J. Nonlinear Var. Anal. 7 (2023), No. 5, pp. 715-726
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.7.2023.5.04

CONVERGENCE RATES FOR NONLINEAR INVERSE PROBLEMS OF PARAMETER IDENTIFICATION USING BREGMAN DISTANCES

DINH NHO HÀO[1], AKHTAR A. KHAN[2,*], SIMEON REICH[3]

[1]*Hanoi Institute of Mathematics, VAST, 18 Hoang Quoc Viet Road, 10307 Hanoi, Vietnam*
[2]*School of Mathematics and Statistics, Rochester Institute of Technology, Rochester, New York, 14623, USA*
[3]*Department of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel*

Abstract. Deriving convergence rates constitutes a crucial and profound field of investigation, carrying significant implications in both theoretical and practical contexts. This study focuses on establishing new convergence rates for nonlinear inverse problems concerning the identification of variable parameters in an abstract variational problem. We employ the energy least squares and output least squares methods to study the inverse problem in an optimization framework. The convergence rates are given in terms of the renowned Bregman distance associated with a convex regularizer. An intriguing aspect of the derived convergence rates is that they do not necessitate any smallness condition, making them applicable to a wide array of practical models.

Keywords. Energy least-squares; Output least-squares; Parameter identification; Regularization; Variational problems.

1. INTRODUCTION

A multitude of models in applied sciences result in partial differential equations (PDEs) that incorporate variable parameters corresponding to distinct physical attributes of the underlying model. In this context, the direct problem involves seeking a solution to the PDE. On the other hand, the inverse problem revolves around determining the variable parameters from a measurement of the solution of the underlying PDE. In recent years, the field of inverse problems has emerged as a dynamic and rapidly expanding branch of applied mathematics. The driving force behind this growth is the increasing number of real-world scenarios that can be effectively modeled and studied using the framework of inverse problems. Theoretical aspects of inverse problems pose significant challenges, necessitating a seamless integration of various branches of mathematics. As a result, researchers and practitioners have been motivated to explore and develop innovative techniques to tackle these complex problems, leading to exciting advancements in both theoretical and practical aspects of inverse problems. This multidisciplinary nature of the field enhances its appeal and relevance to a wide range of applications, including engineering, physics, medicine, and environmental sciences, among others. As the complexities of real-world situations continue to grow, the field

*Corresponding author.
E-mail address: hao@math.ac.vn (D. N. Hào), aaksma@rit.edu (A. A. Khan), sreich@technion.ac.il (S. Reich).
Received July 31, 2023; Accepted August 15, 2023.

of inverse problems remains at the forefront of addressing these challenges and finding solutions to a diverse array of problems. To get a glimpse of some of the recent developments, see [2, 3, 4, 5, 6, 10, 11, 14, 15, 18, 22, 23, 31, 33, 35, 36, 37, 41] and the references therein.

Before elaborating on the primary objective and contributions of the present work, we will provide a concise overview of the prototypic inverse problem involving the identification of a variable parameter in the following elliptic boundary value problem (BVP):

$$-\nabla\cdot(q\nabla u)=f \;\text{ in }\Omega,\; u=0 \;\text{ on }\partial\Omega, \tag{1.1}$$

where Ω is a suitable domain in $\mathbb{R}^2$ or $\mathbb{R}^3$ and $\partial\Omega$ is its boundary. The above BVP models various interesting real-world problems and has been extensively studied. In (1.1), $u=u(x)$ can represent the steady-state temperature at a specific point x in a body. In this context, q denotes the variable thermal conductivity coefficient, and f represents the external heat source. Additionally, the same system (1.1) can also be applied to modelling underground steady-state aquifers. In this case, the parameter q corresponds to the aquifer transmissivity coefficient, u represents the hydraulic head, and f denotes the recharge. An important challenge in the context of the above BVP is the inverse problem, which involves estimating the coefficient q from a measurement z of the solution u. Solving this inverse problem is crucial for understanding and predicting various phenomena in both thermal and aquifer systems, such as heat distribution in materials or groundwater flow.

Numerous approaches to addressing the aforementioned inverse problem have been proposed in the literature. Most of these approaches involve two main strategies: interpreting (1.1) as a hyperbolic partial differential equation (PDE) with respect to the variable q, or formulating an optimization problem where the solution yields an estimate of q. The optimization-based approach to reformulating (1.1) can be categorized into two possibilities: formulating an unconstrained optimization problem or treating it as a constrained optimization problem with the PDE itself serving as the constraint. Among the optimization-based techniques, one of the most widely used methods is the output least-squares (OLS) method that minimizes the discrepancy between the model output and the observed data by employing a suitable norm. To be precise, the OLS approach minimizes the functional

$$q\to\|u(q)-z\|_Z^2, \tag{1.2}$$

where z is the data, which belongs to a Hilbert space Z with norm $\|\cdot\|_Z$, and $u(q)$ solves the variational form of (1.1) given by

$$\int_\Omega q\nabla u\cdot\nabla v=\int_\Omega fv,\;\; \text{for all}\, v\in H_0^1(\Omega). \tag{1.3}$$

While the OLS functional is a prevalent choice for tackling inverse problems, its most significant drawback is its nonconvex nature.

To address the challenges associated with the nonconvexity of the OLS objective, Knowles [32] proposed minimizing a coefficient-dependent norm

$$q\to\int q\nabla(u(q)-z)\cdot\nabla(u(q)-z), \tag{1.4}$$

where z is the data (the measurement of u) and $u(q)$ solves (1.3). Knowles [32] proved that the above functional is convex. It is worth mentioning that Zou [42] independently proposed the idea of minimizing the energy without dwelling on its convexity.

Inverse problems are notorious for being ill-posed, which means that their solutions can be highly sensitive to small changes in the input data. Various regularization methods have been developed to address this issue in order to mitigate the adverse effects of ill-posedness. In the context of the OLS method, the objective is to minimize the following regularized OLS loss function:

$$q \to \|u(q) - z\|^2 + \varepsilon \|q\|^2, \tag{1.5}$$

where z represents the observed data, $u(q)$ is the solution of (1.3), $\|q\|^2$ denotes a suitable quadratic regularizer term, and $\varepsilon > 0$ serves as the regularization parameter.

The overall success of regularization methods hinges on carefully selecting the regularization parameter. A crucial aspect of this selection process is finding a balance, as both overly large and excessively small parameters can lead to suboptimal outcomes. This issue becomes more pronounced, particularly when the data are contaminated with noise. When noise is present, it becomes essential to tune the regularization parameter, considering the level of noise in the data.

The investigation of convergence rates, a crucial and fundamental aspect in the realm of inverse problems, seeks to establish an asymptotic correlation between the noise level and the regularization parameter. This correlation, in turn, enables an effective means of selecting the appropriate regularization parameter by considering the level of data contamination. Investigating convergence rates represents a profound and essential area of research, with far-reaching implications both in theory and practical applications.

Before presenting our main findings, we will provide a concise overview of the key results in this particular field. One of the earlier works on convergence rates is the seminal paper by Engl, Kunisch, and Neubauer [13], where, given Hilbert spaces X and Y, a map $F : D(F) \subset X \to Y$, and $y_0 \in Y$, the focus is on an ill-posed problem seeking $x \in D(F)$ such that

$$F(x) = y_0. \tag{1.6}$$

As commonly done in the study of ill-posed problems, in [13], the operator equation (1.6) was studied as the following OLS-based optimization problem:

$$\min_{x \in D(F)} \|F(x) - \bar{y}\|_Y^2 + \alpha \|x - \hat{x}\|_X^2,$$

where $\alpha > 0$ and $\hat{x}$ is an *a priori* estimate of the unknown solution. Let S be the solution set of (1.6). We recall that $\bar{x} \in S$ is called an $\hat{x}$-minimal norm solution, if $\|\bar{x} - \hat{x}\|_X \leq \|x - \hat{x}\|_X$, for every $x \in S$.

The following theorem is the main result established in [13].

Theorem 1.1. *Let X and Y be Hilbert spaces, let $F : D(F) \subset X \to Y$ be a nonlinear map with a convex domain $D(F)$. Let $y_0 \in Y$ and for $\delta > 0$, let y_δ be such that $\|y_\delta - y_0\|_Y \leq \delta$. Let $\bar{x}$ be an $\hat{x}-$minimal norm solution. Assume that the following conditions hold:*

- $(\mathbb{A}_1)$**:** *The map F is Fréchet differentiable with F' denoting the Fréchet derivative.*
- $(\mathbb{A}_2)$**:** *There exists $L > 0$ such that $\|F'(x_0) - F'(z)\|_Y \leq L\|x_0 - z\|_X$, for every $z \in D(F)$.*
- $(\mathbb{A}_3)$**:** *There exists $w \in Y$ with $\bar{x} - \hat{x} = F'(x_0)^* w$, where $F'(x_0)^*$ is the adjoint of $F'(x_0)$.*
- $(\mathbb{A}_4)$**:** $L\|w\|_Y < 1$.

Let $x_{\delta,\alpha}$ be a solution of the following optimization problem with contaminated data:

$$\min \|F(x) - y_\delta\|_Y^2 + \alpha \|x - \hat{x}\|_X^2.$$

Then, for the choice $\alpha \sim \delta$, we have

$$\|x_{\delta,\alpha} - x_0\|_X = O\left(\sqrt{\delta}\right).$$

In their research, Engl, Kunisch, and Neubauer [13] employed Theorem 1.1 to tackle the inverse problem of estimating a variable coefficient in a two-point boundary value problem, along with Hammerstein integral equations. However, the smallness condition presented as $L\|\omega\|_Y < 1$ in Theorem 1.1 turned out to be excessively stringent and challenging to verify in practical situations. Because of this constraint, Theorem 1.1 received significant attention, leading to subsequent advancements focused on relaxing the smallness condition.

To mention one of the many extensions of Theorem 1.1, we note that Hào and Quyen [24] considered the inverse problem of identifying the parameter q in (1.1) and employed the functional (1.4) for the inverse problem. Hào and Quyen [24] proved convergence rates, completely analogous to those given in Theorem 1.1, but without requiring the smallness condition.

Another noteworthy direction for extending Theorem 1.1 involves the introduction of convergence rates using the Bregman distance, wherein the Tikhonov regularization term $\|\cdot\|_X^2$ was replaced due to its tendency to induce over-smoothing effects that are unsuitable for numerous applications. Subsequent research addressed this concern by employing nonsmooth regularizers. Notable contributions in this direction include the works of Burger and Osher [8], Kaltenbacher and Hofmann [30], Kügler and Sincich [34], Resmerita [39], and Resmerita and Scherzer [40], along with other references cited therein.

To describe the key contribution by Resmerita and Scherzer [40], we first need to recall the celebrated notion of the Bregman distance, proposed by Brègman [7]. Let X be a Banach space and let X^* be the dual of X. We denote the pairing between a Banach space X and its dual X^* by $\langle\cdot,\cdot\rangle_X$. Let $f : X \to (-\infty,\infty]$ be a proper convex functional with domain $D(f) := \{q \in X |\, f(q) < +\infty\}$, and let $\partial f(q)$ be the subdifferential of f at $q \in D(f)$ given by

$$\partial f(q) := \{q^* \in X^* |\, f(p) \geq f(q) + \langle q^*, p-q\rangle_X, \text{ for all } p \in X\}.$$

The set $\partial f(q) \neq \emptyset$, if f is continuous at q. Moreover, $\partial f(q)$ is convex and weak* compact.

The Bregman distance between two elements $p,q \in X$, with respect to f and $q^* \in \partial f(q)$, provided that $\partial f(q) \neq \emptyset$, is defined by

$$D_{f,q^*}(p,q) := f(p) - f(q) - \langle q^*, p-q\rangle_X.$$

The Bregman distance, in general, is not a metric on X. However, $D_{f,q^*}(p,q) \geq 0$ for each $q^* \in \partial f(q)$, and $D_{f,q^*}(p,p) = 0$. For a recent re-examination of various aspects related to Bregman functions and distances, see the paper by Reem, Reich, and De Pierro [38].

Let X and Y be Banach spaces, let $F : D(F) \subset X \to Y$ be a nonlinear map with domain $D(F)$, and let $y_0 \in Y$. Resmerita and Scherzer [40] studied an ill-posed problem seeking $x \in D(F)$ such that (1.6) holds. They posed this ill-posed problem as the following OLS-based optimization problem:

$$\min_{x \in D(F)} \|F(x) - y_\delta\|_Y^2 + \alpha R(x), \tag{1.7}$$

where $\alpha > 0$ is the regularization parameter, $y_\delta \in Y$ is the noisy data such that $\|y_0 - y_\delta\|_Y \leq \delta$, and $R : X \to \mathbb{R}$ is a proper convex functional. They made the following assumptions:

$(\mathbb{B}_1)$: There exists an R-minimizing solution $\bar{x}$ of (1.6) defined by

$$F(\bar{x}) = y_0, \quad \text{and} \quad R(\bar{x}) = \min\{R(x) |\, F(x) = y_0\}.$$

($\mathbb{B}_2$): A solution $x_{\delta,\alpha}$ of (1.7) exists.
($\mathbb{B}_3$): F is Fréchet differentiable around $\bar{x}$ with $F'(\bar{x})$ denoting the Fréchet derivative.
($\mathbb{B}_4$): There is a constant $\gamma > 0$ such that for any $x \in D(F) \cap B_r(\bar{x})$, we have

$$\|F(x) - F(\bar{x}) - F'(\bar{x})(x-\bar{x})\|_Y \leq \gamma D_{R,s^*}(x,\bar{x}), \quad \text{for all } s^* \in \partial R(\bar{x}).$$

($\mathbb{B}_5$): There exists $w \in Y^*$ with $F'(\bar{x})^* w \in \partial R(\bar{x})$, where $F'(\bar{x})^*$ is the adjoint of $F'(\bar{x})$.
($\mathbb{B}_6$): $L\|w\|_{Y^*} < 1$.

Then, for $\alpha \sim \delta$, Resmerita and Scherzer [40] showed that $D_{R,F'(\bar{x})^*w}(x_{\delta,\alpha},\bar{x}) = O(\delta)$.

Hào and Quyên [25] continued their investigation of the inverse problem of identifying the parameter q in (1.1), and by employing the functional (1.4) once more, demonstrated convergence rates for a convex regularizer, which were entirely analogous to those given by Resmerița and Scherzer [40]. However, Hào and Quyên's usage of the energy functional eliminated the requirement of imposing the smallness condition.

Based on the preceding discussion, the use of a convex objective functional has effectively eliminated the necessity of the smallness condition in establishing convergence rates. This achievement, however, is currently limited to identification in the scalar PDEs. Conversely, the convergence rates attained using the smallness condition are applicable within a broader and abstract framework.

Inspired by the aforementioned noticeable gap in the existing literature, our research aims to address the convergence rates of the inverse problem of parameter identification in an abstract variational problem without imposing any smallness conditions. To achieve this objective, we leverage the convex energy least squares (ELS) functional, introduced in [20], in order to extend the functional given in (1.4). Prior works on the ELS functional and its extensions have been explored in [17, 19, 21, 27]. Furthermore, we demonstrate that these convergence rates also apply to the OLS formulation.

2. Optimization Formulations for the Inverse Problems

Let B be a Banach space, and let $A \subset B$ be nonempty, closed, and convex. Let V be a Hilbert space continuously embedded into another Hilbert space Z, and let V^* be the dual of V. Let $T : B \times V \times V \to \mathbb{R}$ be a trilinear form with $T(a,u,v)$ symmetric in u, v. Let $m \in V^*$. Assume that for $\alpha > 0$ and $\beta > 0$, the following continuity and ellipticity conditions hold:

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

$$T(a,u,u) \geq \alpha \|u\|_V^2, \ \text{for all } u \in V, \ a \in A. \tag{2.2}$$

We consider the following variational problem: Given $a \in A$, find $u = u(a) \in V$ such that

$$T(a,u,v) = m(v), \ \text{ for every } v \in V. \tag{2.3}$$

In view of the symmetry, continuity, and ellipticity of T, the Riesz representation theorem ensures that for every $a \in A$, the variational problem (2.3) admits a unique solution $u(a)$.

In this work, we study the inverse problem related to the direct problem (2.3): Given a measurement of u, say $z \in Z$, estimate the coefficient a which together with u makes (2.3) true.

This inverse problem is often posed as the OLS-based optimization problem

$$\min_{a \in A} \widehat{J}(a) := \frac{1}{2}\|u(a) - z\|_Z^2, \tag{2.4}$$

which minimizes the gap between the computed solution $u(a)$ and the data $z \in Z$.

The output least-squares (OLS) functional $\widehat{J}$ is generally non-convex, yielding only local minimizers. This non-convexity also negatively affects stability and numerical aspects.

We also recall the following energy output least-squares (ELS) based formulation

$$\min_{a \in A} J(a) := \frac{1}{2} T(a, u(a) - z, u(a) - z), \tag{2.5}$$

which minimizes the energy of (2.3), where $u(a)$ is the computed solution and $z \in V$ stands for the data. It was shown in [20] that the ELS functional is convex in the set A.

We recall the following results concerning the smoothness of the parameter-to-solution map.

Lemma 2.1. [20] *For each $a \in A$, $u(a)$ satisfies $\|u(a)\|_V \leq \alpha^{-1}\|m\|_{V^*}$. For $a, b \in A$, we have*

$$\|u(a) - u(b)\|_V \leq \min\left\{\frac{\beta}{\alpha}\|u(a)\|_V, \frac{\beta}{\alpha}\|u(b)\|_V, \frac{\beta}{\alpha^2}\|m\|_{V^*}\right\}\|b - a\|_B. \tag{2.6}$$

For each a in the interior of A, u is infinitely differentiable at a. The first-order derivative of $u(a)$ at a in the direction δa, denoted by $Du(a)\delta a$ is the unique solution of the variational equation

$$T(a, Du(a)\delta a, v) = -T(\delta a, u(a), v), \text{ for every } v \in V. \tag{2.7}$$

Moreover, the following bounds hold:

$$\|Du(a)\| \leq \frac{\beta}{\alpha}\|u(a)\|_V \leq \frac{\beta}{\alpha^2}\|m\|_{V^*}. \tag{2.8}$$

We make the following assumptions regarding a non-quadratic regularization framework:

($\mathbb{C}_1$): The Banach space B is continuously embedded in a Banach space L. There is another Banach space $\widehat{B}$ that is compactly embedded in L. The set A is a subset of $B \cap \widehat{B}$, closed and bounded in B and also closed in L.

($\mathbb{C}_2$): For bounded sequences $\{u_k\} \subset V$, and $\{a_k\} \subset B$ with $a_k \to a$ in L, for any fixed $v \in V$, we have

$$T(a_k - a, u_k, v) \to 0. \tag{2.9}$$

($\mathbb{C}_3$): $R : \widehat{B} \to \mathbb{R}$ is a positive, convex, $\|\cdot\|_L$-lower-semicontinuous functional satisfying

$$R(a) \geq \tau_1\|a\|_{\widehat{B}} - \tau_2, \quad \text{for every } a \in A,\ \tau_1 > 0,\ \tau_2 > 0. \tag{2.10}$$

The illustrate the above abstract regularization framework, we recall some functional spaces. Given the domain Ω, for $1 \leq p < \infty$, by $L^p(\Omega)$, we denote the space of pth Lebesgue integrable (equivalence classes of) functions. The space $L^\infty(\Omega)$ consists of measurable functions that are bounded almost everywhere (a.e.) on Ω. We also recall that the Sobolev spaces are given by

$$H^1(\Omega) = \left\{y \in L^2(\Omega),\ \partial_{x_i} y \in L^2(\Omega),\ i = 1, \ldots, n\right\},$$
$$H_0^1(\Omega) = \left\{y \in H^1(\Omega),\ y|_{\partial D} = 0\right\},$$

and $H^{-1}(\Omega) = (H_0^1(\Omega))^*$ is the topological dual of $H_0^1(\Omega)$. For $m \in \mathbb{N}$, higher-order Sobolev spaces $H^m(\Omega)$ consist of $L^2(\Omega)$ functions with all their partial derivatives up to order m residing in $L^2(\Omega)$. In accordance, by $H_0^m(\Omega)$, we represent the functions y in $H^m(\Omega)$ whose boundary traces of the derivatives of up to order less than m are zero.

We recall that the total variation of $f \in L^1(\Omega)$ is given by

$$\mathrm{TV}(f) := \sup\left\{-\int_\Omega f(\nabla \cdot g) : g \in \left(C_0^1(\Omega)\right)^2,\ |g(x)| \leq 1 \text{ for all } x \in \Omega\right\},$$

where $|\cdot|$ denotes the Euclidean norm of a vector. 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 the space of functions of bounded variations $\mathrm{BV}(\Omega)$ is defined by $\mathrm{BV}(\Omega) = \left\{ f \in L^1(\Omega) : \mathrm{TV}(f) < \infty \right\}$. The norm on $\mathrm{BV}(\Omega)$ is $\|f\|_{\mathrm{BV}(\Omega)} = \|f\|_{L^1(\Omega)} + \mathrm{TV}(f)$. The functional $\mathrm{TV}(\cdot)$ is a seminorm on $\mathrm{BV}(\Omega)$ and is termed the BV-seminorm.

In the context of the variational problem (1.3), we set $V = H_0^1(\Omega)$, $B = L^\infty(\Omega)$, $L = L^1(\Omega)$, $\widehat{B} = \mathrm{BV}(\Omega)$, and $R(a) = TV(a)$, and define the set of feasible parameters by

$$A := \{a \in L^\infty(\Omega) |\ 0 < c_1 \leq a(x) \leq c_2, \text{ almost everywhere}\}, \tag{2.11}$$

where c_1 and c_2 are constants. 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 with respect to the $L^1(\Omega)$-norm (see [1, 16]).

We now return to the abstract framework. In the following discussion, we set $X := B \cap \widehat{B}$, and assume that it is a Banach space equipped with the norm $\|\cdot\|_X = \|\cdot\|_B + \|\cdot\|_{\widehat{B}}$. Then, $B^* \subset X^*$ and $\widehat{B}^* \subset X^*$. We denote by $\bar{X}$ and $\widehat{X}$, the space X equipped with the norms $\|\cdot\|_B$ and $\|\cdot\|_{\widehat{B}}$, respectively. Note that for any $\ell \in \widehat{X}^*$, we have

$$\langle \ell, h \rangle_{\widehat{X}} = \langle \ell, h \rangle_X, \quad \text{for all } h \in X. \tag{2.12}$$

The above preparation leads us to considering the following regularized ELS-based optimization problem:

$$\min_{a \in A} J_\kappa(a) = \frac{1}{2} T(a, u(a) - z, u(a) - z) + \kappa R(a), \tag{2.13}$$

where $\kappa > 0$ is a regularization parameter and R is the regularization map satisfying $(\mathbb{C}_3)$.

In the same vein, we also consider the regularized OLS-based optimization problem:

$$\min_{a \in A} \widehat{J}_\kappa(a) = \frac{1}{2} \|u(a) - z\|_Z^2 + \kappa R(a). \tag{2.14}$$

We now summarize existence results and optimality conditions for (2.13) and (2.14).

Theorem 2.1. [29] *The optimization problems* (2.13) *and* (2.14) *have nonempty solution sets, where every minimizer of* (2.13) *is a global minimizer. Moreover, $\bar{a} \in A$ is a minimizer of* (2.13) *if and only if it solves the following variational inequality that seeks $\bar{a} \in A$ such that*

$$T(b - \bar{a}, u(\bar{a}) + z, z - u(\bar{a})) \geq 2\kappa [R(\bar{a}) - R(b)], \quad \text{for every } b \in A. \tag{2.15}$$

Furthermore, if $\bar{a} \in A$ is a minimizer of (2.14)*, then it solves the following variational inequality of finding $\bar{a} \in A$ such that*

$$\langle Du(\bar{a}), b - \bar{a} \rangle \geq \kappa \left[R(\bar{a}) - R(b)\right], \quad \text{for every } b \in A. \tag{2.16}$$

3. New Convergence Rates for the Inverse Problem

To derive new convergence rates, we assume that instead of z, the contaminated data $z_\delta \in V$ is available and satisfies

$$\|z - z_\delta\|_V \leq \delta. \tag{3.1}$$

We will now consider the following regularized ELS functional with the perturbed data z_δ:

$$J_\delta^\kappa(a) := \frac{1}{2} T(a, u(a) - z_\delta, u(a) - z_\delta) + \frac{\kappa}{2} R(a), \tag{3.2}$$

where $\kappa > 0$ is a regularization parameter, R is the regularization map satisfying ($\mathbb{C}_3$), and $u(a) \in V$ solves the variational problem (2.3).

Given the data $z \in V$, we recall that a solution $\bar{a} \in A$ is called R-minimizing if it solves the following optimization problem:

$$\min R(a), \quad \text{subject to } a \in S_A := \{a \in A|\ u(a) = z\}.$$

We have the following result concerning the convergence rates.

Theorem 3.1. *Let $\bar{a}$ be an R-minimizing solution. Assume that there exists $w^* \in V^*$ such that*

$$s^* := Du(\bar{a})^* w^* \in \partial R(\bar{a}). \tag{3.3}$$

where $Du(\bar{a})^*$ is the adjoint of $Du(\bar{a})$. Let $a_{\kappa,\delta}$ be a solution of (3.2). Then, for the choice $\kappa \sim \delta$, the following estimates hold:

$$D_{R,s^*}(a_{\kappa,\delta}, \bar{a}) = O(\delta),$$
$$\|u(a_{\kappa,\delta}) - z_\delta\|_V = O(\delta).$$

Proof. First of all, note that with $s^* := Du(\bar{a})^* w^* \in \partial R(\bar{a})$, we have

$$\begin{aligned}
D_{R,s^*}(a_\kappa^{\delta,\bar{q}}, \bar{a}) &= R(a_\kappa^\delta) - R(\bar{a}) + \langle s^*, a_\kappa^\delta - \bar{a}\rangle_{\widehat{X}} \\
&= R(a_\kappa^\delta) - R(\bar{a}) + \langle Du(\bar{a})^* w^*, a_\kappa^\delta - \bar{a}\rangle_X \\
&= R(a_\kappa^\delta) - R(\bar{a}) + \langle Du(\bar{a})^* w^*, a_\kappa^\delta - \bar{a}\rangle_B \\
&= R(a_\kappa^\delta) - R(\bar{a}) + \langle w^*, Du(\bar{a})(a_\kappa^\delta - \bar{a})\rangle_V \\
&= R(a_\kappa^\delta) - R(\bar{a}) + \langle w, Du(\bar{a})(a_\kappa^\delta - \bar{a})\rangle_V,
\end{aligned} \tag{3.4}$$

where $w \in V$ is the Riesz element associated to $w^* \in V^*$. Here we abused the notation by using the same notation $\langle \cdot, \cdot \rangle_V$ for the inner product in V and the pairing between V and V^*.

Since $a_{\kappa,\delta} \in \mathbb{K}$ is a regularized solution, the following inequality holds:

$$J_\delta(a_{\kappa,\delta}) + \kappa R(a_{\kappa,\delta}) \le J_\delta(a) + \kappa R(a), \quad \text{for every } a \in A, \tag{3.5}$$

where

$$J_\delta(a) = \frac{1}{2} T(a, u(a) - z_\delta, u(a) - z_\delta).$$

By using the ellipticity of T, we obtain

$$J_\delta(a_{\kappa,\delta}) = \frac{1}{2} T(a_{\kappa,\delta}, u(a_{\kappa,\delta}) - z_\delta, u(a_{\kappa,\delta}) - z_\delta) \ge \frac{\alpha}{2} \|u(a_{\kappa,\delta}) - z_\delta\|_V^2. \tag{3.6}$$

Setting $a = \bar{a}$ in (3.5), we obtain

$$J_\delta(a_{\kappa,\delta}) + \kappa R(a_{\kappa,\delta}) \le J_\delta(\bar{a}) + \kappa R(\bar{a}).$$

Next, by using the definition of the Bregman distance, we have

$$\begin{aligned}
J_\delta(a_{\kappa,\delta}) + \kappa D_R(a_{\kappa,\delta}, \bar{a}) &= J_\delta(a_{\kappa,\delta}) + \kappa \left[R(a_{\kappa,\delta}) - R(\bar{a}) + \langle Du(\bar{a})^* w, \bar{a} - a_{\kappa,\delta}\rangle_{\widehat{X}}\right] \\
&\le J_\delta(\bar{a}) + \kappa \langle w, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})\rangle_V \\
&\le c_1 \delta^2 + \kappa \langle w, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})\rangle_V,
\end{aligned} \tag{3.7}$$

where we used (2.1), (2.12), (3.4), and the attainability $u(\bar{a}) = z$. Also, with $c_1 = \frac{\beta}{2}\|\bar{a}\|_B$, we have

$$J_\delta(\bar{a}) = \frac{1}{2}T(\bar{a}, u(\bar{a}) - z_\delta, u(\bar{a}) - z_\delta) \leq \frac{\beta}{2}\|\bar{a}\|_B \|u(\bar{a}) - z_\delta\|_V^2 \leq c_1\delta^2.$$

Given $w \in V$, we now consider the variational problem of finding $\widetilde{w} \in V$ such that

$$T(\bar{a}, \widetilde{w}, v) = \langle w, v\rangle_V, \quad \text{for every } v \in V. \tag{3.8}$$

By the standard arguments, (3.8) is uniquely solvable and $\|\widetilde{w}\|_V$ is bounded by a constant involving a bound on $\|w\|_V$. We take $v = Du(\bar{a})(\bar{a} - a_{\kappa,\delta})$ in (3.8), which leads to

$$T(\bar{a}, \widetilde{w}, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})) = \langle w, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})\rangle_V.$$

Using the symmetry of the trilinear form, we have

$$T(\bar{a}, Du(\bar{a})(\bar{a} - a_{\kappa,\delta}), \widetilde{w}) = \langle w, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})\rangle_V,$$

which, by leveraging the derivative characterization (2.7) can be written as follows:

$$T(a_{\kappa,\delta} - \bar{a}, u(\bar{a}), \widetilde{w}) = \langle w, Du(\bar{a})(\bar{a} - a_{\kappa,\delta})\rangle_V. \tag{3.9}$$

Since the following two variational problems are uniquely solvable

$$\begin{aligned} T(\bar{a}, u(\bar{a}), v) &= \langle m, v\rangle, \quad \text{for every } v \in V,\\ T(a_{\kappa,\delta}, u(a_{\kappa,\delta}), v) &= \langle m, v\rangle, \quad \text{for every } v \in V, \end{aligned}$$

we deduce that

$$T(\bar{a}, u(\bar{a}), \widetilde{w}) = T(a_{\kappa,\delta}, u(a_{\kappa,\delta}), \widetilde{w}).$$

Equipped with the above identity, we return to (3.9) and perform a simple calculation as follows:

$$\begin{aligned} T(a_{\kappa,\delta} - \bar{a}, u(\bar{a}), \widetilde{w}) &= T(a_{\kappa,\delta}, u(\bar{a}), \widetilde{w}) - T(\bar{a}, u(\bar{a}), \widetilde{w})\\ &= T(a_{\kappa,\delta}, u(\bar{a}), \widetilde{w}) - T(a_{\kappa,\delta}, u(a_{\kappa,\delta}), \widetilde{w})\\ &= T(a_{\kappa,\delta}, u(\bar{a}) - z_\delta, \widetilde{w}) + T(a_{\kappa,\delta}, z_\delta - u(a_{\kappa,\delta}), \widetilde{w})\\ &\leq \beta\|a_{\kappa,\delta}\|_B \|u(\bar{a}) - z_\delta\|_V \| \|\widetilde{w}\|_V + \beta\|a_{\kappa,\delta}\|_B \|u(a_{\kappa,\delta}) - z_\delta\|_V \| \|\widetilde{w}\|_V\\ &\leq \beta\|a_{\kappa,\delta}\|_B \|u(\bar{a}) - z_\delta\|_V \|\widetilde{w}\|_V\\ &\quad + \beta\|a_{\kappa,\delta}\|_B \|u(a_{\kappa,\delta}) - z_\delta\|_V \|\widetilde{w}\|_V\\ &\leq \beta\delta\|a_{\kappa,\delta}\|_B \|\widetilde{w}\|_V + \frac{\kappa\beta^2}{\alpha}\|a_{\kappa,\delta}\|_B^2 \|\widetilde{w}\|_V^2 + \frac{\alpha}{4\kappa}\|u(a_{\kappa,\delta}) - z_\delta\|_V^2\\ &\leq c\delta + c\kappa + \frac{\alpha}{4\kappa}\|u(a_{\kappa,\delta}) - z_\delta\|_V^2, \end{aligned} \tag{3.10}$$

where c is a positive constant.

We now combine (3.6), (3.7), (3.9), and (3.10) to obtain

$$\frac{\alpha}{2}\|u(a_{\kappa,\delta}) - z_\delta\|_{L^2(\Omega;V)}^2 + \kappa D_R(a_{\kappa,\delta}, \bar{a}) \leq c_1\delta^2 + c\kappa\delta + c\kappa^2 + \frac{\alpha}{4}\|u(a_{\kappa,\delta}) - z_\delta\|_V^2,$$

which yields

$$\begin{aligned} D_R(a_{\kappa,\delta}, \bar{a}) &\leq C\frac{(\delta^2 + \kappa\delta + \kappa^2)}{\kappa},\\ \|u(a_{\kappa,\delta}) - z_\delta\|_{L^2(\Omega;V)}^2 &\leq C(\delta^2 + \kappa\delta + \kappa^2), \end{aligned}$$

where $C > 0$ is a constant. Thus, as a consequence of the above inequalities, for $\kappa \sim \delta$, we have

$$D_R(a_{\kappa,\delta}, \bar{a}) = O(\delta),$$
$$\|u(a_{\kappa,\delta}) - z_\delta\|^2_{L^2(\Omega;V)} = O(\delta^2),$$

which completes the proof. □

Remark 3.1. The OLS functional, defined in (2.4), assumes that the data z belong to the Hilbert space Z. On the other hand, the ELS functional requires the data z to be in the space V. From a practical standpoint, this means that the ELS formulation demands data with higher regularity compared to the OLS formulation. However, it is worth noting that the above proof demonstrates that the estimates indeed hold for the OLS functional when it is defined using the norm of the Hilbert space V. We also observe that the abstract framework can be readily adapted to incorporate scenarios where regularization is accomplished through a combination of a quadratic regularizer and a convex regularizer, as exemplified in the studies conducted by Chavent and Kunisch [9] and Hào and Quyen [26].

4. Concluding Remarks

We investigated the nonlinear inverse problem of estimating variable parameters in general variational problems. Our primary contribution lies in deriving novel convergence rates for this inverse problem, eliminating the need for smallness conditions. To achieve this, we employ the convex energy least squares functional as a fundamental technical tool, leveraging its effectiveness in extending the OLS formulation. It would be of interest to verify the convergence rates computationally. Furthermore, extending them to stochastic inverse problems (see [12, 28]) and to more general identification problems in variational inequalities is also an important topic to pursue.

Acknowledgments

The research of D. N. Hào was supported by the VAST grant NVCC01.06/22-23. The research of Akhtar Khan was supported by the FEAD AY-22-23 Grant. Simeon Reich was partially supported by the Israel Science Foundation (Grant 820/17), by the Fund for the Promotion of Research at the Technion (Grant 2001893), and by the Technion General Research Fund (Grant 2016723).

References

[1] R. Acar, C. R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, Inverse Problems 10 (6) (1994) 1217–1229.

[2] A. S. Ackleh, S. Aizicovici, S. Reich, Parameter identification in nonlocal nonlinear evolution equations, Numer. Funct. Anal. Optim. 21 (5–6) (2000) 553–570.

[3] H. Ammari, P. Garapon, F. Jouve, Separation of scales in elasticity imaging: a numerical study, J. Comput. Math. 28 (3) (2010) 354–370.

[4] H. T. Banks, S. Reich, I. G. Rosen, An approximation theory for the identification of nonlinear distributed parameter systems, SIAM J. Control Optim. 28 (3) (1990) 552–569.

[5] C. Boehm, M. Ulbrich, A semismooth Newton-CG method for constrained parameter identification in seismic tomography, SIAM J. Sci. Comput. 37 (5) (2015) S334–S364.

[6] R. Boiger, B. Kaltenbacher, An online parameter identification method for time dependent partial differential equations, Inverse Problems 32 (4) (2016) 045006, 28

[7] L. M. Bregman, A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming, Z. Vycisl. Mat i Mat. Fiz. 7 (1967) 620–631.

[8] M. Burger, S. Osher, Convergence rates of convex variational regularization, Inverse Problems 20 (5) (2004) 1411–1421.

[9] G. Chavent, K. Kunisch, Regularization of linear least squares problems by total bounded variation, ESAIM Control Optim. Calc. Var. 2 (1997) 359–376.

[10] C. Clason, L^∞ fitting for inverse problems with uniform noise, Inverse Problems 28 (10) (2012) 104007, 18.

[11] E. Crossen, M. S. Gockenbach, B. Jadamba, A. A. Khan, B. Winkler, An equation error approach for the elasticity imaging inverse problem for predicting tumor location, Comput. Math. Appl. 67 (1) (2014) 122–135.

[12] J. Dippon, J. Gwinner, A. A. Khan, M. Sama, A new regularized stochastic approximation framework for stochastic inverse problems, Nonlinear Anal. Real World Appl. 73 (2023) Paper No. 103869, 29.

[13] H. W. Engl, K. Kunisch, A. Neubauer, Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems, Inverse Problems 5 (4) (1989) 523–540.

[14] R. O. Evstigneev, M. Y. Medvedik, Y. G. Smirnov, Inverse problem of determining parameters of inhomogeneity of a body from acoustic field measurements, Comput. Math. Math. Phys. 56 (3) (2016) 483–490.

[15] A. Gholami, A. Mang, G. Biros, An inverse problem formulation for parameter estimation of a reaction-diffusion model of low grade gliomas, J. Math. Biol. 72 (1-2) (2016) 409–433.

[16] E. Giusti, Minimal surfaces and functions of bounded variation, vol. 80 of Monographs in Mathematics, Birkhauser Verlag, Basel, 1984.

[17] M. S. Gockenbach, B. Jadamba, A. A. Khan, Numerical estimation of discontinuous coefficients by the method of equation error, Int. J. Math. Comput. Sci. 1 (3) (2006) 343–359.

[18] M. S. Gockenbach, B. Jadamba, A. A. Khan, C. Tammer, B. Winkler, Proximal methods for the elastography inverse problem of tumor identification using an equation error approach, in: Advances in variational and hemivariational inequalities, vol. 33 of Adv. Mech. Math., Springer, Cham, 2015, pp. 173–197.

[19] M. S. Gockenbach, A. A. Khan, Identification of Lame parameters in linear elasticity: a fixed point approach, J. Ind. Manag. Optim. 1 (4) (2005) 487–497.

[20] M. S. Gockenbach, A. A. Khan, An abstract framework for elliptic inverse problems: Part 1. an output least-squares approach, Mathematics and Mechanics of Solids 12 (3) (2007) 259–276.

[21] M. S. Gockenbach, A. A. Khan, An abstract framework for elliptic inverse problems. II. An augmented Lagrangian approach, Math. Mech. Solids 14 (6) (2009) 517–539.

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

[23] W. Hager, C. Ngo, M. Yashtini, H.-C. Zhang, An alternating direction approximate Newton algorithm for ill-conditioned inverse problems with application to parallel MRI, J. Oper. Res. Soc. China 3 (2) (2015) 139–162.

[24] D. N. Hào, T. N. T. Quyen, Convergence rates for Tikhonov regularization of coefficient identification problems in Laplace-type equations, Inverse Problems 26 (12) (2010) 125014, 23.

[25] D. N. Hào, T. N. T. Quyen, Convergence rates for total variation regularization of coefficient identification problems in elliptic equations I, Inverse Problems 27 (7) (2011) 075008, 28.

[26] D. N. Hào, T. N. T. Quyen, Convergence rates for total variation regularization of coefficient identification problems in elliptic equations II, J. Math. Anal. Appl. 388 (1) (2012) 593–616.

[27] B. Jadamba, A. A. Khan, G. Rus, M. Sama, B. Winkler, A new convex inversion framework for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location, SIAM J. Appl. Math. 74 (5) (2014) 1486–1510.

[28] B. Jadamba, A. A. Khan, M. Sama, H.-J. Starkloff, C. Tammer, A convex optimization framework for the inverse problem of identifying a random parameter in a stochastic partial differential equation, SIAM/ASA Journal on Uncertainty Quantification 9 (2) (2021) 922–952.

[29] B. Jadamba, A. A. Khan, M. Sama, C. Tammer, On convex modified output least-squares for elliptic inverse problems: stability, regularization, applications, and numerics, Optimization 66 (6) (2017) 983–1012.

[30] B. Kaltenbacher, B. Hofmann, Convergence rates for the iteratively regularized Gauss-Newton method in Banach spaces, Inverse Problems 26 (3) (2010) 035007, 21.

[31] S. Kindermann, L. D. Mutimbu, E. Resmerita, A numerical study of heuristic parameter choice rules for total variation regularization, J. Inverse Ill-Posed Probl. 22 (1) (2014) 63–94.
[32] I. Knowles, Parameter identification for elliptic problems, J. Comput. Appl. Math. 131 (1-2) (2001) 175–194.
[33] P. Kuchment, D. Steinhauer, Stabilizing inverse problems by internal data. II: non-local internal data and generic linearized uniqueness, Anal. Math. Phys. 5 (4) (2015) 391–425.
[34] P. Kugler, E. Sincich, Logarithmic convergence rates for the identification of a nonlinear Robin coefficient, J. Math. Anal. Appl. 359 (2) (2009) 451–463.
[35] T. Liu, A wavelet multiscale-homotopy method for the parameter identification problem of partial differential equations, Comput. Math. Appl. 71 (7) (2016) 1519–1523.
[36] S. Manservisi, M. Gunzburger, A variational inequality formulation of an inverse elasticity problem, Appl. Numer. Math. 34 (1) (2000) 99–126.
[37] A. Neubauer, T. Hein, B. Hofmann, S. Kindermann, U. Tautenhahn, Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces, Appl. Anal. 89 (11) (2010) 1729–1743.
[38] D. Reem, S. Reich, A. De Pierro, Re-examination of Bregman functions and new properties of their divergences, Optimization 68 (1) (2019) 279–348.
[39] E. Resmerita, Regularization of ill-posed problems in Banach spaces: convergence rates, Inverse Problems 21 (4) (2005) 1303–1314.
[40] E. Resmerita, O. Scherzer, Error estimates for non-quadratic regularization and the relation to enhancement, Inverse Problems 22 (3) (2006) 801–814.
[41] Y. Xu, J. Zou, Convergence of an adaptive finite element method for distributed flux reconstruction, Math. Comp. 84 (296) (2015) 2645–2663.
[42] J. Zou, Numerical methods for elliptic inverse problems, Int. J. Comput. Math. 70 (2) (1998) 211–232.