Exact Recovery of Sparse Signals Using Orthogonal Matching Pursuit: How Many Iterations Do We Need?

Jian Wang and Byonghyo Shim, *Senior Member, IEEE*

Abstract—Orthogonal matching pursuit (OMP) is a greedy algorithm widely used for the recovery of sparse signals from compressed measurements. In this paper, we analyze the number of iterations required for the OMP algorithm to perform exact recovery of sparse signals. Our analysis shows that OMP can accurately recover all K-sparse signals within $\lceil 2.8\, K \rceil$ iterations when the measurement matrix satisfies a restricted isometry property (RIP). Our result improves upon the recent result of Zhang and also bridges the gap between Zhang's result and the fundamental limit of OMP at which exact recovery of K-sparse signals cannot be uniformly guaranteed.

Index Terms—Sparse signal recovery, orthogonal matching pursuit (OMP), restricted isometry property (RIP).

TABLE I
THE OMP ALGORITHM

Input	$\mathbf{\Phi}$, $\mathbf{y}$, and maximum iteration number $k_{\max}$.
Initialize	iteration counter $k = 0$,
	estimated support $T^0 = \emptyset$,
	and residual vector $\mathbf{r}^0 = \mathbf{y}$.
While	$k < k_{\max}$, **do**
	$k = k + 1$.
	Identify $t^k = \arg\max_{i \in \Omega \setminus T^{k-1}} \lvert \langle \phi_i, \mathbf{r}^{k-1} \rangle \rvert$.
	Enlarge $T^k = T^{k-1} \cup t^k$.
	Estimate $\hat{\mathbf{x}}^k = \arg\min_{\mathbf{u}: supp(\mathbf{u})=T^k} \lVert \mathbf{y} - \mathbf{\Phi u} \rVert_2$.
	Update $\mathbf{r}^k = \mathbf{y} - \mathbf{\Phi}\hat{\mathbf{x}}^k$.
End	
Output	T^k and $\hat{\mathbf{x}}^k$

I. INTRODUCTION

RECENTLY, there has been a growing interest in recovering sparse signals from compressed measurements [1]–[8]. The main goal of this task is to accurately estimate a high dimensional K-sparse vector $\mathbf{x} \in \mathbb{R}^n$ ($\lVert \mathbf{x} \rVert_0 \leq K$) from a small number of linear measurements $\mathbf{y} \in \mathbb{R}^m$ ($m \ll n$). The relationship between the signal vector and the measurements is given by

$$\mathbf{y} = \mathbf{\Phi x}, \qquad (1)$$

where $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ is often called the measurement (sensing) matrix. A key finding in the sparse recovery problem is that one can recover the original vector $\mathbf{x}$ with far fewer measurements than traditional approaches use, as long as the signal to be recovered is sparse and the measurement matrix roughly preserves the energy of the signal of interest. Among many algorithms designed to recover the sparse signal, orthogonal matching pursuit (OMP) algorithm has received much attention for its competitive

performance as well as practical benefits, such as implementation simplicity and low computational complexity [9], [10]. In essence, the OMP algorithm estimates the input sparse vector $\mathbf{x}$ and its support (index set of nonzero elements) in an iterative fashion, generating a series of locally optimal updates fitting the measurement data. Specifically, at each iteration the index of column that is mostly correlated with the modified measurements (often called residual) is chosen as a new element of the estimated support set. The vestiges of columns in the estimated support are then eliminated from the measurements, yielding a new residual for the next iteration. See Table I for a detailed description of the OMP algorithm.

Over the years, the OMP algorithm has long been considered as a heuristic algorithm hard to be analyzed. Recently, however, much research has been devoted to discovering the condition of OMP ensuring exact recovery of sparse signals. In one direction, studies to identify the recovery condition using probabilistic analyses have been proposed. Tropp and Gilbert showed that when the measurement matrix $\mathbf{\Phi}$ is generated i.i.d. at random, and the measurement size is on the order of $K \log n$, OMP ensures the accurate recovery of every fixed K-sparse signal with overwhelming probability [11]. Another line of work is to characterize exact recovery conditions of OMP using properties of measurement matrices, such as the mutual incoherence property (MIP) and the restricted isometry property (RIP) [13]. A measurement matrix $\mathbf{\Phi}$ is said to satisfy the RIP of order K if there exists a constant $\delta(\mathbf{\Phi}) \in (0, 1)$ such that

$$(1 - \delta(\mathbf{\Phi})) \lVert \mathbf{x} \rVert_2^2 \leq \lVert \mathbf{\Phi x} \rVert_2^2 \leq (1 + \delta(\mathbf{\Phi})) \lVert \mathbf{x} \rVert_2^2 \quad (2)$$

for any K-sparse vector $\mathbf{x}$. In particular, the minimum of all constants $\delta(\mathbf{\Phi})$ satisfying (2) is called the restricted isometry constant (RIC) and denoted by $\delta_K(\mathbf{\Phi})$. In the sequel, we use

J. Wang is with School of information and Control, Nanjing University of Information Science and Technology, Nanjing 210044, China, and also with Department of Electrical and Computer Engineering, Seoul National University, Seoul 151-744, Korea (e-mail: wangjianeee@gmail.com).

B. Shim is with Institute of New Media and Communications and School of Electrical and Computer Engineering, Seoul National University, Seoul 151-744, Korea (e-mail: bshim@snu.ac.kr).

δ_K instead of $\delta_K(\boldsymbol{\Phi})$ for brevity. In [14], Davenport and Wakin showed that OMP ensures exact reconstruction of any K-sparse signal under

$$\delta_{K+1} < \frac{1}{3\sqrt{K}}. \tag{3}$$

Since then, many efforts have been made to improve this condition [15]–[19]. Recently, Mo has improved the condition to [20]

$$\delta_{K+1} < \frac{1}{\sqrt{K+1}}, \tag{4}$$

which is in fact sharp since there exist measurement matrices $\boldsymbol{\Phi}$ with $\delta_{K+1} = \frac{1}{\sqrt{K+1}}$, for which OMP fails to recover the original K-sparse signal from its measurements in K iterations [17], [20]. Therefore, in order to uniformly recover all K-sparse signals in K iterations of OMP, the RIC should at least be inversely proportional to $\sqrt{K}$.

While aforementioned studies of OMP have focused on the scenario where the number of iterations is limited to the sparsity K, there have been recent works investigating the behavior of OMP when it performs more than K iterations [21]–[24] or when it chooses more than one index per iteration [25]–[28]. Both in theoretical performance guarantees and empirical simulations, these approaches provide better results and also offers new insights into this seemingly simple-minded yet clever algorithm. Livshitz showed that with proper choices of α and β ($\alpha \sim 2 \cdot 10^5$ and $\beta \sim 10^{-6}$), OMP accurately reconstructs K-sparse signals in $\lfloor \alpha K^{1.2} \rfloor$ iterations under [22]

$$\delta_{\alpha K^{1.2}} = \beta K^{-0.2}. \tag{5}$$

Although the RIC decays slowly with K (when compared to that in the results of [14]–[20]), and thus offers significant benefits in the measurement size, it is not easy to enjoy the benefits in practice, since it requires too many iterations. Recently, it has been shown by Zhang that OMP recovers any K-sparse signal with $30K$ iterations under [21]

$$\delta_{31K} < \frac{1}{3}. \tag{6}$$

The significance of this result is that when running $30K$ iterations, OMP can recover K-sparse signals accurately with the RIC being an absolute constant independent of K, exhibiting the reconstruction capability comparable to the state of the art sparse recovery algorithms (e.g., Basis Pursuit [29] and CoSaMP [4]). In the sequel, to distinguish the OMP algorithm running $\lceil cK \rceil$ ($c > 1$) iterations from the conventional OMP algorithm running K iterations, we denote it as OMP_{cK}.

In this paper, we go further to investigate how many iterations of OMP_{cK} would be enough to guarantee exact recovery of sparse signals, given that the RIC is an absolute constant. Note that running fewer number of iterations offers many computational benefits in practice. Our main result, described in Theorem 2, is that OMP_{cK} accurately recovers all K-sparse signals $\mathbf{x}$ from the measurements $\mathbf{y} = \boldsymbol{\Phi}\mathbf{x}$ if c satisfies a condition expressed in terms of the RIC. The main significance of this result is that as long as

$$c > 4 \log 2 \approx 2.8, \tag{7}$$

there always exists an RIC, which is an absolute constant, such that the underlying condition is fulfilled. Therefore, the required number of iterations of OMP for exact recovery of sparse signals can be as few as $\lceil 2.8 K \rceil$.

It is well known that with overwhelming probability, random matrices (e.g., random Gaussian, Bernoulli, and partial Fourier matrices) satisfy the RIP when the number of measurements scales nearly linearly with the sparsity K [13], [30]. In view of this, our result implies that if a K-sparse signal is measured by random matrices, it can be recovered with the nearly optimal number of measurements via the OMP algorithm running only $\lceil 2.8 K \rceil$ iterations.

We summarize notations used throughout this paper. $\Omega = \{1, 2, \ldots, n\}$. $T = supp(\mathbf{x}) = \{i | i \in \Omega, x_i \neq 0\}$ is the set of nonzero positions in $\mathbf{x}$. For $S \subseteq \Omega$, $T \setminus S$ is the set of all elements contained in T but not in S. $|S|$ is the cardinality of S. If $|S| \neq 0$, $\mathbf{x}_S \in \mathbb{R}^{|S|}$ is the restriction of the vector $\mathbf{x}$ to those elements indexed by S. Similarly, $\boldsymbol{\Phi}_S \in \mathbb{R}^{m \times |S|}$ is a submatrix of $\boldsymbol{\Phi}$ that only contains columns indexed by S. If $\boldsymbol{\Phi}_S$ has full column rank, $\boldsymbol{\Phi}_S^\dagger = (\boldsymbol{\Phi}_S' \boldsymbol{\Phi}_S)^{-1} \boldsymbol{\Phi}_S'$ is the Moore-Penrose pseudoinverse of $\boldsymbol{\Phi}_S$ where $\boldsymbol{\Phi}_S'$ is the transpose of the matrix $\boldsymbol{\Phi}_S$. $\mathcal{P}_S = \boldsymbol{\Phi}_S \boldsymbol{\Phi}_S^\dagger$ is the projection onto $span(\boldsymbol{\Phi}_S)$ (i.e., the span of columns in $\boldsymbol{\Phi}_S$). In particular, if $S = \varnothing$, $\mathbf{x}_\varnothing$ is a 0-by-1 empty vector, $\boldsymbol{\Phi}_\varnothing$ is an m-by-0 empty matrix, and $\boldsymbol{\Phi}_\varnothing \mathbf{x}_\varnothing$ is an m-by-1 zero matrix [31] (for further discussion on empty matrices, see, e.g., [32]).

II. EXACT RECOVERY OF SPARSE SIGNALS VIA OMP

Suppose $\lceil cK \rceil$ ($c > 1$) is the number of iterations ensuring selection of all support indices of the K-sparse signal $\mathbf{x}$ (i.e., $T \subseteq T^{\lceil cK \rceil}$). Then the estimated support set $T^{\lceil cK \rceil}$ may contain indices not in T. Even in this situation, the final result is unaffected and the original signal $\mathbf{x}$ is recovered accurately (i.e., $\widehat{\mathbf{x}}^{\lceil cK \rceil} = \mathbf{x}$) because

$$\widehat{\mathbf{x}}^{\lceil cK \rceil} = \arg \min_{\mathbf{u}: supp(\mathbf{u}) = T^{\lceil cK \rceil}} \| \mathbf{y} - \boldsymbol{\Phi}\mathbf{u} \|_2$$

and

$$\begin{aligned}
\left(\widehat{\mathbf{x}}^{\lceil cK \rceil}\right)_{T^{\lceil cK \rceil}} &= \boldsymbol{\Phi}_{T^{\lceil cK \rceil}}^\dagger \mathbf{y} = \boldsymbol{\Phi}_{T^{\lceil cK \rceil}}^\dagger \boldsymbol{\Phi}_T \mathbf{x}_T \\
&\overset{(a)}{=} \boldsymbol{\Phi}_{T^{\lceil cK \rceil}}^\dagger \boldsymbol{\Phi}_{T^{\lceil cK \rceil}} \mathbf{x}_{T^{\lceil cK \rceil}} \\
&= \mathbf{x}_{T^{\lceil cK \rceil}}, \tag{8}
\end{aligned}$$

where (a) is from the fact that $\mathbf{x}$ is supported on T and hence $\mathbf{x}_{T^{\lceil cK \rceil} \setminus T} = \mathbf{0}$. This simple property allows us to investigate OMP running more than K iterations. While running more iterations than the sparsity level would be beneficial in obtaining better recovery bound, at the same time it induces additional computational burden. In fact, since the dimension of the matrix to be inverted increases by one per iteration (see Table I), both the operation cost and running time increase cubically with the number of iterations. Therefore, it is of importance to investigate the lower bound for the number of iterations ensuring accurate identification of the whole support, i.e., the lower bound for c that ensures

$$T \subseteq T^{\lceil cK \rceil}. \tag{9}$$

Our main result is formally described in the following theorem.

Theorem 1: Let $\mathbf{x} \in \mathbb{R}^n$ be any K-sparse signal supported on T and $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ be the measurement matrix. Furthermore, let $N^k := |T \setminus T^k|$ be the number of remaining support indices after k ($0 \le k \le \lceil cK \rceil$) iterations of OMP_{cK}. Then if $\mathbf{\Phi}$ obeys the RIP of order $s := |T \cup T^{k+\lfloor cN^k \rfloor}|$ and

$$c \ge -\frac{4(1+\delta_1)}{1-\delta_s} \log\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\delta_{N^k} + \delta_s}{1 + \delta_{N^k}}} \right), \quad (10)$$

then OMP_{cK} satisfies $T \subseteq T^{k+\lceil cN^k \rceil}$.

The proof of Theorem 1 will be given in Section III. The key point of Theorem 1 is that after performing k ($0 \le k \le \lceil cK \rceil$) iterations, OMP_{cK} selects the remaining N^k support indices within $\lceil cN^k \rceil$ additional iterations, as long as the condition in (10) is fulfilled. In particular, when $k=0$, $N^k = |T \setminus T^k| = |T \setminus \varnothing| = K$ and $T \subseteq T^{\lceil cK \rceil}$ holds true under

$$c \ge -\frac{4(1+\delta_1)}{1-\delta_{|T \cup T^{\lfloor cK \rfloor}|}} \log\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\delta_K + \delta_{|T \cup T^{\lfloor cK \rfloor}|}}{1 + \delta_K}} \right). \tag{11}$$

Further, from monotonicity of the RIC (i.e., $\delta_{K_1} \le \delta_{K_2}$ for $K_1 \le K_2$), one can easily show that (11) is guaranteed by

$$c \ge -\frac{4(1+\delta)}{1-\delta} \log\left(\frac{1}{2} - \sqrt{\frac{\delta}{2+2\delta}} \right)$$

$$\text{where } \delta := \delta_{\lfloor (c+1)K \rfloor}. \tag{12}$$

Hence, we obtain a simpler version of Theorem 1.

Theorem 2: Let $\mathbf{x} \in \mathbb{R}^n$ be any K-sparse signal and let $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ be the measurement matrix satisfying the RIP of order $\lfloor (c+1)K \rfloor$. Then if c satisfies (12), OMP_{cK} perfectly recovers the signal $\mathbf{x}$ from the measurements $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$.

In Fig. 1, we compare the lower bound of c in (12) with the result of Zhang ([21], Theorem 2.1). In both cases, one can observe that the lower bound of c increases monotonically with the RIC. In particular, the lower bound c of this work is uniformly smaller than that of [21] for the whole range of RIC. For example, it requires $30\,K$ iterations to recover K-sparse signals with $\delta_{31K} = \frac{1}{3}$ in [21]. Whereas, it requires only $\lceil 15.4\,K \rceil$ iterations in our new result.[1]

Another interesting point we observe from Fig. 1 is the difference in the critical value of c such that $\delta_{\lfloor (s+1)K \rfloor} = 0$. In [21], the lower bound of c ensuring $\delta_{\lfloor (s+1)K \rfloor} > 0$ is

$$c > 4 \log 20 \approx 12, \tag{13}$$

while that in our work is

$$c > 4 \log 2 \approx 2.8. \tag{14}$$

This means that if OMP_{cK} runs at least $\lceil 2.8\,K \rceil$ iterations, there always exists an absolute constant $\delta_{\lfloor (c+1)K \rfloor} \in (0,1)$ satisfying (12), and under this condition OMP performs the exact
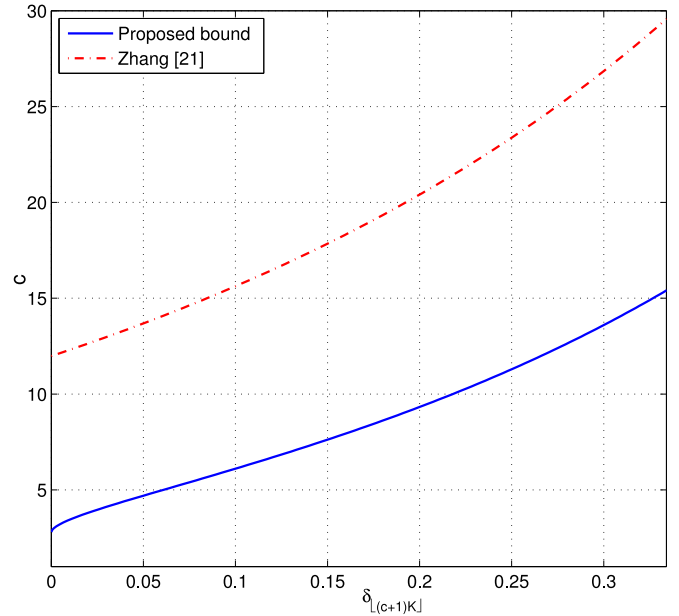

Fig. 1. Comparison between the proposed bound and Zhang's result [21].

recovery of all K-sparse signals. In fact, by applying $c = 2.8$ to (12), we obtain the upper bound of the RIC as

$$\delta_{\lceil 3.8K \rceil} \le 2 \cdot 10^{-5}. \tag{15}$$

Remark 1 (RIP Condition): It is worth mentioning that the value of c and the upper bound of the RIC are closely related. In fact, we can obtain better (larger) RIC bounds by using larger values of c. For example, when using $c = 30$ in Theorem 2, we can obtain the upper bound of the RIC as

$$\delta_{31K} \le \frac{1}{2}, \tag{16}$$

which is better (less restrictive) than the result $\delta_{31K} < \frac{1}{3}$ [21].

Remark 2 (Why Smaller c?): Running fewer number of iterations of OMP_{cK} offers many benefits. Noting that the number of indices chosen by OMP_{cK} should not exceed the number of measurements (i.e., $\lceil cK \rceil \le m$),[2] a smaller value of c directly leads to a wider range of sparsity K when m is fixed. On the other hand, when K is fixed, running fewer number of iterations is also beneficial since larger c may require larger number of measurements. Furthermore, identifying a small value of c is of importance for the noisy case because running too many iterations will degrade the denoising performance of the algorithm [33]–[36].

Remark 3 (Fundamental Limit of c): While Theorem 2 demonstrates that OMP_{cK} can uniformly recover all K-sparse signals using at least $\lceil 2.8\,K \rceil$ iterations, it should be noted that the number cannot be smaller or equal to K, given that the RIC is an absolute constant. In fact, to ensure exact recovery with the conventional OMP algorithm (i.e., OMP_{cK} with $c = 1$), the RIC should be at least inversely proportional to $\sqrt{K}$ [15], [16], which therefore places a fundamental limit to the recovery

---

[1]From Fig. 1, the RIP condition associated with $\lceil 15.4\,K \rceil$ iterations is $\delta_{\lfloor 16.4\,K \rfloor} \le \frac{1}{3}$ in our result, which is less restrictive than $\delta_{31K} < \frac{1}{3}$ [21].

[2]Otherwise the signal estimation step (i.e., the least squares (LS) projection) in the OMP algorithm cannot be performed.

performance of OMP_{cK}. Our result bridges the gap between the result of Zhang [21] and this fundamental limit.

Remark 4 (Measurement Size): It is well known that many random measurement matrices satisfy the RIP with overwhelming probability when the number of measurements scales linearly with the sparsity. For example, a random matrix $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ with entries drawn i.i.d. from Gaussian distribution $\mathcal{N}\left(0, \frac{1}{m}\right)$ obeys the RIP with $\delta_K = \varepsilon \in (0, 1)$ with overwhelming probability if $m \geq \frac{CK \log \frac{n}{K}}{\varepsilon^2}$ for some constant $C > 0$ [13], [30]. Therefore, our result implies that with high probability, OMP_{cK} can recover K-sparse signals in $\lceil 2.8 K \rceil$ iterations when the number of random Gaussian measurements is on the order of $K \log \frac{n}{K}$. This is essentially an encouraging result since by running slightly more than the sparsity level K, OMP_{cK} can accurately reconstruct all K-sparse signals with the same order of measurements as required by the state of the art sparse recovery algorithms (e.g., Basis Pursuit [29], [37] and CoSaMP [4]). This is in contrast to the conventional OMP algorithm, for which the RIC should be at least inversely proportional to $\sqrt{K}$ in order to ensure exact recovery of all K-sparse signals [15], [16], and hence the required number of measurements should be on the order of $K^2 \log \frac{n}{K}$.

Remark 5 (Comparison With [24]): Our result is closely related to the recent work of Livshitz and Temlyakov [24]. In their work, authors considered random sparse signals and showed that with high probability, these signals can be recovered with $\lceil (1 + \epsilon) K \rceil$ iterations of OMP_{cK} [24]. Our result has two key distinctions over this work. Firstly, while each nonzero component of sparse signals are upper bounded in [24], our result does not impose any constraint on the nonzero components of input signals. Secondly, and more importantly, the analysis in [24] relies on the assumption that $K \geq \delta_{2K}^{-1/2}$, which essentially applies to the situation where the sparsity K is nontrivial. In contrast, our analysis works for input signals with arbitrary sparsity levels.

III. PROOF OF THEOREM 1

A. Preliminaries

For notational simplicity, we denote $\Gamma^k = T \backslash T^k$ so that $N^k = |\Gamma^k|$. Also, without loss of generality, we assume that $\Gamma^k = \{1, \ldots, N^k\}$ and that nonzero elements in $\mathbf{x}_{\Gamma^k}$ are arranged in descending order of their magnitudes (i.e., $|x_1| \geq \cdots \geq |x_{N^k}|$). Now, we define the subset of Γ^k as

$$
\Gamma_\tau^k = \begin{cases} \varnothing & \tau = 0, \\ \{1, 2, \ldots, 2^\tau - 1\} & \tau = 1, 2, \ldots, \lfloor \log_2 N^k \rfloor, \\ \Gamma^k & \tau = \lfloor \log_2 N^k \rfloor + 1. \end{cases}
\tag{17}
$$

See Fig. 2 for the illustration of Γ_τ^k. Notice that the last set $\Gamma_{\lfloor \log_2 N^k \rfloor + 1}^k$ $(= \Gamma^k)$ may have less than $2^{\lfloor \log_2 N^k \rfloor + 1} - 1$ elements. For example, if $\Gamma^k = \{1, 2, \ldots, 10\}$, then $\Gamma_0^k = \varnothing, \Gamma_1^k = \{1\}, \Gamma_2^k = \{1, 2, 3\}, \Gamma_3^k = \{1, \ldots, 7\}$, and $\Gamma_4^k = \{1, 2, \ldots, 10\}$ has less than $2^4 - 1$ $(= 15)$ elements.
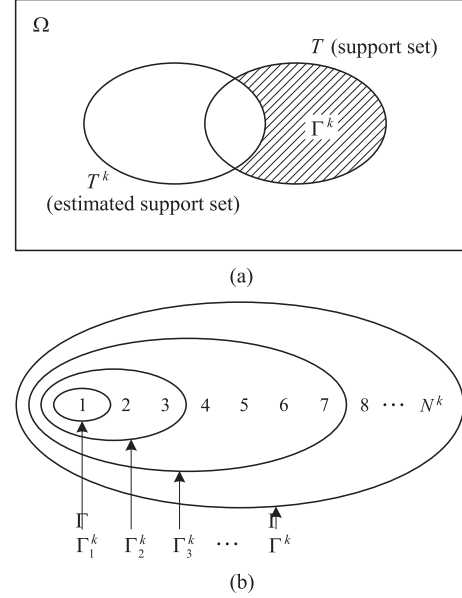

Fig. 2. Illustration of sets T, T^k, and Γ^k. (a) Set diagram of T, T^k, and Γ^k. (b) Illustration of indices in Γ^k.

For given set Γ^k and constant $\sigma > 1$, let $L \in \{1, \ldots, \lfloor \log_2 N^k \rfloor + 1\}$ be the minimum integer satisfying

$$
\| \mathbf{x}_{\Gamma^k \backslash \Gamma_0^k} \|_2^2 < \sigma \| \mathbf{x}_{\Gamma^k \backslash \Gamma_1^k} \|_2^2,
\tag{18a}
$$

$$
\| \mathbf{x}_{\Gamma^k \backslash \Gamma_1^k} \|_2^2 < \sigma \| \mathbf{x}_{\Gamma^k \backslash \Gamma_2^k} \|_2^2,
\tag{18b}
$$

$$
\vdots
$$

$$
\| \mathbf{x}_{\Gamma^k \backslash \Gamma_{L-2}^k} \|_2^2 < \sigma \| \mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k} \|_2^2,
\tag{18c}
$$

$$
\| \mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k} \|_2^2 \geq \sigma \| \mathbf{x}_{\Gamma^k \backslash \Gamma_L^k} \|_2^2.
\tag{18d}
$$

Moreover, if (18d) holds true for all $L \geq 1$, then we simply take $L = 1$ and ignore (18a)–(18c). We remark that L always exists because $\| \mathbf{x}_{\Gamma^k \backslash \Gamma_{\lfloor \log_2 N^k \rfloor + 1}^k} \|_2^2 = 0$ so that (18d) holds true at least for $L = \lfloor \log_2 N^k \rfloor + 1$.

B. Main Idea

The proof of Theorem 1 is based on mathematical induction in N^k, the number of remaining support indices after k iterations of OMP_{cK}. It is inspired by the work of Zhang [21]. First, we check the case where $N^k = 0$. This case is trivial since it implies that all support indices have already been selected $(T \subseteq T^k)$ so that no more iteration is needed. Then, we assume that the argument holds up to an integer $\gamma - 1$ $(\gamma \geq 1)$. In other words, we assume that whenever $N^k \leq \gamma - 1$, it requires at most $\lceil cN^k \rceil$ additional iterations to select all remaining indices in T. Under this inductive assumption, we will show that if $N^k = \gamma$, it also requires at most $\lceil c\gamma \rceil$ additional iterations to select the remaining γ indices in T (i.e., $T \subseteq T^{k + \lceil c\gamma \rceil}$).

We now proceed to the proof of the induction step $(N^k = \gamma)$.
1) First, we show that after a specified number of additional iterations, a substantial amount of indices in Γ^k can be

γ support indices in Γ^k are chosen within $\lceil c\gamma \rceil$ iterations

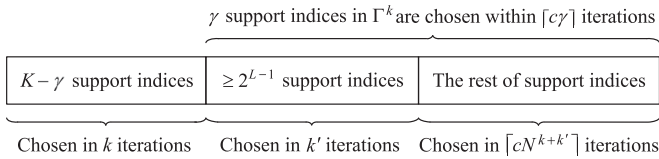| $K - \gamma$ support indices	$\geq 2^{L-1}$ support indices	The rest of support indices
Chosen in k iterations	Chosen in k' iterations	Chosen in $\lceil cN^{k+k'} \rceil$ iterations

Fig. 3. Illustration of the induction step when $N^k = \gamma$.

chosen, and the number of remaining support indices is upper bounded. More precisely, we show that OMP_{cK} chooses at least 2^{L-1} support indices of Γ^k in

$$k' := \lceil c2^{L-1} \rceil - 1 \qquad (19)$$

additional iterations (where L is defined in (18a)–(18d)) so that the number of remaining support indices (after $k + k'$ iterations) satisfies

$$N^{k+k'} \leq \gamma - 2^{L-1}. \qquad (20)$$

2) Second, since (20) directly implies $N^{k+k'} \leq \gamma - 1$, by induction hypothesis it requires at most $\lceil cN^{k+k'} \rceil$ additional iterations to choose the rest of support indices. In summary, the total number of iterations of OMP_{cK} to choose all support indices is no more than

$$k + k' + \lceil cN^{k+k'} \rceil \overset{(a)}{\leq} k + k' + \lceil c(\gamma - 2^{L-1}) \rceil$$

$$\overset{(b)}{=} k + (\lceil c2^{L-1} \rceil - 1) + \lceil c(\gamma - 2^{L-1}) \rceil$$

$$\overset{(c)}{\leq} k + \lceil c\gamma \rceil, \qquad (21)$$

where (a) is from (20), (b) follows from (19), and (c) holds true because $\lceil a \rceil + \lceil b \rceil - 1 \leq \lceil a + b \rceil$.

Therefore, we can conclude that all support indices are chosen in $k + \lceil c\gamma \rceil$ iterations of OMP_{cK} $\left(T \subseteq T^{k+\lceil c\gamma \rceil} \right)$, which establishes the induction step. An illustration of the induction step is given in Fig. 3. Now, what remains is the proof of (20).

C. Sketch of Proof for (20)

Before we proceed, we explain the key idea to prove this claim. Instead of proving (20) directly, we show that a sufficient condition of (20) is true. That is,

$$\| \mathbf{x}_{\Gamma^{k+k'}} \|_2^2 < \| \mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}} \|_2^2. \qquad (22)$$

We first explain why (22) is a sufficient condition of (20). Consider $\mathbf{x}_{\Gamma^{k+k'}}$ and $\mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}}$, which are two truncated vectors of $\mathbf{x}_{\Gamma^k}$. From the definition of Γ^k_τ in (17), we have $\Gamma^k_{L-1} = \{1, 2, \dots, 2^{L-1} - 1\}$ and $|\Gamma^k \setminus \Gamma^k_{L-1}| = |\{2^{L-1}, \dots, \gamma\}| = \gamma - 2^{L-1} + 1$. Thus, we can obtain an alternative form of (20) as $N^{k+k'} \leq |\Gamma^k \setminus \Gamma^k_{L-1}| - 1$. Equivalently,

$$N^{k+k'} < |\Gamma^k \setminus \Gamma^k_{L-1}|. \qquad (23)$$

Further, since $|x_i|, i = 1, 2, \dots, \gamma$, are arranged in descending order of their magnitudes (i.e., $|x_1| \geq \cdots \geq |x_\gamma|$), $\mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}} = (x_{2^{L-1}}, \dots, x_\gamma)'$ consists of $\gamma - 2^{L-1} + 1$ smallest elements (in magnitude) of $\mathbf{x}_{\Gamma^k}$. Then it is easy to see that (22) ensures (23), and thus (22) becomes a sufficient condition of (20).

Now, what remains is the proof of (22). To the end, we build an upper bound for $\|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2$ and a lower bound for $\|\mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}}\|_2^2$, and then relate them to get a condition for (22).

Proposition 1: Let $\eta := \exp(-\frac{c(1 - \delta_{|T \cup T^{k+\lfloor c\gamma \rfloor}|})}{4(1+\delta_1)})$. Then for any σ such that $\sigma\eta < 1$, we have

$$\|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2 \leq \frac{\|\mathbf{r}^{k+k'}\|_2^2}{1 - \delta_{|T \cup T^{k+\lfloor c\gamma \rfloor}|}}, \qquad (24)$$

$$\|\mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}}\|_2^2 > \frac{\sigma(1 - \sigma\eta)\|\mathbf{r}^{k+k'}\|_2^2}{(1 + \delta_\gamma)(1 - \eta)}. \qquad (25)$$

The proof is left to Appendix A. From Proposition 1, it is clear that (22) holds true whenever

$$\frac{1}{1 - \delta_{|T \cup T^{k+\lfloor c\gamma \rfloor}|}} \leq \frac{\sigma(1 - \sigma\eta)}{(1 + \delta_\gamma)(1 - \eta)}. \qquad (26)$$

Noting that $\sigma(1 - \sigma\eta) = \frac{1}{\eta}(-(\sigma\eta - \frac{1}{2})^2 + \frac{1}{4})$, by choosing $\sigma\eta = \frac{1}{2}$ we have $\sigma(1 - \sigma\eta) = \frac{1}{4\eta}$, and hence (26) becomes

$$4\eta(1 - \eta)(1 + \delta_\gamma) \leq 1 - \delta_{|T \cup T^{k+\lfloor c\gamma \rfloor}|}.$$

Equivalently,

$$c \geq -\frac{4(1 + \delta_1)}{1 - \delta_s} \log\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\delta_\gamma + \delta_s}{1 + \delta_\gamma}} \right)$$

where $s := |T \cup T^{k+\lfloor c\gamma \rfloor}|$.

Thus completes the proof.

IV. Conclusion

In this paper, we have investigated the recovery performance of OMP when the number of iterations exceeds the sparsity K of input signals. We have established a lower bound on the number of iterations of OMP, expressed in terms of RIC, that guarantees the exact recovery of sparse signals. Our result demonstrates that OMP can accurately recover any K-sparse signal in $\lceil 2.8K \rceil$ iterations with the RIC being an absolute constant. This result bridges the gap between the recent result of [21] and the fundamental limit of the OMP algorithm at which exact sparse recovery cannot be uniformly ensured. Considering that a large number of iterations leads to a high computational complexity and also imposes a strict limitation on the sparsity level $\left(K \leq \frac{m}{c} \right)$, the reduction on the number of iterations offers computational benefits as well as relaxations in the measurement size and the sparsity range of underlying signals to be recovered.

APPENDIX A
PROOF OF PROPOSITION 1

A. Upper bound for $\|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2$

Since $\mathbf{r}^{k+k'} = \mathbf{y} - \mathbf{\Phi}\widehat{\mathbf{x}}^{k+k'} = \mathbf{\Phi}\left(\mathbf{x} - \widehat{\mathbf{x}}^{k+k'}\right)$,

$$\|\mathbf{r}^{k+k'}\|_2^2 \overset{(a)}{\geq} \left(1 - \delta_{|T \cup T^{k+k'}|}\right)\|\mathbf{x} - \widehat{\mathbf{x}}^{k+k'}\|_2^2$$

$$\overset{(b)}{\geq} \left(1 - \delta_{|T \cup T^{k+k'}|}\right)\|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2$$

$$\overset{(c)}{\geq} \left(1 - \delta_{|T \cup T^{k+\lfloor c\gamma \rfloor}|}\right)\|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2,$$

where (a) is from the RIP, (b) is due to $\|\mathbf{x} - \widehat{\mathbf{x}}^{k+k'}\|_2^2 \geq \|\mathbf{x}_{T \setminus T^{k+k'}}\|_2^2 = \|\mathbf{x}_{\Gamma^{k+k'}}\|_2^2$, and (c) is because

$$k + k' \overset{(d)}{=} k + \lceil c2^{L-1} \rceil - 1$$

$$\overset{(e)}{\leq} k + \lceil c\gamma \rceil - 1$$

$$\leq k + \lfloor c\gamma \rfloor, \tag{A.1}$$

where (d) is from (19) and (e) is because $2^{L-1} \leq \gamma$.

B. Lower Bound for $\|\mathbf{x}_{\Gamma^k \setminus \Gamma^k_{L-1}}\|_2^2$

Compared to the upper bound analysis, the lower bound analysis requires a little more effort. The following lemmas will be used in our analysis.

Lemma 1: The residual $\mathbf{r}^k$ satisfies

$$\|\mathbf{r}^k\|_2^2 \leq (1 + \delta_{N^k}) \|\mathbf{x}_{\Gamma^k}\|_2^2.$$

Proof: From Table I, the residual of OMP can be expressed as $\mathbf{r}^k = \mathbf{y} - \mathbf{\Phi}\widehat{\mathbf{x}}^k = \mathbf{y} - \mathcal{P}_{T^k}\mathbf{y}$. Since $\mathbf{r}^k \perp \mathcal{P}_{T^k}\mathbf{y}$,

$$\|\mathbf{r}^k\|_2^2 = \|\mathbf{y}\|_2^2 - \|\mathcal{P}_{T^k}\mathbf{y}\|_2^2$$

$$\overset{(a)}{\leq} \|\mathbf{y}\|_2^2 - \|\mathcal{P}_{T^k \cap T}\mathbf{y}\|_2^2$$

$$\overset{(b)}{=} \|\mathbf{y} - \mathcal{P}_{T^k \cap T}\mathbf{y}\|_2^2, \tag{A.2}$$

where (a) is due to $(T^k \cap T) \subseteq T^k$ so that $span(\mathbf{\Phi}_{T^k}) \supseteq span(\mathbf{\Phi}_{T^k \cap T})$ and $\|\mathcal{P}_{T^k}\mathbf{y}\|_2^2 \geq |\mathcal{P}_{T^k \cap T}\mathbf{y}\|_2^2$ and (b) is because $(\mathbf{y} - \mathcal{P}_{T^k \cap T}\mathbf{y}) \perp \mathcal{P}_{T^k \cap T}\mathbf{y}$. Noting that $\mathcal{P}_{T^k \cap T}\mathbf{y}$ is the projection of $\mathbf{y}$ onto $span(\mathbf{\Phi}_{T^k \cap T})$, we have

$$\|\mathbf{y} - \mathcal{P}_{T^k \cap T}\mathbf{y}\|_2^2 = \min_{\mathbf{u}:supp(\mathbf{u})=T^k \cap T} \|\mathbf{y} - \mathbf{\Phi}\mathbf{u}\|_2^2$$

$$\leq \|\mathbf{y} - \mathbf{\Phi}_{T^k \cap T}\mathbf{x}_{T^k \cap T}\|_2^2 = \|\mathbf{\Phi}_{\Gamma^k}\mathbf{x}_{\Gamma^k}\|_2^2$$

$$\leq (1 + \delta_{N^k}) \|\mathbf{x}_{\Gamma^k}\|_2^2,$$

where the last inequality is from the RIP $(|\Gamma^k| = N^k)$.

This, together with (A.2), establishes the lemma. $\square$

The second lemma provides a lower bound for $\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2$ in the $(l+1)$-th $(l \geq k)$ iteration of OMP.

Lemma 2: For given integer $l \geq k$, we have

$$\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2$$

$$\geq \frac{1 - \delta_{|\Gamma^k_\tau \cup T^l|}}{(1 + \delta_1) |\Gamma^k_\tau|} \left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma^k_\tau}\mathbf{x}_{\Gamma^k \setminus \Gamma^k_\tau}\|_2^2 \right), \tag{A.3}$$

where $\tau = 1, \ldots, \lfloor \log_2 N^k \rfloor + 1$.

Proof: The proof consists of two parts. First, we show that the residual power difference of OMP satisfies

$$\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq \frac{\|\mathbf{\Phi}'\mathbf{r}^l\|_\infty^2}{1 + \delta_1}. \tag{A.4}$$

Second, we show that

$$\|\mathbf{\Phi}'\mathbf{r}^l\|_\infty^2 \geq \frac{1 - \delta_{|\Gamma^k_\tau \cup T^l|}}{|\Gamma^k_\tau|} \left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma^k_\tau}\mathbf{x}_{\Gamma^k \setminus \Gamma^k_\tau}\|_2^2 \right). \tag{A.5}$$

The lemma is established by combining (A.4) and (A.5).

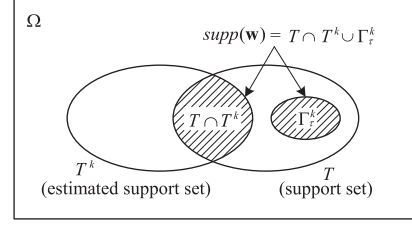Fig. 4. Illustration of $supp(\mathbf{w})$.

1) Proof of (A.4): Observe that the residual of OMP satisfies

$$\mathbf{r}^l - \mathbf{r}^{l+1} = (\mathbf{y} - \mathcal{P}_{T^l}\mathbf{y}) - (\mathbf{y} - \mathcal{P}_{T^{l+1}}\mathbf{y})$$

$$\overset{(a)}{=} (\mathcal{P}_{T^{l+1}} - \mathcal{P}_{T^{l+1}}\mathcal{P}_{T^l}) \mathbf{y}$$

$$= \mathcal{P}_{T^{l+1}} (\mathbf{y} - \mathcal{P}_{T^l}\mathbf{y}) = \mathcal{P}_{T^{l+1}}\mathbf{r}^l,$$

where (a) is because $span(\mathbf{\Phi}_{T^l}) \subseteq span(\mathbf{\Phi}_{T^{l+1}})$ so that $\mathcal{P}_{T^l}\mathbf{y} = \mathcal{P}_{T^{l+1}}\mathcal{P}_{T^l}\mathbf{y}$. Recalling that t^{l+1} is the index chosen at the $(l+1)$-th iteration and $T^{l+1} = T^l \cup t^{l+1}$, we have $span(\mathbf{\Phi}_{T^{l+1}}) \supseteq span(\phi_{t^{l+1}})$, and hence

$$\|\mathbf{r}^l - \mathbf{r}^{l+1}\|_2^2 = \|\mathcal{P}_{T^{l+1}}\mathbf{r}^l\|_2^2 \geq \|\mathcal{P}_{t^{l+1}}\mathbf{r}^l\|_2^2.$$

Further, noting that $\|\mathbf{r}^l - \mathbf{r}^{l+1}\|_2^2 = \|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2$ and that $\mathcal{P}_{t^{l+1}} = \mathcal{P}'_{t^{l+1}} = \left(\phi^\dagger_{t^{l+1}}\right)' \phi'_{t^{l+1}}$, we have

$$\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq \|\mathcal{P}_{t^{l+1}}\mathbf{r}^l\|_2^2 = \|\left(\phi^\dagger_{t^{l+1}}\right)' \phi'_{t^{l+1}}\mathbf{r}^l\|_2^2$$

$$\overset{(a)}{=} (\phi'_{t^{l+1}}\mathbf{r}^l)^2 \|\phi^\dagger_{t^{l+1}}\|_2^2$$

$$\overset{(b)}{\geq} \frac{(\phi'_{t^{l+1}}\mathbf{r}^l)^2 \|\phi^\dagger_{t^{l+1}} \phi_{t^{l+1}}\|_2^2}{\|\phi_{t^{l+1}}\|_2^2}$$

$$\overset{(c)}{\geq} \frac{(\phi'_{t^{l+1}}\mathbf{r}^l)^2}{1 + \delta_1} = \frac{\|\mathbf{\Phi}'\mathbf{r}^l\|_\infty^2}{1 + \delta_1},$$

where (a) is because $\phi'_{t^{l+1}}\mathbf{r}^l$ is a scalar, (b) is from the norm inequality, and (c) is due to the RIP.

2) Proof of (A.5): First, let $\mathbf{w} \in \mathbb{R}^n$ be the vector such that

$$\mathbf{w}_S = \begin{cases} \mathbf{x}_S & S = T \cap T^k \cup \Gamma^k_\tau, \\ \mathbf{0} & S = \Omega \setminus \left(T \cap T^k \cup \Gamma^k_\tau\right), \end{cases}$$

where $\tau \in \{1, 2, \ldots, \lfloor \log_2 N^k \rfloor + 1\}$. An illustration of $supp(\mathbf{w})$ is provided in Fig. 4. Then, by noting that $supp(\mathbf{\Phi}'\mathbf{r}^l) = \Omega \setminus T^l$, we have

$$\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} \rangle = \langle (\mathbf{\Phi}'\mathbf{r}^l)_{\Omega \setminus T^l}, \mathbf{w}_{\Omega \setminus T^l} \rangle$$

$$\overset{(a)}{\leq} \|(\mathbf{\Phi}'\mathbf{r}^l)_{\Omega \setminus T^l}\|_\infty \|\mathbf{w}_{\Omega \setminus T^l}\|_1$$

$$\overset{(b)}{\leq} \sqrt{|\Omega \setminus T^l|} \|\mathbf{\Phi}'\mathbf{r}^l\|_\infty \|\mathbf{w}_{\Omega \setminus T^l}\|_2$$

$$\overset{(c)}{\leq} \sqrt{|\Gamma^k_\tau|} \|\mathbf{\Phi}'\mathbf{r}^l\|_\infty \|\mathbf{w}_{\Omega \setminus T^l}\|_2, \tag{A.6}$$

where (a) is from Hölder's inequality, (b) follows from the norm inequality $\left(\|\mathbf{u}\|_1 \leq \sqrt{\|\mathbf{u}\|_0}\|\mathbf{u}\|_2\right)$, and (c) is because some indices in Γ^k_τ may be identified in iterations $k+1, \ldots, l$ so that $|\Omega \setminus T^l| \leq |\Gamma^k_\tau|$. Since $supp(\widehat{\mathbf{x}}^l) = T^l$ and $supp(\mathbf{\Phi}'\mathbf{r}^l) = \Omega \setminus$

T^l, it is clear that $\langle \mathbf{\Phi}'\mathbf{r}^l, \widehat{\mathbf{x}}^l \rangle = \mathbf{0}$ and $\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} \rangle = \langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle$. Thus, (A.6) can be rewritten as

$$\|\mathbf{\Phi}'\mathbf{r}^l\|_\infty \geq \frac{\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} \rangle}{\sqrt{|\Gamma_\tau^k|}\|\mathbf{w}_{\Omega \setminus T^l}\|_2} = \frac{\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle}{\sqrt{|\Gamma_\tau^k|}\|\mathbf{w}_{\Omega \setminus T^l}\|_2}. \quad \text{(A.7)}$$

Next, we build a lower bound for $\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle$. Note that

$$2\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle = 2\langle \mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right), \mathbf{r}^l \rangle$$

$$= \|\mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^l - \mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2^2$$

$$\overset{(a)}{=} \|\mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}\left(\mathbf{x} - \mathbf{w} \right)\|_2^2$$

$$= \|\mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2^2 + \|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2, \quad \text{(A.8)}$$

where (a) holds because $\mathbf{r}^l + \mathbf{\Phi}\widehat{\mathbf{x}}^l = \mathbf{y} = \mathbf{\Phi}\mathbf{x}$. When $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \geq 0$,[3] using $a^2 + b^2 \geq 2ab$ in (A.8) yields

$$\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle \geq \|\mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2 \sqrt{\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2}.$$

Moreover, since $supp\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right) = \left(T \cap T^k \cup \Gamma_\tau^k \right) \cup T^l \subseteq \Gamma_\tau^k \cup T^l$,

$$\|\mathbf{\Phi}\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)\|_2 \overset{(a)}{\geq} \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}\|\mathbf{w} - \widehat{\mathbf{x}}^l\|_2$$

$$\geq \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}\|\left(\mathbf{w} - \widehat{\mathbf{x}}^l \right)_{\Omega \setminus T^l}\|_2$$

$$\overset{(b)}{=} \sqrt{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}\|\mathbf{w}_{\Omega \setminus T^l}\|_2,$$

where (a) is from the RIP and (b) is uses $\left(\widehat{\mathbf{x}}^l \right)_{\Omega \setminus T^l} = \mathbf{0}$. Hence,

$$\langle \mathbf{\Phi}'\mathbf{r}^l, \mathbf{w} - \widehat{\mathbf{x}}^l \rangle \geq \|\mathbf{w}_{\Omega \setminus T^l}\|_2$$

$$\times \sqrt{\left(1 - \delta_{|\Gamma_\tau^k \cup T^l|} \right)\left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right)}. \quad \text{(A.9)}$$

Finally, plugging (A.9) into (A.7), we obtain (A.5). $\qquad \square$

The consequence of Lemma 2 is the following lemma, which is crucial for proving the lower bound for $\|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2$.

Lemma 3: For any integer $l' \geq l \geq k$ and $\tau \in \{1, \ldots, \lfloor \log_2 N^k \rfloor + 1\}$, the residual of OMP satisfies

$$\|\mathbf{r}^{l'}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2$$

$$\leq C_{\tau,l,l'}\left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right)$$

where $C_{\tau,l,l'} = \exp\left(-\frac{\left(1 - \delta_{|\Gamma_\tau^k \cup T^{l'-1}|} \right)(l'-l)}{(1+\delta_1)|\Gamma_\tau^k|} \right)$.

Proof: Using $a \geq 1 - e^{-a}$ with $a = \frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{(1+\delta_1)|\Gamma_\tau^k|}$, we have

$$\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{(1+\delta_1)|\Gamma_\tau^k|} \geq 1 - \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{(1+\delta_1)|\Gamma_\tau^k|} \right) > 0. \quad \text{(A.10)}$$

Then, we can rewrite (A.3) in Lemma 2 as[4]

$$\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq \left(1 - \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{(1+\delta_1)|\Gamma_\tau^k|} \right) \right)$$

$$\times \left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right). \quad \text{(A.11)}$$

Subtracting both sides of (A.11) by $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{T \setminus \Gamma_\tau^k}\mathbf{x}_{T \setminus \Gamma_\tau^k}\|_2^2$, we have

$$\|\mathbf{r}^{l+1}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^l|}}{(1+\delta_1)|\Gamma_\tau^k|} \right)$$

$$\times \left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right), \quad \text{(A.12)}$$

and thus

$$\|\mathbf{r}^{l+2}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^{l+1}|}}{(1+\delta_1)|\Gamma_\tau^k|} \right)$$

$$\times \left(\|\mathbf{r}^{l+1}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right), \quad \text{(A.13)}$$

$$\vdots$$

$$\|\mathbf{r}^{l'}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \leq \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^{l'-1}|}}{(1+\delta_1)|\Gamma_\tau^k|} \right)$$

$$\times \|\mathbf{r}^{l'-1}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \Big). \quad \text{(A.14)}$$

From (A.12)–(A.14), we further have

$$\|\mathbf{r}^{l'}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2$$

$$\leq \prod_{i=l}^{l'-1} \exp\left(-\frac{1 - \delta_{|\Gamma_\tau^k \cup T^i|}}{(1+\delta_1)|\Gamma_\tau^k|} \right) \left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right)$$

$$\leq C_{\tau,l,l'}\left(\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \right),$$

where $C_{\tau,l,l'} = \exp\left(-\frac{\left(1 - \delta_{|\Gamma_\tau^k \cup T^{l'-1}|} \right)(l'-l)}{(1+\delta_1)|\Gamma_\tau^k|} \right)$. $\qquad \square$

Now, we are ready to identify the lower bound for $\|\mathbf{x}_{\Gamma^k \setminus \Gamma_{L-1}^k}\|_2^2$. First, let $k_0 = k$ and

$$k_i = k + \sum_{\tau=1}^{i} \left\lceil \frac{c}{4}|\Gamma_\tau^k| \right\rceil, \quad i = 1, \ldots, L,$$

where Γ_τ^k and L are defined in (17) and (18a)–(18d), respectively. Then, by applying Lemma 3 with $l' = k_i$ and $l = k_{i-1}$, $i = 1, \ldots, L$, we have

$$\|\mathbf{r}^{k_i}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_i^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_i^k}\|_2^2$$

$$\leq C_{i,k_{i-1},k_i}\left(\|\mathbf{r}^{k_{i-1}}\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_i^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_i^k}\|_2^2 \right), \quad \text{(A.15)}$$

where $C_{i,k_{i-1},k_i} = \exp\left(-\frac{\left(1 - \delta_{|\Gamma_i^k \cup T^{k_i-1}|} \right)(k_i - k_{i-1})}{(1+\delta_1)|\Gamma_i^k|} \right)$.

---

[3]We only need to consider $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \geq 0$ because if $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 < 0$, (A.5) holds trivially since $\|\mathbf{\Phi}'\mathbf{r}^l\|_\infty \geq 0$.

[4]Note that $\|\mathbf{r}^l\|_2^2 - \|\mathbf{r}^{l+1}\|_2^2 \geq 0$ due to orthogonal projection at each iteration of OMP. When $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 \geq 0$, (A.11) directly follows from (A.3) and (A.10). When $\|\mathbf{r}^l\|_2^2 - \|\mathbf{\Phi}_{\Gamma^k \setminus \Gamma_\tau^k}\mathbf{x}_{\Gamma^k \setminus \Gamma_\tau^k}\|_2^2 < 0$, (A.11) holds trivially because in this case its right-hand side is negative.

We next build an upper bound for the constant C_{i,k_{i-1},k_i} in (A.15). Since

$$\frac{k_i - k_{i-1}}{|\Gamma_i^k|} = \frac{\left\lceil \frac{c}{4}|\Gamma_i^k| \right\rceil}{|\Gamma_i^k|} \geq \frac{c}{4}, \quad i = 1, 2, \ldots, L,$$

and also noting that $k_i - 1 \leq k_L$, we can rewrite C_{i,k_{i-1},k_i} as

$$C_{i,k_{i-1},k_i} \leq \exp\left(-\frac{c\left(1 - \delta_{|T \cup T^{k_L}|}\right)}{4(1 + \delta_1)} \right), \qquad (A.16)$$

where we have used monotonicity of the RIC ($|\Gamma_i^k \cup T^{k_i - 1}| \leq |\Gamma^k \cup T^{k_i - 1}| \leq |T \cup T^{k_L}|$ for $i = 1, \ldots, L$).

Further, we find an upper bound for k_L in (A.16). Recalling from (17) that $|\Gamma_\tau^k| \leq 2^\tau - 1$ for $\tau = 1, 2, \ldots, L$, we have

$$k_L = k + \sum_{\tau=1}^{L} \left\lceil \frac{c}{4}|\Gamma_\tau^k| \right\rceil \leq k + \sum_{\tau=1}^{L} \left\lceil \frac{c}{4}(2^\tau - 1) \right\rceil. \quad (A.17)$$

Since (10) directly implies that $c \geq 4\log 2 \geq 2$ under the RIP assumption in Theorem 1, one can show that (see Appendix B)

$$\sum_{\tau=1}^{L} \left\lceil \frac{c}{4}(2^\tau - 1) \right\rceil \leq \left\lceil c2^{L-1} \right\rceil - 1, \qquad (A.18)$$

Hence, (A.17) becomes

$$k_L \leq k + \left\lceil c2^{L-1} \right\rceil - 1 = k + k' \overset{(a)}{\leq} k + \lfloor c\gamma \rfloor, \quad (A.19)$$

where (a) is from (A.1). This, together with (A.16), implies

$$C_{i,k_{i-1},k_i} \leq \exp\left(-\frac{c\left(1 - \delta_{|T \cup T^{k + \lfloor c\gamma \rfloor}|}\right)}{4(1 + \delta_1)} \right). \qquad (A.20)$$

Now we can construct an upper bound for $\|\mathbf{r}^{k_i}\|_2^2$ using (A.15) and (A.20). By denoting

$$\eta := \exp\left(-\frac{c\left(1 - \delta_{|T \cup T^{k + \lfloor c\gamma \rfloor}|}\right)}{4(1 + \delta_1)} \right),$$

we rewrite (A.15) as

$$\|\mathbf{r}^{k_i}\|_2^2 \leq \eta \|\mathbf{r}^{k_{i-1}}\|_2^2 + (1 - \eta)\left\|\mathbf{\Phi}_{\Gamma^k \backslash \Gamma_i^k} \mathbf{x}_{\Gamma^k \backslash \Gamma_i^k}\right\|_2^2,$$
$$i = 1, \ldots, L.$$

Some additional manipulations yield the following result,

$$\|\mathbf{r}^{k_L}\|_2^2 \leq \eta^L \|\mathbf{r}^k\|_2^2 + (1 - \eta)\sum_{\tau=1}^{L} \eta^{L-\tau} \left\|\mathbf{\Phi}_{\Gamma^k \backslash \Gamma_\tau^k} \mathbf{x}_{\Gamma^k \backslash \Gamma_\tau^k}\right\|_2^2$$

$$\overset{(a)}{\leq} \eta^L \|\mathbf{r}^k\|_2^2 + (1 - \eta)(1 + \delta_\gamma)\sum_{\tau=1}^{L} \eta^{L-\tau} \left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_\tau^k}\right\|_2^2$$

$$\overset{(b)}{\leq} \left(\eta^L \left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_0^k}\right\|_2^2 + (1 - \eta)\sum_{\tau=1}^{L} \eta^{L-\tau} \left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_\tau^k}\right\|_2^2 \right)(1 + \delta_\gamma)$$

$$\overset{(c)}{\leq} \left((\sigma\eta)^L + (1 - \eta)\sum_{\tau=1}^{L} (\sigma\eta)^{L-\tau} \right)\frac{(1 + \delta_\gamma)\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k}\right\|_2^2}{\sigma}$$

$$\overset{(d)}{<} \left((1 - \eta)\sum_{\tau=L}^{\infty} (\sigma\eta)^\tau + (1 - \eta)\sum_{\tau=0}^{L-1} (\sigma\eta)^\tau \right)$$

$$\times \frac{(1 + \delta_\gamma)\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k}\right\|_2^2}{\sigma}$$

$$= \frac{(1 + \delta_\gamma)(1 - \eta)\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k}\right\|_2^2}{\sigma(1 - \sigma\eta)},$$

where (a) is due to the RIP ($|\Gamma^k \backslash \Gamma_\tau^k| \leq |\Gamma^k| = \gamma$ for $\tau = 1, \ldots, L$), (b) is from Lemma 1 (i.e., $\|\mathbf{r}^k\|_2^2 \leq (1 + \delta_{N^k})\|\mathbf{x}_{\Gamma^k}\|_2^2 = (1 + \delta_\gamma)\|\mathbf{x}_{\Gamma^k \backslash \Gamma_0^k}\|_2^2$), (c) follows from

$$\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_\tau^k}\right\|_2^2 \leq \sigma^{L-1-\tau}\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k}\right\|_2^2, \quad \tau = 0, 1, \ldots, L,$$

which is a direct consequence of (18a)–(18d), and (d) is because $\sigma > 1$ and $\eta < 1$ so that $(\sigma\eta)^L < \frac{1-\eta}{1-\sigma\eta} \cdot (\sigma\eta)^L = (1 - \eta)\sum_{\tau=L}^{\infty} (\sigma\eta)^\tau$ when $0 < \sigma\eta < 1$.

Finally, since $k_L \leq k + k'$ by (A.19), and also noting that the residual power of OMP is always non-increasing, we have

$$\|\mathbf{r}^{k+k'}\|_2^2 \leq \|\mathbf{r}^{k_L}\|_2^2 \leq \frac{(1 + \delta_\gamma)(1 - \eta)\left\|\mathbf{x}_{\Gamma^k \backslash \Gamma_{L-1}^k}\right\|_2^2}{\sigma(1 - \sigma\eta)},$$

which is the desired result.

APPENDIX B
PROOF OF (A.18)

Proof: We prove (A.18) by mathematical induction on L. First, when $L = 1$, (A.18) becomes $\left\lceil \frac{c}{4} \right\rceil \leq \lceil c \rceil - 1$, which is simply true since $c \geq 2$. Next, we assume that (A.18) holds up to an integer ℓ so that

$$\sum_{\tau=1}^{\ell} \left\lceil \frac{c}{4}(2^\tau - 1) \right\rceil \leq \left\lceil c2^{\ell-1} \right\rceil - 1. \qquad (B.1)$$

Then if $L = \ell + 1$,

$$\sum_{\tau=1}^{\ell+1} \left\lceil \frac{c}{4}(2^\tau - 1) \right\rceil \leq \left\lceil c2^{\ell-1} \right\rceil - 1 + \left\lceil \frac{c}{4}(2^{\ell+1} - 1) \right\rceil$$

$$\overset{(a)}{\leq} \left\lceil c2^{\ell-1} \right\rceil + \left\lceil c2^{\ell-1} - \tfrac{1}{2} \right\rceil - 1$$

$$\overset{(b)}{\leq} \left\lceil c2^{\ell} \right\rceil - 1,$$

where (a) is because $c \geq 2$ and (b) uses Hermite's identity [38] ($\lceil ax \rceil = \lceil x \rceil + \left\lceil x - \frac{1}{a} \right\rceil \cdots + \left\lceil x - \frac{a-1}{a} \right\rceil$ with $a = 2$ and $x = c2^{\ell-1}$), which completes the proof. $\square$

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[3] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[4] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, Mar. 2009.

[5] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[6] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.

[7] C. Soussen, R. Gribonval, J. Idier, and C. Herzet, "Joint k-step analysis of orthogonal matching pursuit and orthogonal least squares," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 3158–3174, May 2013.

[8] L. Chen and Y. Gu, "Oracle-order recovery performance of greedy pursuits with replacement against general perturbations," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4625–4636, Sep. 2013.

[9] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 1993, vol. 1, pp. 40–44.

[10] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[11] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[12] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[13] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[14] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.

[15] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit algorithm," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3654–3656, Jun. 2012.

[16] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4973–4976, Sep. 2012.

[17] J. Wen, X. Zhu, and D. Li, "Improved bounds on restricted isometry constant for orthogonal matching pursuit," *Electron. Lett.*, vol. 49, no. 23, pp. 1487–1489, 2013.

[18] M. Yang and F. Hoog, "New coherence and RIP analysis for weak orthogonal matching pursuit," in *Proc. IEEE Workshop Statist. Signal Process. (SSP)*, Jun. 2014, pp. 376–379.

[19] L. Chang and J. Wu, "An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5702–5715, Sep. 2014.

[20] Q. Mo, "A sharp restricted isometry constant bound of orthogonal matching pursuit," 2015, DOI: arXiv:1501.01708.

[21] T. Zhang, "Sparse recovery with orthogonal matching pursuit under RIP," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6215–6221, Sep. 2011.

[22] E. D. Livshits, "On the efficiency of the orthogonal matching pursuit in compressed sensing," *Sbornik: Mathematics*, vol. 203, no. 2, p. 183, 2012.

[23] S. Foucart, "Stability and robustness of weak orthogonal matching pursuits," *Recent Advances in Harmonic Analysis and Applications*, New York, NY, USA: Springer, 2013, pp. 395–405.

[24] E. D. Livshitz and V. N. Temlyakov, "Sparse approximation and recovery by greedy algorithms," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3989–4000, Jul. 2014.

[25] E. Liu and V. N. Temlyakov, "The orthogonal super greedy algorithm and applications in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2040–2047, Apr. 2012.

[26] S. Huang and J. Zhu, "Recovery of sparse signals using OMP and its variants: Convergence analysis based on RIP," *Inverse Problems*, vol. 27, no. 3, p. 035003, 2011.

[27] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6202–6216, Dec. 2012.

[28] J. Wang, S. Kwon, P. Li, and B. Shim, "Recovery of sparse signals via generalized orthogonal matching pursuit: A new analysis," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 1076–1089, Feb. 2016.

[29] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, pp. 129–159, 2001.

[30] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Construct. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[31] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton, NJ, USA: Princeton Univ. Press, 2005.

[32] C. De Boor, "An empty exercise," *ACM Signum Newslett.*, vol. 25, no. 4, pp. 2–6, Nov. 1990.

[33] J. Ding, L. Chen, and Y. Gu, "Perturbation analysis of orthogonal matching pursuit," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 398–410, Jan. 2013.

[34] Y. Dong, P. Fan, and K. Ben Letaief, "High-speed railway wireless communications: Efficiency versus fairness," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 925–930, Feb. 2014.

[35] J. Wang, "Support recovery with orthogonal matching pursuit in the presence of noise," *IEEE Trans. Signal Process.*, vol. 63, no. 21, pp. 5868–5877, Nov. 2015.

[36] J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo, "Sharp conditions for exact support recovery of sparse signals with noise via OMP," 2015, DOI: arXiv:1512.07248.

[37] J. Wen, D. Li, and F. Zhu, "Stable recovery of sparse signals via ℓ_p-minimization," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 1, pp. 161–176, 2015.

[38] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*. Reading, MA, USA: Addison-Wesley, 1989.
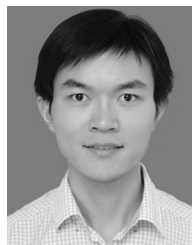
Jian Wang (S'11) received the B.S. degree in material engineering and the M.S. degree in information and communication engineering from Harbin Institute of Technology, China, and the Ph.D. degree in electrical and computer engineering from Korea University, Korea, in 2006, 2009, and 2013, respectively. From 2013 to 2015, he held positions as Postdoctoral Research Associate at Department of Statistics & Biostatistics, Department of Computer Science, Rutgers University, Piscataway, NJ, USA, and Department of Electrical & Computer Engineering, Duke University, Durham, NC 27708, USA. He is currently a research assistant professor in Seoul National University, Seoul 151-744, Korea. His research interests include compressed sensing, signal processing for wireless communications, and statistical learning.

Byonghyo Shim (SM'09) received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Korea, in 1995 and 1997, respectively. He received the M.S. degree in mathematics and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign (UIUC), USA, in 2004 and 2005, respectively.

From 1997 and 2000, he was with the Department of Electronics Engineering, Korean Air Force Academy as an Officer (First Lieutenant) and an Academic Full-time Instructor. From 2005 to 2007, he was with Qualcomm Inc., San Diego, CA, USA, as a Staff Engineer. From 2007 to 2014, he was with the School of Information and Communication, Korea University, Seoul, as an Associate Professor. Since September 2014, he has been with the Department of Electrical and Computer Engineering, Seoul National University, where he is presently an Associate Professor. His research interests include wireless communications, statistical signal processing, estimation and detection, compressive sensing, and information theory.

Dr. Shim was the recipient of the 2005 M. E. Van Valkenburg Research Award from the Electrical and Computer Engineering Department of the University of Illinois and 2010 Hadong Young Engineer Award from IEIE. He is currently an Associate Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, *Journal of Communications and Networks*, and a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC).