



ON THE EXISTENCE OF A VARIATIONAL REGULARIZATION PARAMETER UNDER MOROZOV'S DISCREPANCY PRINCIPLE FOR NONLINEAR INVERSE PROBLEMS

LIANG DING^{*1}, LONG LI¹, WEIMIN HAN² AND WEI WANG³

¹Department of Mathematics, Northeast Forestry University, China

²Department of Mathematics, University of Iowa, USA

³College of Data Science, Jiaying University, China

(Communicated by Shuai Lu)

ABSTRACT. Morozov's discrepancy principle is commonly adopted in Tikhonov regularization for choosing the regularization parameter. Nevertheless, for a general nonlinear inverse problem, the discrepancy $\|F(x_\alpha^\delta) - y^\delta\|_Y$ does not depend continuously on α and it is questionable whether there exists a regularization parameter α such that $\tau_1 \delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2 \delta$ ($1 \leq \tau_1 < \tau_2$). In this paper, we prove the existence of α under Morozov's discrepancy principle if $\tau_2 \geq (3 + 2\gamma)\tau_1$ or $\tau_1 \leq \tau_2/(3 + 2\gamma)$, where $\gamma > 0$ is a parameter in a tangential cone condition for the nonlinear operator F . Furthermore, we present results on the convergence of the regularized solutions under Morozov's discrepancy principle. Numerical results are reported on the efficiency of the proposed approach.

1. Introduction. In this paper, we consider to solve an operator equation of the form

$$F(x) = y, \quad (1)$$

where $F : X \rightarrow Y$ is a nonlinear operator between two reflexive Banach spaces X and Y . We assume $y^\delta \in Y$ is known with $\|y^\delta - y\|_Y \leq \delta$ for a given $\delta \geq 0$. The commonly adopted technique to solve problem (1) is the Tikhonov regularization of the form

$$\min_{x \in X} \left\{ \mathcal{J}_\alpha^\delta(x) = \frac{1}{2} \|F(x) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x) \right\} \quad (2)$$

with $\alpha > 0$ and $\mathcal{R}$ is a convex penalty term, cf. the monographs [20, 36, 39] and the special issues [7, 14, 28, 29] for many developments on regularization properties and minimization schemes.

In practice, it is crucial to choose an appropriate regularization parameter α for problem (2). Morozov's discrepancy principle (MDP) is the most commonly

2020 *Mathematics Subject Classification.* Primary:49J27, 49J52; Secondary: 65J15, 65J20.

Key words and phrases. Non-linear inverse problem, Morozov's discrepancy principle, tangential cone condition, convex penalty.

The first author is supported by the Fundamental Research Funds for the Central Universities (no. 2572021DJ03).

*Corresponding author: Liang Ding.

adopted technique to determine the regularization parameter $\alpha = \alpha(\delta, y^\delta) > 0$ such that

$$\tau_1 \delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2 \delta, \quad (3)$$

where $1 \leq \tau_1 < \tau_2$ and x_α^δ is a minimizer of $\mathcal{J}_\alpha^\delta(x)$ in (2). If the minimizer of $\mathcal{J}_\alpha^\delta(x)$ is unique, there exists a regularization parameter α such that (3) holds ([1, 3, 41]). However, due to the nonlinearity of F , there may exist multiple minimizers of $\mathcal{J}_\alpha^\delta(x)$, the discrepancy $\|F(x_\alpha^\delta) - y^\delta\|_Y$ does not exhibit continuous dependence on α and one can not ensure the existence of α such that (3) holds ([37, 41]). Indeed, a weakness of MDP lies in the fact that a regularization parameter α satisfying (3) might not exist for a general nonlinear operator equation. In some applications of nonlinear inverse problems, one has to assume the existence of α such that (3) holds. Few results are available on the existence of α under MDP.

1.1. Related works. The first theoretical analysis on MDP for Tikhonov regularization dates back to 1966 ([35]). Several numerical algorithms have been proposed to compute the regularization parameter α for the classical quadratic Tikhonov regularization ([21, 33, 34]). Subsequently, MDP is extended to general convex regularizations. If $\mathcal{R}(x)$ is convex, then $\|F(x_\alpha^\delta) - y^\delta\|_Y$ depends continuously on α , and the existence of α follows. In [9], MDP is applied to the Tikhonov regularization with convex penalty terms more general than the classical quadratic one. In [30], two iterative parameter choice methods by MDP are proposed for non-smooth Tikhonov regularization with general convex penalty terms.

We note that the results cited in the previous paragraph are limited to linear ill-posed problems. Due to the existence of possibly multiple minimizers, the techniques for linear ill-posed problems can not be extended to nonlinear ill-posed operator equations directly. Special regularization techniques are needed to explore the existence of the regularization parameter α which is determined by MDP. In [37], for the classical quadratic Tikhonov regularization, it is shown that if $\tau_1 = 1$ and

$$\tau_2 > 1 + \frac{L + 2\|F'(x^\dagger)\|}{L} \max \left\{ \frac{4}{Q^2}, 1 \right\},$$

the existence of α can be guaranteed, where $x^\dagger$ is an $\mathcal{R}$ -minimum solution of the equation (1), and the parameters L and Q depend on the source condition and the lower bound of $x^\dagger$. In [3], for a general convex penalty term, the existence of α is proved under the following condition: there is no $\alpha > 0$ with minimizers $x_\alpha^{\delta,1}, x_\alpha^{\delta,2}$ such that

$$\|F(x_\alpha^{\delta,1}) - y^\delta\|_Y < \tau_1 \delta < \tau_2 \delta < \|F(x_\alpha^{\delta,2}) - y^\delta\|_Y.$$

In [4], convergence rates are investigated via variational inequalities, where the regularization parameter is determined by MDP. A relation between uniqueness of minimizers in Tikhonov-type regularization and Morozov-like discrepancy principles is shown in [1]. In [17], the convex function $\mathcal{R}(x)$ is replaced by a non-convex penalty term $\alpha \ell_1(x) - \beta \ell_2(x)$, and the existence of the regularization parameter α is demonstrated. It is proven that there exists at least one α such that

$$\delta \leq \|F(x_{\alpha,\beta}^\delta) - y^\delta\|_Y \leq (\max \{ \tau^2 \delta^2, (3 + 2\gamma) \delta^2 + \delta(2 + 2\gamma) \|y^\delta - F(0)\|_Y \})^{\frac{1}{2}}. \quad (4)$$

However, (4) does not give the same order for the lower and the upper bound. The upper bound is only $O(\sqrt{\delta})$.

1.2. Contribution and organization. In this paper, we assume $\mathcal{R}$ is convex and the operator F satisfies a tangential cone condition: there exists a constant $\gamma > 0$ such that for arbitrary $x_1, x_2 \in \mathcal{L}_\alpha^\delta$,

$$\|F(x_2) - F(x_1) - F'(x_1)(x_2 - x_1)\|_Y \leq \gamma \|F(x_2) - F(x_1)\|_Y,$$

where $\mathcal{L}_\alpha^\delta$ represents the set of all minimizers of the functional $\mathcal{J}_\alpha^\delta(x)$ defined in (2). We will prove that there exists at least one α such that

$$\tau_1 \delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq c\delta, \quad (5)$$

where $c := \max\{\tau_2, \tau_1(3 + 2\gamma)\}$ or

$$c\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2 \delta, \quad (6)$$

where $c := \min\{\tau_1, \tau_2/(3 + 2\gamma)\}$. This implies that if $\tau_2 \geq \tau_1(3 + 2\gamma)$ or $\tau_1 \leq \tau_2/(3 + 2\gamma)$, the existence of α can be guaranteed. Based on (5) and (6), we will investigate the well-posedness and convergence rate of the regularized solution of the problem (2). We will show that with the parameter choice rule (5) and (6), $\alpha \rightarrow 0$ as the noise level $\delta \rightarrow 0$.

For the traditional MDP for linear ill-posed problems, the existence of α is guaranteed. However, for nonlinear ill-posed problems, the discrepancy term $\|F(x_\alpha^\delta) - y^\delta\|_Y$ is not continuous with respect to α . So it is challenging to choose appropriate values of τ_1 and τ_2 to ensure the existence of α . We will prove the modified MDP (5) can ensure the existence of α and we can apply Algorithm 1 to determine the value of α . We will present results from two numerical experiments to show that the upper bound $c\delta$ in (5) is actually needed, cf. Fig. 6 and Fig. 10.

Furthermore, it is quite natural to extend the current work to the principle of sequence differences [2, 24]. In the future, we will combine the stochastic regularization technique [45] to extend the above results to a stochastic setting.

An outline of the rest of this paper is as follows. In the next section we introduce the notation and review some results on the Tikhonov regularization with a general convex penalty term. In Section 3, we investigate the existence of α determined by MDP (5). In Section 4, we present results on the well-posedness and the convergence rate of the regularized solution under MDP (5). In addition, we show that MDP (5) can guarantee $\alpha(\delta, y^\delta) \rightarrow 0$ as $\delta \rightarrow 0$. Finally, numerical results are presented in Section 5 on a nonlinear compressive sensing problem and a nonlinear one-dimensional gravity inversion problem.

2. Preliminaries. We first briefly introduce some notation for the Tikhonov regularization. Denote by

$$x_\alpha^\delta \in \arg \min_{x \in X} \left\{ \frac{1}{2} \|F(x) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x) \right\} \quad (7)$$

a minimizer of the regularization function $\mathcal{J}_\alpha^\delta(x)$ defined in (2). Let $\mathcal{L}_\alpha^\delta$ be the set of all such minimizers x_α^δ . An element $x^\dagger \in X$ is called an $\mathcal{R}$ -minimum solution of the equation (1) if

$$F(x^\dagger) = y \text{ and } \mathcal{R}(x^\dagger) = \min_{x \in X} \{\mathcal{R}(x) \mid F(x) = y\}.$$

Throughout this paper, we will assume the operator F and the data y^δ have the following properties.

Condition 2.1. (i) $F : X \rightarrow Y$ is Fréchet differential.

(ii) $F : X \rightarrow Y$ is weakly sequentially closed, i.e., $x_n \rightharpoonup x$ in X and $F(x_n) \rightharpoonup y$ in Y imply that $F(x) = y$.

(iii) There exists a constant $\gamma > 0$ such that

$$\|F(x_2) - F(x_1) - F'(x_1)(x_2 - x_1)\|_Y \leq \gamma \|F(x_2) - F(x_1)\|_Y \quad \forall x_1, x_2 \in \mathcal{L}_\alpha^\delta. \quad (8)$$

(iv) There exist $\delta > 0$ and $\tau_2 > \tau_1 \geq 1$ such that

$$\|y - y^\delta\|_Y \leq \tau_1 \delta < \tau_2 \delta \leq \|F(0) - y^\delta\|_Y. \quad (9)$$

The condition (8) was used in [19, pp. 278–279], [31, p. 6], [40, pp. 69–70]. It follows from (8) that

$$\|F'(x_1)(x_2 - x_1)\|_Y \leq (1 + \gamma) \|F(x_2) - F(x_1)\|_Y,$$

which has been adopted by several researchers. In (9), we require $\|F(0) - y^\delta\|_Y \geq \tau \delta$, which is a reasonable assumption. Indeed, if $\|F(0) - y^\delta\|_Y < \tau \delta$, then 0 can be viewed as a good approximation to the regularized solution x_α^δ . Moreover, practically, it is almost impossible to recover a solution from an observed data of a size in the same order as the noise.

Next, we assume the convex penalty term $\mathcal{R}$ has the following properties.

Condition 2.2. (i) $\mathcal{R}(x)$ is proper, and $\mathcal{R}(x) = 0$ if and only if $x = 0$.

(ii) $\mathcal{R}(x)$ satisfy the one-homogeneity.

(iii) (Coercivity) $\|x\|_X \rightarrow \infty$ implies $\mathcal{R}(x) \rightarrow \infty$.

(iv) (Weak lower semi-continuity) $x_n \rightharpoonup x$ in X implies $\liminf_n \mathcal{R}(x_n) \geq \mathcal{R}(x)$.

(v) (Radon-Riesz property) $x_n \rightharpoonup x$ in X and $\mathcal{R}(x_n) \rightarrow \mathcal{R}(x)$ imply $x_n \rightarrow x$ in X .

Definition 2.3. (Bregman distance) A Bregman distance $D_{\mathcal{R}}^\xi : X \times X \rightarrow \mathbb{R}^+$ of two element $x_1, x_2 \in X$ is defined by

$$D_{\mathcal{R}}^\xi(x_1, x_2) := \mathcal{R}(x_1) - \mathcal{R}(x_2) - \langle \xi, x_1 - x_2 \rangle,$$

where $\xi \in \partial \mathcal{R}(x_2)$ and $\partial \mathcal{R}(x_2)$ denotes the subgradient of $\mathcal{R}(x)$ at x_2 .

Definition 2.4. (Morozov's discrepancy principle) Given $\tau_2 > \tau_1 \geq 1$, choose $\alpha = \alpha(\delta, y^\delta) > 0$ such that

$$\tau_1 \delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2 \delta \quad (10)$$

holds for an element $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$.

The following first order necessary condition for problem (2) is standard ([16, Lemma 3.1]).

Lemma 2.5. If $\hat{x}$ is a minimizer of $\mathcal{J}_\alpha^\delta(x)$, then

$$\langle F'(\hat{x})(F(\hat{x}) - y^\delta), z - \hat{x} \rangle \geq \alpha \mathcal{R}(\hat{x}) - \alpha \mathcal{R}(z), \quad \forall z \in X. \quad (11)$$

3. Existence of regularization parameter. In this section, we discuss the existence of the regularization parameter α determined by MDP. In the rest of this paper, we view α as the regularization parameter. For fixed noise level $\delta > 0$, define functions

$$\begin{aligned} G(x_\alpha^\delta) &:= \frac{1}{2} \|F(x_\alpha^\delta) - y^\delta\|_Y^2, \\ m(\alpha) &:= \mathcal{J}_\alpha^\delta(x_\alpha^\delta) = \min \mathcal{J}_\alpha^\delta(x) \end{aligned}$$

for $\alpha \in (0, \infty)$.

Next, we recall some properties of $\mathcal{R}(x_\alpha^\delta)$, $G(x_\alpha^\delta)$ and $m(\alpha)$, cf. [42] for their proofs.

Lemma 3.1. *$\mathcal{R}(x_\alpha^\delta)$ is non-increasing with respect to the regularization parameter α , whereas $G(x_\alpha^\delta)$ and $m(\alpha)$ are non-decreasing with respect to $\alpha \in (0, \infty)$.*

Lemma 3.2. *Under Condition 2.1, there exist $\alpha_1, \alpha_2 \in \mathbb{R}^+$ such that*

$$\|F(x_{\alpha_1}^\delta) - y^\delta\|_Y < \tau_1 \delta < \tau_2 \delta < \|F(x_{\alpha_2}^\delta) - y^\delta\|_Y.$$

Thanks to Condition 2.1 and 2.2, we have the following standard result, cf. [19] for its proof.

Lemma 3.3. *Let $\alpha_n \rightarrow \alpha > 0$ as $n \rightarrow \infty$. Denote by $x_n := x_{\alpha_n}^\delta$ a minimizer of $\mathcal{J}_{\alpha_n}^\delta(x)$. Then $\{x_n\}$ contains a convergent subsequence $\{x_{n_k}\}$ such that $x_{n_k} \rightarrow x_\alpha^\delta$ in X , where x_α^δ is a minimizer of $\mathcal{J}_\alpha^\delta(x)$.*

Proposition 3.4. *The function $m(\alpha)$ is continuous with respect to α .*

Although $m(\alpha)$ is continuous, the functions $G(x_\alpha^\delta)$ and $\mathcal{R}(x_\alpha^\delta)$ are not necessarily continuous with respect to α , as shown in Fig. 6 for a nonlinear compressive sensing problem. If (10) does not hold for any parameter α , then $\|F(x_\alpha^\delta) - y^\delta\|_Y$ has a jump at a certain parameter. Actually, we have the following lemma, which can be proved similar to that of Theorem 2.3 in [42].

Lemma 3.5. *Denote $G(x) := \frac{1}{2}\|F(x) - y^\delta\|_Y$. For each $\alpha > 0$, there exist $x_{\alpha}^{\delta,1}, x_{\alpha}^{\delta,2} \in \mathcal{L}_\alpha^\delta$ such that*

$$G(x_{\alpha}^{\delta,1}) = \inf_{x \in \mathcal{L}_\alpha^\delta} G(x), \quad G(x_{\alpha}^{\delta,2}) = \sup_{x \in \mathcal{L}_\alpha^\delta} G(x).$$

A minimizer of $\mathcal{J}_\alpha^\delta(x)$ may be non-unique. Hence, the inequality (10) may not hold. More precisely, there may exist a regularization parameter α such that

$$\|F(x_{\alpha}^{\delta,1}) - y^\delta\|_Y < \tau_1 \delta < \tau_2 \delta < \|F(x_{\alpha}^{\delta,2}) - y^\delta\|_Y. \quad (12)$$

To prove the existence of a regularization parameter α satisfying (10), it is sufficient to ensure that there are no minimizers $x_{\alpha}^{\delta,1}, x_{\alpha}^{\delta,2}$ of $\mathcal{J}_\alpha^\delta(x)$ for (12) to be valid. It can be shown that if there is no parameter α such that (12) holds, then there exists α satisfying (10). The proof is similar to that of Theorem 3.10 in [3].

Lemma 3.6. *Assume Conditions 2.1 and 2.2. If there is no $\alpha > 0$ with minimizers $x_{\alpha}^{\delta,1}, x_{\alpha}^{\delta,2} \in \mathcal{L}_\alpha^\delta$ such that (12) is valid, then there exist $\alpha = \alpha(\delta, y^\delta) > 0$ and $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$ such that (10) holds.*

Next, we investigate the selection range of the corresponding regularization parameters when τ_1 or τ_2 is fixed. Initially, we provide the estimates for $\|F(x_{\alpha}^{\delta,1}) - y^\delta\|_Y$ and $\|F(x_{\alpha}^{\delta,2}) - y^\delta\|_Y$ if $x_{\alpha}^{\delta,1}$ and $x_{\alpha}^{\delta,2}$ satisfy (12).

3.1. The estimate for $\|F(x_{\alpha}^{\delta,2}) - y^\delta\|_Y$.

Lemma 3.7. *Assume $x_{\alpha}^{\delta,1}, x_{\alpha}^{\delta,2}$ satisfy (12). Then*

$$\tau_2 \delta < \|F(x_{\alpha}^{\delta,2}) - y^\delta\|_Y \leq \tau_1 (3 + 2\gamma) \delta.$$

Proof. Denote $x_1 := x_\alpha^{\delta,1}$, $x_2 := x_\alpha^{\delta,2}$. From Lemma 2.5, we have

$$\langle F'(x_1)^* (F(x_1) - y^\delta), z - x_1 \rangle \geq \alpha \mathcal{R}(x_1) - \alpha \mathcal{R}(z), \quad \forall z \in X. \quad (13)$$

Thus,

$$\langle F'(x_1)^* (F(x_1) - y^\delta), x_2 - x_1 \rangle \geq \alpha \mathcal{R}(x_1) - \alpha \mathcal{R}(x_2),$$

which can be rewritten as

$$\alpha \mathcal{R}(x_1) - \alpha \mathcal{R}(x_2) \leq \|F(x_1) - y^\delta\|_Y \|F'(x_1)(x_2 - x_1)\|_Y.$$

It follows from Condition 2.1 (iv) that

$$\alpha \mathcal{R}(x_1) - \alpha \mathcal{R}(x_2) \leq \|F(x_1) - y^\delta\|_Y (1 + \gamma) \|F(x_2) - F(x_1)\|_Y. \quad (14)$$

In addition,

$$\begin{aligned} \frac{1}{2} \|F(x_2) - y^\delta\|_Y^2 &= \frac{1}{2} \|F(x_2) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x_2) - \alpha \mathcal{R}(x_2) \\ &= \frac{1}{2} \|F(x_1) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x_1) - \alpha \mathcal{R}(x_2). \end{aligned} \quad (15)$$

A combination of (14) and (15) leads to

$$\begin{aligned} \frac{1}{2} \|F(x_2) - y^\delta\|_Y^2 &\leq \frac{1}{2} \tau_1^2 \delta^2 + \tau_1 \delta (1 + \gamma) \|F(x_2) - y^\delta + y^\delta - F(x_1)\|_Y \\ &\leq \frac{2\gamma + 3}{2} \tau_1^2 \delta^2 + \tau_1 \delta (1 + \gamma) \|F(x_2) - y^\delta\|_Y, \end{aligned} \quad (16)$$

which is rewritten as

$$\|F(x_2) - y^\delta\|_Y^2 - (2 + 2\gamma) \tau_1 \delta \|F(x_2) - y^\delta\|_Y - (2\gamma + 3) \tau_1^2 \delta^2 \leq 0. \quad (17)$$

The inequality (17) implies that

$$-\tau_1 \delta \leq \|F(x_2) - y^\delta\|_Y \leq (2\gamma + 3) \tau_1 \delta.$$

This proves the lemma. $\square$

From Lemmas 3.6 and 3.7, we deduce the following result.

Theorem 3.8. *Assume Conditions 2.1 and 2.2. Then there exist a regularization parameter α and a minimizer $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$ such that*

$$\tau_1 \delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq c \delta, \quad c := \max \{ \tau_2, (3 + 2\gamma) \tau_1 \}. \quad (18)$$

Proof. By Lemma 3.7, if there exists a regularization parameter α such that $x_\alpha^{\delta,1}$, $x_\alpha^{\delta,2}$ satisfy (12), then

$$\tau_2 \delta < \|F(x_\alpha^{\delta,2}) - y^\delta\|_Y \leq \tau_1 (3 + 2\gamma) \delta.$$

This implies that there exists a regularization parameter α such that (18) holds. On the contrary, if there is no regularization parameter α with minimizers $x_\alpha^{\delta,1}$, $x_\alpha^{\delta,2}$ such that (12) is valid, by Lemma 3.6, there exist α and $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$ such that (10) holds. Obviously, (10) implies (18). Hence, there exists a regularization parameter α such that (18) holds, which proves the theorem. $\square$

In the following, we will show that for MDP, $\alpha \equiv \alpha(\delta, y^\delta) \rightarrow 0$ as the noise level $\delta \rightarrow 0$.

Lemma 3.9. *Let $\alpha > 0$ be fixed. If $F(x^*) = y$ and*

$$x^* \in \arg \min_{x \in X} \left\{ \frac{1}{2} \|F(x) - y\|_Y^2 + \alpha \mathcal{R}(x) \right\},$$

then $x^* = 0$.

Proof. By the assumption, we have

$$\begin{aligned} \alpha \mathcal{R}(x^*) &= \frac{1}{2} \|F(x^*) - y\|_Y^2 + \alpha \mathcal{R}(x^*) \\ &\leq \frac{1}{2} \|F((1-t)x^*) - y\|_Y^2 + \alpha \mathcal{R}((1-t)x^*), \quad \forall t \in (0, 1). \end{aligned} \quad (19)$$

Since $\mathcal{R}((1-t)x^*) = (1-t)\mathcal{R}(x^*)$, we can rewrite (19) as

$$0 \leq \alpha \mathcal{R}(x^*) \leq \frac{1}{2} \frac{\|F((1-t)x^*) - y\|_Y^2}{t}, \quad \forall t \in (0, 1).$$

Since

$$\begin{aligned} F((1-t)x^*) - y &= F(x^*) + F'(x^*)(-tx^*) + o(\| -tx^* \|_X) - y \\ &= F'(x^*)(-tx^*) + o(t), \end{aligned}$$

we have

$$0 \leq \alpha \mathcal{R}(x^*) \leq \frac{t}{2} \|F'(x^*)x^* + o(t)\|_Y^2, \quad \forall t \in (0, 1).$$

Then take the limit as $t \rightarrow 0^+$ to obtain $\mathcal{R}(x^*) = 0$. By Condition 2.2, $\mathcal{R}(x^*) = 0$ implies that $x^* = 0$. $\square$

Theorem 3.10. *Let $\delta_n \rightarrow 0$ and $y^{\delta_n} \rightarrow y$ as $n \rightarrow \infty$. Let $\alpha_n := \alpha(\delta_n, y^{\delta_n})$ be the regularization parameter obtained from MDP (18) with δ replaced by δ_n . Then $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. We argue by contradiction. Suppose α_n does not converge to 0, i.e., $\exists \alpha_0 > 0$, $\forall N \in \mathbb{N}^+$, $\exists n_0 > N$ such that $|\alpha_{n_0} - 0| \geq \alpha_0$. This implies that there exists a subsequence of $\{\alpha_n\}$, still denoted by $\{\alpha_n\}$ such that $\alpha_n \geq \alpha_0$. Denote

$$x_{\alpha_0}^{\delta_n} \in \arg \min_{x \in X} \left\{ \frac{1}{2} \|F(x) - y^{\delta_n}\|_Y^2 + \alpha_0 \mathcal{R}(x) \right\}$$

and

$$x_n := x_{\alpha_n}^{\delta_n} \in \arg \min_{x \in X} \left\{ \frac{1}{2} \|F(x) - y^{\delta_n}\|_Y^2 + \alpha_n \mathcal{R}(x) \right\}.$$

By Lemma 3.1, $\|F(x_{\alpha_0}^{\delta_n}) - y^{\delta_n}\|_Y^2$ is non-decreasing with respect to α . Hence,

$$\begin{aligned} \frac{1}{2} \|F(x_{\alpha_0}^{\delta_n}) - y^{\delta_n}\|_Y^2 &\leq \frac{1}{2} \|F(x_n) - y^{\delta_n}\|_Y^2 \\ &\leq \frac{1}{2} \max\{\tau_2^2 \delta_n^2, (3 + 2\gamma)^2 \tau_1^2 \delta_n^2\} \rightarrow 0. \end{aligned} \quad (20)$$

By the definition of $x_{\alpha_0}^{\delta_n}$, there exist $x^* \in X$ and a subsequence $x_{\alpha_0}^{\delta_{n_k}}$ such that $x_{\alpha_0}^{\delta_{n_k}} \rightharpoonup x^*$ in X . By Condition 2.2 (ii), $\{x_{\alpha_0}^{\delta_{n_k}}\}$ is bounded in X , otherwise $\mathcal{R}(x_{\alpha_0}^{\delta_{n_k}}) \rightarrow \infty$, contradicting the definition of the minimizer. Since F is weakly sequentially closed and $y^{\delta_n} \rightarrow y$, $F(x_{\alpha_0}^{\delta_{n_k}}) - y^{\delta_{n_k}} \rightharpoonup F(x^*) - y$, it follows from the weak lower semi-continuity of the norm that

$$0 \leq \frac{1}{2} \|F(x^*) - y\|_Y^2 \leq \liminf_{k \rightarrow \infty} \frac{1}{2} \|F(x_{\alpha_0}^{\delta_{n_k}}) - y^{\delta_{n_k}}\|_Y^2$$

$$\leq \frac{1}{2} \max\{\tau_2^2 \delta_{n_k}^2, (3 + 2\gamma)^2 \tau_1^2 \delta_{n_k}^2\} \rightarrow 0. \quad (21)$$

Hence, $F(x^*) = y$. In addition,

$$\begin{aligned} \frac{1}{2} \|F(x^*) - y\|_Y^2 + \alpha_0 \mathcal{R}(x^*) &\leq \liminf_{k \rightarrow \infty} \left\{ \frac{1}{2} \|F(x_{\alpha_0}^{\delta_{n_k}}) - y^{\delta_{n_k}}\|_Y^2 + \alpha_0 \mathcal{R}(x_{\alpha_0}^{\delta_{n_k}}) \right\} \\ &\leq \liminf_{k \rightarrow \infty} \left\{ \frac{1}{2} \|F(x) - y^{\delta_{n_k}}\|_Y^2 + \alpha_0 \mathcal{R}(x) \right\} \\ &= \frac{1}{2} \|F(x) - y\|_Y^2 + \alpha_0 \mathcal{R}(x) \end{aligned} \quad (22)$$

for any $x \in X$. Thus, x^* is a minimizer of $\frac{1}{2} \|F(x) - y\|_Y^2 + \alpha_0 \mathcal{R}(x)$. By Lemma 3.9, this implies that $x^* = 0$. Then $y = F(0)$. Hence, $\|F(0) - y^\delta\|_Y = \|y - y^\delta\|_Y \leq \delta$, contradicting to $\|F(0) - y^\delta\|_Y > \delta$ in Condition 2.1 (iv). This proves the theorem. $\square$

For a numerical realization of MDP (18), we can make use of an iterative algorithm described in [37], cf. Algorithm 1. Even though α is determined from the upper bound: $\|F(x_\alpha^\delta) - y^\delta\|_Y = c\delta$, we can still obtain the convergence of the regularized solution. Numerical simulation results reported in Section 5 show that we can obtain better inversion results if α is determined by Algorithm 1.

Algorithm 1 Iterative algorithm for α under MDP (18)

Choose $\tau > 0$, $0 < q < 1$, $j = 0$, $\alpha_0 > 0$ with $\|F(x_{\alpha_0}^\delta) - y^\delta\|_Y > c\delta$
 For $j = 1, 2, \dots$,
 if $\|F(x_{\alpha_j}^\delta) - y^\delta\|_Y > c\delta$
 $\alpha_j = q\alpha_{j-1}$
 compute $x_{\alpha_j}^\delta$
 $\alpha_j^{\max} = \alpha_{j-1}$, $\alpha_j^{\min} = \alpha_j$
 else
 $\alpha_j = (\alpha_{j-1}^{\min} + \alpha_{j-1}^{\max}) / 2$
 compute $x_{\alpha_j}^\delta$ by iterative soft thresholding algorithm
 if $\|F(x_{\alpha_j}^\delta) - y^\delta\|_Y > c\delta$ then $\alpha_j^{\max} = \alpha_j$
 if $\|F(x_{\alpha_j}^\delta) - y^\delta\|_Y < \tau_1 \delta$ then $\alpha_j^{\min} = \alpha_j$
 end

We note that the upper bound $c\delta$ from Theorem 3.8 is $O(\delta)$. Then the convergence of the regularized solution is standard ([19]). Indeed, we can show the regularized solution x_α^δ defined by (7) converges to an $\mathcal{R}$ -minimum solution of problem $F(x) = y$.

3.2. The estimate for $\|F(x_\alpha^{\delta,1}) - y^\delta\|_Y$.

Lemma 3.11. Assume $x_\alpha^{\delta,1}$, $x_\alpha^{\delta,2}$ satisfy (12). Then

$$\frac{\tau_2}{2\gamma + 3} \delta < \|F(x_\alpha^{\delta,1}) - y^\delta\|_Y \leq \tau_1 \delta.$$

Proof. Denote $x_1 := x_\alpha^{\delta,1}$, $x_2 := x_\alpha^{\delta,2}$. From Lemma 3.7, we have

$$\alpha \mathcal{R}(x_2) - \alpha \mathcal{R}(x_1) \geq -\|F(x_1) - y^\delta\|_Y \|F'(x_1)(x_2 - x_1)\|_Y.$$

It follows from Condition 2.1 (iv) that

$$\alpha \mathcal{R}(x_2) - \alpha \mathcal{R}(x_1) \geq -\|F(x_1) - y^\delta\|_Y (1 + \gamma) \|F(x_2) - F(x_1)\|_Y. \quad (23)$$

In addition,

$$\begin{aligned} \frac{1}{2}\|F(x_1) - y^\delta\|_Y^2 &= \frac{1}{2}\|F(x_1) - y^\delta\|_Y^2 + \alpha\mathcal{R}(x_1) - \alpha\mathcal{R}(x_1) \\ &= \frac{1}{2}\|F(x_2) - y^\delta\|_Y^2 + \alpha\mathcal{R}(x_2) - \alpha\mathcal{R}(x_1). \end{aligned} \quad (24)$$

A combination of (23) and (24) implies that

$$\begin{aligned} \frac{1}{2}\|F(x_1) - y^\delta\|_Y^2 &\geq \frac{1}{2}\tau_2^2\delta^2 - \|F(x_1) - y^\delta\|_Y(1 + \gamma)\|F(x_2) - y^\delta + y^\delta - F(x_1)\|_Y \\ &\geq \frac{1}{2}\tau_2^2\delta^2 - \|F(x_1) - y^\delta\|_Y(1 + \gamma)(\tau_2\delta + \|F(x_1) - y^\delta\|_Y), \end{aligned} \quad (25)$$

which is rewritten as

$$(3 + 2\gamma)\|F(x_1) - y^\delta\|_Y^2 + (2 + 2\gamma)\tau_2\delta\|F(x_1) - y^\delta\|_Y - \tau_2^2\delta^2 \geq 0. \quad (26)$$

The inequality (26) implies that $\|F(x_1) - y^\delta\|_Y \geq \frac{\tau_2}{2\gamma+3}\delta$. This proves the lemma. $\square$

From Lemmas 3.6 and 3.11, we deduce the following result.

Theorem 3.12. *Assume Conditions 2.1 and 2.2. Then there exist a regularization parameter α and a minimizer $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$ such that*

$$c\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2\delta, \quad c := \min \left\{ \tau_1, \frac{\tau_2}{3 + 2\gamma} \right\}. \quad (27)$$

Proof. By Lemma 3.11, if there exists a regularization parameter α such that $x_\alpha^{\delta,1}$, $x_\alpha^{\delta,2}$ satisfy (12), then

$$\frac{\tau_2}{3 + 2\gamma}\delta < \|F(x_\alpha^{\delta,1}) - y^\delta\|_Y \leq \tau_2\delta.$$

This implies that there exists a regularization parameter α such that (27) holds. On the contrary, if there is no regularization parameter α with minimizers $x_\alpha^{\delta,1}$, $x_\alpha^{\delta,2}$ such that (12) is valid, by Lemma 3.6, there exist α and $x_\alpha^\delta \in \mathcal{L}_\alpha^\delta$ such that (10) holds. It is obvious that (10) implies that (27) holds. Hence, there exists a regularization parameter α such that (27) holds, which proves the theorem. $\square$

Theorem 3.13. *Let $\delta_n \rightarrow 0$ and $y^{\delta_n} \rightarrow y$ as $n \rightarrow \infty$. Let $\alpha_n := \alpha(\delta_n, y^{\delta_n})$ be the regularization parameter obtained from MDP (27) with δ replaced by δ_n . Then $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$.*

The proof of Theorem 3.13 exhibits a structure akin to that of Theorem 3.10, with the distinction that the parameter τ_2 replaces $c := \max \{\tau_2, (3 + 2\gamma)\tau_1\}$. Consequently, it follows that Theorem 3.13 is evidently valid.

4. Regularization properties. In this section, we examine the well-posedness of the regularization method under MDP (18) and (27). We prove that the regularized solution x_α^δ defined by (7) converges to an $\mathcal{R}$ -minimum solution of the problem $F(x) = y$. In addition, we discuss the convergence rate of x_α^δ . The proofs are along the lines of standard quadratic Tikhonov regularization ([19]) and sparsity regularization ([23, 27, 39, 43]). However, analysis of the convergence rate is different from that in [15] since the regularization parameter α is now determined by (18).

Theorem 4.1. (Convergence) *Let $x_{\alpha_n}^{\delta_n}$ be a minimizer of $\mathcal{J}_{\alpha_n}^{\delta_n}(x)$ defined by (2) with the data y^{δ_n} satisfying $\|y - y^{\delta_n}\| \leq \delta_n$, where $\delta_n \rightarrow 0$ as $n \rightarrow \infty$ and y^{δ_n} belongs to the range of F . Let α_n be chosen by the Morozov's discrepancy principle (18). Then there exists a subsequence of $\{x_{\alpha_n}^{\delta_n}\}$, still denoted by $\{x_{\alpha_n}^{\delta_n}\}$, that converges to an $\mathcal{R}$ -minimum solution $x^\dagger$ in X . In addition, if the $\mathcal{R}$ -minimum solution $x^\dagger$ is unique, then*

$$\lim_{n \rightarrow \infty} \|x_{\alpha_n}^{\delta_n} - x^\dagger\|_X = 0.$$

Proof. Denote $y_n := y^{\delta_n}$, $x_n := x_{\alpha_n}^{\delta_n}$. By the definition of x_n , we obtain

$$\begin{aligned} \frac{1}{2} \|F(x_n) - y_n\|_Y^2 + \alpha_n \mathcal{R}(x_n) &\leq \frac{1}{2} \|F(x^\dagger) - y_n\|_Y^2 + \alpha_n \mathcal{R}(x^\dagger) \\ &\leq \frac{1}{2} \delta_n^2 + \alpha_n \mathcal{R}(x^\dagger). \end{aligned} \quad (28)$$

By (18), this implies that $\mathcal{R}(x_n) \leq \mathcal{R}(x^\dagger)$. Hence, the sequence $\{\mathcal{R}(x_n)\}$ is bounded. Denote

$$c := \max\{\tau_2, (3 + 2\gamma)\tau_1\}.$$

We have $c\delta_n \rightarrow 0$ as $\delta_n \rightarrow 0$. In addition,

$$\|F(x_n) - y\|_Y \leq \|F(x_n) - y_n\|_Y + \|y - y_n\|_Y \leq c\delta_n + \delta_n. \quad (29)$$

Then

$$\lim_{n \rightarrow \infty} F(x_n) = y. \quad (30)$$

On the other hand, it follows from (28) that

$$\limsup_{n \rightarrow \infty} \mathcal{R}(x_n) \leq \mathcal{R}(x^\dagger). \quad (31)$$

Since $\mathcal{R}(x_n)$ is bounded, there exist an element $x^* \in X$ and a subsequence of $\{x_n\}$, still denoted by $\{x_n\}$, such that $x_n \rightharpoonup x^*$ in X . Together with (30), it follows that

$$\|F(x^*) - y\|_Y \leq \liminf_{n \rightarrow \infty} \|F(x_n) - y\|_Y = 0.$$

Hence, $F(x^*) = y$. Meanwhile, by Condition 2.2 (ii), we have

$$\mathcal{R}(x^*) \leq \liminf_n \mathcal{R}(x_n) \leq \mathcal{R}(x^\dagger). \quad (32)$$

By the definition of $x^\dagger$, x^* is an $\mathcal{R}$ -minimum solution. If the $\mathcal{R}$ -minimum solution is unique, then $x^* = x^\dagger$. A combination of (31) and (32) implies $\mathcal{R}(x_n) \rightarrow \mathcal{R}(x^\dagger)$. Thus, $\mathcal{R}(x_n) \rightarrow \mathcal{R}(x^\dagger)$. Then $\lim_{n \rightarrow \infty} \|x_n - x^\dagger\|_X = 0$ by Condition 2.2 (iii). $\square$

Next, we present a result on the linear convergence rate $O(\delta)$ of the regularized solution x_α^δ under the source condition and Bregman distance. For this purpose, we introduce an assumption on an $\mathcal{R}$ -minimum solution $x^\dagger$ of the problem $F(x) = y$.

Remark 4.2. The convergence outcomes presented in Theorems 4.1 remain applicable to MDP (27), with the distinction that $c := \max\{\tau_2, (3 + 2\gamma)\tau_1\}$ is substituted with τ_2 .

Assumption 4.3. *There exists an $\omega \in Y$ such that*

$$F'(x^\dagger)^* \omega \in \partial \mathcal{R}(x^\dagger). \quad (33)$$

Theorem 4.4. (Convergence rate) *Let x_α^δ be defined by (7). Assume that there exist parameters α satisfying MDP (18). Then, under Assumption 4.3, there exists $d \in D_{\mathcal{R}}^\xi(x_\alpha^\delta, x^\dagger)$ such that $d = O(\delta)$.*

Proof. Due to MDP (18) and the minimization property of x_α^δ ,

$$\begin{aligned} \frac{1}{2}\delta^2 + \alpha\mathcal{R}(x_\alpha^\delta) &\leq \frac{1}{2}\|F(x_\alpha^\delta) - y^\delta\|_Y^2 + \alpha\mathcal{R}(x_\alpha^\delta) \\ &\leq \frac{1}{2}\|F(x^\dagger) - y^\delta\|_Y^2 + \alpha\mathcal{R}(x^\dagger) \\ &\leq \frac{1}{2}\delta^2 + \alpha\mathcal{R}(x^\dagger). \end{aligned}$$

Then $\mathcal{R}(x_\alpha^\delta) \leq \mathcal{R}(x^\dagger)$. Consider the Bregman distance

$$d = \mathcal{R}(x_\alpha^\delta) - \mathcal{R}(x^\dagger) - \langle F'(x^\dagger)^* \omega, x_\alpha^\delta - x^\dagger \rangle.$$

Since $\mathcal{R}(x_\alpha^\delta) \leq \mathcal{R}(x^\dagger)$,

$$d \leq -\langle F'(x^\dagger)^* \omega, x_\alpha^\delta - x^\dagger \rangle \leq \|\omega\| \|F'(x^\dagger)(x_\alpha^\delta - x^\dagger)\|_Y. \quad (34)$$

A combination of (34) and the nonlinear condition (8) implies that

$$\begin{aligned} d &\leq \|\omega\|(1 + \gamma)\|F(x_\alpha^\delta) - F(x^\dagger)\|_Y \\ &\leq \|\omega\|(1 + \gamma)(\|F(x_\alpha^\delta) - y^\delta\|_Y + \|y - y^\delta\|_Y). \end{aligned}$$

By MDP (18), $\|F(x_\alpha^\delta) - y^\delta\|_Y \leq c\delta$. Thus

$$d \leq \|\omega\|(1 + \gamma)(c + 1)\delta,$$

i.e., $d = O(\delta)$. □

Remark 4.5. Regrettably, the convergence rate results from Theorem 4.4 cannot be extended to MDP (27) using a similar proof. Although $\tau_1, \tau_2 > 1$, the value of $\tau_2/(3\gamma + 2)$ may still be less than 1. In such instances, the proof method employed in Theorem 4.4 does not hold.

5. Numerical experiments. In this section, we present two numerical examples to demonstrate the efficiency of MDP (18). The relative error (Rerror) is utilized to measure the accuracy of the reconstruction x^* :

$$\text{Rerror} := \frac{\|x^* - x^\dagger\|_{\ell_2}}{\|x^\dagger\|_{\ell_2}},$$

where $x^\dagger$ is a true solution. The exact data $y^\dagger$ is generated by $y^\dagger = F(x^\dagger)$. The noise level δ is defined by $\delta := \|F(x^\dagger) - y^\delta\|_Y$. All numerical calculations were done using MATLAB R2010 on an i7-6500U 2.50GHz workstation with 8Gb RAM.

5.1. A nonlinear compressive sensing problem. First, we solve a nonlinear compressive sensing (CS) problem ([6, 8, 12, 44]) to demonstrate the efficiency of MDP for the ℓ_1 -regularization. The research of nonlinear CS is important not only in theoretical analysis but also in many applications where the observation system is often nonlinear. For example, in diffraction imaging, a charge coupled device (CCD) records the amplitude of the Fourier transform of the original signal. So one only obtains the nonlinear measurements of the original signal. Fortunately, as shown in [8], if the system satisfies some nonlinear conditions then recovery is still possible. Under the nonlinear CS frame, the measurement system is nonlinear. Assume, therefore, that the observation model is

$$y = F(x) + \delta, \quad (35)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a nonlinear operator, $x \in \mathbb{R}^n$ and $\delta \in \mathbb{R}^m$ is a noise. It can be shown that if the linearization of F at an exact solution $x^\dagger$ satisfies the

restricted isometry property (RIP), then the convergence property of the iterative hard thresholding algorithm (IHTA) is guaranteed ([8]). Next we illustrate the efficiency of the proposed algorithm by a nonlinear CS example of the form ([8])

$$y = F(x) := \hat{a}(A\hat{b}(x)), \quad (36)$$

where A is a CS matrix, $\hat{a}$ and $\hat{b}$ are nonlinear operators, respectively. Here, $\hat{a}$ encodes nonlinearity after mixing by A as well as nonlinear “crosstalk” between mixed elements, and $\hat{b}$ encodes the same system properties for the inputs before mixing. For simplicity, we write $\hat{a}(x) = x^c$ and $\hat{b}(x) = x + b(x)$, where b is nonlinear map. In particular, let $b(x) = x^d$, where $c, d \in \mathbb{N}^+$ and x^c and x^d should be understood in a componentwise sense.

Next, we consider the nonlinearity of the operator F . In [8], it is shown that the Jacobian matrix of $\hat{a}(A\hat{b}(x))$ is of the form

$$F'(x) = [a'_x(A\hat{b}(x))][A(I + b'_x(x))].$$

There exists a constant $\gamma > 0$ such that

$$\|F(x_1) - F(x_2) - F'(x_2)(x_1 - x_2)\|_{\gamma(\mathbb{R}^n, \mathbb{R}^m)} \leq \gamma \|x_1 - x_2\|_{\ell_2},$$

cf. [8, Lemmas 3 & 4]. For example, we let $c = 1$, $d = 3$, then $F(x) = A(x + x^3)$. Assume A is bounded, then obviously,

$$\|x_1 - x_2\|_{\ell_2} \leq K \|F(x_1) - F(x_2)\|_{\ell_2}.$$

Consequently, F satisfies the nonlinear condition (2.2).

In the experiment, an ST- $(\alpha\ell_1 - \beta\ell_2)$ ($\beta = 0$) algorithm for nonlinear ill-posed problems is used to compute the iterative solution, cf. [18] for details of the algorithm. The ST- $(\alpha\ell_1 - \beta\ell_2)$ algorithm reduces to the form

$$x^{k+1} = \mathbb{S}_{\alpha/\lambda} \left(x^k - \frac{1}{\lambda} F'(x^k)^* (F(x^k) - y^\delta) \right).$$

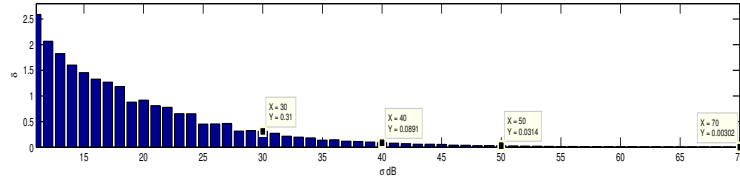


FIGURE 1. The noise level δ vs. σ .

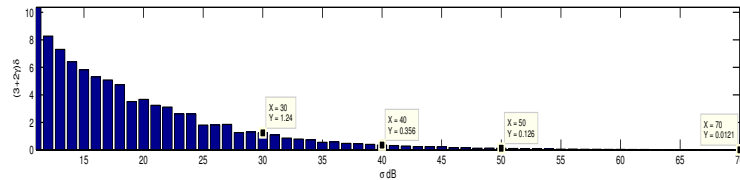


FIGURE 2. The upper bound $c\delta$ vs. σ .

Let $n = 200$, $m = 0.4n$, $p = 0.2m$, $\tau_1 = 1$ and $\tau_2 = 2$, where p is the number of the impulses in the true solution. White Gaussian noise is added to the exact

data $y^\dagger$ by calling $y^\delta = \text{awgn}(F(x^\dagger), \sigma)$ in MATLAB, where σ is the added noise, measured in dB, which measures the ratio between the true (noise free) data $y^\dagger$ or $F(x^\dagger)$ and Gaussian noise. Meanwhile, we let $\gamma = 1/2$, $c = 1$ and $d = 3$. For the sparsity regularization of linear ill-posed problems, the value of $\|A_{m \times n}\|_{\ell_2}$ needs to be less than 1 ([13]). This condition is also needed for the nonlinear CS problem (35). Since the value of $\|A_{m \times n}\|_{\ell_2}$ is around 20, we re-scale the matrix $A_{m \times n}$ by $A_{m \times n} \rightarrow 0.05A_{m \times n}$. The initial value $x^0 = \text{zeros}(n, 1)$.

In Fig. 1 and Fig. 2, we show relations among the noise level δ , $c\delta$ and σ . Since the Gaussian noise is random, strict monotonicity is not maintained numerically for the noise level and the upper bound in (5) with respect to σ . Nevertheless, overall, the approximate monotonicity property is observed.

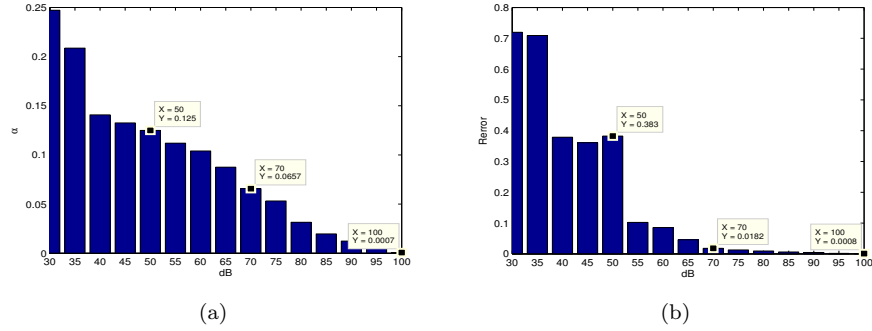


FIGURE 3. (a) α vs. σ , α being determined by $\|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$; (b) Error vs. σ .

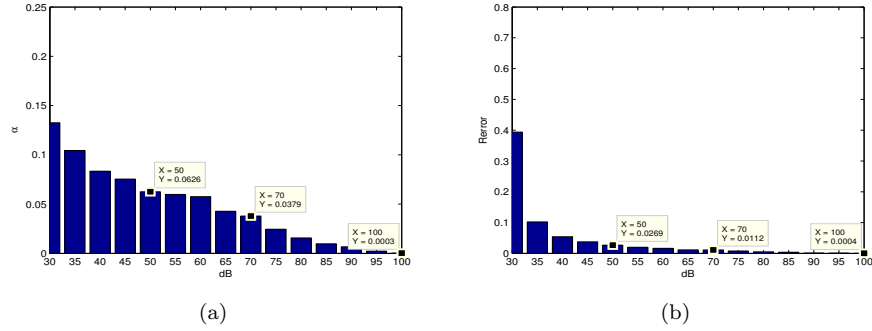


FIGURE 4. (a) α vs. σ , α being determined by Algorithm 1; (b) Error vs. σ .

To test sharpness of the upper bound in (5), let $\tau_1 = 1$ and choose α with $\|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$. The value of α satisfying $\|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$ can be obtained only approximately. In computational practice, this value of α can be found approximately by a sorting from a given set of values $\alpha_1, \alpha_2, \dots, \alpha_n$. For example, with a good initial guess α_0 , we let $\alpha_n = \alpha_0 + 0.001 * n$ (or $\alpha_n = \alpha_0 - 0.001 * n$) to find a satisfactory approximate value of α of a solution of $\|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$.

In Fig. 3 (a), we observe that $\alpha \rightarrow 0$ as the noise level $\delta \rightarrow 0$. In Fig. 3 (b), we see that Error decreases with respect to σ . This illustrates the convergence of the regularized solution with respect to the noise level δ . Although the upper bound in (5) appears to be loose, if the noise levels are small enough, we can still recover satisfactory results.

Next, we test the convergence of the regularized solutions where α is determined by Algorithm 1. In Fig. 4, it is shown that $\alpha \rightarrow 0$ as the noise level $\delta \rightarrow 0$ and Error converge to 0 as $\delta \rightarrow 0$.

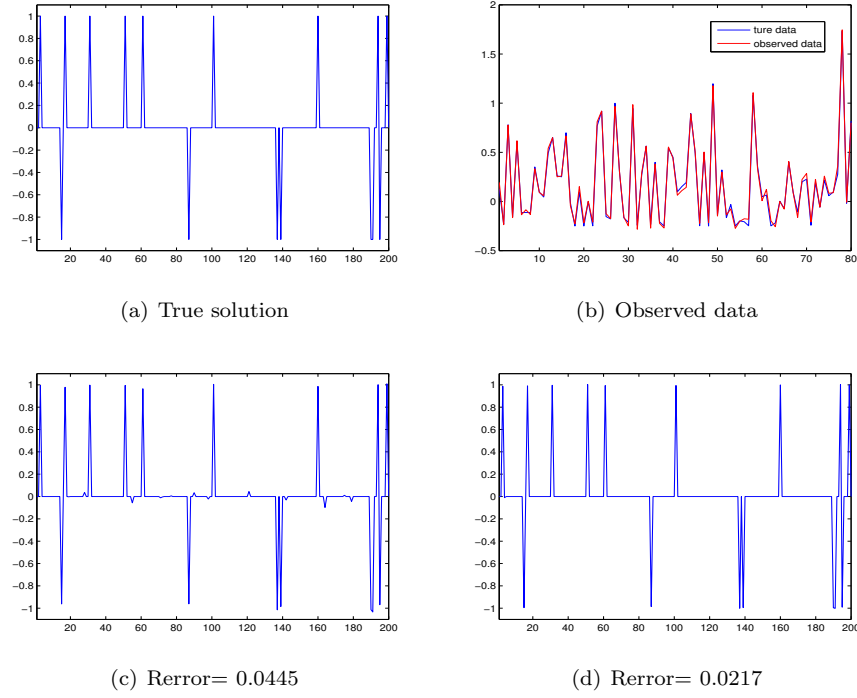


FIGURE 5. (a) True solution; (b) Observed data; (c) Recovered solution with $\alpha : \|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$; (d) Recovered solution with α is determined by Algorithm 1.

Graphs of the reconstruction x^* corresponding to $\sigma = 60\text{dB}$ ($\delta = 0.014$ and $c\delta = 0.056$) can be seen in Fig. 5. In Fig. 5 (c), the reconstruction x^* is computed with $\alpha : \|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} = c\delta$. Error is 4.45%. In Fig. 5 (d), the reconstruction x^* is computed with α determined by Algorithm 1. Error is 2.17%. It is shown that we can obtain better results if the parameter α is determined by Algorithm 1.

Next we show if τ_1 and τ_2 are chosen small, the existence of α can not be guaranteed under the condition (3). We let $\sigma = 30\text{dB}$, then $\delta = 0.2903$ and $c\delta = 1.1612$. In Fig. 6 (a), it is shown if $1 \leq \tau_1 < \tau_2 \leq 2$, then $\tau_1\delta \geq 0.2903$ and $\tau_2\delta \leq 0.5806$. Consequently, there is no α such that the traditional MDP (3) holds. Even with larger values, e.g. $2.208 \leq \tau_1 < \tau_2 \leq 4.681$, it is challenging to find α such that (3) holds, as shown in Fig. 6 (b). Indeed, for the traditional MDP with linear ill-posed problems, one commonly tries $\alpha_j = \alpha 2^{-j}$, $j = 1, 2, \dots$. We increase j and calculate $x_{\alpha_j}^\delta$ until (3) is valid.

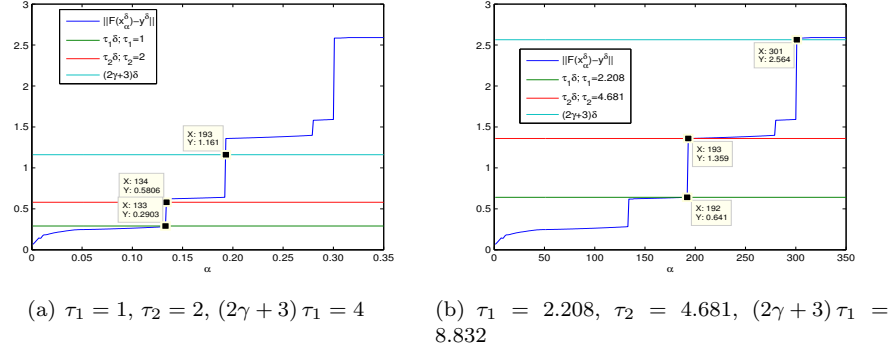


FIGURE 6. The value of the discrepancy $\|F(x) - y^\delta\|_{\ell_2}$ by the ST- $(\alpha\ell_1 - \beta\ell_2)$ ($\beta = 0$) algorithm with different values of α .

However, for nonlinear ill-posed problems, the discrepancy is not continuous with respect to α . If $\alpha_j = 0.2$, then $\alpha_{j+1} = 0.1$, we can not find an α such that $\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} \leq \tau_2\delta$ holds (cf. Fig. 6). Certainly, we can try $\alpha_j = \alpha 3^{-j}$ or $\alpha_j = \alpha 4^{-j}$, etc. Nevertheless, it requires more computational time. In addition, one still does not have the theoretical assurance on the existence of α . So it is challenging to choose appropriate values of τ_1 and τ_2 to ensure the existence of α . Although the upper bound $c\delta$ in (5) seems loose, it guarantees the existence of α . We can determine an α by Algorithm 1 such that the modified MDP (5) holds. So the upper bound $c\delta$ appears to be natural.

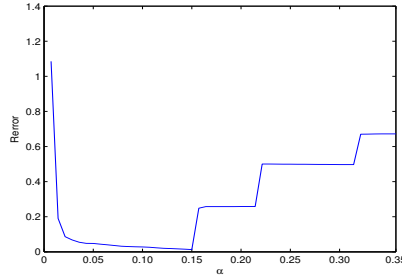


FIGURE 7. The relative error of reconstruction x^* by the ST- $(\alpha\ell_1 - \beta\ell_2)$ ($\beta = 0$) algorithm with different α .

Finally, Fig. 7 illustrates the relationship between the relative error and the regularization parameter. The diagram indicates that the relative error of the reconstructed solution is minimized when $\alpha = 0.15$. At this point, the regularization parameter α satisfies $\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} \leq c\delta$, but it does not fulfill $\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_{\ell_2} \leq \tau_2\delta$, as shown in Fig. 6 (a). This further emphasizes that a more suitable regularization parameter can be identified using Algorithm 1.

5.2. A nonlinear one-dimensional gravity inversion problem. We subsequently present results from a nonlinear one-dimensional gravity inversion problem [38] to demonstrate the efficiency of MDP for the Tikhonov regularization. For this

one-dimensional gravity inversion problem, the operator F exhibits Lipschitz continuity for its first-order derivative but fails to meet the tangential cone condition. Actually, in the proof of Lemma 3.7, the tangent cone condition is used to enlarge $\|F'(x_1)(x_2 - x_1)\|_Y$ into $\|F(x_2) - F(x_1)\|_Y$ in the inequality. In other words, Lemma 3.7 can be proved as long as the operator F satisfies

$$\|F'(x_1)(x_2 - x_1)\|_Y \leq c\|F(x_2) - F(x_1)\|_Y,$$

where c is a constant. Therefore, even if the tangent cone condition is not satisfied, effective results may still be obtained when the above inequality can be established. The primary objective of this example is to assess the applicability of the method discussed in this paper to other nonlinear inversion problems, even when these problems do not adhere to the tangential cone condition.

Investigating one-dimensional gravity anomaly inversion is of significant importance not only for theoretical analyses of potential fields but also for its practical applications in geophysical exploration, where the nonlinear effects of complex geological structures frequently influence observational data. For instance, in the context of geophysical exploration, gravimeters measure the vertical component of the gravitational field [11], which is induced by subsurface density anomalies. Due to the inherent characteristics of signal superposition and attenuation, where the field signal diminishes inversely with the square of the distance, actual observational data can be conceptualized as a nonlinear integral transform of the underlying density distribution. Thus, we can only extract one-dimensional nonlinear projection information regarding the subsurface media. Within the sparse constraint inversion proposed by [32], high-precision inversion remains attainable if the density anomalies exhibit local sparsity conditions, such as those characteristics of a finite layer model. The forward modeling of gravity anomalies is grounded in the gravitational theory about spherical bodies. Specifically, for a sphere of radius r , exhibiting a density contrast $\Delta\rho$, relative to its surrounding matrix and located at a depth d , the vertical gravity anomaly recorded at the surface measurement point x is given by

$$g(x) = \frac{4\pi}{3} \frac{G\Delta\rho r^3 d}{\left(d^2 + (x - x_0)^2\right)^{3/2}}, \quad (37)$$

where G denotes the gravitational constant ($6.674 \times 10^{-11} m^3 kg^{-1} s^{-2}$), and x_0 represents the horizontal position of the sphere. In the case of two spheres, the cumulative gravity anomaly experienced at a given point is the summation of the anomalies produced by each sphere

$$g_{total}(x) = g_1(x) + g_2(x), \quad (38)$$

where $g_1(x)$ denote the gravitational influence of the first spherical body at point x , while $g_2(x)$ signifies the gravitational influence of the second spherical body at the same point x . These functions collectively represent the contributions to the gravity anomaly at x induced by each respective sphere. For this analysis, we examine a simplified scenario of one-dimensional gravity inversion, where two spheres contribute to the gravity anomalies, as illustrated in Fig. 7. The spheres under consideration possess a radius of $r_1 = r_2 = 100m$, a density contrast of $\Delta\rho = 300kg/m^3$, and depths from the surface to their centers of $d_1 = 150m$ and $d_2 = 200m$, respectively. This configuration can be interpreted as a model representing iron ore (a mineral characterized by high iron content) situated within sedimentary

rock formations. Our objective is to ascertain the positional and depth-related information of the spheres based on the observed gravity anomaly data.

Next, we reformulate the problems above into the context of nonlinear ill-posed problems. We denote this relationship as

$$F(x) = g^\delta,$$

where F is a nonlinear operator associated with the gravity system, g^δ represents the observed gravity anomaly data, and x is the parameter to be determined, which includes information on location and depth.

We employ the Landweber iterative algorithm in the experiment to resolve these nonlinear ill-posed problems, as detailed in [25]. The Landweber iterative algorithm can be expressed in a simplified form

$$x^{k+1} = x^k - \omega_k F'(x^k)^* (F(x^k) - g^\delta).$$

To enhance the stability of the algorithm, we incorporate a Tikhonov regularization term [19, 21, 26] into the functional. To prevent the solution from getting trapped in local minimum and to accelerate convergence, we apply hard-thresholding to the regularization parameter and adapt the step size using the Frobenius norm of the Jacobian matrix [5, 10].

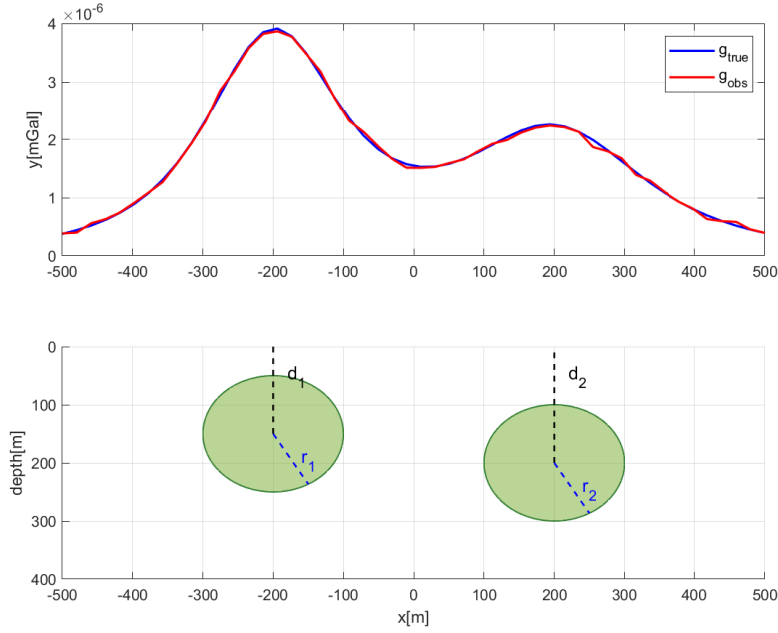


FIGURE 8. Comparative analysis of observed Gravity values and theoretical values.

We consider a scenario with 50 measurement points, a side line range $[-500, 500]$, and a true parameter set $[-200, 150, 200, 200]$ representing the positions and depths of two spheres. We establish a maximum iteration $maxiter = 500$, an initial parameter guess $[-150, 100, 150, 250]$, and introduce 2% Gaussian noise to simulate the disturbances typically encountered in actual observational processes.

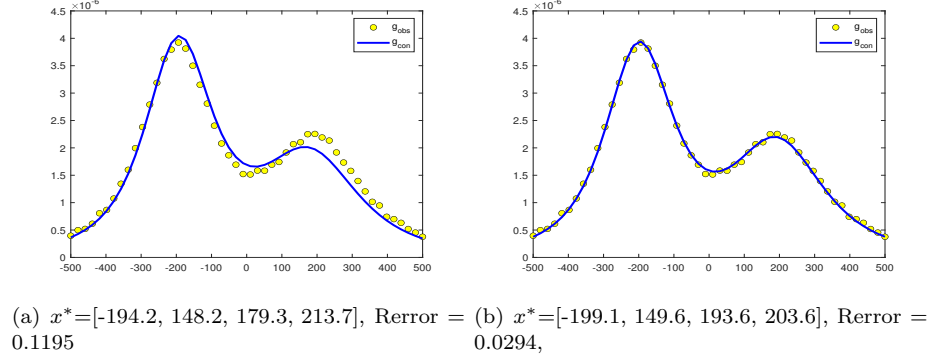


FIGURE 9. (a) Reconstructed data by recovered solution with $\alpha : \|F(x_\alpha^\delta) - y^\delta\|_Y = c\delta$; (b) Reconstructed data by recovered solution with α is determined by Algorithm 1.

The theoretical measurement data, actual measurement data with 2% Gaussian noise ($\delta = 2.7421 \times 10^{-7}$ and $c\delta = 1.2087 \times 10^{-6}$), and the true location parameters of the two underground spheres are illustrated in Fig. 8. Fig. 9 shows the theoretical measurement data generated from the reconstruction solutions obtained using various methods for selecting the regularization parameter. In Fig. 9 (a), the reconstruction x^* is calculated with $\alpha : \|F(x_\alpha^\delta) - y^\delta\|_Y = c\delta$ and Error is 11.95%. In Fig. 9 (b), the reconstruction x^* is derived using α determined by Algorithm 1 and Error is 2.94%. This demonstrates that better outcomes can be achieved when the parameter α is selected using Algorithm 1.

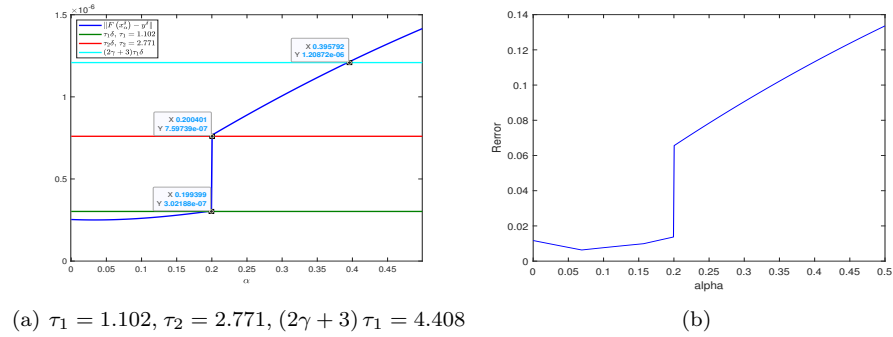


FIGURE 10. (a) The value of the discrepancy $\|F(x) - y^\delta\|_Y$ by the Landweber iterative algorithm with different α ; (b) The relative error of reconstruction x^* by the Landweber iterative algorithm with different α .

Additionally, we analyze the scenario where a small value is chosen for τ_2 . In this case, the existence of α can not be guaranteed under condition (3). Maintaining the 2% Gaussian noise yields $\delta = 2.7421 \times 10^{-7}$ and $c\delta = 1.2087 \times 10^{-6}$. As shown in Fig. 10 (a), if $1.102 \leq \tau_1 < \tau_2 \leq 2.771$, then $\tau_2\delta < 7.5974 \times 10^{-7}$. Consequently, no α satisfies the traditional MDP (3). This observation further emphasizes that

the discrepancy may not be continuous with respect to α for some nonlinear ill-posed problems. The upper bound $c\delta$ in (5) guarantees the existence of α . We can determine an appropriate α using Algorithm 1 such that the modified MDP (5) holds.

Finally, Fig. 10 (b) illustrates the relationship between the relative error and the regularization parameter. The diagram indicates that the relative error of the reconstructed solution is minimized when $\alpha = 0.07$. At this point, though the regularization parameter α does not satisfy both $\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq c\delta$ and $\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2\delta$, as shown in Fig. 10, Fig. 9 (a) still emphasizes that a suitable regularization parameter can be identified by Algorithm 1.

6. Conclusion. For Tikhonov regularization involving a nonlinear operator and a general convex penalty, we demonstrate the existence of a regularization parameter α such that

$$\tau_1\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq c\delta,$$

where $c := \max\{\tau_2, (3 + 2\gamma)\tau_1\}$ or

$$c\delta \leq \|F(x_\alpha^\delta) - y^\delta\|_Y \leq \tau_2\delta,$$

where $c := \min\{\tau_1, \tau_2/(3 + 2\gamma)\}$. Furthermore, we establish that $\alpha \equiv \alpha(\delta, y^\delta) \rightarrow 0$ as the noise level $\delta \rightarrow 0$. It is shown that this regularization method is well-posed when α is determined using Morozov's Discrepancy Principle. We also prove the convergence of the regularized solution x_α^δ , as defined by (7), to an $\mathcal{R}$ -minimum solution of the problem $F(x) = y$. This convergence occurs under Morozov's discrepancy principle, yielding a linear convergence rate $O(\delta)$ for the regularized solution x_α^δ under the source condition and utilizing Bregman distance. Numerical simulation results illustrate the effectiveness of the regularization parameter selection method, showing that the relative error decreases as δ decreases. Although the upper bound in (5) and the lower bound in (6) may appear loose, the existence of α is assured. Moreover, we can still achieve satisfactory results if the noise levels are sufficiently low. A promising direction for future research involves extending the methodologies and results presented in this work to stochastic setting.

REFERENCES

- [1] V. Albani and A. De Cezaro, [A connection between uniqueness of minimizers in Tikhonov-type regularization and Morozov-like discrepancy principles](#), *Inverse Problems and Imaging*, **13** (2019), 211-229.
- [2] S. W. Anzengruber, B. Hofmann and P. Mathé, [Regularization properties of the sequential discrepancy principle for Tikhonov regularization in Banach spaces](#), *Applicable Analysis*, **93** (2014), 1382-1400.
- [3] S. W. Anzengruber and R. Ramlau, [Morozov's discrepancy principle for Tikhonov-type functionals with nonlinear operators](#), *Inverse Problems*, **26** (2010), 025001.
- [4] S. W. Anzengruber and R. Ramlau, [Convergence rates for Morozov's discrepancy principle using variational inequalities](#), *Inverse Problems*, **27** (2011), 105007.
- [5] K. Axiotis and M. Sviridenko, [Sparse convex optimization via adaptively regularized hard thresholding](#), *Journal of Machine Learning Research*, **22** (2021), 1-47.
- [6] A. Beck and Y. C. Eldar, [Sparsity constrained nonlinear optimization: Optimality conditions and algorithms](#), *SIAM Journal on Optimization*, **23** (2013), 1480-1509.
- [7] M. Benning and M. Burger, [Modern regularization methods for inverse problems](#), *Acta Numerica*, (2018), 1-111.
- [8] T. Blumensath, [Compressed sensing with nonlinear observations and related nonlinear optimization problems](#), *IEEE Transactions on Information Theory*, **59** (2013), 3466-3474.
- [9] T. Bonesky, [Morozov's discrepancy principle and Tikhonov-type functionals](#), *Inverse Problems*, **25** (2009), 015015.

- [10] K. Bredies and D. A. Lorenz, [Iterated hard shrinkage for minimization problems with sparsity constraints](#), *SIAM Journal on Scientific Computing*, **30** (2008), 657-683.
- [11] R. J. Blakely, *Potential Theory in Gravity and Magnetic Applications*, Cambridge University Press, 1996.
- [12] S. P. Chepuri and G. Leus, [Sparsity-promoting sensor selection for non-linear measurement models](#), *IEEE Transactions on Signal Processing*, **63** (2015), 684-698.
- [13] I. Daubechies, M. Defrise and C. De Mol, [An iterative thresholding algorithm for linear inverse problems with a sparsity constraint](#), *Communications on Pure and Applied Mathematics*, **57** (2004), 1413-1457.
- [14] I. Daubechies, M. Defrise and C. De Mol, [Sparsity-enforcing regularisation and ISTA revisited](#), *Inverse Problems*, **32** (2016), 104001.
- [15] L. Ding and W. Han, [\$\alpha\ell_1 - \beta\ell_2\$ regularization for sparse recovery](#), *Inverse Problems*, **35** (2019), 125009.
- [16] L. Ding and W. Han, [A projected gradient method for \$\alpha\ell_1 - \beta\ell_2\$ sparsity regularization](#), *Inverse Problems*, **36** (2020), 125012.
- [17] L. Ding and W. Han, [Morozov's discrepancy principle for \$\alpha\ell_1 - \beta\ell_2\$ sparsity regularization](#), *Inverse Problems and Imaging*, **17** (2023), 157-179.
- [18] L. Ding and W. Han, [\$\alpha\ell_1 - \beta\ell_2\$ sparsity regularization for nonlinear ill-posed problems](#), *Journal of Computational and Applied Mathematics*, **450** (2024), 115987.
- [19] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Mathematics and its Applications Vol. 375, Dordrecht: Kluwer, 1996.
- [20] M. Fornasier, eds, *Theoretical Foundations and Numerical Methods for Sparse Recovery*, De Gruyter, 2010.
- [21] A. Frommer and P. Maass, [Fast cg-based methods for Tikhonov regularization](#), *SIAM Journal on Scientific Computing*, **5** (1999), 1831-1850.
- [22] M. Grasmair, [Well-posedness and convergence rates for sparse regularization with sublinear \$\ell^q\$ penalty term](#), *Inverse Problems and Imaging*, **3** (2009), 383-387.
- [23] M. Grasmair, M. Haltmeier and O. Scherzer, [Sparse regularization with \$\ell^q\$ penalty term](#), *Inverse Problems*, **24** (2008), 055020.
- [24] Q. Huang, R. Gong, Q. Jin and Y. Zhang, [A Tikhonov regularization method for Cauchy problem based on a new relaxation model](#), *Nonlinear Analysis: Real World Applications*, **74** (2023), 103935.
- [25] M. Hanke, A. Neubauer and O. Scherzer, [A convergence analysis of the Landweber iteration for nonlinear ill-posed problems](#), *Numerische Mathematik*, **72** (1995), 21-37.
- [26] K. Ito and B. Jin, *A Convergence Analysis of the Landweber Iteration for Nonlinear Ill-Posed Problems*, Series on Applied Mathematics Vol. 22, 2015.
- [27] B. Jin, D. A. Lorenz and S. Schiffer, [Elastic-net regularization: error estimates and active set methods](#), *Inverse Problems*, **25** (2009), 115022.
- [28] B. Jin and P. Maass, [Sparsity regularization for parameter identification problems](#), *Inverse Problems*, **28** (2012), 123001.
- [29] B. Jin, P. Maass and O. Scherzer, [Sparsity regularization in inverse problems](#), *Inverse Problems*, **33** (2017), 060301.
- [30] B. Jin, Y. Zhao and J. Zou, [Iterative parameter choice by discrepancy principle](#), *IMA Journal of Numerical Analysis*, **32** (2012), 1714-1732.
- [31] B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, De Gruyter, 2012.
- [32] Y. Li and D. W. Oldenburg, [3-D inversion of gravity data](#), *Geophysics*, **63** (1998), 109-119.
- [33] A. K. Louis, *Inverse und Schlecht Gestellte Probleme*, Teubner: Stuttgart, 1989.
- [34] P. Maass, S. V. Pereverzev, R. Ramlau and S. G. Solodky, [An adaptive discretization scheme for Tikhonov-regularization with a posteriori parameter selection](#), *Numerische Mathematik*, **87** (2001), 485-502.
- [35] V. A. Morozov, On the solution of functional equations by the method of regularization, *Soviet Math. Dokl.*, **7** (1966), 414-417.
- [36] V. A. Morozov, *Regularization Methods for Ill-Posed Problems*, CRC Press, Boca Raton, Florida, 1993.
- [37] R. Ramlau, [Morozov's discrepancy principle for Tikhonov regularization of nonlinear operators](#), *Numerical Functional Analysis and Optimization*, **23** (2002), 147-172.

- [38] N. A. B. Riis, A. M. A. Alghamdi, F. Uribe, S. L. Christensen, B. M. Afkham and P. C. Hansen, [Cuqipy: i. computational uncertainty quantification for inverse problems in python](#), *Inverse Problems*, **40** (2024), 045009.
- [39] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier and F. Lenzen, *Variational Methods in imaging*, Applied Mathematical Sciences Vol. 167, New York: Springer, 2009.
- [40] T. Scherzer, B. Kaltenbacher, B. Hofmann and K. Kazimierski, *Regularization Methods in Banach Spaces*, De Gruyter, 2012.
- [41] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston Wiley: New York, 1977.
- [42] A. N. Tikhonov, A. S. Leonov and A. G. Yagola, *Nonlinear Ill-Posed Problems*. London: Chapman & Hall, 1998.
- [43] W. Wang, S. Lu, H. Mao and J. Cheng, [Multi-parameter Tikhonov regularization with the \$\ell_0\$ sparsity constrain](#), *Inverse Problems*, **29** (2013), 065018.
- [44] S. Yang, M. Wang, P. Li, L. Jin, B. Wu and L. Jiao, [Compressive hyperspectral imaging via sparse tensor and nonlinear compressed sensing](#), *IEEE Transactions on Geoscience and Remote Sensing*, **53** (2015), 5943-5957.
- [45] Y. Zhang and C. Chen, [Stochastic linear regularization methods: Random discrepancy principle and applications](#), *Inverse Problems*, **40** (2024), 025007.

Received June 2025; revised September 2025; early access September 2025.