# Improved RIP-Based Bounds for Guaranteed Performance of Two Compressed Sensing Algorithms *

Yun-Bin Zhao† and Zhi-Quan Luo‡

**Abstract** Iterative hard thresholding (IHT) and compressive sampling matching pursuit (CoSaMP) are two mainstream compressed sensing algorithms using the hard thresholding operator. The guaranteed performance of the two algorithms for signal recovery was mainly analyzed in terms of the restricted isometry property (RIP) of sensing matrices. At present, the best known bound using RIP of order $3k$ for guaranteed performance of IHT (with unit stepsize) is $\delta_{3k} < 1/\sqrt{3} \approx 0.5774$, and the bound for CoSaMP using RIP of order $4k$ is $\delta_{4k} < 0.4782$. A fundamental question in this area is whether such theoretical results can be further improved. The purpose of this paper is to affirmatively answer this question and to rigorously show that the above-mentioned RIP bound for guaranteed performance of IHT can be significantly improved to $\delta_{3k} < (\sqrt{5}-1)/2 \approx 0.618$, and the bound for CoSaMP can be improved to $\delta_{4k} < 0.5102$.

**Keywords** Iterative hard thresholding (IHT), compressive sampling matching pursuit (CoSaMP), compressed sensing, guaranteed performance, restricted isometry property (RIP)

**AMS(2010):** 90C30, 65F10, 94A12, 15A29

## 1 Introduction

One of the important tasks in signal processing is to recover (reconstruct) an unknown signal from the linear and nonadaptive measurements acquired for the signal. In many scenarios, the signal can be sparsely represented or approximated in certain transformed domains or over redundant bases [7, 18, 20, 21, 22, 31]. The compressed sensing theory indicates that only small number of measurements might be needed to recover a sparse signal or the significant information of a compressible signal [11, 12, 13, 19]. The recovery of a signal with sparsity usually amounts to a sparse optimization problem, and the numerical methods for solving such a problem are often called compressed sensing algorithms (see, e.g., [12, 21, 22, 24, 42]).

Let $\|z\|_0$ denote the '$\ell_0$-norm' counting the number of nonzero entries of the vector $z \in \mathbb{R}^n$. Let $A$ be an $m \times n$ sensing matrix with $m < n$, and let $y := Ax + \nu$ be the measurements of the target sparse signal $x \in \mathbb{R}^n$ with measurement errors $\nu \in \mathbb{R}^m$. To recover a sparse signal, it is essential and sufficient to find the significant information of the signal which is usually interpreted as a few largest magnitudes of the signal. Thus a typical model for sparse signal recovery can be formulated as the following minimization problem with a sparsity constraint:

$$\min_{z\in\mathbb{R}^n}\{\|Az-y\|_2^2 : \ \|z\|_0 \le k\}, \tag{1}$$

*The research was supported by the National Natural Science Foundation of China (NSFC) under the grant 12071307 and 61571384.

†(Corresponding author) Shenzhen Research Institute of Big Data, Chinese University of Hong Kong, Shenzhen, Guangdong, China (Email: yunbinzhao@cuhk.edu.cn).

‡Shenzhen Research Institute of Big Data, Chinese University of Hong Kong, Shenzhen, Guangdong, China (Email: luozq@cuhk.edu.cn).

where $k$ is a given integer number. The solution $z^*$ of the above problem is $k$-sparse and best fits the linear measurements of $x$. The compressed sensing theory claims that the recovery $z^* = x$ can be achieved under sparsity and other assumptions, and a rapid development of compressed sensing algorithms has been achieved over the past two decades (1) (e.g., [21, 22, 24, 39, 42]). The problem (1) is not only essential to signal recovery, but also important to related problems such as low-rank matrix recovery [14, 17, 25], variable selections in statistics [3, 34, 35], and other sparse optimization problems [2, 30, 42].

The iterative hard thresholding (IHT) [4, 6, 23] is a basic iterative method for the problem (1). The IHT admits several modifications such as the hard thresholding pursuit (HTP) [26], IHT with a fixed stepsize [27], normalized iterative hard thresholding (NIHT) [5], graded IHT [8, 9], and the recent Newton-step-based hard thresholding [33, 47]. More sophisticated methods with hard thresholding were also studied, including the compressive sampling matching pursuit (CoSaMP) [28, 36] and subspace pursuit (SP) [15, 16, 29, 41]. Focused on IHT and CoSaMP, this paper is aiming at achieving a remarkable improvement on the existing results concerning the guaranteed success of the two algorithms for sparse signal recovery.

The guaranteed performance (as well as stability and convergence) of IHT and CoSaMP was widely investigated under the restricted isometry property (RIP) of sensing matrices. The RIP and the associated restricted isometry constant of order $K$, denoted by $\delta_K$, were first introduced by Candès and Tao [12, 13]. The first result for IHT was established by Blumensath and Davies [4] under the RIP condition $\delta_{3k} < 1/\sqrt{32}$, which was improved to $\delta_{3k} < 1/\sqrt{8}$ in [6]. It is important to further relax such a restrictive bound from a theoretical point of view. A remarkable progress in this aspect was made by Foucaut [26] (see also in [24]), who showed that the RIP bound for guaranteed performance of IHT can be improved to $\delta_{3k} < 1/\sqrt{3} \approx 0.5774$. The same bound was also shown for the HTP which is a combination of IHT and orthogonal projection [26]. In this paper, we show that the RIP bound for the theoretical performance of IHT remains not tight, and it can be further improved to $\delta_{3k} < (\sqrt{5}-1)/2 \approx 0.618$. Some evidences observed in this paper point to the conjecture that this new bound might be optimal, i.e., the tightest one for IHT. As for the algorithm CoSaMP, the first result for its guaranteed performance in signal recovery was established in [36] under the condition $\delta_{4k} \leq 0.1$. (Their proof actually implies that their results are valid under the bound $\delta_{4k} < 0.17157$.) This initial result for CoSaMP was significantly improved to $\delta_{4k} < 0.4782$ by Foucart and Rauhut in [24]. In this paper, we further improve the RIP bound for CoSaMP to $\delta_{4k} < 0.5102$. Such improvements are far from being trivial and are achieved by establishing a deep property of the hard thresholding operator. The main contribution of the paper is summarized in the table below:

| Algorithms | Existing results | New results |
|---|---|---|
| IHT | $\delta_{3k} < 0.5774$ | $\delta_{3k} < 0.618$ |
| CoSaMP | $\delta_{4k} < 0.4782$ | $\delta_{4k} < 0.5102$ |

The paper is organized as follows. In Section 2, we describe the IHT and CoSaMP algorithms. In Section 3, we establish a deep property of the hard thresholding operator and use it to show the improved RIP bound for IHT. The improved result for CoSaMP is established in Section 4.

## 2 Algorithms: IHT and CoSaMP

*Notation.* Let us first introduce some notations used in this paper. $\mathbb{R}^n$ denotes the $n$-dimensional Euclidean space and all vectors are understood as column vectors unless otherwise specified. A vector $x \in \mathbb{R}^n$ is said to be $k$-sparse if $\|x\|_0 \leq k$. Given a vector $z \in \mathbb{R}^n$, the operator $\mathcal{H}_k(z) \in \mathbb{R}^n$, called the hard thresholding operator, retains the $k$ largest magnitudes of $z$ and sets other entries to zeros. We use $L_k(z)$ to denote the index set of the $k$ largest absolute entries of the vector $z$, and we use $\text{supp}(z) = \{i : z_i \neq 0\}$ to denote the support of the vector $z$, i.e., the index set of nonzero entries of $z$. The symbols $z^T$ and $A^T$ denote the transpose of the vector $z$ and matrix $A$, respectively. For a given set $S \subseteq \{1, 2, \ldots, n\}$, $|S|$ denotes the cardinality of $S$, and $\overline{S} = \{1, 2, \ldots, n\} \backslash S$ denotes the complement set of $S$. The set difference of $S$ and $U$ is denoted by $S \backslash U = \{i : i \in S, i \notin U\}$. Given $S \subseteq \{1, \ldots, n\}$ and a vector $x \in \mathbb{R}^n$, the vector $x_S \in \mathbb{R}^n$ is obtained by retaining the entries of $x$ indexed by $S$ and zeroing out other entries of $x$.

We now describe the algorithms IHT and CoSaMP. The IHT [4, 6, 23] is a simple iterative scheme for the problem (1), which is stated as follows.

---

**Algorithm 1** Iterative Hard Thresholding (IHT)

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

S1 Choose an initial $k$-sparse vector $x^0$, typically $x^0 = 0$;

S2 Let

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

and repeat until a stopping criterion is met.

Output: the $k$-sparse vector $\hat{x}$.

---

More sophisticated algorithms than IHT can be obtained by integrating orthogonal projections (also called a pursuit step) (see, e.g., [24, 26]). In particular, using hard thresholding and orthogonal projections, Needell and Tropp [36] introduced the so-called compressive sampling matching pursuit (CoSaMP). The CoSaMP was closely related to an earlier greedy method called regularized orthogonal matching pursuit proposed by Needell and Vershynin [37, 38].

---

**Algorithm 2** Compressive Sampling Matching Pursuit (CoSaMP)

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

S1 Choose an initial $k$-sparse vector $x^0$, typically $x^0 = 0$;

S2 Let

$$U^{p+1} = \text{supp}(x^p) \cup L_{2k}(A^T(y - Ax^p)), \tag{CP1}$$

$$z^{p+1} = \arg\min_{z \in \mathbb{R}^n} \{\|y - Az\|_2 : \text{ supp}(z) \subseteq U^{p+1}\}, \tag{CP2}$$

$$x^{p+1} = \mathcal{H}_k(z^{p+1}), \tag{CP3}$$

and repeat until a stopping criterion is met.

Output: the $k$-sparse vector $\hat{x}$.

---

The step (CP2) in CoSaMP is an orthogonal projection which seeks a vector that best fits the measurements over the given support. As pointed out in [43, 44], the orthogonal projection may generally stabilize or speed up the IHT framework.

In the remaining of this paper, we establish improved RIP-based bounds for guaranteed success of IHT and CoSaMP for sparse signal recovery. Such improvements are vital for compressed sensing theory and applications. The RIP condition directly clarifies the scenarios in which the algorithms are guaranteed to recover the signal. Thus an improved RIP bound actually identifies a broader class of signal recovery problems that can be successfully solved by the compressed sensing algorithms. Moreover, the relaxed RIP-based bound can dramatically impact on the number of measurements required for signal recovery. As shown in [1, 13, 32], for Gaussian random sensing matrix $A$ of size $m \times n$ ($m \ll n$), there is a universal constant $C^* > 0$ such that the RIC of $A/\sqrt{m}$ satisfies $\delta_{2k} \leq \delta^* < 1$ with probability at least $1 - \xi$ provided that

$$m \geq C^*(\delta^*)^{-2}(k(1 + \ln(n/k)) + \ln(2\xi - 1)).$$

From this result, it can be seen that the higher the bound $\delta^*$, the less number of measurements are required to ensure the measurement matrix possessing the RIP in high probability.

# 3 Improved RIP bound for IHT

To establish the main results of this paper, we first need to characterize a deep property of the hard thresholding operator $\mathcal{H}_k(\cdot)$. This property will eventually lead to the improved RIP-based bounds that ensure the success of signal recovery with IHT and CoSaMP.

**Lemma 3.1** *For any vector $z \in \mathbb{R}^n$ and any $k$-sparse vector $x \in \mathbb{R}^n$, one has*

$$\|(x - \mathcal{H}_k(z))_{S\setminus S^*}\|_2 \leq \|(x - z)_{S\setminus S^*}\|_2 + \|(x - z)_{S^*\setminus S}\|_2, \tag{2}$$

*where $S = supp(x)$ and $S^* = supp(\mathcal{H}_k(z))$.*

*Proof.* By the definition of $\mathcal{H}_k(\cdot)$, we immediately see that $\|z - \mathcal{H}_k(z)\|_2^2 \leq \|z - d\|_2^2$ for any $k$-sparse vector $d$. In particular, setting $d = z_S$, where $S = \text{supp}(x)$, yields

$$\|z - \mathcal{H}_k(z)\|_2^2 \leq \|z - z_S\|_2^2 = \|z_{\overline{S}}\|_2^2 = \|(z - x)_{\overline{S}}\|_2^2,$$

where the last equality follows from $x_{\overline{S}} = 0$. Note that

$$\|z - \mathcal{H}_k(z)\|_2^2 = \|z - x\|_2^2 + \|x - \mathcal{H}_k(z)\|_2^2 - 2(x - \mathcal{H}_k(z))^T(x - z).$$

Therefore,

$$\|x - \mathcal{H}_k(z)\|_2^2 \leq -\|(z - x)_S\|_2^2 + 2(x - \mathcal{H}_k(z))^T(x - z). \tag{3}$$

Note that $\text{supp}(x - \mathcal{H}_k(z)) \subseteq S \cup S^*$ which can be decomposed into three disjoint sets $S\setminus S^*$, $S^*\setminus S$ and $S^* \cap S$. We also note that $(\mathcal{H}_k(z))_i = z_i$ for every $i \in S^*$, and thus $(\mathcal{H}_k(z))_{S^*\setminus S} = z_{S^*\setminus S}$ and $(\mathcal{H}_k(z))_{S\cap S^*} = z_{S^*\cap S}$. The left-hand side of (3) can be written as

$$\begin{aligned}\|x - \mathcal{H}_k(z)\|_2^2 &= \|[x - \mathcal{H}_k(z)]_{S\setminus S^*}\|_2^2 + \|[x - \mathcal{H}_k(z)]_{S^*\setminus S}\|_2^2 + \|[x - \mathcal{H}_k(z)]_{S^*\cap S}\|_2^2 \\ &= \|[x - \mathcal{H}_k(z)]_{S\setminus S^*}\|_2^2 + \|(x - z)_{S^*\setminus S}\|_2^2 + \|(x - z)_{S^*\cap S}\|_2^2.\end{aligned}$$

The right-hand side of (3) is bounded as

$$\begin{aligned}&-\|(z - x)_S\|_2^2 + 2(x - \mathcal{H}_k(z))^T(x - z) \\ &\quad = -\|(z - x)_S\|_2^2 + 2[(x - \mathcal{H}_k(z))_{S\setminus S^*}]^T(x - z)_{S\setminus S^*} + 2\|(x - z)_{S^*\setminus S}\|_2^2 \\ &\qquad + 2\|(x - z)_{S^*\cap S}\|_2^2 \\ &\quad \leq -\|(z - x)_{S\setminus S^*}\|_2^2 + 2\|[x - \mathcal{H}_k(z)]_{S\setminus S^*}\|_2\|(x - z)_{S\setminus S^*}\|_2 + 2\|(x - z)_{S^*\setminus S}\|_2^2 \\ &\qquad + \|(x - z)_{S^*\cap S}\|_2^2.\end{aligned}$$

Therefore, by substituting the above two relations into (3) and cancelling and rearranging terms, we obtain

$$\begin{aligned}\|[x-\mathcal{H}_k(z)]_{S\backslash S^*}\|_2^2 \leq & -\|(z-x)_{S\backslash S^*}\|_2^2 + 2\|(x-\mathcal{H}_k(z))_{S\backslash S^*}\|_2\|(x-z)_{S\backslash S^*}\|_2 \\ & + \|(x-z)_{S^*\backslash S}\|_2^2.\end{aligned}$$

Thus $\|(x-\mathcal{H}_k(z))_{S\backslash S^*}\|_2$ is smaller than or equal to the largest real root of the quadratic equation

$$Q(r) := r^2 - 2r\|(z-x)_{S\backslash S^*}\|_2 + \|(x-z)_{S\backslash S^*}\|_2^2 - \|(x-z)_{S^*\backslash S}\|_2^2 = 0,$$

to which the largest real root is given by

$$r^* = \|(x-z)_{S\backslash S^*}\|_2 + \|(x-z)_{S^*\backslash S}\|_2.$$

Thus we immediately obtain the inequality (2). □

The next useful result is key to our later analysis.

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

$$\|x-\mathcal{H}_k(z)\|_2 \leq \frac{\sqrt{5}+1}{2}\|(x-z)_{S\cup S^*}\|_2, \tag{4}$$

*where $S = supp(x)$ and $S^* = supp\ (\mathcal{H}_k(z))$.*

*Proof.* By Lemma 3.1, $\|(x-\mathcal{H}_k(z))_{S\backslash S^*}\|_2 \leq \Delta_1 + \Delta_2$, where $\Delta_1$ and $\Delta_2$ are defined as

$$\Delta_1 = \|(x-z)_{S^*\backslash S}\|_2, \quad \Delta_2 = \|(x-z)_{S\backslash S^*}\|_2.$$

Thus,

$$\begin{aligned}\|x-\mathcal{H}_k(z)\|_2^2 &= \|(x-\mathcal{H}_k(z))_{S\cup S^*}\|_2^2 \\ &= \|(x-\mathcal{H}_k(z))_{S^*}\|_2^2 + \|(x-\mathcal{H}_k(z))_{S\backslash S^*}\|_2^2 \\ &\leq \|(x-\mathcal{H}_k(z))_{S^*}\|_2^2 + (\Delta_1+\Delta_2)^2 \\ &= \|(x-\mathcal{H}_k(z))_{S^*\backslash S}\|_2^2 + \|(x-\mathcal{H}_k(z))_{S^*\cap S}\|_2^2 + (\Delta_1+\Delta_2)^2 \\ &= \|(x-z)_{S^*\backslash S}\|_2^2 + \|(x-z)_{S^*\cap S}\|_2^2 + (\Delta_1+\Delta_2)^2.\end{aligned} \tag{5}$$

Let $C := \|(x-z)_{S^*\cup S}\|_2$. We see that $C^2 = \|(x-z)_{S^*\cap S}\|_2^2 + \Delta_1^2 + \Delta_2^2$, and thus

$$\|(x-z)_{S^*\backslash S}\|_2^2 + \|(x-z)_{S^*\cap S}\|_2^2 = \Delta_1^2 + \|(x-z)_{S^*\cap S}\|_2^2 = C^2 - \Delta_2^2.$$

Substituting this relation into (5) yields

$$\|x-\mathcal{H}_k(z)\|_2^2 \leq C^2 + \Delta_1^2 + 2\Delta_1\Delta_2.$$

When $\Delta_1 = 0$, the above inequality immediately implies the bound (4). Thus, without loss of generality, we assume that $\Delta_1 \neq 0$. Denote by $r = \Delta_2/\Delta_1$. Substituting $\Delta_2 = r\Delta_1$ into the above inequality leads to

$$\|x-\mathcal{H}_k(z)\|_2^2 \leq (1+2r)\Delta_1^2 + C^2. \tag{6}$$

We also note that $\Delta_1^2 + \Delta_2^2 \leq C^2$ which together with $\Delta_2 = r\Delta_1$ implies that $\Delta_1^2 \leq C^2/(1+r^2)$. Thus it follows from (6) that

$$\|x-\mathcal{H}_k(z)\|_2^2 \leq \left(1+\frac{1+2r}{1+r^2}\right)C^2 = g(r)C^2, \tag{7}$$

where

$$g(r) := 1 + \frac{1+2r}{1+r^2} = \frac{2(1+r)+r^2}{1+r^2}.$$

Consider the maximum of $g(r)$ over the interval $[0,\infty)$. If $r=0$, then $g(0)=2$. When $r\to\infty$, we see that $g(r)\to 1$. Note that $g(r)$ has a unique stationary point in $[0,\infty)$, i.e., the equation $0 = g'(r) = \frac{2(1-r-r^2)}{(1+r^2)^2}$ has a unique solution given by $r^* = \frac{\sqrt{5}-1}{2} \approx 0.618$ at which

$$g(r^*) = 1 + \frac{1+2r^*}{1+(r^*)^2} = \frac{5+\sqrt{5}}{5-\sqrt{5}} = \left(\frac{\sqrt{5}+1}{2}\right)^2.$$

Thus the maximum value of $g(r)$ over the interval $[0,\infty)$ is given by

$$\max\{g(0), g(r^*), g(\infty)\} = g(r^*) = \left(\frac{\sqrt{5}+1}{2}\right)^2.$$

Therefore it follows from (7) that $\|x-\mathcal{H}_k(z)\|_2 \le \sqrt{g(r^*)}C = (\sqrt{5}+1)C/2$, which is the desired relation (4). □

It is worth mentioning that the result in Lemma 3.2 can be also established by using Theorem 1 in [40] with a different analysis. We now point out that the bound (4) is the tightest of its kind (and hence it cannot be improved further). It is this tightness that makes it possible to establish improved theoretical results for the performance of IHT and CoSaMP.

**Example 3.3** [Tightness of (4)]. Let $0 < \tau < k$ be two given integer numbers. Consider two classes of vectors in $\mathbb{R}^n$ $(n > k+\tau)$ of the following form:

$$z = (\overbrace{1,\ldots,1}^{k}, \overbrace{\varepsilon,\ldots,\varepsilon}^{\tau}, 1/2,\ldots,1/2)^T \in \mathbb{R}^n, \tag{8}$$

$$x = (\overbrace{0,\ldots,0}^{\tau}, \overbrace{1,\ldots,1}^{k-\tau}, \overbrace{\alpha+\varepsilon,\ldots,\alpha+\varepsilon}^{\tau}, 0,\ldots,0)^T \in \mathbb{R}^n, \tag{9}$$

where $\alpha \ge 0$ and $0 < \varepsilon \le 1$ are two parameters.

For the vectors $z$ and $x$ given by (8) and (9), respectively, we see that $x$ is $k$-sparse and $\mathcal{H}_k(z)$ is determined as

$$\mathcal{H}_k(z) = (\overbrace{1,\ldots,1}^{k}, \overbrace{0,\ldots,0}^{n-k})^T \in \mathbb{R}^n.$$

Clearly, $S^* = \mathrm{supp}(\mathcal{H}_k(z)) = \{1,\ldots,k\}$ and $S = \mathrm{supp}(x) = \{\tau+1,\ldots,\tau+k\}$. Thus

$$S^* \cup S = \{1,2,\ldots,k+\tau\}, \ S^* \cap S = \{\tau+1,\ldots,k\},$$

and hence

$$\|(x-z)_{S^*\cup S}\|_2^2 = \tau(1+\alpha^2), \ \|x-\mathcal{H}_k(z)\|_2^2 = \tau\left[1+(\alpha+\varepsilon)^2\right].$$

Consider the ratio

$$\frac{\|x-\mathcal{H}_k(z)\|_2^2}{\|(x-z)_{S^*\cup S}\|_2^2} = \frac{1+(\alpha+\varepsilon)^2}{1+\alpha^2} =: g(\alpha,\varepsilon).$$

We now find the maximum value of the function $g(\alpha,\varepsilon)$ with respect to $\alpha \in [0,\infty)$. It is easy to check that there exists a unique stationary point of $g(\alpha,\varepsilon)$ with respect to $\alpha \in [0,\infty)$. In fact, let $\frac{\partial g(\alpha,\varepsilon)}{\partial \alpha}=0$ which yields $\alpha^2+\alpha\varepsilon-1=0$. Thus the unique stationary point of $g(\alpha,\varepsilon)$ in $[0,\infty)$ is $\alpha^*=\frac{\sqrt{4+\varepsilon^2}-\varepsilon}{2}$, at which

$$g(\alpha^*,\varepsilon)=\frac{1+(\alpha^*+\varepsilon)^2}{1+(\alpha^*)^2}=\frac{1+(\frac{\sqrt{4+\varepsilon^2}+\varepsilon}{2})^2}{1+(\frac{\sqrt{4+\varepsilon^2}-\varepsilon}{2})^2}=\frac{1+\sqrt{\frac{\varepsilon^2}{4+\varepsilon^2}}}{1-\sqrt{\frac{\varepsilon^2}{4+\varepsilon^2}}}=g_1(g_2(\varepsilon)),$$

where the functions $g_1$ and $g_2$ are defined as follows:

$$g_2(\varepsilon)=\sqrt{\frac{\varepsilon^2}{4+\varepsilon^2}},\quad g_1(t)=\frac{1+t}{1-t},$$

where $0\le t<1$. Clearly, $g_1$ and $g_2$ are increasing functions and $g_2(\varepsilon)<1$. Thus $g(\alpha^*,\varepsilon)$ is an increasing function of $\varepsilon$ over $(0,1]$. Therefore, as $\varepsilon$ takes a value close to 1, the maximum of the function is achieved at $\varepsilon=1$. Note that $g(\alpha^*,\varepsilon)=\frac{\sqrt{4+\varepsilon^2}+\varepsilon}{\sqrt{4+\varepsilon^2}-\varepsilon}$. Thus

$$\lim_{\varepsilon\to 1} g(\alpha^*,\varepsilon)=\frac{\sqrt{5}+1}{\sqrt{5}-1}=\left(\frac{\sqrt{5}+1}{2}\right)^2\ge 1+\varepsilon^2$$

for any $\varepsilon\in(0,1]$. As $g(0,\varepsilon)=1+\varepsilon^2$ and $g(\infty,\varepsilon):=\lim_{\alpha\to\infty} g(\alpha,\varepsilon)=1$, the maximum of $g(\alpha,\varepsilon)$ in $[0,\infty)$ is determined as follows:

$$\max_{\alpha\in[0,\infty)} g(\alpha,\varepsilon)=\max\{g(0,\varepsilon),g(\infty,\varepsilon),g(\alpha^*\varepsilon)\}=\max\{1+\varepsilon^2,1,g(\alpha^*,\varepsilon)\}=g(\alpha^*,\varepsilon),$$

which tends to $(\frac{\sqrt{5}+1}{2})^2$ as $\varepsilon\to 1$. This implies that the bound (4) is tight since the ratio $g(\alpha,\varepsilon)$ can approach to $(\frac{\sqrt{5}+1}{2})^2$ for any level of accuracy provided that $\alpha$ and $\varepsilon$ are suitably chosen. In particular, this ratio can achieve the exact value $(\frac{\sqrt{5}+1}{2})^2$ by taking $\varepsilon=1$ and $\alpha=\frac{\sqrt{5}-1}{2}$. In other words, the equality in (4) can be achieved at the following specific vectors:

$$z=(\overbrace{1,\dots,1}^{k+\tau},1/2,\dots,1/2)^T\in\mathbb{R}^n,$$

$$x=(\overbrace{0,\dots,0}^{\tau},\overbrace{1,\dots,1}^{k-\tau},\overbrace{\eta,\dots,\eta}^{\tau},0,\dots,0)^T\in\mathbb{R}^n,$$

where $\eta=(\sqrt{5}+1)/2$. □

In the rest of the paper, we will use the following concept of restricted isometry constant (RIC) and its several useful properties summarized in Lemma 3.5.

**Definition 3.4** [12] *Let $A$ be a given $m\times n$ matrix with $m<n$. The restricted isometry constant (RIC), denoted $\delta_q:=\delta_q(A)$, is the smallest number $\delta\ge 0$ such that*

$$(1-\delta)\|x\|_2^2\le\|Ax\|_2^2\le(1+\delta)\|x\|_2^2$$

*for all $q$-sparse vectors $x\in\mathbb{R}^n$. If $\delta_q<1$, then $A$ is said to satisfy the restricted isometry property (RIP) of order $q$.*

From the definition, we see that $\delta_{q_1}\le\delta_{q_2}$ for $q_1\le q_2$. Implied directly from the above definition are the following properties which are widely utilized in the compressed sensing literature.

**Lemma 3.5** [12, 26, 36] *(i) Let $u, v \in \mathbb{R}^n$ be $s$-sparse and $t$-sparse vectors, respectively. If $supp(u) \cap supp(v) = \emptyset$, then*

$$|u^T A^T A v| \le \delta_{s+t} \|u\|_2 \|v\|_2.$$

*(ii) Let $v \in \mathbb{R}^n$ be a vector and $S \subseteq \{1, 2, \dots, n\}$ be an index set. If $|S \cup supp(v)| \le t$, then*

$$\|[(I - A^T A)v]_S\|_2 \le \delta_t \|v\|_2.$$

*(iii) Let $\Lambda \subseteq \{1, \dots, n\}$ be an index set. If $\Lambda \cap supp(u) = \emptyset$ and $|\Lambda \cup supp(u)| \le t$, then*

$$\|(A^T A u)_\Lambda\|_2 \le \delta_t \|u\|_2.$$

Item (iii) follows from (ii). In fact, when $\Lambda \cap \text{supp}(u) = \emptyset$ which means $u_\Lambda = 0$, one has $\|(A^T A u)_\Lambda\|_2 = \|[(I - A^T A)u]_\Lambda\|_2 \le \delta_t \|u\|_2$.

It should be pointed out that a practical signal $x$ may not necessarily be $k$-sparse. In this case, the $k$-sparse vector $x_S$, where $S = L_k(x)$, is the best $k$-term approximation of $x$. Note that the measurements of the signal $x$ can be represented as

$$y = Ax + \nu = Ax_S + \nu'$$

where $\nu' = Ax_{\overline{S}} + \nu$. This means the measurements $y$ of $x$ with measurement errors $\nu$ can be seen as the measurements of $x_S$ with measurement errors $\nu'$. Thus when the signal is not exactly $k$-sparse, the recovery by solving the model (1) is actually made for the best $k$-term approximation of the signal. We are now ready to show the main result for IHT.

**Theorem 3.6** *Suppose that the sensing matrix $A$ satisfies*

$$\delta_{3k} < \frac{\sqrt{5} - 1}{2} \approx 0.618. \tag{10}$$

*Let $y = Ax + \nu$ be the measurements of $x$ with measurement errors $\nu$ and $S = L_k(x)$. Then the iterates $\{x^p\}$, generated by the IHT, approximate $x$ with error*

$$\|x^p - x_S\|_2 \le \rho^p \|x^0 - x_S\|_2 + \frac{\sqrt{5}+1}{2(1-\rho)} \|A^T \nu'\|_2, \tag{11}$$

*where $\nu' = Ax_{\overline{S}} + \nu$, and the constant $\rho$ is given by*

$$\rho := \left( \frac{\sqrt{5}+1}{2} \right) \delta_{3k} < 1. \tag{12}$$

*Proof.* Denote by $u^p := x^p + A^T(y - Ax^p)$. By the structure of the IHT, $S^{p+1} := \text{supp}(x^{p+1}) = \text{supp}(\mathcal{H}_k(u^p))$. By Lemma 3.2, one has

$$\|x_S - x^{p+1}\|_2 = \|x_S - \mathcal{H}_k(u^p)\|_2 \le \frac{\sqrt{5}+1}{2} \|(x_S - u^p)_{S^{p+1} \cup S}\|_2. \tag{13}$$

We now estimate the term $\|(x_S - u^p)_{S^{p+1} \cup S}\|_2$ which can be bounded as

$$\begin{aligned}
\|(x_S - u^p)_{S^{p+1} \cup S}\|_2 &= \|(x_S - x^p - A^T(y - Ax^p))_{S^{p+1} \cup S}\|_2 \\
&= \|(x_S - x^p - A^T(Ax_S + \nu' - Ax^p))_{S^{p+1} \cup S}\|_2 \\
&= \|[(I - A^T A)(x_S - x^p) - A^T \nu']_{S^{p+1} \cup S}\|_2 \\
&\le \|[(I - A^T A)(x_S - x^p)]_{S^{p+1} \cup S}\|_2 + \|A^T \nu'\|_2 \\
&\le \delta_{3k} \|x_S - x^p\|_2 + \|A^T \nu'\|_2,
\end{aligned}$$

where the last inequality follows from Lemma 3.5 (ii) with $|\mathrm{supp}(x_S - x^p) \cup (S^{p+1} \cup S)| \leq 3k$. Substituting this into (13) yields

$$\begin{aligned} \|x_S - x^{p+1}\|_2 &\leq \frac{\sqrt{5}+1}{2}(\delta_{3k}\|x_S - x^p\|_2 + \|A^T\nu'\|_2) \\ &= \rho\|x_S - x^p\|_2 + \frac{\sqrt{5}+1}{2}\|A^T\nu'\|_2, \end{aligned} \tag{14}$$

where the constant

$$\rho := \left(\frac{\sqrt{5}+1}{2}\right)\delta_{3k} < \left(\frac{\sqrt{5}+1}{2}\right)\left(\frac{\sqrt{5}-1}{2}\right) = 1,$$

where the inequality follows from the condition (10). Therefore, the error bound (11) immediately follows from (14). □

**Remark 3.7** Condition (10) is more relaxed than the best known bound $\delta_{3k} < 1/\sqrt{3} \approx 0.5774$ for IHT established by Foucart [26]. The tightness of (4), as indicated by Example 3.3, is essential to obtain an improved result for IHT. Example 3.3 together with the argument in the proof of Theorem 3.6 points to the conjecture that the new bound $\delta_{3k} < (\sqrt{5}-1)/2$ might be tight for IHT. It is also worth mentioning that some researchers developed the RIP-based bounds for guaranteed success of compressed sensing algorithms according to the geometric rate $\rho \leq 0.5$ instead of $\rho < 1$. From the analysis above, if we require that the geometric rate $\rho$ given in (12) be less than 0.5, namely, $\rho = \left(\frac{\sqrt{5}+1}{2}\right)\delta_{3k} \leq 0.5$, which is guaranteed by $\delta_{3k} \leq (\sqrt{5}-1)/4 \approx 0.309$, then it immediately follows from (14) that

$$\|x^p - x_S\|_2 \leq 0.5^p\|x^0 - x_S\|_2 + (\sqrt{5}+1)\|A^T\nu'\|_2.$$

Shen and Li [40] showed that $\delta_{3k} \leq 0.22$ is a sufficient condition for the guaranteed success of IHT in terms of geometric rate 0.5. In addition, it is implied from the proof of Theorem 6.18 in [24] that $\delta_{3k} \leq 1/(2\sqrt{3}) \approx 0.2886$ is sufficient for guaranteed success of IHT in terms of geometric rate 0.5. Clearly, our result, i.e., $\delta_{3k} < 0.309$, improves these existing results for IHT in terms of geometric rate 0.5.

We also mention that the estimation (11) can also imply the finite convergence and stability of IHT through a standard treatment by using Lemma 6.23 in [24]. We do not state such results here, and we only focus on the establishment of the key estimation like (11) which ensures the convergence of an algorithm and the success of signal recovery with the algorithm.

## 4 Improved RIP bound for CoSaMP

As pointed out in [43, 44], the hard thresholding operator may cause numerical oscillation in some situations due to the fact that performing such a thresholding is independent of the reduction of the objective function $\|y - Az\|_2$. The orthogonal projection is one of the techniques which may alleviate the oscillation problem. Thus it is widely used in traditional hard-thresholding-based algorithms and in the latest algorithmic developments [33, 43, 44]. In this section, we show an improved performance result for CoSaMP. Before doing this, we first state a few useful technical results.

**Lemma 4.1** *Given three constants $\alpha_1, \alpha_2, \alpha_3 \geq 0$ where $\alpha_1 < 1$, if $t$ satisfies the condition $0 \leq t - \alpha_3 \leq \alpha_1\sqrt{t^2 + \alpha_2^2}$, then*

$$t \leq \left(\frac{\alpha_1}{\sqrt{1-\alpha_1^2}}\right)\alpha_2 + \left(\frac{1}{1-\alpha_1}\right)\alpha_3.$$

*Proof.* Under the conditions of the Lemma, $t$ satisfies the condition $(t-\alpha_3)^2 \leq \alpha_1^2(t^2+\alpha_2^2)$, i.e.,

$$\phi(t) := (1-\alpha_1^2)t^2 - 2t\alpha_3 + \alpha_3^2 - \alpha_1^2\alpha_2^2 \leq 0.$$

Thus $t$ is less than or equal to the largest real root of the quadratic equation $\phi(t)=0$. That is,

$$\begin{aligned} t &\leq \frac{2\alpha_3 + \sqrt{4\alpha_3^2 - 4(1-\alpha_1^2)(\alpha_3^2-\alpha_1^2\alpha_2^2)}}{2(1-\alpha_1^2)} \\ &= \frac{\alpha_3 + \sqrt{\alpha_1^2\alpha_3^2 + (1-\alpha_1^2)\alpha_1^2\alpha_2^2}}{1-\alpha_1^2} \\ &\leq \frac{\alpha_3 + \alpha_1\alpha_3 + \alpha_1\alpha_2\sqrt{1-\alpha_1^2}}{1-\alpha_1^2} \\ &= \left(\frac{\alpha_1}{\sqrt{1-\alpha_1^2}}\right)\alpha_2 + \frac{\alpha_3}{1-\alpha_1}, \end{aligned}$$

as desired. □

A fundamental property of orthogonal projection is given as follows. A similar property can be found in the literature, however, the following one is more general than the existing ones.

**Lemma 4.2** *Let $y = Ax+\nu$ be the measurements of the signal $x$, where $\nu$ is a noisy vector. Let $S, \Lambda \subseteq \{1,\ldots,n\}$ be two nonempty index sets and $|S| \leq \tau$, where $\tau$ is an integer number. Let $x^*$ be the solution to the orthogonal projection problem*

$$x^* = \arg\min_{z\in\mathbb{R}^n}\{\|y-Az\|_2 : \ supp(z) \subseteq \Lambda\}. \tag{15}$$

*Let $\Gamma$ be any given index set satisfying $\Lambda \subseteq \Gamma \subseteq \left\{i : \left[A^T(y-Ax^*)\right]_i = 0\right\}$. If $\delta_{|\Gamma|+\tau} < 1$, then*

$$\|(x_S - x^*)_\Gamma\|_2 \leq \frac{\delta_{|\Gamma|+\tau}\|(x_S-x^*)_{\overline{\Gamma}}\|_2}{\sqrt{1-\delta_{|\Gamma|+\tau}^2}} + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+\tau}}, \tag{16}$$

*and hence*

$$\|x_S - x^*\|_2 \leq \frac{\|(x_S-x^*)_{\overline{\Gamma}}\|_2}{\sqrt{1-\delta_{|\Gamma|+\tau}^2}} + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+\tau}}, \tag{17}$$

*where $\nu' = Ax_{\overline{S}} + \nu$.*

*Proof.* Since $x^*$ is the optimal solution to the problem (15), by optimality, we immediately see that $[A^T(y-Ax^*)]_\Lambda = 0$ and thus the set $\left\{i : [A^T(y-Ax^*)]_i = 0\right\}$ is nonempty since it contains $\Lambda$ as a subset. By the definition of $\Gamma$, we have $[A^T(y-Ax^*)]_\Gamma = 0$ which, together with $y = Ax_S + \nu'$ where $\nu' = Ax_{\overline{S}} + \nu$, implies that

$$\begin{aligned} 0 &= \left[A^TA(x_S-x^*) + A^T\nu'\right]_\Gamma \\ &= \left[(A^TA - I)(x_S-x^*)\right]_\Gamma + (x_S-x^*)_\Gamma + [A^T\nu']_\Gamma. \end{aligned} \tag{18}$$

As $\text{supp}(x_S - x^*) \subseteq S \cup \Lambda$ and $\Lambda \subseteq \Gamma$, we see that

$$|\text{supp}(x_S-x^*) \cup \Gamma| \leq |(S\cup\Lambda)\cup\Gamma| = |S\cup\Gamma| \leq |\Gamma| + \tau.$$

Thus it follows from (18) and Lemma 3.5 (ii) that

$$\begin{aligned}\|(x_S - x^*)_\Gamma\|_2 &\le \|[(A^TA - I)(x_S - x^*)]_\Gamma\|_2 + \|A^T\nu'\|_2 \\ &\le \delta_{|\Gamma|+\tau}\|x_S - x^*\|_2 + \|A^T\nu'\|_2 \\ &= \delta_{|\Gamma|+\tau}\sqrt{\|(x_S - x^*)_\Gamma\|_2^2 + \|(x_S - x^*)_{\overline{\Gamma}}\|_2^2} + \|A^T\nu'\|_2. \end{aligned} \tag{19}$$

If $\|(x_S - x^*)_\Gamma\|_2 \le \|A^T\nu'\|_2$, the desired relations (16) and (17) hold trivially. Otherwise if $\|(x_S - x^*)_\Gamma\|_2 > \|A^T\nu\|_2$, then by setting $\alpha_1 = \delta_{|\Gamma|+\tau} < 1, \alpha_2 = \|(x_S - x^*)_{\overline{\Gamma}}\|_2, \alpha_3 = \|A^T\nu'\|_2$, and $t = \|(x_S - x^*)_\Gamma\|_2$, it follows from (19) and Lemma 4.1 that

$$\|(x_S - x^*)_\Gamma\|_2 \le \frac{\delta_{|\Gamma|+\tau}\|(x_S - x^*)_{\overline{\Gamma}}\|_2}{\sqrt{1-\delta^2_{|\Gamma|+\tau}}} + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+\tau}}. \tag{20}$$

Note that

$$\sqrt{(a+b)^2 + c^2} \le \sqrt{a^2+c^2} + b \tag{21}$$

for any $a, b, c \ge 0$. It follows from (20) and (21) that

$$\begin{aligned}\|x_S - x^*\|_2^2 &= \|(x_S - x^*)_\Gamma\|_2^2 + \|(x_S - x^*)_{\overline{\Gamma}}\|_2^2 \\ &\le \left(\frac{\delta_{|\Gamma|+\tau}\|(x_S - x^*)_{\overline{\Gamma}}\|_2}{\sqrt{1-\delta^2_{|\Gamma|+\tau}}} + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+k}}\right)^2 + \|(x_S - x^*)_{\overline{\Gamma}}\|_2^2 \\ &\le \left(\sqrt{\frac{\delta^2_{|\Gamma|+\tau}\|(x_S - x^*)_{\overline{\Gamma}}\|_2^2}{1-\delta^2_{|\Gamma|+\tau}} + \|(x_S - x^*)_{\overline{\Gamma}}\|_2^2} + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+\tau}}\right)^2 \\ &= \left(\frac{1}{\sqrt{1-\delta^2_{|\Gamma|+\tau}}}\|(x_S - x^*)_{\overline{\Gamma}}\|_2 + \frac{\|A^T\nu'\|_2}{1-\delta_{|\Gamma|+\tau}}\right)^2, \end{aligned}$$

as desired. □

We now show the next technical result which together with Lemma 4.2 eventually yields an improved RIP-based bound for the guaranteed success of CoSaMP.

**Lemma 4.3** *Let $y = Ax + \nu$ be the measurements of $x$, where $\nu$ is a noise vector, and let $S = L_k(x)$. Given a $k$-sparse vector $x^p$ with $S^p = supp(x^p)$ and the index set*

$$T = L_\beta(A^T(y - Ax^p)),$$

*where $\beta \ge 2k$ is an integer number, if $\delta_{2k+\beta} < 1$, then one has*

$$\|(x^p - x_S)_{\overline{T}}\|_2 \le \sqrt{2}\left(\delta_{2k+\beta}\|x^p - x_S\|_2 + \|A^T\nu'\|_2\right), \tag{22}$$

*where $\nu' = Ax_{\overline{S}} + \nu$.*

*Proof.* Let $S, S^p$, and $T$ be defined as in the lemma. If $S \cup S^p \subseteq T$, then $\overline{T} \subseteq \overline{S \cup S^p}$ which implies that $\|(x^p - x_S)_{\overline{T}}\|_2 \le \|(x^p - x_S)_{\overline{S\cup S^p}}\|_2 = 0$. Thus the relation (22) holds trivially. We only need to show (22) for the case $S \cup S^p \not\subseteq T$. Thus in the remaining proof, we assume that $S \cup S^p \not\subseteq T$. It is convenient to define

$$\Omega := \|[A^T(y - Ax^p)]_{(S\cup S^p)\setminus T}\|_2. \tag{23}$$

As the cardinality $|T| = \beta \geq 2k \geq |S \cup S^p|$, we see that

$$\begin{aligned}|(S \cup S^p)\backslash T| &= |S \cup S^p| - |(S \cup S^p) \cap T| \\ &\leq |T| - |(S \cup S^p) \cap T| = |T\backslash(S \cup S^p)|. \end{aligned} \tag{24}$$

This means the number of elements in $(S\cup S^p)\backslash T$ is less than or equal to the number of elements in $T\backslash(S\cup S^p)$. By the definition of $T$, the entries of the vector $A^T(y-Ax^p)$ supported on $(S\cup S^p)\backslash T$ are not among the $\beta$ largest absolute entries of the vector. This together with (24) implies that

$$\|[A^T(y - Ax^p)]_{(S\cup S^p)\backslash T}\|_2 \leq \|[A^T(y - Ax^p)]_{T\backslash(S\cup S^p)}\|_2. \tag{25}$$

Denote by

$$\Omega^* = \|[(x_S - x^p) - A^T(y - Ax^p)]_{(S\cup S^p)\Delta T}\|_2,$$

where $\Delta$ denotes the symmetric difference of two sets, i.e., $(S\cup S^p)\Delta T = ((S\cup S^p)\backslash T)\cup(T\backslash(S\cup S^p))$. Note that

$$|(S \cup S^p) \cup ((S \cup S^p)\Delta T)| \leq |(S \cup S^p) \cup T| \leq 2k + \beta.$$

As $y = Ax_S + \nu'$, by Lemma 3.5 (iii), we have

$$\begin{aligned}\Omega^* &= \|[(I - A^TA)(x_S - x^p) + A^T\nu']_{(S\cup S^p)\Delta T}\|_2 \\ &\leq \|[(I - A^TA)(x_S - x^p)]_{(S\cup S^p)\Delta T}\|_2 + \|(A^T\nu')_{(S\cup S^p)\Delta T}\|_2 \\ &\leq \delta_{2k+\beta}\|x_S - x^p\|_2 + \|A^T\nu'\|_2. \end{aligned} \tag{26}$$

Let $\widehat{\Omega} = \|[A^T(y - Ax^p)]_{T\backslash(S\cup S^p)}\|_2$ and

$$W = (x_S - x^p)_{(S\cup S^p)\backslash T} - [A^T(y - Ax^p)]_{(S\cup S^p)\backslash T}.$$

Then by the definition of $\Omega^*$, we see that

$$\begin{aligned}(\Omega^*)^2 &= \|[(x_S - x^p) - A^T(y - Ax^p)]_{(S\cup S^p)\backslash T}\|_2^2 \\ &\quad + \|[(x_S - x^p) - A^T(y - Ax^p)]_{T\backslash(S\cup S^p)}\|_2^2 \\ &= \|[(x_S - x^p) - A^T(y - Ax^p)]_{(S\cup S^p)\backslash T}\|_2^2 + \|[A^T(y - Ax^p)]_{T\backslash(S\cup S^p)}\|_2^2 \\ &= \|W\|_2^2 + \widehat{\Omega}^2 \end{aligned} \tag{27}$$

where the second equality follows from the fact $(x_S - x^p)_{T\backslash(S\cup S^p)} = 0$. There are only two cases.

Case 1: $\widehat{\Omega} = 0$. Then it follows from (25) that $[A^T(y - Ax^p)]_{(S\cup S^p)\backslash T} = 0$. In this case, we see that $W = (x_S - x^p)_{(S\cup S^p)\backslash T}$. Since $\|(x_S - x^p)_{\overline{T}}\|_2 = \|(x_S - x^p)_{(S\cup S^p)\backslash T}\|_2$, we immediately have that

$$\|(x_S - x^p)_{\overline{T}}\|_2 = \|W\|_2 \leq \Omega^* \leq \delta_{2k+\beta}\|x_S - x^p\|_2 + \|A^T\nu'\|_2,$$

where $\|W\|_2 \leq \Omega^*$ follows from (27) and the last inequality above follows from (26). Thus the bound (22) holds trivially for this case.

Case 2: $\widehat{\Omega} \neq 0$. Then let $\gamma$ be the ratio of $\|W\|_2$ and $\widehat{\Omega}$, i.e. $\|W\|_2 = \gamma\widehat{\Omega}$. Substituting this into (27), we immediately obtain

$$\widehat{\Omega} = \frac{\Omega^*}{\sqrt{1+\gamma^2}}, \quad \|W\|_2 = \frac{\gamma\Omega^*}{\sqrt{1+\gamma^2}}. \tag{28}$$

Therefore, by the definition of $\Omega$ in (23), we have

$$\begin{aligned}\Omega^2 &= \|(x_S - x^p)_{(S\cup S^p)\setminus T} - \left[(x_S - x^p)_{(S\cup S^p)\setminus T} - [A^T(y - Ax^p)]_{(S\cup S^p)\setminus T}\right] \|_2^2 \\ &= \|(x_S - x^p)_{(S\cup S^p)\setminus T} - W\|_2^2 \\ &= \|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2^2 - 2\left[(x_S - x^p)_{(S\cup S^p)\setminus T}\right]^T W + \|W\|_2^2 \\ &\geq \|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2^2 - 2\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2\|W\|_2 + \|W\|_2^2. \end{aligned} \tag{29}$$

By (25) and (28), the left-hand side of the above inequality can be bounded as

$$\Omega^2 \leq \widehat{\Omega}^2 = \frac{(\Omega^*)^2}{1+\gamma^2}.$$

Thus substituting $\|W\|_2$ in (28) into (29) leads to

$$\frac{(\Omega^*)^2}{1+\gamma^2} \geq \|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2^2 + \frac{\gamma^2(\Omega^*)^2}{1+\gamma^2} - 2\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2 \frac{\gamma\Omega^*}{\sqrt{1+\gamma^2}}.$$

Simplifying the above inequality yields

$$\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2^2 - \frac{2\gamma\Omega^*}{\sqrt{1+\gamma^2}}\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2 + \frac{(\gamma^2-1)(\Omega^*)^2}{1+\gamma^2} \leq 0,$$

which is a quadratic inequality of $\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2$, and thus

$$\begin{aligned}\|(x_S - x^p)_{(S\cup S^p)\setminus T}\|_2 &\leq \frac{1}{2}\left(\frac{2\gamma\Omega^*}{\sqrt{1+\gamma^2}} + \sqrt{\frac{4\gamma^2(\Omega^*)^2}{1+\gamma^2} - \frac{4(\gamma^2-1)(\Omega^*)^2}{1+\gamma^2}}\right) \\ &= \frac{1+\gamma}{\sqrt{1+\gamma 2}}\Omega^* \leq \left(\max_{0\leq\gamma<\infty} \frac{1+\gamma}{\sqrt{1+\gamma 2}}\right)\Omega^* = \sqrt{2}\Omega^*,\end{aligned}$$

where the maximum of the univariate function of $\gamma$ achieves at $\gamma = 1$. By combining (26) and the inequality above, we immediately obtain the desired inequality in the lemma. □

The main result for CoSaMP is summarized as follows.

**Theorem 4.4** *If the restricted isometry constant of the sensing matrix $A$ satisfies that*

$$\delta_{4k} < \sqrt{\frac{2}{\sqrt{13+4\sqrt{5}}+3}} \approx 0.5102, \tag{30}$$

*then the iterates $\{x^p\}$, generated by the CoSaMP, satisfy that*

$$\|x_S - x^p\|_2 \leq \rho^p\|x_S - x^0\|_2 + \frac{C}{1-\rho}\|A^T\nu'\|_2,$$

*where the constants $\rho$ and $C$ are given as*

$$\rho = \delta_{4k}\sqrt{\frac{2+(\sqrt{5}+1)\delta_{4k}^2}{1-\delta_{4k}^2}} < 1 \tag{31}$$

*and*

$$C = \sqrt{\frac{2+(\sqrt{5}+1)\delta_{4k}^2}{1-\delta_{4k}^2}} + \frac{\sqrt{5}+1}{2(1-\delta_{4k})}.$$

*Proof.* Let $U^{p+1}, z^{p+1}$ and $x^{p+1}$ are given, respectively, by (CP1)–(CP3) of CoSaMP. From the structure of CoSaMP, we see that $S^p = \text{supp}(x^p) \subseteq U^{p+1}$ and $S^{p+1} = \text{supp}(x^{p+1}) = \text{supp}(\mathcal{H}_k(z^{p+1})) \subseteq U^{p+1}$ (so $(x^{p+1})_{U^{p+1}} = x^{p+1}$). By Lemma 3.2, we have

$$\begin{aligned}
\|(x_S - x^{p+1})_{U^{p+1}}\|_2 &= \|x_{S\cap U^{p+1}} - x^{p+1}\|_2 = \|x_{S\cap U^{p+1}} - \mathcal{H}_k(z^{p+1})\|_2 \\
&\leq \frac{\sqrt{5}+1}{2}\|(x_{S\cap U^{p+1}} - z^{p+1})_{(S\cap U^{p+1})\cup S^{p+1}}\|_2 \\
&\leq \eta\|(x_{S\cap U^{p+1}} - z^{p+1})_{U^{p+1}}\|_2 \\
&\qquad\qquad (\text{since } (S\cap U^{p+1})\cup S^{p+1} \subseteq U^{p+1}) \\
&= \eta\|(x_S - z^{p+1})_{U^{p+1}}\|_2,
\end{aligned}$$

where $\eta = (\sqrt{5}+1)/2$. Also, since $\text{supp}(x^{p+1}) \subseteq U^{p+1}$, we have that $(x^{p+1})_{\overline{U^{p+1}}} = 0 = (z^{p+1})_{\overline{U^{p+1}}}$. This together with the above relation implies that

$$\begin{aligned}
\|x_S - x^{p+1}\|_2^2 &= \|(x_S - x^{p+1})_{\overline{U^{p+1}}}\|_2^2 + \|(x_S - x^{p+1})_{U^{p+1}}\|_2^2 \\
&= \|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2^2 + \|(x_S - x^{p+1})_{U^{p+1}}\|_2^2 \\
&\leq \|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2^2 + \left(\eta\|(x_S - z^{p+1})_{U^{p+1}}\|_2\right)^2. \qquad (32)
\end{aligned}$$

Consider the step (CP2) of the CoSaMP. Applying Lemma 4.2 with $\Gamma = \Lambda = U^{p+1}$, $S = L_k(x)$, $x^* = z^{p+1}$, $\tau = k$ and $|\Gamma| = |U^{p+1}| \leq 3k$, we conclude that if $\delta_{4k} < 1$, one has

$$\|(x_S - z^{p+1})_{U^{p+1}}\|_2 \leq \frac{\delta_{4k}\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2}{\sqrt{1-\delta_{4k}^2}} + \frac{\|A^T\nu'\|_2}{1-\delta_{4k}}. \qquad (33)$$

By combining (32) and (33) and using the inequality (21), we obtain

$$\begin{aligned}
\|x_S - x^{p+1}\|_2^2 &\leq \|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2^2 + \left(\frac{\eta\delta_{4k}\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2}{\sqrt{1-\delta_{4k}^2}} + \frac{\eta\|A^T\nu'\|_2}{1-\delta_{4k}}\right)^2 \\
&\leq \left(\sqrt{\left(1 + \frac{\eta^2\delta_{4k}^2}{1-\delta_{4k}^2}\right)\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2^2} + \frac{\eta\|A^T\nu'\|_2}{1-\delta_{4k}}\right)^2 \\
&= \left(\sqrt{\frac{1+\eta\delta_{4k}^2}{1-\delta_{4k}^2}}\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2 + \frac{\eta\|A^T\nu'\|_2}{1-\delta_{4k}}\right)^2, \qquad (34)
\end{aligned}$$

where the equality follows from the fact $\eta^2 - 1 = \eta$. In the remainder of this proof, it is sufficient to bound the term $\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2$ in terms of $\|x_S - x^p\|_2$. Note that $(z^{p+1})_{\overline{U^{p+1}}} = 0$ and $S^p = \text{supp}(x^p) \subseteq U^{p+1}$ which implies that $(x^p)_{\overline{U^{p+1}}} = 0$. Thus

$$\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2 = \|(x_S)_{\overline{U^{p+1}}}\|_2 = \|(x_S - x^p)_{\overline{U^{p+1}}}\|_2. \qquad (35)$$

Setting $\beta = 2k$ and $T = \text{supp}[\mathcal{H}_\beta(A^T(y - Ax^p))]$, it follows from Lemma 4.3 that the CoSaMP satisfies the relation

$$\|(x_S - x^p)_{\overline{T}}\|_2 \leq \sqrt{2}(\delta_{4k}\|x^p - x_S\|_2 + \|A^T\nu'\|_2). \qquad (36)$$

Note that $T \subseteq U^{p+1}$ which implies that $\overline{U^{p+1}} \subseteq \overline{T}$. Thus

$$\|(x_S - x^p)_{\overline{U^{p+1}}}\|_2 \leq \|(x_S - x^p)_{\overline{T}}\|_2. \qquad (37)$$

Merging the three relations (35)–(37) leads to

$$\|(x_S - z^{p+1})_{\overline{U^{p+1}}}\|_2 \le \sqrt{2}(\delta_{4k}\|x^p - x_S\|_2 + \|A^T\nu'\|_2).$$

Therefore, it follows from (34) that

$$\begin{aligned}\|x_S - x^{p+1}\|_2 &\le \sqrt{\frac{2(1+\eta\delta_{4k}^2)}{1-\delta_{4k}^2}}(\delta_{4k}\|x_S - x^p\|_2 + \|A^T\nu'\|_2) + \frac{\eta}{1-\delta_{4k}}\|A^T\nu'\|_2 \\ &= \rho\|x_S - x^p\|_2 + C\|A^T\nu'\|_2,\end{aligned}$$

where the constants $\rho$ and $C$ are given by

$$\rho = \delta_{4k}\sqrt{\frac{2+(\sqrt{5}+1)\delta_{4k}^2}{1-\delta_{4k}^2}},$$

$$C = \sqrt{\frac{2(1+\eta\delta_{4k}^2)}{1-\delta_{4k}^2}} + \frac{\eta}{1-\delta_{4k}} = \sqrt{\frac{2+(\sqrt{5}+1)\delta_{4k}^2}{1-\delta_{4k}^2}} + \frac{\sqrt{5}+1}{2(1-\delta_{4k})}.$$

We now prove that if $\delta_{4k}$ satisfies (30), then the constant $\rho < 1$. In fact, to ensure $\rho < 1$, it is sufficient to ensure that

$$\delta_{4k}\sqrt{\frac{2(1+\eta\delta_{4k}^2)}{1-\delta_{4k}^2}} < 1,$$

which, by squaring, can be written as

$$2\eta\delta_{4k}^4 + 3\delta_{4k}^2 - 1 < 0.$$

Thus

$$\delta_{4k}^2 < \frac{\sqrt{9+4(\sqrt{5}+1)}-3}{2(\sqrt{5}+1)} = \frac{2}{\sqrt{9+4(\sqrt{5}+1)}+3},$$

and hence

$$\delta_{4k} < \sqrt{\frac{2}{\sqrt{13+4\sqrt{5}}+3}} \approx 0.5102.$$

The proof is complete. □

**Remark 4.5** The existing bound $\delta_{4k} < 0.4782$ for guaranteed performance of CoSaMP was shown by Foucart and Rauhut (see, e.g., Theorem 6.18 in [24]). Their bound was improved to $\delta_{4k} < 0.5102$ in the theorem. Moveover, in terms of geometric rate 0.5, Foucart and Rauhut's bound is $\delta_{4k} < 0.299$, which was slightly improved to $\delta_{4k} < 0.301$ by Shen and Li [40]. Let us find out our new bound if the geometric rate $\rho \le 0.5$ is required, where $\rho$ is given by (31). Let

$$\rho = \delta_{4k}\sqrt{\frac{2(1+\eta\delta_{4k}^2)}{1-\delta_{4k}^2}} \le 0.5. \tag{38}$$

It is not difficult to verify that the relation (38) is satisfied if

$$\delta_{4k} \le \sqrt{\frac{2}{\sqrt{81+16(\sqrt{5}+1)}+9}} \approx 0.3122,$$

which guarantees that

$$\|x^p - x_S\|_2 \leq 0.5^p \|x^0 - x_S\|_2 + \xi \|A^T \nu'\|_2, \tag{39}$$

where $\xi$ is a certain univariate constant. Thus our result for CoSaMP in terms of geometric rate 0.5 also improves the existing results in [24, 40].

**Remark 4.6** Finally, we make a few remarks before closing the paper.

(i) The general IHT adopts the iterative scheme

$$x^{p+1} = \mathcal{H}_k(x^p + \lambda_p A^T(y - Ax^p)),$$

where $\lambda_p$ is a steplength which is either fixed or updated from iteration to iteration (see [5, 24, 27]). In this paper, we only consider the case $\lambda_p \equiv 1$. Although our analysis provides a possible idea to enhance the existing results for IHT with other stepsizes, we are still not clear at the moment whether an improved result for IHT with $\lambda_p \neq 1$ can be obtained by our argument method. In fact, when a steplength is involved, several key estimations such as those in Lemma 3.5 cannot directly apply to the analysis of the algorithms for this case. To ensure the success of IHT for solving a sparse recovery problem, both some assumptions imposed on the measurement matrix and the steplength are needed. The existing results indicate that the analysis in this case becomes more complicated, and the choice of $\lambda_p$ usually relies on the RIC.

(ii) Several modifications of IHT were proposed in [26], and the HTP is one of those modifications. It was shown that $\delta_{3k} < 1/\sqrt{3}$ is a sufficient condition for guaranteed performance of the HTP with unit steplength, and this bound remains the best known one for this method in terms of the RIP of order $3k$. It seems that the analysis in this paper for IHT and CoSaMP does not help improving the result for HTP. It is also worth pointing out that although the structure of the efficient subspace pursuit (SP) [16] is quite similar to that of CoSaMP, it remains not clear at the moment whether the best known bound, $\delta_{4k} < 0.4858$, shown by Song, Xia and Liu [41] for SP can be improved via the analysis in this paper.

(iii) The analysis of the algorithm, in theory, can be performed independently in terms of any order $K$ of RIC provided that $K \geq k$, where $k$ is the sparsity level of the signal. Depending on the structure of the algorithm and the argument approach, however, some order of RIC is particularly suited for the analysis of a specific algorithm. For instance, $\delta_{3k}$ and $\delta_{4k}$ are convenient for the analysis of IHT and CoSaMP, respectively, and $\delta_{2k}$ is convenient for $\ell_1$-minimization. The derived results with different orders are often independent in the sense that one cannot imply the other. For instance, the IHT with steplength $\lambda_p = 3/4$ can guarantee to recover a $k$-sparse signal if $\delta_{2k} < 1/3$ (see [27] or p.173 in [24]). This condition does not imply the bound $\delta_{3k} < 0.618$ derived in this paper and via versa. However, some implications might still be made when comparing different recovery conditions. Taking $\ell_1$-minimization as an example, it is known that $\delta_{2k} < \sqrt{2}/2 \approx 0.707$ is a sufficient condition for $\ell_1$-minimization to recover the $k$-sparse signal [10]. Since $\delta_{2k} \leq \delta_{3k} \leq \delta_{4k}$, the sufficient conditions (in terms of $\delta_{3k}$ and $\delta_{4k}$) developed in this paper for IHT and CoSaMP must also imply that $\ell_1$-minimization guarantees to recover the $k$-sparse signal. Specifically, since the condition $\delta_{3k} < 0.618$ implies $\delta_{2k} < 0.618$, both IHT and $\ell_1$-minimization will succeed in recovering the $k$-sparse signal from a theoretical point of view. However, these algorithms may perform very differently from a numerical point of view. In practical applications, one must choose between algorithms by taking into account the computational complexity of an algorithm. It is well known that $\ell_1$-minimization, as a convex optimization problem, admits much higher computational complexity than IHT and CoSaMP which are much easier to implement and work faster than $\ell_1$-minimization, especially when solving a large scaled recovery problem.

(iv) The RIC is not the only tool for the analysis of compressed sensing algorithms. The earlier analysis was usually performed via the mutual coherence (see [21, 24] the references therein). The mutual coherence is easy to compute, but the derived results are often conservative. As a result, the RIP and other analysis tools, including the null space property of measurement matrices (e.g., [24]) and the range space property of transposed measurement matrices (e.g., [45, 46]), are also introduced to analyze the algorithms. A common goal of these analyses is to identify the largest class of recovery problems that are guaranteed to be solved by an algorithm. In terms of $\delta_K$, finding the tightest RIP-based bound $\delta_K < \delta^*$ for guaranteed success of an algorithm is equivalent to identifying the largest family of recovery problems for which the algorithm can successfully solve. The bound established in this paper for IHT seems tight, while it remains unclear whether the improved bound for CoSaMP is tight or not. However, any improved bound must be an important step moving towards the tightest one.

## 5 Conclusions

The deep property of hard thresholding operator in Lemma 3.2 makes it possible to improve the performance theory for the two compressed sensing algorithms IHT and CoSaMP. The new RIP-based bound $\delta_{3k} < (\sqrt{5}-1)/2$ shown for IHT in this paper remarkably improves the existing bounds for this algorithm. While the improvement of the performance result for CoSaMP is much more challenging, the existing RIP-based bound for this algorithm was eventually improved to $\delta_{4k} < 0.5102$ in this paper.

## References


[1] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, A simple proof of the restricted isometry property for random matrices, *Constr. Approx.*, 28 (2008), pp. 253–263.

[2] A. Beck and Y.C. Eldar, Sparsity constrained nonlinear optimization: Optimality conditions and algorithms, *SIAM J. Optim.*, 23 (2013), no.3, pp. 1480–1509.

[3] A. Bertsimas, R. King and R. Mazumder, Best subset selection via a modern optimization Lens, *Ann. Statist.*, 44 (2016), pp. 813–852.

[4] T. Blumensath and M. Davies, Iterative thresholding for sparse approximations. *J. Fourier Anal. Appl.* 14 (2008), no.5-6, pp. 629–654.

[5] T. Blumensath and M. Davies, Normalized iterative hard thresholding: guaranteed stability and performance. *IEEE J. Sel. Top. Signal Process.* 4 (2010), no.2, pp. 298–309.

[6] T. Blumensath and M. Davies, Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* 27 (2009), no.3, pp. 265–274.

[7] H. Boche, G. Caire, R. Calderbank, G. Kutyniok, R. Mathar, and P. Petersen, *Compressed Sensing and Its Applications*: Third International MATHEON Conference 2017, Birkhauser, 2019.

[8] J.-U. Bouchot, A generalized class of hard thresholding algorithms for sparse signal recovery. In: Fasshauer G., Schumaker L. (eds) Approximation Theory XIV: San Antonio 2013. Springer Proceedings in Mathematics & Statistics, 83 (2014), pp. 45–63.

[9] J.-U., Bouchot, S. Foucart and P. Hitczenki, Hard thresholding pursuit algorithms: Number of iterations, *Appl. Comput. Harmon. Anal.*, 41 (2016), no.2, pp. 412–435.

[10] T. Cai and A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, *IEEE Trans. Inform Theory*, 60 (2014), pp. 122–132.

[11] E.J. Candès, *Compresive Sampling*, Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.

[12] E.J. Candès and T. Tao, Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51(2005), no.12, pp. 4203–4215.

[13] E.J. Candès and T. Tao, Near optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inform. Theory* 52 (2006), no.12, pp.5406–5425.

[14] E.J. Candès and Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Trans. Inform. Theory*, 57 (2011), no.4, pp. 2342–2359.

[15] L.-H. Chang and J.-Y. Wu, An improved RIP-based performance guarantee for sparse signal reconstruction via subspace pursuit, IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM), 2014, pp. 405–408.

[16] W. Dai, and O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Trans. Inform. Theory*, 55 (2009), no.5, pp. 2230–2249.

[17] M.A. Davenport and J. Romberg, An overview of low-rank matrix recovery from incomplete observations, *IEEE J. Sel. Topics Signal Process.*, 10 (2016), no.4, pp. 608–622.

[18] D.L. Donoho and I. Johnstone, Idea spatial adaptation via wavelet shrinkage, *Biomatrika*, 81 (1994), no.3, pp. 425–455.

[19] D.L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory*, 52 (2006), no.4, pp. 1289–1306.

[20] M. Elad, Why simple shringkage is still relevant for redundant representation, *IEEE Trans. Inform. Theory*, 52 (2006), no.12, pp. 5559–5569.

[21] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, New York, 2010.

[22] Y.C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, Cambridge, UK, 2012.

[23] M. Fornasier and R. Rauhut, Iterative thresholding algorithms, *Appl. Comput. Harmon. Anal.*, 25 (2008), no.2, pp. 187–208.

[24] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer, NY, 2013.

[25] S. Foucart and S. Subramanian, Iterative hard thresholding for low-rank recovery from rank-one projections, *Linear Algebra Appl.,* 572 (2019), pp. 117–134.

[26] S. Foucart, Hard thresholding pursuit: an algorithm for compressive sensing. *SIAM J. Numer. Anal.* 49 (2011), no.6, pp. 2543–2563.

[27] R. Garg and R. Khandekar, Gradient descent with sparsification: An iterative algorithm for sparse recovery with restricted isometry property. In Proceedings of the 26th Annual International Conference on Machine Learning, ICML 09, pp. 337–344, ACM, New York, NY, USA, 2009.

[28] R. Giryes and M. Elad, RIP-based near-oracle performance guarantees for SP, CoSaMP and IHT, *IEEE Trans. Signal Process.*, 60 (2012), no. 3, pp. 1465–1568.

[29] K. Lee, Y. Bresler and M. Junge, Oblique pursuits for compressed sensing, *IEEE Trans. Inf. Theory*, 59 (2013), no.9, pp. 6111–6141.

[30] H. Liu and R.F. Barber, Between hard and soft thresholding: Optimal iterative thresholding algorithms, arXiv, July 2019.

[31] A. Majumdar, *Compressed Sensing for Magnetic Resonance Image Reconstruction*, Cambridge University Press, Cambridge, UK, 2015.

[32] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, Uniform uncertainty principle for Bernoulli and subgaussian ensembles. *Constr. Approx.*, 28 (2008), pp. 277–289.

[33] N. Meng and Y.-B. Zhao, Newton-Step-Based Hard Thresholding Algorithms for Sparse Signal Recovery, *IEEE Trans. Signal process.*, 68 (2020), pp. 6594–6606

[34] A. Miller, *Subset Selection in Regression*, CRC Press, Washington, 2002.

[35] M. Ndaoud and A.B. Tsybakov, Optimal variable selection and adaptive noisy compressed sensing, *IEEE Trans. Inform. Theory* 66 (2020), no. 4, 2517–2532.

[36] D. Needell, J. Tropp, CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.*, 26(2009), no.3, pp. 301–321.

[37] D. Needell, R. Vershynin, Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit. *Found. Comput. Math.*, 9(2009), no.3, pp.317–334

[38] D. Needell, R. Vershynin, Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE J. Sel. Top. Signal Process.*, 4 (2010), no.2, pp.310–316.

[39] N. Nguyen, D. Needell and T. Woolf, Linear convergence of stochastic iterative greedy algorithms with sparse constraints, *IEEE Trans. Inform. Theory*, 63 (2017), no. 11, pp. 6869–6895.

[40] J. Shen and P. Li, A tight bound of hard thresholding, *J. Machine Learning Res.*, 18 (2018), pp. 1–42.

[41] C.-B. Song, S.-T. Xia and X.-J. Liu, Improved analysis for subspace pursuit algorithm in terms of restricted isometry constant, *IEEE Signal Process. Lett.*, 21 (2014), no. 11, pp. 1365–1369,

[42] Y.-B. Zhao, *Sparse Optimization Theory and Methods*, CRC Press, Boca Raton, FL, 2018.

[43] Y.-B. Zhao, Optimal $k$-thresholding algirithms for sparse optimization problems, *SIAM J. Optim.*, 30 (2020), no. 1, pp. 31–55.

[44] Y.-B. Zhao and Z.-Q. Luo, Analysis of optimal thresholding algorithms for compressed sensing, *Signal Process.*, 187 (2021), 108148.

[45] Y.-B. Zhao, RSP-Based analysis for sparsest and least $\ell_1$-norm solutions to underdetermined linear systems, *IEEE Trans. Signal Process.*, 61 (2013), no.22, pp. 5777–5788.

[46] Y.-B. Zhao, H. Jiang and Z.-Q. Luo, Weak stability of $\ell_1$-minimization methods in sparse data reconstruction, *Math. Oper. Res.*, 44 (2019), no.1, pp.173–195.

[47] S. Zhou, N.H. Xiu and H.D. Qi, Global and quadratic convergence of Newton hard-thresholding pursuit, *J Mach Learn Res.*, 22 (2021), no. 12, 1–45.