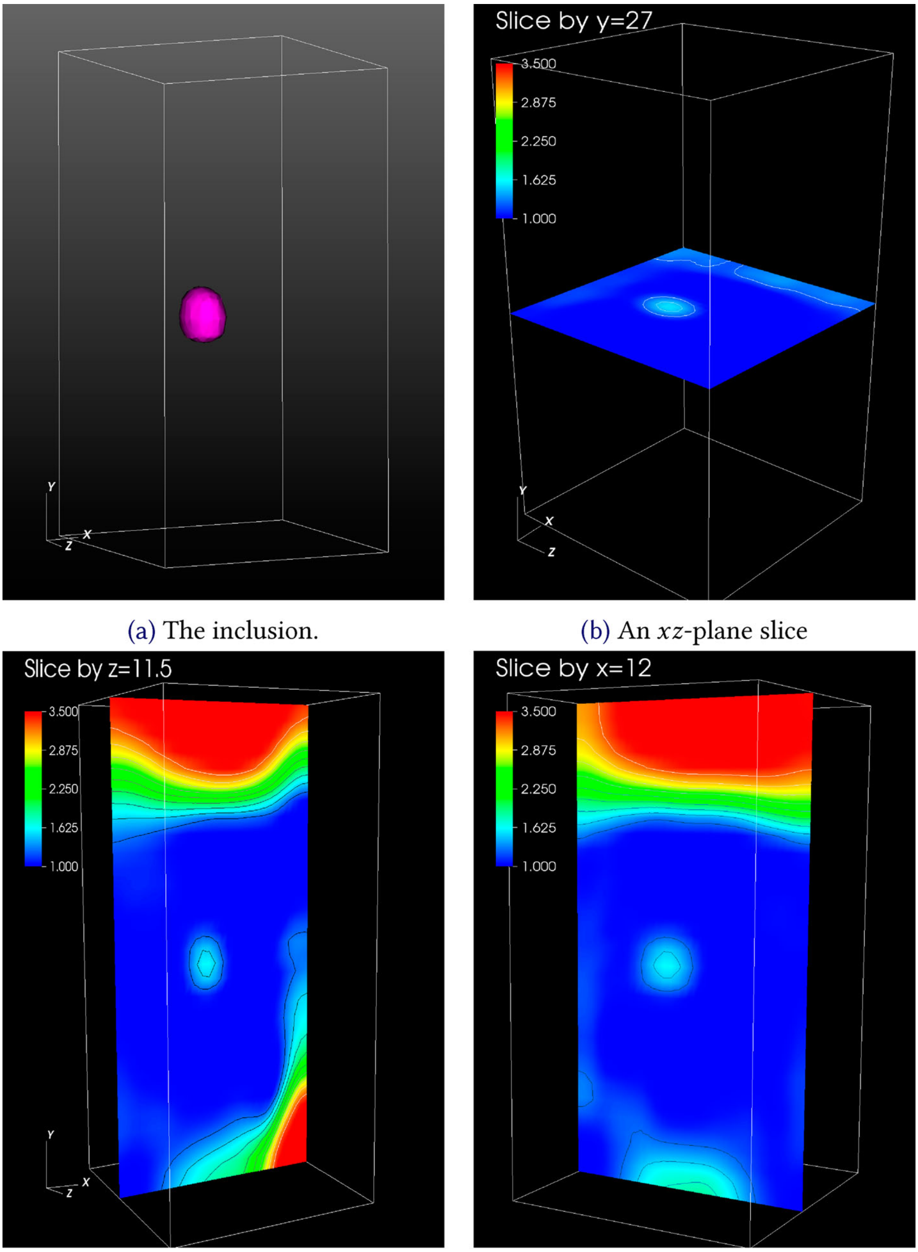
available corresponds to a mesh of size $41 \times 61 \times 41$). Another important aspect, when dealing with phantom data, is the choice of boundary conditions. Following [44,45], we considered Dirichlet conditions in the top and bottom boundaries. We set the vertical direction/component of the remaining boundary conditions to be Dirichlet and allow lateral components to be traction free. We solve the problem with IPOPT, and take $l_b(x) = 1$ and $u_b(x) = 3.5$ as the lower and the upper bound. Since we are dealing with a discontinuous parameter, we have chosen the standard TV regularizer, see [44]).

The 3D simulation results are shown in Figures 3 and 4. In Figure 3, we see the parameter μ values along five lines on two central slices (by planes $x = 12$ and $y = 25$). Figure 4 highlight the inclusion from different perspectives, and plane slices in Figure 4 are coloured by the values of the identified parameter μ. The developed scheme isolates the inclusion very well, even from a data that is projected into a coarser mesh. In particular, one can recognize the cylindrical shape of the inclusion in Figure 4(a). Certain artifacts that are close to the boundaries and a stand-off layer correspond to the stiffer bottom part. The background contrast is approximately 1, corresponding to the lower bound, whereas the inclusion contrast is between 1.3 and 1.5.

5. Identifying a constant coefficient in saddle point problem (15)

We shall now revisit Example 1.3 to identify $\mu(x) = 1$ in the saddle point problem (15a) by taking $a(\mu, u, v) = \int_\Omega \mu \nabla u \cdot \nabla v \, \mathrm{d}x$. Following [31, Section 5], we take $\Omega = (0,1) \times (0,1)$, while the curve Γ_0 is defined by $\Gamma_0 = \{(x, z(x)) : x \in (0,1)\}$, where $z(x) = -x^3 + 1.5x^2 + 0.25$. The subdomains below and above are defined by Ω^- and Ω^+, respectively. We denote the restriction of u on Ω^+ by u^+ and on Ω^- by u^-. We have

$$p = \frac{\partial u^-}{\partial u} - \frac{\partial u^+}{\partial u} \in H^{1/2}(\Gamma_0).$$



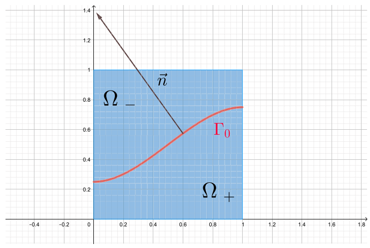
(a) The inclusion. (b) An xz-plane slice (c) An xy-plane slice (d) An yz-plane slice

Figure 4. A 3D visualization of the inclusion. (a) The inclusion. (b) An *xz*-plane slice. (c) An *xy*-plane slice and An *yz*-plane slice.

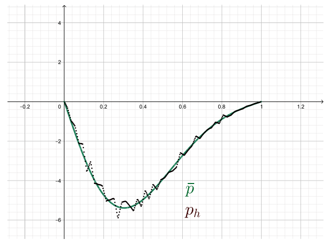
As in [31, Section 5], we take

$$u(x, y) = \sin(\pi x) \sin \frac{\pi(y-1)(y-z(x))}{z(x)},$$

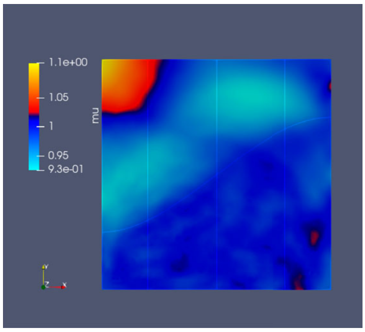
and consider the problem with $f = -\Delta u$ on Ω^+ and $f = 0$ on Ω^-.



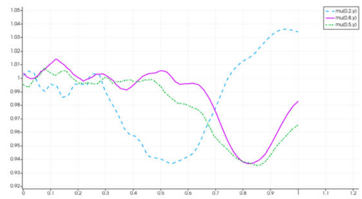
(a) Geometrical description of the domain.



(b) Comparison between analytical $\bar{p}$ and numerical p_h.



(c) Identification. ($h = \sqrt{2}/50$, κ =1e-05, ε =1e-04).



(d) Three lines of vertical parameter reconstruction.

Figure 5. Numerical Results for Example 1.3. (a) Geometrical description of the domain. (b) Comparison between analytical $\bar{p}$ and numerical p_h. (c) Identification. ($h = \sqrt{2}/50$, κ = 1e−05, ε = 1e−04) and (d) Three lines of vertical parameter reconstruction.

Table 4. Error behaviour.

h	$\|\mu - \mu_h\|_{L^2(\Omega)}/\|\mu\|_{L^2(\Omega)}$
$\sqrt{2}/10$	4.019e−02
$\sqrt{2}/20$	5.006e−02
$\sqrt{2}/30$	3.949e−02
$\sqrt{2}/40$	5.405e−02
$\sqrt{2}/50$	3.768e−02

Following [31], we take ε = 1e−04 in our numerical computations. We use the L^2-regularization, that is, $c(p,q) = \int_0^1 pq\,\mathrm{d}x$, for every $p, q \in L^2(0,1)$. This is the simplest of the four choices of the regularization used in [31]. Based on the numerical experimentation, we choose the regularization parameter for the OLS to be κ =1-e05, which provides the best performance for the considered set of discretization parameters. We give the numerical results in Table 4. Numerical approximation of $\bar{p}$ in Figure 5, which corresponds to the optimal μ obtained by solving the inverse problem, is quite comparable with the corresponding solution in [31, Figure 5.2]. The preliminary computations seem to provide a reasonable reconstruction; it is in the range of 3–5% in the lower part Ω^+, which is slightly better than in the upper part Ω, where the primary error source is in the right top corner. We refer the reader to [31] for any missing details and additional information Example 1.3.

6. Concluding remarks

We studied the inverse problem of parameter identification in general saddle point problems by using the OLS objective. We advocated for the usefulness of incorporating the regularization saddle point problems into the OLS formulation. It would be of natural interest to develop error estimates involving the discretization parameter and the regularization parameters. Such error estimates have been developed for the direct problem. However, there is no parallel study for inverse problems. Detailed numerical studies also need to be carried out for the cases when the Inf-Sup condition is violated. In recent years, a great deal of attention has been given to identification in the stochastic PDEs (see [1]) or to Bayesian inverse problems (see [46]), and it is of interest to study the elasticity imaging inverse problem when the sought elasticity parameters are random variables.

Acknowledgments

We are grateful to the reviewers for the careful reading and helpful suggestions that brought substantial improvements to the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research of Baasansuren Jadamba and Akhtar Khan is supported by the National Science Foundation under Award No. 1720067. Miguel Sama is partially supported by the Ministerio de Ciencia, Innovacion y Universidades (MCIU), Agencia Estatal de Investigacion (AEI) (Spain), and Fondo Europeo de Desarrollo Regional (FEDER) under project PGC2018-096899-B-I00 (MCIU/AEI/FEDER, UE) and grant number 2020-MAT11 from ETSI Industriales, UNED.

References

[1] Aboulaich R, Fikal N, El Guarmah E, et al. Stochastic finite element method for torso conductivity uncertainties quantification in electrocardiography inverse problem. Math Model Nat Phenom. 2016;11:1–19.
[2] Ammari H, Garapon P, Jouve F. Separation of scales in elasticity imaging: a numerical study. J Comput Math. 2010;28:354–370.
[3] Beigl A, Scherzer O, Sogn J, et al. Preconditioning inverse problems for hyperbolic equations with applications to photoacoustic tomography; 2019.
[4] Boehm C, Ulbrich M. A semismooth Newton-CG method for constrained parameter identification in seismic tomography. SIAM J Sci Comput. 2015;37:S334–S364.
[5] Boiger R, Kaltenbacher B. An online parameter identification method for time dependent partial differential equations. Inverse Probl. 2016;32:045006, 28.
[6] Cahill N, Jadamba B, Khan AA, et al. A first-order adjoint and a second-order hybrid method for an energy output least squares elastography inverse problem of identifying tumor location. Boundary Value Probl. 2013;263:1–14.

[7] Clason C. L^{∞} fitting for inverse problems with uniform noise. Inverse Probl. 2012;28:104007, 18.
[8] Crossen E, Gockenbach MS, Jadamba B, et al. An equation error approach for the elasticity imaging inverse problem for predicting tumor location. Comput Math Appl. 2014;67:122–135.
[9] Doyley MM, Jadamba B, Khan AA, et al. 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. 2014;35:984–1017.
[10] Evstigneev RO, Medvedik MY, Smirnov YG. Inverse problem of determining parameters of inhomogeneity of a body from acoustic field measurements. Comput Math Math Phys. 2016;56:483–490.
[11] Gholami A, Mang A, Biros G. An inverse problem formulation for parameter estimation of a reaction-diffusion model of low grade gliomas. J Math Biol. 2016;72:409–433.
[12] Gockenbach MS, Jadamba B, Khan AA, et al. Proximal methods for the elastography inverse problem of tumor identification using an equation error approach. In: Advances in variational and hemivariational inequalities. Cham: Springer; 2015. p. 173–197.
[13] Guchhait S, Banerjee B. Constitutive error based material parameter estimation procedure for hyperelastic material. Comput Methods Appl Mech Engrg. 2015;297:455–475.
[14] Hager W, Ngo C, Yashtini M, et al. An alternating direction approximate Newton algorithm for ill-conditioned inverse problems with application to parallel MRI. J Oper Res Soc China. 2015;3:139–162.
[15] Kindermann S, Mutimbu LD, Resmerita E. A numerical study of heuristic parameter choice rules for total variation regularization. J Inverse Ill-Posed Probl. 2014;22:63–94.
[16] Kuchment P, Steinhauer D. Stabilizing inverse problems by internal data. II: non-local internal data and generic linearized uniqueness. Anal Math Phys. 2015;5:391–425.
[17] Liu T. A wavelet multiscale-homotopy method for the parameter identification problem of partial differential equations. Comput Math Appl. 2016;71:1519–1523.
[18] Mang A, Gholami A, Davatzikos C, et al. PDE-constrained optimization in medical image analysis. Optim Eng. 2018;19:765–812.
[19] Manservisi S, Gunzburger M. A variational inequality formulation of an inverse elasticity problem. Appl Numer Math. 2000;34:99–126.
[20] Neubauer A, Hein T, Hofmann B, et al. Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces. Appl Anal. 2010;89:1729–1743.
[21] Scheufele K, Mang A, Gholami A, et al. Coupling brain-tumor biophysical models and diffeomorphic image registration. Comput Methods Appl Mech Engrg. 2019;347:533–567.
[22] Subramanian S, Scheufele K, Mehl M, et al. Where did the tumor start? an inverse solver with sparse localization for tumor growth models. Inverse Probl. 2020. (in press).
[23] Girault V, Raviart PA. Finite element methods for Navier-Stokes equations. Berlin: Springer-Verlag; 1986. p. x+374.
[24] Braess D. Finite elements. Cambridge: Cambridge University Press; 2007. p. xviii+365.
[25] Bochev PB, Lehoucq RB. Regularization and stabilization of discrete saddle-point variational problems. Electron Trans Numer Anal. 2006;22:97–113.
[26] Brezzi F, Fortin M. Mixed and hybrid finite element methods. New York: Springer-Verlag; 1991. p. x+350.
[27] Bercovier M. Perturbation of mixed variational problems. Application to mixed finite element methods. RAIRO Anal Numér. 1978;12:211–236.

[28] Barbosa HJC, Hughes TJR. Circumventing the Babuška-Brezzi condition in mixed finite element approximations of elliptic variational inequalities. Comput Methods Appl Mech Engrg. 1992;97:193–210.

[29] Bochev PB, Gunzburger MD, Lehoucq RB. On stabilized finite element methods for the Stokes problem in the small time step limit. Internat J Numer Methods Fluids. 2007;53:573–597.

[30] Pestana J, Wathen AJ. Natural preconditioning and iterative methods for saddle point systems. SIAM Rev. 2015;57:71–91.

[31] Ito K, Kunisch K, Peichl G. On the regularization and approximation of saddle point problems without inf-sup condition. Comput Appl Math. 2002;21:245–274.

[32] Glowinski R, Li CH, Lions JL. A numerical approach to the exact boundary controllability of the wave equation. I. Dirichlet controls: description of the numerical methods. Japan J Appl Math. 1990;7:1–76.

[33] Kunisch K, Peichl G. Shape optimization for mixed boundary value problems based on an embedding domain method. Dynam Contin Discrete Impuls Systems. 1998;4:439–478.

[34] Jadamba B, Khan AA, Rus G, et al. 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. 2014;74:1486–1510.

[35] Nashed MZ, Scherzer O. Least squares and bounded variation regularization with nondifferentiable functionals. Numer Funct Anal Optim. 1998;19:873–901.

[36] Cioaca A, Alexe M, Sandu A. Second-order adjoints for solving PDE-constrained optimization problems. Optim Methods Softw. 2012;27:625–653.

[37] Knopoff DA, Fernández DR, Torres GA, et al. Adjoint method for a tumor growth PDE-constrained optimization problem. Comput Math Appl. 2013;66:1104–1119.

[38] Kourounis D, Durlofsky LJ, Jansen JD, et al. Adjoint formulation and constraint handling for gradient-based optimization of compositional reservoir flow. Comput Geosci. 2014;18:117–137.

[39] Namdar Zanganeh M, Kraaijevanger JFBM, Buurman HW, et al. Challenges in adjoint-based optimization of a foam EOR process. Comput Geosci. 2014;18:563–577.

[40] Oberai AA, Gokhale NH, Feijóo GR. Solution of inverse problems in elasticity imaging using the adjoint method. Inverse Probl. 2003;19:297–313.

[41] Jadamba B, Khan AA, Oberai A, et al. 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;25:1768–1787.

[42] Hecht F. New development in freefem++. J Numer Math. 2012;20:251–266.

[43] Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Program. 2006;106:25–57.

[44] Richards MS, Barbone PE, Oberai AA. Quantitative three-dimensional elasticity imaging from quasi-static deformation: a phantom study. Phys Med Biol. 2009;54:757.

[45] Richards MS. Quantitative three dimensional elasticity imaging [PhD thesis]. Boston University; 2007.

[46] Kaipio J, Somersalo E. Statistical and computational inverse problems. New York: Springer-Verlag; 2005. xvi+339.