Statistical Theory and Related Fields

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

A new result on recovery sparse signals using orthogonal matching pursuit

Xueping Chen, Jianzhong Liu & Jiandong Chen

To cite this article: Xueping Chen, Jianzhong Liu & Jiandong Chen (2022) A new result on recovery sparse signals using orthogonal matching pursuit, Statistical Theory and Related Fields, 6:3, 220-226, DOI: 10.1080/24754269.2022.2048445

To link to this article: https://doi.org/10.1080/24754269.2022.2048445

Published online: 13 Mar 2022.

STATISTICAL THEORY AND RELATED FIELDS
2022, VOL. 6, NO. 3, 220–226
https://doi.org/10.1080/24754269.2022.2048445

Taylor & Francis
Taylor & Francis Group

A new result on recovery sparse signals using orthogonal matching pursuit

Xueping Chen, Jianzhong Liu and Jiandong Chen

School of Mathematics and Physics, Jiangsu University of Technology, Jiangsu, People's Republic of China

ABSTRACT
Orthogonal matching pursuit (OMP) algorithm is a classical greedy algorithm widely used in compressed sensing. In this paper, by exploiting the Wielandt inequality and some properties of orthogonal projection matrix, we obtained a new number of iterations required for the OMP algorithm to perform exact recovery of sparse signals, which improves significantly upon the latest results as we know.

ARTICLE HISTORY
Received 19 November 2020
Revised 4 February 2022
Accepted 23 February 2022

KEYWORDS
Compressed sensing; orthogonal matching pursuit; Wielandt inequality; orthogonal projection matrix

1. Introduction

Orthogonal matching pursuit (OMP) has received growing attention due to its simplicity and competitive reconstruction performance recently. Consider the following compressed linear model:

$$y = \Phi x, \tag{1}$$

where $x \in \mathbf{C}^n$ is a K-sparse signal (i.e., $\|x\|_0 \le K$), $\Phi = [\phi_1, \phi_2, \dots, \phi_n] \in \mathbf{C}^{m\times n}$ is a known measurement matrix with $m \ll n$ and $y \in \mathbf{C}^m$ is the observation signal. It has been demonstrated that under some appropriate conditions on Φ, OMP can reliably recover the signal x based on a set of compressive observations y by iteratively identifying the support of the sparse signal according to the maximum correlation between columns of measurement matrix and the current residual. See Table 1 for a detailed description of the OMP algorithm (Cai & Wang, 2011; Chang & Wu, 2014; Tropp & Gilbert, 2007; Wang & Shim, 2016; Wen et al., 2020, 2017; Wu et al., 2013). In Table 1, $supp(x)$ is the set of nonzero positions in x. r^k denotes the residual after the kth iteration of OMP and T^k the estimated support set within kth iteration of OMP.

In compressed sensing, a commonly used framework for analysing the recovery performance is the restricted isometry property (RIP) (Cai et al., 2010; Candes & Tao, 2005; Chang & Wu, 2014). A matrix Φ is said to satisfy the RIP of order K if there exists a constant $\delta \in [0, 1)$ such that

$$(1-\delta)\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1+\delta)\|x\|_2^2, \tag{2}$$

for all K-sparse signal x. In particular, the minimum of all constants δ satisfying (2) is called the K-order Restricted Isometry Constant (RIC) and denoted by δ_K. Over the years, many RIP-based conditions have been proposed to guarantee exact recovery of any K-sparse signals via OMP in K iterations. It has been shown in Davenport and Wakin (2010) that $\delta_{K+1} < (3\sqrt{K})^{-1}$ is sufficient for OMP to recover any K-sparse signals x in K iterations. The sufficient reconstruction condition of OMP is then improved to $\delta_{K+1} < (1 + 2\sqrt{K})^{-1}$ by Huang and Zhu (2011). Mo (2015) demonstrated that $\delta_{K+1} < (\sqrt{K+1})^{-1}$ is a sharp condition for exact recovery of any K-sparse signal with OMP in K iterations. Our recent work Liu et al. (2017) provides some sufficient conditions for recovering restricted classes of K-sparse signals with a more relaxed bound on RIC.

Obviously, running fewer number of iterations of OMP offers many benefits and many efforts have been made to improve the condition (Cai et al., 2009; Chang & Wu, 2014; Li & Wen, 2019; Wen et al., 2017; Wu et al., 2013). In Livshitz (2012), Livshitz showed that with proper choices of α and β ($\alpha \sim 2 \times 10^5, \beta \sim 10^{-6}$), OMP accurately reconstructs K-sparse signals in $\alpha K^{1.2}$ iterations under $\delta_{\alpha K^{1.2}} = \beta K^{-0.2}$. It has been shown by Zhang (2011) that OMP recovers any K-sparse signal in 30 K iterations under $\delta_{31K} \le 3^{-1}$. Livshitz and Temlyakov (2014) considered random sparse signals and showed that with high probability, these signals can be recovered within $\lceil (1+\epsilon)K \rceil$ iterations of OMP for any $\epsilon > 0$. Recently, Wang and Shim (2016) showed that if

$$c \ge -\frac{4(1+\delta)}{1-\delta} \ln\left(\frac{1}{2} - \sqrt{\frac{\delta}{2+2\delta}}\right), \tag{3}$$

and $\delta_{\lfloor (c+1)K \rfloor} \le \delta$, OMP can recover the K-sparse signals in $\lceil cK \rceil$ iterations. It is the best result as we know in the literature.

CONTACT Xueping Chen chenxueping@jsut.edu.cn

Table 1. Orthogonal matching pursuit.

Input:	Φ, y, and maximum iteration number $k_{\max}$
Initialize:	$k = 0, r^0 = y, T^0 = \emptyset$
while	$k < k_{\max}$, do
	$k = k + 1$,
	$t^k = \arg\max_{1 \le i \le n} \lvert\langle r^{k-1}, \phi_i \rangle\rvert$,
	$T^k = T^{k-1} \cup \{t^k\}$,
	$\widehat{x}^k = \arg\min_{u \in \mathbb{C}^n : \mathrm{supp}(u) \subset T^k} \lVert y - \Phi u \rVert_2$,
	$r^k = y - \Phi \widehat{x}^k$.
Output:	T^k and $\widehat{x}^k$

In this paper, we present a new result on how many iterations of OMP would be enough to guarantee exact recovery of sparse signals:

$$c \geq -\frac{4(1+\delta_1)}{1-\delta} \ln\left(\frac{1-\delta}{2}\right),$$

which improves significantly upon the results proposed in Wang and Shim (2016).

We first give some notation. Let $\mathbf{N} = \{0, 1, 2, \ldots\}$, $\mathbf{N}^+ = \{1, 2, \ldots\}$ and $\Omega_n = \{1, 2, \ldots, n\}$. $[\cdot]$ and $\lceil\cdot\rceil$ denote floor and ceiling function, respectively. For any two sets Λ and Γ, let $\Lambda \setminus \Gamma = \{i : i \in \Lambda, i \notin \Gamma\}$, and $|\Lambda|$ is the cardinality of Λ. For $\Lambda \subset \Omega_n$ and $\Lambda \neq \emptyset$, Φ_Λ denotes the submatrix of Φ that contains only the columns indexed by Λ and x_Λ denotes the subvector of x that contains only the entries indexed by Λ, and $\mathrm{span}(\Phi_\Lambda)$ represents the span of columns in Φ_Λ. Let $P_\Lambda = \Phi_\Lambda(\Phi_\Lambda^* \Phi_\Lambda)^+ \Phi_\Lambda^*$ stand for an orthogonal projection matrix onto $\mathrm{span}(\Phi_\Lambda)$, where Φ^* is the conjugate transpose of the matrix Φ, and $(\Phi_\Lambda^* \Phi_\Lambda)^+$ is Moore–Penrose pseudo inverse of $\Phi_\Lambda^* \Phi_\Lambda$. $P_\Lambda^\perp = I_m - P_\Lambda$ is an orthogonal projection matrix onto the orthogonal complement of $\mathrm{span}(\Phi_\Lambda)$, where I_m denotes the identity matrix. In particular, if $\Lambda = \emptyset$, then $x_\emptyset$ is a 0-by-1 empty vector, $\Phi_\emptyset$ is an m-by-0 empty matrix, $\Phi_\emptyset x_\emptyset$ is an m-by-1 zero matrix and $\mathrm{span}(\Phi_\emptyset) := \{0\}$. For further details on empty matrices, see, e.g., Bernstein (2005).

2. Main results

For notational simplicity, we denote $\Gamma^k = T \setminus T^k$.

Theorem 2.1: *For any $\delta \in (0, 1)$, let*

$$\alpha = -\frac{4(1+\delta_1)}{1-\delta} \ln\left(\frac{1-\delta}{2}\right),$$

and $\kappa = [\alpha K]$. If the measurement matrix Φ in (1) satisfies the RIP of order $K + \kappa$ and $\delta_{K+\kappa} \leq \delta$, then $|\Gamma^\kappa| = 0$.

Remark 2.1 (Performance of Theorem 2.1): From Theorem 2.1, if

$$c \geq -\frac{4(1+\delta_1)}{1-\delta} \ln\left(\frac{1-\delta}{2}\right), \tag{4}$$

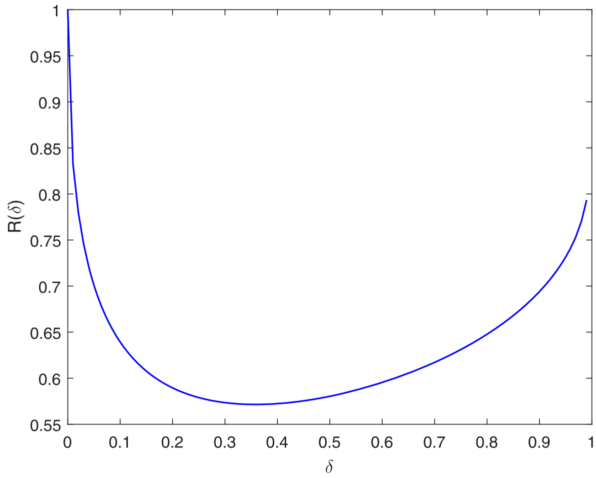

Figure 1. Performance of Theorem 2.1.

and $\delta_{[(c+1)K]} \leq \delta$, then OMP perfectly recovers the signal x from the measurements $y = \Phi x$ in $[cK]$ iteration. In the following, we compare the lower bound in (4) with the result of Wang and Shim (2016), which has been showed in (3). We first establish an upper bound for the ratio of (4) to (3) by using monotonicity property of the RIC ($\delta_1 \leq \delta_{K+\kappa}$) as

$$R(\delta) := \frac{\ln\left(\frac{1-\delta}{2}\right)}{\ln\left(\frac{1}{2} - \sqrt{\frac{\delta}{2+2\delta}}\right)}.$$

It is easy to check that $R(\delta) < 1$ for $0 < \delta < 1$, which means the lower bound of c in this paper is uniformly smaller than the one proposed in Wang and Shim (2016). See Figure 1, for example, we have $R(3^{-1}) = 0.57$ and $R(2^{-1}) = 0.58$. For $0.015 < \delta < 0.993$, we have $0.57 < R(\delta) < 0.8$. That is, nearly 98% of the values of the bounds proposed in this paper is less than 0.8 times of the one in (3).

Remark 2.2 (Main differences with the previous work): Our methods have several key distinctions in construction of the more efficient lower and upper bounds of resident in OMP. First, Wang and Shim (2016) obtained the upper bound of residual in OMP algorithm based on the following fundamental set:

$$\{\lVert x_\Gamma \rVert_2 : \Gamma \subset \Gamma^k\}.$$

We modified the above set to

$$\{\lVert \Phi_\Gamma x_\Gamma \rVert_2 : \Gamma \subset \Gamma^k\},$$

which leads to a more efficient upper bound of resident in OMP algorithm (see inequality (17) of Section 3.3). Second, Livshitz and Temlyakov (2014), Wang and Shim (2016), and Zhang (2011) only use RIP to obtain the lower bound of resident. However, in this paper, we not only used RIP, but also utilized the Wielandt inequality (Wang & Ip, 1999) to derive more efficient lower bound of resident (see inequality (16) of Section 3.3). The details can be found in the following sections.

3. Proof of the main theorem

3.1. Preliminaries

In the following, we introduce the subset Γ_τ^k of Γ^k based on the magnitude of elements in the set $\{\|\Phi_\Gamma x_\Gamma\|_2 : \Gamma \subset \Gamma^k\}$. For $k, \tau \in N$ and $\Gamma^k \neq \emptyset$, let

$$\aleph_\tau = \{\Lambda : \Lambda \subset \Gamma^k, |\Lambda| \leq 2^\tau - 1\}, \tag{5}$$

$$f(\tau) = \min\{\|\Phi_{\Gamma^k\setminus\Lambda} x_{\Gamma^k\setminus\Lambda}\| : \Lambda \in \aleph_\tau\}. \tag{6}$$

It is clear that $\aleph_\tau \subset \aleph_{\tau+1}$, and $f(\tau) \geq f(\tau+1)$ with $\tau \in \mathbf{N}$.

Definition 3.1: If subset Γ_τ^k in $\aleph_\tau$ satisfies: (a) $\|\Phi_{\Gamma^k\setminus\Gamma_\tau^k} x_{\Gamma^k\setminus\Gamma_\tau^k}\|_2 = f(\tau)$ and (b) $|\Gamma_\tau^k| = \max\{|\Lambda| : \Lambda \in \aleph_\tau, \|\Phi_{\Gamma^k\setminus\Lambda} x_{\Gamma^k\setminus\Lambda}\|_2 = f(\tau)\}$, then Γ_τ^k is said to be the $\Phi - k - \tau$ characteristic set of x.

It is easy to verify that the characteristic set Γ_τ^k has the following properties:

- **(P.1)** $|\Gamma_\tau^k| \leq 2^\tau - 1$, for $\tau \in \mathbf{N}$;
- **(P.2)** If $\Gamma \subset \Gamma^k$ and $\|\Phi_\Gamma x_\Gamma\|_2 < \|\Phi_{\Gamma^k\setminus\Gamma_\tau^k} x_{\Gamma^k\setminus\Gamma_\tau^k}\|_2$, then $|\Gamma| \leq |\Gamma^k| - 2^\tau$;
- **(P.3)** $0 = |\Gamma_0^k| \leq |\Gamma_1^k| \leq \cdots$, and $\Gamma_{J_k}^k = \Gamma_{J_k+1}^k = \cdots = \Gamma^k$, with $J_k = 1 + [\log_2 |\Gamma^k|]$.

In fact, (P.1) is obvious as $\Gamma_\tau^k \in \aleph_\tau$. For (P.2), since $\Gamma \subset \Gamma^k$, Γ can be rewritten as $\Gamma = \Gamma^k \setminus (\Gamma^k \setminus \Gamma)$. From (6) and $\|\Phi_{\Gamma^k\setminus\Gamma_\tau^k} x_{\Gamma^k\setminus\Gamma_\tau^k}\|_2 = f(\tau)$, we know that if $\|\Phi_\Gamma x_\Gamma\|_2 < \|\Phi_{\Gamma^k\setminus\Gamma_\tau^k} x_{\Gamma^k\setminus\Gamma_\tau^k}\|_2$, then $\Gamma^k \setminus \Gamma \notin \aleph_\tau$. Hence $|\Gamma^k \setminus \Gamma| = |\Gamma^k| - |\Gamma| \geq 2^\tau$. That is $|\Gamma| \leq |\Gamma^k| - 2^\tau$. For (P.3), if $f(\tau) = f(\tau+1)$, it is obvious that $|\Gamma_\tau^k| \leq |\Gamma_{\tau+1}^k|$ based on Definition 3.1. If $f(\tau) > f(\tau+1)$, then

$$\|\Phi_{\Gamma^k\setminus\Gamma_{\tau+1}^k} x_{\Gamma^k\setminus\Gamma_{\tau+1}^k}\|_2 < \|\Phi_{\Gamma^k\setminus\Gamma_\tau^k} x_{\Gamma^k\setminus\Gamma_\tau^k}\|_2.$$

From (P.2),

$$|\Gamma^k| - |\Gamma_{\tau+1}^k| = |\Gamma^k \setminus \Gamma_{\tau+1}^k| \leq |\Gamma^k| - 2^\tau.$$

Thus $|\Gamma_{\tau+1}^k| \geq 2^\tau$. From (P.1), $|\Gamma_\tau^k| \leq 2^\tau - 1 < |\Gamma_{\tau+1}^k|$. Otherwise, it is easy to verify that $|\Gamma^k| \leq 2^{J_k} - 1 \leq 2^\tau - 1$ for $J_k \leq \tau$. Thus we have $\Gamma_\tau^k = \Gamma^k$ based on Definition 3.1.

From the fact $|\Gamma_0^k| = 0$, $|\Gamma_{J_k}^k| = |\Gamma^k|$, and (P.3), there exists $0 \leq \nu_k \leq J_k - 1$ such that $\Gamma_0^k = \cdots = \Gamma_{\nu_k}^k = \emptyset$, and $\Gamma_{\nu_k+1}^k \neq \emptyset$. Then it follows from (P.3) that $\Gamma_{\nu_k+i}^k \neq \emptyset, i = 1, 2, \ldots$.

Let $\sigma = (1-\delta)^{-1}$, thus $\sigma > 1$. It can be verified that there exists the maximum value of $\sigma^i \|\Phi_{\Gamma^k\setminus\Gamma_i^k} x_{\Gamma^k\setminus\Gamma_i^k}\|_2^2$ over $i \in \mathbf{N}$. Since $\|\Phi_{\Gamma^k\setminus\Gamma_i^k} x_{\Gamma^k\setminus\Gamma_i^k}\|_2^2 = 0$ based on the fact that $\Gamma_i^k = \Gamma^k$ from (P.3) for $i \geq J_k$. Hence, we give the following definition.

Definition 3.2: Let $\Gamma^k \neq \emptyset$, integer $L(k)$ is said to be the $k - \sigma$ character of x if it satisfies the following condition:

$$\sigma^{L(k)} \|\Phi_{\Gamma^k\setminus\Gamma_{L(k)}^k} x_{\Gamma^k\setminus\Gamma_{L(k)}^k}\|_2^2 = \max_{i\in\mathbf{N}} \sigma^i \|\Phi_{\Gamma^k\setminus\Gamma_i^k} x_{\Gamma^k\setminus\Gamma_i^k}\|_2^2.$$

By Definition 3.2, it is easy to check that

$$\|\Phi_{\Gamma^k\setminus\Gamma_{L(k)}^k} x_{\Gamma^k\setminus\Gamma_{L(k)}^k}\|_2^2 \geq \sigma^{-L(k)} \|\Phi_{\Gamma^k} x_{\Gamma^k}\| > 0$$

and

$$\|\Phi_{\Gamma^k\setminus\Gamma_i^k} x_{\Gamma^k\setminus\Gamma_i^k}\|_2^2 \leq \sigma^{L(k)-i} \|\Phi_{\Gamma^k\setminus\Gamma_{L(k)}^k} x_{\Gamma^k\setminus\Gamma_{L(k)}^k}\|_2^2, i \in \mathbf{N}. \tag{7}$$

Recalling that $\|\Phi_{\Gamma^k\setminus\Gamma_i^k} x_{\Gamma^k\setminus\Gamma_i^k}\|_2^2 = 0$ for $i \geq J_k$ and $\sigma > 1$, we have $\nu_k \leq L(k) \leq J_k - 1$. Thus

$$2^{L(k)} \leq 2^{J_k-1} \leq |\Gamma^k|. \tag{8}$$

Based on $k - \sigma$ character $L(k)$, we define the following integer sequence $\{\kappa_i\}$ as

$$\kappa_i = \begin{cases} 0 & \text{for } i = 0, \\ \kappa_{i-1} & \text{for } |\Gamma^{\kappa_{i-1}}| = 0, \\ \kappa_{i-1} + [\alpha 2^{L(\kappa_{i-1})}] & \text{for } |\Gamma^{\kappa_{i-1}}| \neq 0. \end{cases} \tag{9}$$

Obviously, the sequence $\{\kappa_i\}$ is monotonically non-decreasing, i.e., $\kappa_0 \leq \kappa_1 \leq \cdots$.

3.2. Main idea

Inspired by Wang and Shim (2016) and Zhang (2011), we show that after k iterations, OMP can select a substantial amount of indices in Γ^k for a specified number of additional iterations, and the rate of the number of additional iterations to the number of chosen indices is upper bounded by the constant α. More precisely, we give the following important proposition, which is the basis of Theorem 2.1.

Proposition 3.1: *Suppose* $\Gamma^k \neq \emptyset, k + [\alpha|\Gamma^k|] \leq \kappa$, *and* $\delta_{K+\kappa} \leq \delta$, *then we have*

$$|\Gamma^{k+[\alpha 2^{L(k)}]}| \leq |\Gamma^k| - 2^{L(k)}, \tag{10}$$

where $L(k)$ *is the* $k - \sigma$ *character of* x.

Now, we can prove Theorem 2.1 based on Proposition 3.1. The proof can be divided into two steps.

Step 1: We first prove that

$$\kappa_i + [\alpha|\Gamma^{\kappa_i}|] \leq \kappa, \quad i \in \mathbf{N}. \tag{11}$$

Obviously, (11) holds when $i = 0$ since $\kappa_0 + [\alpha|\Gamma^0|] = [\alpha|\Gamma^0|] \leq [\alpha K] = \kappa$. In the following, we assume that

$\kappa_{i-1} + \lceil \alpha |\Gamma^{\kappa_{i-1}}| \rceil \le \kappa$. If $|\Gamma^{\kappa_{i-1}}| = 0$, then $|\Gamma^{\kappa_i}| = 0$ from $\Gamma^{\kappa_i} \subset \Gamma^{\kappa_{i-1}}$. It follows that

$$\kappa_i + \lceil \alpha |\Gamma^{\kappa_i}| \rceil = \kappa_i = \kappa_{i-1} + \lceil \alpha |\Gamma^{\kappa_{i-1}}| \rceil \le \kappa.$$

If $|\Gamma^{\kappa_{i-1}}| \neq 0$, by substituting $k = \kappa_{i-1}$ into (10), we have

$$|\Gamma^{\kappa_i}| = |\Gamma^{\kappa_{i-1} + \lceil \alpha 2^{L(\kappa_{i-1})} \rceil}| \le |\Gamma^{\kappa_{i-1}}| - 2^{L(\kappa_{i-1})}.$$

Thus

$$\begin{aligned}
&\kappa_i + \lceil \alpha |\Gamma^{\kappa_i}| \rceil \\
&\quad = \lceil \kappa_i + \alpha |\Gamma^{\kappa_i}| \rceil \\
&\quad \le \lceil \kappa_{i-1} + \alpha 2^{L(\kappa_{i-1})} + \alpha (|\Gamma^{\kappa_{i-1}}| - 2^{L(\kappa_{i-1})}) \rceil \\
&\quad = \kappa_{i-1} + \lceil \alpha |\Gamma^{\kappa_{i-1}}| \rceil \\
&\quad \le \kappa.
\end{aligned}$$

This completes the proof of (11).

Step 2: We prove that there exists a constant $s \in \mathbf{N}$ such that $|\Gamma^{\kappa_s}| = 0$. On one hand, from (11) we have

$$\kappa_i \le \kappa_i + \lceil \alpha |\Gamma^{\kappa_i}| \rceil \le \kappa, \quad i \in \mathbf{N}. \tag{12}$$

Thus $\{\kappa_i\}$ is a monotonically non-decreasing and bounded integer sequence. On the other hand, it is clear that $\alpha \ge 4 \ln 2 > 2$. Then from (9), one can easily show that $\kappa_{i-1} < \kappa_i$ for $|\Gamma^{\kappa_{i-1}}| \neq 0$. Therefore, there must exist a constant $s \in \mathbf{N}$ such that $|\Gamma^{\kappa_s}| = 0$. From (12), we have $\kappa_s \le \kappa$, then $|\Gamma^{\kappa}| \le |\Gamma^{\kappa_s}|$. Hence $|\Gamma^{\kappa}| = 0$, which completes the proof.

3.3. *Sketch of proof of Proposition 3.1*

We here give a sketch of the proof of Proposition 3.1, the remaining details of (16) and (17) can be found in the Appendix. For notational simplicity, let $k + \lceil \alpha 2^{L(k)} \rceil = n_1$. From $\Gamma^k \neq \emptyset$, $k + \lceil \alpha |\Gamma^k| \rceil \le \kappa$, and (8), we have

$$n_1 = k + \lceil \alpha 2^{L(k)} \rceil \le k + \lceil \alpha |\Gamma^k| \rceil \le \kappa, \tag{13}$$

and (10) can be rewritten as

$$|\Gamma^{n_1}| \le |\Gamma^k| - 2^{L(k)}. \tag{14}$$

By (P.2), a sufficient condition of the above inequality is

$$\|\Phi_{\Gamma^{n_1}} x_{\Gamma^{n_1}}\|_2^2 < \|\Phi_{\Gamma_k \setminus \Gamma^k_{L(k)}} x_{\Gamma_k \setminus \Gamma^k_{L(k)}}\|_2^2. \tag{15}$$

Now, what remains is the proof of (15), which is based on the analysis of the residual of OMP. First, by exploiting Wielandt inequality and RIP, we construct a lower bound for $\|r^{n_1}\|_2^2$, that is,

$$\|r^{n_1}\|_2^2 \ge (1 - \delta^2) \|\Phi_{\Gamma^{n_1}} x_{\Gamma^{n_1}}\|_2^2. \tag{16}$$

Next, by exploiting some properties of orthogonal projection matrix and RIP, we construct an upper bound for $\|r^{n_1}\|_2^2$,

$$\|r^{n_1}\|_2^2 < (1 - \delta^2) \|\Phi_{\Gamma_k \setminus \Gamma^k_{L(k)}} x_{\Gamma_k \setminus \Gamma^k_{L(k)}}\|_2^2. \tag{17}$$

From (16) and (17), it is easy to verify (15) holds. Hence the remains is the proof of (16) and (17) (see Appendices).

4. Conclusion

In this paper, we analyse the number of iterations required for OMP to exactly recover sparse signals. Our analysis shows that OMP can recover any K-sparse signals in $\lceil cK \rceil$ iterations ($c \ge -\frac{4(1+\delta_1)}{1-\delta} \ln \frac{1-\delta}{2}$), which is uniformly smaller than the one proposed in Wang and Shim (2016). For example, to accurately recover K-sparse signals, it has been shown in Wang and Shim (2016) that OMP requires 30 K iterations with $\delta_{31\,\mathrm{K}} \le 2^{-1}$ while our result shows that OMP requires only $\lceil 16.8\,\mathrm{K} \rceil$ iterations with $\delta_{\lceil 17.8\,\mathrm{K} \rceil} \le 2^{-1}$. In practical application, a large number of iterations often lead to a high computational complexity. Our result provides a theoretical basis for the reduction of the number of iterations required for OMP.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The corresponding author gratefully acknowledges support from the National Natural Science Foundation of China No. 11971204, Natural Science Foundation of Jiangsu Province of China No. BK20200108, and the Zhongwu Youth Innovative Talent Program of Jiangsu University of Technology.

ORCID

Xueping Chen http://orcid.org/0000-0003-0499-0972

References

Bernstein, D. S. (2005). *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory*. Princeton University Press.

Cai, T., & Wang, L. (2011). Orthogonal matching pursuit for sparse signal recovery with noise. *IEEE Transactions on Information Theory*, *57*(7), 4680–4688. https://doi.org/10.1109/TIT.2011.2146090

Cai, T., Wang, L., & Xu, G. (2010). New bounds for restricted isometry constants. *IEEE Transactions on Information Theory*, *56*(9), 4388–4394. https://doi.org/10.1109/TIT.2010.2054730

Cai, T., Xu, G., & Zhang, J. (2009). On recovery of sparse signals via l_1 minimization. *IEEE Transactions on Information Theory*, *55*(7), 3388–3397. https://doi.org/10.1109/TIT.2009.2021377

Candes, E. J., & Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory*, *51*(12), 4203–4215. https://doi.org/10.1109/TIT.2005.858979

Chang, L. H., & Wu, J. Y. (2014). An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, *60*(9), 405–408. https://doi.org/10.1109/TIT.18

Davenport, M. A., & Wakin, M. B. (2010). Analysis of orthogonal matching pursuit using the restricted isometry property. *IEEE Transactions on Information Theory*, *56*(9), 4395–4401. https://doi.org/10.1109/TIT.2010.2054653

Huang, S., & Zhu, J. (2011). Recovery of sparse signals using OMP and its variants: convergence analysis based on RIP. *Inverse Problems*, *27*(3), 035003. https://doi.org/10.1088/0266-5611/27/3/035003

Li, H. F., & Wen, J. M. (2019). A new analysis for support recovery with block orthogonal matching pursuit. *IEEE Signal Processing Letters*, *26*(2), 247–251. https://doi.org/10.1109/LSP.2018.2885919

Liu, C., Fang, Y., & Liu, J. Z. (2017). Some new results about sufficient conditions for exact support recovery of sparse signals via orthogonal matching pursuit. *IEEE Transactions on Signal Processing*, *65*(17), 4511–4524. https://doi.org/10.1109/TSP.2017.2711543

Livshitz, E. D. (2012). On the efficiency of the orthogonal matching pursuit in compressed sensing. *Sbornik Mathematics*, *203*(2), 33–44. https://doi.org/10.4213/sm

Livshitz, E. D., & Temlyakov, V. N. (2014). Sparse approximation and recovery by greedy algorithms. *IEEE Transactions on Information Theory*, *60*(7), 3989–4000. https://doi.org/10.1109/TIT.2014.2320932

Mo, Q. (2015). A sharp restricted isometry constant bound of orthogonal matching pursuit. arXiv: 1501.01708 [Online]. Available: http://arxiv.org/pdf/1501.01708v1.pdf.

Tropp, J. A., & Gilbert, A. C. (2007). Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, *53*(12), 4655–4666. https://doi.org/10.1109/TIT.2007.909108

Wang, J., & Shim, B. (2016). Exact recovery of sparse signals using orthogonal matching pursuit: How many iterations do we need. *IEEE Transactions on Signal Processing*, *64*(16), 4194–4202. https://doi.org/10.1109/TSP.2016.2568162

Wang, S. G., & Ip, W. -C. (1999). A matrix version of the Wielandt inequality and its application to statistics. *Linear Algebra Its Appl*, *2*, 118–120. https://doi.org/10.1016/S0024-3795(99)00117-2

Wen, J., Zhang, R., & Yu, W. (2020). Signal-dependent performance analysis of orthogonal matching pursuit for exact sparse recovery. *IEEE Transactions on Signal Processing*, *68*, 5031–5046. https://doi.org/10.1109/TSP.78

Wen, J., Zhou, Z., Wang, J., Tang, X., & Mo, Q. (2017). A sharp condition for exact support recovery with orthogonal matching pursuit. *IEEE Transactions on Signal Processing*, *65*(6), 1370–1382. https://doi.org/10.1109/TSP.2016.2634550

Wu, R., Huang, W., & Chen, D.-R. (2013). The exact support recovery of sparse signals with noise via orthogonal matching pursuit. *IEEE Signal Processing Letters*, *20*(4), 403–406. https://doi.org/10.1109/LSP.2012.2233734

Zhang, T. (2011). Sparse recovery with orthogonal matching pursuit under RIP. *IEEE Transactions on Information Theory*, *57*(9), 6215–6221. https://doi.org/10.1109/TIT.2011.2162263

Appendices

Appendix 1. Proof of (16)

The proof of (16) is based on the Wielandt inequality (Wang & Ip, 1999), which is presented in the following Lemma.

Lemma A.1 (Wielandt inequality): *Let* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ *be an n-order positive-definite matrix with* $aI_n \le A \le bI_n$, $(a > 0)$. *Then we have*

$$A_{21}A_{11}^{-1}A_{12} \le \left(\frac{b-a}{b+a}\right)^2 A_{22}. \tag{A1}$$

We now proceed to the proof of (16). The conclusion holds naturally if $\Gamma^{n_1} = \emptyset$. In the following, we prove that (16) still holds under $\Gamma^{n_1} \ne \emptyset$. Without loss of generality, we assume that $T^{n_1} = \{1, 2, \dots, n_1\}$, and $\Gamma^{n_1} = \{n_1+1, n_1+2, \dots, n_1+n_2\}$, where $n_2 = |\Gamma^{n_1}|$. Notice that

$$\Phi^*_{T^{n_1}\cup\Gamma^{n_1}}\Phi_{T^{n_1}\cup\Gamma^{n_1}} = \begin{pmatrix} \Phi^*_{T^{n_1}}\Phi_{T^{n_1}} & \Phi^*_{T^{n_1}}\Phi_{\Gamma^{n_1}} \\ \Phi^*_{\Gamma^{n_1}}\Phi_{T^{n_1}} & \Phi^*_{\Gamma^{n_1}}\Phi_{\Gamma^{n_1}} \end{pmatrix}. \tag{A2}$$

By (13), we have $n_1 + n_2 \le K + \kappa$. Furthermore, from the monotonicity of the RIC and $\delta_{K+\kappa} \le \delta$, we have $\delta_{n_1+n_2} \le \delta_{K+\kappa} \le \delta$. It implies that

$$(1-\delta)I_{n_1+n_2} \le \Phi^*_{T^{n_1}\cup\Gamma^{n_1}}\Phi_{T^{n_1}\cup\Gamma^{n_1}} \le (1+\delta)I_{n_1+n_2}. \tag{A3}$$

Since (A2), (A3), and Lemma A.1, we have

$$\begin{aligned}\Phi^*_{\Gamma^{n_1}}P_{T^{n_1}}\Phi_{\Gamma^{n_1}} &= \Phi^*_{\Gamma^{n_1}}\Phi_{T^{n_1}}(\Phi^*_{T^{n_1}}\Phi_{T^{n_1}})^{-1}\Phi^*_{T^{n_1}}\Phi_{\Gamma^{n_1}} \\ &\le \delta^2\Phi^*_{\Gamma^{n_1}}\Phi_{\Gamma^{n_1}}. \end{aligned}\tag{A4}$$

Hence by (A4), it follows that

$$\begin{aligned}\|r^{n_1}\|_2^2 &= \|P^{\perp}_{T^{n_1}}\Phi_{\Gamma^{n_1}}x_{\Gamma^{n_1}}\|_2^2 \\ &= x^*_{\Gamma^{n_1}}\Phi^*_{\Gamma^{n_1}}P^{\perp}_{T^{n_1}}\Phi_{\Gamma^{n_1}}x_{\Gamma^{n_1}} \\ &= x^*_{\Gamma^{n_1}}\Phi^*_{\Gamma^{n_1}}\Phi_{\Gamma^{n_1}}x_{\Gamma^{n_1}} - x^*_{\Gamma^{n_1}}\Phi^*_{\Gamma^{n_1}}P_{T^{n_1}}\Phi_{\Gamma^{n_1}}x_{\Gamma^{n_1}} \\ &\ge (1-\delta^2)\|\Phi_{\Gamma^{n_1}}x_{\Gamma^{n_1}}\|_2^2. \end{aligned}$$

This completes the proof.

Appendix 2. Proof of (17)

The proof of (17) is based on the following Lemma A.2, which generalized the results of Wang and Shim (2016) to the complex field.

Lemma A.2: *Suppose* $\Gamma^k \ne \emptyset$, *and* Γ *be a nonempty subset of* Γ^k, *the residual of OMP satisfies*

$$\|r^{l'}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 \le C_{\Gamma,l,l'}(\|r^l\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2), \tag{A5}$$

where $C_{\Gamma,l,l'} = \exp\left(-\frac{(1-\delta_{|T^{l'-1}\cup\Gamma|})(l'-l)}{(1+\delta_1)|\Gamma|}\right)$ *for any integer* $l' \ge l \ge k$.

Proof: Since $l \le l'$, it follows that $\|r^{l'}\|_2^2 \le \|r^l\|_2^2$. Notice that $0 < C_{\Gamma,l,l'} \le 1$. Thus if $l' = l$ or $\|r^{l'}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 \le 0$, then Lemma A.2 holds evidently. In the following, we assume that $l' > l$ and

$$\|r^{l'}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 > 0. \tag{A6}$$

It implies that

$$\begin{aligned}\|r^i\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 \ge \|r^{l'}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 > 0, \\ i = 0, 1, \dots, l'. \end{aligned}\tag{A7}$$

We later proved that the residual of OMP satisfies

$$\begin{aligned}&\|r^{i+1}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 \\ &\le \exp\left(-\frac{1-\delta_{|T^{l'-1}\cup\Gamma|}}{(1+\delta_1)|\Gamma|}\right)\left(\|r^i\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right), \end{aligned}\tag{A8}$$

for $i = l, l+1, \dots, l'-1$. Then (A5) can be obtained by (A8) as follows,

$$\begin{aligned}&\|r^{l'}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2 \\ &\le \exp\left(-\frac{1-\delta_{|T^{l'-1}\cup\Gamma|}}{(1+\delta_1)|\Gamma|}\right)\left(\|r^{l'-1}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right) \\ &\le \exp\left(-\frac{2(1-\delta_{|T^{l'-1}\cup\Gamma|})}{(1+\delta_1)|\Gamma|}\right) \end{aligned}$$

$$\times\left(\|r^{l'-2}\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right)$$
$$\le\cdots$$
$$\le\exp\left(-\frac{(1-\delta_{|T^{l'-1}\cup\Gamma|})(l'-l)}{(1+\delta_1)|\Gamma|}\right)$$
$$\times\left(\|r^{l}\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right).$$

(A8) can be proved as follows. It is easy to check that $P^{\perp}_{T^{i+1}}=P^{\perp}_{T^{i+1}}P^{\perp}_{T^i}$. Thus

$$r^{i+1}=P^{\perp}_{T^{i+1}}y=P^{\perp}_{T^{i+1}}P^{\perp}_{T^i}y=P^{\perp}_{T^{i+1}}r^i.$$

Furthermore,

$$\|r^i\|_2^2-\|r^{i+1}\|_2^2=\|r^i\|_2^2-\|P^{\perp}_{T^{i+1}}r^i\|_2^2\overset{(a)}{=}\|P_{T^{i+1}}r^i\|_2^2$$
$$\overset{(b)}{\ge}\|P_{\{t^{i+1}\}}r^i\|_2^2\overset{(c)}{=}\frac{|\langle r^i,\phi_{t^{i+1}}\rangle|^2}{\phi^*_{t^{i+1}}\phi_{t^{i+1}}}\overset{(d)}{\ge}\frac{|\langle r^i,\phi_{t^{i+1}}\rangle|^2}{1+\delta_1},\qquad\text{(A9)}$$

where (a), (b), (c), (d) above can be obtained as follows: (a) is from that fact that $\|P^{\perp}_{T^{i+1}}r^i\|_2^2+\|P_{T^{i+1}}r^i\|_2^2=\|r^i\|_2^2$; (b) is due to $t^{i+1}\in T^{i+1}$; (c) is because

$$\|P_{\{t^{i+1}\}}r^i\|_2^2=(r^i)^*P_{\{t^{i+1}\}}r^i=(r^i)^*\phi_{t^{i+1}}(\phi^*_{t^{i+1}}\phi_{t^{i+1}})^+\phi^*_{t^{i+1}}r^i$$

and $(\phi^*_{t^{i+1}}\phi_{t^{i+1}})^+=\frac{1}{\phi^*_{t^{i+1}}\phi_{t^{i+1}}}$; (d) is followed from the definition of RIP.

Now we construct a lower bound for $|\langle r^i,\phi_{t^{i+1}}\rangle|^2$ below.

Notice that $P^{\perp}_{T^i}\Phi_\Lambda x_\Lambda=0$ for $\Lambda\subset T^i$ and $T\setminus\Gamma^k\subset T^k\subset T^i$. Thus we have

$$r^i=P^{\perp}_{T^i}y=P^{\perp}_{T^i}\Phi_Tx_T=P^{\perp}_{T^i}(\Phi_{\Gamma^k}x_{\Gamma^k}+\Phi_{T\setminus\Gamma^k}x_{T\setminus\Gamma^k})$$
$$=P^{\perp}_{T^i}\Phi_{\Gamma^k}x_{\Gamma^k}.\qquad\text{(A10)}$$

From (A10), it follows that

$$\|r^i-P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2$$
$$=\|r^i-P^{\perp}_{T^i}\Phi_\Gamma x_\Gamma+P^{\perp}_{T^i}\Phi_{\Gamma\cap T^i}x_{\Gamma\cap T^i}\|_2^2$$
$$=\|r^i-P^{\perp}_{T^i}\Phi_\Gamma x_\Gamma\|_2^2=\|P^{\perp}_{T^i}(\Phi_{\Gamma^k}x_{\Gamma^k}-\Phi_\Gamma x_\Gamma)\|_2^2$$
$$\le\|\Phi_{\Gamma^k}x_{\Gamma^k}-\Phi_\Gamma x_\Gamma\|_2^2=\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2.\qquad\text{(A11)}$$

Notice that $P^{\perp}_{T^i}r^i=r^i$. From (A11), we have

$$2\mathrm{Re}\langle r^i,\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle=2\mathrm{Re}\langle P^{\perp}_{T^i}r^i,\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle$$
$$=2\mathrm{Re}\langle r^i,P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle$$
$$=\|r^i\|_2^2+\|P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2-\|r^i-P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2$$
$$\ge\|r^i\|_2^2+\|P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2.\qquad\text{(A12)}$$

Hence by (A7) and (A12), we have

$$2\mathrm{Re}\langle r^i,\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle\ge\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma})\|_2^2>0,\qquad\text{(A13)}$$

and $\Gamma\setminus T^i\ne\emptyset$.

Now we prove

$$\|P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2\ge(1-\delta_{|T^i\cup\Gamma|})\|x_{\Gamma\setminus T^i}\|_2^2.\qquad\text{(A14)}$$

In fact, if $i=0$, then $T^0=\emptyset$ and $P^{\perp}_{T^0}=I_m$. From RIP, we have

$$\|P^{\perp}_{T^0}\Phi_{\Gamma\setminus T^0}x_{\Gamma\setminus T^0}\|_2^2=\|\Phi_\Gamma x_\Gamma\|_2^2\ge(1-\delta_{|\Gamma|})\|x_\Gamma\|_2^2$$
$$=(1-\delta_{|T^0\cup\Gamma|})\|x_{\Gamma\setminus T^0}\|_2^2.$$

For $i\ne0$, we have

$$\|P^{\perp}_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2$$
$$=\|\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}-P_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2$$
$$=\|\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}-\Phi_{T^i}(\Phi^*_{T^i}\Phi_{T^i})^+\Phi^*_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2$$
$$=\left\|\begin{pmatrix}\Phi_{\Gamma\setminus T^i} & \Phi_{T^i}\end{pmatrix}\begin{pmatrix}x_{\Gamma\setminus T^i}\\ -(\Phi^*_{T^i}\Phi_{T^i})^+\Phi^*_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\end{pmatrix}\right\|_2^2$$
$$\ge(1-\delta_{|T^i\cup\Gamma|})(\|x_{\Gamma\setminus T^i}\|_2^2$$
$$+\|(\Phi^*_{T^i}\Phi_{T^i})^+\Phi^*_{T^i}\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\|_2^2)$$
$$\ge(1-\delta_{|T^i\cup\Gamma|})\|x_{\Gamma\setminus T^i}\|_2^2.$$

Then from (A12), (A14) and the arithmetic-geometric mean inequality, it can be verified that

$$\mathrm{Re}\langle r^i,\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle$$
$$\ge\sqrt{1-\delta_{|T^i\cup\Gamma|}}\|x_{\Gamma\setminus T^i}\|_2\sqrt{\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2}.\qquad\text{(A15)}$$

On the other hand, we have $|\langle r^i,\phi_\tau\rangle|\le|\langle r^i,\phi_{t^{i+1}}\rangle|$ for $\tau\in\Omega_n$. Thus

$$\mathrm{Re}\langle r^i,\Phi_{\Gamma\setminus T^i}x_{\Gamma\setminus T^i}\rangle=\mathrm{Re}\left\langle r^i,\sum_{\tau\in\Gamma\setminus T^i}x_\tau\phi_\tau\right\rangle$$
$$\le\left|\left\langle r^i,\sum_{\tau\in\Gamma\setminus T^i}x_\tau\phi_\tau\right\rangle\right|$$
$$=\left|\sum_{\tau\in\Gamma\setminus T^i}x_\tau\langle r^i,\phi_\tau\rangle\right|$$
$$\le\|x_{\Gamma\setminus T^i}\|_1|\langle r^i,\phi_{t^{i+1}}\rangle|.\qquad\text{(A16)}$$

Notice that $\frac{\|x_{\Gamma\setminus T^i}\|_1^2}{\|x_{\Gamma\setminus T^i}\|_2^2}\le|\Gamma\setminus T^i|\le|\Gamma|$. Combining (A15) and (A16), we obtain

$$|\langle r^i,\phi_{t^{i+1}}\rangle|^2$$
$$\ge\frac{(1-\delta_{|T^i\cup\Gamma|})\|x_{\Gamma\setminus T^i}\|_2^2}{\|x_{\Gamma\setminus T^i}\|_1^2}(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2)$$
$$\ge\frac{(1-\delta_{|T^i\cup\Gamma|})}{|\Gamma|}(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2).\qquad\text{(A17)}$$

This, together with (A9), implies

$$\|r^i\|_2^2-\|r^{i+1}\|_2^2\ge\frac{1-\delta_{|T^i\cup\Gamma|}}{(1+\delta_1)|\Gamma|}(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2).\qquad\text{(A18)}$$

Hence (A18) can be rewritten as

$$\|r^{i+1}\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2$$
$$\le\left(1-\frac{1-\delta_{|T^i\cup\Gamma|}}{(1+\delta_1)|\Gamma|}\right)\left(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right).\qquad\text{(A19)}$$

Let $a=\frac{1-\delta_{|T^i\cup\Gamma|}}{(1+\delta_1)|\Gamma|}$, then $1-a\le\exp(-a)$. From monotonicity of the RIC, we can obtain that $\delta_{|T^i\cup\Gamma|}\le\delta_{|T^{l'-1}\cup\Gamma|}$ for $i\le l'-1$. Thus from (A19), we have

$$\|r^{i+1}\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2$$
$$\le\exp\left(-\frac{1-\delta_{|T^i\cup\Gamma|}}{(1+\delta_1)|\Gamma|}\right)\left(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right)$$
$$\le\exp\left(-\frac{1-\delta_{|T^{l'-1}\cup\Gamma|}}{(1+\delta_1)|\Gamma|}\right)\left(\|r^i\|_2^2-\|\Phi_{\Gamma^k\setminus\Gamma}x_{\Gamma^k\setminus\Gamma}\|_2^2\right),$$

for $i = l, l+1, \ldots, l'-1$. This completes the proof of (A8). ■

We now proceed to the proof of (17). Let $k_0 = k$ and

$$k_i = k + \sum_{\tau=\nu_k+1}^{\nu_k+i} \left\lceil \frac{\alpha}{4}|\Gamma_\tau^k| \right\rceil, \quad i = 1, 2, \ldots, L(k) + 1 - \nu_k.$$

Let $l' = k_i, l = k_{i-1}$ and $\Gamma = \Gamma^k_{\nu_k+i}$ in (A5). Notice that $k_i - k_{i-1} = \lceil \frac{\alpha}{4}|\Gamma^k_{\nu_k+i}|\rceil$. Then it follows that

$$\begin{aligned}
&\|r^{k_i}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2 \\
&\quad \le C_{\Gamma^k_{\nu_k+i},k_{i-1},k_i}\left(\|r^{k_{i-1}}\|_2^2 - \|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2\right),
\end{aligned} \tag{A20}$$

where

$$C_{\Gamma^k_{\nu_k+i},k_{i-1},k_i} = \exp\left(-\frac{(1-\delta_{|T^{k_{i-1}}\cup\Gamma^k_{\nu_k+i}|})\lceil\frac{\alpha}{4}|\Gamma^k_{\nu_k+i}|\rceil}{(1+\delta_1)|\Gamma^k_{\nu_k+i}|}\right).$$

In the following, we construct an upper bound for $C_{\Gamma^k_{\nu_k+i},k_{i-1},k_i}$. Recall that $\alpha \ge 4\ln 2 \ge 2$. Then from Appendix 2 in Wang and Shim (2016), we have

$$\sum_{\tau=1}^{l} \left\lceil \frac{\alpha}{4}(2^l - 1) \right\rceil \le \lceil \alpha 2^{l-1} \rceil - 1. \tag{A21}$$

By (P.1) and (A21), it follows that

$$\begin{aligned}
k_{L(k)+1-\nu_k} &= k + \sum_{\tau=\nu_k+1}^{L(k)+1} \left\lceil \frac{\alpha}{4}|\Gamma_\tau^k| \right\rceil \\
&\le k + \sum_{\tau=1}^{L(k)+1} \left\lceil \frac{\alpha}{4}(2^\tau - 1) \right\rceil \\
&\le k + \lceil \alpha 2^{L(k)} \rceil - 1 \\
&\le k + [\alpha 2^{L(k)}] \\
&= n_1.
\end{aligned} \tag{A22}$$

Notice that $\Gamma^k_\tau \subset \Gamma^k \subset T$, and $T^{k_{i-1}} \subset T^{k_{L(k)+1-\nu_k}}$ for $i = 1, 2, \ldots, L(k) + 1 - \nu_k$. From (13) and (A22), we obtain

$$|\Gamma^k_{\nu_k+i} \cup T^{k_{i-1}}| \le |T \cup T^{k_{L(k)+1-\nu_k}}| \le |T \cup T^{n_1}| \le K + \kappa.$$

Furthermore, from monotonicity of the RIC and $\delta_{K+\kappa} \le \delta$, we have

$$\frac{1-\delta_{|\Gamma^k_{\nu_k+i}\cup T^{k_{i-1}}|}}{1+\delta_1} \ge \frac{1-\delta_{K+\kappa}}{1+\delta_1} \ge \frac{1-\delta}{1+\delta_1}. \tag{A23}$$

Hence,

$$\begin{aligned}
C_{\Gamma^k_{\nu_k+i},k_{i-1},k_i} &= \exp\left(-\frac{(1-\delta_{|T^{k_{i-1}}\cup\Gamma^k_{\nu_k+i}|})\lceil\frac{\alpha}{4}|\Gamma^k_{\nu_k+i}|\rceil}{(1+\delta_1)|\Gamma^k_{\nu_k+i}\|}\right) \\
&\le \exp\left(-\frac{(1-\delta)\frac{\alpha}{4}|\Gamma^k_{\nu_k+i}|}{(1+\delta_1)|\Gamma^k_{\nu_k+i}|}\right) \\
&= (1-\delta)/2.
\end{aligned} \tag{A24}$$

Together with (A20), it implies that

$$\|r^{k_i}\|_2^2 \le \frac{1-\delta}{2}\|r^{k_{i-1}}\|_2^2 + \frac{1+\delta}{2}\|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2, \tag{A25}$$

for $i = 1, 2, \ldots, L(k) + 1 - \nu_k$. Note that we only need to consider $\|r^{k_{i-1}}\|_2^2 \ge \|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2$ since if $\|r^{k_{i-1}}\|_2^2 < \|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2$, (A25) holds from

$$\begin{aligned}
\|r^{k_i}\|_2^2 \le \|r^{k_{i-1}}\|_2^2 &= \frac{1-\delta}{2}\|r^{k_{i-1}}\|_2^2 + \frac{1+\delta}{2}\|r^{k_{i-1}}\|_2^2 \\
&\le \frac{1-\delta}{2}\|r^{k_{i-1}}\|_2^2 \\
&\quad + \frac{1+\delta}{2}\|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k+i}}\|_2^2.
\end{aligned}$$

By substituting (7) into (A25), we further obtain

$$\begin{aligned}
\|r^{k_i}\|_2^2 &\le \frac{1-\delta}{2}\|r^{k_{i-1}}\|_2^2 \\
&\quad + \frac{1+\delta}{2}\sigma^{L(k)-\nu_k-i}\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2.
\end{aligned} \tag{A26}$$

From $\Gamma^k_{\nu_k} = \emptyset$ and (7), we have

$$\begin{aligned}
\|r^k\|_2^2 &= \|P^\perp_{T^k}\Phi_{\Gamma^k} x_{\Gamma^k}\|_2^2 \\
&= \|P^\perp_{T^k}\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k}}\|_2^2 \\
&\le \|\Phi_{\Gamma^k\setminus\Gamma^k_{\nu_k}} x_{\Gamma^k\setminus\Gamma^k_{\nu_k}}\|_2^2 \\
&\le \sigma^{L(k)-\nu_k}\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2.
\end{aligned} \tag{A27}$$

Notice that $\sigma = \frac{1}{1-\delta}$ and $\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2 > 0$, Then by (A22), (A26), (A27), we obtain

$$\begin{aligned}
\|r^{n_1}\|_2^2 &\le \|r^{k_{L(k)+1-\nu_k}}\|_2^2 \\
&\le \left(\frac{1-\delta}{2}\right)^{L(k)+1-\nu_k} (\|r^k\|_2^2 \\
&\quad + \frac{1+\delta}{2}\sum_{\tau=1}^{L(k)+1-\nu_k}\left(\frac{1-\delta}{2}\right)^{-\tau} \\
&\quad \times \sigma^{L(k)-\nu_k-\tau}\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2) \\
&\le \left(\frac{1-\delta}{2}\right)^{L(k)+1-\nu_k} (\sigma^{L(k)-\nu_k} \\
&\quad + \left(\frac{1+\delta}{2}\right)\sum_{\tau=1}^{L(k)+1-\nu_k}\left(\frac{1-\delta}{2}\right)^{-\tau} \\
&\quad \times \sigma^{L(k)-\nu_k-\tau})\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2 \\
&= (1-\delta^2 - \frac{\delta(1-\delta)}{2^{L(k)+1-\nu_k}})\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2 \\
&< (1-\delta^2)\|\Phi_{\Gamma^k\setminus\Gamma^k_{L(k)}} x_{\Gamma^k\setminus\Gamma^k_{L(k)}}\|_2^2.
\end{aligned}$$

Thus (17) holds. This completes the proof.