On the Recovery Limit of Sparse Signals using Orthogonal Matching Pursuit

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

***Abstract*—Orthogonal matching pursuit (OMP) is a greedy search algorithm popularly being used for the recovery of compressive sensed sparse signals. In this paper, we show that if the isometry constant δ_{K+1} of the sensing matrix $\mathbf{\Phi}$ satisfies $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$ then the OMP algorithm can perfectly recover K-sparse signals from the compressed measurements $\mathbf{y} = \mathbf{\Phi x}$. Our bound offers a substantial improvement over the recent result of Davenport and Wakin, and also closes gap between the recovery bound and fundamental limit over which the perfect recovery of the OMP cannot be guaranteed.**

***Index Terms*—Compressed sensing (CS), sparse signal, orthogonal matching pursuit (OMP), restricted isometry property (RIP).**

I. Introduction

A. Orthogonal Matching Pursuit

As a sampling paradigm guaranteeing the reconstruction of sparse signal with sampling rate significantly lower than the Nyquist rate, compressive sensing (CS) has received considerable attention in recent years [1]–[9]. For a given matrix $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ ($n > m$), the CS recovery algorithm generates an estimate of K-sparse vector $\mathbf{x} \in \mathbb{R}^n$ from a set of linear measurements

$$\mathbf{y} = \mathbf{\Phi x}. \tag{1}$$

Although the system is under-determined and hence the inverse problem is in general ill-posed, due to the prior information of signal sparsity, $\mathbf{x}$ can be perfectly reconstructed via properly designed recovery algorithm. Among many recovery algorithms in the literature, greedy search methods sequentially investigating the support of the sparse signal have generated a great deal of interest for practical benefits. In each iteration of greedy search algorithms, correlations between each column of $\mathbf{\Phi}$ and the modified measurements (so called *residual*) are compared to identify the element of the support. Algorithms contained in this category include orthogonal matching pursuit (OMP) [1], regularized orthogonal matching pursuit (ROMP) [10], stagewise orthogonal matching pursuit (StOMP) [11], and compressive sampling matching pursuit (CoSaMP) [9]. As a canonical method in this family, the OMP algorithm has received special attention due to its simplicity and competitive reconstruction performance. In fact, it has been shown that the OMP is reliable for reconstructing both sparse and near-sparse signals [3] while achieving the complexity expressed in a linear function of sparsity level $O(nmK)$ [12].

TABLE I
OMP Algorithm

Input:	measurements $\mathbf{y}$
	sensing matrix $\mathbf{\Phi}$
	sparsity K.
Initialize:	iteration count $k = 0$
	residual vector $\mathbf{r}^0 = \mathbf{y}$
	estimated support set $T^0 = \emptyset$.
While $k < K$	
	$k = k + 1$.
	(Identify) $t^k = \arg\max_j \lvert\langle \mathbf{r}^{k-1}, \varphi_j \rangle\rvert$.
	(Augment) $T^k = T^{k-1} \cup \{t^k\}$.
	(Estimate) $\hat{\mathbf{x}}_{T^k} = \arg\min_{\mathbf{x}} \lVert \mathbf{y} - \mathbf{\Phi}_{T^k}\mathbf{x} \rVert_2$.
	(Update) $\mathbf{r}^k = \mathbf{y} - \mathbf{\Phi}_{T^k}\hat{\mathbf{x}}_{T^k}$.
End	
Output:	$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}:\mathrm{supp}(\mathbf{x}) = T^K} \lVert \mathbf{y} - \mathbf{\Phi x} \rVert_2$.

In many CS recovery algorithms, including the OMP algorithm, a sufficient condition guaranteeing the perfect recovery of the sparse signal $\mathbf{x}$ is expressed in terms of *restricted isometry property* (RIP). A sensing matrix $\mathbf{\Phi}$ satisfies the RIP of order K if there exists a constant δ, such that [12]

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

for any K-sparse vector $\mathbf{x}$ ($\lVert \mathbf{x} \rVert_0 \le K$). In particular, the minimum of all constants δ satisfying (2) is called the isometry constant δ_K. Over the years, many efforts have been made to find out the condition (upper bound of the isometry constant) guaranteeing the exact recovery of sparse signals. In [2], Wakin and Davenport have shown that the OMP can reconstruct K-sparse signals if

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

In this paper, we present an improved condition guaranteeing the perfect recovery of the OMP algorithm. Our main conclusion states that the RIP of order $K + 1$ with $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$ is a sufficient condition guaranteeing the perfect recovery of K-sparse signals via the OMP algorithm. Our result, based on the mathematical induction, is built on an observation that the general step of the OMP process is in essence the same as the initial step since the residual is considered as a new measurement preserving the sparsity

J. Wang and B. Shim are with School of Information and Communication, Korea University, Seoul, Korea (email: {jwang,bshim}@isl.korea.ac.kr).

This work is supported by Basic Science Research Program through National Research Foundation (NRF) funded by the Ministry of Education, Science and Technology (MEST) (No. 2011-0012525) and second Brain Korea 21 project.

level of the input vector. Interestingly, therefore, a condition guaranteeing to choose the correct index in the initial step is readily extended to the general (induction) step and the resulting condition becomes equivalent to the initial condition.

Our result is formally described in the following theorem.

*Theorem 1 (***Improved recovery bound of the OMP***):* For any K-sparse vector $\mathbf{x}$, the OMP algorithm perfectly recovers $\mathbf{x}$ from $\mathbf{y} = \mathbf{\Phi x}$ if the isometry constant δ_{K+1} satisfies

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

B. The Sharpness of the Proposed Bound

In the following example, we show that the proposed bound is near optimal by showing that, even with slight relaxation of the proposed bound ($\delta_{K+1} = \frac{1}{\sqrt{K}}$), there exists a matrix for which the OMP algorithm fail to recover the K-sparse signal.

Example 1 (The OMP may fail under $\delta_{K+1} = \frac{1}{\sqrt{K}}$): Consider the problem of recovering K-sparse vector $\mathbf{x} = (1, 1, \cdots, 1, 0)' \in \mathbb{R}^{K+1}$ ($K \geq 2$) from the measurements $\mathbf{y} = \mathbf{\Phi x}$. If the sensing matrix $\mathbf{\Phi}$ satisfies

$$\mathbf{\Phi}'\mathbf{\Phi} = \begin{pmatrix} 1 & b & \cdots & b \\ b & 1 & \cdots & b \\ & & \cdots & \\ b & b & \cdots & 1 \end{pmatrix}_{(K+1)\times(K+1)}.$$

Then, one can show that eigenvalues of $\mathbf{\Phi}'\mathbf{\Phi}$ are

$$\lambda_1 = \cdots = \lambda_K = 1 - b \;\text{ and }\; \lambda_{K+1} = 1 + Kb.$$

Further, let $b = -\frac{1}{K\sqrt{K}}$, then

$$\mathbf{\Phi}'\mathbf{\Phi} = \begin{pmatrix} 1 & -\frac{1}{K\sqrt{K}} & \cdots & -\frac{1}{K\sqrt{K}} \\ -\frac{1}{K\sqrt{K}} & 1 & \cdots & -\frac{1}{K\sqrt{K}} \\ & & \cdots & \\ -\frac{1}{K\sqrt{K}} & -\frac{1}{K\sqrt{K}} & \cdots & 1 \end{pmatrix} \tag{5}$$

and hence $\lambda_1 = \lambda_2 = \cdots = \lambda_K = 1 + \frac{1}{K\sqrt{K}}$ and $\lambda_{K+1} = 1 - \frac{1}{\sqrt{K}}$. Clearly,

$$\lambda_{\min}(\mathbf{\Phi}'\mathbf{\Phi}) = 1 - \frac{1}{\sqrt{K}} \;\text{ and }\; \lambda_{\max}(\mathbf{\Phi}'\mathbf{\Phi}) = 1 + \frac{1}{K\sqrt{K}}.$$

From [12, Remark 1], we have

$$\delta_{K+1} = \max\{\lambda_{\max}(\mathbf{\Phi}'\mathbf{\Phi}) - 1, 1 - \lambda_{\min}(\mathbf{\Phi}'\mathbf{\Phi})\} = \frac{1}{\sqrt{K}}. \tag{6}$$

Also, noting that $\varphi_i = \mathbf{\Phi e}_i$ for $i \in T = \{1, 2, \cdots, K\}$, we have

$$\begin{aligned} |\langle \varphi_i, \mathbf{y}\rangle| &= |\langle \mathbf{\Phi e}_i, \mathbf{\Phi x}\rangle| & (7)\\ &= |\langle \mathbf{\Phi}'\mathbf{\Phi e}_i, \mathbf{x}\rangle| & (8)\\ &= 1 - \frac{K-1}{K\sqrt{K}} & (9) \end{aligned}$$

where (9) follows from the fact that $\mathbf{\Phi}'\mathbf{\Phi e}_i$ is the i-th column of $\mathbf{\Phi}'\mathbf{\Phi}$ (see (5)) and $\mathbf{x} = (1, 1, \cdots, 1, 0)'$ is the K-sparse vector. In a similar way, we have

$$\begin{aligned} |\langle \varphi_{K+1}, \mathbf{y}\rangle| &= |\langle \mathbf{\Phi e}_{K+1}, \mathbf{\Phi x}\rangle| & (10)\\ &= |\langle \mathbf{\Phi}'\mathbf{\Phi e}_{K+1}, \mathbf{x}\rangle| & (11)\\ &= \left|-\frac{K}{K\sqrt{K}}\right| = \frac{1}{\sqrt{K}} & (12) \end{aligned}$$

where (12) is due to the fact that $K + 1 \notin T$. The OMP algorithm fails in the first iteration if

$$|\langle \varphi_{K+1}, \mathbf{y}\rangle| \geq \max_{i\in\{1,\cdots,K+1\}} |\langle \varphi_i, \mathbf{y}\rangle|.$$

In other words,

$$|\langle \varphi_{K+1}, \mathbf{y}\rangle| \geq |\langle \varphi_i, \mathbf{y}\rangle| \tag{13}$$

for all $i \in T$.

Plugging (12) and (9) into (13), we have

$$\frac{1}{\sqrt{K}} \geq 1 - \frac{K-1}{K\sqrt{K}}.$$

One can check that this inequality holds true for $K = 2$ and hence the OMP algorithm fails to identify a correct index in this case.

C. Dimension of the Sensing Matrices ensuring Perfect Reconstruction

It is clear that if the RIP bound is relaxed, more sensing matrices satisfies the RIP condition and hence the set of sensing matrices for which exact recovery of sparse signal is possible gets larger. While observing this in a direct way is not easy (since the maximum and minimum singular values of $\binom{n}{K}$ sub-matrices of $\mathbf{\Phi}$ need to be tested), one can indirectly check it using the results derived from the RIP. To be specific, if $m \times n$ random matrix $\mathbf{\Phi}$ whose entries are independent and identically distributed Gaussian random variables $\mathcal{N}(0, 1/m)$, then $\mathbf{\Phi}$ obeys the RIP condition ($\delta_S \leq \epsilon$) with overwhelming probability under [5]

$$m \geq \frac{\rho S \log(n/S)}{\epsilon^2} \tag{14}$$

where ρ is a constant. Hence, from (4) and (14), we have

$$m \geq \rho(K+1)(\sqrt{K}+1)^2 \log\frac{n}{K+1}. \tag{15}$$

When compared with the result derived using (3), we observe that the minimum dimension of the measurements m ensuring exact reconstruction of K-sparse signal is smaller (roughly $1/9$ for large K).

II. Proof of Theorem 1

A. Notations

We summarize the notations used throughout the rest of this paper. $T = supp(\mathbf{x}) = \{\, i \mid x_i \neq 0\}$ is the set of non-zero positions in $\mathbf{x}$. For $D \subseteq \{1, 2, \cdots, n\}$, $|D|$ is the cardinality of D. $T\backslash D$ is the set of all elements contained in T but not in D. $\mathbf{\Phi}_D \in \mathbb{R}^{m\times|D|}$ is a submatrix of $\mathbf{\Phi}$ that only contains columns indexed by D. $\mathbf{x}_D \in \mathbb{R}^{|D|}$ is a restriction of the vector $\mathbf{x}$ to the elements indexed by D. $span(\mathbf{\Phi}_D)$ is the span (range) of columns in $\mathbf{\Phi}_D$. $\mathbf{\Phi}'_D$ is the transpose of the matrix $\mathbf{\Phi}_D$ and $\mathbf{\Phi}_D^{\dagger} = (\mathbf{\Phi}'_D\mathbf{\Phi}_D)^{-1}\mathbf{\Phi}'_D$ is the pseudoinverse of $\mathbf{\Phi}_D$. $\mathbf{P}_D = \mathbf{\Phi}_D\mathbf{\Phi}_D^{\dagger}$ is the orthogonal projection onto $span(\mathbf{\Phi}_D)$. $\mathbf{P}_D^{\perp} = \mathbf{I} - \mathbf{P}_D$ is the orthogonal projection onto the orthogonal complement of $span(\mathbf{\Phi}_D)$.

B. Proof of Theorem 1

Before presenting the proof of Theorem 1, we provide lemmas useful in our analysis.

Lemma 2 (Extension of Definition 2 in [12]): For $I \subset \{1, 2, \cdots, n\}$, if $\delta_{|I|} < 1$, then

$$(1 - \delta_{|I|})\|\mathbf{v}\|_2 \le \|\mathbf{\Phi}'_I\mathbf{\Phi}_I\mathbf{v}_I\|_2 \le (1 + \delta_{|I|})\|\mathbf{v}\|_2$$

holds for any vector $\mathbf{v}$ supported on I.

Lemma 3 (Lemma 1 in [12]): For disjoint sets $I_1, I_2 \subset \{1, 2, \cdots, n\}$, if $\delta_{|I_1|+|I_2|} < 1$, then

$$\left\|\mathbf{\Phi}'_{I_1}\mathbf{\Phi}\mathbf{v}\right\|_2 = \left\|\mathbf{\Phi}'_{I_1}\mathbf{\Phi}_{I_2}\mathbf{v}_{I_2}\right\|_2 \le \delta_{|I_1|+|I_2|}\|\mathbf{v}\|_2$$

holds for any vector $\mathbf{v}$ supported on I_2.

Lemma 4 (Monotonicity of the isometry constant [12]): If the sensing matrix satisfies the RIP of both orders K_1 and K_2, then

$$\delta_{K_1} \le \delta_{K_2}$$

for any $K_1 \le K_2$. This property is referred to as the monotonicity of the isometry constant.

Proof of Theorem 1

As mentioned, our proof is based on the mathematical induction. In order to prove the theorem, we first provide a condition under which the OMP algorithm chooses a correct index in the first iteration. We next show that the residual in the general iteration preserves the sparsity level of the input signal. As a result, the condition for the first iteration can be readily extended to the general iteration and the theorem is established.

Let φ_i be the i-th column of $\mathbf{\Phi}$ and t^k be the index of the column maximally correlated with the residual $\mathbf{r}^{k-1}$. Since $\mathbf{r}^{k-1} = \mathbf{y}$ for $k = 1$, we have

$$t^1 = \arg\max_i |\langle\varphi_i, \mathbf{y}\rangle|. \tag{16}$$

Lemma 5 (Initial step condition): For the K-sparse signal $\mathbf{x}$ with the support T, the index chosen in the first iteration of the OMP algorithm belongs to the support (i.e., $t^1 \in T$) if the isometry constant δ_{K+1} of a matrix $\mathbf{\Phi}$ satisfies

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

Proof: From (16), we have

$$\begin{aligned} |\langle\varphi_{t^1}, \mathbf{y}\rangle| &= \max_i |\langle\varphi_i, \mathbf{y}\rangle| && (17)\\ &= \|\mathbf{\Phi}'_T\mathbf{y}\|_\infty && (18)\\ &\ge \frac{1}{\sqrt{K}}\|\mathbf{\Phi}'_T\mathbf{y}\|_2 && (19)\end{aligned}$$

where (19) is from the norm inequality. Now that $\mathbf{y} = \mathbf{\Phi}_T\mathbf{x}_T$, we have

$$\begin{aligned} |\langle\varphi_{t^1}, \mathbf{y}\rangle| &\ge \frac{1}{\sqrt{K}}\|\mathbf{\Phi}'_T\mathbf{\Phi}_T\mathbf{x}_T\|_2 && (20)\\ &\ge \frac{1}{\sqrt{K}}(1 - \delta_K)\|\mathbf{x}_T\|_2 && (21)\end{aligned}$$

where (21) follows from Lemma 2.

Now, suppose that t^1 is not belonging to the support of $\mathbf{x}$ (i.e., $t^1 \notin T$). Then

$$\begin{aligned} |\langle\varphi_{t^1}, \mathbf{y}\rangle| &= \|\varphi'_{t^1}\mathbf{\Phi}_T\mathbf{x}_T\|_2 && (22)\\ &\le \delta_{K+1}\|\mathbf{x}_T\|_2 && (23)\end{aligned}$$

where (23) is from Lemma 3. This case, however, will never occur if

$$\frac{1}{\sqrt{K}}(1 - \delta_K)\|\mathbf{x}_T\|_2 > \delta_{K+1}\|\mathbf{x}_T\|_2 \tag{24}$$

or

$$\sqrt{K}\delta_{K+1} + \delta_K < 1. \tag{25}$$

Now that $\delta_K \le \delta_{K+1}$ by the monotonicity of the isometry constant, (25) is guaranteed when $\sqrt{K}\delta_{K+1} + \delta_{K+1} < 1$ or

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

In summary, if $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$, then $t^1 \in T$ for the first iteration of the OMP algorithm. ■

Next, in the induction step, we show that the $(k+1)$-th iteration is successful under the assumption that the former k iterations are successful ($t^{k+1} \in T \backslash T^k$). Following two lemmas validate our claim.

Lemma 6 (Inductive step): Suppose that initial k iterations ($1 \le k \le K - 1$) of the OMP algorithm are successful (i.e., $T^k = \{t^1, t^2, \cdots, t^k\} \in T$), then the $(k+1)$-th iteration is also successful (i.e., $t^{k+1} \in T$) under $\delta_{K+1} < \frac{1}{\sqrt{K}+1}$.

Proof: Recall from Table I that the residual at the k-th iteration of the OMP is expressed as

$$\mathbf{r}^k = \mathbf{y} - \mathbf{\Phi}_{T^k}\hat{\mathbf{x}}_{T^k}. \tag{26}$$

Since $\mathbf{y} = \mathbf{\Phi}_T\mathbf{x}_T$ and also $\mathbf{\Phi}_{T^k}$ is a submatrix of $\mathbf{\Phi}_T$,

$$\mathbf{r}^k \in span(\mathbf{\Phi}_T)$$

and hence $\mathbf{r}^k$ can be expressed as a linear combination of the $|T|$ $(= K)$ columns of $\mathbf{\Phi}_T$. Thus, we can express $\mathbf{r}^k$ as

$$\mathbf{r}^k = \mathbf{\Phi}\mathbf{x}^k$$

where the support of $\mathbf{x}^k$ is contained in the support of $\mathbf{x}$. In other words, $\mathbf{r}^k$ is a measurement of K-sparse signal $\mathbf{x}^k$ using the sensing matrix $\mathbf{\Phi}$. From this observation together with the Lemma 5, we conclude that if $T^k \in T$, then $t^{k+1} \in T$ under (4). ■

In order to guarantee "true success" in the $(k+1)$-iteration, other than $t^{k+1} \in T$, we should also show that already selected index cannot be selected again in the upcoming iterations (i.e., $t^{k+1} \notin T^k$).

Lemma 7: For $1 \leq k \leq K-1$, the index t^{k+1} selected at the $(k+1)$-th iteration of the OMP algorithm is not in T^k.

Proof: Recalling that the task of *estimation* step of the OMP algorithm is to solve the least squares problem, one can easily show that $\hat{\mathbf{x}}_{T^k} = \mathbf{\Phi}_{T^k}^{\dagger}\mathbf{y}$ and $\mathbf{r}^k = \mathbf{P}_{T^k}^{\perp}\mathbf{y}$. Since the residual $\mathbf{r}^k$ is orthogonal to the columns φ_i selected thus far (i.e., $\langle \varphi_i, \mathbf{r}^k \rangle = 0$ for $i \in T^k$), we conclude that $t^{k+1} \notin T^k$. ■

It is important to note that since $\mathbf{r}^k \neq \mathbf{0}$ and $\mathbf{r}^k \in span\,(\mathbf{\Phi}_T)$, there exists $i \in T$ such that $\langle \varphi_i, \mathbf{r}^k \rangle \neq 0$, which in turn implies that there should exist at least one index $t^{k+1} \in T \backslash T^k$ such that $\langle \varphi_{t^{k+1}}, \mathbf{r}^k \rangle \neq 0$.

Now we establish the Theorem 1 by simple combination of lemmas 5, 6, and 7.

Proof of main theorem: From Lemma 5, 6, and 7, we conclude that $t^{k+1} \in T \backslash T^k$ under $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ and hence all indices in T will be chosen in K iterations. Therefore, $T^K = T$ and also

$$\begin{aligned} \hat{\mathbf{x}}_{T^K} &= \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}_{T^K}\mathbf{x}\|_2 & (27)\\ &= \mathbf{\Phi}_{T^K}^{\dagger}\mathbf{y} & (28)\\ &= \mathbf{\Phi}_{T}^{\dagger}\mathbf{y} & (29)\\ &= (\mathbf{\Phi}_T^t\mathbf{\Phi}_T)^{-1}\mathbf{\Phi}_T'\mathbf{\Phi}_T\mathbf{x}_T & (30)\\ &= \mathbf{x}_T, & (31) \end{aligned}$$

which completes the proof. ■

III. Discussions

In this section, we discuss the near optimality of the proposed bound. First, we compare our bound with that of Davenport and Wakin (henceforth referred to as D-W bound) [2]. Two key ingredients leading to the improvement are 1) contradiction based construction of the success condition in the first iteration (**Lemma 5**) and 2) observation that the residual in the general iteration preserves the sparsity level of the input signal (**Lemma 6**). We next show that Lemma 6 and 7 can be used to improve the D-W bound.

A. Revisiting D-W bound

While the D-W bound of the OMP algorithm guaranteeing the perfect recovery of K-sparse signal is $\delta_{K+1} < \frac{1}{3\sqrt{K}}$, the bound obtained from Theorem 1 is $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$. Loosely speaking, the proposed bound is about three times larger than the D-W bound, in particular for large K. The reader might wonder what makes this difference. In order to understand the root cause of the difference, we need to investigate major steps in two approaches. In what follows, we use the notation $\mathbf{h} = \mathbf{\Phi}'\mathbf{y}$ in [2] and hence $\mathbf{h}_j = \langle \varphi_j, \mathbf{y} \rangle$.

Lemma 8 (Initial step condition in [2]): For the K-sparse signal $\mathbf{x}$ with the support T, the index chosen in the first iteration of the OMP algorithm belongs to the support if the isometry constant δ_{K+1} of a matrix $\mathbf{\Phi}$ satisfies

$$\delta_{K+1} < \frac{1}{2\sqrt{K}}.$$

Proof: Since $\mathbf{h}_j = \langle \varphi_j, \mathbf{y} \rangle = \langle \mathbf{\Phi}\mathbf{e}_j, \mathbf{\Phi}\mathbf{x} \rangle$ and $\mathbf{x}_j = \langle \mathbf{e}_j, \mathbf{x} \rangle$, by [2, Lemma 3.1] we have

$$|\mathbf{h}_j - \mathbf{x}_j| = |\langle \mathbf{\Phi}\mathbf{e}_j, \mathbf{\Phi}\mathbf{x} \rangle - \langle \mathbf{e}_j, \mathbf{x} \rangle| \leq \delta_{K+1}\|\mathbf{x}\|_2 \qquad (32)$$

where $j \in \{1, 2, \cdots, n\}$ and $\mathbf{e}_j$ is a unit vector supported on j.

Now consider $j = t^1$ (recall that $t^1 = \arg\max_i |\mathbf{h}_i|$ is the index chosen in the first iteration of the OMP algorithm). Suppose t^1 does not belong to the support ($t^1 \notin T$), then $\mathbf{x}_{t^1} = 0$ and hence (32) becomes

$$|\mathbf{h}_{t^1}| \leq \delta_{K+1}\|\mathbf{x}\|_2. \qquad (33)$$

This case, however, will never occur if

$$\|\mathbf{x}\|_\infty = \max_i |\mathbf{x}_i| > 2\delta_{K+1}\|\mathbf{x}\|_2. \qquad (34)$$

This claim is justified as follows. First, $\|\mathbf{x}\|_\infty > 2\delta_{K+1}\|\mathbf{x}\|_2$ implies that there exists at least one index j such that $|\mathbf{x}_j| > 2\delta_{K+1}\|\mathbf{x}\|_2$. Next, using the triangle inequality and (32), we have

$$\begin{aligned} |\mathbf{h}_j| &\geq |\mathbf{x}_j| - |\mathbf{h}_j - \mathbf{x}_j| & (35)\\ &> 2\delta_{K+1}\|\mathbf{x}\|_2 - \delta_{K+1}\|\mathbf{x}\|_2 = \delta_{K+1}\|\mathbf{x}\|_2 & (36) \end{aligned}$$

Since $|\mathbf{h}_{t^1}| = \|\mathbf{h}\|_\infty = \max_i |\mathbf{h}_i| \geq |\mathbf{h}_j|$, it is clear that

$$|\mathbf{h}_{t^1}| > \delta_{K+1}\|\mathbf{x}\|_2, \qquad (37)$$

which contradicts (33) and thus if $\|\mathbf{x}\|_\infty > 2\delta_{K+1}\|\mathbf{x}\|_2$ then $t^1 \in T$.

Moreover, since $\|\mathbf{x}\|_\infty \geq \frac{\|\mathbf{x}\|_2}{\sqrt{\|\mathbf{x}\|_0}} = \frac{\|\mathbf{x}\|_2}{\sqrt{K}}$, we have $\|\mathbf{x}\|_\infty > 2\delta_{K+1}\|\mathbf{x}\|_2$ as long as $\frac{1}{\sqrt{K}} > 2\delta_{K+1}$. Hence, we conclude that if $\delta_{K+1} < \frac{1}{2\sqrt{K}}$, then $t^1 \in T$. ■

One can observe that the bound in Lemma 8 is more stringent (smaller) than the bound in Lemma 5 and the ratio between two is around two. The main reason of this difference is that, while Lemma 5 is based on the inequality (21) characterizing the condition of the index corresponding to the maximal correlation, Lemma 8 relies on inequality (34), which in nature allows multiple indices.

Another key distinction is that our proof is built on the observation that the residual of each iteration of OMP can be modeled as a new measurement preserving sparsity level of the original signal (see Lemma 6). As a result, whole iterations are essentially governed by the condition that ensures the success of the first iterations. This observation is beneficial in two respects: 1) the analysis can be greatly simplified and 2) the loss in the bound, if any, caused by the non-initial iterations can be prevented. Indeed, by combining Lemma 6, 7, and Lemma 8, the initial condition can be used as a final condition and an improved D-W bound is obtained.[1] The improved condition is formally stated as follows.

Theorem 9 (Improved D-W bound): For any K-sparse vector $\mathbf{x}$, the OMP algorithm perfectly recovers $\mathbf{x}$ from $\mathbf{y} = \mathbf{\Phi}\mathbf{x}$ if the isometry constant δ_{K+1} satisfies $\delta_{K+1} < \frac{1}{2\sqrt{K}}$.

Proof: Immediate from Lemma 6, 7, and 8. ■

[1] In the construction of the D-W bound, the condition in the initial step is $\delta_{K+1} < \frac{1}{2\sqrt{K}}$ and that in the inductive step is $\delta_{K+1} < \frac{1}{3\sqrt{K}}$ so that the resulting condition is governed by $\delta_{K+1} < \frac{1}{3\sqrt{K}}$.

IV. Conclusion

In this paper, we presented the sufficient condition ensuring the exact reconstruction of sparse signals via the OMP algorithm. We showed that if the isometry constant δ_{K+1} of the sensing matrix $\mathbf{\Phi}$ satisfies $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ then the OMP algorithm can perfectly recover K-sparse signals from compressed measurements $\mathbf{y} = \mathbf{\Phi x}$. Our bound offers a substantial improvement over the previous result, and narrows down the gap between the recovery bound and the fundamental limit over which the perfect recovery of the OMP cannot be guaranteed. One important point we would like to mention is that while the OMP algorithm cannot guarantee the perfect recovery of the sparse signal for the matrix with $\delta_{K+1} \geq \frac{1}{\sqrt{K}}$, it does not necessarily mean that the OMP algorithm is not useful for such scenario. Even in this case, by allowing marginal performance degradation, the OMP algorithm will serve as a cost effective solution. Future works will be directed toward analyzing "lossy" bound of the OMP algorithm and its variants.

References

[1] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, pp. 4655–4666, Dec. 2007.

[2] M. A. Davenport and M. B. Wakin, "Analysis of Orthogonal Matching Pursuit using the restricted isometry property," *IEEE Trans. Inform. Theory*, vol. 56, pp. 4395–4401, Sept. 2010.

[3] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[4] E. Liu and V. N. Temlyakov, "The Orthogonal super greedy algorithm and applications in compressed sensing," *IEEE Trans. Inform. Theory*, vol. 58, pp. 2040–2047, April 2012.

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

[6] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Trans on Signal Process.*, vol. 58, pp. 1300–1308, March 2010.

[7] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans on Inform. Theory*, vol. 48, pp. 2558–2567, Sept. 2002.

[8] J. Wang, S. Kwon, and B. Shim, "Near optimal bound of orthogonal matching pursuit using restricted isometric constant," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–8, 2012.

[9] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[10] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE J. Sel. Topic Signal Process.*, vol. 4, no. 2, pp. 310–316, April 2010.

[11] D. L. Donoho, I. Drori, Y. Tsaig, and J. L. Starck, *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit*, Citeseer, 2006.

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