Sparsity regularization of the diffusion coefficient identification problem: well-posedness and convergence rates

Pham Quy Muoi†

March 13, 2014

† Department of Mathematics, Danang University of Education - University of Danang
459 Ton Duc Thang, Danang, Vietnam
Email:pqmuoi@ud.edu.vn

Abstract

In this paper, we investigate sparsity regularization for the diffusion coefficient identification problem. Here, the regularization method is incorporated with the energy functional approach. The advantages of our approach are to deal with convex minimization problems. Therefore, the well-posedness of the problem is obtained without requiring regularity property of the parameter. The convexity of regularized problems also allows to use the fast algorithms developed recently. Furthermore, the convergence rates of the method are obtained under a simple source condition.

The main results of the paper are the well-posedness and convergence rates of sparsity regularization. We also obtain some new results of the continuity and the differentiability of related operators.

2010 Mathematics Subject Classification: 47A52, 49N30, 49N45
Keywords: *Sparsity regularization, Diffusion coefficient identification problem, Nonlinear inverse problems, Well-posedness, Convergence rate.*

1 Introduction

The diffusion coefficient identification problem is to identify the coefficient σ in the equation

$$-\operatorname{div}(\sigma\nabla\phi) = y \text{ in } \Omega,\ \phi = 0 \text{ on } \partial\Omega \tag{1}$$

from noisy data $\phi^\delta \in H_0^1(\Omega)$ of ϕ such that

$$\left\|\phi^* - \phi^\delta\right\|_{H^1(\Omega)} \leq \delta. \quad (\delta > 0)$$

This problem has attracted great attention of many researchers. For surveys on this problem, we refer to [18, 40, 14, 27, 34, 28, 11, 38, 1, 9] and the references therein. It is well-known that the problem is ill-posed and thus need to be regularized. There have been several regularization methods proposed. Among of them, Tikhonov regularization [18, 13] and the total variational regularization [40, 8] are most popular.

In some applications, the coefficient σ^*, which needs to be recovered, has a sparse presentation, i.e. the number of nonzero components of $\sigma^* - \sigma^0$ are finite in an orthonormal basis (or frame) of $L^2(\Omega)$. The sparsity of $\sigma^* - \sigma^0$ promotes to use sparsity regularization.

Sparsity regularization has been of interest by many researchers for the last years. The well-posedness and some convergence rates of the method have been analyzed for linear inverse problems [12] as well as for nonlinear inverse problems [17, 6, 37]. Some numerical algorithms have also been proposed [29, 12, 4, 3, 33, 2]. It is shown that sparsity regularization is simple for use and very efficient for inverse problems with sparse solutions. This method has been investigated and applied very successfully to some fields such as for compressive imaging [36, 39] and electrical impedance tomography [23, 16, 24].

Note that it is possible to apply the least squares approach in [17] or more general misfit term in [15] for our problem. However, it is not clear that the solution operator $F_D(\cdot)\,y$, which maps each parameter σ into the solution of (1), is weakly sequentially closed in $L^2(\Omega)$ without additional conditions. Therefore, if the such approaches in [17, 15] are applied, then it needs further conditions. Moreover, the approaches lead to a non-convex minimization problem and the source conditions are difficult to be checked for the problem, see e.g. [18].

To overcome this shortcoming, we use the energy functional approach incorporating with sparsity regularization, i.e. considering the minimization problem

$$\min_{\sigma\in\mathcal{A}} F_{\phi^\delta}(\sigma) + \alpha\Phi\left(\sigma-\sigma^0\right), \tag{2}$$

where

- $\mathcal{A}$ is an admissible set defined by

$$\mathcal{A} = \left\{\sigma \in L^\infty(\Omega) : \lambda \le \sigma \le \lambda^{-1} \text{ a.e. on } \Omega \text{ and } \operatorname{supp}\left(\sigma-\sigma^0\right) \subset \Omega' \subset\subset \Omega\right\}. \tag{3}$$

  Here $\lambda \in (0,1)$ is a given constant ,σ^0 is the background value of σ, and Ω' is an open set with the smooth boundary that contained compactly in Ω.

- $\alpha > 0$ is a regularization parameter.

- $F_{\phi^\delta}(\sigma)$ and $\Phi(\vartheta)$ are defined by

$$F_{\phi^\delta}(\sigma) := \int_\Omega \sigma\left|\nabla\left(F_D(\sigma)y-\phi^\delta\right)\right|^2 dx, \tag{4}$$

$$\Phi(\vartheta) := \sum \omega_k\left|\langle\vartheta,\varphi_k\rangle\right|^p, \quad (1\le p\le 2) \tag{5}$$

  where $\{\varphi_k\}$ is an orthonormal basis (or frame) of $L^2(\Omega)$, $\langle\cdot,\cdot\rangle$ is the scalar product in $L^2(\Omega)$, and $\omega_k \ge \omega_{min} > 0$ for all k.

Let us give a short discussion on the admissible set $\mathcal{A}$. The first condition $\sigma \in [\lambda, \lambda^{-1}]$ is required to obtain the unique existence of the solution of (1). The second condition, $\operatorname{supp}\left(\sigma-\sigma^0\right) \subset \Omega' \subset\subset \Omega$, means that $\sigma-\sigma^0$ has a sparse expansion in a suitable basis $\{\varphi_k\}$ [31]. Under this assumption, sparsity regularization should be applied [12, 17]. Note that our assumption is satisfied if the values of the parameter (we need to recover) are known on the boundary $\partial\Omega$. In such a situation any $\sigma^0 \in L^\infty(\Omega)$ that has the same values of the parameter on the boundary is satisfied. This assumption is often used in many papers such as [40, 26, 18].

Note that the energy functional approach was first introduced by Zou [40] and then was used by Knowles in [26]. However, the authors in those papers did not consider the well-posedness and convergence rates of Tikhonov-type regularization, they have used the energy functional for finding the numerical solution using some numerical algorithms. Recently, Hao and Quyen have used this approach incorporating with either Tikhonov regularization or the total variation regularization for some problems [18, 19, 20, 21] and the well-posedness and convergence rates are obtained under some conditions.

This paper is motivated by the numerical results in [29, 31]. In those papers, we have applied some numerical algorithms for problem (2). The algorithms have worked very well and very effective, but the well-posedness of problem (2) has not been analyzed. Here, we will investigate the differentiability of $F_{\phi^\delta}(\cdot)$ and the well-posedness of problem (2) as well as convergence rates. We will prove that problem (2) is convex and well-posed, and if there exists w^* such that $\xi = \left(F'_D(\sigma^+)y\right)^* w^* \in \partial\Phi\left(\sigma^+-\sigma^0\right)$, then the convergence rates

$$D_\xi\left(\sigma^p_{\alpha,\delta},\sigma^+\right) = O(\delta) \text{ and } \left\|\sigma^p_{\alpha,\delta}-\sigma^+\right\|_{L^2(\Omega)} = O\left(\sqrt{\delta}\right) \quad (1<p\le 2),$$

are obtained as $\delta\to 0$ and $\alpha\sim\delta$. Note that $\sigma^p_{\alpha,\delta}$ is a minimizer of (2), σ^+ is a Φ-minimizing solution of the diffusion coefficient identification problem and $D_\xi\left(\sigma^p_{\alpha,\delta},\sigma^+\right)$ is the Bregman distance of two elements $\sigma^p_{\alpha,\delta}$ and σ^+ [20, 21].

Comparing the standard conditions in [17] and the references therein, our source condition is very simple and does not require the smallness. Furthermore, the objective functional in (2) is now convex and thus its global minimizers are easy to find by available efficient algorithms [29, 31].

2 Auxiliary Results

We recall that a function ϕ in $H_0^1(\Omega)$ is a weak solution of (1) if the identity

$$\int_\Omega \sigma \nabla \phi \cdot \nabla v dx = \int_\Omega y v dx \tag{6}$$

holds for all $v \in H_0^1(\Omega)$.

If $\sigma \in \mathcal{A}$ and $y \in L^2(\Omega)$, then there is a unique weak solution $\phi \in H_0^1(\Omega)$ of (1) [18], which satisfies the inequality

$$\|\phi\|_{H^1(\Omega)} \le \frac{1}{C} \|y\|_{L^2(\Omega)}, \tag{7}$$

where $C > 0$ is a constant depending only on Ω and λ.

In the next sections, two following inequalities are used:

- For any $\eta \in H_0^1(\Omega)$ and $\sigma \in \mathcal{A}$, in virtue of the Poincaré-Friedrichs inequality we have

$$\int_\Omega \sigma |\nabla \eta|^2 dx \ge C \|\eta\|_{H^1(\Omega)}^2 \tag{8}$$

  with $C > 0$ defined by (7).

- For any $y \in L^r(\Omega), r \ge 2$ with a bounded set $\Omega \subset \mathbb{R}^d$, we have

$$\|y\|_{L^2(\Omega)} \le |\Omega|^{\frac{1}{2}-\frac{1}{r}} \|y\|_{L^r(\Omega)}. \tag{9}$$

We shall endow the set $\mathcal{A}$ with the $L^q(\Omega)-$norm, $q \in [1, \infty)$ and define the nonlinear coefficient-to-solution mapping $F_D(\cdot) y : \mathcal{A} \subset L^q(\Omega) \to H_0^1(\Omega)$ which maps the coefficient $\sigma \in \mathcal{A}$ to the solution $u = F_D(\sigma) y$ of problem (1).

Before considering sparsity regularization for the problem, we analyze some properties of $F_D(\cdot) y$ and $F_{\phi^\delta}(\cdot)$ with respect to the L^q-norm. These properties are needed for investigating the well-posedness and convergence rates of the method as well as numerical algorithms. They are derived by exploiting Meyers' gradient estimate [30], which has recently been employed by [35, 24].

Theorem 1 (Meyers' theorem) *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ $(d \ge 2)$. Assume that $\sigma \in L^\infty(\Omega)$ satisfies $\lambda < \sigma < \lambda^{-1}$ for some fixed $\lambda \in (0,1)$. For $z \in (L^r(\Omega))^d$ and $y \in L^r(\Omega)$, let $\phi \in H^1(\Omega)$ be a weak solution of the equation*

$$-\operatorname{div}(\sigma \nabla \phi) = -\operatorname{div}(z) + y \text{ in } \Omega.$$

Then, there exists a constant $Q \in (2, +\infty)$ depending on λ and d only, $Q \to 2$ as $\lambda \to 0$ and $Q \to \infty$ as $\lambda \to 1$, such that for any $2 < r < Q$, $\phi \in W_{loc}^{1,r}(\Omega)$ and for any $\Omega' \subset\subset \Omega$

$$\|\nabla \phi\|_{L^r(\Omega')} \le C' \left(\|\phi\|_{H^1(\Omega)} + \|z\|_{L^r(\Omega)} + \|y\|_{L^r(\Omega)} \right),$$

where the constant C' depends on λ, d, r, Ω' and Ω.

Using this result, we can show that the mappings $F_D(\cdot) y$ and $F_{\phi^\delta}(\cdot)$ are continuous and continuous Fréchet differentiable on the set $\mathcal{A}$ with respect to the L^q-norm. These results are shown in the following lemmas.

Lemma 2 *Let $q \in \left(\frac{2Q}{Q-2}, \infty\right]$, $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$ and $y \in L^r(\Omega)$. For $\sigma, \sigma + \vartheta \in \mathcal{A}$, we have*

$$\|\nabla F_D(\sigma + \vartheta) y - \nabla F_D(\sigma) y\|_{L^2(\Omega)} \le C \|\vartheta\|_{L^q(\Omega')} \|y\|_{L^r(\Omega)},$$

where C is a positive constant depending on λ, d, r, Ω' and Ω.

Proof. The weak solution formulas of $F_D(\sigma) y$ and $F_D(\sigma + \vartheta) y$ give

$$\int_\Omega \sigma \nabla F_D(\sigma) y \cdot \nabla v dx = \int_\Omega (\sigma + \vartheta) \nabla F_D(\sigma + \vartheta) y \cdot \nabla v dx, \ \forall v \in H_0^1(\Omega),$$

i.e.

$$\int_{\Omega}\sigma\nabla\left(F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\right)\cdot\nabla v dx=-\int_{\Omega}\vartheta\nabla F_D\left(\sigma+\vartheta\right)y\cdot\nabla v dx,\ \forall v\in H_0^1\left(\Omega\right).$$

Taking $v=F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\in H_0^1\left(\Omega\right)$ in the last equation, we obtain

$$\begin{aligned}\int_{\Omega}\sigma\left|\nabla\left(F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\right)\right|^2 dx&=-\int_{\Omega}\vartheta\nabla F_D\left(\sigma+\vartheta\right)y\cdot\nabla\left(F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\right)dx\\&=-\int_{\Omega'}\vartheta\nabla F_D\left(\sigma+\vartheta\right)y\cdot\nabla\left(F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\right)dx\\&\leq\|\vartheta\|_{L^q(\Omega')}\|\nabla F_D\left(\sigma+\vartheta\right)y\|_{L^r(\Omega')}\|\nabla\left(F_D\left(\sigma+\vartheta\right)y-F_D\left(\sigma\right)y\right)\|_{L^2(\Omega)},\end{aligned}$$

where $\frac{1}{q}+\frac{1}{r}=\frac{1}{2}$. The assumption $q\in\left(\frac{2Q}{Q-2},\infty\right]$ implies that $r\in(2,Q)$. By Theorem 1, there exist constants C and C' such that

$$\|\nabla F_D\left(\sigma+\vartheta\right)y\|_{L^r(\Omega')}\leq C'\left(\|F_D\left(\sigma+\vartheta\right)y\|_{H^1(\Omega)}+\|y\|_{L^r(\Omega)}\right)\overset{(7),(9)}{\leq}C\|y\|_{L^r(\Omega)}.$$

It follows that there exists a constant C such that

$$\|\nabla F_D\left(\sigma+\vartheta\right)y-\nabla F_D\left(\sigma\right)y\|_{L^2(\Omega)}\leq C\|\vartheta\|_{L^q(\Omega')}\|y\|_{L^r(\Omega)}.$$

■

Remark 3 *1) Note that for $\sigma,\sigma+\vartheta\in\mathcal{A}$ and $1\leq q_1\leq q_2$, we have*

$$|\Omega|^{-1/q_1}\|\vartheta\|_{L^{q_1}(\Omega)}\leq|\Omega|^{-1/q_2}\|\vartheta\|_{L^{q_2}(\Omega)},$$

and

$$\|\vartheta\|^{q_2}_{L^{q_2}(\Omega)}\leq\left(2\lambda^{-1}\right)^{q_2-q_1}\|\vartheta\|^{q_1}_{L^{q_1}(\Omega)}.$$

This means that the convergence of ϑ to zero with respect to the $L^{q_1}\left(\Omega\right)-$norm and the $L^{q_2}\left(\Omega\right)-$norm are equivalent.

2) By the above lemma, $F_D\left(\cdot\right)y$ is Lipschitz continuous on $\mathcal{A}$ with respect to the $L^q\left(\Omega\right)-$norm for $q\in\left(\frac{2Q}{Q-2},\infty\right]$. Furthermore, by the above remark, it implies that $F_D\left(\cdot\right)y$ is continuous on $\mathcal{A}$ with respect to the $L^q\left(\Omega\right)$-norm for any $q\geq 1$.

Next, we consider the differentiability of $F_D\left(\cdot\right)y$ and $F_{\phi^\delta}\left(\cdot\right)$. Here the continuous Fréchet differentiability is understood as in [24].

Lemma 4 *Let $q\in\left(\frac{2Q}{Q-2},\infty\right]$, $\frac{1}{q}+\frac{1}{r}=\frac{1}{2}$ and $y\in L^{r+\epsilon}\left(\Omega\right)$ with some $\epsilon>0$. Then, the mapping $F_D\left(\cdot\right)y:\mathcal{A}\subset L^q\left(\Omega\right)\to H_0^1\left(\Omega\right)$ is continuously Fréchet differentiable on $\mathcal{A}$ and for each $\sigma\in\mathcal{A}$, the Fréchet derivative $F_D'\left(\sigma\right)y$ of $F_D\left(\cdot\right)y$ has the property that the differential $\eta:=F_D'\left(\sigma\right)y\left(\vartheta\right)$, with any $\vartheta\in L^\infty\left(\Omega'\right)$ extended by zero outside Ω', is the (unique) weak solution of the Dirichlet problem*

$$-\operatorname{div}\left(\sigma\nabla\eta\right)=\operatorname{div}\left(\vartheta\nabla F_D\left(\sigma\right)y\right)\ \text{ in }\Omega,\ \eta=0\ \text{ on }\partial\Omega$$

in the sense that it satisfies the equation

$$\int_{\Omega}\sigma\nabla F_D'\left(\sigma\right)y\left(\vartheta\right)\cdot\nabla v dx=-\int_{\Omega}\vartheta\nabla F_D\left(\sigma\right)y\cdot\nabla v dx\tag{10}$$

for all $v\in H_0^1\left(\Omega\right)$. Moreover,

$$\|F_D'\left(\sigma\right)y\left(\vartheta\right)\|_{H^1(\Omega)}\leq C_1\|y\|_{L^r(\Omega)}\|\vartheta\|_{L^q(\Omega')},\forall\vartheta\in L^\infty\left(\Omega'\right),\tag{11}$$

where C_1 is a positive constant depending on λ,d,r,Ω' and Ω.

Proof. Note that variational equation (10) has the unique solution $\eta := \eta(\vartheta) = F_D'(\sigma)\, y(\vartheta) \in H_0^1(\Omega)$ with $\sigma \in \mathcal{A}$. We first show that for a fixed σ in $\mathcal{A}$, $\eta = \eta(\vartheta)$ defines a bounded linear operator from $L^q(\Omega')$ to $H_0^1(\Omega)$ for any $q \in \left(\frac{2Q}{Q-2}, \infty\right]$. From (10), η is a linear operator of ϑ. By the weak solution formula of η and the generalized Hölder inequality, we have

$$\begin{aligned}\int_\Omega \sigma \nabla\eta \cdot \nabla\eta dx &= -\int_\Omega \vartheta \nabla F_D(\sigma)\, y \cdot \nabla\eta dx \\ &= -\int_{\Omega'} \vartheta \nabla F_D(\sigma)\, y \cdot \nabla\eta dx \\ &\le \|\vartheta\|_{L^q(\Omega')} \|\nabla F_D(\sigma)\, y\|_{L^r(\Omega')} \|\nabla\eta\|_{L^2(\Omega)}.\end{aligned}$$

From the last inequality and (8), there exists a constant C such that

$$\|\eta\|_{H^1(\Omega)} \le C \|\vartheta\|_{L^q(\Omega')} \|\nabla F_D(\sigma)\, y\|_{L^r(\Omega')}. \tag{12}$$

Besides, the assumption $q \in \left(\frac{2Q}{Q-2}, \infty\right]$ implies $r \in (2, Q)$. By Theorem 1, (7) and (9), there exist positive constants C, C', C'' such that

$$\begin{aligned}\|\nabla F_D(\sigma)\, y\|_{L^r(\Omega')} &\le C' \left(\|F_D(\sigma)\, y\|_{H^1(\Omega)} + \|y\|_{L^r(\Omega)}\right) \\ &\le C' \left(\frac{1}{C} \|y\|_{L^2(\Omega)} + \|y\|_{L^r(\Omega)}\right) \\ &\le C'' \|y\|_{L^r(\Omega)}. \end{aligned} \tag{13}$$

Thus, due to two last inequalities, η is a bounded linear operator from $L^q(\Omega') \to H_0^1(\Omega)$ and there exists a positive constant C_1 such that

$$\|F_D'(\sigma)\, y(\vartheta)\|_{H^1(\Omega)} \le C_1 \|y\|_{L^r(\Omega)} \|\vartheta\|_{L^q(\Omega')}, \forall \vartheta \in L^\infty(\Omega').$$

We now show that $F_D(\cdot)\, y$ is Fréchet differentiable. Note that the function $R := F_D(\sigma+\vartheta)\, y - F_D(\sigma)\, y - \eta \in H_0^1(\Omega)$ is the weak solution of the equation

$$-\operatorname{div}((\sigma+\vartheta)\nabla R) = \operatorname{div}(\vartheta\nabla\eta) \text{ in } \Omega.$$

Taking R as the test function in the weak solution formula of R gives

$$\begin{aligned}\int_\Omega (\sigma+\vartheta) |\nabla R|^2 dx = -\int_\Omega \vartheta\nabla\eta \cdot \nabla R dx &= -\int_{\Omega'} \vartheta\nabla\eta \cdot \nabla R dx \\ &\le \|\vartheta\|_{L^q(\Omega')} \|\nabla\eta\|_{L^r(\Omega')} \|\nabla R\|_{L^2(\Omega)}.\end{aligned}$$

This implies that

$$\frac{\|R\|_{H^1(\Omega)}}{\|\vartheta\|_{L^q(\Omega')}} \le C \|\nabla\eta\|_{L^r(\Omega')}. \tag{14}$$

To show that $F_D(\cdot)\, y : \mathcal{A} \subset L^q(\Omega) \to H_0^1(\Omega)$ is continuously Fréchet differentiable and its differential $F_D'(\sigma)\, y(\vartheta)$ is η, we need to prove that $\|\nabla\eta\|_{L^r(\Omega')}$ converges to zero as $\|\vartheta\|_{L^q(\Omega')}$ converges to zero.
By Theorem 1, there exists a positive constant C such that

$$\|\nabla\eta\|_{L^r(\Omega')} \le C \left(\|\eta\|_{H^1(\Omega)} + \|\vartheta\nabla F_D(\sigma)\, y\|_{L^r(\Omega')}\right).$$

Since $\|\eta\|_{H^1(\Omega)}$ converges to zero as $\|\vartheta\|_{L^q(\Omega')}$ converges to zero by (12), we need to prove that

$$\|\vartheta\nabla F_D(\sigma)\, y\|_{L^r(\Omega')} \to 0.$$

Take any small $\epsilon_1 \in (0, \epsilon)$ such that $r' = r + \epsilon_1 \in (r, Q)$. Using Hölder's inequality, we deduce

$$\begin{aligned} \int_{\Omega'} |\vartheta \nabla F_D(\sigma) y|^r \, dx &= \int_{\Omega'} |\vartheta|^r |\nabla F_D(\sigma) y|^r \, dx \\ &\leq \left(\int_{\Omega'} |\vartheta|^{\frac{rr'}{r'-r}} dx \right)^{1-\frac{r}{r'}} \left(\int_{\Omega'} |\nabla F_D(\sigma) y|^{r'} dx \right)^{\frac{r}{r'}}. \\ &\leq C_2 \|y\|^r_{L^{r'}(\Omega)} \left(\int_{\Omega'} |\vartheta|^{\frac{rr'}{r'-r}} dx \right)^{1-\frac{r}{r'}}, \end{aligned} \tag{15}$$

where we have applied Theorem 1 to the term $\|\nabla F_D(\sigma) y\|_{L^{r'}(\Omega')}$, see (13). By Remark 3, the convergence of ϑ to zero with respect to the $L^{q_1}(\Omega)$−norm and the $L^{q_2}(\Omega)$−norm ($q_1, q_2 \in [1, \infty)$) are equivalent. Therefore, $\|\vartheta \nabla F_D(\sigma) y\|_{L^r(\Omega')}$ converges to zero as $\|\vartheta\|_{L^q(\Omega')}$ converges to zero. ■

Remark 5 *1) If $y \in L^r(\Omega)$, then from the proof above we conclude that $F_D(\cdot) y : \mathcal{A} \subset L^q(\Omega) \to H^1_0(\Omega)$ is Gâuteaux differentiable.*

2) This lemma improves the known results on the differentiability of $F_D(\cdot) y$ with respect to the L^∞−norm in [26, 18]. There, the authors have shown that $F_D(\cdot) y : \mathcal{A} \subset L^\infty(\Omega) \to H^1_0(\Omega)$ is the Fréchet differentiable under the condition $y \in L^\infty(\Omega)$ [26] or $y \in L^2(\Omega)$ [18].

Lemma 6 *For $\phi \in H^1_0(\Omega)$, the functional $F_\phi(\cdot) : \mathcal{A} \subset L^q(\Omega) \to \mathbb{R}$ defined by*

$$F_\phi(\sigma) = \int_\Omega \sigma |\nabla (F_D(\sigma) y - \phi)|^2 \, dx$$

has the following properties

1) For $q \geq 1$ and $y \in L^r(\Omega)$, $F_\phi(\cdot)$ is continuous with respect to the L^q−norm.

2) For $q \in \left(\frac{2Q}{Q-2}, \infty\right]$, $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$, and $y \in L^{r+\epsilon}(\Omega)$ with $\epsilon > 0$, if $\|\nabla \phi\|_{L^r(\Omega')} < \infty$, then $F_\phi(\cdot)$ is Fréchet differentiable with respect to the L^q-norm and

$$F'_\phi(\sigma) \vartheta = -\int_\Omega \vartheta \left(|\nabla F_D(\sigma) y|^2 - |\nabla \phi|^2 \right) dx.$$

Furthermore, $F_\phi(\cdot)$ is convex on the convex set $\mathcal{A}$ and the norm of operators $\|F''_\phi(\sigma)\|$ are uniformly bounded, i.e. there exists a constant M such that

$$\|F''_\phi(\sigma)\| \leq M, \forall \sigma \in \mathcal{A}.$$

Proof. 1) We first prove for $q \in \left(\frac{2Q}{Q-2}, \infty\right]$. For $\sigma, \sigma + \vartheta \in \mathcal{A}$, we have

$$\begin{aligned} & F_\phi(\sigma + \vartheta) - F_\phi(\sigma) \\ &= \int_\Omega (\sigma + \vartheta) |\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 - \sigma |\nabla (F_D(\sigma) y - \phi)|^2 \, dx \\ &= \int_\Omega \sigma \left(|\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 - |\nabla (F_D(\sigma) y - \phi)|^2 \right) dx + \int_\Omega \vartheta |\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 \, dx. \end{aligned}$$

Using the triangle inequality, generalized Hölder inequality and Theorem 1, the second term is estimated by

$$\begin{aligned} & \int_\Omega \vartheta |\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 \, dx = \int_{\Omega'} \vartheta |\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 \, dx \\ & \leq \|\vartheta\|_{L^q(\Omega')} \|\nabla (F_D(\sigma + \vartheta) y - \phi)\|_{L^2(\Omega)} \left(\|\nabla F_D(\sigma + \vartheta) y\|_{L^r(\Omega')} + \|\nabla \phi\|_{L^r(\Omega')} \right) \\ & \leq C \|\vartheta\|_{L^q(\Omega')}. \end{aligned}$$

On the other hand, by Lemma 2 the first term is estimated by

$$\begin{aligned} & \int_\Omega \sigma \left(|\nabla (F_D(\sigma + \vartheta) y - \phi)|^2 - |\nabla (F_D(\sigma) y - \phi)|^2 \right) dx \\ & \leq \lambda^{-1} \int_\Omega \nabla (F_D(\sigma + \vartheta) y - F_D(\sigma) y) \cdot \nabla (F_D(\sigma + \vartheta) y + F_D(\sigma) y - 2\phi) \, dx \\ & \leq C \|\nabla (F_D(\sigma + \vartheta) y - F_D(\sigma) y)\|_{L^2(\Omega)} \leq C' \|\vartheta\|_{L^q(\Omega')}. \end{aligned}$$

Therefore, $F_\phi(\cdot)$ is Lipschitz continuous on $\mathcal{A}$ with respect to the $L^q(\Omega')$-norm for $q \in \left(\frac{2Q}{Q-2}, \infty\right]$.

Finally, by Remark 3 F_ϕ is continuous on $\mathcal{A}$ with respect to the $L^q(\Omega')$ −norm for $q \geq 1$.

2) From Lemma 4, it implies that $F_\phi(\cdot)$ is Fréchet differentiable and

$$F'_\phi(\sigma)\vartheta = \int_\Omega \vartheta \left|\nabla\left(F_D(\sigma)y - \phi\right)\right|^2 dx + 2\int_\Omega \sigma\nabla\left(F_D(\sigma)y-\phi\right)\cdot\nabla F'_D(\sigma)\vartheta dx.$$

Since $F_D(\sigma)y - \phi \in H_0^1(\Omega)$ and (10), the last equation yields

$$\begin{aligned} F'_\phi(\sigma)\vartheta &= \int_\Omega \vartheta\left|\nabla\left(F_D(\sigma)y-\phi\right)\right|^2 dx - 2\int_\Omega \vartheta\nabla F_D(\sigma)y\cdot\nabla\left(F_D(\sigma)y-\phi\right)dx \\ &= -\int_\Omega \vartheta\left(\left|\nabla F_D(\sigma)y\right|^2 - \left|\nabla\phi\right|^2\right)dx. \end{aligned}$$

For $\vartheta \in L^\infty(\Omega')$ and extended by zero outside Ω', the second derivative of $F_\phi(\cdot)$ is given by

$$F''_\phi(\sigma)(\vartheta,\vartheta) = -2\int_\Omega \vartheta\nabla F_D(\sigma)y\cdot\nabla F'_D(\sigma)y(\vartheta)dx = 2\int_\Omega \sigma\left|\nabla F'_D(\sigma)y(\vartheta)\right|^2 dx \geq 0.$$

Therefore, $F_\phi(\cdot)$ is convex. Furthermore, by Lemma 4, it implies that $\|F''_\phi(\sigma)\|$ is uniformly bounded on $\mathcal{A}$. ■

Remark 7 *The uniform boundedness of $\|F''_\phi(\sigma)\|$ on $\mathcal{A}$ implies that $F'_\phi(\cdot)$ is Lipschitz continuous with respect to the L^q−norms with $q \in \left(\frac{2Q}{Q-2}, \infty\right]$.*

3 The Well-posedness

We now assume that there exists some $\sigma^* \in \mathcal{A}$ such that $\phi^* = F_D(\sigma^*)y$ and only noisy data $\phi^\delta \in H_0^1(\Omega)$ of ϕ^* such that

$$\left\|\phi^* - \phi^\delta\right\|_{H^1(\Omega)} \leq \delta$$

with $\delta > 0$ are given. Our problem is to reconstruct σ^* from ϕ^δ. Because of the ill-posedness of the problem and the assumption of sparsity of $\sigma^* - \sigma^0$, using sparsity regularization incorporated with the energy functional approach leads to considering the minimization problem (2).

Note that noisy data is normally assumed to be in $L^2(\Omega)$ and thus the assumption $\phi^\delta \in H_0^1(\Omega)$ seem to be not a real situation. In fact this assumption is not too tight since if noisy data is in $L^2(\Omega)$, we can obtain new noisy data in $H_0^1(\Omega)$ by using some methods of data processing. The assumption $\phi^\delta \in H_0^1(\Omega)$ is also used in some papers, e.g. in [18].

We now analyze the well-posedness of problem (2), which consists of the existence, stability and convergence. Note that the well-posedness of problem (2) is still valid if the sparsity promoting penalty (5) is replaced by any convex, lower semicontinuous functional in L^q. Because the set $\mathcal{A}$ is weakly compact.

Before proving the main results, we introduce some properties of the functional (5) and the notion of Φ-minimizing solution.

Lemma 8 *The functional Φ defined by (5) has the following properties*

1) Φ is non-negative, convex and weakly lower semi-continuous.

2) There exists a positive constant C such that for any $\sigma \in L^2(\Omega)$,

$$\Phi(\sigma) \geq \omega_{min} C^{p/2} \|\sigma\|^p.$$

This implies that Φ is weakly coercive, i.e. $\Phi(\sigma) \to \infty$ as $\|\sigma\| \to \infty$.

3) If $\{\sigma^n\}_{n\in\mathbb{N}} \subset L^2(\Omega)$ weakly converges to $\sigma \in L^2(\Omega)$ and $\Phi(\sigma^n)$ converges to $\Phi(\sigma)$, then $\Phi(\sigma^n - \sigma)$ converges to zero.

Proof. Φ is non-negative, convex and weakly lower semi-continuous because it is the sum of non-negative, convex and weakly continuous functionals. The proofs of 2) and 3) can be found in [17, Remark 3.] and [17, Lemma 2.], respectively. ■

Lemma 9 *The set*

$$\Pi(\phi^*) := \{\sigma \in \mathcal{A} : F_D(\sigma)\, y = \phi^*\}$$

is nonempty, convex, bounded and closed with respect to the $L^2(\Omega)$*-norm. Thus, there exists a solution* σ^+ *of the problem*

$$\min_{\sigma\in\Pi(\phi^*)} \Phi\left(\sigma - \sigma^0\right)$$

which is called a Φ*-minimizing solution of the diffusion coefficient identification problem. The* Φ*-minimizing solution is unique if* $p > 1$.

Proof. It is trivial that the set $\Pi(\phi^*)$ is nonempty, convex and bounded. The closeness of $\Pi(\phi^*)$ in the $L^2(\Omega)$ −norm is proven similarly as that of [18, Lemma 2.1].

We now prove that there exists at least a Φ-minimizing solution. Suppose that there does not exist a Φ-minimizing solution in $\Pi(\phi^*)$. There exists a sequence $\{\sigma^k\} \subset \Pi(\phi^*)$ such that $\Phi\left(\sigma^k - \sigma^0\right) \to c$ and

$$c < \Phi\left(\sigma - \sigma^0\right) \text{ for all } \sigma \in \Pi(\phi^*). \tag{16}$$

Since $\Pi(\phi^*)$ is weakly compact, there exists a subsequence of $\{\sigma^k\}$, denoted by $\{\sigma^k\}$ again, which weakly converges to $\tilde{\sigma} \in \Pi(\phi^*)$. From the weakly lower semi-continuity of Φ, it follows that

$$\Phi\left(\tilde{\sigma} - \sigma^0\right) \le \lim_{k\to\infty} \inf \Phi\left(\sigma^k - \sigma^0\right) = c.$$

This gives a contradiction to (16).

For $p > 1$, $\Phi(\cdot)$ is strictly convex and thus the Φ-minimizing solution is unique. ■

Theorem 10 (Existence) *Problem (2) has at least one solution.*

Proof. Since the functional $F_{\phi^\delta}(\cdot)$ is convex and continuous with respect to the $L^2(\Omega)$-norm, it is weakly lower semi-continuous. Besides, $\Phi(\cdot)$ is also convex and weakly lower semi-continuous with respect to the $L^2(\Omega)$-norm (see Lemma 8). Therefore, the objective functional of problem (2) is convex and weakly lower semi-continuous on $\mathcal{A}$. On the other hand, since $\mathcal{A}$ is nonempty, convex, bounded and closed with respect to the $L^2(\Omega)$-norm, it is weakly compact. Therefore, there exists at least one solution of (2). ■

Theorem 11 (Stability) *For a fixed regularization* $\alpha > 0$, *let the sequence* $\{\phi^n\}$ *converge to* ϕ^δ *in* $H_0^1(\Omega)$ *and*

$$\sigma^n \in \underset{\sigma\in\mathcal{A}}{\operatorname{argmin}}\, F_{\phi^n}(\sigma) + \alpha\Phi\left(\sigma - \sigma^0\right).$$

Then, there exist a subsequence $\{\sigma^{n_k}\}$ *of* $\{\sigma^n\}$ *and a minimizer* $\sigma^p_{\alpha,\delta}$ *of (2) such that*

$$\left\|\sigma^{n_k} - \sigma^p_{\alpha,\delta}\right\|_{L^2(\Omega)} \to 0.$$

In addition, if the minimizer $\sigma^p_{\alpha,\delta}$ *is unique, then the sequence* $\{\sigma^n\}$ *converges to* $\sigma^p_{\alpha,\delta}$ *with respect to the* $L^2(\Omega)$*-norm.*

Proof. By the definition of σ^n, we have

$$\begin{aligned} F_{\phi^n}(\sigma^n) + \alpha\Phi\left(\sigma^n - \sigma^0\right) &\le F_{\phi^n}(\sigma) + \alpha\Phi\left(\sigma - \sigma^0\right) \\ &\le \lambda^{-1}\left(\|F_D(\sigma)\, y\|^2_{H^1(\Omega)} + C\right) + \alpha\Phi\left(\sigma - \sigma^0\right) \end{aligned} \tag{17}$$

for any $\sigma \in \mathcal{A}$, where the constant C is independent of n such that $\|\phi^n\|^2_{H^1(\Omega)} \le C$ for all n. This follows that $\{\Phi\left(\sigma^n - \sigma^0\right)\}$ is bounded. Since Φ is weakly coercive in $L^2(\Omega)$ (see Lemma 8), the sequence $\{\sigma^n\}$ is also bounded in $L^2(\Omega)$. Therefore, there exist a subsequence of $\{\sigma^n\}$ denoted by $\{\sigma^{n_k}\}$ and an element $\sigma^p_{\alpha,\delta} \in L^2(\Omega)$ such that $\{\sigma^{n_k}\}$ weakly converges to $\sigma^p_{\alpha,\delta}$ in $L^2(\Omega)$. Since $\mathcal{A}$ is a convex closed set in $L^2(\Omega)$, $\sigma^p_{\alpha,\delta} \in \mathcal{A}$. On the other hand, since $F_{\phi^\delta}(\cdot)$ and $\Phi(\cdot)$ are weakly lower semi-continuous, we have

$$F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) \le \lim_k \inf F_{\phi^\delta}(\sigma^{n_k}) \tag{18}$$

and

$$\Phi\left(\sigma_{\alpha,\delta}^{p}-\sigma^{0}\right)\leq\liminf_{k}\Phi\left(\sigma^{n_k}-\sigma^{0}\right). \tag{19}$$

Furthermore, we have

$$F_{\phi^\delta}\left(\sigma^{n_k}\right)=F_{\phi^{n_k}}\left(\sigma^{n_k}\right)+\left(2\int_\Omega\sigma^{n_k}\nabla F_D\left(\sigma^{n_k}\right)y\cdot\nabla\left(\phi^{n_k}-\phi^\delta\right)dx\right.$$
$$\left.-\int_\Omega\sigma^{n_k}\left|\nabla\left(\phi^{n_k}-\phi^\delta\right)\right|^2dx\right). \tag{20}$$

Since $\phi^{n_k}\to\phi^\delta$ in $H^1(\Omega)$, the term in brackets on the right-hand side of (20) converges to zero as $k\to\infty$. Therefore,

$$\liminf_{k}F_{\phi^\delta}\left(\sigma^{n_k}\right)=\liminf_{k}F_{\phi^{n_k}}\left(\sigma^{n_k}\right),\ \limsup_{k}F_{\phi^\delta}\left(\sigma^{n_k}\right)=\limsup_{k}F_{\phi^{n_k}}\left(\sigma^{n_k}\right). \tag{21}$$

From (21), (17), (18) and (19), we obtain

$$\begin{aligned}F_{\phi^\delta}\left(\sigma_{\alpha,\delta}^{p}\right)+\alpha\Phi\left(\sigma_{\alpha,\delta}^{p}-\sigma^0\right)&\overset{(18),(19)}{\leq}\liminf_{k}F_{\phi^\delta}\left(\sigma^{n_k}\right)+\alpha\liminf_{k}\Phi\left(\sigma^{n_k}-\sigma^0\right)\\&\overset{(21)}{\leq}\liminf_{k}\left(F_{\phi^{n_k}}\left(\sigma^{n_k}\right)+\alpha\Phi\left(\sigma^{n_k}-\sigma^0\right)\right)\\&\leq\limsup_{k}\left(F_{\phi^{n_k}}\left(\sigma^{n_k}\right)+\alpha\Phi\left(\sigma^{n_k}-\sigma^0\right)\right)\\&\overset{(17)}{\leq}\limsup_{k}\left(F_{\phi^{n_k}}\left(\sigma\right)+\alpha\Phi\left(\sigma-\sigma^0\right)\right)\\&=F_{\phi^\delta}\left(\sigma\right)+\alpha\Phi\left(\sigma-\sigma^0\right)\end{aligned} \tag{22}$$

for all $\sigma\in\mathcal{A}$. It means that $\sigma_{\alpha,\delta}^{p}$ is a minimizer of (2).

From (22), setting $\sigma=\sigma_{\alpha,\delta}^{p}$, we get

$$\lim_{k}\left(F_{\phi^\delta}\left(\sigma^{n_k}\right)+\alpha\Phi\left(\sigma^{n_k}-\sigma^0\right)\right)=F_{\phi^\delta}\left(\sigma_{\alpha,\delta}^{p}\right)+\alpha\Phi\left(\sigma_{\alpha,\delta}^{p}-\sigma^0\right).$$

Together with (18) and (19), we deduce that $\Phi\left(\sigma^{n_k}-\sigma^0\right)\to\Phi\left(\sigma_{\alpha,\delta}^{p}-\sigma^0\right)$. Finally, since $\{\sigma^{n_k}\}$ weakly converges to $\sigma_{\alpha,\delta}^{p}$ and $\Phi\left(\sigma^{n_k}-\sigma^0\right)\to\Phi\left(\sigma_{\alpha,\delta}^{p}-\sigma^0\right)$ as $k\to\infty$, we conclude that $\Phi\left(\sigma^{n_k}-\sigma_{\alpha,\delta}^{p}\right)\to0$ as $k\to0$, and thus $\left\|\sigma^{n_k}-\sigma_{\alpha,\delta}^{p}\right\|_{L^2(\Omega)}\to0$ as $k\to\infty$ by Lemma 8.

In the case the minimizer $\sigma_{\alpha,\delta}^{p}$ is unique, the convergence of the original sequence $\{\sigma^n\}$ to $\sigma_{\alpha,\delta}^{p}$ follows by a subsequence argument. ∎

Theorem 12 (Convergence) *Assume that the operator equation $F_D(\sigma)y=\phi^*$ attains a solution in $\mathcal{A}$ and that $\alpha:\mathbb{R}_{>0}\to\mathbb{R}_{>0}$ satisfies*

$$\alpha(\delta)\to0\ \text{and}\ \frac{\delta^2}{\alpha(\delta)}\to0\ \text{as}\ \delta\to0.$$

Let $\delta_n\to0$ and $\|\phi^n-\phi^*\|_{H^1(\Omega)}\leq\delta_n$. Moreover, let $\alpha_n=\alpha(\delta_n)$ and

$$\sigma^n\in\operatorname*{argmin}_{\sigma\in\mathcal{A}}F_{\phi^n}(\sigma)+\alpha_n\Phi\left(\sigma-\sigma^0\right).$$

Then, there exist a Φ-minimizing solution σ^+ of $F_D(\sigma)y=\phi^*$ and a subsequence of $\{\sigma^n\}$ converging to σ^+ on $\mathcal{A}$ with respect to the $L^2(\Omega)-$norm.

Proof. Let $\tilde\sigma\in\mathcal{A}$ be a solution of $F_D(\sigma)y=\phi^*$. The definition of σ^n implies that

$$\begin{aligned}F_{\phi^n}\left(\sigma^n\right)+\alpha_n\Phi\left(\sigma^n-\sigma^0\right)&\leq F_{\phi^n}\left(\tilde\sigma\right)+\alpha_n\Phi\left(\tilde\sigma-\sigma^0\right)\\&\leq\frac{1}{\lambda}\int_\Omega\left|\nabla\left(F_D\left(\tilde\sigma\right)y-\phi^n\right)\right|^2+\alpha_n\Phi\left(\tilde\sigma-\sigma^0\right)\\&\leq\frac{1}{\lambda}\left\|\phi^*-\phi^n\right\|_{H^1(\Omega)}^2+\alpha_n\Phi\left(\tilde\sigma-\sigma^0\right)\\&\leq\frac{1}{\lambda}\delta_n^2+\alpha_n\Phi\left(\tilde\sigma-\sigma^0\right).\end{aligned} \tag{23}$$

In particular, when $\delta \to 0$ and $\alpha \sim \delta^2$, it follows that

$$F_{\phi^n}(\sigma^n) \to 0 \text{ and } \limsup_n \Phi\left(\sigma^n - \sigma^0\right) \le \Phi\left(\tilde{\sigma} - \sigma^0\right). \tag{24}$$

This implies that $\{\Phi\left(\sigma^n - \sigma^0\right)\}$ is bounded. Since $\Phi(\cdot)$ is weakly coercive, $\{\sigma^n\}$ is bounded, too. Therefore, there exist a subsequence $\{\sigma^{n_k}\}$ of $\{\sigma^n\}$ and $\sigma^+ \in \mathcal{A}$ such that σ^{n_k} weakly converges to σ^+. From (24), we deduce

$$\begin{aligned} F_{\phi^*}(\sigma^{n_k}) &= \int_\Omega \sigma^{n_k} \left|\nabla\left(F_D(\sigma^{n_k})\, y - \phi^*\right)\right|^2 \\ &\le \int_\Omega \sigma^{n_k} \left|\nabla\left(F_D(\sigma^{n_k})\, y - \phi^{n_k}\right)\right|^2 + \int_\Omega \sigma^{n_k} \left|\nabla\left(\phi^{n_k} - \phi^*\right)\right|^2 \\ &\le F_{\phi^{n_k}}(\sigma^{n_k}) + \lambda^{-1} \left\|\phi^{n_k} - \phi^*\right\|^2_{H^1(\Omega)} \to 0 \ (k \to \infty). \end{aligned}$$

Since $F_{\phi^*}(\cdot)$ is weakly lower semi-continuous,

$$0 \le F_{\phi^*}\left(\sigma^+\right) \le \liminf_k F_{\phi^*}(\sigma^{n_k}) = 0.$$

Thus, $F_{\phi^*}(\sigma^+) = 0$. It implies that $\|F_D(\sigma^+)\, y - \phi^*\|_{H^1(\Omega)} = 0$. Hence σ^+ is a solution of the equation $F_D(\sigma)\, y = \phi^*$.

Moreover, since $\Phi(\cdot)$ is weakly lower semi-continuous in $L^2(\Omega)$, by using (24) we get

$$\Phi\left(\sigma^+ - \sigma^0\right) \le \liminf_k \Phi\left(\sigma^{n_k} - \sigma^0\right) \le \limsup_k \Phi\left(\sigma^{n_k} - \sigma^0\right) \le \Phi\left(\tilde{\sigma} - \sigma^0\right). \tag{25}$$

It implies that σ^+ is a Φ-minimizing solution. Finally, choosing $\tilde{\sigma} = \sigma^+$ in (25), we have $\Phi\left(\sigma^{n_k} - \sigma^0\right) \to \Phi\left(\sigma^+ - \sigma^0\right)$ as $k \to \infty$. Since $\{\sigma^{n_k} - \sigma^0\}$ weakly converges to $\sigma^+ - \sigma^0$ in $L^2(\Omega)$ and $\Phi\left(\sigma^{n_k} - \sigma^0\right) \to \Phi\left(\sigma^+ - \sigma^0\right)$ as $k \to \infty$, $\Phi(\sigma^{n_k} - \sigma^+) \to 0$ as $k \to 0$ and thus $\|\sigma^{n_k} - \sigma^+\|_{L^2(\Omega)} \to 0$.

In the case the minimizer σ^+ is unique, the convergence of the original sequence $\{\sigma^n - \sigma^0\}$ to $\sigma^+ - \sigma^0$ follows by a subsequence argument. ■

4 Convergence Rates

As shown before, for $\sigma \in \mathcal{A}$, the operator

$$F'_D(\sigma)\, y(\cdot) : L^q(\Omega') \to H^1_0(\Omega) \text{ with } q \in \left(\frac{2Q}{Q-2}, \infty\right]$$

is continuous and linear. Denote by

$$\left(F'_D(\sigma)\, y\right)^*(\cdot) : H^{-1}(\Omega) = \left(H^1_0(\Omega)\right)^* \to L^{q_1}(\Omega') \text{ with } \frac{1}{q} + \frac{1}{q_1} = 1,$$

the dual operator of $F'_D(\sigma)\, y$. Then,

$$\left\langle \left(F'_D(\sigma)\, y\right)^*(w^*), \vartheta \right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))} = \left\langle w^*, F'_D(\sigma)\, y(\vartheta) \right\rangle_{\left(H^{-1}(\Omega), H^1_0(\Omega)\right)}. \tag{26}$$

To obtain convergence rates of sparsity regularization, an important tool is the Bregman distance relating to a proper convex functional. We briefly introduce this notion here. For a detail discussion on the Bregman distance, we refer to [5, 10, 25, 7].

Let X be a Banach space with its dual space X^* and $R : X \to (-\infty, +\infty]$ be a proper convex functional with $dom(R) := \{x \in X : R(x) < +\infty\} \ne \emptyset$. The subdifferential of R at $x \in dom(R)$ is defined by

$$\partial R(x) := \{x^* \in X^* : R(y) \ge R(x) + \langle x^*, y - x\rangle_{(X^*, X)} \text{ for all } y \in X\}.$$

Then, for a fixed element $x^* \in \partial R(x)$, the expression

$$D^R_{x^*}(y, x) := R(y) - R(x) - \langle x^*, y - x\rangle_{(X^*, X)}$$

is called the *Bregman distance of two elements* $y, x \in X$ *with respect to* R *and* x^*. In the following, we denote $D_{x^*}(y,x)$ instead of $D^R_{x^*}(y,x)$ for simplicity.

Since $\partial R(x)$ might be empty or multi-valued, Bregman distance might be not defined or multi-valued. However, for a continuously differentiable functional, there is a unique element in the subdifferential and consequently, a unique Bregman distance. In this case, the distance is just the difference at the point y between $R(\cdot)$ and the first order Taylor series approximation to $R(\cdot)$ at x. Furthermore, if $R(y)$ is strictly convex, $D_{x^*}(y,x)$ is also strictly convex in y for each fixed x, and therefore $D_{x^*}(y,x)=0$ if and only if $y = x$.

Note that $D_{x^*}(y,x)$ is not a distance in the usual metric sense since, in general, $D(y,x) \neq D(x,y)$ and the triangle inequality does not hold. However, it is a measurement of closeness in the sense that $D_{x^*}(y,x) \geq 0$ and $D_{x^*}(y,x) = 0$ if $y = x$.

Convergence rates of sparsity regularization are given in the following theorem.

Theorem 13 *For* $q \in \left(\frac{2Q}{Q-2}, \infty\right]$, $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$ *and* $y \in L^r(\Omega)$. *Assume that* $\left\|\phi^\delta - \phi^*\right\|_{H^1(\Omega)} \leq \delta$ *and* $\sigma^p_{\alpha,\delta}$ *is a solution of (2). Moreover, assume that there exists a function* $w^* \in H^{-1}(\Omega)$ *such that*

$$\xi := \left(F_D'(\sigma^+)\, y\right)^* (w^*) \in \partial\Phi\left(\sigma^+ - \sigma^0\right). \tag{27}$$

Then,

$$D_\xi\left(\sigma^p_{\alpha,\delta}, \sigma^+\right) = O(\delta) \text{ and } \left\|F_D\left(\sigma^p_{\alpha,\delta}\right) y - \phi^\delta\right\|_{H^1(\Omega)} = O(\delta)$$

as $\delta \to 0$ *and* $\alpha \sim \delta$. *In particular, for* $p \in (1,2]$, *we have*

$$\left\|\sigma^p_{\alpha,\delta} - \sigma^+\right\|_{L^2(\Omega)} = O\left(\sqrt{\delta}\right).$$

Proof. The proof follows the ideas of Hao and Quyen in [18, 19]. By the definition of $\sigma^p_{\alpha,\delta}$, we get

$$F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) + \alpha\Phi\left(\sigma^p_{\alpha,\delta} - \sigma^0\right) \leq F_{\phi^\delta}\left(\sigma^+\right) + \alpha\Phi\left(\sigma^+ - \sigma^0\right). \tag{28}$$

Then, we have

$$\begin{aligned}
&F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) + \alpha D_\xi\left(\sigma^p_{\alpha,\delta}, \sigma^+\right)\\
&= F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) + \alpha\left(\Phi\left(\sigma^p_{\alpha,\delta} - \sigma^0\right) - \Phi\left(\sigma^+ - \sigma^0\right) - \left\langle \xi, \sigma^p_{\alpha,\delta} - \sigma^+\right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))}\right)\\
&\leq F_{\phi^\delta}\left(\sigma^+\right) - \alpha\left\langle \xi, \sigma^p_{\alpha,\delta} - \sigma^+\right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))}\\
&\leq \frac{1}{\lambda}\delta^2 - \alpha\left\langle \xi, \sigma^p_{\alpha,\delta} - \sigma^+\right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))}.
\end{aligned} \tag{29}$$

From (26) and (27), we get

$$\left\langle \xi, \sigma^p_{\alpha,\delta} - \sigma^+\right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))} = \left\langle w^*, F_D'\left(\sigma^+\right) y\left(\sigma^p_{\alpha,\delta} - \sigma^+\right)\right\rangle_{\left(H^{-1}(\Omega), H^1_0(\Omega)\right)}. \tag{30}$$

By Riesz's representation theorem, there exists an element $w \in H^1_0(\Omega)$ such that

$$\left\langle w^*, F_D'\left(\sigma^+\right) y\left(\sigma^p_{\alpha,\delta} - \sigma^+\right)\right\rangle_{\left(H^{-1}(\Omega), H^1_0(\Omega)\right)} = \left\langle w, F_D'\left(\sigma^+\right) y\left(\sigma^p_{\alpha,\delta} - \sigma^+\right)\right\rangle_{H^1_0(\Omega)}. \tag{31}$$

Since $\sigma^+ \geq \lambda > 0$, the scalar product

$$[\phi, v]_{H^1_0(\Omega)} := \int_\Omega \sigma^+ \nabla\phi \cdot \nabla v dx, \text{ for all } \phi, v \in H^1_0(\Omega)$$

is equivalent to $\langle \phi, v\rangle_{H^1_0(\Omega)}$ on $H^1_0(\Omega)$. Therefore, there exists an element $\hat{w} \in H^1_0(\Omega)$ independent of $\sigma^p_{\alpha,\delta}$ such that

$$\left\langle w, F_D'\left(\sigma^+\right) y\left(\sigma^p_{\alpha,\delta} - \sigma^+\right)\right\rangle_{H^1_0(\Omega)} = \int_\Omega \sigma^+ \nabla\hat{w} \cdot \nabla F_D'\left(\sigma^+\right) y\left(\sigma^p_{\alpha,\delta} - \sigma^+\right) dx. \tag{32}$$

From (30), (31) and (32), we have

$$\left\langle \xi, \sigma^p_{\alpha,\delta} - \sigma^+ \right\rangle_{(L^{q_1}(\Omega'), L^q(\Omega'))} = \int_\Omega \sigma^+ \nabla \hat{w} \cdot \nabla F'_D\left(\sigma^+\right) y \left(\sigma^p_{\alpha,\delta} - \sigma^+\right) dx =: \Lambda.$$

From the weak solution formulas of $F_D\left(\sigma^+\right) y$ and $F'_D\left(\sigma^+\right) y \left(\sigma^p_{\alpha,\delta} - \sigma^+\right)$ (see (6) and (10)), we deduce

$$\begin{aligned}
\alpha\Lambda &= \alpha \int_\Omega \sigma^+ \nabla \hat{w} \cdot \nabla F'_D\left(\sigma^+\right) y \left(\sigma^p_{\alpha,\delta} - \sigma^+\right) dx \\
&= -\alpha \int_\Omega \left(\sigma^p_{\alpha,\delta} - \sigma^+\right) \nabla \hat{w} \cdot \nabla F_D\left(\sigma^+\right) y dx \\
&= \alpha \int_\Omega \sigma^+ \nabla \hat{w} \cdot \nabla F_D\left(\sigma^+\right) y dx - \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla F_D\left(\sigma^+\right) y dx \\
&= \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla F_D\left(\sigma^p_{\alpha,\delta}\right) y dx - \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla F_D\left(\sigma^+\right) y dx \\
&= \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla \left(F_D\left(\sigma^p_{\alpha,\delta}\right) y - F_D\left(\sigma^+\right) y\right) dx \\
&= \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla \left(F_D\left(\sigma^p_{\alpha,\delta}\right) y - \phi^\delta\right) dx + \alpha \int_\Omega \sigma^p_{\alpha,\delta} \nabla \hat{w} \cdot \nabla \left(\phi^\delta - \phi^*\right) dx.
\end{aligned}$$

Using the Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}
\alpha\left|\Lambda\right| &\le \alpha \left(\int_\Omega \sigma^p_{\alpha,\delta} \left|\nabla \hat{w}\right|^2 dx\right)^{1/2} \left(\int_\Omega \sigma^p_{\alpha,\delta} \left|\nabla \left(F_D\left(\sigma^p_{\alpha,\delta}\right) y - \phi^\delta\right)\right|^2 dx\right)^{1/2} \\
&\quad + \alpha \left(\int_\Omega \left(\sigma^p_{\alpha,\delta}\right)^2 \left|\nabla \hat{w}\right|^2 dx\right)^{1/2} \left(\int_\Omega \left|\nabla \left(\phi^\delta - \phi^*\right)\right|^2 dx\right)^{1/2} \\
&\le \alpha \left(\frac{1}{\lambda} \int_\Omega \left|\nabla \hat{w}\right|^2 dx\right)^{1/2} \left(J_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right)\right)^{1/2} + \frac{\alpha}{\lambda} \left(\int_\Omega \left|\nabla \hat{w}\right|^2 dx\right)^{1/2} \left\|\phi^\delta - \phi^*\right\|_{H^1(\Omega)} \\
&\le \frac{\alpha^2}{2\lambda} \int_\Omega \left|\nabla \hat{w}\right|^2 dx + \frac{1}{2} F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) + \frac{\alpha\delta}{\lambda} \left(\int_\Omega \left|\nabla \hat{w}\right|^2 dx\right)^{1/2}.
\end{aligned} \tag{33}$$

Here, we used the inequality $ab \le \frac{\alpha a^2}{2} + \frac{b^2}{2\alpha}$ for the first term. Together with (29), we deduce

$$\frac{1}{2} F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) + \alpha D_\xi\left(\sigma^p_{\alpha,\delta}, \sigma^+\right) \le \frac{1}{\lambda}\delta^2 + \frac{\alpha^2}{2\lambda} C_1^2 + \frac{\alpha\delta}{\lambda} C_1, \tag{34}$$

with $C_1 = \left(\int_\Omega \left|\nabla \hat{w}\right|^2 dx\right)^{1/2}$. This inequality implies that

$$D_\xi\left(\sigma^p_{\alpha,\delta}, \sigma^+\right) = O(\delta) \text{ as } \alpha \to 0 \text{ and } \alpha \sim \delta.$$

By (8) and (34), we have

$$\left\|F_D\left(\sigma^p_{\alpha,\delta}\right) y - \phi^\delta\right\|^2_{H^1(\Omega)} \le \frac{1}{C} F_{\phi^\delta}\left(\sigma^p_{\alpha,\delta}\right) = O\left(\delta^2\right) \text{ as } \delta \to 0 \text{ and } \alpha \sim \delta.$$

In particular, for $p \in (1, 2]$ there exists a constant $C_p > 0$ such that $D_\xi\left(\sigma^p_{\alpha,\delta}, \sigma^+\right) \ge C_p \left\|\sigma^p_{\alpha,\delta} - \sigma^+\right\|^2_{L^2(\Omega)}$, see [17, Lemma 10]. Therefore, we have

$$\left\|\sigma^p_{\alpha,\delta} - \sigma^+\right\|_{L^2(\Omega)} = O\left(\sqrt{\delta}\right).$$

■

Remark 14 *Our source condition is very simple if we compare with the source conditions in [22, 17, 13, 32]. Especially, we do not need the smallness requirement in the source condition.*

5 Conclusion

In this paper, sparsity regularization incorporated with the energy functional approach was analyzed for the diffusion coefficient identification problem. The regularized problem was proven to be well-posed and convergence rates of the method was obtained under a simple source condition. An advantage of the new approach is to work with a convex energy functional. Another advantage is that the source condition of obtaining convergence rates are very simple.

6 Acknowledgement

The author would like to thank all three implicit referees for their very valuable comments. I have learned a lots from these comments and made the paper clearer.

References

[1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 10:1217–1229, 1996.

[2] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.

[3] T. Bonesky, K. Bredies, D. A. Lorenz, and P. Maass. A generalized conditional gradient method for nonlinear operator equations with sparsity constraints. *Inverse Problems*, 33:2041–2058, 2007.

[4] K. Bredies, D. A. Lorenz, and P. Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *Computational Optimization and Application*, 42(2):173–193, 2009.

[5] L. M. Bregman. The relaxation of finding the common points of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys*, 7:200–217, 1967.

[6] M. Burger, J. Flemming, and B. Hofmann. Convergence rates in ℓ^1-regularization if the sparsity assumption fails. *Inverse Problems*, 29(2):025013, 2013.

[7] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20:1411–1421, 2004.

[8] T. F. Chan and X. Tai. Identification of discontinuous coefficients in ellptic problems using total variation regularization. *SIAM J. Sci. Comput.*, 25(3):881–904, 2003.

[9] T. F. Chan and X. Tai. Level set and total variation regularization for elliptic inverse problems with discontinuous coefficients. *Journal of Computational Physics*, 193:40–66, 2003.

[10] G. Chen and M. Teboulle. Convergence analysis of a proximal-like minimization algorithm using Bregman functions. *SIAM J. Optim.*, 3(3):538–543, 1993.

[11] Z. Chen and J. Zou. An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. *SIAM Journal of Control and Optimization*, 37:892–910, 1999.

[12] I. Daubechies, M. Defrise, and C. Demol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1541, 2004.

[13] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[14] R. Falk. Error estimates for the numerical identification of a variable coefficient. *Mathematics of Computation*, 40(3):537–546, 1983.

[15] J. Flemming. *Generalized Tikhonov regularization Basic theory and comprehensive results on convergence rates.* PhD thesis, PhD thesis, TU Chemnitz, 2011.

[16] M. Gehre, T. Kluth, A. Lipponen, B. Jin, A. Seppänen, J. P Kaipio, and P. Maass. Sparsity reconstruction in electrical impedance tomography: an experimental evaluation. *Journal of Computational and Applied Mathematics*, 236(8):2126–2136, 2012.

[17] M. Grasmair, M. Haltmeier, and O. Scherer. Sparsity regularization with l^q penalty term. *Inverse Problems*, 24:055020, 2008.

[18] D. N. Hào and T. N. T. Quyen. Convergence rates for Tikhonov regularization of coefficient identification problems in Laplace-type equation. *Inverse Problems*, 26:125014, 2010.

[19] D. N. Hào and T. N. T. Quyen. Convergence rates for total variation regularization of coefficient identification problems in elliptic equations I. *Inverse Problems*, 27:075008, 2011.

[20] D. N. Hào and T. N. T. Quyen. Convergence rates for Tikhonov regularization of a two-coefficient identification problem in an elliptic boundary value problem. *Numerische Mathematik*, 120(1):45–77, 2012.

[21] D. N. Hào and T. N. T. Quyen. Convergence rates for total variation regularization of coefficient identification problems in elliptic equations ii. *Journal of Mathematical Analysis and Applications*, 388(1):593–616, 2012.

[22] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23:987–1010, 2007.

[23] B. Jin, T. Khan, and P. Maass. A reconstruction algorithm for electrical impedance tomography based on sparsity regularization. *International Journal for Numerical Methods in Engineering*, 89(3):337–353, 2012.

[24] B. Jin and P. Maass. An analysis of electrical impedance tomography with applications to tikhonov regularization. *ESAIM: Control, Optimisation and Calculus of Variations*, 1(1), 2011.

[25] K. C. Kiwiel. Proximal minimization methods with generalized Bregman functions. *SIAM J. Control Optim.*, 35:1142–1168, 1997.

[26] I. Knowles. Parameter identification for elliptic problems. *Journal of Computational and Applied Mathematics*, 131:175–194, 2001.

[27] R. V. Kohn and B. Lowe. A variational method for parameter identification. *Mathematical Modeling and Numerical Analysis*, 22:119–158, 1988.

[28] C. Kravaris and J. H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM Journal of Control and Optimization*, 23:217–241, 1985.

[29] D. A. Lorenz, P. Maass, and P. Q. Muoi. Gradient descent methods based on quadratic approximations of Tikhonov functionals with sparsity constraints: theory and numerical comparison of stepsize rules. *Electronic Transactions on Numerical Analysis*, 39:437–463, 2012.

[30] N. G. Meyers. An L^p- estimate for the gradient of the solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa (3)*, 17:189–206, 1963.

[31] P. Q. Muoi, D. N. Hào, P. Maass, and M. Pidcock. Semismooth Newton and quasi-Newton methods in weighted l^1-regularization. *Journal of Inverse and Ill-Posed Problems*, 21(5):665–693, 2013.

[32] A. Neubauer. On enhanced convergence rates for Tikhonov regularization of nonlinear ill-posed problems in Banach spaces. *Inverse Problems*, 25:065009, 1989.

[33] R. Ramlau and G. Teschke. A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints. *Numer. Math.*, 104:177–203, 2006.

[34] G. R. Richter. Numerical identification of a spatially varying diffusion coefficient. *Mathematics of Computation*, 36:375–386, 1981.

[35] L. Rondi and F. Santosa. Enhanced electrical impedance tomography via the Mumford-Shah functional. *ESAIM Control Optim. Calc. Var.*, 6:517–538, 2001.

[36] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk. A new compressive imaging camera architecture using optical-domain compression. In *in Proc. of Computational Imaging IV at SPIE Electronic Imaging*, pages 43–52, 2006.

[37] F. Werner and T. Hohage. Convergence rates in expectation for Tikhonov-type regularization of inverse problems with poisson data. *Inverse Problems*, 28(10):104004, 2012.

[38] L. W. White and J. Zhou. Continuity and uniqueness of regularized output least squares optimal estimators. *Journal of Mathematical Analysis and Applications*, 196:53–83, 1994.

[39] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing. *SIAM Journal on Imaging Sciences*, 1(1):143–168, 2008.

[40] J. Zou. Numerical methods for elliptic inverse problems. *International Journal of Computer Mathematics*, 70:211–232, 1998.