AIMS Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(9): 16811–16831.
DOI: 10.3934/math.2022923
Received: 02 March 2022
Revised: 23 June 2022
Accepted: 03 July 2022
Published: 14 July 2022

Research article

Compressive hard thresholding pursuit algorithm for sparse signal recovery

Liping Geng[1], **Jinchuan Zhou**[1,*], **Zhongfeng Sun**[1] **and Jingyong Tang**[2]

[1] Department of Statistics, School of Mathematics and Statistics, Shandong University of Technology, Zibo 255000, Shandong, China

[2] School of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, Henan, China

* **Correspondence:** Email: jinchuanzhou@163.com.

Abstract: Hard Thresholding Pursuit (HTP) is one of the important and efficient algorithms for reconstructing sparse signals. Unfortunately, the hard thresholding operator is independent of the objective function and hence leads to numerical oscillation in the course of iterations. To alleviate this drawback, the hard thresholding operator should be applied to a compressible vector. Motivated by this idea, we propose a new algorithm called Compressive Hard Thresholding Pursuit (CHTP) by introducing a compressive step first to the standard HTP. Convergence analysis and stability of CHTP are established in terms of the restricted isometry property of a sensing matrix. Numerical experiments show that CHTP is competitive with other mainstream algorithms such as the HTP, Orthogonal Matching Pursuit (OMP) and Subspace Pursuit (SP) algorithms both in the sparse signal reconstruction ability and average recovery runtime.

Keywords: compressive sensing; sparse signal recovery; restricted isometry property; compressive hard thresholding pursuit; convergence analysis
Mathematics Subject Classification: 90C26, 65F10, 15A29, 94A12.

1. Introduction

One of the important problems in signal processing is to recover an unknown sparse signal from a few measurements, which can be expressed as follows:

$$\begin{aligned} \min \quad & \frac{1}{2}\|y - Ax\|_2^2 \\ \text{s.t.} \quad & \|x\|_0 \le s, \end{aligned} \tag{1.1}$$

where $\|x\|_0$ represents the total number of nonzero entries of $x \in \mathbb{R}^N$, $A \in \mathbb{R}^{m\times N}$ is the measurement matrix with $m \ll N$, and y is the measurement vector. This model has been widely applied in many important areas, including machine learning, compressed sensing, signal processing, pattern recognition, wireless communication, etc. Note that the l_0-norm is not continuous and has combinatorial structure. Hence, the problem (1.1) is known to be NP-hard.

Over the past decades, many efficient algorithms have been proposed for solving the model (1.1), including convex relaxation methods, greedy methods, and thresholding-based methods, to name a few. For example, basis pursuit [1, 2], l_p algorithms [3, 4], alternating projections [5] and conjugate gradient adaptive filtering [6] have been proposed by using convex optimization techniques. Greedy methods include orthogonal matching pursuit (OMP) [7, 8, 9], compressive sampling matching pursuit (CoSaMP) [10] and subspace pursuit (SP) [11, 12]. Thresholding-based methods provide a simple way to ensure the feasibility of iteration, which includes iterative hard thresholding [13, 14, 15], hard thresholding pursuit [16, 17, 18], soft thresholding [19] and Newton-step hard thresholding [20]. Recently, Zhao [21, 22] proposed a new thresholding operator called s-optimal thresholding by connecting the s-thresholding directly to the reduction of the objective function. For more information on theoretical analysis and applications of algorithms for solving sparse signal recovery, see [23, 24, 25, 26].

In this paper, we mainly focus on the hard thresholding algorithms due to their simple structure and low computational cost. One of the key steps in a hard thresholding algorithm is

$$x^{n+1} = \mathcal{H}_s(x^n + A^T(y - Ax^n)),$$

where $\mathcal{H}_s(\cdot)$ stands for the hard thresholding operator, keeping the s-largest absolute components of the vector and zeroing out the rest of the components. The main aim in this step is to ensure the feasibility of iteration. However, three possible questions exist for this step. First, the objective function can increase in some iterations, i.e., $\{\|y - Ax^n\|_2\}$ is not a decreasing sequence. Second, if $x^n + A^T(y - Ax^n)$ is a dense vector, only keeping the s-part would lose too much important information. Third, even for a non-dense vector, it is possible that some useful indexes may be missed, as the difference between the s-largest components and $(s + j)$-largest (with $j > 0$) components of $x^n + A^T(y - Ax^n)$ is very small. To the best of our knowledge, the first question can be solved by using the s-optimal thresholding operator introduced by Zhao in [21], and the second question has been discussed recently in [27] by using partial gradient technology, i.e., full gradients $A^T(y - Ax^n)$ are replaced by partial ones $\mathcal{H}_q(A^T(y - Ax^n))$ with $q \geq s$.

Inspired by these works, we continue to study the aforementioned third question in this paper. Precisely, a little greater signal basis, say β, which is greater than s, is selected in each iteration first. This undoubtedly increases the chance of correctly identifying the support set of a signal, particularly for the case of the s-largest components close to the β-largest components. A β-sparse vector u^{n+1} is attained by solving a least-square problem over the subspace determined by the index set $U^{n+1} := \mathcal{L}_\beta(x^n + A^T(y - Ax^n))$, where $\mathcal{L}_\beta(x)$ stands for the index set of the β largest absolute entries of x. Subsequently, the hard thresholding operator $\mathcal{H}_s(\cdot)$ will be applied to u^{n+1} for reconstructing the s-sparse vector. Finally, to further improve numerical performance, we choose an iteration to seek a vector that best fits the measurements over the prescribed support set by using a pursuit step, also referred to as debiasing or orthogonal projection. This leads to the Compressive Hard Thresholding Pursuit (CHTP) algorithm. Numerical experiments demonstrate that CHTP has better performances when β is slightly greater than s.

The notation used in this paper is standard. For a given set $S \subseteq \{1, 2, \ldots, N\}$, let us denote by $|S|$ the cardinality of S and by $\bar{S} := \{1, 2 \ldots, N\} \backslash S$ the complement of S. For a fixed vector x, x_S is obtained by retaining the elements of x indexed in S and zeroing out the rest of the elements. The support of x is defined as $supp(x) := \{i | x_i \neq 0\}$. Given two sets S_1 and S_2, the symmetrical difference of S_1 and S_2 is denoted by $S_1 \Delta S_2 := (S_1 \backslash S_2) \cup (S_2 \backslash S_1)$. Clearly,

$$\|x_{S_1 \backslash S_2}\|_2 + \|x_{S_2 \backslash S_1}\|_2 \leq \sqrt{2} \|x_{S_1 \Delta S_2}\|_2, \quad \forall x \in \mathbb{R}^N.$$

The paper is organized as follows. The CHTP algorithm is proposed in Section 2. Error estimation and convergence analysis of CHTP are given in Section 3. Numerical results are reported in Section 4.

2. Compressive Hard Thresholding Pursuit

In order to solve problem (1.1), Blumensath and Davies [13] proposed the following Iterative Hard Thresholding (IHT) algorithm:

$$x^{n+1} := \mathcal{H}_s(x^n + A^T(y - Ax^n)).$$

The negative gradient is used as the search direction, and the hard thresholding operator is employed to ensure sparse feasibility. Furthermore, by combining the IHT and CoSaMP algorithms, Foucart [17] proposed an algorithm called Hard Thresholding Pursuit (HTP), which has stronger numerical performance than IHT.

Algorithm: Hard Thresholding Pursuit (HTP).

Input: a measurement matrix A, a measurement vector y and a sparsity level s. Perform the following steps:

S1. Start with an s-sparse $x^0 \in \mathbb{R}^N$, typically $x^0 = 0$;

S2. Repeat

$$U^{n+1} := \mathcal{L}_s(x^n + A^T(y - Ax^n)); \tag{2.1}$$
$$x^{n+1} := \arg min\left\{\|y - Az\|_2 : supp(z) \subseteq U^{n+1}\right\}$$

until a stopping criterion is met.

Output: the s-sparse vector x^*.

In Step (2.1) of HTP, we get $|U^{n+1}| = s$, which results in only up to s positions being taken as the basis at each iteration in the process of restoring vectors. However, it should be noted that the selected basis of s positions could not accurately represent the full gradient information. From the view of theoretical analysis, the less gradient information of the selected basis there is, the more difficult it is for HTP to achieve good numerical performance. In order to further improve the efficiency of HTP, we propose the following CHTP algorithm. It allows us to choose more elements of gradient information at each iteration. In particular, if $\beta = s$, then $x^{n+1} = u^{n+1}$ in CHTP. Hence, CHTP reduces to HTP in this special case.

Algorithm: Compressive Hard Thresholding Pursuit (CHTP).

Input: a measurement matrix A, a measurement vector y and a sparsity level s. Perform the following steps:

S1. Start with an s-sparse $x^0 \in \mathbb{R}^N$, typically $x^0 = 0$;

S2. Repeat

$$\begin{aligned} U^{n+1} &:= \mathcal{L}_\beta(x^n + A^T(y - Ax^n)), \\ u^{n+1} &:= \arg min\Big\{\|y - Az\|_2 : supp(z) \subseteq U^{n+1}\Big\}, \\ T^{n+1} &:= \mathcal{L}_s(u^{n+1}), \\ x^{n+1} &:= \arg\min\Big\{\|y - Az\|_2 : supp(z) \subseteq T^{n+1}\Big\}, \end{aligned}$$

where $\beta \geq s$, until a stopping criterion is met.

Output: the s-sparse vector x^*.

The detailed discussion on numerical experiments including special stopping criteria is given in Section 4.

3. Convergence analysis

Theoretical analysis for CHTP is carried out in terms of the concept of the restricted isometry property (RIP) of a measurement matrix. First, recall that for a given matrix A and two positive integers p, q, the norm $\|A\|_{p\to q}$ is defined as

$$\|A\|_{p\to q} := \max_{\|x\|_p \leq 1} \|Ax\|_q.$$

Definition 3.1. *[1] Let $A \in \mathbb{R}^{m\times N}$ be a matrix with $m < N$. The s-th restricted isometry constant (RIC) of A, denoted by δ_s, is defined as*

$$\delta_s := \max_{S\subseteq\{1,2,\ldots,N\},|S|\leq s} \|A_S^T A_S - I\|_{2\to 2},$$

where A_S is the submatrix of A created by deleting the columns not in the index set S.

The s-th restricted isometry constant can be defined equivalently as the smallest number $\delta > 0$ such that

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2, \quad \forall x \text{ satisfying } \|x\|_0 \leq s. \tag{3.1}$$

In particular, if $\delta_s < 1$, we say that the matrix A has the restricted isometry property (RIP) of order s.

The following inequalities follow from the definition of RIC and play an important role in the theoretical analysis of CHTP.

Lemma 3.1. *[23] Given a vector $v \in \mathbb{R}^N$ and a set $S \subseteq \{1, 2, \ldots, N\}$, one has the following statements.*

(i). $\|((I - A^T A)v)_S\|_2 \leq \delta_t \|v\|_2$ if $|S \cup supp(v)| \leq t$.
(ii). $\|(A^T v)_S\|_2 \leq \sqrt{1 + \delta_t}\|v\|_2$ if $|S| \leq t$.

The following lemma comes from [28], providing an estimate on the Lipschitz constant of the hard thresholding operator. Moreover, [28, Example 2.3] indicates that the constant $(\sqrt{5} + 1)/2$ is tight.

Lemma 3.2. *For any vector $z \in \mathbb{R}^N$ and for any s-sparse vector $x \in \mathbb{R}^N$, one has*

$$\|x - \mathcal{H}_s(z)\|_2 \leq \frac{\sqrt{5} + 1}{2}\|(x - z)_{S\cup Z}\|_2,$$

where $S := supp(x)$ and $Z := supp(\mathcal{H}_s(z))$.

The following result is inspired by [28, Lemma 3.3].

Lemma 3.3. *Let $y = Ax + e$. Given an s-sparse vector z and $\beta \geq s$, define*

$$V := \mathcal{L}_\beta(z + A^T(y - Az)).$$

Then,

$$\|x_{S\setminus V}\|_2 \leq \sqrt{2}(\delta_{2s+\beta}\|x_S - z\|_2 + \sqrt{1 + \delta_{s+\beta}}\|e'\|_2),$$

where $S := \mathcal{L}_s(x)$ and $e' := Ax_{\overline{S}} + e$.

Proof. The case of $S \subseteq V$ holds trivially due to $x_{S\setminus V} = 0$. Now, let us consider the remaining case of $S \subsetneq V$. Define

$$\Phi_{S\setminus V} := \left\|[z + A^T(y - Az)]_{S\setminus V}\right\|_2 \quad \text{and} \quad \Phi_{V\setminus S} := \left\|[z + A^T(y - Az)]_{V\setminus S}\right\|_2. \tag{3.2}$$

Since $|V| = \beta \geq s = |S|$,

$$|S\setminus V| = |S - S \cap V| \leq |V - S \cap V| = |V\setminus S|. \tag{3.3}$$

Meanwhile, from the definition of V, we know that the entries of the vector $z + A^T(y - Az)$ supported on $S\setminus V$ are not greater than the β largest absolute entries, that is,

$$(z + A^T(y - Az))_i \leq (z + A^T(y - Az))_j, \quad \text{for any } i \in S\setminus V,\ j \in V\setminus S.$$

Together with (3.3), this leads to

$$\Phi_{S\setminus V} = \left\|(z + A^T(y - Az))_{S\setminus V}\right\|_2 \leq \left\|(z + A^T(y - Az))_{V\setminus S}\right\|_2 = \Phi_{V\setminus S}. \tag{3.4}$$

Define

$$\Phi_{S\Delta V} := \left\|[(x_S - z) - A^T(y - Az)]_{S\Delta V}\right\|_2.$$

Since $y = Ax_S + e'$ with $e' = Ax_{\bar{S}} + e$,

$$\begin{aligned}\Phi_{S\Delta V} &= \left\|((x_S - z) - A^T(Ax_S + e' - Az))_{S\Delta V}\right\|_2 \\ &= \left\|((I - A^T A)(x_S - z) - A^T e')_{S\Delta V}\right\|_2 \\ &\leq \left\|((I - A^T A)(x_S - z))_{S\Delta V}\right\|_2 + \left\|(A^T e')_{S\Delta V}\right\|_2\end{aligned}$$

$$\leq \quad \delta_{2s+\beta}\|x_S - z\|_2 + \sqrt{1+\delta_{s+\beta}}\|e'\|_2, \tag{3.5}$$

where the last inequality follows from Lemma 3.1 and the fact that

$$|supp(x_S - z) \cup (S\Delta V)| \leq |supp(z) \cup S \cup V| \leq 2s + \beta.$$

Due to $(S\setminus V) \cap (V\setminus S) = \emptyset$, we have

$$\begin{aligned}
\Phi^2_{S\Delta V} &= \left\|((x_S - z) - A^T(y - Az))_{S\Delta V}\right\|_2^2 \\
&= \left\|((x_S - z) - A^T(y - Az))_{S\setminus V}\right\|_2^2 + \left\|((x_S - z) - A^T(y - Az))_{V\setminus S}\right\|_2^2 \\
&= \left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2^2 + \left\|(z + A^T(y - Az))_{V\setminus S}\right\|_2^2 \\
&= \|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\|_2^2 + \Phi^2_{V\setminus S}.
\end{aligned} \tag{3.6}$$

Let us discuss the following two cases separately.
Case 1. $\Phi_{V\setminus S} = 0$. Then, $\Phi_{S\setminus V} = 0$ by (3.4). It follows from (3.2) that $(z + A^T(y - Az))_{S\Delta V} = 0$. So,

$$\begin{aligned}
\|x_{S\setminus V}\|_2 &= \|x_{S\setminus V} - (z + A^T(y - Az))_{S\Delta V}\|_2 \\
&= \Phi_{S\Delta V} \\
&\leq \delta_{2s+\beta}\|x_S - z\|_2 + \sqrt{1+\delta_{s+\beta}}\|e'\|_2,
\end{aligned}$$

where the last inequality results from (3.5).
Case 2. $\Phi_{V\setminus S} > 0$. Let

$$\alpha := \frac{\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\|_2}{\Phi_{V\setminus S}}. \tag{3.7}$$

It follows from (3.6) that

$$\Phi_{S\Delta V} = \sqrt{\left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2^2 + \Phi^2_{V\setminus S}} = \sqrt{1+\alpha^2}\,\Phi_{V\setminus S}. \tag{3.8}$$

Combining (3.7) and (3.8) yields

$$\left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2 = \alpha\Phi_{V\setminus S} = \frac{\alpha}{\sqrt{1+\alpha^2}}\Phi_{S\Delta V}.$$

Furthermore,

$$\begin{aligned}
\Phi^2_{S\setminus V} &= \|(z + A^T(y - Az))_{S\setminus V}\|_2^2 \\
&= \|x_{S\setminus V} - (x_S - z - A^T(y - Az))_{S\setminus V}\|_2^2 \\
&= \|x_{S\setminus V}\|_2^2 - 2\langle x_{S\setminus V}, x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\rangle \\
&\quad + \left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2^2 \\
&\geq \|x_{S\setminus V}\|_2^2 - 2\|x_{S\setminus V}\|_2 \left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2 \\
&\quad + \left\|x_{S\setminus V} - (z + A^T(y - Az))_{S\setminus V}\right\|_2^2
\end{aligned}$$

$$= \|x_{S\setminus V}\|_2^2 - 2\frac{\alpha}{\sqrt{1+\alpha^2}}\Phi_{S\Delta V}\|x_{S\setminus V}\|_2 + \frac{\alpha^2}{1+\alpha^2}\Phi_{S\Delta V}^2. \tag{3.9}$$

On the other hand, it follows from (3.4) and (3.8) that

$$\Phi_{S\setminus V}^2 \le \Phi_{V\setminus S}^2 = \frac{1}{1+\alpha^2}\Phi_{S\Delta V}^2. \tag{3.10}$$

Putting (3.9) and (3.10) together yields

$$\|x_{S\setminus V}\|_2^2 - \frac{2\alpha\Phi_{S\Delta V}}{\sqrt{1+\alpha^2}}\|x_{S\setminus V}\|_2 + \frac{(\alpha^2-1)\Phi_{S\Delta V}^2}{1+\alpha^2} \le 0.$$

Thus,

$$\begin{aligned}\|x_{S\setminus V}\|_2 &\le \frac{1}{2}\left(\frac{2\alpha\Phi_{S\Delta V}}{\sqrt{1+\alpha^2}} + \sqrt{\frac{4\alpha^2\Phi_{S\Delta V}^2}{1+\alpha^2} - \frac{4(\alpha^2-1)\Phi_{S\Delta V}^2}{1+\alpha^2}}\right)\\ &= \frac{1+\alpha}{\sqrt{1+\alpha^2}}\Phi_{S\Delta V}\\ &\le \sqrt{2}\,\Phi_{S\Delta V}, \end{aligned} \tag{3.11}$$

where the last inequality results from

$$\left(\frac{1+\alpha}{\sqrt{1+\alpha^2}}\right)^2 = 1 + \frac{2\alpha}{1+\alpha^2} \le 2.$$

Taking into account (3.5) and (3.11) yields

$$\|x_{S\setminus V}\|_2 \le \sqrt{2}\Phi_{S\Delta V} \le \sqrt{2}(\delta_{2s+\beta}\|x_S - z\|_2 + \sqrt{1+\delta_{s+\beta}}\|e'\|_2).$$

□

Lemma 3.4. *Let v be a vector satisfying $supp(v) \subseteq T$ and $|T| \le s$. For any vector $x \in \mathbb{R}^N$, let $y = Ax_S + e'$ with $e' = Ax_{\overline{S}} + e$ and $S = \mathcal{L}_s(x)$. If z^* satisfies*

$$z^* = \arg\min_z\{\|y - Az\|_2 : supp(z) \subseteq T\},$$

then

$$\|x_S - z^*\|_2 \le \frac{\|(x_S - v)_{\overline{T}}\|_2}{\sqrt{1-\delta_{2s}^2}} + \frac{\sqrt{1+\delta_s}\|e'\|_2}{1-\delta_{2s}}.$$

Proof. According to the definition of z^*, we have

$$[A^T(y - Az^*)]_T = 0.$$

Hence,

$$[A^T(Ax_S + e' - Az^*)]_T = [A^TA(x_S - z^*)]_T + (A^Te')_T = 0.$$

By the triangle inequality and Lemma 3.1, we get

$$\begin{aligned}\|(x_S - z^*)_T\|_2 &= \|(x_S - z^*)_T - [A^TA(x_S - z^*)]_T - (A^Te')_T\|_2 \\ &\leq \|[(I - A^TA)(x_S - z^*)]_T\|_2 + \|(A^Te')_T\|_2 \\ &\leq \delta_{2s}\|(x_S - z^*)\|_2 + \|(A^Te')_T\|_2,\end{aligned}$$

where the last inequality is due to $|supp(x_S - z^*) \cup T| \leq 2s$. This means that

$$\begin{aligned}\|x_S - z^*\|_2^2 &= \|(x_S - z^*)_{\overline{T}}\|_2^2 + \|(x_S - z^*)_T\|_2^2 \\ &\leq \|(x_S - z^*)_{\overline{T}}\|_2^2 + \left(\delta_{2s}\|(x_S - z^*)\|_2 + \|(A^Te')_T\|_2\right)^2 \\ &= \|(x_S - z^*)_{\overline{T}}\|_2^2 + \delta_{2s}^2\|(x_S - z^*)\|_2^2 \\ &\quad +2\delta_{2s}\|(A^Te')_T\|_2\|x_S - z^*\|_2 + \|(A^Te')_T\|_2^2,\end{aligned}$$

i.e.,

$$(1 - \delta_{2s}^2)\|x_S - z^*\|_2^2 - 2\delta_{2s}\|(A^Te')_T\|_2\|x_S - z^*\|_2 - \|(x_S - z^*)_{\overline{T}}\|_2^2 - \|(A^Te')_T\|_2^2 \leq 0.$$

This is a quadratic inequality on $\|x_S - z^*\|_2$. Hence,

$$\begin{aligned}\|x_S - z^*\|_2 &\leq \frac{2\delta_{2s}\|(A^Te')_T\|_2 + \sqrt{4\delta_{2s}^2\|(A^Te')_T\|_2^2 + 4(1 - \delta_{2s}^2)\left(\|(x_S - z^*)_{\overline{T}}\|_2^2 + \|(A^Te')_T\|_2^2\right)}}{2(1 - \delta_{2s}^2)} \\ &= \frac{\delta_{2s}\|(A^Te')_T\|_2 + \sqrt{(1 - \delta_{2s}^2)\|(x_S - z^*)_{\overline{T}}\|_2^2 + \|(A^Te')_T\|_2^2}}{1 - \delta_{2s}^2} \\ &\leq \frac{\delta_{2s}\|(A^Te')_T\|_2 + \sqrt{1 - \delta_{2s}^2}\|(x_S - z^*)_{\overline{T}}\|_2 + \|(A^Te')_T\|_2}{1 - \delta_{2s}^2} \\ &= \frac{\|(x_S - z^*)_{\overline{T}}\|_2}{\sqrt{1 - \delta_{2s}^2}} + \frac{\|(A^Te')_T\|_2}{1 - \delta_{2s}},\end{aligned} \tag{3.12}$$

where the second inequality comes from the fact that $\sqrt{a^2 + b^2} \leq a + b$ for all $a, b \geq 0$. By Lemma 3.1 and $|T| \leq s$, one has

$$\|(A^Te')_T\|_2 \leq \sqrt{1 + \delta_s}\|e'\|_2. \tag{3.13}$$

Note that $z^*_{\bar{T}} = v_{\bar{T}} = 0$. It follows from (3.12) and (3.13) that

$$\|x_S - z^*\|_2 \leq \frac{\|(x_S - z^*)_{\overline{T}}\|_2}{\sqrt{1 - \delta_{2s}^2}} + \frac{\sqrt{1 + \delta_s}\|e'\|_2}{1 - \delta_{2s}} = \frac{\|(x_S - v)_{\overline{T}}\|_2}{\sqrt{1 - \delta_{2s}^2}} + \frac{\sqrt{1 + \delta_s}\|e'\|_2}{1 - \delta_{2s}}.$$

This completes the proof. □

Error bounds and convergence analysis of CHTP are summarized as follows.

Theorem 3.1. *Let $y = Ax + e$ be the measurement of the signal x and e be the measurement error. If*

$$\delta_{2s+\beta} < \sqrt{\frac{1}{\sqrt{4+\sqrt{5}}+2}}, \tag{3.14}$$

then the iterate $\{x^n\}$ generated by CHTP approximates x with

$$\|x^n - x_S\|_2 \le \rho^n \|x^0 - x_S\|_2 + \tau \frac{1-\rho^n}{1-\rho} \|Ax_{\overline{S}} + e\|_2,$$

where

$$\rho := \sqrt{\frac{\delta_{2s+\beta}^2 \left(2 + (\sqrt{5}+1)\delta_{s+\beta}^2\right)}{(1-\delta_{s+\beta}^2)(1-\delta_{2s}^2)}} < 1, \tag{3.15}$$

$$\tau := \sqrt{\frac{2+(\sqrt{5}+1)\delta_{s+\beta}^2}{(1-\delta_{s+\beta})(1-\delta_{2s}^2)}} + \frac{(\sqrt{5}+1)\sqrt{1+\delta_\beta}}{2(1-\delta_{s+\beta})\sqrt{1-\delta_{2s}^2}} + \frac{\sqrt{1+\delta_s}}{1-\delta_{2s}}. \tag{3.16}$$

Proof. For the convenience of discussion, we define $t^{n+1} := \mathcal{H}_s(u^{n+1})$. Then,

$$supp(t^{n+1}) \subseteq T^{n+1} = \mathcal{L}_s(u^{n+1}) \subseteq U^{n+1}.$$

It is clear that $(t^{n+1})_{U^{n+1}} = t^{n+1}$. Hence,

$$\begin{aligned} \|(x_S - t^{n+1})_{U^{n+1}}\|_2 &= \|x_{S\cap U^{n+1}} - t^{n+1}\|_2 \\ &= \|x_{S\cap U^{n+1}} - \mathcal{H}_s(u^{n+1})\|_2 \\ &\le \frac{\sqrt{5}+1}{2} \|(x_{S\cap U^{n+1}} - u^{n+1})_{(S\cap U^{n+1})\cup T^{n+1}}\|_2 \\ &\le \frac{\sqrt{5}+1}{2} \|(x_S - u^{n+1})_{U^{n+1}}\|_2, \end{aligned} \tag{3.17}$$

where the first inequality comes from Lemma 3.2, and the second inequality follows from the fact that $[(S \cap U^{n+1}) \cup T^{n+1}] \subseteq U^{n+1}$.

Since $supp(t^{n+1}) \subseteq U^{n+1}$ and $supp(u^{n+1}) \subseteq U^{n+1}$, $(t^{n+1})_{\overline{U^{n+1}}} = (u^{n+1})_{\overline{U^{n+1}}} = 0$. Thus,

$$\begin{aligned} \|x_S - t^{n+1}\|_2^2 &= \|(x_S - t^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \|(x_S - t^{n+1})_{U_{n+1}}\|_2^2 \\ &= \|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \|(x_S - t^{n+1})_{U^{n+1}}\|_2^2, \\ &\le \left\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\right\|_2^2 + \left(\xi \left\|(x_S - u^{n+1})_{U^{n+1}}\right\|_2\right)^2, \end{aligned} \tag{3.18}$$

where $\xi := (\sqrt{5}+1)/2$, and the last step is due to (3.17).

From Step 2 in CHTP, we know that the following orthogonality condition holds:

$$\langle Au^{n+1} - y, Az\rangle = 0, \quad \text{for any } supp(z) \subseteq U^{n+1},$$

where $y = Ax_S + e'$ and $e' = Ax_{\overline{S}} + e$. Hence, for any z satisfying $supp(z) \subseteq U^{n+1}$, we get

$$\begin{aligned}
&\langle u^{n+1} - x_S, A^T Az\rangle = \langle Au^{n+1} - Ax_S, Az\rangle \\
&= \langle Au^{n+1} - (y - e'), Az\rangle = \langle Au^{n+1} - y, Az\rangle + \langle e', Az\rangle \\
&= \langle e', Az\rangle = \langle A^T e', z\rangle.
\end{aligned} \tag{3.19}$$

Furthermore,

$$\begin{aligned}
\|(x_S - u^{n+1})_{U^{n+1}}\|_2^2 &= \left\langle (u^{n+1} - x_S)_{U^{n+1}}, (u^{n+1} - x_S)_{U^{n+1}} \right\rangle \\
&= \left\langle (u^{n+1} - x_S)_{U^{n+1}}, [(I - A^T A)(u^{n+1} - x_S)]_{U^{n+1}} \right\rangle \\
&\quad + \left\langle (u^{n+1} - x_S)_{U^{n+1}}, [A^T A(u^{n+1} - x_S)]_{U^{n+1}} \right\rangle.
\end{aligned} \tag{3.20}$$

Lemma 3.1 and $\left|supp(u^{n+1} - x_S) \cup U^{n+1}\right| \leq s + \beta$ lead to

$$\left\langle (u^{n+1} - x_S)_{U^{n+1}}, [(I - A^T A)(u^{n+1} - x_S)]_{U^{n+1}} \right\rangle \leq \delta_{s+\beta} \|x_S - u^{n+1}\|_2 \|(x_S - u^{n+1})_{U^{n+1}}\|_2. \tag{3.21}$$

Note that

$$\begin{aligned}
\left\langle (u^{n+1} - x_S)_{U^{n+1}}, [A^T A(u^{n+1} - x_S)]_{U^{n+1}} \right\rangle &= \left\langle (u^{n+1} - x_S)_{U^{n+1}}, A^T A(u^{n+1} - x_S) \right\rangle \\
&= \left\langle u^{n+1} - x_S, A^T A(u^{n+1} - x_S)_{U^{n+1}} \right\rangle \\
&= \left\langle A^T e', (u^{n+1} - x_S)_{U^{n+1}} \right\rangle \\
&= \left\langle (A^T e')_{U^{n+1}}, (u^{n+1} - x_S)_{U^{n+1}} \right\rangle \\
&\leq \|(A^T e')_{U^{n+1}}\|_2 \|(u^{n+1} - x_S)_{U^{n+1}}\|_2,
\end{aligned} \tag{3.22}$$

where the third equality comes from (3.19) since $supp((u^{n+1} - x_S)_{U^{n+1}}) \subset U^{n+1}$. Putting (3.20), (3.21) and (3.22) together yields

$$\|(x_S - u^{n+1})_{U^{n+1}}\|_2 \leq \delta_{s+\beta} \|x_S - u^{n+1}\|_2 + \|(A^T e')_{U^{n+1}}\|_2. \tag{3.23}$$

Therefore,

$$\begin{aligned}
&\|x_S - u^{n+1}\|_2^2 \\
= \; &\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \|(x_S - u^{n+1})_{U^{n+1}}\|_2^2 \\
\leq \; &\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \delta_{s+\beta}^2 \|x_S - u^{n+1}\|_2^2 + 2\delta_{s+\beta} \|(A^T e')_{U^{n+1}}\|_2 \|x_S - u^{n+1}\|_2 + \|(A^T e')_{U^{n+1}}\|_2^2,
\end{aligned}$$

i.e.,

$$(1 - \delta_{s+\beta}^2)\|x_S - u^{n+1}\|_2^2 - 2\delta_{s+\beta}\|(A^T e')_{U^{n+1}}\|_2 \|x_S - u^{n+1}\|_2 - \left(\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \|(A^T e')_{U^{n+1}}\|_2^2\right) \leq 0.$$

This is a quadratic inequality on $\|x_S - u^{n+1}\|_2$. So,

$$\|x_S - u^{n+1}\|_2$$

$$
\begin{aligned}
&\le \frac{2\delta_{s+\beta}\|(A^T e')_{U^{n+1}}\|_2 + \sqrt{4(1-\delta_{s+\beta}^2)\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + 4\|(A^T e')_{U^{n+1}}\|_2^2}}{2(1-\delta_{s+\beta}^2)} \\
&\le \frac{\sqrt{1-\delta_{s+\beta}^2}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + (1+\delta_{s+\beta})\|(A^T e')_{U^{n+1}}\|_2}{1-\delta_{s+\beta}^2} \\
&= \frac{\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2}{\sqrt{1-\delta_{s+\beta}^2}} + \frac{\|(A^T e')_{U^{n+1}}\|_2}{1-\delta_{s+\beta}}, \qquad (3.24)
\end{aligned}
$$

where the second inequality comes from the fact that $\sqrt{a^2+b^2} \le a+b$ for all $a, b \ge 0$.

Due to $supp(u^{n+1}) \subseteq U^{n+1}$, we obtain $(u^{n+1})_{\overline{U^{n+1}}} = 0$. According to Lemma 3.3, we have

$$
\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 = \|x_{S \setminus U^{n+1}}\|_2 \le \sqrt{2}\delta_{2s+\beta}\|x_S - x^n\|_2 + \sqrt{2(1+\delta_{s+\beta})}\|e'\|_2. \qquad (3.25)
$$

Combining (3.23) and (3.24) yields

$$
\begin{aligned}
&\|(x_S - u^{n+1})_{U^{n+1}}\|_2 \\
\le\ & \frac{\delta_{s+\beta}}{\sqrt{1-\delta_{s+\beta}^2}}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + \frac{\delta_{s+\beta}}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2 + \|(A^T e')_{U^{n+1}}\|_2 \\
=\ & \frac{\delta_{s+\beta}}{\sqrt{1-\delta_{s+\beta}^2}}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + \frac{1}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2. \qquad (3.26)
\end{aligned}
$$

Hence, it follows from (3.18) and (3.26) that

$$
\begin{aligned}
&\|x_S - t^{n+1}\|_2^2 \\
\le\ & \|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2^2 + \left(\frac{\xi\delta_{s+\beta}}{\sqrt{1-\delta_{s+\beta}^2}}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + \frac{\xi}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2\right)^2 \\
\le\ & \left(\sqrt{\frac{1+\xi\delta_{s+\beta}^2}{1-\delta_{s+\beta}^2}}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + \frac{\xi}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2\right)^2,
\end{aligned}
$$

where the last step is due to the fact that $a^2 + (b+c)^2 \le (\sqrt{a^2+b^2}+c)^2$ for all $a, b, c \ge 0$, and $\xi^2 - 1 = \xi$ since $\xi = (\sqrt{5}+1)/2$. Taking the square root of both sides of this inequality and using (3.25) yields

$$
\begin{aligned}
\|x_S - t^{n+1}\|_2 &\le \sqrt{\frac{1+\xi\delta_{s+\beta}^2}{1-\delta_{s+\beta}^2}}\|(x_S - u^{n+1})_{\overline{U^{n+1}}}\|_2 + \frac{\xi}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2 \\
&\le \sqrt{\frac{1+\xi\delta_{s+\beta}^2}{1-\delta_{s+\beta}^2}}\left(\sqrt{2}\delta_{2s+\beta}\|x_S - x^n\|_2 + \sqrt{2(1+\delta_{s+\beta})}\|e'\|_2\right) + \frac{\xi}{1-\delta_{s+\beta}}\|(A^T e')_{U^{n+1}}\|_2
\end{aligned}
$$

$$\le \sqrt{\frac{2\delta_{2s+\beta}^2\left(1+\xi\delta_{s+\beta}^2\right)}{1-\delta_{s+\beta}^2}}\|x_S - x^n\|_2 + \left(\sqrt{\frac{2(1+\xi\delta_{s+\beta}^2)}{1-\delta_{s+\beta}}} + \frac{\xi\sqrt{1+\delta_\beta}}{1-\delta_{s+\beta}}\right)\|e'\|_2, \tag{3.27}$$

where the third inequality follows from the fact that $\|(A^T e')_{U^{n+1}}\|_2 \le \sqrt{1+\delta_\beta}\|e'\|_2$ by Lemma 3.1.

Note that $supp(t^{n+1}) \subseteq T^{n+1}$, $|T^{n+1}| \le s$, and

$$x^{n+1} = \arg\min_z \left\{\|y - Az\|_2 : supp(z) \subseteq T^{n+1}\right\}.$$

According to Lemma 3.4, we have

$$\begin{aligned}\|x^{n+1} - x_S\|_2 &\le \frac{1}{\sqrt{1-\delta_{2s}^2}}\|(x_S - t^{n+1})_{\overline{T^{n+1}}}\|_2 + \frac{\sqrt{1+\delta_s}}{1-\delta_{2s}}\|e'\|_2\\ &\le \frac{1}{\sqrt{1-\delta_{2s}^2}}\|x_S - t^{n+1}\|_2 + \frac{\sqrt{1+\delta_s}}{1-\delta_{2s}}\|e'\|_2.\end{aligned}$$

Combining this with (3.27), we obtain

$$\begin{aligned}\|x_S - x^{n+1}\|_2 &\le \sqrt{\frac{2\delta_{2s+\beta}^2\left(1+\xi\delta_{s+\beta}^2\right)}{(1-\delta_{s+\beta}^2)(1-\delta_{2s}^2)}}\|x_S - x^n\|_2\\ &\quad + \left(\sqrt{\frac{2(1+\xi\delta_{s+\beta}^2)}{(1-\delta_{s+\beta})(1-\delta_{2s}^2)}} + \frac{\xi\sqrt{1+\delta_\beta}}{(1-\delta_{s+\beta})(\sqrt{1-\delta_{2s}^2})} + \frac{\sqrt{1+\delta_s}}{1-\delta_{2s}}\right)\|e'\|_2\\ &= \rho\|x_S - x^n\|_2 + \tau\|e'\|_2,\end{aligned}$$

where ρ, τ are given in (3.15) and (3.16), respectively. Hence,

$$\begin{aligned}\|x_S - x^{n+1}\|_2 &\le \rho\|x_S - x^n\|_2 + \tau\|e'\|_2\\ &\le \rho(\rho\|x_S - x^{n-1}\|_2 + \tau\|e'\|_2) + \tau\|e'\|_2\\ &\le \cdots\cdots\cdots\\ &\le \rho^{n+1}\|x_S - x^0\|_2 + \tau\frac{1-\rho^{n+1}}{1-\rho}\|e'\|_2.\end{aligned}$$

Since $\delta_{s+\beta} \le \delta_{2s+\beta}$ and $\delta_{2s} \le \delta_{2s+\beta}$, it is easy to get

$$\rho = \sqrt{\frac{\delta_{2s+\beta}^2\left(2+(\sqrt{5}+1)\delta_{s+\beta}^2\right)}{(1-\delta_{s+\beta}^2)(1-\delta_{2s}^2)}} \le \sqrt{\frac{\delta_{2s+\beta}^2\left(2+(\sqrt{5}+1)\delta_{2s+\beta}^2\right)}{(1-\delta_{2s+\beta}^2)(1-\delta_{2s+\beta}^2)}}.$$

To ensure $\rho < 1$, it suffices to require the right side of the above equation less than 1, which is guaranteed by the following RIP bound:

$$\delta_{2s+\beta} < \sqrt{\frac{1}{\sqrt{4+\sqrt{5}}+2}}.$$

This completes the proof. □

The corresponding result for the noiseless case can be obtained immediately.

Corollary 3.1. *Let $y = Ax$ be the measurement of an s-sparse signal x. If*

$$\delta_{2s+\beta} < \sqrt{\frac{1}{\sqrt{4+\sqrt{5}}+2}},$$

then the iterative sequence $\{x^n\}$ generated by CHTP approximates x with

$$\|x^n - x\|_2 \leq \rho^n \|x^0 - x\|_2,$$

where ρ is given in (3.15).

In the noiseless setting, the foregoing result shows that a sparse signal can be identified by CHTP in a finite number of iterations.

Theorem 3.2. *If*

$$\delta_{2s+\beta} < \sqrt{\frac{1}{\sqrt{4+\sqrt{5}}+2}},$$

then any s-sparse vector $x \in \mathbb{R}^N$ can be recovered by CHTP with $y = Ax$ in at most

$$n = \left\lceil \frac{ln\left(\frac{\sqrt{2}\delta_{2s+\beta}\|x^0-x\|_2}{\theta}\right)}{ln\left(\frac{1}{\rho}\right)} \right\rceil + 1$$

iterations, where ρ is given by (3.15), $\theta := \min_{i\in S} |x_i|$ and $S := supp(x)$.

Proof. It is sufficient to show that

$$|(x^n + A^TA(x - x^n))_k| > |(x^n + A^TA(x - x^n))_t|, \quad \forall k \in S,\ t \in \bar{S}. \tag{3.28}$$

If (3.28) holds, we can obtain $S \subseteq U^{n+1}$ from the definition of U^{n+1}, which leads to $u^{n+1} = x$. Furthermore, we can derive $T^{n+1} = S$ and $x^{n+1} = x$ from the CHTP algorithm directly.

Next, we aim to prove (3.28). According to Lemma 3.1 and Corollary 3.1, we yield

$$\begin{aligned} |((I - A^TA)(x^n - x))_k| + |((I - A^TA)(x^n - x))_t| &\leq \sqrt{2}\|((I - A^TA)(x^n - x))_{\{k,t\}}\|_2 \\ &\leq \sqrt{2}\delta_{2s+1}\|x^n - x\|_2 \\ &\leq \sqrt{2}\delta_{2s+\beta}\|x^n - x\|_2 \\ &\leq \sqrt{2}\delta_{2s+\beta}\rho^n\|x^0 - x\|_2. \end{aligned}$$

It follows that

$$|(x^n + A^TA(x - x^n))_k| = |x_k + ((I - A^TA)(x^n - x))_k|$$

$$
\begin{aligned}
&\geq \ \theta - |((I - A^T A)(x^n - x))_k| \\
&\geq \ \theta - \sqrt{2}\delta_{2s+\beta}\rho^n \|x^0 - x\|_2 + |((I - A^T A)(x^n - x))_t| \\
&= \ \theta - \sqrt{2}\delta_{2s+\beta}\rho^n \|x^0 - x\|_2 + |(x^n + A^T A(x - x^n))_t|,
\end{aligned} \tag{3.29}
$$

where $\theta = \min_{i \in S} |x_i|$, and the last step is due to the fact $x_t = 0$ for $t \in \bar{S}$. Taking

$$
n = \left\lceil \frac{ln\left(\frac{\sqrt{2}\delta_{2s+\beta}\|x^0 - x\|_2}{\theta}\right)}{ln\left(\frac{1}{\rho}\right)} \right\rceil + 1,
$$

we get $\sqrt{2}\delta_{2s+\beta}\rho^n \|x^0 - x\|_2 < \theta$. Combining this with (3.29), we conclude that (3.28) holds. This completes the proof. □

Now, let us discuss the stability of CHTP. Recall first that the error of the best s-term approximation of a vector x is defined as

$$
\sigma_s(x)_p := \inf_v \{\|x - v\|_p : \|v\|_0 \leq s\}.
$$

In particular, according to [23, Theorem 2.5], we have

$$
\sigma_s(z)_2 \leq \frac{1}{2\sqrt{s}} \|z\|_1 \quad \forall z. \tag{3.30}
$$

Theorem 3.3. *Let*

$$
\delta_{2s+\beta} < \sqrt{\frac{1}{\sqrt{4 + \sqrt{5}} + 2}}.
$$

Then, for all $x \in \mathbb{R}^N$ and $e \in \mathbb{R}^m$, the iterate $\{x^n\}$ generated by CHTP with $y = Ax + e$ satisfies

$$
\|x - x^n\|_2 \leq \left(\frac{1}{2\sqrt{t}} + \tau \frac{1 - \rho^n}{1 - \rho} \sqrt{\frac{1 + \delta_t}{t}} \right) \sigma_k(x)_1 + \tau \frac{1 - \rho^n}{1 - \rho} \|e\|_2 + \rho^n \|x_S - x^0\|_2,
$$

where ρ is given by (3.15), and

$$
k := \begin{cases} \frac{s}{2} & \text{if } s \text{ is even} \\ \left[\frac{s}{2}\right] + 1 & \text{if } s \text{ is odd} \end{cases} \quad \text{and} \quad t := s - k.
$$

Moreover, every cluster point x^* of the sequence $\{x^n\}$ satisfies

$$
\|x - x^*\|_2 \leq \left(\frac{1}{2\sqrt{t}} + \frac{\tau}{1 - \rho} \sqrt{\frac{1 + \delta_t}{t}} \right) \sigma_k(x)_1 + \frac{\tau}{1 - \rho} \|e\|_2.
$$

Proof. Let $T_1 := \mathcal{L}_k(x)$, $T_2 := \mathcal{L}_t(x_{\overline{T_1}})$, and

$$
T_3 := \mathcal{L}_t(x_{\overline{T_1 \cup T_2}}), \cdots, T_{l-1} := \mathcal{L}_t(x_{\overline{T_1 \cup T_2 \cup \ldots T_{l-2}}}), \ \ T_l := \mathcal{L}_r(x_{\overline{T_1 \cup T_2 \cup \ldots T_{l-1}}}),
$$

where l, r satisfy $(l-2)t + k + r = N$ and $r \le t$. According to the structure of T_j, it follows from [23, Lemma 6.10] that

$$\|x_{T_j}\|_2 \le \frac{1}{\sqrt{t}}\|x_{T_{j-1}}\|_1, \quad j = 3, \ldots, l. \tag{3.31}$$

Taking into account Theorem 3.1, one has

$$\begin{aligned}\|x_S - x^n\|_2 &\le \rho^n\|x_S - x^0\|_2 + \tau\frac{1-\rho^n}{1-\rho}\|Ax_{\overline{S}} + e\|_2\\ &\le \rho^n\|x_S - x^0\|_2 + \tau\frac{1-\rho^n}{1-\rho}\Big(\|Ax_{T_3}\|_2 + \|Ax_{T_4}\|_2 + \ldots + \|Ax_{T_l}\|_2 + \|e\|_2\Big).\end{aligned}$$

It follows from (3.1) and (3.31) that

$$\sum_{i=3}^{l}\|Ax_{T_i}\|_2 \le \sqrt{1+\delta_t}\sum_{i=3}^{l}\|x_{T_i}\|_2 \le \sqrt{1+\delta_t}\frac{1}{\sqrt{t}}\sum_{i=2}^{l-1}\|x_{T_i}\|_1 \le \sqrt{1+\delta_t}\frac{1}{\sqrt{t}}\|x_{\overline{T_1}}\|_1.$$

Hence,

$$\begin{aligned}\|x_S - x^n\|_2 &\le \rho^n\|x_S - x^0\|_2 + \tau\frac{1-\rho^n}{1-\rho}\left(\sqrt{\frac{1+\delta_t}{t}}\|x_{\overline{T_1}}\|_1 + \|e\|_2\right)\\ &= \rho^n\|x_S - x^0\|_2 + \tau\frac{1-\rho^n}{1-\rho}\left(\sqrt{\frac{1+\delta_t}{t}}\sigma_k(x)_1 + \|e\|_2\right).\end{aligned} \tag{3.32}$$

Using (3.30), we get

$$\|x_{\overline{S}}\|_2 = \sigma_t(x_{\overline{T_1}})_2 \le \frac{1}{2\sqrt{t}}\|x_{\overline{T_1}}\|_1 = \frac{1}{2\sqrt{t}}\sigma_k(x)_1. \tag{3.33}$$

Putting (3.32) and (3.33) together implies

$$\begin{aligned}\|x - x^n\|_2 &= \|x_{\overline{S}} + x_S - x^n\|_2\\ &\le \|x_{\overline{S}}\|_2 + \|x_S - x^n\|_2\\ &\le \frac{1}{2\sqrt{t}}\sigma_k(x)_1 + \rho^n\|x_S - x^0\|_2 + \tau\frac{1-\rho^n}{1-\rho}\Big(\sqrt{\frac{1+\delta_t}{t}}\sigma_k(x)_1 + \|e\|_2\Big)\\ &= \left(\frac{1}{2\sqrt{t}} + \tau\frac{1-\rho^n}{1-\rho}\sqrt{\frac{1+\delta_t}{t}}\right)\sigma_k(x)_1 + \tau\frac{1-\rho^n}{1-\rho}\|e\|_2 + \rho^n\|x_S - x_0\|_2.\end{aligned}$$

Since $\rho < 1$, taking $n \to \infty$ in the above inequality yields

$$\|x - x^*\|_2 \le \left(\frac{1}{2\sqrt{t}} + \frac{\tau}{1-\rho}\sqrt{\frac{1+\delta_t}{t}}\right)\sigma_k(x)_1 + \frac{\tau}{1-\rho}\|e\|_2.$$

This completes the proof. □

4. Numerical experiments

In this section, numerical experiments are carried out to check the effectiveness of CHTP. All experiments were conducted on a personal computer with the processor Intel(R) Core(TM) i5-7200U, CPU @ 2.50 GHz and 4 GB memory. The s-sparse standard Gaussian vector $x^* \in \mathbb{R}^N$ and standard Gaussian matrix $A \in \mathbb{R}^{m\times N}$ with $(m, N) = (400, 800)$ are randomly generated, and the positions of non-zero elements of x^* are uniformly distributed. According to [23, Theorem 9.27], if $0 < \eta, \varepsilon < 1$, and

$$m \geq 2\eta^{-2}[s(1 + \ln(N/s)) + \ln(2\varepsilon^{-1})],$$

then the restricted isometry constant δ_s of $\frac{1}{\sqrt{m}}A$ satisfies

$$\delta_s \leq [2 + \eta g(N, s)] \cdot \eta g(N, s),$$

with probability greater than or equal to $1 - \varepsilon$, where

$$g(N, s) := 1 + [2(1 + \ln(N/s))]^{-1/2}.$$

Hence, to satisfy the restricted isometry bound (3.14) with high probability, we replace (1.1) by the following model throughout the numerical experiments:

$$\min_x \left\{\frac{1}{2}\|\widetilde{y} - \widetilde{A}x\|_2^2 : \|x\|_0 \leq s\right\},$$

where $\widetilde{A} := A/\sqrt{m}$, and $\widetilde{y} := y/\sqrt{m}$. Moreover, the measurement vector y is given by $y = Ax^*$ in the noiseless setting and $y = Ax^* + 0.005e$ in the noisy setting, where $e \in \mathbb{R}^m$ is a standard Gaussian random vector. Unless otherwise stated, all algorithms adopt $x^0 = 0$ as the initial point and the condition

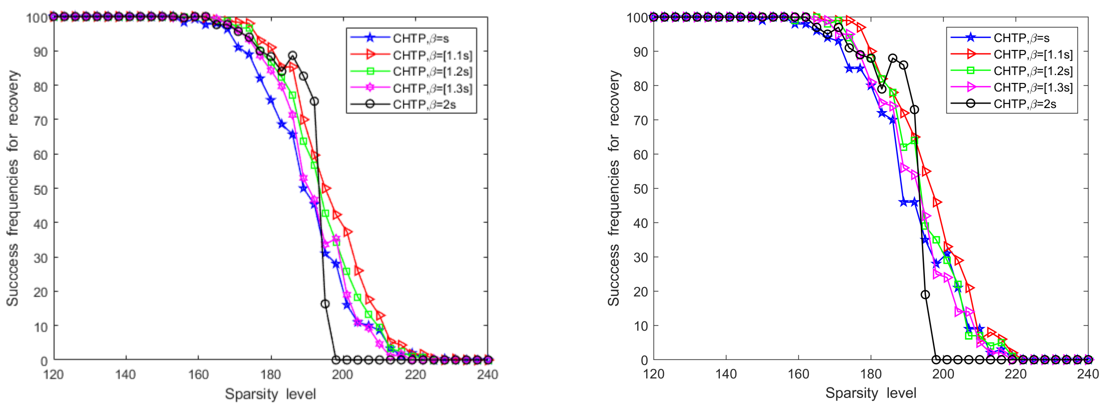
$$\|x^n - x^*\|_2/\|x^*\|_2 \leq 10^{-3}$$

as the successful reconstruction criterion. OMP is performed with s iterations, while other algorithms are performed with at most 50 iterations. For each sparsity level s, 100 random problem instances are tested to investigate the recovery abilities of all algorithms, as shown in Figures 1–3.

The first experiment is performed to illustrate the effect of the parameter β involved in CHTP on the reconstruction ability. Pick the following different values of β in $\{s, [1.1s], [1.2s], [1.3s], 2s\}$ and $s \in \{120, 123, \ldots, 240\}$, where $[\cdot]$ denotes the round function. Numerical results with accurate measurements and inaccurate measurements are shown in Figure 1(a) and 1(b), respectively. Since the results shown in Figure 1(b) are similar to those of Figure 1(a), this indicates that CHTP is stable for weak noise. It is also observed that the recovery ability of CHTP is sensitive to the parameter β, and it becomes worse with the increase of β as $\beta \leq [1.3s]$. In particular, CHTP reduces to HTP as $\beta = s$. Thus, CHTP with $\beta = [1.1s]$, $[1.2s]$ and $[1.3s]$ performs better than HTP in both noiseless and noisy settings. In addition, it should be noted that CHTP with $\beta = 2s$ performs better than HTP as $s \leq 190$, and the success rate of the former suddenly drops to zero as sparsity level s approaches $m/2$. This phenomenon is similar to the performance of SP in Figure 2. Simulations indicate that the performance of the hard thresholding operator can be improved by introducing a compressive step first.

To investigate the effectiveness of CHTP, we compare it with other four mainstream algorithms: HTP, OMP, CoSaMP and SP. The corresponding results are displayed in Figures 2 and 3. Based on the above discussion on the performances of CHTP, the parameter β in this experiment is set as $[1.1s]$. Sparsity level s ranged from 120 to 258 with stepsize 3 in Figure 2(a), and it ranged from 1 to 298 with stepsize 3 in Figure 2(b). The comparison of recovery abilities for different algorithms is displayed in Figure 2(a) with accurate measurements and in Figure 2(b) with inaccurate measurements. Moreover, Figure 2(c) is an enlarged image of Figure 2(b). Figure 2 indicates that CHTP is competitive with HTP, OMP and SP in both noiseless and noisy scenarios. Furthermore, the recovery ability of CHTP can remarkably exceed that of CoSaMP. Figure 2(b) indicates that the sparse signal reconstruction abilities are significantly weakened in the presence of noise for all algorithms as the sparsity level $s \leq 10$.

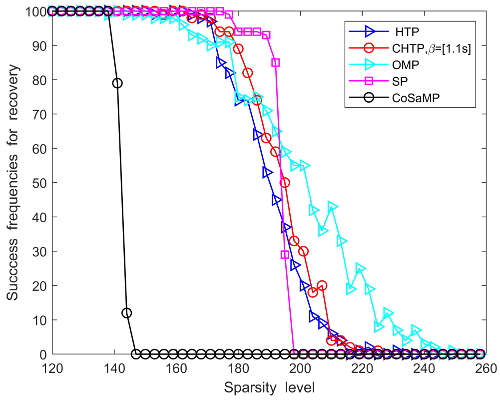
The next problem we examine is the average number of iterations and average recovery runtime for different algorithms with accurate measurements. Due to the number of iterations of OMP being set as s, we just compare all algorithms except OMP in Figure 3, in which $s \leq 120$ is considered to ensure the success rates of all algorithms are 100%. Figure 3 shows that the larger s is taken, the more average number of iterations and computational time are required for all algorithms. From Figure 3(b), we see that the average recovery runtime for all algorithms are close to each other as $s < 50$. Although CHTP needs more iterations than other algorithms in Figure 3(a), it consumes less time than CoSaMP and SP as $s \geq 50$ in Figure 3(b). From this point of view, CHTP is competitive with the other algorithms mentioned above.



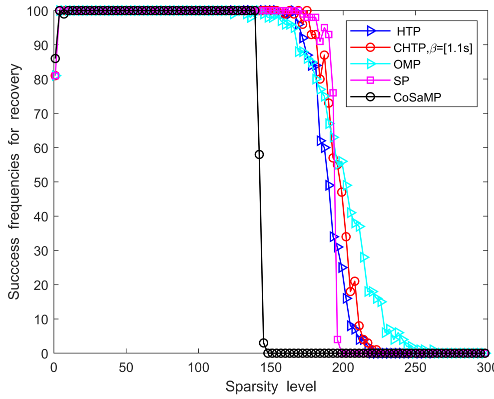
(a) Accurate measurements

(b) Inaccurate measurements

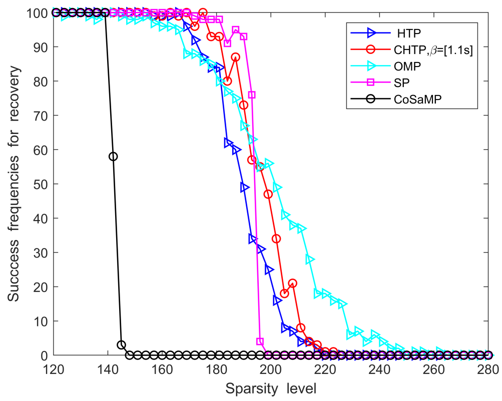
Figure 1. Successful recovery of CHTP with different β value.



(a) Accurate measurements

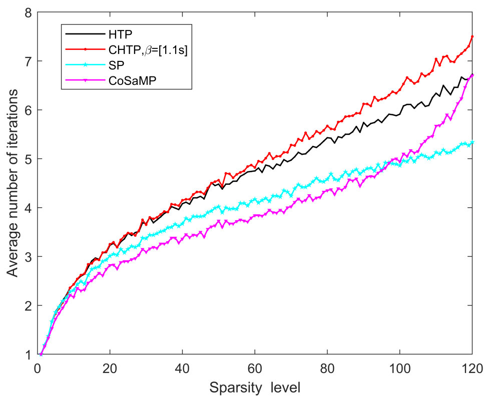


(b) Inaccurate measurements

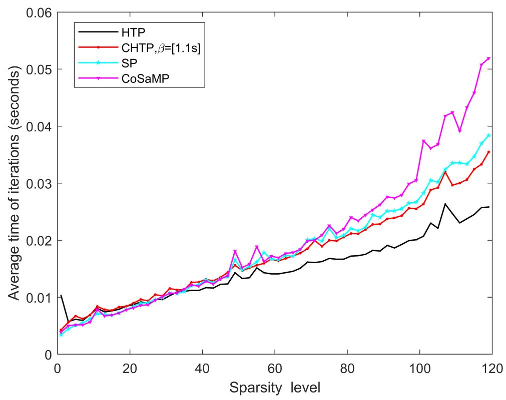


(c) Local enlarged image of (b)

Figure 2. Comparison of successful recovery performances for different algorithms.



(a) Average number of iterations



(b) Average runtime

Figure 3. Comparison of the average number of iterations and recovery runtime for different algorithms.

5. Conclusions

In [27, 21], the authors point out that the hard thresholding operator is independent of the objective function, and this inherent drawback may cause the increase of the objective function in the iterative process. Furthermore, they suggest that the hard thresholding operator should be applied to a compressible vector to overcome this drawback. Motivated by this idea, we proposed the CHTP algorithm in this paper, which reduces to the standard HTP as $\beta = s$. To minimize the negative effect of the hard thresholding operator on the objective function, the orthogonal projection was used twice in CHTP. The convergence analysis of CHTP was established by utilizing the restricted isometry property of the sensing matrix. Numerical experiments indicated that the performance of CHTP with $\beta = \lceil 1.1s \rceil$ is better than in other cases. Simulations showed that CHTP is competitive with other popular algorithms such as HTP, OMP and SP, both in the sparse signal reconstruction ability and the average recovery runtime.

Acknowledgments

The second author's work is supported by the National Natural Science Foundation of China (11771255), Young Innovation Teams of Shandong Province (2019KJI013) and Shandong Province Natural Science Foundation (ZR2021MA066). The fourth author's work is supported by the Natural Science Foundation of Henan Province (222300420520) and Key Scientific Research Projects of Higher Education of Henan Province (22A110020).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. E. J. Candes, T. Tao, Decoding by linear programming, *IEEE T. Inform. Theory,* **51** (2005), 4203–4215. http://dx.doi.org/10.1109/TIT.2005.858979
2. D. L. Donoho, Y. Tsaig, Fast solution of l_1-norm minimization problems when the solution may be sparse, *IEEE T. Inform. Theory,* **54** (2008), 4789–4812. http://dx.doi.org/10.1109/TIT.2008.929958
3. H. Ge, W. Chen, M. K. Ng, New RIP Bounds for recovery of sparse signals with partial support information via weighted ℓ_p-minimization, *IEEE T. Inform. Theory,* **66** (2020), 3914–3928. http://dx.doi.org/10.1109/TIT.2020.2966436
4. A. H. Wan, Uniform RIP conditions for recovery of sparse signals by ℓ_p, $(0 < p \leq 1)$ minimization, *IEEE T. Signal Proces.,* **28** (2020), 5379–5394. http://dx.doi.org/10.1109/TSP.2020.3022822
5. H. Cai, J. F. Cai, T. Wang, G. Yin, Accelerated structured alternating projections for robust spectrally sparse signal recovery, *IEEE T. Signal Proces.,* **69** (2021), 809–821. http://dx.doi.org/10.1109/TSP.2021.3049618

6. C. H. Lee, B. D. Rao, H. Garudadri, A sparse conjugate gradient adaptive filter, *IEEE Signal Proc. Let.,* **27** (2020), 1000–1004. http://dx.doi.org/10.1109/LSP.2020.3000459
7. T. Cai, L. Wang, Orthogonal matching pursuit for sparse signal recovery with noise, *IEEE T. Inform. Theory,* **57** (2011), 4680–4688. http://dx.doi.org/10.1109/TIT.2011.2146090
8. J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE T. Inform. Theory,* **53** (2007), 4655–4666. http://dx.doi.org/10.1109/TIT.2007.909108
9. J. Wen, R. Zhang, W. Yu, Signal-dependent performance analysis of orthogonal matching pursuit for exact sparse recovery, *IEEE T. Signal Proces.,* **68** (2020), 5031–5046. http://dx.doi.org/10.1109/TSP.2020.3016571
10. D. Needell, J. A. Tropp, CoSaMP: Iterative signal recovery from incomplete and inaccurate samples, *Appl. Comput. Harmon. A.,* **26** (2009), 301–321. http://dx.doi.org/10.1145/1859204.1859229
11. W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE T. Infrom. Theory,* **55** (2009), 2230–2249. http://dx.doi.org/10.1109/TIT.2009.2016006
12. N. Han, S. Li, J. Lu, Orthogonal subspace based fast iterative thresholding algorithms for joint sparsity recovery, *IEEE Signal Proces. Let.,* **28** (2021), 1320–1324. http://dx.doi.org/10.1109/LSP.2021.3089434
13. T. Blumensath, M. E. Davies, Iterative hard thresholding for compressed sensing, *Appl. Comput. Harmon. Anal.,* **27** (2009), 265–274. http://dx.doi.org/10.1016/j.acha.2009.04.002
14. M. Fornasier, H. Rauhut, Iterative thresholding algorithms, *Appl. Comput. Harmon. Anal.,* **25** (2008), 187–208. http://dx.doi.org/10.1016/j.acha.2007.10.005
15. S. Foucart, G. Lecue, An IHT algorithm for sparse recovery from subexponential measurements, *IEEE Signal Proces. Let.,* **24** (2017), 1280–1283. http://dx.doi.org/10.1109/LSP.2017.2721500
16. J. L. Bouchot, S. Foucart, P. Hitczenki, Hard thresholding pursuit algorithms: Number of iterations, *Appl. Comput. Harmon. Anal.,* **41** (2016), 412–435. http://dx.doi.org/10.1016/j.acha.2016.03.002
17. S. Foucart, Hard thresholding pursuit: An algorithm for compressive sensing, *SIAM J. Numer. Anal.,* **49** (2011), 2543–2563. http://dx.doi.org/10.1137/100806278
18. X. T. Yuan, P. Li, T. Zhang, Gradient hard thresholding pursuit, *J. Mach. Learn. Res.,* **18** (2018), 1–43.
19. D. L. Donoho, De-noising by soft-thresholding, *IEEE T. Inform. Theory,* **41** (1995), 613–627. http://dx.doi.org/10.1109/18.382009
20. N. Meng, Y. B. Zhao, Newton-step-based hard thresholding algorithms for sparse signal recovery, *IEEE T. Signal Proces.,* **68** (2020), 6594–6606. http://dx.doi.org/10.1109/TSP.2020.3037996
21. Y. B. Zhao, Optimal k-thresholding algorithms for sparse optimization problems, *SIAM J. Optimiz.,* **30** (2020), 31–55. https://doi.org/10.1137/18M1219187
22. Y. B. Zhao, Z. Q. Luo, Analysis of optimal thresholding algorithms for compressed sensing, *IEEE T. Signal Proces.,* **187** (2021), 108–148. https://doi.org/10.1016/j.sigpro.2021.108148
23. S. Foucart, H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer, NY, 2013. https://doi.org/10.1007/978-0-8176-4948-7

24. J. L. Shen, S. Mousavi, Exact support and vector recovery of constrained sparse vectors via constrained matching pursuit, *arXiv Preprint*, 2020. Available from: https://arxiv.org/abs/1903.07236.
25. Y. B. Zhao, *Sparse Optimization Theory and Methods*, Boca Raton, FL, CRC Press, 2018. https://doi.org/10.1201/9781315113142
26. Y. B. Zhao, Z. Q. Luo, Constructing new weighted l_1 algorithms for the sparsest points of polyhedral sets, *Math. Oper. Res.,* **42** (2016), 57–76. https://doi.org/10.1287/moor.2016.0791
27. N. Meng, Y. B. Zhao, M. Kocvara, Partial gradient optimal thresholding algorithms for a class of sparse optimization problems, *arXiv Preprint*, 2022. https://arxiv.org/abs/2107.04319v2
28. Y. B. Zhao, Z. Q. Luo, Improved RIP-based bounds for guaranteed performance of two compressed sensing algorithms, *arXiv Preprint*, 2020. https://arxiv.org/abs/2007.01451