Improved Sufficient Conditions for Sparse Recovery with Generalized Orthogonal Matching Pursuit

Jinming Wen, Zhengchun Zhou, Dongfang Li and Xiaohu Tang

Abstract

The generalized orthogonal matching pursuit (gOMP), also called the orthogonal multi-matching pursuit (OMMP), is a natural generation of the well-known orthogonal matching pursuit (OMP) in the sense that multiple \( N \geq 1 \) indices are identified per iteration. Sufficient conditions for sparse recovery with OMP and gOMP under restricted isometry property of a sensing matrix have received much attention in recent years. In this paper, we show that if the restricted isometry constant \( \delta_{NK+1} < 1/\sqrt{K/N + 1} \) of the sensing matrix \( A \) satisfies \( \delta_{NK+1} < 1/\sqrt{K/N + 1} \), then under some conditions on the signal-to-noise ratio (SNR), gOMP recovers at least one index in the support of the \( K \)-sparse signal \( x \) from an observation vector \( y = Ax + v \) (where \( v \) is a noise vector) in each iteration. Surprisingly, this condition do not require \( N \leq K \) which is needed in Wang, et al. 2012 and Liu, et al. 2012. Thus, \( N \) can have more choices. When \( N = 1 \), this sufficient condition turns to be a sufficient condition for support recovery with OMP. We show that it is weaker than that in Wang 2015 in terms of both SNR and RIP. Moreover, in the noise free case, we obtain that \( \delta_{NK+1} < 1/\sqrt{K/N + 1} \) is a sufficient condition for recovering the \( K \)-sparse signal \( x \) with gOMP in \( K \) iterations which is better than the best known one in terms of \( \delta_{NK+1} \). In particular, this condition is sharp when \( N = 1 \).

Index Terms

Compressed sensing, restricted isometry constant, orthogonal matching pursuit, orthogonal multi-matching pursuit, support recovery.

I. INTRODUCTION

In compressed sensing (CS) setting, we usually observe the following linear model [4], [9], [6]:

\[
y = Ax + v,
\]

(1)

where \( x \in \mathbb{R}^n \) is a \( K \)-sparse unknown signal (i.e., |supp\((x)\)| \( \leq K \), where supp\((x)\) = \{i : \( x_i \neq 0 \)\} is the support of \( x \), and |supp\((x)\)| is the cardinality of supp\((x)\)), \( y \in \mathbb{R}^m \) is an observation vector, \( A \in \mathbb{R}^{m \times n} \) (with \( m \ll n \)) is a known sensing matrix and \( v \in \mathbb{R}^m \) is a noise vector.

One of the central goals in CS is to recover the unknown sparse vector \( x \) in (1) using some efficient algorithms based on \( A \) and \( y \). The state-of-the-art CS theory (see, e.g., [4], [3], [9]) show that this can be done under appropriate conditions on \( A \). A commonly used framework for characterizing such conditions is the so-called restricted isometry property (RIP) [4]. For any \( m \times n \) matrix \( A \) and any integer \( K \) with \( 1 \leq K \leq m \), the restricted isometry constant (RIC) \( \delta_K \) of order \( K \) is defined as the smallest constant such that

\[
1 - \delta_K \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_K )\|x\|_2^2
\]

(2)

for all \( K \)-sparse vectors \( x \).

The orthogonal matching pursuit (OMP) [18] is a well-known greedy algorithm for recovering sparse signals. The generalized orthogonal matching pursuit (gOMP) [20], also called orthogonal multi-matching pursuit (OMMP) in [12], is a natural generalization of OMP in the sense that multiple \( N \geq 1 \) indices are identified per iteration. Simulations in [20] and [12] indicate that, compared with the original OMP, the gOMP has better sparse recovery performance with less CPU time. The gOMP is described in Algorithm 1, where \( A_S \) denote the submatrix of \( A \) that contains the columns indexed by set \( S \subset \{1, 2, \ldots, n\} \), and \( x_S \) denote the subvector of \( x \) that only contains the entries indexed by \( S \). Note that when \( N = 1 \), gOMP turns to OMP.

Many RIC-based conditions have been proposed to guarantee exact recovery of \( K \)-sparse signals with gOMP in the noise free case (i.e., \( v = 0 \)) for general \( N \). It were respectively shown in [20] and [12] that \( \delta_{NK} < 1/(\sqrt{K/N} + 3) \) and \( \delta_{NK} < 1/(2 + \sqrt{2})\sqrt{K/N} \) are sufficient conditions for gOMP to recover \( x \) in \( K \) iterations. Later, the condition was improved to \( \delta_{NK} < 1/(\sqrt{K/N} + 2) \) and \( \delta_{NK+1} < 1/\sqrt{K/N} + 1 \) in [15]. Recently, it was further improved to \( \delta_{NK} < 1/(\sqrt{K/N} + 1.27) \)

J. Wen is with ENS de Lyon, Laboratoire LIP (U. Lyon, CNRS, ENSL, INRIA, UCBL), Lyon, France, 69007, (e-mail: jwen@math.mcgill.ca)
Z. Zhou is with the School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China, (e-mail: zze@home.swjtu.edu.cn).
D. Li is with the School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China, and Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong, (e-mail: dfli@hust.edu.cn).
X. Tang is with the Information Security and National Computing Grid Laboratory, Southwest Jiaotong University, Chengdu 610031, China (e-mail: xhtang@swjtu.edu.cn).

This research was supported by "Programme Avenir Lyon Saint-Etienne de l’Université de Lyon" in the framework of the programme "Investissements d’Avenir" (ANR-11-IDEX-0007), ANR through the HPAC project under Grant ANR 11 BS02 013, NSFC (No. 11571128), a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 11302915)

Manuscript received; revised.
in [16]. It is worthwhile pointing out that there are more results on sparse recovery with OMP (the special case of gOMP for $N = 1$), see, e.g., [8], [12], [21] and [5]. To the best of our knowledge, the best sufficient condition for recovering $K$-sparse signals in $K$ iterations is $\delta_{K+1} < 1/\sqrt{K+1}$ in [13]. On the other hand, it was proved in [23] that for any given positive integer $K \geq 2$ and any given $t$ satisfies $1/\sqrt{K+1} \leq t < 1$, there always exist a $K$-sparse $x$ and a matrix $A$ satisfies the RIP of order $K+1$ with $\delta_{K+1} = t$ such that OMP may fail to recover $x$ in $K$ iterations (here we want to point out that this necessary condition also works for $K = 1$). Thus, $\delta_{K+1} < 1/\sqrt{K+1}$ is a sharp condition for exact recovery of $K$-sparse signals with OMP in $K$ iterations.

Sufficient conditions of support recovery of $K$-sparse signals in the presence of noise with gOMP have also been widely studied (see [2], [1], [7], [11], [17], [5], and the references therein). As one of the latest results, it was proved in [11] that under some conditions on the minimum magnitude of the nonzero elements $x_i$, $\delta_{N,K+1} < 1/\sqrt{K/N+1}$ is a sufficient condition for support recovery under the $l_2 (||v||_2 \leq \epsilon$ for some constant $\epsilon$, see, e.g. [2]) and $l_\infty (||A^Tv||_\infty \leq \epsilon$, see, e.g. [1]) bounded noises.

In this paper, we investigate the RIP based sufficient conditions for support recovery in the presence of noise with gOMP. Instead of considering $l_2$ and $l_\infty$ bounded noises separately (see, e.g., [7], [11], [17] and [5]), we follow [10] and [19] which use signal-to-noise ratio (SNR) and minimum-to-average ratio (MAR) to measure $v$ and $x$, which are respectively defined by

$$\text{SNR} = \begin{cases} \frac{||Ax||_2^2}{||v||_2^2} & v \neq 0 \\ +\infty & v = 0 \end{cases}$$

and

$$\text{MAR} = \frac{\min_{i \in \Omega} |x_i|^2}{||x||_2^2/K}.$$  

We show that under some conditions on SNR, gOMP is ensured to recover at least one index in the support of $x$ in each iteration if $\delta_{N,K+1} < 1/\sqrt{K/N+1}$. As consequences, we have the following contributions:

- Unlike [20] and [12], which require $N \leq \min(K, m/K)$, our condition on $N$ is only $N \leq (m - 1)/K$ which ensures that the assumption $\delta_{N,K+1} < 1/\sqrt{K/N+1}$ makes sense. This allows more choices of $N$ for gOMP.

- The sufficient condition for support recovery with gOMP turns to the condition for OMP when $N = 1$, and it is weaker than that in [19] in terms of both SNR and RIP.

- In the noise free case, we obtain that $\delta_{N,K+1} < 1/\sqrt{K/N+1}$ is a sufficient condition for recovering $K$-sparse signals with gOMP in $K$ iterations. This improves the best known condition $\delta_{N,K+1} < 1/((\sqrt{K/N})+1)$ in [15]. In particular, when $N = 1$, our sufficient condition $\delta_{K+1} < 1/\sqrt{K+1}$ is exactly the one in [13]. As aforementioned, this condition is sharp according to the result in [23].

The rest of the paper is organized as follows. We give some useful notation and lemmas in section II. We present our main results in section III. Finally, this paper is summarized in section IV.

### II. Notation and Useful Lemmas

#### A. Notation

Throughout this paper, we adopt the following notations unless otherwise stated. Let $\mathbb{R}$ be the real field. Boldface lowercase letters denote column vectors, and boldface uppercase letters denote matrices, e.g., $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{m \times n}$. Let $0$ denote a zero column vector. Let $\Omega$ be the support of $X$ and $|\Omega|$ be the cardinality of $\Omega$. Let set $S \subset \{1, 2, \ldots, n\}$, and $\Omega \setminus S = \{i| i \in \Omega, i \notin S\}$. Let $\Omega^c$ and $S^c$ be the complement of $\Omega$ and $S$, i.e., $\Omega^c = \{1, 2, \ldots, n\} \setminus \Omega$, and $S^c = \{1, 2, \ldots, n\} \setminus S$. Let $A_S$ be the submatrix of $A$ that only contains the columns indexed by $S$, and $x_S$ be the subvector of $x$ that only contains the entries indexed by $S$, and $A_S^T$ be the transpose of $A_S$. For full column rank matrix $A_S$, let $P_S = A_S(A_S^T A_S)^{-1} A_S^T$ and $P_S^T = I - P_S$ denote the projector and orthogonal complement projector on the column space of $A_S$, respectively.

---

**Algorithm 1 gOMP**

Input: measurements $y \in \mathbb{R}^m$, sensing matrix $A \in \mathbb{R}^{m \times n}$, sparsity $K$, number of indexes to be chosen per iteration $N$.

Initialize: $k = 0$, $r^0 = y$, $S_0 = \emptyset$.

1: while $k < K$ and $||r^k||_2 > 0$ do
2: \hspace{2em} $k = k + 1$
3: \hspace{4em} Choose indexes $i_1, \ldots, i_N$ corresponding to the $N$ largest magnitude of $|A^T r^{k-1}|$, $S_k = S_{k-1} \cup \{i_1, \ldots, i_N\}$. Let $\hat{x}_{S_k} = \arg\min \{||y - A_{S_k} x||_2\}$.
4: \hspace{4em} $r^{k} = y - A_{S_k} \hat{x}_{S_k}$
5: end while

Output: $\hat{x} = \arg\min_{x, \text{supp}(x) = S_k} ||y - Ax||_2$. 

---


B. Useful lemmas

We now introduce some lemmas that will be useful in the sequel.

Lemma 1 ([4]): If a matrix $A \in \mathbb{R}^{m \times n}$ satisfies the RIP of orders $K_1$ and $K_2$ with $K_1 < K_2$, then $\delta_{K_1} \leq \delta_{K_2}$.

Lemma 2 ([16]): Let $S_1, S_2$ be two subsets of \{1, 2, ..., $n$\} with $|S_2 \setminus S_1| \geq 1$. If a matrix $A \in \mathbb{R}^{m \times n}$ satisfies the RIP of order $|S_1 \cup S_2|$, then for any vector $x \in \mathbb{R}^{S_1 \setminus S_1}$, 
\[
(1 - \delta_{|S_1 \cup S_2|}) \|x\|_2^2 \leq \|P_{S_1} A_{S_2 \setminus S_1} x\|_2^2 \leq (1 + \delta_{|S_1 \cup S_2|}) \|x\|_2^2.
\]

Lemma 3 ([14]): Let $A$ satisfy the RIP of order $K$ and $S$ be a subset of \{1, 2, ..., $n$\} with $|S| \leq K$, then for any $x \in \mathbb{R}^m$, 
\[
\|A^T_S x\|_2^2 \leq (1 + \delta_K) \|x\|_2^2.
\]

The following lemma is the key to the proof of our main results in the next section.

Lemma 4: Suppose that $x$ in (1) is a sparse vector with support $\Omega$. Let set $S \subseteq \{1, 2, \ldots, n\}$ satisfy $|S| = Kn$ and $|\Omega \cap S| = l$ for some integers $N$, $k$, and $l$ with $0 \leq k \leq l \leq |\Omega| - 1$ and $N(k + 1) + |\Omega| - k \leq m$. Let $W \subseteq \Omega$ be a set with $|W| = N$ and $W \cap S = \emptyset$. If the matrix $A$ in (1) satisfies the RIP of order $N(k + 1) + |\Omega| - l$ with $0 \leq \delta_{N(k+1)+|\Omega|-l} < 1$. Then
\[
\max_{i \in \Omega \setminus S} |A^T_{i\Omega \setminus S} P_S^A A_{i\Omega \setminus S} x_{i\Omega \setminus S}| - \frac{1}{N} \sum_{j \in W} |A^T_{j\Omega \setminus S} P_S^A A_{j\Omega \setminus S} x_{j\Omega \setminus S}| 
\geq \frac{(1 - \sqrt{(|\Omega| - l)/N + 1}) \delta_{N(k+1)+|\Omega|-l}) \|x_{i\Omega \setminus S}\|_2}{\sqrt{|\Omega| - l}}.
\]

Proof: See Appendix A.

Remark 1: The condition $N(k+1)+|\Omega|-k \leq m$ in Lemma 4 is to ensure that $A$ satisfies the RIP of order $N(k+1)+|\Omega|-l$ makes sense.

If $N = 1$, then from Lemma 4, $l = k$. Hence, we have the following result which is Lemma 1 in [22].

Corollary 1: Suppose that $x$ in (1) is a sparse vector with support $\Omega$. Let $S$ be a subset of $\Omega$ with $|S| < |\Omega|$. If $A$ in (1) satisfies the RIP of order $|\Omega| + 1$ with $0 \leq \delta_{|\Omega|+1} < 1$. Then,
\[
\|A^T_{i\Omega \setminus S} P_S^A A_{i\Omega \setminus S} x_{i\Omega \setminus S}\| - \|A^T_{j\Omega \setminus S} P_S^A A_{j\Omega \setminus S} x_{j\Omega \setminus S}\| \geq \frac{(1 - \sqrt{|\Omega| - |S| + 1}) \delta_{N(k+1)+|\Omega|-k}) \|x_{i\Omega \setminus S}\|_2}{\sqrt{|\Omega| - |S|}}.
\]

Remark 2: By Lemma 4 and Lemma 1, it is easy to verify that the righthand side of (5) is positive if the matrix $A$ in (1) satisfies the RIP of order $N(k+1) + |\Omega| - k$ with
\[
\delta_{N(k+1)+|\Omega|-k} < \frac{1}{\sqrt{(|\Omega| - k)/N + 1}}.
\]

III. Main results

In this section, we will show that if $A$ in (1) satisfies the RIP of order $NK + 1$ with
\[
\delta_{NK + 1} < \frac{1}{\sqrt{K/N + 1}},
\]
then gOMP is ensured to recover at least one index in the support of $K$-sparse signals in each iteration under some conditions on SNR and MAR. In particular, in the noise free case (i.e., $v = 0$), we will show that (6) is a sufficient condition for recovering $K$-sparse signals in $K$ steps with gOMP. To achieve these goals, we need the following result.

Theorem 1: Suppose that $x$ in (1) is a sparse vector with support $\Omega$, and $A$ in (1) satisfies the RIP of order $N(k+1)+|\Omega|-k$ with
\[
\delta_{N(k+1)+|\Omega|-k} < \frac{1}{\sqrt{|\Omega|/N + 1}}.
\]
for some integers $k$ and $N$ satisfying $0 \leq k \leq |\Omega| - 1$ and $N(k + 1) + |\Omega| - k \leq m$. Then gOMP selects at least one index in $\Omega$ in each iteration of the first $k + 1$ ones before all the indexes in $\Omega$ have been selected and before gOMP terminates provided that
\[
\sqrt{\text{SNR}} > \frac{\sqrt{2K(1 + \delta_{N(k+1)+|\Omega|-k})}}{(1 - \sqrt{|\Omega|/N + 1})\delta_{N(k+1)+|\Omega|-k}} \sqrt{\text{MAR}}.
\]

Proof: See Appendix B.
If \( x \) in (1) is a \( K \)-sparse vector with support \( \Omega \), then by Lemma 1, (7) also holds if (6) holds. Thus, by Theorem 1, we can easily obtain the following result.

**Theorem 2:** Suppose that \( x \) in (1) is a \( K \)-sparse vector with support \( \Omega \), and \( A \) in (1) satisfies the RIP of order \( NK + 1 \) with \( \delta_{NK + 1} \) satisfying (6) for an integer \( N \) with \( 1 \leq N \leq (m - 1)/K \). Then gOMP either recovers at least \( k_0 \) indexes in \( \Omega \) if gOMP terminates after performing \( k_0 \) iterations with \( 1 \leq k_0 < K \) or recovers \( \Omega \) in \( K \) iterations provided that

\[
\sqrt{\text{SNR}} > \frac{\sqrt{2K(1 + \delta_{NK + 1})}}{(1 - \sqrt{|\Omega|/N + 1}\delta_{NK + 1})\sqrt{\text{MAR}}}.
\]

When \( N = 1 \), gOMP turns to OMP. In this case, the following result for OMP can be directly obtained from Theorem 2 which significantly improves a known one in [19].

**Corollary 2:** Suppose that \( x \) in (1) is a \( K \)-sparse vector with support \( \Omega \), and \( A \) in (1) satisfies the RIP of order \( K + 1 \) with \( \delta_{K + 1} < \frac{1}{\sqrt{K + 1}} \).

Then OMP either recovers at least \( k_0 \) indexes in \( \Omega \) if it terminates after performing \( k_0 \) iterations with \( 1 \leq k_0 < K \) or it recovers \( \Omega \) in \( K \) iterations provided that

\[
\sqrt{\text{SNR}} > \frac{\sqrt{2K(1 + \delta_{K + 1})}}{(1 - \sqrt{|\Omega|/N + 1}\delta_{K + 1})\sqrt{\text{MAR}}}.
\]

**Remark 3:** With the same notations as in Corollary 2, it was shown in [19, Theorem 3.1] that OMP either recovers at least \( k_0 \) indexes in \( \Omega \) if OMP terminates after performing \( k_0 \) iterations with \( 1 \leq k_0 < K \) or recovers \( \Omega \) in \( K \) iterations provided that

\[
\delta_{K + 1} < \frac{1}{\sqrt{K + 1}}
\]

and

\[
\sqrt{\text{SNR}} > \frac{2\sqrt{K(1 + \delta_{K + 1})}}{(1 - \sqrt{|\Omega|/N + 1}\delta_{K + 1})\sqrt{\text{MAR}}}.
\]

It is easy to see that our sufficient condition given in Corollary 2 is weaker than that in [19, Theorem 3.1] in terms of both RIC and SNR. Furthermore, by the necessary condition on SNR given by [19, Theorem 3.2], the constraint on SNR (see (11)) is very tight. But the necessary condition is out of the scope of this paper.

Notice that gOMP may terminate before performing \( K \) iterations, and in this case \( \Omega \) is not guaranteed to be recovered by gOMP by Theorem 2 under (6) and (9). However, in the noise-free case, (6) is a sufficient condition for recovering \( x \) with gOMP in \( K \) iterations. Specifically, we have the following result.

**Theorem 3:** Suppose that \( v \) in (1) satisfies \( v = 0 \), \( x \) in (1) is a \( K \)-sparse vector with support \( \Omega \), and \( A \) in (1) satisfies the RIP of order \( NK + 1 \) with \( \delta_{NK + 1} \) satisfying (6) for integer \( N \) with \( 1 \leq N \leq (m - 1)/K \). Then gOMP recovers \( x \) in \( K \) iterations.

**Proof:** The result follows directly from Theorem 2 and Lemma 5 below.

**Remark 4:** In the noise-free case, the best known condition on \( \delta_{NK + 1} \) such that gOMP recovers \( x \) in \( K \) iterations is \( \delta_{NK + 1} < 1/(\sqrt{K/N} + 1) \) in [15]. Obviously, our sufficient condition given in Theorem 3 is better.

**Lemma 5:** Suppose that \( x \) in (1) is a sparse vector with support \( \Omega \), and \( A \) in (1) satisfies the RIP of order \( N(k + 1) + |\Omega| - k \) with \( \delta_{N(k + 1) + |\Omega| - k} \) satisfying (7) for some integers \( k \) and \( N \) with \( 1 \leq k \leq |\Omega| - 1 \) and \( 1 \leq N \leq (m - 1)/K \). If there exists an integer \( k_0 \) with \( 0 < k_0 \leq k \) and \( |\Omega \cap S_{k_0}| \geq k_0 \) such that \( ||r^{k_0}||_2 = 0 \) (see Algorithm 1 for the definitions of \( S_{k_0} \) and \( r^{k_0} \)). Then \( \Omega \subseteq S_{k_0} \).

**Proof:** We prove this lemma by contradiction. Suppose that \( \Omega \not\subseteq S_{k_0} \) and let \( \Gamma = \Omega \cup S_{k_0} \). Let \( \tilde{x}, \hat{x} \in \mathbb{R}^{|\Gamma|} \) satisfy \( \tilde{x}_i = x_i \) for \( i \in \Omega \) and \( \tilde{x}_i = 0 \) for \( i \notin \Omega \), and \( \hat{x}_i = (\hat{x}_{S_{k_0}})_i \) for \( i \in S_{k_0} \) and \( \hat{x}_i = 0 \) for \( i \notin S_{k_0} \), where \( \hat{x}_{S_{k_0}} \) is the vector generated by Algorithm 1. Since \( ||r^{k_0}||_2 = 0 \), by line 6 of Algorithm 1, \( A_{S_{k_0}} \hat{x}_{S_{k_0}} = y \), we have

\[
A_{\Gamma} \hat{x} = A_{\Omega} x_{\Omega} = Ax = y = A_{S_{k_0}} \hat{x}_{S_{k_0}} = A_{\Gamma} \hat{x}.
\]

Note that \( |\Omega \cap S_{k_0}| \geq k_0 \) and \( \Gamma = \Omega \cup S_{k_0} \). It then follows that

\[
|\Gamma| = |\Omega| + |S_{k_0}| - |\Omega \cap S_{k_0}| \leq |\Omega| + Nk - k \leq N(k + 1) + |\Omega| - k.
\]

Combing with (12) and (7), we obtain \( \hat{x} = \tilde{x} \).

On the other hand, by the definitions of \( \tilde{x} \) and \( \hat{x} \), and the assumption that \( \Omega \not\subseteq S_{k_0} \), there exists \( j \in (\Omega \setminus S_{k_0}) \) such that \( \tilde{x}_j \neq 0 \) and \( \hat{x}_j = 0 \). This implies that \( \hat{x} \neq \tilde{x} \) and leads to contradiction with \( \hat{x} = \tilde{x} \). Completing the proof.

**Corollary 3:** Suppose that \( v \) in (1) satisfies \( v = 0 \), \( x \) in (1) is a \( K \)-sparse vector with support \( \Omega \), and \( A \) in (1) satisfies the RIP of order \( K + 1 \) with \( \delta_{K + 1} < \frac{1}{\sqrt{K + 1}} \). Then OMP recovers \( x \) in \( K \) iterations.
Remark 5: It was shown in [23] that for any given positive integer \( K \geq 1 \) and any \( 1/\sqrt{K+1} \leq t < 1 \), there always exist a \( K \)-sparse \( x \) and a matrix \( A \) satisfies the RIP of order \( K+1 \) with \( \delta_{K+1} = t \) such that OMP may fail to recover \( x \) in \( K \) iterations. Thus, in the noise free case, \( \delta_{K+1} < \frac{1}{\sqrt{K+1}} \) given in Corollary 3 is a sharp condition for OMP to recover \( K \)-sparse signals in \( K \) iterations. However, for general \( N \), it is still open whether \( \delta_{N,K+1} < \frac{1}{\sqrt{K/N+1}} \) given in Theorem 3 is sharp or not for gOMP to recover \( K \)-sparse signals in \( K \) iterations. The reader is kindly invited to attack this problem.

IV. Conclusion

In this paper, we have studied the sufficient conditions for sparse recovery with gOMP based on the RIP of the sensing matrix. We have shown that under some conditions on SNR, \( \delta_{N,K+1} < 1/\sqrt{K/N+1} \) is a sufficient condition for support recovery of a \( K \)-sparse signal \( x \) with gOMP. Surprisingly, our sufficient condition do not require \( N \leq K \) (which is needed in [20] and [12]), which provides more choices for \( N \). When \( N = 1 \), the sufficient condition is for support recovery with OMP and it is better than that in [19] in terms of both RIC and SNR. In the noise free case, we have also showed that \( \delta_{N,K+1} < 1/\sqrt{K/N+1} \) is a sufficient condition for recovering \( x \) with gOMP in \( K \) iterations, which is much better than the best known one in terms of \( \delta_{N,K+1} \) in [15]. Moreover, the condition is sharp when \( N = 1 \).

APPENDIX A

Proof of Lemma 4

Since \( |\Omega \cap S| = l \) for \( 0 \leq l \leq |\Omega| - 1 \), \( \|x_{\Omega \setminus S}\|_1 \neq 0 \). Thus, we have

\[
\max_{i \in \Omega \setminus S} |A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}| \geq \frac{1}{\|x_{\Omega \setminus S}\|_1} \left( \sum_{j \in \Omega \setminus S} |x_j| \right) \max_{i \in \Omega \setminus S} |A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}|
\]

\[
\geq \frac{1}{\|x_{\Omega \setminus S}\|_2} \left( \sum_{j \in \Omega \setminus S} |x_j| \right) \max_{i \in \Omega \setminus S} |A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}|
\]

\[
\geq \frac{1}{\|x_{\Omega \setminus S}\|_2} \left( \sum_{j \in \Omega \setminus S} (x_j A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}) \right)
\]

\[
\geq \frac{1}{\|x_{\Omega \setminus S}\|_2} \left( \sum_{j \in \Omega \setminus S} (x_j A_i^T (P_{\perp S})^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}) \right)
\]

\[
= \frac{1}{\|x_{\Omega \setminus S}\|_2} \|P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}\|_2^2,
\]

where (a) follows from \( |\text{supp}(x_{\Omega \setminus S})| = |\Omega| - l \) and Cauchy-Schwarz inequality; (b) is because for each \( j \in \Omega \setminus S \),

\[
\max_{i \in \Omega \setminus S} |A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}| \geq A_j^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S};
\]

And (c) is from

\[
(P_{\perp S})^T P_{\perp S} = P_{\perp S}^T P_{\perp S} = P_{\perp S}. \tag{13}
\]

Thus,

\[
\|P_{\perp S}^T A_{\Omega \setminus S} x_{\Omega \setminus S}\|_2^2 \leq \sqrt{|\Omega| - l} \|x_{\Omega \setminus S}\|_2 \max_{i \in \Omega \setminus S} |A_i^T P_{\perp S} A_{\Omega \setminus S} x_{\Omega \setminus S}|. \tag{14}
\]

Let

\[
\alpha = -\frac{\sqrt{|\Omega| - l}/N + 1 - 1}{\sqrt{|\Omega| - l}/N}, \tag{15}
\]

then by some simple calculations, we obtain

\[
\frac{2\alpha}{1 - \alpha^2} = -\sqrt{|\Omega| - l}/N, \quad \frac{1 + \alpha^2}{1 - \alpha^2} = \sqrt{|\Omega| - l}/N + 1. \tag{16}
\]
To simplify notation, let \( W = \{j_1, j_2, \ldots, j_N\} \) (note that \(|W| = N\)), and define \( \overline{e}_W \in \mathbb{R}^N \) with

\[
\overline{e}_i = \begin{cases} 
1 & \text{if } A_{j_i}^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} \geq 0 \\
-1 & \text{if } A_{j_i}^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} < 0
\end{cases}
\]

for \( 1 \leq i \leq N \). Then, it is easy to see that

\[
\overline{e}_W^T A_W^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} = \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|. 
\]

(17)

Furthermore, we define

\[
B = P_S^\perp A_{(\Omega \setminus S) \cup W} = P_S^\perp A_{(\Omega \cup W) \setminus S}, 
\]

\[
u = \begin{bmatrix} x_{\Omega \setminus S} & 0 \end{bmatrix}^T \in \mathbb{R}^{|\Omega \setminus S|+N} 
\]

\[
w = \begin{bmatrix} \alpha \|x_{\Omega \setminus S}\|_2 \overline{e}_W \end{bmatrix}^T \in \mathbb{R}^{|\Omega \setminus S|+N}, 
\]

where (a) is because \( W \cap \Omega = \emptyset \). Then,

\[
Bu = P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}, 
\]

(19)

and

\[
\|u + w\|_2^2 = (1 + \alpha^2)\|x_{\Omega \setminus S}\|_2^2, 
\]

(20)

\[
\|\alpha^2 u - w\|_2^2 = \alpha^2 (1 + \alpha^2)\|x_{\Omega \setminus S}\|_2^2. 
\]

(21)

Thus,

\[
w^T B^T Bu \\
\overset{(a)}{=} w^T A_{(\Omega \setminus S) \cup W}^T (P_S^\perp)^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} \\
\overset{(b)}{=} w^T A_{(\Omega \setminus S) \cup W}^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} \\
\overset{(c)}{=} \|x_{\Omega \setminus S}\|_2 \overline{e}_W^T A_W^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S} \\
\overset{(d)}{=} \frac{\|x_{\Omega \setminus S}\|_2}{\sqrt{N}} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|,
\]

where (a) follows from (18) and (19); and (b) follows from (13), and (c) is from (17). Therefore, we have

\[
\|B(u + w)\|_2^2 \\
= \|Bu\|_2^2 + \|Bw\|_2^2 + 2w^T B^T Bu \\
= \|Bu\|_2^2 + \|Bw\|_2^2 + \frac{2\alpha \|x_{\Omega \setminus S}\|_2}{\sqrt{N}} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|,
\]

and

\[
\|B(\alpha^2 u - w)\|_2^2 = \alpha^4 \|Bu\|_2^2 + \|Bw\|_2^2 \\
- \frac{2\alpha^3 \|x_{\Omega \setminus S}\|_2}{\sqrt{N}} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|.
\]
Thus, by the aforementioned equations, we have
\[
\|B(u + w)\|_2^2 - \|B(\alpha^2 u - w)\|_2^2
= (1 - \alpha^4) \|Bu\|_2^2
+ 2\alpha(1 + \alpha^2) \frac{\|x_{\Omega \setminus S}\|_2}{\sqrt{N}} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|
= (1 - \alpha^4)
\times (\|Bu\|_2^2 + \frac{2\alpha}{1 - \alpha^2} \frac{\|x_{\Omega \setminus S}\|_2}{\sqrt{N}} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|)
= (1 - \alpha^4)
\times (\|Bu\|_2^2 - \frac{\sqrt{|\Omega| - l}}{N} \|x_{\Omega \setminus S}\|_2 \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|),
\]
where the last equality follows from the first equality in (16).

It is not hard to check that
\[
\|B(u + w)\|_2^2 - \|B(\alpha^2 u - w)\|_2^2
\geq (1 - \delta_{N(k+1)+|\Omega|-l}) \|(u + w)\|_2^2
- (1 + \delta_{N(k+1)+|\Omega|-l}) \|(\alpha^2 u - w)\|_2^2
\]
\[= (1 - \delta_{N(k+1)+|\Omega|-l})(1 + \alpha^2) \|x_{\Omega \setminus S}\|_2^2
- (1 + \delta_{N(k+1)+|\Omega|-l}) \alpha^2(1 + \alpha^2) \|x_{\Omega \setminus S}\|_2^2
= (1 + \alpha^2) \|x_{\Omega \setminus S}\|_2^2 \left(1 - \delta_{N(k+1)+|\Omega|-l}ight)
- (1 + \delta_{N(k+1)+|\Omega|-l}) \alpha^2
\]
\[= (1 + \alpha^2) \|x_{\Omega \setminus S}\|_2^2 \left(1 - \delta_{N(k+1)+|\Omega|-l}ight)
= (1 - \alpha^4) \|x_{\Omega \setminus S}\|_2^2 \left(1 - \sqrt{|\Omega| - l}/N + 1\delta_{N(k+1)+|\Omega|-l}\right),
\]
where (a) follows from Lemma 2 and (18), (b) follows from (20) and (21), and (c) follows from the second equality in (16).

By (19), (22), (23) and the fact that $1 - \alpha^4 > 0$, we have
\[
\frac{\|P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}\|_2^2}{N}
- \frac{\sqrt{|\Omega| - l}}{N} \|x_{\Omega \setminus S}\|_2 \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}|
\geq \|x_{\Omega \setminus S}\|_2^2 \left(1 - \sqrt{|\Omega| - l}/N + 1\delta_{N(k+1)+|\Omega|-l}\right).
\]

Thus, combining with (14), we obtain
\[
\sqrt{|\Omega| - l} \|x_{\Omega \setminus S}\|_2
\times \left( \max_{i \in \Omega \setminus S} |A_i^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}| - \frac{1}{N} \sum_{j \in W} |A_j^T P_S^\perp A_{\Omega \setminus S} x_{\Omega \setminus S}| \right)
\geq \|x_{\Omega \setminus S}\|_2^2 \left(1 - \sqrt{|\Omega| - l}/N + 1\delta_{N(k+1)+|\Omega|-l}\right).
\]

Therefore, (5) holds. □

**Appendix B**

**Proof of Theorem 1**

We prove the result by induction. Suppose that gOMP selects at least one correct indexes in the first $\tilde{k} - 1$ (with $1 \leq \tilde{k} \leq k+1$) iterations, then $l = |S_{\tilde{k}-1} \cap \Omega| \geq \tilde{k} - 1$. We assume $\Omega \not\subseteq S_{\tilde{k}-1}$ and Algorithm 1 performs at least $\tilde{k}$ iterations. Then, we need to show that $|(S_{\tilde{k}} \setminus S_{\tilde{k}-1}) \cap \Omega| \geq 1$. Since $S_0 = 0$, the induction assumption $|\Omega| > |S_{\tilde{k}-1} \cap \Omega| \geq \tilde{k} - 1$ holds with $\tilde{k} = 1$. So, the proof for the first iteration is contained in the case that $\tilde{k} = 1$. 

Let $j_1, j_2, \ldots, j_n \in \Omega^c$ such that
\[
|A_{j_1}^T r^{k-1}| \geq |A_{j_2}^T r^{k-1}| \geq \ldots \geq |A_{j_n}^T r^{k-1}|
\]
where we define
\[
W = \{j_1, j_2, \ldots, j_n\}. \tag{26}
\]
Then to show $|(S_k \setminus S_{k-1}) \cap \Omega| \geq 1$, we only need to show
\[
\max_{i \in \Omega} |A_i^T r^{k-1}| > |A_{j_n}^T r^{k-1}|.
\]
By (25), it suffices to show
\[
\max_{i \in \Omega} |A_i^T r^{k-1}| > \frac{1}{N} \sum_{j \in W} |A_j^T r^{k-1}|. \tag{27}
\]
By lines 4 and 5 of Algorithm 1, we have
\[
r^{k-1} = y - A_{S_{k-1}} \hat{x}_{S_{k-1}} = (I - A_{S_{k-1}} (A_{S_{k-1}}^T A_{S_{k-1}})^{-1} A_{S_{k-1}}^T) y = (a) P_{S_{k-1}}^\perp (Ax + v) = (b) P_{S_{k-1}}^\perp (A_{\Omega} x_\Omega + v) = P_{S_{k-1}}^\perp (A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}} + A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}} + v) = (c) P_{S_{k-1}}^\perp A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}} + P_{S_{k-1}}^\perp v, \tag{28}
\]
where (a), (b) and (c) follows from the definition of $P_{S_{k-1}}^\perp$, the fact that $\Omega = \text{supp}(x)$ and $P_{S_{k-1}}^\perp A_{S_{k-1}} = 0$, respectively.

By lines 3 and 4 of Algorithm 1, for each $i \in S_{k-1}$, $|A_i^T r^{k-1}| = 0$. Thus, by (28)
\[
\max_{i \in \Omega} |A_i^T r^{k-1}| = \max_{i \in \Omega \setminus S_{k-1}} \{|A_i^T P_{S_{k-1}}^\perp A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}}| - |A_i^T P_{S_{k-1}}^\perp v|\}. \tag{29}
\]
Similarly, by (28),
\[
\frac{1}{N} \sum_{j \in W} |A_j^T r^{k-1}| \leq \frac{1}{N} \sum_{j \in W} |A_j^T P_{S_{k-1}}^\perp A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}}| + \max_{j \in W} |A_j^T P_{S_{k-1}}^\perp v|. \tag{30}
\]
To simplify notation, we define
\[
\beta_1 = \max_{i \in \Omega \setminus S_{k-1}} |A_i^T P_{S_{k-1}}^\perp A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}}| - \frac{1}{N} \sum_{j \in W} |A_j^T P_{S_{k-1}}^\perp A_{\Omega \setminus S_{k-1}} x_{\Omega \setminus S_{k-1}}|, \tag{31}
\]
and
\[
\beta_2 = \max_{i \in \Omega \setminus S_{k-1}} |A_i^T P_{S_{k-1}}^\perp v| + \max_{j \in W} |A_j^T P_{S_{k-1}}^\perp v|. \tag{32}
\]
Then, by (29)-(32), to show (27), it suffices to show
\[
\beta_1 > \beta_2. \tag{33}
\]
In the following, we give an upper bound on $\beta_2$. Clearly there exist $i_0 \in \Omega \setminus S_{k-1}$ and $j_0 \in W$ such that
\[
\max_{i \in \Omega \setminus S_{k-1}} |A_i^T P_{S_{k-1}}^\perp v| = |A_{i_0}^T P_{S_{k-1}}^\perp v|;
\]
\[
\max_{j \in W} |A_j^T P_{S_{k-1}}^\perp v| = |A_{j_0}^T P_{S_{k-1}}^\perp v|.
Therefore

\[
\beta_2 = \| A_{i_{0}:j_{0}}^T P_{S_{k-1}}^\perp v \|_1 \leq \sqrt{2} \| A_{i_{0}:j_{0}}^T P_{S_{k-1}}^\perp v \|_2 \\
\leq \sqrt{2(1 + \delta_N(k+1)+|\Omega|-k)} \| v \|_2
\]

(34)

where (a) is because \( A_{i_{0}:j_{0}}^T P_{S_{k-1}}^\perp v \) is a \( 2 \times 1 \) vector, (b) follows from Lemma 3, and

\[
\| P_{S_{k-1}}^\perp v \|_2 \leq \| P_{S_{k-1}}^\perp v \|_2 \leq \| v \|_2 \leq \epsilon.
\]

In the following, we give a lower bound on \( \beta_1 \). By line 3 of Algorithm 1, \( |S_{k-1}| = (\bar{k}-1)N \). By the induction assumption

\[
0 \leq \bar{k} - 1 \leq |\Omega \cap S_{k-1}| = l \leq |\Omega| - 1.
\]

(35)

By the definition of set \( W \), \( W \subset \Omega^c \) and \( |W| = N \). Thus, by Lemma 1 and 4, and (31), we obtain

\[
\beta_1 \geq \frac{(1 - \sqrt{|\Omega| - l)/N + [\delta_N(k+1)+|\Omega|-k]}) \| x_{\Omega \setminus S_{k-1}} \|_2}{\sqrt{|\Omega| - l}}
\]

\[
\geq \frac{(1 - \sqrt{|\Omega| - l)/N + [\delta_N(k+1)+|\Omega|-k]}) \| x_{\Omega \setminus S_{k-1}} \|_2}{\sqrt{|\Omega| - l}}
\]

\[
\geq \frac{(1 - \sqrt{|\Omega|/N + [\delta_N(k+1)+|\Omega|-k]}) \| x_{\Omega \setminus S_{k-1}} \|_2}{\sqrt{|\Omega| - l}},
\]

(36)

where the second and third inequality respectively follows from (35) and the fact that \( \bar{k} \leq k + 1 \).

By [19, 25], we have

\[
\| x_{\Omega \setminus S_{k-1}} \|_2 \geq \sqrt{\frac{|\Omega| - l}{K(1 + \delta_N(k+1)+|\Omega|-k)}} \sqrt{\text{MAR} \cdot \text{SNR}} \| v \|_2.
\]

(37)

In fact, by (3) and (4), we have

\[
\| x_{\Omega \setminus S_{k-1}} \|_2 \geq \sqrt{|\Omega| - l} \min_{i \in \Omega} |x_i|
\]

\[
\geq \sqrt{|\Omega| - l} \left( \frac{\sqrt{\text{MAR}} \| x \|_2}{\sqrt{K}} \right)
\]

\[
\geq \sqrt{\frac{|\Omega| - l}{K(1 + \delta_N(k+1)+|\Omega|-k)}} \sqrt{\text{MAR}} \| A x \|_2
\]

\[
\geq \sqrt{\frac{|\Omega| - l}{K(1 + \delta_N(k+1)+|\Omega|-k)}} \sqrt{\text{MAR} \cdot \text{SNR}} \| v \|_2,
\]

where (a) is from (4), (b) is from

\[
\| A x \|_2 = \| A x_{\Omega} \|_2 \leq \sqrt{1 + \delta_N(k+1)+|\Omega|-k} \| x_{\Omega} \|_2
\]

\[
= \sqrt{1 + \delta_N(k+1)+|\Omega|-k} \| x \|_2.
\]

And (c) follows from (3).

By (36) and (37), we have

\[
\beta_1 \geq \frac{(1 - \sqrt{|\Omega|/N + [\delta_N(k+1)+|\Omega|-k]}) \sqrt{\text{MAR} \cdot \text{SNR}} \| v \|_2}{\sqrt{K(1 + \delta_N(k+1)+|\Omega|-k)}},
\]

Thus, by (34), (33) can be guaranteed by

\[
(1 - \sqrt{|\Omega|/N + [\delta_N(k+1)+|\Omega|-k]}) \sqrt{\text{MAR} \cdot \text{SNR}} \| v \|_2
\]

\[
\geq \sqrt{2(1 + \delta_N(k+1)+|\Omega|-k)} \| v \|_2,
\]

which is equivalent to (8). By induction, the theorem holds. \( \Box \)
REFERENCES