Canad. Math. Bull. Vol. **61** (1), 2018 pp. 40–54
http://dx.doi.org/10.4153/CMB-2017-009-6

A Sharp Bound on RIC in Generalized Orthogonal Matching Pursuit

Wengu Chen and Huanmin Ge

Abstract. The generalized orthogonal matching pursuit (gOMP) algorithm has received much attention in recent years as a natural extension of the orthogonal matching pursuit (OMP). It is used to recover sparse signals in compressive sensing. In this paper, a new bound is obtained for the exact reconstruction of every K-sparse signal via the gOMP algorithm in the noiseless case. That is, if the restricted isometry constant (RIC) δ_{NK+1} of the sensing matrix A satisfies

$$\delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N}+1}},$$

then the gOMP can perfectly recover every K-sparse signal x from $y = Ax$. Furthermore, the bound is proved to be sharp. In the noisy case, the above bound on RIC combining with an extra condition on the minimum magnitude of the nonzero components of K-sparse signals can guarantee that the gOMP selects all of the support indices of the K-sparse signals.

1 Introduction

It is well known that compressive sensing acquires signals at a rate greatly below Nyquist rate. It has attracted growing attention in recent years [1, 2, 4, 5, 7, 15, 18]. The main aim of compressive sensing is to reconstruct signals from inaccurate and incomplete measurements from the model

$$y = Ax + e,$$

where $y \in \mathbb{R}^m$ is a measurement vector, the matrix $A \in \mathbb{R}^{m\times n}$ $(m \ll n)$ is a sensing matrix, the vector $x \in \mathbb{R}^n$ is an unknown sparse signal, and $e \in \mathbb{R}^m$ is a measurement error vector. The goal is to recover the signal x based on y and A. In this paper, A_i $(i = 1, 2, \dots, n)$ denotes the i-th column of A and all columns of A are normalized, *i.e.*, $\|A_i\|_2 = 1$ for $i = 1, 2, \dots, n$. Define the support of the vector x by $T = \operatorname{supp}(x) = \{i | x_i \neq 0\}$ and the size of its support by $|T| = |\operatorname{supp}(x)|$. For a signal x, if $|\operatorname{supp}(x)| \leqslant K$, then x is called K-sparse.

Received by the editors November 15, 2016; revised January 11, 2017.
Published electronically April 17, 2017.
This work was supported by the NSF of China (Nos.11271050, 11371183).
AMS subject classification: **65D15**, 65J22, 68W40.
Keywords: sensing matrix, generalized orthogonal matching pursuit, restricted isometry constant, sparse signal.

For the recovery of the K-sparse signal x, the most intuitive approach is to solve the following optimization problem

$$\min_x \|x\|_0 \quad \text{subject to } Ax - y \in \mathcal{B}, \tag{1.1}$$

where $\|x\|_0$ denotes the l_0 norm of x, *i.e.,* the number of nonzero coordinates, $\mathcal{B}$ is a bounded error set, *i.e.,* $\mathcal{B} = \{e \in \mathbb{R}^m \mid \|e\|_2 \leqslant \varepsilon\}$. Particularly, in the noiseless case, $\mathcal{B} = \{0\}$. Unfortunately, it is well known that the above optimization problem is NP-hard. Therefore, researchers seek computationally efficient methods to approximate the sparse signal x, such as l_1 minimization [8], $l_p(0 < p < 1)$ minimization [10], greedy algorithm [26], and so on.

To ensure that the K-sparse solution is unique, we shall need the restricted isometry property (RIP) introduced by Candès and Tao in [8]. A matrix A satisfies the restricted isometry property of order K if there exists a constant $\delta_K \in [0,1)$ such that

$$(1 - \delta_K)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1 + \delta_K)\|x\|_2^2$$

holds for all K-sparse signals x. And the smallest constant δ_K is called the *restricted isometry constant* (*RIC*). Candès and Tao also proposed that if $\delta_{2K} < 1$, the above optimization problem has a unique K-sparse solution [8]. Candès showed that if $\delta_{2K} < \sqrt{2} - 1$, then the above optimization problem (1.1) is equivalent to the l_1 minimization problem in [7]. Up to now, there have been many results improving the bound on the RIC, such as [2, 3, 5, 6, 20].

Recently, a family of iterative greedy algorithms for recovery of sparse signals has attracted significant attention; it includes orthogonal least square (OLS) [11], orthogonal matching pursuit (OMP) [27], generalized orthogonal matching pursuit (gOMP) [28], regularized orthogonal matching pursuit (ROMP) [23], orthogonal multi-matching pursuit (OMMP) [32], stagewise orthogonal matching pursuit (StOMP) [16], subspace pursuit (SP) [13], and compressive sampling matching pursuit (CoSaMP) [22].

Specifically, the OMP algorithm is one of the most effective algorithm, in sparse signal recovery due to its implementation simplicity and competitive recovery performance. In the noiseless case, many efforts have been made to find sufficient conditions based on RIC for OMP to exactly reconstruct every K-sparse signal x within K iterations. Davenport and Wakin demonstrated that OMP can recover exactly the K-sparse signal x under $\delta_{K+1} < 1/(3\sqrt{K})$ [14]. Since then, there have been many papers that improve the condition in [12, 17, 19, 21, 29, 30]. Recently, Mo improved the sufficient condition to $\delta_{K+1} < 1/\sqrt{K+1}$, and proved that this condition is sharp [19]. In the presence of noise, Shen and Li proved that OMP can exactly recover the support of the K-sparse signal x under $\delta_{K+1} < 1/(\sqrt{K} + 3)$ and some assumption on the minimum magnitude of the nonzero elements of x in [25]. Later, these sufficient conditions on RIC upper bound and minimum magnitude of the nonzero elements of K-sparse signal x were improved in [9, 31].

Wang, Kwon, and Shim introduced generalized orthogonal matching pursuit [28], which is a natural extension of OMP. It is well known that the OMP algorithm only selects one correct index at each iteration. However, the gOMP algorithm selects N $(N \geqslant 1)$ indices that contain at least one correct index from the support of x in

Input	measurements $y \in \mathbb{R}^m$, sensing matrix $A \in \mathbb{R}^{m\times n}$, sparse level K, number of indices for each selection N ($N \leqslant K$ and $N \leqslant \frac{m}{K}$).
Initialize	iteration count $k = 0$, residual vector $r^0 = y$, estimated support set $\Lambda^0 = \varnothing$.
While	$\|r^k\|_2 > \epsilon$ and $k < \min\{K, \frac{m}{K}\}$ do $k = k + 1$. (Identification step) Select indices set T^k corresponding to N largest (in magnitude) in $A' r^{k-1}$. (Augmentation step) $\Lambda^k = \Lambda^{k-1} \cup T^k$. (Estimation step) $\widehat{x}_{\Lambda^k} = \arg\min_u \|y - A_{\Lambda^k} u\|_2$. (Residual Update step) $r^k = y - A_{\Lambda^k}\widehat{x}_{\Lambda^k}$.
End Output	the estimated signal $\widehat{x} = \arg\min_{u:\mathrm{supp}(u)=\Lambda^k} \|y - Au\|_2$.

Table 1: The gOMP algorithm

each iteration. Therefore, the number of iterations for the gOMP algorithm is much smaller when compared to the OMP algorithm. Wang, Kwon, and Shim obtained that a sufficient condition

$$\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K} + 3\sqrt{N}}$$

can ensure the reconstruction of any K-sparse signal [28]. Later, Satpathi et al. improved the sufficient condition to $\delta_{NK} < \sqrt{N}/(\sqrt{K} + 2\sqrt{N})$ in [24]. They also refined the bound further to

$$\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K} + \sqrt{N}},$$

which reduces to $\delta_{K+1} < 1/(\sqrt{K} + 1)$ of OMP in [21, 29] for $N = 1$.

Motivated by the mentioned papers, we further investigate the recovery of any K-sparse signal by the gOMP. In this paper, we demonstrate that the condition

$$\delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N} + 1}}$$

is sufficient to perfectly reconstruct any K-sparse signal via the gOMP in the noiseless case. As $N = 1$, the sufficient condition is $\delta_{K+1} < 1/(\sqrt{K+1})$ which is a sharp bound for OMP [19]. Moreover, for any given $K \in \mathbb{N}^+$, we construct a matrix A satisfying

$$\delta_{NK+1} = \frac{1}{\sqrt{\frac{K}{N} + 1}}$$

such that the gOMP may fail to recover some K-sparse signal x. That is, the above bound $\delta_{NK+1} < 1/\sqrt{K/N + 1}$ is sharp for the gOMP.

The frame of the gOMP is listed in Table 1.

The rest of the paper is organized as follows. In Section 2, we give notation and prove some basic lemmas that will be used. The main results and their proofs are given in Section 3.

2 Notation and Preliminaries

Throughout this paper, let Γ be an index set and let Γ^c be the complementary set of Γ. The standard notation $\|x\|_\infty = \max_{i=1,2,\ldots,n}|x_i|$ denotes the l_∞−norm of the vector $x \in \mathbb{R}^n$. $x_\Gamma \in \mathbb{R}^{|\Gamma|}$ denotes the vector composed of components of $x \in \mathbb{R}^n$ indexed by $i \in \Gamma$, *i.e.*, $(x_\Gamma)_i = x_i$ $(i \in \Gamma)$. Define $\widetilde{x}_\Gamma \in \mathbb{R}^n$ by

$$(\widetilde{x}_\Gamma)_i = \begin{cases} x_i & i \in \Gamma, \\ 0 & \text{otherwise}, \end{cases}$$

where $i = 1, 2, \ldots, n$. Denote by A_Γ a submatrix of A corresponding to Γ that consists of all columns with index $i \in \Gamma$ of A and the usual inner product of $\mathbb{R}^n$ with $\langle \cdot, \cdot \rangle$. Let $e_i \in \mathbb{R}^n$ be the i-th coordinate unit vector.

Let α_N^{k+1} be the N-th largest correlation in magnitude between r^k and A_i $(i \in (T \cup \Lambda^k)^c)$, and let β_1^{k+1} be the largest correlation in magnitude between r^k and A_i $(i \in (T - \Lambda^k))$ in the $(k+1)$-th iteration of the gOMP algorithm. Let $W_{k+1} \subseteq (T \cup \Lambda^k)^c$ be the set of N indices that correspond to N largest correlation in magnitude between r^k and A_i $(i \in (T \cup \Lambda^k)^c)$.

Suppose $A^\dagger_{\Lambda^k}$ represents the pseudo-inverse of A_{Λ^k}. When A_{Λ^k} is full column rank $(|\Lambda^k| \leqslant m)$, $A^\dagger_{\Lambda^k} = (A'_{\Lambda^k} A_{\Lambda^k})^{-1} A'_{\Lambda^k}$. Moreover, $P_{\Lambda^k} = A_{\Lambda^k} A^\dagger_{\Lambda^k}$ and $P^\perp_{\Lambda^k} = I - P_{\Lambda^k}$ denote two orthogonal projection operators that project a given vector orthogonally onto the spanned space by all columns of A_{Λ^k} and onto its orthogonal complement, respectively.

First, we recall the following lemma, that is, the monotonicity of the restricted isometry constant in [8,13].

Lemma 2.1 *For any $K_1 \leqslant K_2$, if the sensing matrix A satisfies the RIP of order K_2, then $\delta_{K_1} \leqslant \delta_{K_2}$.*

Next, we show the main lemma that plays the key role during our analysis.

Lemma 2.2 *For any S, $C > 0$, let $t = \pm\frac{\sqrt{S+1}-1}{\sqrt{S}}$ and*

$$t_i = \begin{cases} -\frac{C}{2}(1-t^2) & \langle Ax, Ae_i\rangle \geqslant 0, \\ +\frac{C}{2}(1-t^2) & \langle Ax, Ae_i\rangle < 0, \end{cases}$$

where $i \in W \subseteq \{1, 2, \ldots, n\}$ that is a nonempty subset. Then we have $t^2 < 1$ and

$$\Big\| A\Big(x + \sum_{i\in W} t_i e_i\Big)\Big\|_2^2 - \Big\| A\Big(t^2 x - \sum_{i\in W} t_i e_i\Big)\Big\|_2^2 = (1-t^4)\Big(\langle Ax, Ax\rangle - C\sum_{i\in W}\big|\langle Ax, Ae_i\rangle\big|\Big).$$

Proof For $t = \pm\frac{\sqrt{S+1}-1}{\sqrt{S}}$, we have that

$$t^2 = \frac{(\sqrt{S+1}-1)^2}{S} = \frac{\sqrt{S+1}-1}{\sqrt{S+1}+1} < 1.$$

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

The lemma is established by the following chain of equalities and the definition of t_i $(i \in W)$:

$$
\begin{aligned}
&\|A(x+\sum_{i\in W} t_i e_i)\|_2^2 - \|A(t^2x - \sum_{i\in W} t_i e_i)\|_2^2 \\
&= \langle Ax, Ax\rangle + 2\sum_{i\in W} t_i\langle Ax, Ae_i\rangle + 2\sum_{i,j\in W, i\neq j} t_i t_j\langle Ae_i, Ae_j\rangle + \sum_{i\in W} t_i^2\langle Ae_i, Ae_i\rangle \\
&\quad - \Big(t^4\langle Ax, Ax\rangle - 2t^2\sum_{i\in W} t_i\langle Ax, Ae_i\rangle \\
&\qquad\qquad + 2\sum_{i,j\in W, i\neq j} t_i t_j\langle Ae_i, Ae_j\rangle + \sum_{i\in W} t_i^2\langle Ae_i, Ae_i\rangle\Big) \\
&= (1-t^4)\langle Ax, Ax\rangle + 2(1+t^2)\sum_{i\in W} t_i\langle Ax, Ae_i\rangle \\
&= (1-t^4)\Big(\langle Ax, Ax\rangle - \frac{2}{1-t^2}\sum_{i\in W} |t_i||\langle Ax, Ae_i\rangle|\Big) \\
&= (1-t^4)\Big(\langle Ax, Ax\rangle - \frac{2}{1-t^2}(1-t^2)\frac{C}{2}\sum_{i\in W} |\langle Ax, Ae_i\rangle|\Big) \\
&= (1-t^4)\Big(\langle Ax, Ax\rangle - C\sum_{i\in W} |\langle Ax, Ae_i\rangle|\Big).
\end{aligned}
$$

We have already completed the proof of the Lemma 2.2. ∎

Remark 2.3 Lemma 2.2 is a generalization of [19, Lemma II.1].

3 Main Results

It is well known that if at least one index of N indices selected is correct in every iteration, the gOMP makes a success; *i.e.,* in each iteration, there exists $\beta_1^k > \alpha_N^k$ $(1 \leqslant k \leqslant K)$. The following theorems show that a sufficient condition guarantees the gOMP algorithm success. The proof of these theorems mainly uses Lemmas 2.1 and 2.2. Without loss of generality, we assume $\|x\|_2 = 1$ in the proof of Theorem 3.1 and $\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2 = 1$ in the proof of Theorem 3.3 .

Theorem 3.1 *Suppose x is a K-sparse signal and the restricted isometry constant δ_{K+N} of the sensing matrix A satisfies*

$$\delta_{K+N} < \frac{1}{\sqrt{\frac{K}{N}+1}}. \tag{3.1}$$

Then the gOMP algorithm makes a success in the first iteration.

Remark 3.2 In [28], the authors proved that

$$\delta_{K+N} < \frac{\sqrt{N}}{\sqrt{K}+\sqrt{N}}$$

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

is sufficient to make a success in the first iteration of the gOMP. It is clear that

$$\delta_{K+N} < \frac{\sqrt{N}}{\sqrt{K}+\sqrt{N}} < \frac{1}{\sqrt{\frac{K}{N}+1}};$$

i.e., the sufficient condition (3.1) is weaker than that in [28].

Proof In the first iteration, by the definition of α_N^1, it satisfies

$$\alpha_N^1 = \min\{\,|\langle Ae_i, Ax\rangle|\,|i \in W_1\} \leqslant \frac{\sum_{i\in W_1}|\langle Ae_i, Ax\rangle|}{N}, \tag{3.2}$$

where $W_1 \subseteq T^c$.

For β_1^1, which is the largest correlation in magnitude in $A_T' Ax$, we have

$$\begin{aligned}\langle Ax, Ax\rangle &= \langle A\sum_{i\in T} x_i e_i, Ax\rangle = \sum_{i\in T} x_i\langle Ae_i, Ax\rangle \leqslant \sum_{i\in T}|x_i||\langle Ae_i, Ax\rangle| \\ &\leqslant \beta_1^1\|x\|_1 \leqslant \beta_1^1\sqrt{K}\|x\|_2 = \beta_1^1\sqrt{K}.\end{aligned} \tag{3.3}$$

Let

$$t = -\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}}} \quad \text{and} \quad t_i = \begin{cases} -\frac{\sqrt{K}}{2N}(1-t^2) & \langle Ax, Ae_i\rangle \geqslant 0,\\ +\frac{\sqrt{K}}{2N}(1-t^2) & \langle Ax, Ae_i\rangle < 0,\end{cases}$$

where $i \in W_1 \subseteq T^c$ with $|W_1| = N$. Then we have that

$$t^2 = \Big(\sqrt{\frac{K}{N}+1}-1\Big)\Big/\Big(\sqrt{\frac{K}{N}+1}+1\Big) < 1,$$

$$\sum_{i\in W_1} t_i^2 = \Big(\frac{\sqrt{K}}{2N}(1-t^2)\Big)^2 N = \frac{K}{4N}\Big(1-\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}+1}+1}\Big)^2 = t^2. \tag{3.4}$$

By (3.2), (3.3), and Lemma 2.2, we obtain

$$\begin{aligned}(1-t^4)\sqrt{K}(\beta_1^1-\alpha_N^1) &\geqslant (1-t^4)\Big(\langle Ax, Ax\rangle - \sqrt{K}\frac{\sum_{i\in W_1}|\langle Ae_i, Ax\rangle|}{N}\Big)\\ &= \Big\|A\Big(x+\sum_{i\in W_1} t_i e_i\Big)\Big\|_2^2 - \Big\|A\Big(t^2 x - \sum_{i\in W_1} t_i e_i\Big)\Big\|_2^2.\end{aligned}$$

Because the sensing matrix A satisfies the RIP of order $K+N$ with δ_{K+N}, $\|x\|_2 = 1$ with $\text{supp}(x) \subseteq T$, $W_1 \subseteq T^c$, it follows from (3.4) that

$$\begin{aligned}&\Big\|A\Big(x+\sum_{i\in W_1} t_i e_i\Big)\Big\|_2^2 - \Big\|A\Big(t^2 x - \sum_{i\in W_1} t_i e_i\Big)\Big\|_2^2\\ &\geqslant (1-\delta_{K+N})\Big(\|x+\sum_{i\in W_1} t_i e_i\|_2^2\Big) - (1+\delta_{K+N})\Big(\|t^2 x - \sum_{i\in W_1} t_i e_i\|_2^2\Big)\\ &= (1-\delta_{K+N})\Big(\|x\|_2^2 + \sum_{i\in W_1} t_i^2\Big) - (1+\delta_{K+N})\Big(t^4\|x\|_2^2 + \sum_{i\in W_1} t_i^2\Big)\\ &= (1-\delta_{K+N})(1+t^2) - (1+\delta_{K+N})(t^4+t^2)\\ &= (1-t^4) - \delta_{K+N}(1+t^2)^2 \quad = (1+t^2)^2\Big(\frac{1-t^2}{1+t^2} - \delta_{K+N}\Big).\end{aligned}$$

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

It follows from the definition of t that

$$\frac{1-t^2}{1+t^2}=\frac{1-\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}+1}+1}}{1+\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}+1}+1}}=\frac{1}{\sqrt{\frac{K}{N}+1}}.$$

Therefore, by the condition $\delta_{K+N}<\frac{1}{\sqrt{\frac{K}{N}+1}}$, we obtain

$$\begin{aligned}(1-t^4)\sqrt{K}(\beta_1^1-\alpha_N^1)&\geqslant(1+t^2)^2\Big(\frac{1-t^2}{1+t^2}-\delta_{K+N}\Big)\\&\geqslant(1+t^2)^2\Big(\frac{1}{\sqrt{\frac{K}{N}+1}}-\delta_{K+N}\Big)>0,\end{aligned}$$

i.e., $\beta_1^1>\alpha_N^1$, which represents the gOMP selects at least one index from the support T.

As mentioned, if $\delta_{K+N}<1/(\sqrt{\frac{K}{N}+1})$, then the gOMP algorithm makes a success in the first iteration. ∎

Theorem 3.3 *If the gOMP algorithm has performed k iterations successfully, where $1\leqslant k<K$ and the sensing matrix A satisfies the RIP of order $NK+1$ with RIC δ_{NK+1} fulfilling*

$$\delta_{NK+1}<\frac{1}{\sqrt{\frac{K}{N}+1}},$$

then in the $(k+1)$-th iteration, the gOMP will succeed.

Proof For the gOMP algorithm, $r^k=P_{\Lambda^k}^{\perp}y$ is orthogonal to each column of A_{Λ^k}. Then

$$\begin{aligned}r^k&=P_{\Lambda^k}^{\perp}y=P_{\Lambda^k}^{\perp}A_Tx_T=P_{\Lambda^k}^{\perp}(A_{T-\Lambda^k}x_{T-\Lambda^k}+A_{T\cap\Lambda^k}x_{T\cap\Lambda^k})\\&=P_{\Lambda^k}^{\perp}A_{T-\Lambda^k}x_{T-\Lambda^k}=A_{T-\Lambda^k}x_{T-\Lambda^k}-P_{\Lambda^k}A_{T-\Lambda^k}x_{T-\Lambda^k}\\&=A_{T-\Lambda^k}x_{T-\Lambda^k}-A_{\Lambda^k}z_{\Lambda^k}=A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k},\end{aligned}$$

where we used the fact that $P_{\Lambda^k}A_{T-\Lambda^k}x_{T-\Lambda^k}\in\operatorname{span}(A_{\Lambda^k})$, so $P_{\Lambda^k}A_{T-\Lambda^k}x_{T-\Lambda^k}$ can be written as $A_{\Lambda^k}z_{\Lambda^k}$ for some $z_{\Lambda^k}\in\mathbb{R}^{|\Lambda^k|}$ and $\omega_{T\cup\Lambda^k}$ is given by

$$\omega_{T\cup\Lambda^k}=\begin{pmatrix}x_{T-\Lambda^k}\\-z_{\Lambda^k}\end{pmatrix}.$$

By the definition of α_N^{k+1} and β_1^{k+1}, we have that

$$\begin{aligned}(3.5)\quad \alpha_N^{k+1}&=\min\{|\langle Ae_i,r^k\rangle|| i\in W_{k+1},W_{k+1}\subseteq(T\cup\Lambda^k)^c\}\leqslant\frac{\sum_{i\in W_{k+1}}|\langle Ae_i,r^k\rangle|}{N}\\&=\frac{\sum_{i\in W_{k+1}}|\langle Ae_i,A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\rangle|}{N}=\frac{\sum_{i\in W_{k+1}}|\langle Ae_i,A\widetilde{\omega}_{T\cup\Lambda^k}\rangle|}{N}\end{aligned}$$

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

and

$$\begin{aligned}(3.6)\quad \beta_1^{k+1} &= \|A'_{T-\Lambda^k} r^k\|_\infty = \|[A_{T-\Lambda^k}\ A_{T\cap\Lambda^k}]' A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty \\ &= \|A'_T A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty \\ &= \|[A_T\ A_{\Lambda^k-T}]' A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty = \|A'_{T\cup\Lambda^k} A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty.\end{aligned}$$

Notice the fact that

$$\begin{aligned}(3.7)\qquad \|A'_T A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty &\geqslant \frac{1}{\sqrt{K}}\|A'_T A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_2 \\ &= \frac{1}{\sqrt{K}}\|A'_{T\cup\Lambda^k} A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_2.\end{aligned}$$

By the hypothesis of $\|\omega_{T\cup\Lambda^k}\|_2 = 1$, (3.6), and (3.7), it follows that

$$\begin{aligned}(3.8)\qquad \langle A\widetilde{\omega}_{T\cup\Lambda^k}, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle &= \langle A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}, A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\rangle \\ &= \langle A'_{T\cup\Lambda^k} A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}, \omega_{T\cup\Lambda^k}\rangle \\ &\leqslant \|A'_{T\cup\Lambda^k} A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_2\|\omega_{T\cup\Lambda^k}\|_2 \\ &\leqslant \sqrt{K}\beta_1^{k+1}\|\omega_{T\cup\Lambda^k}\|_2 = \sqrt{K}\beta_1^{k+1}.\end{aligned}$$

As in the proof of Theorem 3.1, let

$$t = -\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}}} \quad \text{and} \quad t_i = \begin{cases} -\frac{\sqrt{K}}{2N}(1-t^2) & \langle Ax, Ae_i\rangle \geqslant 0, \\ +\frac{\sqrt{K}}{2N}(1-t^2) & \langle Ax, Ae_i\rangle < 0, \end{cases}$$

where $i \in W_{k+1} \subseteq (\Lambda^k \cup T)^c$. By (3.5), (3.8), and Lemma 2.2, we obtain

$$\begin{aligned}&(1-t^4)\sqrt{K}(\beta_1^{k+1} - \alpha_N^{k+1}) \\ &\quad\geqslant (1-t^4)\Big(\langle A\widetilde{\omega}_{T\cup\Lambda^k}, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle - \sqrt{K}\frac{\sum_{i\in W_{k+1}}|\langle Ae_i, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle|}{N}\Big) \\ &\quad= \|A(\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i)\|_2^2 - \Big\|A\Big(t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i\Big)\Big\|_2^2.\end{aligned}$$

Let $l = |T\cap\Lambda^k|$. Then $k \leqslant l \leqslant K$ and $Nk + K - l + N \leqslant NK + 1$. Since A satisfies RIP of order $NK+1$ with δ_{NK+1}, $\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2 = 1$ with $\operatorname{supp}(\widetilde{\omega}_{T\cup\Lambda^k}) \subseteq T\cup\Lambda^k$, $W_{k+1} \subseteq (T\cup\Lambda^k)^c$, it follows from Lemma 2.1 that

$$\begin{aligned}&\|A(\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i)\|_2^2 - \|A(t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i)\|_2^2 \\ &\quad\geqslant (1-\delta_{Nk+K-l+N})\Big(\Big\|\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i\Big\|_2^2\Big) \\ &\qquad - (1+\delta_{Nk+K-l+N})\Big(\|t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i\|_2^2\Big) \\ &\quad= (1-\delta_{Nk+K-l+N})\Big(\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2 + \sum_{i\in W_{k+1}} t_i^2\Big) \\ &\qquad - (1+\delta_{Nk+K-l+N})\Big(t^4\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2 + \sum_{i\in W_{k+1}} t_i^2\Big)\end{aligned}$$

$$= (1 - \delta_{Nk+K-l+N})(1+t^2) - (1+\delta_{Nk+K-l+N})(t^4+t^2)$$
$$= (1+t^2)^2\Big(\frac{1-t^2}{1+t^2} - \delta_{Nk+K-l+N}\Big) \geqslant (1+t^2)^2\Big(\frac{1-t^2}{1+t^2} - \delta_{NK+1}\Big).$$

Since

$$\frac{1-t^2}{1+t^2} = \frac{1}{\sqrt{\frac{K}{N}+1}}$$

and the condition $\delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N}+1}}$, we obtain

$$(1-t^4)\sqrt{K}(\beta_1^{k+1} - \alpha_N^{k+1}) \geqslant (1+t^2)^2\Big(\frac{1-t^2}{1+t^2} - \delta_{NK+1}\Big)$$
$$\geqslant (1+t^2)^2\Big(\frac{1}{\sqrt{\frac{K}{N}+1}} - \delta_{NK+1}\Big) > 0,$$

i.e., $\beta_1^{k+1} > \alpha_N^{k+1}$, which ensures that the set Λ^{k+1} contains at least one correct index in the $(k+1)$-th iteration of the gOMP algorithm.

As mentioned, we have completed the proof of the theorem. ■

Now combining the condition for success in the first iteration in Theorem 3.1 with that in non-initial iterations in Theorem 3.3, we obtain overall sufficient condition of the gOMP algorithm guaranteeing the perfect recovery of K-sparse signals in the following theorem.

Theorem 3.4 *Suppose x is a K-sparse signal and the sensing matrix A satisfies RIP of order $KN+1$ with the RIC δ_{NK+1} fulfilling*

$$\delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N}+1}}.$$

Then the gOMP algorithm can recover the signal x exactly.

Proof For $N \geqslant 1$, $K \geqslant 1$, and $N \leqslant \min\{K, \frac{m}{K}\}$, $K+N \leqslant NK+1$. It follows Lemma 2.1 that

$$\delta_{K+N} \leqslant \delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N}+1}}.$$

By Theorems 3.1 and 3.3, the gOMP algorithm can recover perfectly any K-sparse signals under the sufficient condition $\delta_{NK+1} < 1/\sqrt{K/N+1}$ from $y = Ax$. ■

Remark 3.5 The condition $\delta_{NK+1} < 1/\sqrt{K/N+1}$ is weaker than the sufficient condition $\delta_{NK+1} < \sqrt{N}/(\sqrt{K}+\sqrt{N})$ in [31].

Remark 3.6 If $N=1$, this sufficient condition is consistent with the sharp condition $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ of OMP in [19].

In the following theorem, we show that the proposed bound $\delta_{NK+1} < 1/\sqrt{K/N+1}$ is optimal.

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

Theorem 3.7 *For any given $K \in \mathbb{N}^+$, there are a K-sparse signal x and a matrix A satisfying*

$$\delta_{NK+1} = \frac{1}{\sqrt{\frac{K}{N}+1}}$$

such that the gOMP may fail.

Proof For any given positive integer K, let $A \in \mathbb{R}^{(NK+1)\times(NK+1)}$ be

$$A = \begin{pmatrix} \sqrt{\frac{K}{K+N}} I_K & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} \frac{1}{b} & \cdots & \frac{1}{b} \\ \vdots & \ddots & \vdots \\ \frac{1}{b} & \cdots & \frac{1}{b} \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} & I_{NK+1-N-K} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} \\ \begin{matrix} \frac{1}{b} & \cdots & \frac{1}{b} \\ \vdots & \ddots & \vdots \\ \frac{1}{b} & \cdots & \frac{1}{b} \end{matrix} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} & I_N \end{pmatrix},$$

where $b = \sqrt{K(K+N)}$. Then we have that

$$A'A = \begin{pmatrix} \frac{K}{K+N} I_K & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} \frac{1}{K+N} & \cdots & \frac{1}{K+N} \\ \vdots & \ddots & \vdots \\ \frac{1}{K+N} & \cdots & \frac{1}{K+N} \end{matrix} \\ \begin{matrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{matrix} & I_{NK+1-N-K} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \vdots & 0 \\ 0 & \cdots & 0 \end{matrix} \\ \begin{matrix} \frac{1}{K+N} & \cdots & \frac{1}{K+N} \\ \vdots & \ddots & \vdots \\ \frac{1}{K+N} & \cdots & \frac{1}{K+N} \end{matrix} & \begin{matrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} 1+\frac{1}{K+N} & \cdots & \frac{1}{K+N} \\ \vdots & \ddots & \vdots \\ \frac{1}{K+N} & \cdots & 1+\frac{1}{K+N} \end{matrix} \end{pmatrix}.$$

Moreover, by direct calculation, we obtain that

$$A'A - \lambda I = (1-\lambda)^{NK-K}\Big(\frac{K}{K+N} - \lambda\Big)^{K-1}\Big(\lambda^2 - 2\lambda + \frac{K}{K+N}\Big).$$

It is clear that $\frac{K}{K+N}$ and 1 are eigenvalues of $A'A$ with multiplicity of $K-1$ and $NK-K$ respectively; $1 \pm 1/\sqrt{K/N+1}$ also are eigenvalues of $A'A$. Therefore, we have

$$\delta_{NK+1} = \frac{1}{\sqrt{\frac{K}{N}+1}}.$$

Consider K-sparse signal $x = (1,1,\dots,1,0\dots,0)' \in \mathbb{R}^{NK+1}$, *i.e.*, $T = \operatorname{supp}(x) = \{1,2,\dots,K\}$. As $i \in T$, we have

$$|\langle Ae_i, y\rangle| = |\langle Ae_i, Ax\rangle| = |\langle A'Ae_i, x\rangle| = \frac{K}{K+N}.$$

For $i \in \{K+1,\dots,NK+1-N\}$, it follows immediately that

$$|\langle Ae_i, y\rangle| = |\langle Ae_i, Ax\rangle| = |\langle A'Ae_i, x\rangle| = 0.$$

If $i \in \{NK+2-N,\dots,NK+1\}$, we have

$$|\langle Ae_i, y\rangle| = |\langle Ae_i, Ax\rangle| = |\langle A'Ae_i, x\rangle| = \frac{K}{K+N}.$$

Therefore, we have $\beta_1^1 = \frac{K}{K+N}$ and $\alpha_N^1 = \frac{K}{K+N}$ by the definitions of β_1^1 and α_N^1, that is, $\beta_1^1 = \alpha_N^1$. This implies the gOMP may fail to identify at least one correct index in the first iteration. So the gOMP algorithm may fail for the given matrix A and the K-sparse signal x. ∎

Finally, we show a sufficient condition guarantees exact support identification by the gOMP algorithm from $y = Ax + e$. This sufficient condition is in terms of the RIC δ_{NK+1} and the minimum magnitude of the nonzero entries of K-sparse signal x. Here, we only consider l_2 bounded noise, *i.e.*, $\|e\|_2 \leqslant \varepsilon$.

Theorem 3.8 *Suppose $\|e\|_2 \leqslant \varepsilon$ and the sensing matrix A satisfies*

$$\delta_{NK+1} < \frac{1}{\sqrt{\frac{K}{N}+1}}. \tag{3.9}$$

Moreover, assume all the nonzero components x_i satisfy

$$|x_i| > \frac{2\sqrt{K}\varepsilon/\sqrt{\frac{K}{N}+1}}{1/\sqrt{\frac{K}{N}+1}-\delta_{NK+1}}. \tag{3.10}$$

Then there exists an integer $1 \leqslant k_0 \leqslant K$ such that the gOMP algorithm with the stopping rule $\|r^k\|_2 \leqslant \varepsilon$ recovers the correct support of any K-sparse signals x, that is, $T \subseteq \Lambda^{k_0}$. Meanwhile, for $\widehat{x} = \arg\min_{x:\operatorname{supp}(x)=\Lambda^{k_0}} \|y - Ax\|_2$, there are

$$\|x - \widehat{x}\|_2 \leqslant \frac{\varepsilon}{\sqrt{1-\delta_{NK+1}}}.$$

Proof Suppose that the gOMP performed k iterations successfully. Consider the $(k+1)$-th iteration. First, we observe that

$$r^k = P_{\Lambda^k}^{\perp} y = P_{\Lambda^k}^{\perp} A_T x_T + P_{\Lambda^k}^{\perp} e = A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k} + (I - P_{\Lambda^k})e = A\widetilde{\omega}_{T\cup\Lambda^k} + (I - P_{\Lambda^k})e$$

for some $\omega_{T\cup\Lambda^k}$ as in the proof of Theorem 3.3. Consider the following two cases to prove the theorem.

Case 1: $T - \Lambda^k = \varnothing$. In this case, there is $T \subseteq \Lambda^k$. Then the correct support T of the original K-sparse signal x has already been selected.

Case 2: $T - \Lambda^k \neq \varnothing$, i.e., $|T - \Lambda^k| \geqslant 1$. By the definitions of α_N^{k+1} and β_1^{k+1}, we obtain that

$$\begin{aligned}\alpha_N^{k+1} &= \min\{|\langle Ae_i, r^k\rangle| | i \in W_{k+1}, W_{k+1} \subseteq (T\cup\Lambda^k)^c\} \\ &\leqslant \frac{\sum_{i\in W_{k+1}} |\langle Ae_i, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle| + \sum_{i\in W_{k+1}} |\langle Ae_i, (I-P_{\Lambda^k})e\rangle|}{N}\end{aligned} \tag{3.11}$$

https://doi.org/10.4153/CMB-2017-009-6 Published online by Cambridge University Press

and

$$\beta_1^{k+1} = \|A'_{T-\Lambda^k} r^k\|_\infty = \|A'_T r^k\|_\infty = \|A'_{T\cup\Lambda^k} r^k\|_\infty \geqslant \|A'_{T\cup\Lambda^k} A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_\infty - \|A'_{T\cup\Lambda^k}(I-P_{\Lambda^k})e\|_\infty. \tag{3.12}$$

Let

$$t = -\frac{\sqrt{\frac{K}{N}+1}-1}{\sqrt{\frac{K}{N}}} \quad \text{and} \quad t_i = \begin{cases} -\frac{\sqrt{K}}{2N}(1-t^2)\|\omega_{T\cup\Lambda^k}\|_2 & \langle Ax, Ae_i\rangle \geqslant 0, \\ +\frac{\sqrt{K}}{2N}(1-t^2)\|\omega_{T\cup\Lambda^k}\|_2 & \langle Ax, Ae_i\rangle < 0, \end{cases}$$

where $i \in W_{k+1} \subseteq (\Lambda^k \cup T)^c$. Then we have $\sum_{i\in W_{k+1}} t_i^2 = t^2\|\omega_{T\cup\Lambda^k}\|_2^2$. It follows from (3.7), (3.8), (3.11), and (3.12) that

$$\begin{aligned}
&(1-t^4)\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2(\beta_1^{k+1}-\alpha_N^{k+1}) \\
&\geqslant (1-t^4)\Bigg(\langle A\widetilde{\omega}_{T\cup\Lambda^k}, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle - \sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2\|A'_{T\cup\Lambda^k}(I-P_{\Lambda^k})e\|_\infty \\
&\quad - \frac{\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2(\sum_{i\in W_{k+1}}|\langle Ae_i, A\widetilde{\omega}_{T\cup\Lambda^k}\rangle| + \sum_{i\in W_{k+1}}|\langle Ae_i, (I-P_{\Lambda^k})e\rangle|)}{N}\Bigg) \\
&= \|A(\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i)\|_2^2 - \|A(t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i)\|_2^2 \\
&\quad - (1-t^4)\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2\Big(\|A'_{T\cup\Lambda^k}(I-P_{\Lambda^k})e\|_\infty + \frac{\sum_{i\in W_{k+1}}|\langle Ae_i, (I-P_{\Lambda^k})e\rangle|}{N}\Big).
\end{aligned}$$

As in the proof of Theorem 3.3, $l = |T\cap\Lambda^k|$; then $Nk+K-l+N \leqslant NK+1$. Because A satisfies RIP of order $NK+1$ with δ_{NK+1}, $\operatorname{supp}(\widetilde{\omega}_{T\cup\Lambda^k}) \subseteq T\cup\Lambda^k$, $W_{k+1} \subseteq (T\cup\Lambda^k)^c$, it follows from $\sum_{i\in W_{k+1}} t_i^2 = t^2\|\omega_{T\cup\Lambda^k}\|_2^2$ and Lemma 2.1 that

$$\begin{aligned}
&\Big\| A\Big(\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i\Big)\Big\|_2^2 - \Big\| A\Big(t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i\Big)\Big\|_2^2 \\
&\quad \geqslant (1-\delta_{Nk+K-l+N})\Big(\|\widetilde{\omega}_{T\cup\Lambda^k} + \sum_{i\in W_{k+1}} t_i e_i\|_2^2\Big) \\
&\qquad - (1+\delta_{Nk+K-l+N})\Big(\|t^2\widetilde{\omega}_{T\cup\Lambda^k} - \sum_{i\in W_{k+1}} t_i e_i\|_2^2\Big) \\
&\quad = (1-\delta_{Nk+K-l+N})\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2(1+t^2) - (1+\delta_{Nk+K-l+N})\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2(t^4+t^2) \\
&\quad = (1-t^4)\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2 - \delta_{Nk+K-l+N}\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2(1+t^2)^2 \\
&\quad = (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2\Big(\frac{1-t^2}{1+t^2} - \delta_{Nk+K-l+N}\Big) \\
&\quad \geqslant (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2\Big(\frac{1-t^2}{1+t^2} - \delta_{NK+1}\Big).
\end{aligned}$$

Moreover, notice the fact that

$$\|A'(I-P_{\Lambda^k})e\|_\infty = \max_i |\langle Ae_i, (I-P_{\Lambda^k})e\rangle| \leqslant \|Ae_i\|_2\|(I-P_{\Lambda^k})e\|_2 \leqslant \|e\|_2 \leqslant \varepsilon.$$

By the above three inequalities, (3.9), and (3.10), it follows that

$$
\begin{aligned}
&(1-t^4)\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2(\beta_1^{k+1}-\alpha_N^{k+1})\\
&\quad\geqslant (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2\Big(\frac{1-t^2}{1+t^2}-\delta_{NK+1}\Big)-(1-t^4)\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2\\
&\qquad\Big(\|A'_{T\cup\Lambda^k}(I-P_{\Lambda^k})e\|_\infty+\frac{\sum_{i\in W_{k+1}}|\langle Ae_i,(I-P_{\Lambda^k})e\rangle|}{N}\Big)\\
&\quad\geqslant (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2^2\Big(\frac{1-t^2}{1+t^2}-\delta_{NK+1}\Big)-(1-t^4)\sqrt{K}\|\omega_{T\cup\Lambda^k}\|_2\Big(\varepsilon+\frac{N\varepsilon}{N}\Big)\\
&\quad= (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2\Big(\Big(\frac{1-t^2}{1+t^2}-\delta_{NK+1}\Big)\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2-2\sqrt{K}\varepsilon\frac{1-t^2}{1+t^2}\Big)\\
&\quad\geqslant (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2\Big(\Big(\frac{1-t^2}{1+t^2}-\delta_{NK+1}\Big)\|x_{T-\Lambda_k}\|_2-2\sqrt{K}\varepsilon\frac{1-t^2}{1+t^2}\Big)\\
&\quad\geqslant (1+t^2)^2\|\widetilde{\omega}_{T\cup\Lambda^k}\|_2\Big(\Big(\frac{1}{\sqrt{\frac{K}{N}+1}}-\delta_{NK+1}\Big)|T-\Lambda_k|\min_{i\in T-\Lambda_k}|x_i|-\frac{2\sqrt{K}\varepsilon}{\sqrt{\frac{K}{N}+1}}\Big)>0,
\end{aligned}
$$

i.e., $\beta_1^{k+1}>\alpha_N^{k+1}$, which guarantees at least one index selected from the correct support in the $(k+1)$-th iteration.

Now, we shall turn to show that the gOMP algorithm exactly stops under the stopping rule $\|r^k\|\leqslant\varepsilon$ when all the correct indices are selected. That is, we shall prove that there exists $0\leqslant k_0\leqslant K$ such that $T\subseteq\Lambda^{k_0}$. First, assume that $T-\Lambda^k=\varnothing$; then $T\subseteq\Lambda^k$ and $(I-P_{\Lambda^k})Ax=0$. Therefore, it follows that $\|r^k\|_2=\|(I-P_{\Lambda^k})Ax+(I-P_{\Lambda^k})e\|_2=\|(I-P_{\Lambda^k})e\|_2\leqslant\|e\|_2\leqslant\varepsilon$. Second, assume that $T-\Lambda^k\neq\varnothing$; then it follows from the definition of the RIP and (3.10) that

$$
\begin{aligned}
\|r^k\|_2&\geqslant\|A_{T\cup\Lambda^k}\omega_{T\cup\Lambda^k}\|_2-\|(I-P_{\Lambda^k})e\|_2\geqslant\sqrt{1-\delta_{|T\cup\Lambda^k|}}\|x_{T-\Lambda^k}\|_2-\|e\|_2\\
&\geqslant\sqrt{1-\delta_{NK+1}}\min_{i\in T}|x_i|_2-\varepsilon\geqslant(1-\delta_{NK+1})\frac{2\sqrt{K}\varepsilon/\sqrt{\frac{K}{N}+1}}{1/\sqrt{\frac{K}{N}+1}-\delta_{NK+1}}-\varepsilon>\varepsilon,
\end{aligned}
$$

which implies that the gOMP algorithm does not terminate until all the correct indices are selected. By all of the above, there exists $0\leqslant k_0\leqslant K$ such that $T\subseteq\Lambda^{k_0}$.

It remains to estimate an upper bound of $\|x-\widehat{x}\|_2$, where

$$
\widehat{x}=\arg\min_{x:\operatorname{supp}(x)=\Lambda^{k_0}}\|y-Ax\|_2.
$$

Then we have

$$
\begin{aligned}
\|x-\widehat{x}\|_2&\leqslant\frac{1}{\sqrt{1-\delta_{Nk_0}}}\|A(x-\widehat{x})\|_2\leqslant\frac{1}{\sqrt{1-\delta_{Nk_0}}}\|Ax-A_{\Lambda^{k_0}}A^\dagger_{\Lambda^{k_0}}y\|_2\\
&=\frac{1}{\sqrt{1-\delta_{Nk_0}}}\|Ax-A_{\Lambda^{k_0}}A^\dagger_{\Lambda^{k_0}}Ax-A_{\Lambda^{k_0}}A^\dagger_{\Lambda^{k_0}}e\|_2\\
&=\frac{\|P_{\Lambda^{k_0}}e\|_2}{\sqrt{1-\delta_{Nk_0}}}\leqslant\frac{\varepsilon}{\sqrt{1-\delta_{NK+1}}},
\end{aligned}
$$

where we apply the fact $T = \operatorname{supp}(x) \subseteq \Lambda^{k_0}$, $Ax - A_{\Lambda^{k_0}} A^{\dagger}_{\Lambda^{k_0}} Ax = 0$ and $\|P_{\Lambda^{k_0}} e\|_2 \leqslant \|e\|_2 \leqslant \varepsilon$. We complete the proof of the theorem. ■

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), no. 3, 253–263. http://dx.doi.org/10.1007/s00365-007-9003-x

[2] T. T. Cai, L. Wang, and G. W. Xu, *Shifting inequality and recovery of sparse signals.* IEEE Trans. Signal Process. **58**(2010), 1300–1308. http://dx.doi.org/10.1109/TSP.2009.2034936

[3] ———, *New bounds for restricted isometry constants.* IEEE Trans. Inf. Theory. **56**(2010), no. 9, 4388–4394. http://dx.doi.org/10.1109/TIT.2010.2054730

[4] T. T. Cai and A. R. Zhang, *Compressed sensing and affine rank minimization under restricted isometry.* IEEE Trans. Signal Process. **61**(2013), no. 13, 3279–3290. http://dx.doi.org/10.1109/TSP.2013.2259164

[5] ———, *Sparse representation of a polytope and recovery of sparse signals and low-rank matrices.* IEEE Trans. Inf. Theory. **60**(2014), no. 1, 122–132. http://dx.doi.org/10.1109/TIT.2013.2288639

[6] ———, *Sharp RIP bound for sparse signal and low rank matrix recovery.* Appl. Comput. Harmon. Anal. **35**(2013), 74–93. http://dx.doi.org/10.1016/j.acha.2012.07.010

[7] E. J. Candès, *The restricted isometry property and its implications for compressed sensing.* C. R. Math. Acad. Sci. Paris **346**(2008), no. 9–10, 589–592. http://dx.doi.org/10.1016/j.crma.2008.03.014

[8] E. J. Candès and T. Tao, *Decoding by linear programming.* IEEE Trans. Inf. Theory **51**(2005), no. 12, 4203–4215. http://dx.doi.org/10.1109/TIT.2005.858979

[9] L.-H. Chang and J.-Y. Wu, *An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit.* IEEE Trans. Inf. Theory. **60**(2014), no. 9, 5702–5715. http://dx.doi.org/10.1109/TIT.2014.2338314

[10] R. Chartrand, *Exact reconstruction of sparse signals via nonconvex minimization.* IEEE Signal Proc. Lett. **14**(2007), no. 10, 707–710.

[11] S. Chen, S. A. Billings, and W. Luo, *Orthogonal least squares methods and their application to non-linear system identification.* Internat. J. Control **50**(1989), no. 5, 1873–1896. http://dx.doi.org/10.1080/00207178908953472

[12] W. Dan and R. H. Wang, *Robustness of orthogonal matching pursuit under restricted isometry property.* Sci. China Math. **57**(2014), no. 3, 627–634. http://dx.doi.org/10.1007/s11425-013-4655-4

[13] W. Dai and O. Milenkovic, *Subspace pursuit for compressive sensing signal reconstruction.* IEEE Trans. Inf. Theory **55**(2009), no. 5, 2230–2249. http://dx.doi.org/10.1109/TIT.2009.2016006

[14] M. A. Davenport and M. B. Wakin, *Analysis of orthogonal matching pursuit using the restricted isometry property.* IEEE Trans. Inf. Theory **56**(2010), 4395–4401. http://dx.doi.org/10.1109/TIT.2010.2054653

[15] D. L. Donoho, *Compressed sensing.* IEEE Trans. Inf. Theory **56**(2006), no. 4, 1289–1306. http://dx.doi.org/10.1109/TIT.2006.871582

[16] D. L. Donoho, I. Drori, Y. Tsaig, and J. L. Starck, *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit.* IEEE Trans. Inf. Theory **58**(2012), no. 2, 1094–1121. http://dx.doi.org/10.1109/TIT.2011.2173241

[17] S. S. Huang and J. B. Zhu, *Recovery of sparse signals using OMP and its variants: Convergence analysis based on RIP.* Inverse Problems. **27**(2011), no. 3, 035003, 14/ http://dx.doi.org/10.1088/0266-5611/27/3/035003

[18] E. Liu and V. N. Temlyakov, *The orthogonal super greedy algorithm and applications in compressed sensing.* IEEE Trans. Inf. Theory **58**(2012), no. 4, 2040–2047. http://dx.doi.org/10.1109/TIT.2011.2177632

[19] Q. Mo, *A sharp restricted isometry constant bound of orthogonal matching pursuit.* arxiv:1501.01708v1

[20] Q. Mo and S. Li, *New bounds on the restricted isometry constant* δ_{2k}. Appl. Comput. Harmon. Anal. **31**(2011), no. 3, 460–468. http://dx.doi.org/10.1016/j.acha.2011.04.005

[21] Q. Mo and Y. Shen, *A remark on the restricted isometry property in orthogonal matching pursuit.* IEEE Trans. Inform. Theory. **58**(2012), no. 6, 3654–3656. http://dx.doi.org/10.1109/TIT.2012.2185923

[22] D. Needell and J. A. Troop, *CoSaMP: Iterative signal recovery from incomplete and inaccurate samples.* Appl. Comput. Harmon. Anal. **26**(2009), no. 3, 301–321. http://dx.doi.org/10.1016/j.acha.2008.07.002

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

[24] S. Satpathi, R. L. Das, and M. Chakraborty, *Improving the bound on the RIP constant in generalized orthogonal matching pursuit.* IEEE Signal Proc. Lett. **20**(2013), no. 20, 1074–1077.

[25] Y. Shen and S. Li, *Sparse signals recovery from noise measurements by orthogonal matching pursuit.* Inverse Probl. Imaging. **9**(2015), no. 1, 231–238. http://dx.doi.org/10.3934/ipi.2015.9.231

[26] J. A. Tropp, *Greed is good: Algorithmic results for sparse approximation.* IEEE Trans. Inf. Theory. **50**(2004), 2231–2242. http://dx.doi.org/10.1109/TIT.2004.834793

[27] J. A. Tropp and A. C. Gilbert, *Signal recovery from random measurements via orthogonal matching pursuit.* IEEE Trans. Inf. Theory. **53**(2007), no. 12, 4655–4666. http://dx.doi.org/10.1109/TIT.2007.909108

[28] J. Wang, S. Kwon, and B. Shim, *Generalized orthogonal matching pursuit.* IEEE Trans. Signal Process. **60**(2012), no. 12, 6202–6216. http://dx.doi.org/10.1109/TSP.2012.2218810

[29] J. Wang and B. Shim, *On the recovery limit of sparse signals using orthogonal matching pursuit.* IEEE Trans. Signal Process. **60**(2012), no. 9, 4973–4976. http://dx.doi.org/10.1109/TSP.2012.2203124

[30] J. M. Wen, X. M. Zhu, and D. F. Li, *Improved bounds on restricted isometry constant for orthogonal matching pursuit.* Electronic letters. **49**(2013), no. 23, 1487–1489.

[31] R. Wu, W. Huang, and D. R. Chen, *The exact support recovery of sparse signals with noise via orthogonal matching pursuit.* IEEE Signal Proc. Lett. **20**(2013), no. 4, 403–406.

[32] Z. Q. Xu, *The performance of orthogonal multi-matching pursuit under RIP.* Computer Science **33**(2015), 495–516.

Institute of Applied Physics and Computational Mathematics, Beijing, 100088, China
e-mail: chenwg@iapcm.ac.cn

Graduate School, China Academy of Engineering Physics, Beijing, 100088, China
e-mail: gehuanmin@163.com