

THE EXACT RECOVERY OF SPARSE SIGNALS VIA ORTHOGONAL MATCHING PURSUIT*

Anping Liao, Jiaxin Xie and Xiaobo Yang

College of Mathematics and Econometrics, Hunan University, Changsha 410082, China

Email: liaop@hnu.edu.cn, xiejiaxin@hnu.edu.cn, xiaoboyang@hnu.edu.cn

Peng Wang

Department of Mathematics, Wuyi University, Jiangmen 529020, China

Email: p_wong@126.com

Abstract

This paper aims to investigate sufficient conditions for the recovery of sparse signals via the orthogonal matching pursuit (OMP) algorithm. In the noiseless case, we present a novel sufficient condition for the exact recovery of all k -sparse signals by the OMP algorithm, and demonstrate that this condition is sharp. In the noisy case, a sufficient condition for recovering the support of k -sparse signal is also presented. Generally, the computation for the restricted isometry constant (RIC) in these sufficient conditions is typically difficult, therefore we provide a new condition which is not only computable but also sufficient for the exact recovery of all k -sparse signals.

Mathematics subject classification: 90C90, 94A12, 65J22, 15A29.

Key words: Compressed sensing, Sparse signal recovery, Restricted orthogonality constant (ROC), Restricted isometry constant (RIC), Orthogonal matching pursuit (OMP).

1. Introduction

Recovery of a sparse signal based on a small number of linear measurements is a fundamental problem in compressed sensing [10]. We consider the following model:

$$y = \Phi\beta + \epsilon, \quad (1.1)$$

where $y \in \mathbb{R}^m$ is an observation vector, $\Phi \in \mathbb{R}^{m \times n}$ is a known sensing matrix and $\epsilon \in \mathbb{R}^m$ is the measurement error vector. Suppose $\Phi = (\phi_1, \phi_2, \dots, \phi_n)$ where ϕ_i denotes the i th column of Φ . Throughout this paper we assume that the columns of Φ are normalized, i.e., $\|\phi_i\|_2 = 1$ for $i = 1, 2, \dots, n$. The goal of compressed sensing is to reconstruct the unknown $\beta \in \mathbb{R}^n$ based on y and Φ .

One of the most commonly used frameworks for the recovery of sparse signals is the Mutual Coherence Property introduced by Donoho and Huo in [11]. For a vector $\beta = (\beta(1), \dots, \beta(n)) \in \mathbb{R}^n$, the support of β is defined as $\text{supp}(\beta) = \{i : \beta(i) \neq 0\}$ and β is said to be k -sparse if $|\text{supp}(\beta)| \leq k$. The mutual coherence is defined by [11].

Definition 1.1. (Mutual Coherence [11]) *The mutual coherence μ of a matrix Φ is defined as*

$$\mu := \max_{i \neq j} |\langle \phi_i, \phi_j \rangle|. \quad (1.2)$$

* Received May 11, 2015 / Revised version received September 18, 2015 / Accepted October 21, 2015 /
Published online January 18, 2016 /

The validity of the OMP algorithm was investigated by Tropp [18] and Cai and Xu [4] in the framework of Mutual Coherence. In the noiseless case, Tropp [18] showed that $\mu < \frac{1}{2k-1}$ is a sufficient condition for the exact recovery of a k -sparse signal β , and Cai and Xu [4] showed that this condition is in fact sharp. When the linear measurement is corrupted by noise, Cai and Wang [2] considered two types of bounded noise. One is ℓ_2 bounded noise, i.e., $\|\epsilon\|_2 \leq \eta_1$, for some constant $\eta_1 > 0$. The other is ℓ_∞ bounded noise, i.e., $\|\Phi^T \epsilon\|_\infty \leq \eta_2$, for some constant $\eta_2 > 0$. In the ℓ_2 bounded noise case, if the conditions

$$\mu < \frac{1}{2k-1} \quad \text{and} \quad |\beta(i)| > \frac{2\eta_1}{1 - (2k-1)\mu} \quad (i \in \text{supp}(\beta))$$

are satisfied, then the support of the k -sparse signal β can be recovered exactly via OMP. In the ℓ_∞ bounded noise case, a similar result was given.

In the framework of restricted isometry property (RIP), the validity of the OMP algorithm was investigated by Mo et al. [14], Wu et al. [17] and Cheng et al. [8]. Their results were related to the restricted isometry constant (RIC), that is defined by [7].

Definition 1.2. Let $\Phi \in \mathbb{R}^{m \times n}$ be a matrix, and let $1 \leq k \leq n$ be an integer. The restricted isometry constant (RIC) of order k is defined as the smallest non-negative number δ_k^Φ such that for all k -sparse vectors $\beta \in \mathbb{R}^n$,

$$(1 - \delta_k^\Phi) \|\beta\|_2^2 \leq \|\Phi\beta\|_2^2 \leq (1 + \delta_k^\Phi) \|\beta\|_2^2.$$

In the noiseless case, Mo and Shen [14] showed that under the condition $\delta_{k+1}^\Phi < \frac{1}{1+\sqrt{k}}$, OMP can exactly recover the k -sparse signal. In the ℓ_2 bounded noisy case, Wu et al. [17] showed that the support of the k -sparse signal β can be recovered exactly via OMP under the conditions

$$\delta_{k+1}^\Phi < \frac{1}{1 + \sqrt{k}} \quad \text{and} \quad |\beta(i)| > \frac{(\sqrt{1 + \delta_{k+1}^\Phi} + 1) \eta_1}{1 - (\sqrt{k} + 1) \delta_{k+1}^\Phi} \quad (i \in \text{supp}(\beta)).$$

In the ℓ_∞ bounded noise case, a similar result was given.

In this paper, some sufficient conditions based on the restricted orthogonality constant (ROC) are given. The following definition can be seen, e.g., in [5, 15].

Definition 1.3. Let $\Phi \in \mathbb{R}^{m \times n}$ be a matrix, and let $1 \leq k_1, k_2 \leq n$ be two integers. The restricted orthogonality constant (ROC) of order (k_1, k_2) is defined as the smallest non-negative number θ_{k_1, k_2}^Φ such that

$$|\langle \Phi\beta_1, \Phi\beta_2 \rangle| \leq \theta_{k_1, k_2}^\Phi \|\beta_1\|_2 \|\beta_2\|_2,$$

for all k_1 -sparse vector β_1 and k_2 -sparse vector β_2 with disjoint supports. We set

$$\theta_{k_1, 0}^\Phi = \theta_{0, k_2}^\Phi = 0.$$

For a matrix Φ with normalized columns, the mutual coherence is a special case of the ROC, i.e., $\mu = \theta_{1,1}^\Phi$. Roughly speaking, the RIC δ_k^Φ and ROC θ_{k_1, k_2}^Φ measure how far subsets of cardinality k of columns of Φ are to an orthonormal system. It is obvious that δ_k^Φ and θ_{k_1, k_2}^Φ are increasing in each of their indices.

In this paper, we establish some more relaxed conditions for sparse signals recovery via OMP. We show that the condition

$$\delta_k^\Phi + \sqrt{k} \theta_{1, k}^\Phi < 1$$

guarantees the exact recovery of all k -sparse signals via OMP in the noiseless case. Furthermore, the upper bound 1 is sharp in the sense that for any $\xi > 0$, the condition $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1 + \xi$ is not sufficient to guarantee such exact recovery by using any recovery method. This is verified by constructing a specific counterexample. Note that $\theta_{1,k}^\Phi \leq \delta_{k+1}^\Phi$ and $\delta_k^\Phi \leq \delta_{k+1}^\Phi$, it is obvious that $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is weaker than $\delta_{k+1}^\Phi < \frac{1}{1+\sqrt{k}}$, thus some related results ([14], Theorem 3.1 etc.) are improved. Because the computation of δ_k^Φ is typically difficult [1] but $\theta_{1,k}^\Phi$ can be effectively computed, we give a computable condition

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1,$$

which can guarantee the exact recovery of all k -sparse signals via OMP in the noiseless case. Furthermore, the condition is proved to be sharp. In the ℓ_2 bounded noisy case, when the conditions

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1 \quad \text{and} \quad |\beta(i)| > \frac{2\eta_1}{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi - \sqrt{k}\theta_{1,k}^\Phi} \quad (i \in \text{supp}(\beta))$$

are satisfied, we show that the support of the k -sparse signal β can be recovered exactly via OMP. In the ℓ_∞ bounded noisy case, a similar result is given. Since $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$ is weaker than $\mu < \frac{1}{2k-1}$, some related results ([2], Theorem 1 etc.) are also improved.

The rest of the paper is organized as follows. In Section 2, we will introduce some notations and preliminary lemmas. In Section 3, we will give some details on OMP. In Section 4, we focus on sparse signals recovery via OMP.

2. Notations and Preliminaries

In this section, we introduce some basic notations and preliminary lemmas that will be used throughout the paper.

For two sets Γ and Λ in $\{1, \dots, n\}$, let $\Gamma \setminus \Lambda = \{i | i \in \Gamma, i \notin \Lambda\}$ and $\Gamma^c = \{1, 2, \dots, n\} \setminus \Gamma$. For any subset $\Lambda \subseteq \{1, 2, \dots, n\}$, denote by Φ_Λ a submatrix of Φ consisting of the columns ϕ_i with $i \in \Lambda$, β_Λ is a sub-vector of β consisting of the components $\beta(i)$ with $i \in \Lambda$ and

$$P_\Lambda = \Phi_\Lambda(\Phi_\Lambda^T \Phi_\Lambda)^{-1} \Phi_\Lambda^T$$

denotes the projection onto the linear space spanned by the elements of Φ_Λ . For a matrix Φ ,

$$\|\Phi\|_{1,1} = \max_{x \neq 0} \frac{\|\Phi x\|_1}{\|x\|_1}$$

is the induced norm, where $\|x\|_1 = \sum_{i=1}^n |x_i|$ for $x = (x(1), \dots, x(n))$. $e_i \in \mathbb{R}^n$ is a vector with 1 in the i th index and zeros elsewhere. We will call i a correct index if the corresponding $\beta(i) \neq 0$ and call i an incorrect index otherwise. $S = \{i : \beta(i) \neq 0\}$ denotes the support of a vector β .

Using the above notations, we state and prove some key technical tools which will be used in the proof of the main results.

Lemma 2.1. ([3]) *For any $\nu \geq 1$ and positive integers k_1, k_2 such that νk_2 is an integer, then*

$$\theta_{k_1, \nu k_2}^\Phi \leq \sqrt{\nu} \theta_{k_1, k_2}^\Phi. \quad (2.1)$$

Lemma 2.2. *If $\mu < \frac{1}{2k-1}$, then $\sqrt{k-1}\theta_{1,k-1}^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$.*

Proof. If $k=1$, then $\theta_{1,1}^\Phi = \mu < \frac{1}{2-1} = 1$. Consequently,

$$\sqrt{k-1}\theta_{1,k-1}^\Phi + \sqrt{k}\theta_{1,k}^\Phi = \theta_{1,1}^\Phi < 1.$$

For $k > 1$, it follows from Lemma 2.1 that $\theta_{1,k}^\Phi \leq \sqrt{k}\theta_{1,1}^\Phi = \sqrt{k}\mu$ and $\theta_{1,k-1}^\Phi \leq \sqrt{k-1}\mu$. Hence

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi \leq k\mu + (k-1)\mu = (2k-1)\mu < 1. \quad \square$$

The following proposition presents a method to compute $\theta_{1,k}^\Phi$ in polynomial time.

Proposition 2.1. *Let $\Phi \in \mathbb{R}^{m \times n}$ has normalized columns. Then*

$$\theta_{1,k}^\Phi = \max_{1 \leq i \leq n} \max_{\substack{\Lambda \subseteq \{1, \dots, n\} \setminus \{i\} \\ |\Lambda| \leq k}} \|\Phi_\Lambda^T \phi_i\|_2. \quad (2.2)$$

Proof. Without loss of generality, let $\alpha = e_i$ and β be a k -sparse vector with unit ℓ_2 norm, and with disjoint supports. Let $\Lambda = \text{supp}(\beta)$ and $C = \max_{1 \leq i \leq n} \max_{\substack{\Lambda \subseteq \{1, \dots, n\} \setminus \{i\} \\ |\Lambda| \leq k}} \|\Phi_\Lambda^T \phi_i\|_2$. Then

$$|\langle \Phi \alpha, \Phi \beta \rangle| = |\langle \phi_i, \Phi_\Lambda \beta \rangle| = |\langle \Phi_\Lambda^T \phi_i, \beta \rangle| \leq \|\Phi_\Lambda^T \phi_i\|_2 \leq C. \quad (2.3)$$

From (2.3) and the definition of $\theta_{1,k}^\Phi$, we have $\theta_{1,k}^\Phi \leq C$.

Conversely, let $\beta_\Lambda = \frac{\Phi_\Lambda^T \phi_i}{\|\Phi_\Lambda^T \phi_i\|_2}$ and $\beta_{\Lambda^c} = 0$. Then

$$|\langle \Phi \alpha, \Phi \beta \rangle| = \|\Phi_\Lambda^T \phi_i\|_2 \leq \theta_{1,k}^\Phi,$$

the inequality is due to the definition of $\theta_{1,k}^\Phi$, then it follows that $\theta_{1,k}^\Phi \geq C$. To sum up, (2.2) is proved. \square

Remark 2.1. For $\alpha = (\alpha(1), \dots, \alpha(n)) \in \mathbb{R}^n$ and $|\alpha(i_1)| \geq |\alpha(i_2)| \geq \dots \geq |\alpha(i_n)|$ with all i_k distinct, we define $\alpha_{\max(k)} = (\alpha(i_1), \alpha(i_2), \dots, \alpha(i_k))$. Then from (2.2), we have

$$\theta_{1,k}^\Phi = \max_{1 \leq i \leq n} \|\Psi_{\max(k)}^{(i)}\|_2,$$

where

$$\Psi^{(i)} = (\phi_i^T \phi_1, \dots, \phi_i^T \phi_{i-1}, 0, \phi_i^T \phi_{i+1}, \dots, \phi_i^T \phi_n)$$

denotes the i th row of $\Phi^T \Phi - I_n$. The main cost for calculating $\theta_{1,k}^\Phi$ is to compute the strict upper triangular part of the matrix $\Phi^T \Phi$. Firstly, we have to calculate $n(n-1)/2$ inner products $\phi_i^T \phi_j$ for $i, j \in \{1, \dots, n\}$ with $i \neq j$. Since each of the inner product needs $2n-1$ flops, it takes $O(n^3)$ flops to calculate $\theta_{1,k}^\Phi$. Thus $\theta_{1,k}^\Phi$ is indeed computable.

Let $S = \{i : \beta(i) \neq 0\}$ be the support of the k -sparse vector β . Define

$$M := \max_{\phi \in \Phi \setminus \Phi_S} \{ \|(\Phi_S^T \Phi_S)^{-1} \Phi_S^T \phi\|_1 \}. \quad (2.4)$$

The condition

$$M < 1$$

is called the Exact Recovery Condition (ERC) in Tropp [18]. It was shown that the ERC is a sufficient condition for OMP to exact recovery of the support of the signal in the noiseless case and it is easy to see that the value of M is not computable as it depends on the unknown support S .

Lemma 2.3. *Let $S = \text{supp}(\beta)$ with $|S| \leq k$. If $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$, then*

$$M \leq \frac{\sqrt{k}\theta_{1,k}^\Phi}{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi} < 1.$$

Proof. Without loss of generality, we suppose $|S| = k$. Apply the usual norm bound

$$M = \max_{\phi \in \Phi \setminus \Phi_S} \{ \|(\Phi_S^T \Phi_S)^{-1} \Phi_S^T \phi\|_1 \} \leq \|(\Phi_S^T \Phi_S)^{-1}\|_{1,1} \max_{\phi \in \Phi \setminus \Phi_S} \|\Phi_S^T \phi\|_1.$$

By Proposition 2.1 and noting that $\Phi_S^T \phi$ has no more than k nonzero elements, we have

$$\max_{\phi \in \Phi \setminus \Phi_S} \|\Phi_S^T \phi\|_1 \leq \max_{\phi \in \Phi \setminus \Phi_S} \sqrt{k} \|\Phi_S^T \phi\|_2 \leq \sqrt{k}\theta_{1,k}^\Phi. \quad (2.5)$$

Let $\Phi_S^T \Phi_S = I_k + \Psi$, each column of Ψ lists the inner products between one atom of Φ_S and the remaining $(k-1)$ atoms. By the property of $\|\cdot\|_{1,1}$,

$$\|\Psi\|_{1,1} = \max_{i \in S} \|\Phi_{S \setminus \{i\}}^T \phi_i\|_1 \leq \max_{i \in S} \sqrt{k-1} \|\Phi_{S \setminus \{i\}}^T \phi_i\|_2 \leq \sqrt{k-1}\theta_{1,k-1}^\Phi. \quad (2.6)$$

Since $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$, we can obtain $\|\Psi\|_{1,1} < 1$. Thus the Neumann series $\sum (-\Psi)^t$ converges to the inverse $(I_k + \Psi)^{-1}$. In this case, we have

$$\begin{aligned} \|(\Phi_S^T \Phi_S)^{-1}\|_{1,1} &= \|(I_k + \Psi)^{-1}\|_{1,1} = \left\| \sum_{t=0}^{\infty} (-\Psi)^t \right\|_{1,1} \\ &\leq \sum_{t=0}^{\infty} \|\Psi\|_{1,1}^t = \frac{1}{1 - \|\Psi\|_{1,1}} \leq \frac{1}{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi}. \end{aligned} \quad (2.7)$$

From (2.5), (2.7) and $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$, we have $M \leq \frac{\sqrt{k}\theta_{1,k}^\Phi}{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi} < 1$. \square

Remark 2.2. Suppose $S = \{s_1, s_2, \dots, s_k\}$. We should emphasize that the first equality in (2.6) only holds for the matrices with normalized columns and does not hold for general matrices. For example, let $\Phi_S = 2I_k$, then we have

$$\Psi = \Phi_S^T \Phi_S - I_k = 3I_k,$$

hence $\|\Psi\|_{1,1} = 3$. However, for any $s_\ell, s_j \in S$ with $s_\ell \neq s_j$, it holds that

$$\phi_{s_\ell}^T \phi_{s_j} = e_\ell^T e_j = 0,$$

so $\Phi_{S \setminus \{i\}}^T \phi_i = 0$ for any $i \in S$ which implies that the first equality in (2.6) does not hold. In fact, for general matrix Φ_S , let

$$\Lambda = \text{diag}(\phi_{s_1}^T \phi_{s_1}, \phi_{s_2}^T \phi_{s_2}, \dots, \phi_{s_k}^T \phi_{s_k}),$$

then the equality in (2.6) is satisfied with $\Psi = \Phi_S^T \Phi_S - \Lambda$.

Lemma 2.4. *Let $S = \text{supp}(\beta)$ with $|S| \leq k$. If $\sqrt{k-1}\theta_{1,k-1}^\Phi < 1$, then*

$$1 - \sqrt{k-1}\theta_{1,k-1}^\Phi \leq \lambda_{\min} \leq \lambda_{\max} \leq 1 + \sqrt{k-1}\theta_{1,k-1}^\Phi,$$

where λ_{\min} and λ_{\max} are the minimum and maximum eigenvalues of $\Phi_S^T \Phi_S$.

Proof. Let $\Phi_S^T \Phi_S = I_k + A$ and λ be any eigenvalue of $\Phi_S^T \Phi_S$. By the Gershgorin Circle Theorem, we have

$$|\lambda - 1| \leq \|A\|_{1,1}.$$

From (2.6), we can get $|\lambda - 1| \leq \sqrt{k-1} \theta_{1,k-1}^\Phi$. \square

Lemma 2.5. ([9]) *Suppose $|S| \leq k$, then the minimum and maximum eigenvalues of $\Phi_S^T \Phi_S$, which are denoted by λ_{\min} and λ_{\max} , satisfy*

$$1 - \delta_k^\Phi \leq \lambda_{\min} \leq \lambda_{\max} \leq 1 + \delta_k^\Phi.$$

In fact, we can prove that

$$\delta_k^\Phi = \max_{\substack{\Lambda \subseteq \{1, \dots, n\} \\ |\Lambda|=k}} \{1 - \lambda_{\min}(\Phi_\Lambda^T \Phi_\Lambda), \lambda_{\max}(\Phi_\Lambda^T \Phi_\Lambda) - 1\}.$$

Lemma 2.6. ([6]) *Suppose $\Phi \in \mathbb{R}^{m \times n}$, then*

$$\theta_{k_1, k_2}^\Phi \leq \delta_{k_1 + k_2}^\Phi.$$

Lemma 2.7. ([2]) *Suppose $\Lambda \subseteq S$, then the minimum eigenvalue of $\Phi_S^T \Phi_S$ is less than or equal to the minimum eigenvalue of $\Phi_{S \setminus \Lambda}^T (I - P_\Lambda) \Phi_{S \setminus \Lambda}$. The maximum eigenvalue of $\Phi_S^T \Phi_S$ is greater than or equal to the maximum eigenvalue of $\Phi_{S \setminus \Lambda}^T (I - P_\Lambda) \Phi_{S \setminus \Lambda}$.*

3. The OMP Algorithm

In this section we present a detailed description of the orthogonal matching pursuit (OMP) algorithm. We assume that the columns of Φ are normalized, i.e., $\|\phi_i\|_2 = 1$ for $i = 1, 2, \dots, n$.

The OMP algorithm can be stated as follows [12, 16, 19].

OMP algorithm.

Step 1. Input: Matrix $\Phi \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$ and the error level η ;

Step 2. Initialization: $r_0 := y$, $\beta_0 := 0$, $\Lambda_0 := \emptyset$, $i := 1$;

Step 3. Find the index λ_i such that

$$|\langle \phi_{\lambda_i}, r_{i-1} \rangle| = \arg \max_{\phi \in \Phi} |\langle \phi, r_{i-1} \rangle|$$

where ϕ is any column of Φ ;

Step 4. Update the support

$$\Lambda_i = \Lambda_{i-1} \cup \{\lambda_i\};$$

Step 5. Update the residual vector

$$r_i = y - \Phi \beta_i \quad \text{or} \quad r_i = (I - P_{\Lambda_i}) y$$

where $\beta_i = \arg \min_{\text{supp}(\beta) \subseteq \Lambda_i} \|y - \Phi \beta\|_2$ and $P_{\Lambda_i} = \Phi_{\Lambda_i} (\Phi_{\Lambda_i}^T \Phi_{\Lambda_i})^{-1} \Phi_{\Lambda_i}^T$;

Step 6. If the stopping rule is satisfied, stop the algorithm. Otherwise, set $i = i + 1$ and return to step 3.

By step 5, the residual vector r_i is orthogonal to the columns of Φ_{Λ_i} , i.e., $\Phi_{\Lambda_i}^T r_i = 0$. It is known, [14, 17], that the condition

$$\|\Phi_{S^c}^T r_{i-1}\|_\infty < \|\Phi_S^T r_{i-1}\|_\infty \quad (3.1)$$

can guarantee that OMP (step 3) selects a correct index, i.e., $\lambda_i \in S$, at i th iteration for $i = 1, 2, \dots, k$. In this paper, our discussions are based on (3.1), we shall establish an upper bound of the left hand of (3.1) as well as establish a lower bound of the right hand of (3.1), and then to get our sufficient conditions.

In this paper, we will consider two types of bounded noise. One is ℓ_2 bounded noise, i.e., $\|\epsilon\|_2 \leq \eta_1$ for some constant $\eta_1 > 0$. The other is ℓ_∞ bounded noise where $\|\Phi^T \epsilon\|_\infty \leq \eta_2$ for some constant $\eta_2 > 0$.

4. Recovery of Sparse Signals with OMP

We establish in this section the main results of this paper. In the noiseless case, a sharp condition and a computable condition for the exact recovery of the k -sparse signal via the OMP algorithm are given. In the noisy case, the support of the k -sparse signal can be recovered exactly under some assumptions.

4.1. ℓ_2 Bounded noise

We will consider the ℓ_2 bounded noise, i.e., $\|\epsilon\|_2 \leq \eta_1$ for some constant $\eta_1 > 0$. A sharp condition and a computable condition are established.

4.1.1. A sharp condition

We will show that the condition $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is sufficient for the exact recovery of all k -sparse signals in the noiseless case, and the condition is in fact sharp. In the noisy case, the following Theorem shows that the support of the k -sparse signal β can be recovered exactly by the OMP algorithm under the conditions

$$\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1 \quad \text{and} \quad |\beta(i)| > \frac{(1 + \sqrt{1 + \delta_k^\Phi})\eta_1}{1 - \delta_k^\Phi - \sqrt{k}\theta_{1,k}^\Phi} (i \in \text{supp}(\beta)).$$

Theorem 4.1. *Consider the signal recovery model (1.1) with $\|\epsilon\|_2 \leq \eta_1$. Let β be a k -sparse signal and $S = \text{supp}(\beta)$. If $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ and*

$$|\beta(j)| > \frac{(1 + \sqrt{1 + \delta_k^\Phi})\eta_1}{1 - \delta_k^\Phi - \sqrt{k}\theta_{1,k}^\Phi} (j \in S),$$

then the OMP algorithm with the stopping rule $\|r_i\|_2 \leq \eta_1$ will exactly recover the support S .

Before proving this Theorem, we will provide some technical analysis of the algorithm firstly. The analysis sheds light on how and when the OMP algorithm works properly. Let S be the support of a k -sparse signal β and suppose the algorithm selects the correct indexes in the first $i - 1$ iteration, i.e., $\Lambda_k \subset S$ for all $k \leq i - 1$. We will present a condition under which the OMP

algorithm selects the correct index at the i th iteration, i.e., $\Lambda_i \subseteq S$. The residual vector r_{i-1} can be written as

$$r_{i-1} = y - \Phi\beta_{i-1} = \Phi(\beta - \beta_{i-1}) + \epsilon.$$

In fact, at the i th iteration, we have

$$\begin{aligned} \|\Phi_{S^c}^T r_{i-1}\|_\infty &= \|\Phi_{S^c}^T \Phi(\beta - \beta_{i-1}) + \Phi_{S^c}^T \epsilon\|_\infty \\ &\leq \|\Phi_{S^c}^T \Phi(\beta - \beta_{i-1})\|_\infty + \|\Phi_{S^c}^T \epsilon\|_\infty \\ &= \max_{t \in S^c} |\langle \Phi(\beta - \beta_{i-1}), \Phi e_t \rangle| + \max_{t \in S^c} |\langle \phi_t, \epsilon \rangle| \\ &\leq \theta_{1,k}^\Phi \|\beta - \beta_{i-1}\|_2 + \eta_1. \end{aligned} \quad (4.1)$$

The second inequality is due to the definition of ROC and the Cauchy-Schwarz inequality.

It is easy to see that there exists a permutation matrix P such that

$$\Phi_S^T r_{i-1} = P \begin{pmatrix} \Phi_{S \setminus \Lambda_{i-1}}^T r_{i-1} \\ \Phi_{\Lambda_{i-1}}^T r_{i-1} \end{pmatrix} = P \begin{pmatrix} \Phi_{S \setminus \Lambda_{i-1}}^T r_{i-1} \\ 0 \end{pmatrix},$$

thereby $\Phi_S^T r_{i-1}$ is a $(k - (i - 1))$ -sparse vector. Without loss of generality, let $P = I$, then

$$\begin{aligned} \|\Phi_S^T r_{i-1}\|_\infty &\geq \frac{1}{\sqrt{k - (i - 1)}} \|\Phi_S^T r_{i-1}\|_2 \\ &= \frac{1}{\sqrt{k - (i - 1)}} \|\Phi_S^T \Phi_S(\beta - \beta_{i-1})_S + \Phi_S^T \epsilon\|_2 \\ &\geq \frac{1}{\sqrt{k - (i - 1)}} (\|\Phi_S^T \Phi_S(\beta - \beta_{i-1})_S\|_2 - \|\Phi_S^T \epsilon\|_2). \end{aligned} \quad (4.2)$$

From Lemma 2.5 we have

$$\|\Phi_S^T \Phi_S(\beta - \beta_{i-1})_S\|_2 \geq (1 - \delta_k^\Phi) \|\beta - \beta_{i-1}\|_2 \quad (4.3)$$

and

$$\|\Phi_S^T \epsilon\|_2 \leq \|\Phi_S\|_2 \|\epsilon\|_2 \leq \sqrt{1 + \delta_k^\Phi} \cdot \eta_1. \quad (4.4)$$

Substituting (4.3), (4.4) into (4.2), yields that

$$\|\Phi_S^T r_{i-1}\|_\infty \geq \frac{1}{\sqrt{k - (i - 1)}} \left((1 - \delta_k^\Phi) \|\beta - \beta_{i-1}\|_2 - \sqrt{1 + \delta_k^\Phi} \cdot \eta_1 \right). \quad (4.5)$$

If

$$\|\beta - \beta_{i-1}\|_2 > \sqrt{k - (i - 1)} \cdot \frac{(1 + \sqrt{1 + \delta_k^\Phi}) \eta_1}{1 - \delta_k^\Phi - \sqrt{k} \theta_{1,k}^\Phi}, \quad (4.6)$$

then from (4.1) and (4.5) we have

$$\|\Phi_{S^c}^T r_{i-1}\|_\infty < \|\Phi_S^T r_{i-1}\|_\infty.$$

Therefore, (4.6) is a sufficient condition for $\Lambda_i \subseteq S$. Now we are ready to prove Theorem 4.1.

Proof. We first consider the case where $|S| = k$. Our proof is constituted by three parts.

Part 1. In this part, we will prove that $\Lambda_i \subseteq S$ ($i = 1, \dots, k$) if $\Lambda_{i-1} \subseteq S$. From $\Lambda_{i-1} \subseteq S$, we have

$$\begin{aligned} \|\beta - \beta_{i-1}\|_2 &\geq \|\beta_{S \setminus \Lambda_{i-1}}\|_2 \geq \sqrt{k - (i-1)} \min_{j \in S} |\beta(j)| \\ &> \sqrt{k - (i-1)} \frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1}{1 - \delta_k^\Phi - \sqrt{k} \theta_{1,k}^\Phi}, \end{aligned}$$

the last inequality is due to

$$|\beta(j)| > \frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1}{1 - \delta_k^\Phi - \sqrt{k} \theta_{1,k}^\Phi} (j \in S).$$

The above inequality shows that (4.6) holds, therefore all the k correct indexes will be selected in the first k steps, i.e.,

$$\Lambda_i \subseteq S (i = 1, \dots, k).$$

Part 2. In this part, we will prove $\|r_k\|_2 \leq \eta_1$. Let P_{Λ_k} denote the projection onto the linear space spanned by Φ_S . Then

$$\|r_k\|_2 = \|(I - P_{\Lambda_k})\epsilon\|_2 \leq \|\epsilon\|_2 \leq \eta_1.$$

So when all the k correct indexes are selected, the ℓ_2 norm of the residual vector will be less than or equal to η_1 and hence the algorithm will stop.

Part 3. In this part, we will show that the OMP algorithm does not stop early, i.e.,

$$\|r_i\|_2 > \eta_1 (i = 1, \dots, k-1).$$

In fact, when $i < k$

$$\begin{aligned} \|r_i\|_2 &= \|\Phi(\beta - \beta_i) + \epsilon\|_2 \geq \|\Phi(\beta - \beta_i)\|_2 - \|\epsilon\|_2 \geq \sqrt{1 - \delta_k^\Phi} \|\beta - \beta_i\|_2 - \eta_1 \\ &> \sqrt{1 - \delta_k^\Phi} \frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1}{1 - \delta_k^\Phi - \sqrt{k} \theta_{1,k}^\Phi} - \eta_1 \geq 2\eta_1 - \eta_1 = \eta_1. \end{aligned}$$

Combining above three parts and noting that $\Lambda_0 \subseteq S$, we have proved that Theorem 4.1 holds for $|S| = k$.

Now we consider the case where signal β has less than k nonzero elements, i.e., $|S| < k$. By the established result above, if

$$|\beta(j)| > \frac{\left(1 + \sqrt{1 + \delta_{|S|}^\Phi}\right) \eta_1}{1 - \delta_{|S|}^\Phi - \sqrt{|S|} \theta_{1,|S|}^\Phi} (j \in S), \quad (4.7)$$

then OMP will exactly recover the support S in $|S|$ iterations, and the algorithm will stop, i.e., $\|r_{|S|}\|_2 \leq \eta_1$.

Moreover, noting that $\delta_{|S|}^\Phi \leq \delta_k^\Phi$ and $\theta_{1,|S|}^\Phi \leq \theta_{1,k}^\Phi$, it is easy to verify that (4.7) can be derived from the condition of Theorem 4.1, i.e.,

$$|\beta(j)| > \frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1}{1 - \delta_k^\Phi - \sqrt{k} \theta_{1,k}^\Phi} (j \in S).$$

This complete the proof. \square

From Theorem 4.1, we have the following corollary which shows that the condition

$$\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$$

is sufficient for the exact recovery of k -sparse signals in the noiseless case.

Corollary 4.1. *Consider the signal recovery model (1.1) with $\|\epsilon\|_2 = 0$. If $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$, then all k -sparse signals can be recovered exactly in k -iterations by the OMP algorithm.*

Remark 4.1. From Corollary 4.1, we know that $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is sufficient to recover exactly all k -sparse signals β in k -iterations by the OMP algorithm. Noting that $\theta_{1,k}^\Phi \leq \delta_{k+1}^\Phi$ (Lemma 2.6) and $\delta_k^\Phi \leq \delta_{k+1}^\Phi$, it is obvious $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is weaker than $\delta_{k+1}^\Phi < \frac{1}{\sqrt{k+1}}$. Thus Theorem 3.1 in [14] is improved.

The following theorem shows that the upper bound 1 is sharp for the exact recovery in the noiseless case.

Theorem 4.2. *Let k be a positive integer. Then there exists a sensing matrix Φ with normalized columns such that $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi = 1$, and for some k -sparse signals $u, v \in \mathbb{R}^n$ with $u \neq v$ satisfying $\Phi u = \Phi v$. In other words, there does not exist any algorithm that can exactly recover all k -sparse signals β based on (Φ, y) with $y = \Phi\beta$. In particular, the OMP algorithm cannot recover all k -sparse signals in the noiseless case.*

Proof. Let $e = (1, 1, \dots, 1)^T \in \mathbb{R}^{2k}$ and

$$\Phi = \sqrt{\frac{2k}{2k-1}} \cdot \begin{pmatrix} \phi^{(1)} \\ \vdots \\ \phi^{(2k-1)} \end{pmatrix} \in \mathbb{R}^{(2k-1) \times 2k}, \quad (4.8)$$

where

$$\phi^{(i)} = \left(\overbrace{-\frac{\sqrt{i^2+i}}{i^2+i}, \dots, -\frac{\sqrt{i^2+i}}{i^2+i}}^i, \frac{\sqrt{i^2+i}}{i+1}, 0, \dots, 0 \right) \in \mathbb{R}^{2k}, \quad i = 1, 2, \dots, 2k-1.$$

are row vectors. It is easy to verify that each column of Φ is of length 1 in ℓ_2 norm and $\Phi e = 0$. By simple calculation, we get

$$\Phi^T \Phi = \begin{pmatrix} 1 & -\frac{1}{2k-1} & \cdots & -\frac{1}{2k-1} \\ -\frac{1}{2k-1} & 1 & \cdots & -\frac{1}{2k-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{2k-1} & -\frac{1}{2k-1} & \cdots & 1 \end{pmatrix}_{2k \times 2k},$$

then for any $\Lambda \subseteq \{1, \dots, 2k\}$ with $|\Lambda| = k$, the eigenvalues $\{\lambda_i\}_{i=1}^k$ of $\Phi_\Lambda^T \Phi_\Lambda$ are

$$\lambda_1 = \cdots = \lambda_{k-1} = 1 + \frac{1}{2k-1} \quad \text{and} \quad \lambda_k = 1 - \frac{k-1}{2k-1}.$$

Therefore,

$$\begin{aligned}\delta_k^\Phi &= \max\{1 - \lambda_{\min}(\Phi_\Lambda^T \Phi_\Lambda), \lambda_{\max}(\Phi_\Lambda^T \Phi_\Lambda) - 1\} \\ &= \max\{1 - \lambda_k, \lambda_{k-1} - 1\} = \frac{k-1}{2k-1}.\end{aligned}\quad (4.9)$$

It follows from Remark 2.1, we have

$$\sqrt{k}\theta_{1,k}^\Phi = \sqrt{k} \left[\underbrace{\left(\frac{-1}{2k-1}\right)^2 + \dots + \left(\frac{-1}{2k-1}\right)^2}_k \right]^{\frac{1}{2}} = \frac{k}{2k-1}.\quad (4.10)$$

Thereby, from (4.9) and (4.10) we have

$$\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi = \frac{k-1}{2k-1} + \frac{k}{2k-1} = 1.$$

Finally, suppose

$$u = (\overbrace{1, \dots, 1}^k, \overbrace{0, \dots, 0}^k)^T, \quad v = (\overbrace{0, \dots, 0}^k, \overbrace{-1, \dots, -1}^k)^T.$$

Since $\Phi(u - v) = 0$, then u, v are both k -sparse such that $\Phi u = \Phi v$. Therefore, it is impossible to recover both u and v only base on $(\Phi u, \Phi)$, which finishes the proof. \square

One of the most commonly used approaches for sparse signal recovery is the ℓ_1 -minimization, i.e., $\min_\beta \{\|\beta\|_1 : A\beta = b\}$ (we consider the noiseless case) and the sufficient conditions in terms of the RIC or ROC for the exact recovery of k -sparse signals had been extensively studied [3, 5, 22]. As was shown in [5, 22], $\delta_k^\Phi + \theta_{k,k}^\Phi < 1$ is a sharp condition for the exact recovery via ℓ_1 -minimization. From Lemma 2.1, we have $\theta_{k,k}^\Phi \leq \sqrt{k}\theta_{1,k}^\Phi$, so if $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ then we can get $\delta_k^\Phi + \theta_{k,k}^\Phi < 1$, thus $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is a sufficient condition for ℓ_1 -minimization to exactly recover k -sparse signal. Moreover, from Theorem 4.2 we know that $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1$ is also a sharp condition for ℓ_1 -minimization. For the OMP algorithm, by simple calculation we know that the matrix Φ constructed as (4.8) satisfies $\delta_k^\Phi + \theta_{k,k}^\Phi = 1$, so $\delta_k^\Phi + \theta_{k,k}^\Phi \geq 1$ is not a sufficient condition for the exact recovery via OMP. What's more, other approaches for sparse signal recovery had been considered as well, e.g., Wen et al. [20] considered the ℓ_p -minimization and Xu [23] studied the orthogonal multi-matching pursuit (OMMP) algorithm.

Recently, Mo [13] and Wen [21] have considered the RIC for the OMP algorithm under the condition that the columns of the sensing matrix Φ are not necessary to being normalized, and a corresponding sharp condition has been established.

Remark 4.2. Theorem 4.2 is different from the Theorem 2.8 given in Cai and Zhang [5], since our analysis is based on the matrix Φ with normalized columns and their counterexample cannot be used to prove Theorem 4.2.

Remark 4.3. Theorem 4.2 shows that the upper bound 1 is sharp in the sense that for any $\xi > 0$, the condition $\delta_k^\Phi + \sqrt{k}\theta_{1,k}^\Phi < 1 + \xi$ fails to guarantee the exact recovery of all k -sparse signals.

4.1.2. A computable condition

It is shown in [1] that the computation of δ_k^Φ can be difficult, while as was already mentioned in Remark 2.1, $\theta_{1,k}^\Phi$ can be efficiently computed, in this section we will give a computable condition

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1.$$

Theorem 4.3. *Consider the signal recovery model (1.1) with $\|\epsilon\|_2 \leq \eta_1$. Let β be a k -sparse signal and $S = \text{supp}(\beta)$. If $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$ and*

$$|\beta(i)| > \frac{2\eta_1}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi} \quad (i \in S),$$

then the OMP algorithm with the stopping rule $\|r_i\|_2 \leq \eta_1$ will exactly recover the support S .

Similar to the above discussions, we first provide some technical analysis of the algorithm. The analysis sheds light on how and when the OMP algorithm works properly. Let S be the support of a k -sparse signal β and suppose the algorithm selects the correct indexes at the first $i-1$ iteration, i.e., $\Lambda_{i-1} \subseteq S$. From the OMP algorithm, we know that

$$P_{\Lambda_{i-1}} = \Phi_{\Lambda_{i-1}} (\Phi_{\Lambda_{i-1}}^T \Phi_{\Lambda_{i-1}})^{-1} \Phi_{\Lambda_{i-1}}^T,$$

then

$$\|I - P_{\Lambda_{i-1}}\|_2 \leq 1, \quad \text{Phi}_{\Lambda_{i-1}}^T (I - P_{\Lambda_{i-1}}) = 0, \quad (I - P_{\Lambda_{i-1}}) \Phi_{\Lambda_{i-1}} = 0. \quad (4.11)$$

Since the residual vector r_{i-1} can be written as $r_{i-1} = (I - P_{\Lambda_{i-1}})y = (I - P_{\Lambda_{i-1}})\Phi_S\beta_S + (I - P_{\Lambda_{i-1}})\epsilon$ and $\|\epsilon\|_2 \leq \eta_1$, we have

$$\begin{aligned} \|\Phi_S^T r_{i-1}\|_\infty &= \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S + \Phi_S^T (I - P_{\Lambda_{i-1}}) \epsilon\|_\infty \\ &\geq \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty - \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \epsilon\|_\infty \\ &= \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty - \max_{t \in S} |\langle \phi_t, (I - P_{i-1}) \epsilon \rangle| \\ &\geq \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty - \eta_1, \end{aligned} \quad (4.12)$$

and

$$\begin{aligned} \|\Phi_{S^c}^T r_{i-1}\|_\infty &= \|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S + \Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \epsilon\|_\infty \\ &\leq \|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty + \|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \epsilon\|_\infty \\ &\leq \|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty + \eta_1. \end{aligned} \quad (4.13)$$

In order to present a computable condition, the following lemma is necessary.

Lemma 4.1. ([2]) *Let M be defined as in (2.4). Then*

$$\|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty \leq M \cdot \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty.$$

Suppose $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$, from Lemma 2.3 and Lemma 4.1, we have

$$\|\Phi_{S^c}^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty \leq \frac{\sqrt{k}\theta_{1,k}^\Phi}{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi} \cdot \|\Phi_S^T (I - P_{\Lambda_{i-1}}) \Phi_S \beta_S\|_\infty. \quad (4.14)$$

From (4.12), (4.13) and (4.14), we know that the condition

$$\|\Phi_S^T(I - P_{\Lambda_{i-1}})\Phi_S\beta_S\|_\infty > \frac{1 - \sqrt{k-1}\theta_{1,k-1}^\Phi}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi} \cdot 2\eta_1 \quad (4.15)$$

implies (3.1). In other words, the previous discussion shows that (4.15) is a sufficient condition under which OMP will make a correct decision. It follows from (4.11) that

$$\begin{aligned} \|\Phi_S^T(I - P_{\Lambda_{i-1}})\Phi_S\beta_S\|_\infty &= \|\Phi_{S \setminus \Lambda_{i-1}}^T(I - P_{\Lambda_{i-1}})\Phi_S\beta_S\|_\infty \\ &= \|\Phi_{S \setminus \Lambda_{i-1}}^T(I - P_{\Lambda_{i-1}})\Phi_{S \setminus \Lambda_{i-1}}\beta_{S \setminus \Lambda_{i-1}}\|_\infty \\ &\geq \frac{1}{\sqrt{k-(i-1)}} \|\Phi_{S \setminus \Lambda_{i-1}}^T(I - P_{\Lambda_{i-1}})\Phi_{S \setminus \Lambda_{i-1}}\beta_{S \setminus \Lambda_{i-1}}\|_2 \\ &\geq \frac{1}{\sqrt{k-(i-1)}} \lambda_{\min} \|\beta_{S \setminus \Lambda_{i-1}}\|_2 \\ &\geq \frac{1}{\sqrt{k-(i-1)}} (1 - \sqrt{k-1}\theta_{1,k-1}^\Phi) \|\beta_{S \setminus \Lambda_{i-1}}\|_2, \end{aligned} \quad (4.16)$$

the last two inequality are due to Lemma 2.7 and Lemma 2.4.

Suppose

$$\|\beta_{S \setminus \Lambda_{i-1}}\|_2 > \sqrt{k-(i-1)} \cdot \frac{2\eta_1}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi}, \quad (4.17)$$

then (4.15) can be obtained from (4.16), thereby (3.1) holds. Therefore, (4.17) is a sufficient condition for OMP to select a correct index at the current step. Based on the above discussions, now we will prove Theorem 4.3.

Proof of Theorem 4.3. Similar to the proof of Theorem 4.1, we only consider the case where signal β just has k -nonzero elements, i.e., $|S| = k$. For $|S| < k$, the arguments in the proof of Theorem 4.1 also works.

One can verify that $|\beta(i)| > \frac{2\eta_1}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi}$ ensures that (4.17) holds, thereby the OMP algorithm selects all the k correct indexes in the first k steps, and after all the k correct indexes are selected the OMP algorithm will stop, i.e., $\|r_k\|_2 \leq \eta_1$. It remains to show that OMP does not stop early.

$$\begin{aligned} \|r_i\|_2 &= \|(I - P_{\Lambda_i})\Phi_S\beta_S + (I - P_{\Lambda_i})\epsilon\|_2 \\ &\geq \|(I - P_{\Lambda_i})\Phi_S\beta_S\|_2 - \|(I - P_{\Lambda_i})\epsilon\|_2 \\ &\geq \|(I - P_{\Lambda_i})\Phi_{S \setminus \Lambda_i}\beta_{S \setminus \Lambda_i}\|_2 - \eta_1. \end{aligned}$$

It follows from Lemma 2.7 and Lemma 2.4 that

$$\begin{aligned} \|(I - P_{\Lambda_i})\Phi_{S \setminus \Lambda_i}\beta_{S \setminus \Lambda_i}\|_2 &\geq \lambda_{\min}^{\frac{1}{2}}(\Phi_S^T\Phi_S)\|\beta_{S \setminus \Lambda_i}\|_2 \\ &\geq (1 - \sqrt{k-1}\theta_{1,k-1}^\Phi)^{\frac{1}{2}}\|\beta_{S \setminus \Lambda_i}\|_2 \\ &> 2\eta_1. \end{aligned}$$

So $\|r_i\|_2 \geq \|(I - P_{\Lambda_i})\Phi_{S \setminus \Lambda_i}\beta_{S \setminus \Lambda_i}\|_2 - \eta_1 > \eta_1$. Thus we complete the proof. \square

Remark 4.4. The nonzero coefficients $\beta(i)$ are required to satisfy

$$|\beta(i)| > \frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right)\eta_1}{1 - \delta_k^\Phi - \sqrt{k}\theta_{1,k}^\Phi}$$

in Theorem 4.1 and the nonzero coefficients $\beta(i)$ are required to satisfy

$$|\beta(i)| > \frac{2\eta_1}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi}$$

in Theorem 4.3. Since

$$\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1 \geq 2\eta_1 \quad \text{and} \quad 1 - \delta_k^\Phi - \sqrt{k}\theta_{1,k}^\Phi \geq 1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi,$$

the size of

$$\frac{\left(1 + \sqrt{1 + \delta_k^\Phi}\right) \eta_1}{1 - \delta_k^\Phi - \sqrt{k}\theta_{1,k}^\Phi} \quad \text{and} \quad \frac{2\eta_1}{1 - \sqrt{k}\theta_{1,k}^\Phi - \sqrt{k-1}\theta_{1,k-1}^\Phi}$$

cannot immediately be told to be one larger than the other, or vice-versa. Therefore, Theorem 4.1 and Theorem 4.3 are independent of each other.

From Theorem 4.3, we have the following corollary.

Corollary 4.2. *Consider the signal recovery model (1.1) with $\|\epsilon\|_2 = 0$. If*

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1,$$

then all k -sparse signals can be recovered exactly in k -iterations by the OMP algorithm.

From Lemma 2.2, we know that $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1$ is weaker than $\mu < \frac{1}{2k-1}$, so some related results ([2], Theorem 1 etc.) are also improved. In fact, the upper bound 1 is sharp for the exact recovery in the noiseless case.

Remark 4.5. The upper bound 1 is in fact sharp in the sense that for any $\xi > 0$,

$$\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi < 1 + \xi$$

does not ensure such exact recovery. The specific counterexample which was constructed in the proof of Theorem 4.2 is useful as well, since

$$\sqrt{k-1}\theta_{1,k-1}^\Phi = \frac{k-1}{2k-1}$$

and $\sqrt{k}\theta_{1,k}^\Phi = \frac{k}{2k-1}$ lead to $\sqrt{k}\theta_{1,k}^\Phi + \sqrt{k-1}\theta_{1,k-1}^\Phi = 1$.

4.2. ℓ_∞ Bounded noise

We now turn to the case where the noise ϵ is assumed to satisfy $\|\Phi^T \epsilon\|_\infty \leq \eta_2$ for some constant $\eta_2 > 0$. Let S be the support of k -sparse signal β and suppose the algorithm selects the correct indexes at the first $i-1$ iteration, i.e., $\Lambda_{i-1} \subseteq S$. We will provide some technical analysis of the algorithm. The analysis sheds light on how and when the OMP algorithm works properly.

Similar to the analysis of ℓ_2 bounded noise, (4.1) is satisfied (where η_1 is replaced by η_2).

Since

$$\Phi_S^T r_{i-1} = \begin{pmatrix} \Phi_{S \setminus \Lambda_{i-1}}^T r_{i-1} \\ \Phi_{\Lambda_{i-1}}^T r_{i-1} \end{pmatrix} = \begin{pmatrix} \Phi_{S \setminus \Lambda_{i-1}}^T r_{i-1} \\ 0 \end{pmatrix}$$

is a $(k - (i - 1))$ -sparse vector, therefore

$$\begin{aligned}
\|\Phi_S^T r_{i-1}\|_\infty &\geq \frac{1}{\sqrt{k - (i - 1)}} \|\Phi_S^T r_{i-1}\|_2 \\
&= \frac{1}{\sqrt{k - (i - 1)}} \|\Phi_S^T \Phi_S (\beta - \beta_{i-1})_S + \Phi_S^T \epsilon\|_2 \\
&\geq \frac{1}{\sqrt{k - (i - 1)}} (\|\Phi_S^T \Phi_S (\beta - \beta_{i-1})_S\|_2 - \|\Phi_S^T \epsilon\|_2) \\
&\geq \frac{1}{\sqrt{k - (i - 1)}} [\lambda_{\min}(\Phi_S^T \Phi_S) \|\beta - \beta_{i-1}\|_2 - \sqrt{k} \|\Phi_S^T \epsilon\|_\infty] \\
&\geq \frac{1}{\sqrt{k - (i - 1)}} (1 - \sqrt{k - 1} \theta_{1,k-1}^\Phi) \|\beta - \beta_{i-1}\|_2 - \frac{\sqrt{k} \eta_2}{\sqrt{k - (i - 1)}}, \tag{4.18}
\end{aligned}$$

the last inequality is due to Lemma 2.4.

Suppose

$$\|\beta - \beta_{i-1}\|_2 > \sqrt{k - (i - 1)} \cdot \frac{(1 + \sqrt{k}) \eta_2}{1 - \sqrt{k - 1} \theta_{1,k-1}^\Phi - \sqrt{k} \theta_{1,k}^\Phi}, \tag{4.19}$$

from (4.1) and (4.18), we know that (3.1) holds. Therefore, (4.19) is a sufficient condition for OMP selects a correct index at the current step.

Theorem 4.4. *Consider the signal recovery model (1.1) with $\|\Phi^T \epsilon\|_\infty \leq \eta_2$. Let β be a k -sparse signal and $S = \text{supp}(\beta)$. If $\sqrt{k} \theta_{1,k}^\Phi + \sqrt{k - 1} \theta_{1,k-1}^\Phi < 1$ and*

$$|\beta(i)| > \frac{2\sqrt{k} \eta_2}{1 - \sqrt{k} \theta_{1,k}^\Phi - \sqrt{k - 1} \theta_{1,k-1}^\Phi} (i \in S),$$

then the OMP algorithm with the stopping rule $\|\Phi^T r_i\|_\infty \leq \eta_2$ will return the support S .

Proof. We only consider the case where signal β just has k -nonzero elements, i.e., $|S| = k$. For the case where $|S| < k$, the argument in the proof of Theorem 4.1 also works. One can verify that

$$|\beta(i)| > \frac{2\sqrt{k} \eta_2}{1 - \sqrt{k} \theta_{1,k}^\Phi - \sqrt{k - 1} \theta_{1,k-1}^\Phi} (i \in S)$$

ensures that (4.19) holds, thereby the OMP algorithm selects all the k correct indexes in the first k steps.

We now turn to the stopping rule. It remains to prove that the OMP algorithm does not stop early, i.e., $\|\Phi^T r_i\|_\infty > \eta_2$ for $i < k$. In fact

$$\begin{aligned}
\|\Phi^T r_i\|_\infty &= \|\Phi^T \Phi (\beta - \beta_i) + \Phi^T \epsilon\|_\infty \\
&\geq \|\Phi^T \Phi_S (\beta - \beta_i)_S\|_\infty - \|\Phi^T \epsilon\|_\infty \\
&\geq \|\Phi_S^T \Phi_S (\beta - \beta_i)_S\|_\infty - \eta_2. \tag{4.20}
\end{aligned}$$

Since $\Phi_S^T \Phi_S (\beta - \beta_i)_S \in \mathbb{R}^{|S|}$ and $|S| = k$, we have

$$\begin{aligned}
\|\Phi_S^T \Phi_S (\beta - \beta_i)_S\|_\infty &\geq \frac{1}{\sqrt{k}} \|\Phi_S^T \Phi_S (\beta - \beta_i)_S\|_2 \\
&\geq \frac{1 - \sqrt{k - 1} \theta_{1,k-1}^\Phi}{\sqrt{k}} \|\beta - \beta_i\|_2 > 2\eta_2, \tag{4.21}
\end{aligned}$$

the second inequality is due to Lemma 2.4. So

$$\|\Phi^T r_i\|_\infty \geq \|\Phi_S^T \Phi_S(\beta - \beta_i)_S\|_\infty - \eta_2 > 2\eta_2 - \eta_2 = \eta_2.$$

Thus we complete the proof. \square

Remark 4.6. Let \bar{S} be the support set that obtained by the OMP algorithm with certain stopping rules being satisfied. Theorem 4.3 presents that under certain conditions the OMP can exactly recover the support S in the ℓ_2 bounded noise case, i.e., $\bar{S} = S$. In the ℓ_∞ bounded noise case, Theorem 4.4 only shows that the OMP can return the support under some conditions, i.e., $S \subseteq \bar{S}$.

Acknowledgments. The authors are very grateful to the anonymous referees for their valuable comments and suggestions. We want to thank Mr. Liang Chen at Hunan University for many useful comments. This work was supported by the National Natural Science Foundation of China under Grant 11271117.

References

- [1] A.S. Bandeira, E. Dobriban, D.G. Mixon and W.F. Sawin, Certifying the restricted isometry property is hard, *IEEE Trans. Inf. Theory*, **59** (2013), 3448–3450.
- [2] T. Cai and L. Wang, Orthogonal matching pursuit for sparse signal recovery with noise, *IEEE Trans. Inf. Theory*, **57** (2011), 4680–4688.
- [3] T. Cai, L. Wang, and G. Xu, Shifting inequality and recovery of sparse signals, *IEEE Trans. Signal Process.*, **58** (2010), 1300–1308.
- [4] T. Cai, L. Wang, and G. Xu, Stable recovery of sparse signals and an oracle inequality, *IEEE Trans. Inf. Theory*, **56** (2010), 3516–3522.
- [5] T. Cai and A. Zhang, Compressed sensing and affine rank minimization under restricted isometry, *IEEE Trans. Signal Process.*, **62** (2013), 3279–3290.
- [6] E. Candès and Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements, *IEEE Trans. Inf. Theory*, **57** (2009), 2342–2359.
- [7] E. Candès and T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory*, **51** (2005), 4203–4215.
- [8] L. Chang and J. Wu, An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit, *IEEE Trans. Inf. Theory*, **60** (2014), 5702–5715.
- [9] W. Dai and O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Trans. Inf. Theory*, **55** (2009), 2230–2249.
- [10] D. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, **52** (2006), 1289–1306.
- [11] D. Donoho and X. Huo, Uncertainty principles and ideal atomic decomposition, *IEEE Trans. Inf. Theory*, **47** (2001), 2845–2862.
- [12] E.D. Livshitz and V.N. Temlyakov, Sparse approximation and recovery by greedy algorithms, *IEEE Trans. Inf. Theory*, **60** (2014), 3989–4000.
- [13] Q. Mo, A sharp restricted isometry constant bound of orthogonal matching pursuit, *arXiv:1501.0178* (2015).
- [14] Q. Mo and Y. Shen, A remark on the restricted isometry property in orthogonal matching pursuit, *IEEE Trans. Inf. Theory*, **58** (2012), 3654–3656.
- [15] K. Mohan and M. Fazel, New restricted isometry results for noisy low-rank recovery, In *Proc. Intl. Symp. Inf. Theory (ISIT)*, (2010), 1573–1577.

- [16] Y.C. Pati, R. Rezaifar, and P.S. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, In *1993 Conf. Rec. Twenty-Seventh Asilomar Conf. on Signals, Systems and Computers*, **1** (1993), 40–44.
- [17] W. Rui, W. Huang, and D. Chen, The exact support recovery of sparse signals with noise via orthogonal matching pursuit, *IEEE Signal Process. Letters*, **20** (2013), 403–406.
- [18] J. Tropp, Greed is good: algorithmic results for sparse approximation, *IEEE Trans. Inf. Theory*, **50** (2004), 2231–2242.
- [19] J. Tropp and A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inf. Theory*, **53** (2007), 4655–4666.
- [20] J. Wen, D. Li and F. Zhu, Stable recovery of sparse signals via ℓ_p -minimization, *Appl. Comput. Harmon. Anal.*, **38** (2015), 161–176.
- [21] J. Wen, Z. Xiao and D. Li, Improved bounds on restricted isometry constant for orthogonal matching pursuit, *Electronics Letters*, **49** (2013), 1487–1489.
- [22] G. Xu and Z. Xu, On the ℓ_1 -norm invariant convex k -sparse decomposition of signals, *J. Oper. Res. Soc. China*, **1** (2013), 537–541.
- [23] Z. Xu, The performance of orthogonal multi-matching pursuit under the restricted isometry property, *J. Comp. Math.*, **33** (2015), 495–516.