J. Nonlinear Funct. Anal. 2024 (2024) 16 https://doi.org/10.23952/jnfa.2024.16

MATHRES

Journal of Nonlinear Functional Analysis

Available online at http://jnfa.mathres.org

HARD THRESHOLD ACCELERATION ALGORITHMS FOR SPARSE SIGNAL RECOVERY

YASHI HU[1], YUE ZHANG[1], LONGTENG ZHANG[2], XIANGTIAN ZHENG[2], PEICHAO DUAN[1,*]

[1]College of Science, Civil Aviation University of China, Tianjin 300300, China

[2]College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China

Abstract. In sparse signal recovery, hard thresholding algorithms play an important role. By combining hard thresholding techniques with inertial acceleration, the heavy-ball-based hard thresholding algorithm (AIHT) and the heavy-ball-based hard thresholding pursuit algorithm (AHTP) have been proposed in recent years. On this basis, we propose two-step inertial acceleration algorithms A^{2}IHT and A^{2}HTP for signal recovery. In simulation experiments, we test the performance of A^{2}IHT and A^{2}HTP in signal recovery and image restoration as examples. The experimental results show that they take less time to recover with the two-step inertial acceleration algorithms, and the recovery effect is improved in some sense.

Keywords. Compressed Sensing; Hard thresholding algorithm; Sparse signal recovery; Two-step inertial acceleration algorithm.

2020 MSC. 49M37, 65F50, 90C30.

1. Introduction

1.1. **Research background.** In the era of big data, the scale of data in images, videos, and sensor data grows exponentially. Traditional sampling and processing methods are inefficient facing massive amounts of data. Conventional sampling methods often demand high sampling rates, leading to the collection and storage of a substantial amount of redundant information, thereby increasing the burden on transmission and processing [17]. Various efficient iterations for data collection and processing are now under the academic spotlight.

Compressed sensing [7], as an innovative theoretical framework, proposes a method to accurately recover signals at significantly lower sampling rates and is now widely used in various fields. Its theoretical framework no longer requires the sampling process to adhere to the Nyquist theorem. Instead, by leveraging the sparsity of signals, it allows data collection at rates

*Corresponding author.

E-mail address: pcduancauc@126.com (P. Duan).

Received April 24, 2024; Accepted July 22, 2024.

much lower than those demanded by the Nyquist theorem. The introduction of this theory not only provides new perspectives for efficient data collection and processing but also challenges conventional signal processing theories in a revolutionary way.

1.2. **Problem model and solutions.** The core idea of commpressed sensing is that when signals in a certain transform domain are sparse, it means that most of its coefficients in that domain are zero. Based on this notion, compressed sensing achieves efficient compression of data by sparsely sampling in the transform domain, allowing accurate recovery of the original signal with relatively fewer observations. In the context of compressed sensing, the goal is to recover a sparse signal from linear measurements. Assume that a signal f has a sparse representation $f = \Psi x$, where $x \in \mathbb{R}^n$ is sparse. If a measurement matrix $\Phi \in \mathbb{R}^{m\times n}(m \ll n)$ is given, resulting in a measurement vector $y = \Phi f$, we can rewrite the equation as $y = \Phi\Psi x$, where $A = \Phi\Psi \in \mathbb{R}^{m\times n}$. The equation becomes $y = Ax$, where A is referred to as the sensing matrix. Originally, the equation $y = \Phi f$ is underdetermined, but by introducing the sparse representation $f = \Psi x$, the equation $y = Ax$ with the minimum sparsity condition on x may have a unique solution [8, 9, 20]. The problem model is formulated as follows

$$\min \|x\|_0 \quad \text{subject to } \; Ax = y. \tag{1.1}$$

Considering the presence of noise, we have $y = Ax + v$, where v is an unknown noise satisfying $\|v\|_2 \le \varepsilon$. The problem model becomes

$$\min \|x\|_0 \quad \text{subject to } \; \|Ax - y\|_2 \le \varepsilon. \tag{1.2}$$

It can be observed that (1.2) is a constrained optimization problem, which was proven to be NP-hard and challenging to solve directly. To make the solution easier, problem (1.2) is usually transformed into the following form

$$\min \|Ax - y\|_2^2 \quad \text{subject to } \; \|x\|_0 \le k, \tag{1.3}$$

where k represents the sparsity of the signal x with the constraint $0 \le k \le n$. For the precise recovery of a sparse signal, certain restrictive conditions need to be satisfied, especially the incoherence, which underscores the independence among the coefficients of the signal in the transform domain.

1.3. **Algorithms.** Sparse signal recovery algorithms can be mainly categorized into three types: Greedy pursuit methods, convex optimization methods, and hard thresholding methods; see, e.g., [1, 2, 3, 4, 5, 6, 10, 12, 13]. Greedy algorithms iteratively select the most contributing parts of the original signal for recovery. Orthogonal Matching Pursuit (OMP) [18] is a commonly used greedy algorithm that progressively selects the most significant parts of the sparse signal concerning the current observations. Optimization methods aim to find the optimal solution by minimizing a loss function. The Least Absolute Shrinkage and Selection Operator (LASSO) [16] algorithm is an optimization method based on minimizing the l_1 norm, introducing an l_1 regularization term to promote sparsity in the signal. Hard thresholding methods are another important category of sparse recovery algorithms, suitable for signals with known sparsity. In the following, we mainly introduce the Iterative Hard Thresholding algorithm (IHT) [2, 3, 11] and the Hard Thresholding Pursuit algorithm (HTP).

IHT algorithm is an iterative algorithm based on a greedy strategy, aiming to approach the sparse signal gradually through hard thresholding operations during iterations. It was initially proposed by Blumensath and Davies [1] for sparse recovery problems. Preliminary analysis indicates good theoretical guarantees for this algorithm. The algorithm starts with the simple intuition, that is, solving a rectangular system $Ax = y$ is transform to solving the square system $A^T Ax = A^T y$. To ensure sparsity, a hard thresholding operator $\mathscr{H}_k$ is introduced in each iteration to retain the largest k components in the vector. The specific operation is as follows.

Algorithm 1.1 Iterative Hard Thresholding Algorithm (IHT)

Step 1. Input (A, y, k), one parameter $\alpha > 0$, one initial point x^0, and a termination condition.
Step 2. At x^p, set $D^{p+1} = x^p + \alpha A^T(y - Ax^p)$.
Step 3. Let $x^{p+1} = \mathscr{H}_k(D^{p+1})$,
• where $\mathscr{H}_k(\cdot)$ represents retaining the k largest absolute values in '$\cdot$' and setting the other parts to zero.
• stop the iteration if the termination condition is met.
Step 4. Repeat the above steps until a certain stopping criterion is met.

In 2011, Foucart [10] combined the IHT, SP, and CoSaMP algorithms to propose the HTP algorithm. Its basic idea is still to select an appropriate candidate set for the support set first and then find the vector that best fits the measurement results on this support set. The specific operation is as follows.

Algorithm 1.2 Hard Thresholding Pursuit Algorithm (HTP).

Step 1. Input (A, y, k), one parameter $\alpha > 0$, one initial point x^0, and a termination condition.
Step 2. At x^p, set $D^{p+1} = x^p + \alpha A^T(y - Ax^p)$.
Step 3. Let $S^{P+1} = \mathscr{L}_k(D^{p+1})$ and $x^{p+1} = \arg\min_{z\in\mathbb{R}^n}\{\|y - Az\|_2 : \mathrm{supp}(z) \subseteq S^{p+1}\}$,
• where $\mathscr{L}_k(\cdot)$ represents the index set of the k largest absolute values in the '$\cdot$' .
• stop iteration if the termination condition is met.
Step 4. Repeat the above steps until a certain stopping criterion is met.

The AHTP and AIHT algorithms proposed in [17] are acceleration algorithms that introduce an inertia term on the basis of the HTP and IHT algorithms, respectively. Compared to the original algorithm, these two algorithms improve the speed of signal recovery. The specific algorithms are as follows.

Algorithm 1.3 Accelerated Iterative Hard Threshold Algorithm (AIHT)

Step 1. Input (A, y, k), two parameters $\alpha > 0$ and $\beta \geq 0$, and two initial points x^0 and x^1.
Step 2. At x^p, set $D^{p+1} = x^p + \alpha A^T(y - Ax^p) + \beta(x^p - x^{p-1})$.
Step 3. Let $x^{p+1} = \mathscr{H}_k(D^{p+1})$.
Step 4. Repeat the above steps until a certain stopping criterion is met.

It can be seen that the search direction at x^p in both IHT and HTP algorithms is given by $A^T(y - Ax^p)$, which is the negative gradient of the residual at x^p. With the help of the momentum term $x^p - x^{p-1}$, the search direction can be modified to $d^p = \alpha A^T(y - Ax^p) + \beta(x^p - x^{p-1})$. The iterative algorithm after updating the search direction is a one-step inertial acceleration

Algorithm 1.4 Accelerated Hard Thresholding Pursuit Algorithm (AHTP)

Step 1. Input (A,y,k), two parameters $\alpha>0$ and $\beta\geq 0$, and two initial points x^0 and x^1.
Step 2. At x^p, set $D^{p+1}=x^p+\alpha A^T(y-Ax^p)+\beta(x^p-x^{p-1})$.
Step 3. Let $S^{p+1}=\mathscr{L}_k(D^{p+1})$ and $x^{p+1}=\arg\min_{z\in\mathbb{R}^n}\left\{\|y-Az\|_2^2 : supp(z)\subseteq S^{p+1}\right\}$.
Step 4. Repeat the above steps until a certain stopping criterion is met.

algorithm for IHT and HTP, which we define as AIHT and AHTP. Similarly, by adding two-step momentum terms x^p-x^{p-1} and $x^{p-1}-x^{p-2}$, we obtain a new search direction $d^p=\alpha A^T(y-Ax^p)+\beta(x^p-x^{p-1})+\gamma(x^{p-1}-x^{p-2})$. Our proposed algorithms are called as A^{2}IHT and A^{2}HTP [14, 15, 17, 19]. Keeping the other parts of the algorithm unchanged, adding momentum terms can accelerate the algorithm without increasing computational complexity, which achieves faster convergence. Currently, adding a momentum term to accelerate the AIHT and AHTP algorithms has been proven to be successful in recovering signals under conditions $\delta_{3k}<0.618$ and $\delta_{3k}<0.577$ [17]. In this article, we prove that the performance of A^{2}IHT can be guaranteed under

$$\delta_{3k}<\frac{0.7255\gamma^{\frac{1}{3}}}{1+0.8517\gamma^{\frac{2}{3}}-\gamma},$$

$0<\gamma<0.0993$, and the performance of A^{2}HTP can be guaranteed under

$$\delta_{3k}<\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\hat{\eta}^{-\frac{1}{3}}\gamma^{\frac{2}{3}}-\gamma}\ (0<\gamma<\frac{\xi^3}{\hat{\eta}},\hat{\eta}=\frac{\sqrt{2}}{\sqrt{1-(\delta_{2k})^2}}).$$

2. Preliminaries

2.1. **Notation.** We use N to represent $\{1,2,...,n\}$ and let $\omega\subseteq N$. The notations $\bar{\omega}:=N\setminus\omega$ and $|\omega|$ mean the complement set and the cardinality of ω, respectively. Set a vector $z\in\mathbb{R}^n$. The support of z is defined as the index set of nonzero elements $supp(z):=\{i\in N : z_i\neq 0\}$. $z_\omega\in\mathbb{R}^n$ is a vector expressed in the form

$$(z_\omega)_i=\begin{cases} z_i,\ i\in\omega,\\ 0,\ i\notin\omega.\end{cases}$$

The k-sparse vector $\mathscr{H}_k(z)$ is the best k-term approximation of $z\in\mathbb{R}^n$. Represented by $\sigma_k(z)_q$, where $q>0$ is an integer, the residual of the best k-term approximation of z is as follows,

$$\sigma_k(z)_q=\min_u\left\{\|z-u\|_q : \|u\|_0\leq k\right\}.$$

2.2. **Definition and lemmas.** First, we give the restricted isometry property (RIP) of the measurement matrix.

Definition 2.1. [23] (*RIP*) Matrix $A\in\mathbb{R}^{m\times n}$ is said to satisfy the restricted isometry property of order k if and only if there exists a constant $\delta_k\in(0,1)$ such that $(1-\delta_k)\|x\|_2^2\leq\|Ax\|_2^2\leq(1+\delta_k)\|x\|_2^2$ holds for all k-sparse vectors x. The constant δ_k is called the k-order restricted isometry of matrix A.

Next, we give two useful lemmas for our main results.

Lemma 2.2. [8, 22] *Let $v \in \mathbb{R}^n$ be a vector. Let $t \in N$ be a positive integer number, and let $W \subseteq N$ be an index set.*

(i) If $|W \cup supp(v)| \leq t$, then $\left\|((I - A^T A)v)_W\right\|_2 \leq \delta_t \|v\|_2$.

(ii) If $|W| \leq t$, then $\left\|(A^T v)_W\right\|_2 \leq \sqrt{1+\delta_t}\,\|v\|_2$.

Lemma 2.3. [21] *For any vector $z \in \mathbb{R}^n$ and any k-sparse vector $x \in \mathbb{R}^n$, the following inequality holds $\|x - \mathscr{H}_k(z)\|_2 \leq \eta \|(x-z)_{W\cup W^*}\|_2$, where $\eta = (\sqrt{5}+1)/2$, $W = supp(x)$, and $W^* = supp(\mathscr{H}_k(z))$.*

3. Main Results

3.1. **Algorithms.** In this section, we propose the two-step inertial acceleration algorithm. In IHT, we choose the appropriate perceptual matrix A, and the search direction $A^T(y - Ax^{p-1})$ at iteration x^p. We improve the algorithm by introducing momentum terms $x^p - x^{p-1}$ and $x^{p-1} - x^{p-2}$ and adjusting the search direction to $d^p = \alpha A^T(y - Ax^p) + \beta(x^p - x^{p-1}) + \gamma(x^{p-1} - x^{p-2})$. Generate $u^p = x^p + d^p$ from this search direction, where x^p represents the current iteration result of the algorithm. Then performe a hard thresholding on u^p to produce the next iterate. Algorithms derived from this idea can be used to solve problems (1.3). The specific operation is as follows.

Algorithm 3.1 The Two-step Inertial Acceleration Algorithm (A^{2}IHT)

Step 1. Input (A, y, k), three parameters $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, and three initial points x^0, x^1, and x^2.
Step 2. For x^p, set

$$u^p = x^p + \alpha A^T(y - Ax^p) + \beta(x^p - x^{p-1}) + \gamma(x^{p-1} - x^{p-2}). \tag{3.1}$$

Step 3. Let $x^{p+1} = \mathscr{H}_k(u^p)$.
Step 4. Repeat the above steps until the result meets the set conditions.

In the same way, adjusting the search direction of HTP can obtain the A^{2}HTP algorithm. The specific operation is as follows.

Algorithm 3.2 The Two-step Inertial Acceleration Pursuit Algorithm (A^{2}HTP)

Step 1. Input (A, y, k), three parameters $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, and three initial points x^0, x^1, and x^2.
Step 2. For x^p, set $u^p = x^p + \alpha A^T(y - Ax^p) + \beta(x^p - x^{p-1}) + \gamma(x^{p-1} - x^{p-2})$.
Step 3. Let $S^{p+1} = \mathscr{L}_k(u^p)$, and

$$x^{p+1} = \underset{z\in\mathbb{R}^n}{\arg\min}\left\{\|y - Az\|_2^2 : supp(z) \subseteq S^{p+1}\right\}. \tag{3.2}$$

Step 4. Repeat the above steps until the result meets the set conditions.

The stopping condition can be set the maximum number of iterations or the minimum error by itself. For example, let $\|y - Ax^p\|_2 \leq \varepsilon$, where $\varepsilon > 0$ is a small tolerance. If the measurements y are accurate enough, then the measurement error $\|v\|_2 = \|y - Ax\|_2$ would be very small.

3.2. **Lemma.** Before analyzing the performance of A^{2}IHT and A^{2}HTP at a RIP of order 3k, we first prove a useful lemma.

Lemma 3.1. *Let* $\{a^p\} \subseteq \mathbb{R}$ *be a sequence of nonnegative real numbers such that*

$$a^{p+1} \leq b_1 a^p + b_2 a^{p-1} + b_3 a^{p-2} + b_4, p \geq 2, \tag{3.3}$$

where b_1, b_2, $b_3 \geq 0$ *and*

$$b_1 < \frac{1}{\sqrt[3]{\frac{19}{27}+\sqrt{\frac{11}{27}}}+\sqrt[3]{\frac{19}{27}-\sqrt{\frac{11}{27}}}+\frac{1}{3}} \approx 0.543689 = \xi, b_2 = {b_1}^2, b_3 = {b_1}^3.$$

Then

$$a^p \leq \theta^{p-2}[a^2 + (\theta - b_1)a^1 + \frac{b_3}{\theta}a^0] + \frac{b_4}{1-\theta} \tag{3.4}$$

with

$$0 \leq \theta := \sqrt[3]{-\frac{n}{2}+\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \sqrt[3]{-\frac{n}{2}-\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \frac{b_1}{3} < 1,$$

where

$$m = \frac{-b_1^2 - 3b_2}{3}, n = \frac{-2b_1^3 - 9b_1b_2 - 27b_3}{27}.$$

Proof. Let

$$q_1 := \sqrt[3]{-\frac{n}{2}+\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \sqrt[3]{-\frac{n}{2}-\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} - \frac{2b_1}{3},$$

$$q_2 := \frac{b_3}{\sqrt[3]{-\frac{n}{2}+\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \sqrt[3]{-\frac{n}{2}-\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \frac{b_1}{3}},$$

$$m = \frac{-b_1^2 - 3b_2}{3}, n = \frac{-2b_1^3 - 9b_1b_2 - 27b_3}{27}.$$

Since b_1, b_2, $b_3 \geq 0$ and $b_1 < 0.543689 = \xi$, $b_2 = b_1^2$, $b_3 = b_1^3$. It is easy to see that

$$q_1 = (\sqrt[3]{\frac{19}{27}+\sqrt{\frac{11}{27}}}+\sqrt[3]{\frac{19}{27}-\sqrt{\frac{11}{27}}}-\frac{2}{3})b_1 \geq 0,$$

$$q_2 = \frac{b_3}{(\sqrt[3]{\frac{19}{27}+\sqrt{\frac{11}{27}}}+\sqrt[3]{\frac{19}{27}-\sqrt{\frac{11}{27}}}+\frac{1}{3})b_1} \geq 0,$$

$$(b_1+q_1)q_1 = b_2 + q_2, (b_1+q_1)q_2 = b_3,$$

and

$$0 \leq \theta = \sqrt[3]{-\frac{n}{2}+\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \sqrt[3]{-\frac{n}{2}-\sqrt{(\frac{n}{2})^2+(\frac{m}{3})^3}} + \frac{b_1}{3} = b_1 + q_1 < 1,$$

where $\theta < 1$ under the condition $b_1 < 0.543689 = \xi$, $b_2 = b_1^2$, and $b_3 = b_1^3$. Thus it follows from (3.3) that

$$a^{p+1} + q_1 a^p + q_2 a^{p-1} \leq (b_1+q_1)a^p + (b_2+q_2)a^{p-1} + b_3 a^{p-2} + b_4$$

$$= \theta(a^p + q_1 a^{p-1} + q_2 a^{p-2}) + b_4,$$

which implies

$$\begin{aligned} a^{p+1} \le a^{p+1} + q_1 a^p + q_2 a^{p-1} &\le \theta^{p-1}(a^2 + q_1 a^1 + q_2 a^0) + b_4(1 + \theta + \ldots + \theta^{p-2}) \\ &\le \theta^{p-1}(a^2 + q_1 a^1 + q_2 a^0) + b_4(1 + \theta + \ldots + \theta^{p-1}) \\ &\le \theta^{p-1}[a^2 + (\theta - b_1)a^1 + \frac{b_3}{\theta} a^0] + \frac{b_4}{1-\theta}. \end{aligned}$$

Thus inequality (3.4) holds. □

3.3. **Analysis of A^{2}IHT and A^{2}HTP.** Now, we explore how to ensure the performance of A^{2}IHT and A^{2}HTP under RIP of order $3k$. Please note that η=$\frac{\sqrt{5}+1}{2}$, $\xi \approx 0.543689$ in the remaining of this paper. We first consider the A^{2}IHT. The main results are as follows.

Theorem 3.2. *Let the RIC, δ_{3k} of the measurement matrix A and the parameters α, β and γ meet the requirements* :

$$0 < \gamma < \frac{\xi^3}{\eta},\ \beta = \eta^{-\frac{1}{3}}\gamma^{\frac{2}{3}} - \gamma,\ \delta_{3k} < \frac{\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\beta},\ \alpha = \frac{-\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} + 1 + \beta}{1-\delta_{3k}} \ \text{or}\ \alpha = \frac{\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} + 1 + \beta}{1+\delta_{3k}}. \tag{3.5}$$

Let $y := Ax + v$ be the measurements of x, where v is the measurement error. Then the sequence $\{x^p\}$ generated by A^{2}IHT holds that

$$\|x_S - x^p\|_2 \le C_1 \tau^{p-1} + C_2 \|v'\|_2, \tag{3.6}$$

where $S = \mathscr{L}_k(x)$, $v' = v + Ax_{\overline{S}}$, and C_1 and C_2 are the quantities defined by

$$C_1 = \|x_S - x^2\|_2 + (\tau - b_1)\|x_S - x^1\|_2 + \frac{b_3}{\tau}\|x_S - x^0\|_2,\ C_2 = \frac{\eta\alpha}{1-\tau}\sqrt{1+\delta_{2k}} \tag{3.7}$$

with

$$\tau := \sqrt[3]{-\frac{n}{2} + \sqrt{(\frac{n}{2})^2 + (\frac{m}{3})^3}} + \sqrt[3]{-\frac{n}{2} - \sqrt{(\frac{n}{2})^2 + (\frac{m}{3})^3}} + \frac{b_1}{3},$$

$$m = \frac{-b_1^2 - 3b_2}{3},\ n = \frac{-2b_1^3 - 9b_1 b_2 - 27 b_3}{27}.$$

The fact $\tau < 1$ is ensured under (3.3), and b_1, b_2, and b_3 are given by

$$b_1 = \eta(|1 - \alpha + \beta| + \alpha\delta_{3k}),\ b_2 = \eta(\beta + \gamma),\ b_3 = \eta\gamma. \tag{3.8}$$

Proof. By (3.1), we have

$$\begin{aligned} u^p - x_S = &(1 - \alpha + \beta)(x^p - x_S) + \alpha(I - A^T A)(x^p - x_S) - \beta(x^{p-1} - x_S) + \gamma(x^{p-1} - x_S) \\ &- \gamma(x^{p-2} - x_S) + \alpha A^T v', \end{aligned} \tag{3.9}$$

where $v' = v + Ax_{\bar{S}}$. Let $V^p := supp(\mathscr{H}_k(u^p))$. By Lemma 2.3 and (3.9), we derive

$$\begin{aligned} \left\|x^{p+1} - x_S\right\|_2 \le &\ \eta \left\|(u^p - x_S)_{S\cup V^p}\right\|_2 \\ \le &\ \eta\,|1 - \alpha + \beta|\,\|x^p - x_S\|_2 + \eta\alpha\left\|[(I - A^T A)(x^p - x_S)]_{S\cup V^p}\right\|_2 \\ &+ \eta\beta\left\|x^{p-1} - x_S\right\|_2 + \eta\gamma\left\|x^{p-1} - x_S\right\|_2 + \eta\gamma\left\|x^{p-2} - x_S\right\|_2 + \eta\alpha\left\|(A^T v')_{S\cup V^p}\right\|_2. \end{aligned} \tag{3.10}$$

By using Lemma 2.2 and combining $|S \cup V^p| \leq 2k$ and $|supp(x^p - x_S) \cup S \cup V^p| \leq 3k$, we have

$$\left\| [(I - A^T A)(x^p - x_S)]_{S \cup V^p} \right\|_2 \leq \delta_{3k} \| x^p - x_S \|_2 \tag{3.11}$$

and

$$\left\| (A^T v')_{S \cup V^p} \right\|_2 \leq \sqrt{1 + \delta_{2k}} \left\| v' \right\|_2. \tag{3.12}$$

Substituting inequalities (3.11) and (3.12) into (3.10), we obtain

$$\begin{aligned} \left\| x^{p+1} - x_S \right\|_2 \leq\ & \eta \left| 1 - \alpha + \beta \right| \| x^p - x_S \|_2 + \eta \alpha \delta_{3k} \| x^p - x_S \|_2 \\ & + \eta \beta \left\| x^{p-1} - x_S \right\|_2 + \eta \gamma \left\| x^{p-1} - x_S \right\|_2 \\ & + \eta \gamma \left\| x^{p-2} - x_S \right\|_2 + \eta \alpha \sqrt{1 + \delta_{2k}} \left\| v' \right\|_2 \\ =\ & b_1 \| x^p - x_S \|_2 + b_2 \left\| x^{p-1} - x_S \right\|_2 + b_3 \left\| x^{p-2} - x_S \right\|_2 + \eta \alpha \sqrt{1 + \delta_{2k}} \left\| v' \right\|_2, \end{aligned} \tag{3.13}$$

where b_1, b_2, and b_3 are given by (3.8). The form of recursive inequality (3.13) is the same as that given by Lemma 3.1. Now we verify that the coefficients of each iteration term in (3.13) satisfy the condition of Lemma 3.1.

First, note the following values

$$\eta = \frac{\sqrt{5} + 1}{2} \approx 1.6180339,\ \eta^{-1} \approx 0.6180339,\ \eta^{-\frac{2}{3}} \approx 0.7255626,\ \eta^{-\frac{1}{3}} \approx 0.85179964,$$

$$\xi \approx 0.543689,\ \xi^2 \approx 0.29559772,\ \xi^3 \approx 0.1607132.$$

Second, from the range of γ in (3.5), we obtain $\gamma < \eta^{-\frac{1}{3}} \gamma^{\frac{2}{3}}$, which means $\beta > 0$.

Third, from the value of β in (3.5), we obtain $\frac{\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}}}{1+\beta} = \frac{\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}}}{1 + \eta^{-\frac{1}{3}} \gamma^{\frac{2}{3}} - \gamma}$. Now, we suppose that $\gamma = j^3 < \frac{\xi^3}{\eta}$. obviously, there exists

$$\frac{\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}}}{1 + \beta} = \frac{\eta^{-\frac{2}{3}} j}{1 + \eta^{-\frac{1}{3}} j^2 - j^3}. \tag{3.14}$$

To analyze the value of (3.14), we define $f(j) = \eta^{-\frac{2}{3}} j - 1 - \eta^{-\frac{1}{3}} j^2 + j^3$. The derivative of $f(j)$ is $f'(j) = \eta^{-\frac{2}{3}} - 2\eta^{-\frac{1}{3}} j + 3j^2$. In the range of real numbers, we know $[f'(j)]_{min} \approx f'(0.283933) \approx 0.483708 > 0$. Thus we can get the following information: $f'(j) > 0$, and $f(j)$ is an incremental function. This means $f(j)_{max} = f(\xi \eta^{-\frac{1}{3}}) \approx -0.9116828 < 0$. Thus we obtain $\eta^{-\frac{2}{3}} j < 1 + \eta^{-\frac{1}{3}} j^2 - j^3$, which infers $\frac{\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}}}{1+\beta} < 1$.

Forth, note the value of α in (3.5). Obviously, $\alpha = \frac{-\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}} + 1 + \beta}{1 - \delta_{3k}}$ or $\alpha = \frac{\eta^{-\frac{2}{3}} \gamma^{\frac{1}{3}} + 1 + \beta}{1 + \delta_{3k}}$, $\alpha > 0$ holds. To sum up, we have $\alpha > 0$, $\beta > 0$, and $\delta_{3k} < 1$.

Next, we study the relationship among b_1, b_2, and b_3. The range for γ in (3.5) is well defined. Thus $b_3 = \eta \gamma < \xi^3$. Substituting the value of β into b_2, we obtain $b_2 = \eta^{\frac{2}{3}} \gamma^{\frac{2}{3}} = b_3^{\frac{2}{3}}$. The range for δ_{3k} and the value of α in (3.5) are well defined. We discuss it in two cases.

Case 1 : $\alpha = \frac{-\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}}$. Indeed, from $\delta_{3k} < \frac{\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\beta}$ in (3.5), we obtain

$$-(1+\beta)\delta_{3k} > -\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} \Rightarrow (1+\beta)(1-\delta_{3k}) > -\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} + 1 + \beta$$
$$\Rightarrow 1+\beta > \frac{-\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}} = \alpha.$$

In this case,

$$b_1 = \eta[1+\beta-\alpha(1-\delta_{3k})] = \eta[1+\beta-(-\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta)] = \eta^{\frac{1}{3}}\gamma^{\frac{1}{3}} = b_3{}^{\frac{1}{3}}.$$

Case 2 : $\alpha = \frac{\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}}$. Observe that

$$(1+\beta)\delta_{3k} < \eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} \Rightarrow (1+\beta)(1+\delta_{3k}) < \eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}} + 1 + \beta$$
$$\Rightarrow 1+\beta < \frac{\eta^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}} = \alpha.$$

In this case,

$$b_1 = \eta(-1+\alpha-\beta+\alpha\delta_{3k}) = \eta[-1-\beta+\alpha(1+\delta_{3k})] = \eta^{\frac{1}{3}}\gamma^{\frac{1}{3}} = b_3{}^{\frac{1}{3}}.$$

Therefore, there exists $b_1 = b_3{}^{\frac{1}{3}}$ under (3.5). In summary, b_1, b_2, and b_3 have the following relationship: $b_3 < \xi^3$, $b_1 = b_3{}^{\frac{1}{3}}$, and $b_2 = b_3{}^{\frac{2}{3}}$, which means that $b_1 < \xi$, $b_2 = b_1{}^2$, and $b_3 = b_1{}^3$. Hence, recursive inequality (3.13) satisfies the condition of Lemma 3.1. And it follows from Lemma 3.1 that

$$\tau = \sqrt[3]{-\frac{n}{2} + \sqrt{\left(\frac{n}{2}\right)^2 + \left(\frac{m}{3}\right)^3}} + \sqrt[3]{-\frac{n}{2} - \sqrt{\left(\frac{n}{2}\right)^2 + \left(\frac{m}{3}\right)^3}} + \frac{b_1}{3} < 1,$$

where $m = \frac{-b_1^2-3b_2}{3}$, $n = \frac{-2b_1^3-9b_1b_2-27b_3}{27}$, and (3.4) holds, where C_1 and C_2 are given in (3.7). $\square$

When the signal x is k-sparse and the measurements are accurate, in which $x = x_S$ and $v = 0$, the above result shows that $\|x - x^p\|_2 \le \tau^{p-1}C_1 \to 0$ as $p \to \infty$, which implies that the iterates generated by A^{2}IHT converges to the sparse signal.

Next, we consider A^{2}HTP. We first recall the following helpful lemma.

Lemma 3.3. [10] *Given the measurements $y := Ax + v$ of x and the index set S^{p+1}, the iteration x^{p+1} generated by the pursuit step (3.2) obeys*

$$\|x^{p+1} - x_S\|_2 \le \frac{1}{\sqrt{1-(\delta_{2k})^2}}\|(x^{p+1}-x_S)_{\overline{S^{p+1}}}\|_2 + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}\|v'\|_2, \tag{3.15}$$

where $S = \mathcal{L}_k(x)$ and $v' = v + Ax_{\overline{S}}$.

The main result concerning the guaranteed success of A^{2}HTP is as follows.

Theorem 3.4. *Suppose that the RIC, δ_{3k} of the measurement matrix A, and the parameters α, β, and γ satisfy*

$$0 < \gamma < \frac{\xi^3}{\hat{\eta}},\ \beta = \hat{\eta}^{-\frac{1}{3}}\gamma^{\frac{2}{3}} - \gamma,\ \delta_{3k} < \frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\beta},\ \alpha = \frac{-\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}}\ \textit{or}\ \alpha = \frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}} \tag{3.16}$$

where $\hat{\eta} = \frac{\sqrt{2}}{\sqrt{1-(\delta_{2k})^2}}$. Let $y := Ax + v$ be the measurements of x, where v is the measurement error. Then the sequence $\{x^p\}$ generated by A^2HTP holds that

$$\|x_S - x^p\|_2 \leq C_3 \hat{\tau}^{p-1} + C_4 \|v'\|_2, \tag{3.17}$$

where $S = \mathscr{L}_k(x)$, $v' = v + Ax_{\overline{S}}$, and C_3, C_4 are the quantities defined by

$$C_3 = \|x_S - x^2\|_2 + (\hat{\tau} - \hat{b}_1)\|x_S - x^1\|_2 + \frac{\hat{b}_3}{\hat{\tau}}\|x_S - x^0\|_2,\ C_4 = \frac{1}{1-\hat{\tau}}(\hat{\eta}\alpha\sqrt{1+\delta_{2k}} + \frac{\sqrt{1+\delta_k}}{1-\delta_{2k}}) \tag{3.18}$$

with

$$\hat{\tau} := \sqrt[3]{-\frac{\hat{n}}{2} + \sqrt{(\frac{\hat{n}}{2})^2 + (\frac{\hat{m}}{3})^3}} + \sqrt[3]{-\frac{\hat{n}}{2} - \sqrt{(\frac{\hat{n}}{2})^2 + (\frac{\hat{m}}{3})^3}} + \frac{\hat{b}_1}{3},$$

$$\hat{m} = \frac{-\hat{b}_1^2 - 3\hat{b}_2}{3},\ \hat{n} = \frac{-2\hat{b}_1^3 - 9\hat{b}_1\hat{b}_2 - 27\hat{b}_3}{27}.$$

The fact $\hat{\tau} < 1$ is ensured under (3.16), and $\hat{b}_1$, $\hat{b}_2$, and $\hat{b}_3$ are given by

$$\hat{b}_1 = \hat{\eta}(|1-\alpha+\beta| + \alpha\delta_{3k}),\ \hat{b}_2 = \hat{\eta}(\beta+\gamma),\ \hat{b}_3 = \hat{\eta}\gamma. \tag{3.19}$$

Proof. Since $S^{p+1} = \mathscr{L}_k(u^p)$ in A^2HTP and $S = \mathscr{L}_k(x)$, we have $\|(u^p)_{S^{p+1}}\|_2^2 \geq \|(u^p)_S\|_2^2$. Removing the entries indexed by $S \cap S^{p+1}$ from the inequality above and taking square root, we have $\|(u^p)_{S^{p+1}\setminus S}\|_2^2 \geq \|(u^p)_{S\setminus S^{p+1}}\|_2^2$. Note that $(x_S)_{S^{p+1}\setminus S} = 0$ and $(x^{p+1})_{S\setminus S^{p+1}} = 0$. According to the inequality above, we obtain by $(x_S - x^{p+1})_{S\setminus S^{p+1}} = (x_S - x^{p+1})_{\overline{S^{p+1}}}$ that

$$\begin{aligned}\|(u^p - x_S)_{S^{p+1}\setminus S}\|_2 &\geq \|(x_S - x^{p+1} + u^p - x_S)_{S\setminus S^{p+1}}\|_2 \\ &\geq \|(x_S - x^{p+1})_{\overline{S^{p+1}}}\|_2 - \|(u^p - x_S)_{S\setminus S^{p+1}}\|_2.\end{aligned}$$

Thus

$$\begin{aligned}\|(x_S - x^{p+1})_{\overline{S^{p+1}}}\|_2 &\leq \|(u^p - x_S)_{S\setminus S^{p+1}}\|_2 + \|(u^p - x_S)_{S^{p+1}\setminus S}\|_2 \\ &\leq \sqrt{2(\|(u^p - x_S)_{S\setminus S^{p+1}}\|_2^2 + \|(u^p - x_S)_{S^{p+1}\setminus S}\|_2^2)} \\ &= \sqrt{2}\|(u^p - x_S)_{S^{p+1}\triangle S}\|_2,\end{aligned} \tag{3.20}$$

where $S^{p+1}\triangle S := (S^{p+1}\setminus S) \cup (S \setminus S^{p+1})$ is the symmetric difference of S^{p+1} and S. And the last equality above follows from the fact that $(S^{p+1}\setminus S) \cap (S\setminus S^{p+1}) = \emptyset$. Note that (3.9) is still valid for A^2HTP. Combining (3.9) and (3.20), one arrives at

$$\begin{aligned}\|(x_S - x^{p+1})_{\overline{S^{p+1}}}\|_2 \leq \sqrt{2}\{&|1-\alpha+\beta|\,\|(x^p - x_S)_{S^{p+1}\triangle S}\|_2 \\ &+\alpha\|[(I - A^TA)(x^p - x_S)]_{S^{p+1}\triangle S}\|_2 + (\beta+\gamma)\|(x^{p-1} - x_S)_{S^{p+1}\triangle S}\|_2 \\ &+\gamma\|(x^{p-2} - x_S)_{S^{p+1}\triangle S}\|_2 + \alpha\|(A^Tv')_{S^{p+1}\triangle S}\|_2\}.\end{aligned} \tag{3.21}$$

By using Lemma 2.2 and combining $|S^{p+1}\triangle S| \leq 2k$ and $|(S^{p+1}\triangle S) \cup supp(x^p - x_S)| \leq 3k$, one has

$$\|[(I - A^TA)(x^p - x_S)]_{S^{p+1}\triangle S}\|_2 \leq \delta_{3k}\|x^p - x_S\|_2 \tag{3.22}$$

and

$$\|(A^Tv')_{S^{p+1}\triangle S}\|_2 \leq \sqrt{1+\delta_{2k}}\|v'\|_2. \tag{3.23}$$

Substituting (3.22) and (3.23) into (3.21), we obtain

$$\begin{aligned}\|(x_S-x^{p+1})_{\overline{S^{p+1}}}\|_2 \le \sqrt{2}\{(|1-\alpha+\beta|+\alpha\delta_{3k})\|x^p-x_S\|_2+(\beta+\gamma)\|x^{p-1}-x_S\|_2\\ +\gamma\|x^{p-2}-x_S\|_2+\alpha\sqrt{1+\delta_{2k}}\|v'\|_2\}. \qquad (3.24)\end{aligned}$$

According to Lemma 3.3, we see that

$$\|x^{p+1}-x_S\|_2 \le \hat{b}_1\|x^p-x_S\|_2+\hat{b}_2\left\|x^{p-1}-x_S\right\|_2+\hat{b}_3\left\|x^{p-2}-x_S\right\|_2+(1-\hat{\tau})C_4\|v'\|_2. \quad (3.25)$$

We next prove that the coefficients of each term in (3.25) satisfy the condition of Lemma 3.1. In view of

$$\hat{\eta}=\frac{\sqrt{2}}{\sqrt{1-(\delta_{2k})^2}}>1,\ \hat{\eta}^{\frac{1}{3}}>1,\text{ and }\hat{\eta}^{\frac{2}{3}}>1,$$

we have $\hat{\eta}^{-1}<1$, $\hat{\eta}^{-\frac{1}{3}}<1$ and $\hat{\eta}^{-\frac{2}{3}}<1$. From the range of γ in (3.16), we obtain $\gamma<\hat{\eta}^{-\frac{1}{3}}\gamma^{\frac{2}{3}}$, which yields $\beta>0$. Similarly, letting $\gamma=\hat{j}^3<\frac{\xi^3}{\hat{\eta}}$, we obtain

$$\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\beta}=\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\hat{\eta}^{-\frac{1}{3}}\gamma^{\frac{2}{3}}-\gamma}=\frac{\hat{\eta}^{-\frac{2}{3}}\hat{j}}{1+\hat{\eta}^{-\frac{1}{3}}\hat{j}^2-\hat{j}^3}. \qquad (3.26)$$

Defining $g(\hat{j})=\hat{\eta}^{-\frac{2}{3}}\hat{j}-1-\hat{\eta}^{-\frac{1}{3}}\hat{j}^2+\hat{j}^3$, we obtain $g'(\hat{j})=\hat{\eta}^{-\frac{2}{3}}-2\hat{\eta}^{-\frac{1}{3}}\hat{j}+3\hat{j}^2$ and $[g'(\hat{j})]_{min}=g'(\frac{\hat{\eta}^{-\frac{1}{3}}}{3})=\frac{2}{3}\hat{\eta}^{-\frac{2}{3}}>0$. This demonstrates that $g(\hat{j})$ is an incremental function and $g(\hat{j})_{max}=g(\xi\hat{\eta}^{-\frac{1}{3}})=\hat{\eta}^{-1}(\xi-\xi^2+\xi^3)-1<0$. Hence we deduce $\hat{\eta}^{-\frac{2}{3}}\hat{j}<1+\hat{\eta}^{-\frac{1}{3}}\hat{j}^2-\hat{j}^3$, which yields $\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}}{1+\beta}<1$. Whether $\alpha=\frac{-\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}}$ or $\alpha=\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}}$, $\alpha>0$ always holds. Thus $\alpha>0$, $\beta>0$, and $\delta_{3k}<1$.

Next, we consider the relationship among $\hat{b}_1$, $\hat{b}_2$, and $\hat{b}_3$. The range of γ and the value of β are determined in (3.16). Thus $\hat{b}_3=\hat{\eta}\gamma<\xi^3$ and $\hat{b}_2=\hat{\eta}^{\frac{2}{3}}\gamma^{\frac{2}{3}}=\hat{b}_3^{\frac{2}{3}}$. The range for δ_{3k} and the value of α in (3.16) are well defined. One considers the following two cases.

Case 1: $\alpha=\frac{-\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}}$. Following Theorem 3.2, we have $1+\beta>\frac{-\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1-\delta_{3k}}=\alpha$. In this case, $\hat{b}_1=\hat{\eta}^{\frac{1}{3}}\gamma^{\frac{1}{3}}=\hat{b}_3^{\frac{1}{3}}$.

Case 2: $\alpha=\frac{\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}}$. Following Theorem 3.2, we have $1+\beta<\frac{-\hat{\eta}^{-\frac{2}{3}}\gamma^{\frac{1}{3}}+1+\beta}{1+\delta_{3k}}=\alpha$. In this case, $\hat{b}_1=\hat{\eta}^{\frac{1}{3}}\gamma^{\frac{1}{3}}=\hat{b}_3^{\frac{1}{3}}$. Under (3.16), there exists $\hat{b}_1=\hat{b}_3^{\frac{1}{3}}$. Thus $\hat{b}_3<\xi^3$, $\hat{b}_1=\hat{b}_3^{\frac{1}{3}}$, and $\hat{b}_2=\hat{b}_3^{\frac{2}{3}}$, which shows $\hat{b}_1<\xi$, $\hat{b}_2=\hat{b}_1^2$, and $\hat{b}_3=\hat{b}_1^3$. Hence (3.25) satisfies the condition of Lemma 3.1. It follows that

$$\hat{\tau}:=\sqrt[3]{-\frac{\hat{n}}{2}+\sqrt{(\frac{\hat{n}}{2})^2+(\frac{\hat{m}}{3})^3}}+\sqrt[3]{-\frac{\hat{n}}{2}-\sqrt{(\frac{\hat{n}}{2})^2+(\frac{\hat{m}}{3})^3}}+\frac{\hat{b}_1}{3}<1,$$

where $\hat{m}=\frac{-\hat{b}_1^2-3\hat{b}_2}{3}$, $\hat{n}=\frac{-2\hat{b}_1^3-9\hat{b}_1\hat{b}_2-27\hat{b}_3}{27}$, and (3.17) holds. □

When the measurements are accurate and the signal is k-sparse, Theorem 3.4 implies that the sequence $\{x^p\}$ defined by the A^{2}HTP converges to the signal as $p\to\infty$. This means the algorithm exactly recovers the signal. In addition, it is also worth emphasizing that computing the RIC of a matrix is generally difficult. Thus, in practical applications, we do not require the parameters

α, β, and γ be chosen to strictly meet the condition (3.5) or (3.16). These parameters can be set to roughly satisfy these conditions. We may simply set $0<\gamma<0.0993$, $\beta=0.8517\gamma^{\frac{2}{3}}-\gamma$, and $\alpha=-1.0364\gamma^{\frac{1}{3}}+1.4285+1.4285\beta$ or $\alpha=0.558\gamma^{\frac{1}{3}}+0.7692+0.7692\beta$ in A^{2}IHT for simplicity, and set $0<\gamma<0.1136$, $\beta=0.8908\gamma^{\frac{2}{3}}-\gamma$, and $\alpha=-1.1338\gamma^{\frac{1}{3}}+1.4285+1.4285\beta$ or $\alpha=0.6105\gamma^{\frac{1}{3}}+0.7692+0.7692\beta$ in A^{2}HTP.

4. Image Restoration and Signal Recovery Experiments

In this section, we present experiments conducted on a laptop equipped with an AMD Ryzen 7 5800H processor with Radeon Graphics (3.20 GHz), 475 internal memory, and 16 GB RAM. In the signal recovery experiment, the matrix A is a randomly generated Gaussian matrix of size 50×100, serving as a measurement matrix. The vectors to be recovered are randomly generated sparse vectors with a sparsity level k of 10. We convert general images into 256×256 grayscale images. The size of the measurement matrix A is 190×256. After applying wavelet transformation to the grayscale image matrix, we specify the sparsity level of its vector as 50.

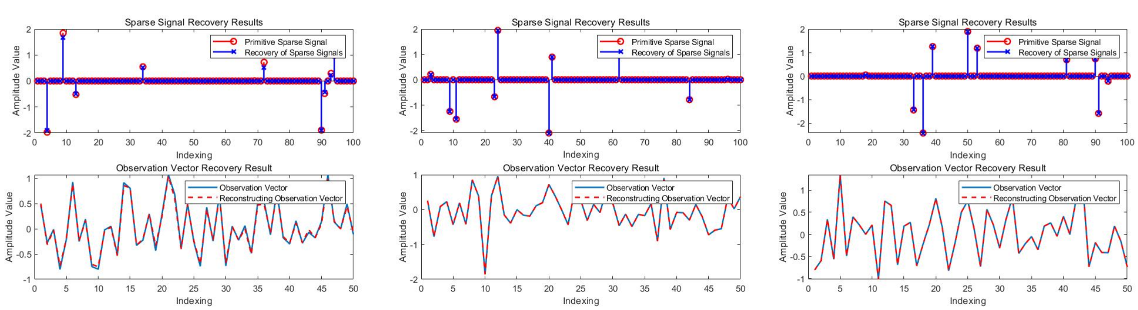
In our both experiments, we employ the Iterative Hard Thresholding (IHT) and Hard Thresholding Pursuit (HTP) algorithms, along with their one-step inertial acceleration algorithms (AIHT, AHTP), and two-step inertial acceleration algorithms (A^{2}IHT, A^{2}HTP).

For algorithms, A^{2}IHT and A^{2}HTP, based on the relationship between α, β, and γ obtained in Section 3, we know that $\alpha=1.08$, $\beta=0.081$, and $\gamma=0.09$ are suitable for A^{2}IHT; and $\alpha=1.04$, $\beta=0.089$, and $\gamma=0.09$ are suitable for A^{2}HTP. According to [17], we know that $\alpha=1.06$, $\beta=0.081$ are suitable for AIHT and AHTP; $\alpha=0.5$ is suitable for IHT, and $\alpha=1$ is suitable for HTP. During the experimental process, we first initialize the measurement vectors as $x^0=x^1=x^2=0$ and then keep iterating until the stopping criterion is met:

$$\frac{\|x_{\text{sparse}}-x_{\text{test}}\|}{\|x_{\text{sparse}}\|}<10^{-4}.$$

Finally, we conduct multiple independent repetitions for each algorithm, and the results are as follows.

4.1. **Signal recovery experiment results.**

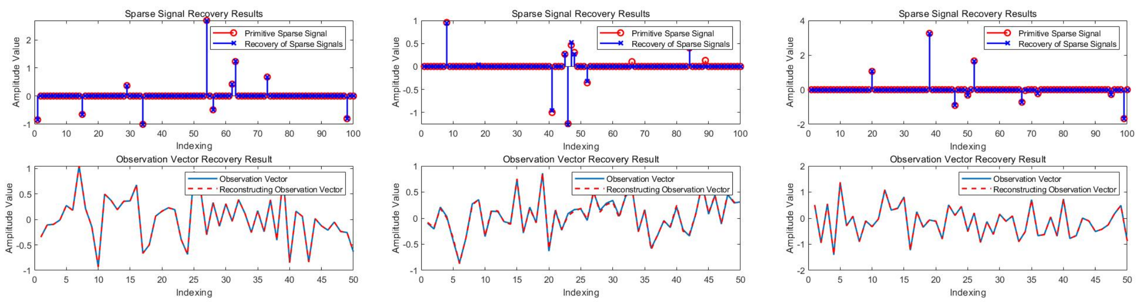


(a) IHT (b) AIHT (c) A^{2}IHT

FIG. 1. Signal recovery of IHT and its accelerated algorithm

Algorithm	CPU Time	Reconstruction Error
IHT	0.024691	0.22868
AIHT	0.016068	2.1585e-08
A^2IHT	0.015091	2.2432e-08

From the experimental results mentioned above, it can be seen that the reconstruction error of the IHT series algorithm is reduced after acceleration, and the time taken is less. Compared to the algorithm without acceleration, the performance has been comprehensively improved.



(a) HTP (b) AHTP (c) A^2HTP

FIG. 2. Signal recovery of HTP and its accelerated algorithms

Algorithm	CPU Time	Reconstruction Error
HTP	0.022365	4.1913e-16
AHTP	0.006380	0.0018313
A^2HTP	0.004779	0.0076893

From the experimental results mentioned above, although the reconstruction error increases after acceleration with the HTP series algorithm, it takes less time and converges faster.

4.2. **Image restoration experiment results.**

(a) original image (b) blurred image

FIG. 3. Original image and blurred image

(a) HTP (b) AHTP (c) A^{2}HTP

FIG. 4. Image Restoration Based on HTP and Its Accelerated Algorithms

(a) IHT (b) AIHT (c) A^{2}IHT

FIG. 5. Image Restoration Based on IHT and Its Accelerated Algorithms

Algorithm	CPU Time	PSNR	NMSE
HTP	5.183147	32.7852	9.4271e-04
AHTP	0.465110	33.7204	7.5999e-04
A^{2}HTP	0.438669	34.1417	6.9268e-04
IHT	0.545246	26.3706	0.0041
AIHT	0.765604	31.8459	0.0012
A^{2}IHT	0.650824	31.2641	0.0013

In the restoration experiment above for 256×256 grayscale images, we intuitively see that the restored images of the IHT series algorithms are relatively blurred, while the image restoration effect of the HTP series algorithms is moderate.

Compared the restoration effects longitudinally, the error of the IHT series algorithms is relatively larger than that of the HTP series algorithms, and the peak signal-to-noise ratio is relatively small. Compared the two types of algorithms horizontally, the accelerated algorithm can better improve the clarity of the original algorithm's restored image, increase the peak signal-to-noise ratio, and reduce the error. Introducing more inertia terms into the accelerated algorithm could reduce the iteration time of the algorithm.

4.3. **Analysis of experimental results.** After signal and image restoration experiments, we intuitively see that the two-inertia algorithm requires less recovery time and is more stable than the one-inertia algorithm, while the recovery effect is similar. Therefore, in the same algorithm background, the two-inertia algorithm performs signal and image restoration faster and more stably.

Funding
This paper was supported by the Innovation and Entrepreneurship Training Program for College Students of Civil Aviation University of China (Grant No. 202310059060).

Acknowledgements
The authors would like to thank the reviewers for their comments, which improved the presentation of this paper significantly.

References

[1] T. Blumensath, E.M. Davies, Iterative hard thresholding for compressed sensing, Appl. Comput. Harmon. Anal. 27 (2009) 265-274.
[2] T. Blumensath, E.M. Davies, Normalized iterative hard thresholding: guaranteed stability and performance, IEEE J. Selected Topics Signal Process. 4 (2010) 298-309.
[3] T. Blumensath, E.M. Davies, Iterative thresholding for sparse approximations, J. Fourier Anal. Appl. 14 (2008) 629-654.
[4] T. Blumensath, Accelerated iterative hard thresholding, Signal Process. 92 (2011) 752-756.
[5] J.D. Blanchard, J. Tanner, K. Wei, CGIHT: conjugate gradient iterative hard thresholding for compressed sensing and matrix completion, Info. Inference 4 (2015) 289-327.
[6] V. Cevher, On accelerated hard thresholding methods for sparse approximation, Proceedings of SPIE - The International Society for Optical Engineering, 2011. DOI: 10.1117/12.894386.
[7] D.L. Donoho, Compressed sensing, IEEE Trans. Info. Theory 52 (2006) 1289-1306.
[8] M. Elad, Sparse and Redundant Rpresentations, Springer, New York. 2010.
[9] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing, Appl. Numer. Harmon. Anal. Birkhäuser New York, 2013. DOI: 10.1007/978-0-8176-4948-7.
[10] S. Foucart, Hard thresholding pursuit: an algorithm for compressive sensing, SIAM J. Numer. Anal. 49 (2011) 2543-2563.
[11] R. Garg, R. Khandekar, Gradient descent with sparsification: An iterative algorithm for sparse recovery with restricted isometry property, International Conference on Machine Learning, 2009. DOI: 10.1145/1553374.1553417.
[12] R. Khanna, A. Kyrillidis, IHT dies hard: provable accelerated iterative hard thresholding, International Conference on Artificial Intelligence and Statistics, 2017. DOI: 10.48550/arXiv.1712.09379.
[13] A. Kyrillidis, V. Cevher, Matrix recipes for hard thresholding methods, J. Math. Imaging Vision. 48 (2014) 235-265.
[14] L. Lessard, B. Recht, A. Packard, Analysis and design of optimization algorithms via integral quadratic constraints, SIAM J. Optim. 26 (2016) 57-95.
[15] B.T. Polyak, Some methods of speeding up the convergence of iteration methods, USSR Comput. Math. Math. Phys. 4 (1964) 1-17.
[16] T. Robert, Regression shrinkage and selection via the lasso, J. Royal Statist. Soc. 58 (2018) 267-288.
[17] Z.F. Sun, J.C. Zhou, Y.B. Zhao, N. Meng, Heavy-Ball-Based hard thresholding algorithms for sparse signal recovery, J. Comput. Appl. Math. 430 (2023) 115264.
[18] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Trans. Info. Theory 53 (2007) 4655-4666.
[19] R. Xin, U.A. Khan, Distributed heavy-ball: A generalization and acceleration of first-order methods with gradient tracking, IEEE Trans. Auto. Control 65 (2019) 2627-2633.
[20] Y.B. Zhao, Sparse Optimization Theory and Methods, CRC Press, 2018.
[21] Y.B. Zhao, Z. Luo, Improved RIP-based bounds for guaranteed performance of two compressed sensing algorithms, Sci. Chin. Math.66 (2023) 1123-1140.
[22] Y.B. Zhao, Optimal k-thresholding algorithms for sparse optimization problems, SIAM J. Optim. 30 (2020) 31-35.
[23] R. Zhang, S. Li, Optimal D-RIP bounds in compressed sensing, Acta Math. Sinica 31 (2015) 755-766.