Applied Mathematics and Computation 503 (2025) 129499

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Full Length Article

Min-min minimization for the fractional ℓ_0-regularized problem

Jun Wang [a], Qiang Ma [b,*], Cheng Zhou [c]

[a] School of Science, Jiangsu University of Science and Technology, Zhenjiang 212100, Jiangsu, China
[b] School of Materials Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212100, Jiangsu, China
[c] Tencent Robotics Y, Shenzhen 518054, Guangdong, China

ARTICLE INFO

MSC:
65K10
68W40
90C26
90C46
90C90

Keywords:
Compressed sensing
ℓ_0-regularization minimization
Fractional nonlinear programming
Quadratic programs with cardinality constraints
Sparse recovery

ABSTRACT

In this paper, we present a novel unconstrained fractional ℓ_0 regularization (FL0R) model to solve cardinality minimization. Firstly, we construct an interesting $\min-\min$ minimization from FL0R by introducing a middle variable of sparsity. Then, we prove that the solution to $\min-\min$ minimization with a given sparsity is one of FL0R. Finally, some numerical examples are presented to illustrate the effectiveness and validity of the new model.

1. Introduction

Compressed Sensing (CS) [1–4] introduced approximately eighteen years ago, has been extensively investigated by many scholars from the fields of optimization theory [5], signal processing [6], and practical engineering [7]. These researchers have progressively deepened their understanding and expanded the theory's application, addressing real-world problems and thereby enriching both the content and the foundational research framework.

The fundamental problem of CS is to recover the sparse signal $\mathbf{x}^o \in \mathbb{R}^n$ from a limited set of corrupted measurements $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e} \in \mathbb{R}^m$, where the coding matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ of full-row rank and $\mathbf{e}$ is an unknown vector of errors with $\|\mathbf{e}\|_2 \le \epsilon$, where ϵ is the error term. Mathematically, we consider the following ℓ_0-minimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{1.1}$$

where the quasinorm $\|\mathbf{x}\|_0 = |\operatorname{supp}(\mathbf{x})|$ measures the number of non-zero elements in $\mathbf{x}$. Although problem (1.1) is a NP-hard [8] combinatorial optimization, it can be solved effectively by greedy algorithms, including orthogonal matching pursuit (OMP) [9], compressive sampling matching pursuit (CoSaMP) [10] and others [11,12]. A seminal work for the cardinality minimization (1.1) was completed by Candes, Donoho and Tao [1–4], which established the equivalence between problem (1.1) and the following ℓ_1 minimization

* Corresponding author.
E-mail address: qma1989-2019@just.edu.cn (Q. Ma).

https://doi.org/10.1016/j.amc.2025.129499
Received 10 October 2024; Received in revised form 22 April 2025; Accepted 24 April 2025

Available online 2 May 2025

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{1.2}$$

with the conditions that the sensing matrix $\mathbf{A}$ meets specific conditions, such as Restricted Isometry Property (RIP) [1] and Null Space Property (NSP) [3]. This pivotal finding has become the cornerstone of CS and continues to be developed [11,12].

In fact, many researchers have directly considered the problem (1.1) and have produced some results [13,14,19,20]. Introducing additional parameter λ, one commonly considers the ℓ_0-regularized minimization

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda\|\mathbf{x}\|_0, \tag{1.3}$$

where the regularization parameter λ balances the data fitting $\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$ and the sparseness $\|\mathbf{x}\|_0$. Many efficient algorithms including iterative hard thresholding (IHT) [21], accelerated iterative hard thresholding (AIHT) [22,23] and the relaxed optimal k-thresholding Pursuit (ROTP) [24] have been developed to solve the problem (1.3). However, the parameter λ not only complicates the theoretical analysis of problem (1.3) but also necessitates their adaptation and optimization to various data types, thus impacting the adaptability of problem (1.3). Hence, a fundamental question arises: how can problem (1.1) be transformed into an unconstrained optimization problem without additional parameters or conditions?

Fortunately, the paper introduces a novel fractional ℓ_0 regularization (`FL0R`) problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2}{(n+1) - \|\mathbf{x}\|_0}, \tag{1.4}$$

which is inspired by the ratio model [25,26]. Unlike problem (1.3), `FL0R` (1.4) presents as a parameter-free unconstrained optimization problem, which poses significantly greater analytical difficulties.

Intuitively, one may obtain from the following inequality

$$\frac{1}{n+1}\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \le \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \le \min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2,$$

which is derived from $1 \le (n+1) - \|\mathbf{x}\|_0 \le n+1$ and

$$\frac{1}{n+1}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \le f(\mathbf{x}) \le \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$$

that the optimal value of `FL0R` is 0 when $\mathbf{x} = \mathbf{A}^+\mathbf{b}$, where $\mathbf{A}^+$ is the generalized Moore-Penrose inverse matrix and

$$f(\mathbf{x}) = \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2}{(n+1) - \|\mathbf{x}\|_0}, \forall \mathbf{x} \in \mathbb{R}^n.$$

However, it is critical to note that the solution to problem $\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ does not promote sparsity and is meaningless in the real application of the problem. For instance, if taking

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix},$$

then we obtain that $\frac{1}{3} \le \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \le 1$. In fact, our goal is to derive a model that guarantees a sparse solution. However, considering the nature of real-world applications, the data fit term cannot be achieved 0 due to inevitable errors. This raises a pertinent question: can the solution to problem `FL0R` indeed be characterized by sparsity?

The paper presents the exact answer to the aforementioned question. Furthermore, based on `FL0R` (1.4), we derive a novel and interesting min-min minimization (2.1). Meanwhile, the derived problem (2.1) allows us to approximate the sparsity level of the original signal without prior knowledge of its exact value. Furthermore, problem (2.1) is solved by some efficient algorithms, including Newton's hard-throwing pursuit (`NHTP`) [19] and Newton method for ℓ_0-regularized (`NL0R`) [20], and our research extends the results of CS without unknown sparsity level of the signal.

The remaining parts of this paper are organized as follows. In Section 2, we derive the min-min problem through a detailed analysis of the fractional regularization problem. Subsequently, we attempt to establish the relationship between the derived problem and the original problem. Finally, we use several examples to demonstrate the validity of the results.

2. Main results

Before providing our results, we first give some notations used throughout this manuscript. Let $\mathbb{R}^n$ be the set of real vectors of size n and $\mathbb{R}^n_{++}$ the set of positive vectors. All real vectors will be column vectors unless otherwise noted. Matrices are bold capital, vectors are bold lower-cases, and scalars or entries are not bold. For instance, $\mathbf{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ is a vector and x_i its i-th component.

Assume that $1 \le \kappa \le n+1$ and define

$$\Omega_\kappa \triangleq \left\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_0 \le n+1-\kappa\right\}.$$

It follows from $n+1-\|\mathbf{x}\|_0 \ge \kappa$ that

$$\frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0} \le \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{\kappa}$$

and therefore

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0} \le \min_{\mathbf{x}\in\Omega_\kappa} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0} \le \min_{\mathbf{x}\in\Omega_\kappa} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{\kappa}.$$

Both sides of the above inequality are minimized with respect to the variable $1 \le \kappa \le n+1$ gives

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) = \min_{1\le\kappa\le n+1} \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \le \min_{1\le\kappa\le n+1} \min_{\mathbf{x}\in\Omega_\kappa} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{\kappa}.$$

Hence, we consider the following unconstrained min – min problem

$$\min_{1\le\kappa\le n+1} \varphi(\kappa) = \min_{1\le\kappa\le n+1} \left(\min_{\|\mathbf{x}\|_0\le n+1-\kappa} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{\kappa} \right), \tag{2.1}$$

where

$$\varphi(\kappa) \triangleq \min_{\|\mathbf{x}\|_0\le n+1-\kappa} \frac{\|\mathbf{Ax}-\mathbf{b}\|_2^2}{\kappa}.$$

Remark 2.1. The parameter-free and unconstrained optimization problem (1.4) is only relative to ℓ_0 regularization (1.3). In addition, FLOR (1.4) is equivalent to minimizing the natural logarithm of $f(\mathbf{x})$, i.e.,

$$\min_{\mathbf{x}\in\mathbb{R}^n} \ln(f(\mathbf{x})) = \ln\left(\|\mathbf{Ax}-\mathbf{b}\|_2^2\right) - \ln\left(n+1-\|\mathbf{x}\|_0\right), \tag{2.2}$$

which is the parameter-free unconstrained optimization. However, problem (2.2) seems to be more complex and difficult to solve than (1.3).

Despite problem (1.4) being very hard to solve, there exist some algorithms including NHTP [19] and NL0R [20] solving the sub-problem

$$\min_{\|\mathbf{x}\|_0\le n+1-\kappa} \|\mathbf{Ax}-\mathbf{b}\|_2^2$$

of the min-min minimization (2.1). Obviously, here is the question: what is the connection between FL0R and problem (2.1)? The subsequent Theorem 2.1 demonstrates the relationship between the solutions of these two problems.

Theorem 2.1. *Assume that $\widehat{\mathbf{x}} \in \mathbb{R}^n$ and $1 \le \kappa^* \le n$ are the optimal solutions of problems (1.4) and (2.1), respectively. In addition, suppose that the following problem*

$$\min_{\|\mathbf{x}\|_0\le n+1-\kappa^*} \|\mathbf{Ax}-\mathbf{b}\|_2^2 = \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 \tag{2.3}$$

possesses a unique optimal solution $\bar{\mathbf{x}}$. Then the following hold.

(i) $f(\bar{\mathbf{x}}) = f(\widehat{\mathbf{x}})$;
(ii) $\|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$.

Proof. We will prove this theorem by analyzing the sparsity of solution $\|\widehat{\mathbf{x}}\|_0$ which is categorized based on the value of $n+1-\kappa^*$.

(C1) Suppose that $n \ge \|\widehat{\mathbf{x}}\|_0 > n+1-\kappa^*$, then

$$1 \le n+1-\|\widehat{\mathbf{x}}\|_0 < n+1-(n+1-\kappa^*) = \kappa^*$$

and hence

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} > \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}. \tag{2.4}$$

In addition, we have that

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*} \ge \frac{\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0}, \tag{2.5}$$

which is due to

$$\frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} = \min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0} \le \min_{1\le\kappa\le n+1} \varphi(\kappa) = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}. \tag{2.6}$$

Combining the inequality (2.5) and (2.4) gives

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} > \frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0}$$

With the help of $1 \le n+1-\|\hat{\mathbf{x}}\|_0 < \kappa^*$, simplifying the above inequality yields

$$\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 > \|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2. \tag{2.7}$$

Denote

$$\Omega_{n+1-\|\hat{\mathbf{x}}\|_0} = \left\{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}\|_0 \le \|\hat{\mathbf{x}}\|_0\right\}$$

and

$$\Omega_{\kappa^*} = \left\{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}\|_0 \le n+1-\kappa^*\right\}.$$

Obviously, we have $\Omega_{\kappa^*} \subseteq \Omega_{n+1-\|\hat{\mathbf{x}}\|_0}$ and thus

$$\bar{\mathbf{x}}, \hat{\mathbf{x}} \in \Omega_{n+1-\|\hat{\mathbf{x}}\|_0}.$$

Now, assume that $\tilde{\mathbf{x}}\in\mathbb{R}^n$ is the optimum solution of the following optimization problem

$$\min_{\mathbf{x}\in\Omega_{n+1-\|\hat{\mathbf{x}}\|_0}} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2 = \|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2. \tag{2.8}$$

For this solution $\tilde{\mathbf{x}}$, we assert that $\tilde{\mathbf{x}} \in \Omega_{n+1-\|\hat{\mathbf{x}}\|_0}$ and $\tilde{\mathbf{x}} \notin \Omega_{\kappa^*}$. In fact, if $\tilde{\mathbf{x}} \in \Omega_{\kappa^*}$, then it follows from (2.3) that

$$\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 \le \|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2.$$

As a result of the fact that $\tilde{\mathbf{x}}$ is solution of problem (2.8), we have the another inequality

$$\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 \le \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2.$$

Consequently, the following equality

$$\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 = \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2$$

holds. However, because of $\hat{\mathbf{x}} \in \Omega_{n+1-\|\hat{\mathbf{x}}\|_0}$, we obtain from problem (2.8) that

$$\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2 \ge \|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 = \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2,$$

which contradicts the inequality (2.7). Thus, we obtain that

$$\tilde{\mathbf{x}} \in \Omega_{n+1-\|\hat{\mathbf{x}}\|_0} - \Omega_{\kappa^*},$$

which means

$$n+1-\kappa^* < \|\tilde{\mathbf{x}}\|_0 \le \|\hat{\mathbf{x}}\|_0.$$

Additionally, the following inequality

$$\kappa^* > n+1-\|\tilde{\mathbf{x}}\|_0 \ge n+1-\|\hat{\mathbf{x}}\|_0 \ge 1$$

gives

$$\frac{\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} \ge \frac{\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\tilde{\mathbf{x}}\|_0}. \tag{2.9}$$

Since that $\hat{\mathbf{x}}\in\mathbb{R}^n$ is the optimal solution of problem (1.4), we obtain

$$\frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} \le \frac{\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\tilde{\mathbf{x}}\|_0}$$

and therefore from (2.9) that

$$\frac{\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} \ge \frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0},$$

which can be reduced to a simple form

$$\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 \geq \|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2. \tag{2.10}$$

Owing to the fact that $\hat{\mathbf{x}}$ is the minimum of problem (2.8), it follows that

$$\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 \leq \|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2$$

and further

$$\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2 = \|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2, \tag{2.11}$$

which comes from (2.10).

By the equality (2.11), calculate the value of φ at the point $n+1-\|\hat{\mathbf{x}}\|_0$,

$$\begin{aligned}\varphi(n+1-\|\hat{\mathbf{x}}\|_0) &= \frac{1}{n+1-\|\hat{\mathbf{x}}\|_0}\min_{\|\mathbf{x}\|_0\leq\|\hat{\mathbf{x}}\|_0}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2\\ &= \frac{1}{n+1-\|\hat{\mathbf{x}}\|_0}\|\mathbf{A}\tilde{\mathbf{x}}-\mathbf{b}\|_2^2\\ &= \frac{1}{n+1-\|\hat{\mathbf{x}}\|_0}\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2 = f(\hat{\mathbf{x}}).\end{aligned}$$

Significantly, we have known from the above inequality, (2.4) and (2.6) that

$$\begin{aligned}f(\hat{\mathbf{x}}) = \frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} &= \min_{\mathbf{x}\in\mathbb{R}^n}\frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0}\\ &\leq \varphi(\kappa^*) = \min_{1\leq\kappa\leq n+1}\varphi(\kappa) = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}\\ &\leq \varphi(n+1-\|\hat{\mathbf{x}}\|_0) = f(\hat{\mathbf{x}})\end{aligned}$$

which implies that

$$\begin{aligned}f(\hat{\mathbf{x}}) = \frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0} &= \varphi(\kappa^*) = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}\\ &= \frac{n+1-\|\bar{\mathbf{x}}\|_0}{\kappa^*}\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0}\\ &\geq \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0} = f(\bar{\mathbf{x}})\\ &\geq f(\hat{\mathbf{x}})\end{aligned} \tag{2.12}$$

where the last inequality holds because that $\hat{\mathbf{x}}$ is the optimal solution of problem (1.4) and the preceding inequality comes from $\|\bar{\mathbf{x}}\|_0 \leq n+1-\kappa^*$. As shown above in (2.12), we acquire $f(\bar{\mathbf{x}}) = f(\hat{\mathbf{x}})$ which ensures (i) and moreover from

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*} = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0}$$

that $\|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$ which verifies (ii).

(C2) Suppose that $1 \leq \|\hat{\mathbf{x}}\|_0 \leq n+1-\kappa^*$. By the fact that $\hat{\mathbf{x}}$ is a feasible solution to problem (2.3), so we have

$$\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2 \geq \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2. \tag{2.13}$$

(C2.1) If the solution $\bar{\mathbf{x}}$ satisfies $1 \leq \|\bar{\mathbf{x}}\|_0 \leq \|\hat{\mathbf{x}}\|_0 \leq n+1-\kappa^*$, then

$$\frac{1}{n+1-\|\bar{\mathbf{x}}\|_0} \leq \frac{1}{n+1-\|\hat{\mathbf{x}}\|_0} \leq \frac{1}{\kappa^*}. \tag{2.14}$$

Moreover, combining problem (2.3) and (2.13) leads

$$\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 = \min_{\mathbf{x}\in\Omega_{n+1-\|\hat{\mathbf{x}}\|_0}}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2.$$

It follows from the above equality and inequality (2.14) that

Table 1
Here, $\widehat{\mathbf{x}} \in \mathbb{R}^n$ and $\bar{\mathbf{x}}$ are the solutions of problems (1.4) and (2.3), respectively. Note that $\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2 = \min_{\|\mathbf{x}\|_0 \le \|\widehat{\mathbf{x}}\|_0} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$ and $\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 = \min_{\|\mathbf{x}\|_0 \le \|\bar{\mathbf{x}}\|_0} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$.

Greater than	Equal to	Less than
$\|\widehat{\mathbf{x}}\|_0 > \|\bar{\mathbf{x}}\|_0$	$\|\widehat{\mathbf{x}}\|_0 = \|\bar{\mathbf{x}}\|_0$	$\|\widehat{\mathbf{x}}\|_0 < \|\bar{\mathbf{x}}\|_0$
$\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2 < \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2$	$\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2 = \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2$	$\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2 > \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2$
$f(\widehat{\mathbf{x}}) = f(\bar{\mathbf{x}})$	$f(\widehat{\mathbf{x}}) = f(\bar{\mathbf{x}})$	$f(\widehat{\mathbf{x}}) = f(\bar{\mathbf{x}})$
$\|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$	$\|\widehat{\mathbf{x}}\|_0 = \|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$	$\|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$

$$\begin{aligned} f(\widehat{\mathbf{x}}) &= \frac{\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} \ge \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} = \varphi(n+1-\|\widehat{\mathbf{x}}\|_0) \\ &= \frac{\min_{\|\mathbf{x}\|_0 \le \|\widehat{\mathbf{x}}\|_0} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} \ge \varphi(\kappa^*) = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*} \\ &\ge \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2 \frac{1}{n+1-\|\bar{\mathbf{x}}\|_0} = f(\bar{\mathbf{x}}) \\ &\ge \min_{\mathbf{x}\in\mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2}{(n+1)-\|\mathbf{x}\|_0} = \frac{\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} = f(\widehat{\mathbf{x}}) \ge \varphi(\kappa^*) \end{aligned} \tag{2.15}$$

and therefore $\varphi(\kappa^*) \ge f(\bar{\mathbf{x}}) \ge \varphi(\kappa^*)$ which gives

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0} = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}.$$

We obtain from the above inequality that $n+1-\|\bar{\mathbf{x}}\|_0 = \kappa^*$ which means that the case (ii) holds. Additionally, it follows from (2.15) that $f(\widehat{\mathbf{x}}) \le f(\bar{\mathbf{x}}) \le f(\widehat{\mathbf{x}})$ which confirms (i).

(C2.2) If the solution $\bar{\mathbf{x}}$ satisfies $1 \le \|\widehat{\mathbf{x}}\|_0 \le \|\bar{\mathbf{x}}\|_0 \le n+1-\kappa^*$, then

$$\frac{1}{n+1-\|\widehat{\mathbf{x}}\|_0} \le \frac{1}{n+1-\|\bar{\mathbf{x}}\|_0} \le \frac{1}{\kappa^*}. \tag{2.16}$$

Now, let us consider again the solution $\check{\mathbf{x}}$ of the problem (2.8). Note that

$$\|\mathbf{A}\check{\mathbf{x}}-\mathbf{b}\|_2^2 = \min_{\mathbf{x}\in\Omega_{n+1-\|\widehat{\mathbf{x}}\|_0}} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2,$$

where $1 \le \|\check{\mathbf{x}}\|_0 \le \|\widehat{\mathbf{x}}\|_0 \le \|\bar{\mathbf{x}}\|_0 \le n+1-\kappa^*$ in this case of (C2.2). With the aid of $\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2 \ge \|\mathbf{A}\check{\mathbf{x}}-\mathbf{b}\|_2^2 \ge \|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2$, we show that

$$\begin{aligned} f(\widehat{\mathbf{x}}) = \frac{\|\mathbf{A}\widehat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} &\ge \frac{\|\mathbf{A}\check{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} = \frac{\min_{\|\mathbf{x}\|_0 \le \|\widehat{\mathbf{x}}\|_0} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2}{n+1-\|\widehat{\mathbf{x}}\|_0} \\ &= \varphi(n+1-\|\widehat{\mathbf{x}}\|_0) \ge \varphi(\kappa^*) = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*} \\ &\ge \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0} = f(\bar{\mathbf{x}}) \ge f(\widehat{\mathbf{x}}), \end{aligned}$$

where the penultimate inequality comes from (2.16), which affirms (i). And then, the equality

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*} = \frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0}$$

can deduce that $\|\bar{\mathbf{x}}\|_0 = n+1-\kappa^*$ which verifies (ii).

To conclude, based on the preceding discussions of the two scenarios of (C1) and (C2), we can establish both the result (i) and (ii), thus completing the proof. □

Theorem 2.1 estimates the relationship between problem (2.1) and (1.4). More specifically, it does demonstrate that the optimal sparse solution $\bar{\mathbf{x}}$ to problem (2.3) is also a solution of problem (1.4), while it does not ensure that the optimal $\widehat{\mathbf{x}}$ is sparse. Table 1 summarizes the detailed results of the cases in Theorem 2.1. However, Table 1 does not provide additional information about κ^*. Note that the issue of κ^*, which is a matter of great concern, deserves further consideration.

Remark 2.2. From the proof of Theorem 2.1, we obtain an intriguing case of $1 \le \|\widehat{\mathbf{x}}\|_0 \le n+1-\kappa^*$. If the solution $\bar{\mathbf{x}}$ satisfies $1 \le \|\bar{\mathbf{x}}\|_0 \le \|\widehat{\mathbf{x}}\|_0 \le n+1-\kappa^*$, it follows from (2.15) that $f(\widehat{\mathbf{x}}) = \varphi(\kappa^*) = f(\bar{\mathbf{x}})$ which implies

$$\frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0}=\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}=\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0}. \tag{2.17}$$

We obtain from (2.13) and (2.17) that

$$\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0}\le\frac{\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\hat{\mathbf{x}}\|_0}=\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{\kappa^*}=\frac{\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2}{n+1-\|\bar{\mathbf{x}}\|_0}$$

and hence $\|\hat{\mathbf{x}}\|_0\le\|\bar{\mathbf{x}}\|_0$. As a consequence, we have

$$\|\hat{\mathbf{x}}\|_0=\|\bar{\mathbf{x}}\|_0=n+1-\kappa^*$$

and $\|\mathbf{A}\bar{\mathbf{x}}-\mathbf{b}\|_2^2=\|\mathbf{A}\hat{\mathbf{x}}-\mathbf{b}\|_2^2$ with the help of $1\le\|\bar{\mathbf{x}}\|_0\le\|\hat{\mathbf{x}}\|_0\le n+1-\kappa^*$.

Hence, if $1\le\|\hat{\mathbf{x}}\|_0\le n+1-\kappa^*$, then we have

$$1\le\|\hat{\mathbf{x}}\|_0\le\|\bar{\mathbf{x}}\|_0\le n+1-\kappa^*,$$

and further we can determine an upper bound $\|\bar{\mathbf{x}}\|_0$ on the sparsity level of the solution of problem (1.4).

Theorem 2.1 indicates that the solution to the problem (2.1) is sparse and concurrently serves as a solution to (1.4). A significant advantage of this finding is the comparative simplicity with which (2.1) can be resolved compared to (1.4).

3. Numerical results

In this section, we present some numerical results to validate the correctness and rationality of Theorem 2.1. It is worth stating that all test code was written and tested in MATLAB R2023b running on the Debian 12.2 (x86_64 6.1.0-13-amd64) with AMD Ryzen 9 7950X (32) (4.5 GHz) and 128 (4×32) GB of UDIMM memory. Floating point arithmetic was done with 64 bits, i.e., machine epsilon is about 2.22×10^{-16}.

First, we examine whether the function $\varphi(\kappa)$ is monotonic. In fact, if $\varphi(\kappa)$ increases or decreases monotonically, then we obtain the minimum value of $\varphi(\kappa)$ at the point $\kappa=1$ or $\kappa=n$. If $1\le\kappa_1<\kappa_2\le n$, then it follows from $n+1-\kappa_2<n+1-\kappa_1$ that

$$\min_{\|\mathbf{x}\|_0\le n+1-\kappa_2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2\ge\min_{\|\mathbf{x}\|_0\le n+1-\kappa_1}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$$

and hence

$$\begin{aligned}\varphi(\kappa_2)=\frac{1}{\kappa_2}\min_{\|\mathbf{x}\|_0\le n+1-\kappa_2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2&\ge\left(\frac{\kappa_1}{\kappa_2}\right)\left(\frac{1}{\kappa_1}\min_{\|\mathbf{x}\|_0\le n+1-\kappa_1}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2\right)\\&=\frac{\kappa_1}{\kappa_2}\varphi(\kappa_1).\end{aligned}$$

Furthermore, we have

$$\frac{\varphi(\kappa_2)-\varphi(\kappa_1)}{\kappa_2-\kappa_1}\ge\frac{\frac{\kappa_1}{\kappa_2}\varphi(\kappa_1)-\varphi(\kappa_1)}{\kappa_2-\kappa_1}=-\frac{\varphi(\kappa_1)}{\kappa_2}$$

and therefore the monotonicity of $\varphi(\kappa)$ can't be guaranteed. Therefore, we need to explore other algorithms to solve the problem (2.1).

To deal with problem

$$\min_{1\le\kappa\le n+1}\frac{1}{\kappa}\left(\min_{\|\mathbf{x}\|_0\le n+1-\kappa}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2\right),$$

it is necessary to search another trick. First, we consider the following problem

$$\min_{\|\mathbf{x}\|_0\le s}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2 \tag{3.1}$$

where the integer $1\le s\le n$, which is equivalent to the following non-convex mixed-integer quadratically constrained quadratic program (MIQCQP) [13,14]

$$\begin{aligned}&\min_{\mathbf{x}\in\mathbb{R}^n}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2\\&\text{s.t.}\begin{cases}\mathbf{1}_n^\top\mathbf{y}\ge n-s,\\\mathbf{x}\odot\mathbf{y}=0,\\\mathbf{y}\in\{0,1\}^n,\end{cases}\end{aligned} \tag{3.2}$$

where $\odot$ denotes the Hadamard product, i.e., $\mathbf{x}\odot\mathbf{y}=0$ means that $x_iy_i=0$, the auxiliary variable vector $\mathbf{y}\in\{0,1\}^n$ of entries y_i is binary variable with the indices $i=1,2,\cdots,n$ and $\mathbf{1}_n=(1,1,\cdots,1)^\top\in\mathbb{R}^n$ is the vector of ones.

Table 2
The first row shows that International domain name abbreviation for countries and regions in the world. The data representation below the first row denotes the sensing matrix $\mathbf{A} \in \mathbb{R}^{12\times 12}$ to be processed.

BE	BR	CN	CU	EG	FR	IN	IL	US	SU	YU	ZR
0.00	5.58	7.00	7.08	4.83	2.17	6.42	3.42	2.50	6.08	5.25	4.75
5.58	0.00	6.50	7.00	5.08	5.75	5.00	5.50	4.92	6.67	6.83	3.00
7.00	6.50	0.00	3.83	8.17	6.67	5.58	6.42	6.25	4.25	4.50	6.08
7.08	7.00	3.83	0.00	5.83	6.92	6.00	6.42	7.33	2.67	3.75	6.67
4.83	5.08	8.17	5.83	0.00	4.92	4.67	5.00	4.50	6.00	5.75	5.00
2.17	5.75	6.67	6.92	4.92	0.00	6.42	3.92	2.25	6.17	5.42	5.58
6.42	5.00	5.58	6.00	4.67	6.42	0.00	6.17	6.33	6.17	6.08	4.83
3.42	5.50	6.42	6.42	5.00	3.92	6.17	0.00	2.75	6.92	5.83	6.17
2.50	4.92	6.25	7.33	4.50	2.25	6.33	2.75	0.00	6.17	6.67	5.67
6.08	6.67	4.25	2.67	6.00	6.17	6.17	6.92	6.17	0.00	3.67	6.50
5.25	6.83	4.50	3.75	5.75	5.42	6.08	5.83	6.67	3.67	0.00	6.92
4.75	3.00	6.08	6.67	5.00	5.58	4.83	6.17	5.67	6.50	6.92	0.00

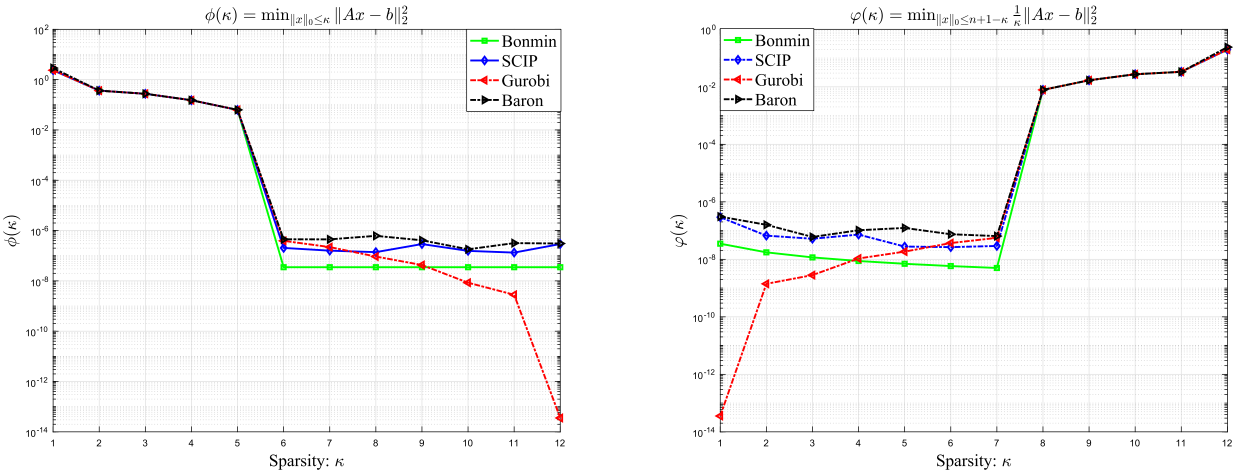

Fig. 1. Plots of the solvers Bonmin, SCIP and Gurobi, Baron with the given data $m = 12$, $n = 12$ and the sparsity $s = 6$ of original signal.

Given the non-convexity of mixed integer problem (3.2), some algorithms including the commercial solvers BARON 24.5.8 [15], GUROBI 11 and open source solvers BONMIN 1.8.8, SCIP 9.0, can quickly attain the optimal solution when the size of the problem is small ($n \le 50$). However, as the size of the problem grows, these four methods are unable to find the optimal solution within a predetermined time (≤ 3600 seconds). Consequently, a necessity arises to further explore efficiency algorithms.

To illustrate the correctness and validity of Theorem 2.1, we provide a practical example solved by the equivalent problem (3.2).

Example 3.1. [16, Experiment 1] Here, to solve problem (2.1) with the sensing matrix $\mathbf{A}$ of the countries data set, which is available at Countries.data and fully described in [17]. The approach employed to obtain the dissimilarity coefficients was to administer a questionnaire in a political science class, wherein students were tasked with evaluating subjective dissimilarities between 12 countries. The resulting average ratings provided by the students were utilized as the final dissimilarity coefficients in Table 2.

We initiate the process by assigning the original signal

$$\mathbf{x}^{\mathrm{o}} = (0,0,1,0,0,0,1,1,0,1,1,1)^{\top} \in \mathbb{R}^{12},$$

followed by setting $\mathbf{b} = \mathbf{A}\mathbf{x}^{\mathrm{o}} + \sigma\mathbf{e}$, where $\mathbf{e} \in \mathbb{R}^{12}$ is an unit vector of i.i.d. standard Gaussian entries and the error term $\sigma = 10^{-3}$. Note that $\|\mathbf{x}^{\mathrm{o}}\|_0 = 6 = s$. Then, we calculate

$$\phi(\kappa) = \|\mathbf{A}\mathbf{x}^{\kappa} - \mathbf{b}\|_2^2, \ \kappa = 1, 2, \cdots, 12,$$

where $\mathbf{x}^{\kappa}$ is the solution of problem (3.1), and therefore obtain

$$\varphi(\kappa) = \frac{\phi(13-\kappa)}{\kappa}, \ \kappa = 1, 2, \cdots, 12.$$

Finally, the scatter plot of data generated by $\phi(\kappa)$ and $\varphi(\kappa)$ are displayed separately, as shown in Fig. 1.

The right of Fig. 1 demonstrates that the optimal solution to the problem (2.1) is achieved at $\kappa^* = 7$, excluding the results of the solver GUROBI. By the (ii) of Theorem 2.1, we have that $\|\bar{\mathbf{x}}\|_0 = 13 - 7 = 6$ which is the same as the sparsity $s = 6$ of the raw signal, where the solution to problem (3.1) of sparsity 6 is

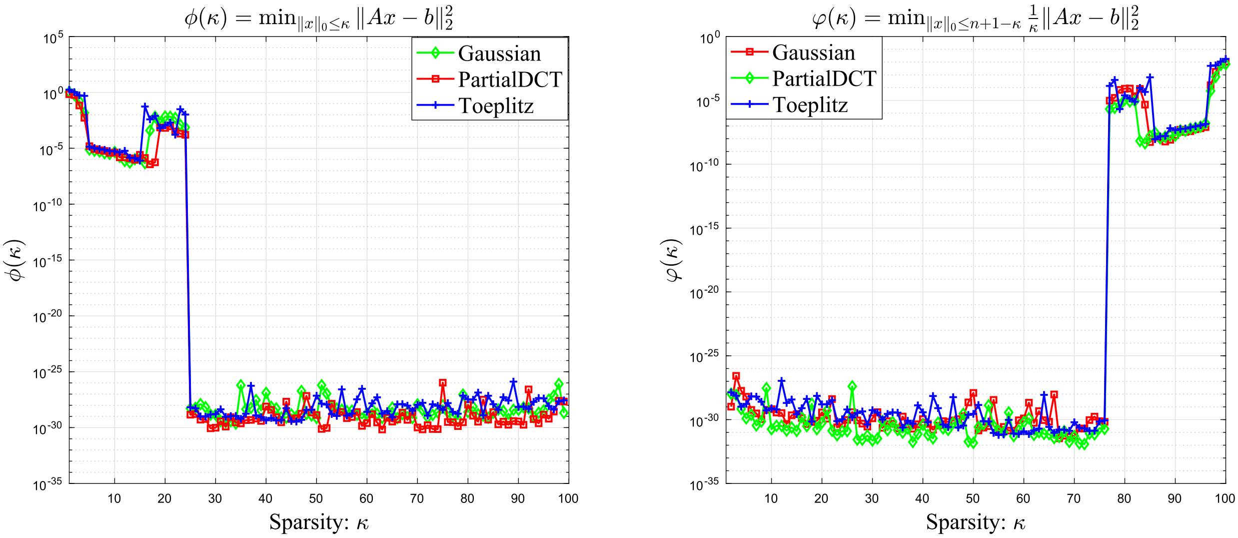

Fig. 2. Plots of the algorithm NHTP with the given data $m = 25$, $n = 100$ and the sparsity $s = 5$ of original signal.

$$\bar{\mathbf{x}} = (0, 0, 0.998962, 0, 0, 0, 1.00008, 0.999037, 0, 1.00204, 0.999239, 1.00093)^\top \in \mathbb{R}^{12}.$$

Moreover, the behavior of the solver GUROBI differs from that of the other three solvers BARON, BONMIN and SCIP, and we speculate on the possible reasons for this discrepancy.

(1) The non-convex mixed-integer nonlinear nature of the MIQCQP (3.2) may lead GUROBI to implement heuristics that favor convergence to the least squares solution;
(2) If the sparsity s is from 7 to 12, the original signal's support set identified by GUROBI results in non-zero elements within the previously unsupported set, steering the solution towards the least squares solution. This phenomenon is particularly relevant because the problem effectively becomes a system of nonlinear equations $\mathbf{Ax} = \mathbf{b}$, which is the sufficient and necessary optimality condition of

$$\min_{\|\mathbf{x}\|_0 \le 12} \|\mathbf{Ax} - \mathbf{b}\|_2^2 = \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|_2^2.$$

Compared to the unconstrained optimization problem (1.4), problem (2.1) appears to be somewhat simpler to resolve. However, if the dimension $n \ge 100$, then the non-convex MIQCQP (3.2) is not easy to be solved by these four solvers. Fortunately, some literatures [16,18–20] present several rapid and efficient algorithms for solving problem (2.1). Here, we adopt the algorithm NHTP which is quadratically convergent under the standard assumption of restricted strong convexity, for solving problem (3.1) as described in [19, Table 1: Framework of NHTP, pp. 13].

To evaluate the capacity of model (2.1) for achieving sparse solutions, we conducted numerical testing using three types of sensing matrices generated by the subfunction

```
compressed_sensing_data(pb, m, n, s, nf)
```

in NHTP [19] with $\mathtt{nf} = 10^{-3}$ and $\mathtt{pb} \in \{\mathtt{Gaussina}, \mathtt{PartialDCT}, \mathtt{Toeplitz}\}$.

(1) The type Guassian denotes that A be a random Gaussian matrix with each element a_{ij} being identically and independently generated from the standard normal distribution.
(2) The type Partial DCT denotes that A be a random partial discrete cosine transform (DCT) matrix generated by

$$a_{ij} = \cos\left(2\pi(j-1)\epsilon_i\right),\ i = 1, 2, \cdots, m\ j = 1, 2, \cdots, n,$$

where ϵ_i is uniformly and independently sampled from $[0, 1]$.
(3) The type Toeplitz Correlation denotes that A be a matrix of

$$A = C\,D,$$

where C is the random Gaussian matrix and $D = \mathrm{Real}\left(\mathrm{G}^{1/2}\right)$ in which the matrix G with each element of $g_{ij} = 0.5^{|i-j|}$.

The results are presented in Figs. 2 and 3.

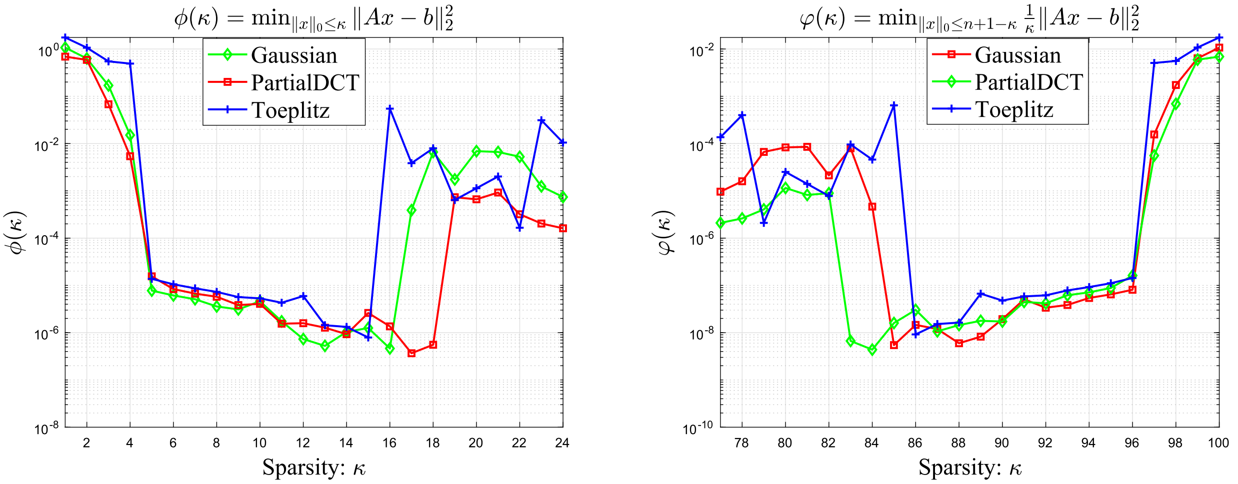

Fig. 3. The truncated plots of the algorithm NHTP with the given data $m = 25$, $n = 100$ and the sparsity $s = 5$ of original signal.

Table 3
The support of the recovery solved by NHTP with three types of sensing matrices, where 6.1233^* means 6.1233×10^{-6} and $1.5544^\#$ denotes 1.5544×10^{-5}. Note that supp($\mathbf{x}^o$) is the support of raw signal $\mathbf{x}^o$ and the abbreviations for Gaussina, PartialDCT, Toeplitz are G, P, T, respectively.

supp($\mathbf{x}^o$)			G			P			T		
G	P	T	95	96	97	95	96	97	95	96	97
26	-	1	26	26	26	10	-	-	1	1	1
-	17	35	36	-	-	17	17	17	35	35	-
57	60	45	57	57	-	60	60	60	45	45	-
61	69	56	61	61	61	69	69	69	56	56	56
66	79	-	66	66	66	79	79	-	-	-	58
76	89	-	76	76	76	89	89	89	91	-	-
-	-	95	-	-	-	-	-	-	95	95	95
Objvals			6.1233^*	6.7602^*	0.0151	8.3511^*	$1.5544^\#$	0.0054	$1.0549^\#$	$1.3762^\#$	0.4912

As depicted in the right of Fig. 2, the curve exhibits three significant jumps from right to left. The first jump occurs at the sparsity $\kappa = 95$, and the second at $\kappa = 90$. The most pronounced drop occurs at $\kappa = 77$, beyond which the value falls below the precision of the machine $\epsilon = 2.2204 \times 10^{-16}$ of the computer and is effectively zero. Consequently, it is crucial to examine the details of the first two stages, as illustrated in Fig. 3.

To deal with problem (1.4), we developed its equivalent model (2.1). However, solving problem (2.1) entails traversing the sparsity range of $\kappa \in \{1, 2, \cdots, n\}$, which is not only time-consuming but also involves substantial computational redundancy. As indicated in Fig. 3 and Table 3, we can obtain the solution to problem (2.1) from $\kappa = 1$, while calculating its objective value until an abnormal value is observed at κ^*. Ultimately, the sparsity of the original signal is estimated by

$$s = n + 1 - \kappa^*.$$

Specifically, we obtain from Fig. 3 and Table 3 that $\kappa^* = 96$ and therefore

$$s = n + 1 - \kappa^* = 100 + 1 - 96 = 5.$$

4. Conclusion

To eliminate the effect of the ℓ_0 regularization parameter λ in problem (1.3), we have the pleasure of introducing a parameter-free fractional ℓ_0 regularization problem (1.4). Subsequently, we demonstrate that this novel problem (1.4) admits an optimal solution with sparsity $s = n + 1 - \kappa^*$ in Theorem 2.1. In particular, the challenging fractional problem (1.4) is transformed into a tractable min-min problem (2.1) solved by several efficient algorithms [16,18–20].

CRediT authorship contribution statement

Jun Wang: Conceptualization, Investigation, Methodology, Writing - original draft; Qiang Ma: Review and editing; ChengZhou: Validation.

Acknowledgements

We are very grateful to anonymous reviewers for their constructive suggestions which improved this paper significantly and the AI tool of Writefull in Overleaf for spelling and grammar check. The work of Qiang Ma has been supported by Scientific Startup Foundation for Doctors of Jiangsu University of Science and Technology (CN) (No. 1062931902), the National Natural Science Foundation of China (NSFC) under Grants (No. 52105350) and State Key Laboratory of Precision Welding & Joining of Materials and Structures (No. AWJ-23R01).

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.amc.2025.129499.

Data availability

No data was used for the research described in the article.

References

[1] E.J. Candés, T. Tao, Decoding by linear programming, IEEE Trans. Inf. Theory 51 (12) (2005) 4203–4215.
[2] E.J. Candés, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inf. Theory 52 (2) (2006) 489–509.
[3] D. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (4) (2006) 1289–1306.
[4] E.J. Candés, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, IEEE Trans. Inf. Theory 52 (12) (2006) 5406–5425.
[5] E.V.D. Berg, M.P. Friedlander, Sparse optimization with least-squares constraints, SIAM J. Optim. 21 (4) (2011) 1201–1229.
[6] A.M. Bruckstein, D.L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM Rev. 51 (2009) 34–81.
[7] M.F. Duarte, M.A. Davenport, D. Takhar, J.N. Laska, T. Sun, K.F. Kelly, R.G. Baraniuk, Single-pixel imaging via compressive sampling, IEEE Signal Process. Mag. 25 (Mar. 2008) 83–91.
[8] B.K. Natarajan, Sparse approximate solutions to linear systems, SIAM J. Comput. 24 (2) (1995) 227–234.
[9] Y.C. Pati, R. Rezaiifar, P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, in: Proc. Asilomar Conf. Signals, Syst. Comput, 1993, pp. 40–44.
[10] D. Needell, J.A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, Appl. Comput. Harmon. Anal. 26 (3) (2009) 301–321.
[11] Y.C. Eldar, G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge Univ. Press, U.K., 2012.
[12] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhauser, Cambridge MA, USA, 2013.
[13] Oleg P. Burdakov, Christian Kanzow, Alexandra Schwartz, Mathematical programs with cardinality constraints: reformulation by complementarity-type conditions and a regularization method, SIAM J. Optim. 26 (1) (2016) 397–425.
[14] Christian Kanzow, Andreas B. Raharja, Alexandra Schwartz, Sequential optimality conditions for cardinality-constrained optimization problems with applications, Comput. Optim. Appl. 80 (2021) 185–211.
[15] Yi Zhang, Nikolaos V. Sahinidis, Solving continuous and discrete nonlinear programs with BARON, Comput. Optim. Appl. 5 (Dec. 2024), https://doi.org/10.1007/s10589-024-00633-0.
[16] N. Krejić, E.H.M. Krulikovski, M. Raydan, An augmented Lagrangian approach for cardinality constrained minimization applied to variable selection problems, Appl. Numer. Math. (19 December 2023).
[17] L. Kaufman, P. Rousseeuw, Finding Groups in Data: an Introduction to Cluster Analysis, Wiley, New York, 1990.
[18] Matteo Lapucci, Tommaso Levato, Marco Sciandrone, Convergent inexact penalty decomposition methods for cardinality-constrained problems, J. Optim. Theory Appl. 188 (2021) 473–496.
[19] Shenglong Zhou, Naihua Xiu, Hou-Duo Qi, Global and quadratic convergence of Newton hard-thresholding pursuit, J. Mach. Learn. Res. 22 (2021) 1–45.
[20] Shenglong Zhou, Lili Pan, Naihua Xiu, Newton method for ℓ_0-regularized optimization, Numer. Algorithms 88 (2021) 1541–1570.
[21] T. Blumensath, M.E. Davies, Iterative thresholding for sparse approximation, J. Fourier Anal. Appl. 14 (2008) 629–654.
[22] T. Blumensath, Accelerated iterative hard thresholding, Signal Process. 92 (2012) 752–756.
[23] F. Wu, W. Bian, Accelerated iterative hard thresholding algorithm for ℓ_0 regularized regression problem, J. Glob. Optim. 76 (2020) 819–840.
[24] Y.-B. Zhao, Optimal k-thresholding algorithms for sparse optimization problems, SIAM J. Optim. 30 (1) (2020) 31–55.
[25] Jun Wang, A wonderful triangle in compressed sensing, Inf. Sci. 611 (2022) 95–106.
[26] Jun Wang, Qiang Ma, The variant of the iterative shrinkage-thresholding algorithm for minimization of the ℓ_1 over ℓ_∞ norms, Signal Process. 211 (2023) 109104.