This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

Some New Results about Sufficient Conditions for Exact Support Recovery of Sparse Signals via Orthogonal Matching Pursuit

Chang Liu, Yong Fang, *Senior Member*, *IEEE*, and Jianzhong Liu

***Abstract*—Support recovery of sparse signals via orthogonal matching pursuit (OMP) has been extensively studied in recent years. In this paper, by exploiting the knowledge about orthogonal projection matrix and Schur complement, we study the sufficient conditions for exact support recovery of sparse signals with OMP in the framework of restricted isometry property (RIP). In the noisy case, we prove that under some constraints on the minimum magnitude of the nonzero elements of the *K*-sparse signal, OMP can exactly recover the support of the signal if the restricted isometry constant δ_{K+1} satisfies $\delta_{K+1} < 1/\sqrt{K+1}$. Our constraints on the minimum magnitude of nonzero elements of the signal are weaker than existing ones. In the noiseless case, although it has been proved that $\delta_{K+1} < 1/\sqrt{K+1}$ is a sharp condition for exactly recovering any *K*-sparse signal with OMP, our result shows that under some constraints on the signal, OMP can also exactly recover the signal if δ_{K+1} satisfies $\delta_{K+1} < \sqrt{2}/2$.**

***Index Terms*—Compressed sensing, orthogonal matching pursuit, restricted isometry property, orthogonal projection matrix, Schur complement**

I. Introduction

A. Overview

ORTHOGONAL matching pursuit (OMP) [1]-[3] is a classical greedy algorithm in compressed sensing which has received special attention due to its simplicity and competitive reconstruction performance. As a fundamental problem in OMP, support recovery of sparse signals from compressed linear measurements has been extensively studied over the years [4]-[9]. Let $x = (x_1, x_2, \cdots, x_N)^T \in \mathbb{C}^N$ (where $\mathbb{C}$ denotes the complex field.) be a *K*-sparse signal (*i.e.*, $|\text{supp}(x)| \le K$, where $\text{supp}(x) = \{i | x_i \ne 0\}$ is the support of x and $|\text{supp}(x)|$ is the cardinality of $\text{supp}(x)$.). $A = [\alpha_1, \alpha_2, \cdots, \alpha_N] \in \mathbb{C}^{n \times N}$ ($n \le N$) is a known sensing matrix. In OMP, it has been demonstrated that under some appropriate conditions on A and the *K*-sparse signal x, the signal can be reliably recovered in *K* iterations based on a set of compressive measurements $y \in \mathbb{C}^n$ obeying the linear model

$$y = Ax + v \tag{1}$$

where $v \in \mathbb{C}^n$ is a noise vector. The goal of support recovery is to identify the positions of nonzero elements of the input signal x from the measurement vector y. There are several types of noise, such as the l_2 bounded noise (*i.e.*, $\|v\|_2 \le \varepsilon$ for some constant $\varepsilon > 0$.), the l_∞ bounded noise (*i.e.*, $\|A^* v\|_\infty \le \varepsilon$ for some constant $\varepsilon > 0$.) and Gaussian noise [5]-[12]. In this paper, we are interested in studying the sufficient conditions for exact support recovery of sparse signals with OMP in the case when the noise is bounded and absent, respectively.

The basic idea of OMP is to iteratively identify the support of the sparse signal by adding one index into the list at a time according to the maximum correlation between columns of the measurement matrix and the current residual [9]. For any set $\Lambda \subset \{1, 2, \cdots, N\}$, let A_Λ denote the submatrix of A that contains only the columns indexed by Λ and x_Λ denote the subvector of x that contains only the entries indexed by Λ. Then OMP can be formally described in Table 1 [13].

TABLE I
ORTHOGONAL MATCHING PURSUIT

Input: A, y, stopping criterion.
Initialize: $k = 0$, $r^0 = y$, $\Lambda_0 = \varnothing$.
while not converged do 1. $k = k + 1$ 2. $\lambda_k = \arg\max_{1 \le i \le N} \left|\left\langle r^{k-1}, \alpha_i \right\rangle\right|$ 3. $\Lambda_k = \Lambda_{k-1} \cup \{\lambda_k\}$ 4. $x^k = \arg\min_{x \in \mathbb{C}^k} \left\|y - A_{\Lambda_k} x\right\|_2$ 5. $r^k = y - A_{\Lambda_k} x^k$
Output: $\hat{x} = x^k$

In compressed sensing, several features of sensing matrix A have been proposed to analyze the recovery performance, e.g., the *restricted isometry property* (RIP), *null space property*

This work was supported by the National Natural Science Foundation of China (NSFC) [61673253, 61271213] and the Ph.D. Programs Foundation of Ministry of Education of China [20133108110014].

Y. Fang and C. Liu are with the School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China (e-mail: yfang@shu.edu.cn; lc92617@163.com).

J. Liu is with the School of Mathematics and Physics, Jiangsu University of Technology, Jiangsu 213001, China (e-mail: ljz@jsut.edu.cn).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

(NSP) and *coherence parameter* [14]. Our results in this paper are all established based on the RIP. A matrix A satisfies the RIP of order m if there exists a constant $\delta_m \in [0,1)$ such that

$$(1-\delta_m)\|x\|_2^2 \le \|Ax\|_2^2 \le (1+\delta_m)\|x\|_2^2 \tag{2}$$

holds for all $x \in \mathbb{C}^N$ with $|\mathrm{supp}(x)| \le m$. In particular, the smallest constant δ_m satisfying (2) is called the restricted isometry constant (RIC).

B. Existing Results

Over the years, many efforts have been made to find out the RIP-based condition guaranteeing the exact recovery of sparse signals via OMP. In the noiseless case (*i.e.*, $\|v\|_2 = 0$), it has been shown in [15] and [16] that $\delta_{K+1} < 1/(\sqrt{K}+1)$ is sufficient for OMP to recover any K-sparse x in K iterations. Later, the conditions have been improved to $\delta_{K+1} < (\sqrt{4K+1}-1)/(2K)$ in [6]. As the latest report, it has been proven in [17] that $\delta_{K+1} < 1/\sqrt{K+1}$ is a sharp condition for exact recovery of any K-sparse signal with OMP. Moreover, there are a few works concerning sufficient conditions for recovering restricted classes of K-sparse signals with a more relaxed bound on RIC, see *e.g.*, [13], [18].

In the noisy case, it has been proven that under some constraints on the minimum magnitude of the nonzero elements of x (*i.e.*, $\min_{i\in\mathrm{supp}(x)} |x_i|$) and δ_{K+1} in (1), OMP can exactly recover $\mathrm{supp}(x)$ in K iterations. Several papers have established the sufficient conditions, such as [5], [6] and [7]. Very recently, Wen et al. have improved the sufficient condition over real field in [8].

C. Paper Contribution

In this paper, we study the RIP-based sufficient conditions for exact support recovery of sparse signals with OMP over complex field in both the noisy and noiseless cases. We conduct the analysis in the complex field because there are numerous applications of OMP in complex settings, such as magnetic resonance imaging, sampling theory and radar imaging community, etc., see e.g., [14]. The main contributions of this paper are summarized as follows.

In the noisy case, our results show that the support of K-sparse signals can be exactly recovered with a certain stopping criterion under either the l_2 or l_∞ bounded noise if A satisfies the RIP of order $K+1$ with $\delta_{K+1} < 1/\sqrt{K+1}$ and the $\min_{i\in\mathrm{supp}(x)} |x_i|$ exceeds a certain lower bound in (1). The proposed sufficient conditions on $\min_{i\in\mathrm{supp}(x)} |x_i|$ are weaker than the latest results proposed in [8].

In the noiseless case, we prove that under some constraints on the signal, OMP can also exactly recover the signal if δ_{K+1} satisfies $\delta_{K+1} < \sqrt{2}/2$.

The rest of the paper is organized as follows. In Section II, we establish the main results of this paper. We prove our results in Section III. Finally, we summarize this paper in Section IV.

Notation: Let $\Omega_m := \{1,2,\cdots,m\}$ with $m \in \mathbb{N}^+$ and denote $\Omega_0 := \varnothing$. For two sets Λ and Γ, let $\Lambda \setminus \Gamma = \{i \mid i \in \Lambda, i \notin \Gamma\}$. Let H denote a subspace of $\mathbb{C}^n$. P_{H} stands for an orthogonal projection matrix onto H. Similarly, $P_{\mathrm{H}}^{\perp} = I_n - P_{\mathrm{H}}$ is an orthogonal projection matrix onto the orthogonal complement of H, where I_n denotes the identity matrix. For any $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_{|\Lambda|}\} \subset \Omega_N$, $H_\Lambda := span\{\alpha_{\lambda_1}, \alpha_{\lambda_2}, \cdots, \alpha_{\lambda_{|\Lambda|}}\}$ represents the span of $\alpha_{\lambda_1}, \alpha_{\lambda_2}, \cdots, \alpha_{\lambda_{|\Lambda|}}$ and denote $H_\varnothing := \{0\}$. Throughout the paper, $(\cdot)^T$ and $(\cdot)^*$ respectively denote matrix transposition and matrix conjugate transposition and $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ respectively denote l_1-norm, l_2-norm, and l_∞-norm of a vector.

For complex vectors $x, y \in \mathbb{C}^n$, the scalar $\langle x, y\rangle = y^* x$ denotes the Euclidean inner product of x and y. For complex matrices $B, C \in \mathbb{C}^{n\times n}$, we write $B > 0$ if B is a positive definite matrix and write $B \ge 0$ if B is a positive semi-definite matrix. Let $B \ge C$ denote $B - C \ge 0$. For a block matrix $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$ where the block B_{11} is a k-order invertible matrix with $0 < k < n$, let $C_k(B)$ denote the Schur complement of B_{11} in B, *i.e.*, $C_k(B) = B_{22} - B_{21}B_{11}^{-1}B_{12}$. Define $C_0(B) = B$.

II. Main Results

In this section, we present some new results about sufficient conditions for exact support recovery of sparse signals with OMP by the following three theorems.

Theorem 1: Under $\|v\|_2 \le \varepsilon$, suppose that A in (1) satisfies the RIP of order $K+1$ with $\delta_{K+1} < 1/\sqrt{K+1}$. Then OMP with the stopping criterion $\|r^k\|_2 \le \varepsilon$ can exactly recover $\mathrm{supp}(x)$ in K iterations provided that

$$\min_{i\in\mathrm{supp}(x)} |x_i| > \frac{\varepsilon}{\sqrt{1-\delta_{K+1}}} + \frac{\sqrt{1+\delta_{K+1}}\,\varepsilon}{1-\sqrt{K+1}\,\delta_{K+1}}. \tag{3}$$

Proof: See Section III-A.

Remark 1: Here, we compare our condition with that in [8], which is the best result so far.

It has been shown in [8, Theorem 1] that if $\|v\|_2 \le \varepsilon$ and A satisfies the RIP with $\delta_{K+1} < 1/\sqrt{K+1}$ in (1), then OMP with the stopping criterion $\left\|r^k\right\|_2 \leq \varepsilon$ can exactly recover $\operatorname{supp}(x)$ in K iterations provided that

$$\min_{i\in\operatorname{supp}(x)} |x_i| > \frac{2\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}. \tag{4s}$$

[8, Theorem 3] also provides a necessary condition on the minimum magnitude of the nonzero elements of x, which is

$$\min_{i\in\operatorname{supp}(x)} |x_i| > \frac{\sqrt{1-\delta_{K+1}}\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}. \tag{4n}$$

To gauge how much improvement has been made over [8] in the sufficient condition and how much further improvement is possible, let

$$e_{4s}(\delta_{K+1}) = \frac{\frac{\varepsilon}{\sqrt{1-\delta_{K+1}}} + \frac{\sqrt{1+\delta_{K+1}}\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}}{\frac{2\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}} = \frac{1}{2}\left(\frac{1-\sqrt{K+1}\delta_{K+1}}{\sqrt{1-\delta_{K+1}}} + \sqrt{1+\delta_{K+1}}\right),$$

$$e_{4n}(\delta_{K+1}) = \frac{\frac{\varepsilon}{\sqrt{1-\delta_{K+1}}} + \frac{\sqrt{1+\delta_{K+1}}\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}}{\frac{\sqrt{1-\delta_{K+1}}\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}} = \frac{1-\sqrt{K+1}\delta_{K+1}}{1-\delta_{K+1}} + \frac{\sqrt{1+\delta_{K+1}}}{\sqrt{1-\delta_{K+1}}}.$$

It is easy to check that $\sqrt{1+\delta_{K+1}} + \sqrt{1-\delta_{K+1}} \leq 2$. Thus we have

$$\begin{aligned} e_{4s}(\delta_{K+1}) &= \frac{1}{2}\left(\frac{1-\sqrt{K+1}\delta_{K+1}}{\sqrt{1-\delta_{K+1}}} + \sqrt{1+\delta_{K+1}}\right) \\ &\leq \frac{1}{2}\left(2 + \frac{1-\sqrt{K+1}\delta_{K+1}}{\sqrt{1-\delta_{K+1}}} - \sqrt{1-\delta_{K+1}}\right) \\ &\leq \frac{1}{2}\left(2 + \frac{1-\delta_{K+1}}{\sqrt{1-\delta_{K+1}}} - \sqrt{1-\delta_{K+1}}\right) = 1. \end{aligned}$$

Therefore, the sufficient condition in (3) is weaker than the one presented in (4s).

By some simple calculations, one can verify that

$$\lim_{\delta_{K+1}\to 0} e_{4s}(\delta_{K+1}) = 1, \quad \lim_{\delta_{K+1}\to \frac{1}{\sqrt{K+1}}} e_{4s}(\delta_{K+1}) = \frac{1}{2}\sqrt{1+\frac{1}{\sqrt{K+1}}},$$

$$\lim_{\delta_{K+1}\to 0} e_{4n}(\delta_{K+1}) = 2, \quad \lim_{\delta_{K+1}\to \frac{1}{\sqrt{K+1}}} e_{4n}(\delta_{K+1}) = \frac{\sqrt{K+1}+1}{\sqrt{K}}.$$

Obviously, when δ_{K+1} is close to $1/\sqrt{K+1}$ and K is large enough, our sufficient condition is nearly half of the one presented in [8, Theorem 1] and is close to the necessary condition presented in [8, Theorem 3]. Figure 1 also confirms our analysis.

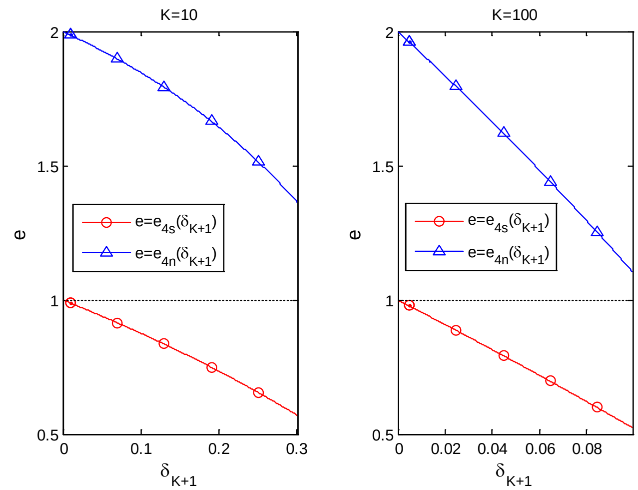

Fig. 1. $e_{4s}(\delta_{K+1})$ and $e_{4n}(\delta_{K+1})$ for different sparse level K. (left) $K=10$, (right) $K=100$.

Theorem 2: Under $\left\|A^* v\right\|_\infty \leq \varepsilon$, suppose that A in (1) satisfies the RIP of order $K+1$ with $\delta_{K+1} < 1/\sqrt{K+1}$. Then OMP with the stopping criterion

$$\left\|A^* r^k\right\|_\infty \leq \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon$$

can exactly recover $\operatorname{supp}(x)$ in K iterations provided that

$$\min_{i\in\operatorname{supp}(x)} |x_i| > \frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}} + \frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{(1-\delta_{K+1})\left(1-\sqrt{K+1}\delta_{K+1}\right)}\varepsilon. \tag{5}$$

Proof: See Section III-B.

Remark 2: Similar to Remark 1, we compare our condition with the one proposed in [8].

It has been shown in [8, Theorem 2] that if $\left\|A^T v\right\|_\infty \leq \varepsilon$ and A satisfies the RIP with $\delta_{K+1} < 1/\sqrt{K+1}$ in (1), then OMP with the stopping criterion

$$\left\|A^T r^k\right\|_\infty \leq \left(1+\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}\sqrt{K}\right)\varepsilon$$

can exactly recover $\operatorname{supp}(x)$ in K iterations provided that

$$\min_{i\in\operatorname{supp}(x)} |x_i| > \frac{2}{1-\sqrt{K+1}\delta_{K+1}}\left(1+\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}\sqrt{K}\right)\varepsilon. \tag{6s}$$

[8, Theorem 4] also provides a necessary condition on the minimum magnitude of the nonzero elements of x, which is

$$\min_{i\in\operatorname{supp}(x)} |x_i| > \frac{2\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}. \tag{6n}$$

Also, let

$$e_{6s}(\delta_{K+1}) = \frac{\frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}} + \frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{(1-\delta_{K+1})\left(1-\sqrt{K+1}\delta_{K+1}\right)}\varepsilon}{\frac{2}{1-\sqrt{K+1}\delta_{K+1}}\left(1+\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}\sqrt{K}\right)\varepsilon}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$$= \frac{\frac{\sqrt{K}\left(1-\sqrt{K+1}\delta_{K+1}\right)}{1-\delta_{K+1}}+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}+1}{2\left(1+\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}\sqrt{K}\right)},$$

$$e_{6n}\left(\delta_{K+1}\right)=\frac{\frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}}+\frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{\left(1-\delta_{K+1}\right)\left(1-\sqrt{K+1}\delta_{K+1}\right)}\varepsilon}{\frac{2\varepsilon}{1-\sqrt{K+1}\delta_{K+1}}}$$

$$=\frac{1}{2}\left(\frac{\sqrt{K}\left(1-\sqrt{K+1}\delta_{K+1}\right)}{1-\delta_{K+1}}+\frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right).$$

It is easy to check that

$$\frac{1}{1-\delta_{K+1}}\le 2\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}$$

when $\delta_{K+1}<1/\sqrt{K+1}$. Thus we have

$$e_{6s}\left(\delta_{K+1}\right)=\frac{\frac{\sqrt{K}\left(1-\sqrt{K+1}\delta_{K+1}\right)}{1-\delta_{K+1}}+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}+1}{2\left(1+\sqrt{\frac{1+\delta_{K+1}}{1-\delta_{K+1}}}\sqrt{K}\right)}$$

$$\le\frac{\frac{\sqrt{K}\left(1-\sqrt{K+1}\delta_{K+1}\right)}{1-\delta_{K+1}}+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}+1}{2+\frac{\sqrt{K}}{1-\delta_{K+1}}}$$

$$=\frac{\sqrt{K}\left(1-\sqrt{K+1}\delta_{K+1}\right)+\sqrt{K}\delta_{K+1}+1-\delta_{K+1}}{\sqrt{K}+2-2\delta_{K+1}}$$

$$=\frac{\sqrt{K}+1+\left(\sqrt{K}-\sqrt{K}\sqrt{K+1}-1\right)\delta_{K+1}}{\sqrt{K}+2-2\delta_{K+1}}$$

$$\le\frac{\sqrt{K}+1-\delta_{K+1}}{\sqrt{K}+2-2\delta_{K+1}}\le 1$$

Therefore, the sufficient condition in (5) is weaker than the one presented in (6s).

By some simple calculations, one can verify that

$$\lim_{\delta_{K+1}\to 0}e_{6s}\left(\delta_{K+1}\right)=\frac{1}{2},$$

$$\lim_{\delta_{K+1}\to\frac{1}{\sqrt{K+1}}}e_{6s}\left(\delta_{K+1}\right)$$

$$=\frac{\sqrt{K+1}+\sqrt{K}-1}{2\left(\sqrt{K+1}-1\right)\left(\sqrt{K+1}+2\right)}$$

$$<\frac{2\left(\sqrt{K+1}+2\right)}{2\left(\sqrt{K+1}-1\right)\left(\sqrt{K+1}+2\right)}=\frac{1}{\sqrt{K+1}-1},$$

$$\lim_{\delta_{K+1}\to 0}e_{6n}\left(\delta_{K+1}\right)=\frac{\sqrt{K}+1}{2},\quad \lim_{\delta_{K+1}\to\frac{1}{\sqrt{K+1}}}e_{6n}\left(\delta_{K+1}\right)=\frac{\sqrt{K+1}+1}{2\sqrt{K}}+\frac{1}{2}.$$

Obviously, when δ_{K+1} is close to zero, our sufficient condition is nearly half of the one in [8, Theorem 2]. When δ_{K+1} is close to $1/\sqrt{K+1}$ and K is large enough, our sufficient condition is close to the necessary condition in [8, Theorem 4]. Figure 2 also confirms our analysis.

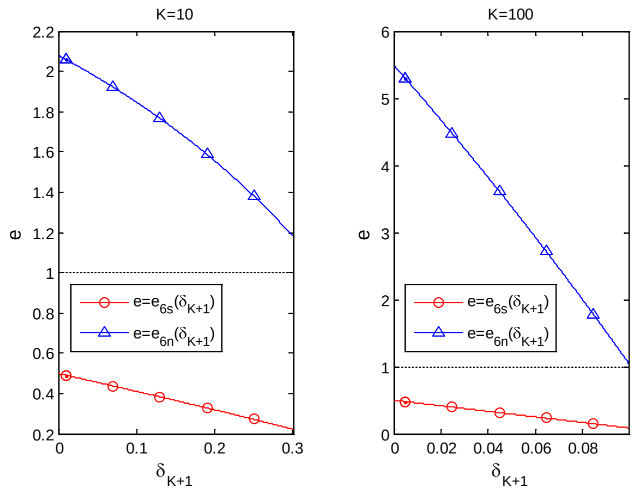

Fig. 2. $e_{6s}\left(\delta_{K+1}\right)$ and $e_{6n}\left(\delta_{K+1}\right)$ for different sparse level K. (left) $K=10$, (right) $K=100$.

In the case that $\operatorname{supp}(x)$ can be exactly recovered in the presence of noise, we are interested in the error between the output signal and the original sparse signal. For this, we have the following proposition.

Proposition 1: In (1), let $\left|\operatorname{supp}(x)\right|=s\le m$. Suppose that A satisfies RIP of order m and $\operatorname{supp}(x)$ can be exactly recovered via OMP in s iterations, *i.e.*, $\Lambda_s=\operatorname{supp}(x)$. Then for $\|v\|_2\le\varepsilon$,

$$\left\|\hat{x}-x_{\Lambda_s}\right\|_2\le\frac{\varepsilon}{\sqrt{1-\delta_m}}.$$

For $\left\|A^*v\right\|_\infty\le\varepsilon$,

$$\left\|\hat{x}-x_{\Lambda_s}\right\|_2\le\frac{\sqrt{s}\varepsilon}{1-\delta_m}.$$

Proof: See Section III-C.

In the noiseless case, although it has been proved that $\delta_{K+1}<1/\sqrt{K+1}$ is a sharp condition for exactly recovering any K-sparse signal with OMP, the papers [13] and [18] have shown that under some constraints on the signal, the condition on RIC for signal recovery with OMP can be weaker. Inspired by the above two works, we have some analogous results. First, we need to introduce the following important notations.

For any $x\in\mathbb{C}^N$, define

$$R(x) := \max\left\{\frac{\|x_\Lambda\|_1}{\|x_\Lambda\|_2} \middle| \Lambda \subset \Omega_N,\ \ \Lambda \neq \varnothing\right\}$$

where $\|x_\Lambda\|_1 / \|x_\Lambda\|_2 = 1$ for $x_\Lambda = 0$. Evidently, if x is a K-sparse signal, then $1 \le R(x) \le \sqrt{K}$ (see Section III (p4)).

Since OMP can always exactly recover signals in $|\mathrm{supp}(x)|$ iterations when $\delta_{K+1} = 0$, we give the following results in the case that $\delta_{K+1} > 0$.

Theorem 3: In the noiseless case, suppose that A in (1) satisfies the RIP of order $K+1$ with $\delta_{K+1} < \sqrt{2}/2$. Then OMP can exactly recover the K-sparse signal x in K iterations provided that

$$R(x) < \frac{\sqrt{1-\delta_{K+1}^2}}{\delta_{K+1}}. \quad (7)$$

Proof: See Section III-D.

Remark 3: One interesting point is why the condition in terms of RIC is $\delta_{K+1} < \sqrt{2}/2$. Recall that $R(x) \ge 1$ for any K-sparse signal x. Thus to ensure that the condition (7) is valid, it is necessary to ensure

$$\frac{\sqrt{1-\delta_{K+1}^2}}{\delta_{K+1}} = \sqrt{\frac{1}{\delta_{K+1}^2} - 1} > 1,$$

i.e., $\delta_{K+1} < \sqrt{2}/2$.

Remark 4: We now demonstrate that Theorem 3 can be adapted to analyze the sufficient condition for recovering α-strongly-decaying K-sparse signals [13] with OMP in the noiseless case.

For any K-sparse signal x, we denote by $x'(j)$ the entries of x ordered by magnitude, i.e.,

$$|x'(1)| \ge |x'(2)| \ge \cdots \ge |x'(K)| \ge 0$$

with $|x'(K+1)| = |x'(K+2)| = \cdots = |x'(N)| = 0$. Then x is called an α-strongly-decaying K-sparse signal ($\alpha > 1$) if

$$|x'(j)| \ge \alpha |x'(j+1)|$$

for all $j \in \{1, 2, \cdots, K-1\}$.

Suppose that x is an α-strongly-decaying K-sparse signal. We have verified that (see Appendix II) $R(x) \le \sqrt{g_K(\alpha)}$ where

$$g_K(\alpha) = \frac{(\alpha^K - 1)(\alpha+1)}{(\alpha^K+1)(\alpha-1)}.$$

Then we have the following result from Theorem 3.

In the noiseless case, suppose that A in (1) satisfies the RIP of order $K+1$ with $\delta_{K+1} < \sqrt{2}/2$. If

$$\sqrt{g_K(\alpha)} < \frac{\sqrt{1-\delta_{K+1}^2}}{\delta_{K+1}},$$

then OMP can exactly recover x in K iterations.

By some simple calculation, we found that the above result is equivalent to [18, Theorem 5] in the case that the block size $d = 1$.

Remark 5: Recall that there are two other common frameworks which can be used to analyze the recovery performance in compressed sensing. One has involved the notation of a coherence parameter $\mu := \max_{i \neq j} |\langle \alpha_i, \alpha_j \rangle|$. It has been shown that OMP can recover any K-sparse signal in the noiseless case in K iterations when the columns of A in (1) have unit norm and $\mu < 1/(2\mathrm{K}-1)$. The other is the NSP (See [14, Definition 4.1]). In the noiseless case, it has been confirmed that the matrix A in (1) satisfying the NSP of order K is a necessary and sufficient condition for exact recovery of any K-sparse vector via Basis Pursuit (See [14, Theorem 4.5]) and is a necessary condition for exact recovery of any K-sparse vector in $|\mathrm{supp}(x)|$ iterations via OMP (See [14, Sect. 5.4]). Inspired by Theorem 3, it is natural to think about the condition for recovering restricted classes of K-sparse signals via OMP based on the coherence parameter or the NSP, which is worth studying in future work.

III. Proof

In this section, the main results given in section II will be proved. First, we need the following evident properties of the sensing matrix A, orthogonal projection matrix and the Schur complement of matrices.

(p1): For any $\alpha, \beta \in \mathbb{C}^n$, it holds that $\langle P_{\mathrm{H}}\alpha, P_{\mathrm{H}}\beta \rangle = \langle P_{\mathrm{H}}\alpha, \beta \rangle = \langle \alpha, P_{\mathrm{H}}\beta \rangle$ and $\|P_{\mathrm{H}}\alpha\|_2 \le \|\alpha\|_2$.

(p2): Suppose that matrix A satisfies RIP of order $K+1$. For any $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_{|\Lambda|}\} \subset \Omega_N$ with $|\Lambda| \le K+1$, it holds that

$$(1-\delta_{K+1})\mathrm{I}_{|\Lambda|} \le A_\Lambda^* A_\Lambda \le (1+\delta_{K+1}) I_{|\Lambda|},$$

$$P_{H_\Lambda} = A_\Lambda \left(A_\Lambda^* A_\Lambda\right)^{-1} A_\Lambda^*$$

and

$$(1-\delta_{K+1})\|\xi\|_2^2 \le \left\|\sum_{i=1}^{|\Lambda|} \xi_i \alpha_{\lambda_i}\right\|_2^2 \le (1+\delta_{K+1})\|\xi\|_2^2,$$

where $\xi = \left(\xi_1, \xi_2, \cdots, \xi_{|\Lambda|}\right)^T$.

(p3) (see [19, Problem 7.7.p41]): For any $B, C \in \mathbb{C}^{n\times n}$, if $B \ge C > 0$, then $C_k(B) \ge C_k(C)$.

(p4): For any $x \in \mathbb{C}^n$, $\|x\|_1 \le \sqrt{n}\|x\|_2$, $\|x\|_2 \le \sqrt{n}\|x\|_\infty$ and $\|x\|_\infty \le \|x\|_2 \le \|x\|_1$.

Next, we present the following important lemmas, which will be used in the proofs of the main results. All the lemmas are based on the assumption that A in (1) satisfies the RIP of order $K+1$. To ease reading, the proofs of the lemmas are relegated to the appendix I.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

Lemma 1: In (1), let $P_{H_{\Omega_s}} y = \sum_{i=1}^{s} \hat{x}_i \alpha_i = A_{\Omega_s} \hat{x}$ with $\hat{x} = (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_s)^T$. Suppose that $Ax = A_{\Omega_s} x_{\Omega_s}$ with $s \le K$ and $\hat{x}_1 \hat{x}_2 \cdots \hat{x}_s \neq 0$. For any $s < j \le N$ and $\Lambda_k = \begin{cases} \emptyset, & k=0 \\ \{\lambda_1, \lambda_2, \cdots, \lambda_k\}, & 1 \le k < s \end{cases}$ with $\{\lambda_1, \lambda_2, \cdots, \lambda_k\} \subset \Omega_s$, define

$$S_{0\Lambda_k} := \max_{1 \le i \le s} \left| \left\langle P_{H_{\Lambda_k}}^{\perp} y, \alpha_i \right\rangle \right|, \quad S_{j\Lambda_k} := \left| \left\langle P_{H_{\Lambda_k}}^{\perp} y, \alpha_j \right\rangle \right|.$$

Then there exists a $\theta_j \in [0, 2\pi]$ such that

$$\begin{aligned} & S_{0\Lambda_k} - S_{j\Lambda_k} \\ & \ge \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} - \delta_{K+1} \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} \sqrt{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1^2 + \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2} \\ & - \operatorname{Re} \left\langle \tilde{\alpha}_j, P_{H_{\Omega_s}}^{\perp} v \right\rangle \end{aligned} \tag{8}$$

where $\tilde{\alpha}_j = e^{i\theta_j} \alpha_j$.

Lemma 2: In (1), if $\left\| A^* v \right\|_\infty \le \varepsilon$, then we have

$$\left| \left\langle \alpha_j, P_{H_{\Omega_s}}^{\perp} v \right\rangle \right| = \left| \left\langle P_{H_{\Omega_s}}^{\perp} \alpha_j, v \right\rangle \right| \le \left(1 + \frac{\sqrt{s} \delta_{K+1}}{1 - \delta_{K+1}} \right) \varepsilon \tag{9}$$

for $s \le K$ and $1 \le j \le N$.

Lemma 3: In (1), let $P_{H_{\Omega_s}} y = \sum_{i=1}^{s} \hat{x}_i \alpha_i$ and suppose that $Ax = \sum_{i=1}^{s} x_i \alpha_i$ with $s \le K$. Then for $\|v\|_2 \le \varepsilon$, we have

$$\min_{1 \le i \le s} |\hat{x}_i| \ge \min_{1 \le i \le s} |x_i| - \frac{\varepsilon}{\sqrt{1 - \delta_{K+1}}}. \tag{10}$$

For $\left\| A^* v \right\|_\infty \le \varepsilon$, we have

$$\min_{1 \le i \le s} |\hat{x}_i| \ge \min_{1 \le i \le s} |x_i| - \frac{\sqrt{K} \varepsilon}{1 - \delta_{K+1}}. \tag{11}$$

Lemma 4: In (1), suppose that $Ax = A_{\Omega_s} x_{\Omega_s} = \sum_{i=1}^{s} x_i \alpha_i$ with $s \le K$ and $x_1 x_2 \cdots x_s \neq 0$. Then for $\|v\|_2 \le \varepsilon$, we have

$$\left\| P_{H_{\Omega_0}}^{\perp} y \right\|_2 \ge \left\| P_{H_{\Omega_1}}^{\perp} y \right\|_2 \ge \left\| P_{H_{\Omega_{s-1}}}^{\perp} y \right\|_2 \ge \sqrt{1 - \delta_{K+1}} \min_{1 \le i \le s} |x_i| - \varepsilon. \tag{12}$$

For $\left\| A^* v \right\|_\infty \le \varepsilon$, we have

$$\left\| A^* P_{H_{\Omega_s}}^{\perp} y \right\|_\infty \le \left(1 + \frac{\sqrt{K} \delta_{K+1}}{1 - \delta_{K+1}} \right) \varepsilon \tag{13}$$

and

$$\left\| A^* P_{\Omega_k}^{\perp} y \right\|_\infty \ge (1 - \delta_{K+1}) \min_{1 \le i \le s} |x_i| - \left(1 + \frac{\sqrt{K-1} \delta_{K+1}}{1 - \delta_{K+1}} \right) \varepsilon \tag{14}$$

for all $k < s$.

In the following, we give the proofs of the main results. Without loss of generality, we may prove the following results in the case that $\operatorname{supp}(x) = \Omega_s = \{1, 2, \cdots, s\}$ with $s \le K$.

A. Proof of Theorem 1

The proof can be divided into two steps. We first show that OMP selects correct indexes in all iterations. Then, we prove that it performs exactly s iterations.

We prove the first step by induction. Suppose that OMP selects correct indexes in the first k iterations, *i.e.*, $\Lambda_k \subset \Omega_s$ with $k < s$. The assumption holds true naturally when $k = 0$ since $\Lambda_0 = \emptyset \subset \Omega_s$. Then, we need to show that OMP selects a correct index in the $k+1$-th iteration, *i.e.*, $\lambda_{k+1} \in \Omega_s$.

Obviously, $r^k = P_{H_{\Lambda_k}}^{\perp} y$. Thus it suffices to show that $S_{0\Lambda_k} - S_{j\Lambda_k} > 0$ for any $k = 0, 1, \cdots, s-1$ and $s < j \le N$ in lemma 1. Let $P_{H_{\Omega_s}} y = \sum_{i=1}^{s} \hat{x}_i \alpha_i$ and $Ax = \sum_{i=1}^{s} x_i \alpha_i$. Then from (3) and (10), we have

$$\begin{aligned} \min_{1 \le i \le s} |\hat{x}_i| & > \frac{\varepsilon}{\sqrt{1 - \delta_{K+1}}} + \frac{\sqrt{1 + \delta_{K+1}} \varepsilon}{1 - \sqrt{K+1} \delta_{K+1}} - \frac{\varepsilon}{\sqrt{1 - \delta_{K+1}}} \\ & \ge \frac{\sqrt{1 + \delta_{K+1}} \varepsilon}{1 - \sqrt{K+1} \delta_{K+1}} > 0. \end{aligned}$$

It follows immediately that $\hat{x}_1 \hat{x}_2 \cdots \hat{x}_s \neq 0$. Since matrix A satisfies RIP of order $K+1$, it is easy to see that $\left\| \tilde{\alpha}_j \right\|_2 = \left\| \alpha_j \right\|_2 \le \sqrt{1 + \delta_{K+1}}$. Thus we have

$$\left| \left\langle \tilde{\alpha}_j, P_{H_{\Omega_s}}^{\perp} v \right\rangle \right| \le \left\| \tilde{\alpha}_j \right\|_2 \left\| P_{H_{\Omega_s}}^{\perp} v \right\|_2 \le \sqrt{1 + \delta_{K+1}} \varepsilon \tag{15}$$

From (8) and (15), we obtain

$$\begin{aligned} & S_{0\Lambda_k} - S_{j\Lambda_k} \\ & \ge \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} - \delta_{K+1} \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} \sqrt{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1^2 + \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2} - \sqrt{1 + \delta_{K+1}} \varepsilon \\ & = \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2 \left(\frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} - \delta_{K+1} \sqrt{1 + \left(\frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} \right)^2} - \frac{\sqrt{1 + \delta_{K+1}} \varepsilon}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2} \right) \end{aligned} \tag{16}$$

Notice that the function $f(x) = x - \delta_{K+1} \sqrt{1 + x^2}$ is monotonously increasing on the interval $[0, +\infty)$. By (p4), it is easy to check that $\frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} \ge \frac{1}{\sqrt{s-k}}$ and $\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2 \ge \sqrt{s-k} \min_{1 \le i \le s} |\hat{x}_i|$.

Thus it follows from (16) that

$S_{0\Lambda_k} - S_{j\Lambda_k}$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$$\geq \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2 \left(\frac{1}{\sqrt{s-k}} - \delta_{K+1} \frac{\sqrt{s-k+1}}{\sqrt{s-k}} - \frac{\sqrt{1+\delta_{K+1}}\,\varepsilon}{\sqrt{s-k}\min\limits_{1\leq i\leq s}|\hat{x}_i|} \right). \quad (17)$$

From (3), (10), and (17), we have

$$\begin{aligned} S_{0\Lambda_k} - S_{j\Lambda_k} &> \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \left(1 - \delta_{K+1}\sqrt{s-k+1} - \left(1 - \sqrt{K+1}\delta_{K+1}\right)\right) \\ &= \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \delta_{K+1}\left(\sqrt{K+1} - \sqrt{s-k+1}\right) \\ &\geq \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \delta_{K+1}\left(\sqrt{K+1} - \sqrt{s+1}\right) \geq 0 \end{aligned}$$

Thus we can infer that OMP can identify a correct index at each iteration.

In the following, we show that OMP with the stopping criterion $\left\| r^k \right\|_2 \leq \varepsilon$ performs exactly s iterations, which is equivalent to show that $\left\| r^k \right\|_2 > \varepsilon$ for $k = 0,1,\cdots,s-1$ and $\left\| r^s \right\|_2 \leq \varepsilon$.

Without loss of generality, we may assume that $\Lambda_k = \Omega_k$ with $0 \leq k \leq s$. Then we can derive from (3) and (12) that

$$\begin{aligned} &\left\| r^0 \right\|_2 \geq \left\| r^1 \right\|_2 \geq \cdots \geq \left\| r^{s-1} \right\|_2 \\ &\geq \sqrt{1-\delta_{K+1}} \min_{1\leq i\leq s}|x_i| - \varepsilon > \frac{\sqrt{1-\delta_{K+1}^2}}{1-\sqrt{K+1}\delta_{K+1}}\varepsilon \geq \frac{\sqrt{1-\delta_{K+1}}}{1-\sqrt{K+1}\delta_{K+1}}\varepsilon \\ &\geq \frac{1-\delta_{K+1}}{1-\sqrt{K+1}\delta_{K+1}}\varepsilon \geq \varepsilon, \end{aligned}$$

i.e., OMP does not terminate before the s-th iteration.

Since $r^s = P_{H_{\Omega_s}}^{\perp} y = P_{H_{\Omega_s}}^{\perp}\left(A_{\Omega_s} x_{\Omega_s} + v\right) = P_{H_{\Omega_s}}^{\perp} v$, it is easy to check that $\left\| r^s \right\|_2 = \left\| P_{H_{\Omega_s}}^{\perp} v \right\|_2 \leq \|v\|_2 \leq \varepsilon$. Therefore, OMP terminates after performing the s-th iteration.

Thus OMP performs s iterations. This completes the proof.

B. Proof of Theorem 2

We first prove that for any α_j with $s < j \leq N$ and any $k \in \mathbb{N}$ with $0 \leq k < s$, if $\Lambda_k \subset \Omega_s$, then $S_{0\Lambda_k} - S_{j\Lambda_k} > 0$. From (5) and (11), we have

$$\min_{1\leq i\leq s}|\hat{x}_i| \geq \min_{1\leq i\leq s}|x_i| - \frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}} > \frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{\left(1-\delta_{K+1}\right)\left(1-\sqrt{K+1}\delta_{K+1}\right)}\varepsilon > 0.$$

Thus $\hat{x}_1\hat{x}_2\cdots\hat{x}_s \neq 0$ holds. It follows from (9) that

$$\left|\left\langle \tilde{\alpha}_j,\ P_{H_{\Omega_s}}^{\perp} v\right\rangle\right| = \left|\left\langle \alpha_j,\ P_{H_{\Omega_s}}^{\perp} v\right\rangle\right| \leq \left(1+\frac{\sqrt{s}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \leq \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon. \quad (18)$$

Using the same argument as in the proof of (17), we can obtain from (8) and (18) that

$$\begin{aligned} &S_{0\Lambda_k} - S_{j\Lambda_k} \\ &\geq \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} - \delta_{K+1} \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} \sqrt{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1^2 + \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2^2} - \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \\ &= \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2 \left(\frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1} - \delta_{K+1}\sqrt{1+\left(\frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_1}\right)^2} - \frac{\left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)}{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}\varepsilon \right) \\ &\geq \left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2 \left(\frac{1}{\sqrt{s-k}} - \delta_{K+1}\frac{\sqrt{s-k+1}}{\sqrt{s-k}} - \frac{\left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)}{\sqrt{s-k}\min\limits_{1\leq i\leq s}|\hat{x}_i|}\varepsilon \right). \quad (19) \end{aligned}$$

Combining (5), (11), and (19), we have

$$\begin{aligned} &S_{0\Lambda_k} - S_{j\Lambda_k} \\ &> \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \left(1 - \delta_{K+1}\sqrt{s-k+1} - \left(1 - \sqrt{K+1}\delta_{K+1}\right)\right) \\ &= \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \delta_{K+1}\left(\sqrt{K+1} - \sqrt{s-k+1}\right) \\ &\geq \frac{\left\| \hat{x}_{\Omega_s \setminus \Lambda_k} \right\|_2}{\sqrt{s-k}} \delta_{K+1}\left(\sqrt{K+1} - \sqrt{s+1}\right) \geq 0 \end{aligned}$$

Thus we can infer that OMP can identify a correct index at each iteration.

Next, we show that OMP with the stopping criterion $\left\| A^* r^k \right\|_\infty \leq \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon$ performs exactly s iterations, which is equivalent to show that

$$\left\| A^* r^k \right\|_\infty > \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon$$

for $k = 0,1,\cdots,s-1$ and

$$\left\| A^* r^s \right\|_\infty \leq \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon.$$

Also, for simplicity, we may take $\Lambda_k = \Omega_k$ with $0 \leq k \leq s$. From (5) and (14), it follows that for any $0 \leq k < s$,

$$\begin{aligned} &\left\| A^* r^k \right\|_\infty = \left\| A^* P_{\Omega_k}^{\perp} y \right\|_\infty \\ &\geq \left(1-\delta_{K+1}\right)\min_{1\leq i\leq s}|x_i| - \left(1+\frac{\sqrt{K-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \\ &> \sqrt{K}\varepsilon + \frac{1-\delta_{K+1}+\sqrt{K}\delta_{K+1}}{1-\sqrt{K+1}\delta_{K+1}}\varepsilon - \left(1+\frac{\sqrt{K-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \\ &= \sqrt{K}\varepsilon + \varepsilon + \frac{\left(\sqrt{K+1}+\sqrt{K}-1\right)\delta_{K+1}}{1-\sqrt{K+1}\delta_{K+1}}\varepsilon - \left(1+\frac{\sqrt{K-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$$\geq \sqrt{K}\varepsilon+\varepsilon+\frac{\left(\sqrt{K+1}+\sqrt{K}-1\right)\delta_{K+1}}{1-\delta_{K+1}}\varepsilon-\left(1+\frac{\sqrt{K-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon$$

$$=\sqrt{K}\varepsilon+\frac{\left(\sqrt{K+1}+\sqrt{K}-\sqrt{K-1}-1\right)\delta_{K+1}}{1-\delta_{K+1}}\varepsilon$$

$$=\left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon+\left(\sqrt{K}-1+\frac{\left(\sqrt{K+1}-\sqrt{K-1}-1\right)\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon. \quad (20)$$

Thus it suffices to prove that for $\delta_{K+1}<1/\sqrt{K+1}$,

$$\sqrt{K}-1+\frac{\left(\sqrt{K+1}-\sqrt{K-1}-1\right)\delta_{K+1}}{1-\delta_{K+1}}\geq 0. \quad (21)$$

It is obvious that (21) holds for $K=1$. For $K>1$, notice that $\sqrt{K+1}-\sqrt{K-1}-1<0$, $\delta_{K+1}/(1-\delta_{K+1})\leq 1/\left(\sqrt{K+1}-1\right)$ and $\sqrt{K}>\sqrt{K-1}/\left(\sqrt{K+1}-1\right)$, thus

$$\sqrt{K}-1+\frac{\left(\sqrt{K+1}-\sqrt{K-1}-1\right)}{1-\delta_{K+1}}\delta_{K+1}\geq\sqrt{K}-\frac{\sqrt{K-1}}{\sqrt{K+1}-1}>0.$$

Therefore, (21) holds, *i.e.*, OMP does not terminate before the s-th iteration.

It follows from (13) that

$$\left\|A^* r^s\right\|_\infty=\left\|A^* P_{H_{\Omega_s}}^{\perp} y\right\|_\infty\leq\left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon.$$

Therefore, OMP terminates after performing the s-th iteration.

Thus OMP performs s iterations. This completes the proof.

C. Proof of Proposition 1

Notice that

$$A_{\Omega_s}\hat{x}=P_{H_{\Omega_s}}y=P_{H_{\Omega_s}}\left(Ax+v\right)$$
$$=P_{H_{\Omega_s}}\left(A_{\Omega_s}x_{\Omega_s}+v\right)=A_{\Omega_s}x_{\Omega_s}+P_{H_{\Omega_s}}v. \quad (22)$$

From (22), (p2), and the fact that A satisfies the RIP of order m, we obtain that for $\|v\|_2\leq\varepsilon$,

$$\sqrt{1-\delta_m}\left\|\hat{x}-x_{\Omega_s}\right\|_2\leq\left\|A_{\Omega_s}\left(\hat{x}-x_{\Omega_s}\right)\right\|_2=\left\|P_{H_{\Omega_s}}v\right\|_2\leq\varepsilon,$$

i.e., $\left\|\hat{x}-x_{\Omega_s}\right\|_2\leq\frac{\varepsilon}{\sqrt{1-\delta_m}}$, and for $\left\|A^* v\right\|_\infty\leq\varepsilon$,

$$\left(1-\delta_m\right)\left\|\hat{x}-x_{\Omega_s}\right\|_2^2$$
$$\leq\left\|A_{\Omega_s}\left(\hat{x}-x_{\Omega_s}\right)\right\|_2^2=\left\|P_{H_{\Omega_s}}v\right\|_2^2=\left\|A_{\Omega_s}\left(A_{\Omega_s}^* A_{\Omega_s}\right)^{-1}A_{\Omega_s}^* v\right\|_2^2$$
$$=v^* A_{\Omega_s}\left(A_{\Omega_s}^* A_{\Omega_s}\right)^{-1}A_{\Omega_s}^* v \quad (23)$$

and

$$\left(1-\delta_m\right)I_s\leq A_{\Omega_s}^* A_{\Omega_s}\leq\left(1+\delta_m\right)I_s. \quad (24)$$

It follows from (24) that

$$\frac{1}{1+\delta_m}I_s\leq\left(A_{\Omega_s}^* A_{\Omega_s}\right)^{-1}\leq\frac{1}{1-\delta_m}I_s. \quad (25)$$

By combining (23) and (25) we have

$$\left(1-\delta_m\right)\left\|\hat{x}-x_{\Omega_s}\right\|_2^2\leq\frac{1}{1-\delta_m}\left\|A_{\Omega_s}^* v\right\|_2^2. \quad (26)$$

Finally, from (26) we have

$$\left\|\hat{x}-x_{\Omega_s}\right\|_2\leq\frac{1}{1-\delta_m}\left\|A_{\Omega_s}^* v\right\|_2\leq\frac{\sqrt{s}}{1-\delta_m}\left\|A_{\Omega_s}^* v\right\|_\infty$$
$$\leq\frac{\sqrt{s}}{1-\delta_m}\left\|A^* v\right\|_\infty\leq\frac{\sqrt{s}}{1-\delta_m}\varepsilon$$

This completes the proof.

D. Proof of Theorem 3

Also, we prove the theorem by induction. Suppose that OMP selects correct indexes in the first k iterations, *i.e.*, $\Lambda_k\subset\Omega_s$ with $k<s$. The assumption holds true naturally when $k=0$ since $\Lambda_0=\varnothing\subset\Omega_s$. Then it suffices to prove that $S_{0\Lambda_k}-S_{j\Lambda_k}>0$ for all $s<j\leq N$. Notice that $\hat{x}=x_{\Omega_s}$ in Lemma 1 when $\|v\|_2=0$ and that $\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}\leq R(x)$. Thus it follows from (7) and (8) that

$$S_{0\Lambda_k}-S_{j\Lambda_k}$$
$$\geq\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2^2}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}-\delta_{K+1}\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}\sqrt{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1^2+\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2^2}$$
$$=\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}\sqrt{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1^2+\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2^2}\left(\frac{1}{\sqrt{1+\left(\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}\right)^2}}-\delta_{K+1}\right)$$
$$\geq\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}\sqrt{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1^2+\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2^2}\left(\frac{1}{\sqrt{1+\mathrm{R}^2(x)}}-\delta_{K+1}\right)$$
$$>\frac{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2}{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1}\sqrt{\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_1^2+\left\|x_{\Omega_s\setminus\Lambda_k}\right\|_2^2}\left(\frac{1}{\sqrt{1+\frac{1-\delta_{K+1}^2}{\delta_{K+1}^2}}}-\delta_{K+1}\right)=0$$

This completes the proof.

IV. Conclusion

In this paper, we study the sufficient conditions for recovering the support of sparse signals via OMP based on RIP in both the noisy and noiseless cases. In the noisy case, our conditions are weaker than that in [8] in terms of the minimum magnitude of the nonzero elements of the signal under both the l_2 and l_∞ bounded noise. In particular, when δ_{K+1} is close to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$1/\sqrt{K+1}$ and K is large enough, our conditions are a significant improvement over the sufficient conditions in previous works and are close to the necessary conditions proposed in [8]. In addition, we also give the upper bound of the error between the recovered signal and the original sparse signal. In the noiseless case, we have shown that under some constraints on the signal, OMP can also exactly recover the signal suppose that δ_{K+1} satisfies $\delta_{K+1} < \sqrt{2}/2$. It is worth mentioning that our perspective may provide a route to study the condition for recovering restricted classes of K-sparse signals (*e.g.*, strongly-decaying signals) via OMP.

APPENDIX I

A. Proof of Lemma 1

Without loss of generality, let $\Lambda_k = \Omega_k$. Then it follows that $H_{\Lambda_k} = H_{\Omega_k}$. If $\left\langle P^{\perp}_{H_{\Omega_k}} y, \alpha_j \right\rangle = e^{i\theta_j} \left| \left\langle P^{\perp}_{H_{\Omega_k}} y, \alpha_j \right\rangle \right|$, then we have $\left\langle P^{\perp}_{H_{\Omega_k}} y, \tilde{\alpha}_j \right\rangle = S_{j\Omega_k}$ where $\tilde{\alpha}_j = e^{i\theta_j} \alpha_j$. For any $J_m = \{ j_1, j_2, \cdots, j_m \} \subset \Omega_s$, mark $\tilde{A}_{J_m,j} = \left[\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_m}, \tilde{\alpha}_j \right]$. Notice that $\tilde{A}_{J_m,j} = A_{J_m \cup \{j\}} U$ with $U = diag\left\{1,1,\cdots,1,e^{i\theta_j}\right\}$. Thus we have $\tilde{A}^*_{J_m,j} \tilde{A}_{J_m,j} = U^* A^*_{J_m \cup \{j\}} A_{J_m \cup \{j\}} U$. It is clear that $\left| J_m \cup \{j\} \right| \le K+1$ and $U^* U = I_{m+1}$. Then by (p2) we obtain

$$\left(1-\delta_{K+1}\right) I_{m+1} \le \tilde{A}^*_{J_m,j} \tilde{A}_{J_m,j} \le \left(1+\delta_{K+1}\right) I_{m+1}. \tag{27}$$

From (p1) and the fact that $P^{\perp}_{H_{\Omega_k}} P^{\perp}_{H_{\Omega_s}} = P^{\perp}_{H_{\Omega_s}} P^{\perp}_{H_{\Omega_k}} = P^{\perp}_{H_{\Omega_s}}$, $P^{\perp}_{H_{\Omega_k}} \alpha_i = 0$ with $1 \le i \le k$ and $P^{\perp}_{H_{\Omega_s}} y = P^{\perp}_{H_{\Omega_s}} \left(Ax+v \right) = P^{\perp}_{H_{\Omega_s}} \left(A_{\Omega_s} x_{\Omega_s} + v \right) = P^{\perp}_{H_{\Omega_s}} v$, it follows that for any $t > 0$,

$$\begin{aligned}
&\left\| \left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} y - P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j \right\|_2^2 \\
&- \left\| \left(t - \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} y + P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j \right\|_2^2 \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\| P^{\perp}_{H_{\Omega_k}} y \right\|_2^2 - 4t \operatorname{Re} \left\langle P^{\perp}_{H_{\Omega_k}} y, \tilde{\alpha}_j \right\rangle \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, \ P^{\perp}_{H_{\Omega_k}} y \right\rangle - 4t S_{j\Omega_k} \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, P^{\perp}_{H_{\Omega_k}} \left(P_{H_{\Omega_s}} y \right) \right\rangle \\
&\quad + \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, P^{\perp}_{H_{\Omega_k}} \left(P^{\perp}_{H_{\Omega_s}} y \right) \right\rangle - 4t S_{j\Omega_k} \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, P^{\perp}_{H_{\Omega_k}} \left(\sum_{i=1}^{s} \hat{x}_i \alpha_i \right) \right\rangle \\
&\quad + \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, P^{\perp}_{H_{\Omega_s}} v \right\rangle - 4t S_{j\Omega_k} \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, \sum_{i=k+1}^{s} \hat{x}_i \alpha_i \right\rangle \\
&\quad + \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle y, P^{\perp}_{H_{\Omega_k}} \left(P^{\perp}_{H_{\Omega_s}} v \right) \right\rangle - 4t S_{j\Omega_k} \\
&= \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle P^{\perp}_{H_{\Omega_k}} y, \sum_{i=k+1}^{s} \hat{x}_i \alpha_i \right\rangle + \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\langle y, P^{\perp}_{H_{\Omega_s}} v \right\rangle - 4t S_{j\Omega_k} .
\end{aligned} \tag{28}$$

Notice that

$$\left| \left\langle P^{\perp}_{H_{\Omega_k}} y, \sum_{i=k+1}^{s} \hat{x}_i \alpha_i \right\rangle \right| \le \left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1 S_{0\Omega_k} \tag{29}$$

and

$$\left\langle y, P^{\perp}_{H_{\Omega_s}} v \right\rangle = \left\langle P^{\perp}_{H_{\Omega_s}} y, P^{\perp}_{H_{\Omega_s}} v \right\rangle = \left\| P^{\perp}_{H_{\Omega_s}} v \right\|_2^2 . \tag{30}$$

Thus it follows from (28), (29) and (30) that

$$\begin{aligned}
&\left\| \left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} y - P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j \right\|_2^2 \\
&- \left\| \left(t - \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} y + P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j \right\|_2^2 \\
&\le 4t \left(S_{0\Omega_k} - S_{j\Omega_k} \right) + \frac{4t}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \left\| P^{\perp}_{H_{\Omega_s}} v \right\|_2^2 .
\end{aligned} \tag{31}$$

Let

$$\xi = \left(\left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) \hat{x}_{k+1}, \cdots, \left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) \hat{x}_s, -1 \right)^T$$

and

$$\eta = \left(\left(t - \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) \hat{x}_{k+1}, \cdots, \left(t - \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) \hat{x}_s, 1 \right)^T .$$

Then

$$\begin{aligned}
&\left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} y - P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j \\
&= \left(t + \frac{1}{\left\| \hat{x}_{\Omega_s \setminus \Omega_k} \right\|_1} \right) P^{\perp}_{H_{\Omega_k}} \left(P_{H_{\Omega_s}} y \right) - P^{\perp}_{H_{\Omega_k}} \tilde{\alpha}_j
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$$+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_k}}^{\perp}\left(P_{H_{\Omega_s}}^{\perp}y\right)$$

$$=P_{H_{\Omega_k}}^{\perp}\left(\sum_{i=k+1}^{s}\hat{x}_i\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\alpha_i-\tilde{\alpha}_j\right)+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_s}}^{\perp}v$$

$$=P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_s}}^{\perp}v.$$

It is easy to check that

$$\left\langle P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi,\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_s}}^{\perp}v\right\rangle$$

$$=\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\left\langle P_{H_{\Omega_s}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi,v\right\rangle$$

$$=-\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\left\langle P_{H_{\Omega_s}}^{\perp}\tilde{\alpha}_j,v\right\rangle$$

Thus we have

$$\left\|\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_k}}^{\perp}y-P_{H_{\Omega_k}}^{\perp}\tilde{\alpha}_j\right\|_2^2$$

$$=\left\|P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_s}}^{\perp}v\right\|_2^2$$

$$=\left\|P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi\right\|_2^2+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)^2\left\|P_{H_{\Omega_s}}^{\perp}v\right\|_2^2$$

$$-2\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\operatorname{Re}\left\langle P_{H_{\Omega_s}}^{\perp}\tilde{\alpha}_j,v\right\rangle$$

$$=\xi^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}A_{\Omega_s\setminus\Omega_k,j}\xi+\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)^2\left\|P_{H_{\Omega_s}}^{\perp}v\right\|_2^2$$

$$-2\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\operatorname{Re}\left\langle\tilde{\alpha}_j,P_{H_{\Omega_s}}^{\perp}v\right\rangle. \quad (32)$$

Similarly, we have

$$\left\|\left(t-\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_k}}^{\perp}y+P_{H_{\Omega_k}}^{\perp}\tilde{\alpha}_j\right\|_2^2$$

$$=\eta^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}A_{\Omega_s\setminus\Omega_k,j}\eta+\left(t-\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)^2\left\|P_{H_{\Omega_s}}^{\perp}v\right\|_2^2$$

$$+2\left(t-\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)\operatorname{Re}\left\langle\tilde{\alpha}_j,P_{H_{\Omega_s}}^{\perp}v\right\rangle. \quad (33)$$

From (32) and (33), it follows that

$$\left\|\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_k}}^{\perp}y-P_{H_{\Omega_k}}^{\perp}\tilde{\alpha}_j\right\|_2^2$$

$$-\left\|\left(t-\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)P_{H_{\Omega_k}}^{\perp}y+P_{H_{\Omega_k}}^{\perp}\tilde{\alpha}_j\right\|_2^2$$

$$=\xi^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}A_{\Omega_s\setminus\Omega_k,j}\xi-\eta^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}A_{\Omega_s\setminus\Omega_k,j}\eta$$

$$+\frac{4t}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\left\|P_{H_{\Omega_s}}^{\perp}v\right\|_2^2-4t\operatorname{Re}\left\langle\tilde{\alpha}_j,P_{H_{\Omega_s}}^{\perp}v\right\rangle. \quad (34)$$

Partition $\tilde{A}_{\Omega_s,j}=\left[\alpha_1,\alpha_2,\cdots,\alpha_s,\tilde{\alpha}_j\right]$ as $\tilde{A}_{\Omega_s,j}=\left[A_{\Omega_k},\tilde{A}_{\Omega_s\setminus\Omega_k,j}\right]$. By some simple calculations, we have

$$\tilde{A}_{\Omega_s,j}^*\tilde{A}_{\Omega_s,j}=\begin{bmatrix}A_{\Omega_k}^*A_{\Omega_k} & A_{\Omega_k}^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}\\ \tilde{A}_{\Omega_s\setminus\Omega_k,j}^*A_{\Omega_k} & \tilde{A}_{\Omega_s\setminus\Omega_k,j}^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}\end{bmatrix}.$$

Thus it follows that

$$C_k\left(\tilde{A}_{\Omega_s,j}^*\tilde{A}_{\Omega_s,j}\right)$$

$$=\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}-\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*A_{\Omega_k}\left(A_{\Omega_k}^*A_{\Omega_k}\right)^{-1}A_{\Omega_k}^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}$$

$$=\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}. \quad (35)$$

By combining (p3) and (27), we obtain

$$\left(1-\delta_{K+1}\right)I_{s-k+1}\le C_k\left(\tilde{A}_{\Omega_s,j}^*\tilde{A}_{\Omega_s,j}\right)\le\left(1+\delta_{K+1}\right)I_{s-k+1}. \quad (36)$$

Therefore, from (p2), (35), and (36), we have

$$\xi^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi$$

$$=\xi^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*\left(I_n-A_{\Omega_k}\left(A_{\Omega_k}^*A_{\Omega_k}\right)^{-1}A_{\Omega_k}^*\right)\tilde{A}_{\Omega_s\setminus\Omega_k,j}\xi$$

$$=\xi^*C_k\left(\tilde{A}_{\Omega_s,j}^*\tilde{A}_{\Omega_s,j}\right)\xi$$

$$\ge\left(1-\delta_{K+1}\right)\xi^*\xi=\left(1-\delta_{K+1}\right)\left[\left(t+\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)^2\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2+1\right]. \quad (37)$$

Similarly, we have

$$\eta^*\tilde{A}_{\Omega_s\setminus\Omega_k,j}^*P_{H_{\Omega_k}}^{\perp}\tilde{A}_{\Omega_s\setminus\Omega_k,j}\eta$$

$$=\eta^*C_k\left(\tilde{A}_{\Omega_s,j}^*\tilde{A}_{\Omega_s,j}\right)\eta$$

$$\le\left(1+\delta_{K+1}\right)\left[\left(t-\frac{1}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\right)^2\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2+1\right]. \quad (38)$$

It follows from (31), (34), (37) and (38) that

$$S_{0\Omega_k}-S_{j\Omega_k}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

$$\geq \frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1} - \frac{\delta_{K+1}}{2}\left(t\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2 + \frac{1}{t}\left(\frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1^2}+1\right)\right)$$
$$-\operatorname{Re}\left\langle \tilde{\alpha}_j, P_{H_{\Omega_s}}^{\perp} v\right\rangle. \quad (39)$$

By using arithmetic-geometric mean inequality to (39), we obtain

$$S_{0\Omega_k} - S_{j\Omega_k}$$
$$\geq \max_{t>0}\left\{\frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1} - \frac{\delta_{K+1}}{2}\left(t\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2 + \frac{1}{t}\left(\frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1^2}+1\right)\right) - \operatorname{Re}\left\langle \tilde{\alpha}_j, P_{H_{\Omega_k}}^{\perp} v\right\rangle\right\}$$
$$= \frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1} - \delta_{K+1}\frac{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2}{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1}\sqrt{\left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_1^2 + \left\|\hat{x}_{\Omega_s\setminus\Omega_k}\right\|_2^2} - \operatorname{Re}\left\langle \tilde{\alpha}_j, P_{H_{\Omega_s}}^{\perp} v\right\rangle.$$

So, the lemma holds.

B. Proof of Lemma 2

When $j \leq s$, the lemma holds evidently since $\left|\left\langle \alpha_j, P_{H_{\Omega_s}}^{\perp} v\right\rangle\right| = \left|\left\langle P_{H_{\Omega_s}}^{\perp}\alpha_j, v\right\rangle\right| = 0$. In the following, we prove the lemma still holds when $j > s$. Mark $P_{H_{\Omega_s}}\alpha_j = \sum_{i=1}^{s} z_i\alpha_i = A_{\Omega_s}Z$ with $Z = (z_1, z_2, \cdots, z_s)^T$. Then by (p2), we have

$$\left\|P_{H_{\Omega_s}}\alpha_j\right\|_2^2 = \left\|A_{\Omega_s}Z\right\|_2^2 = Z^* A_{\Omega_s}^* A_{\Omega_s} Z \geq (1-\delta_{K+1})\|Z\|_2^2 \quad (40)$$

and for any $t > 0$,

$$\left\|P_{H_{\Omega_s}}\alpha_j\right\|_2^2 = \left\langle P_{H_{\Omega_s}}\alpha_j, P_{H_{\Omega_s}}\alpha_j\right\rangle = \operatorname{Re}\left\langle P_{H_{\Omega_s}}\alpha_j, \sum_{i=1}^{s} z_i\alpha_i\right\rangle$$
$$= \operatorname{Re}\left\langle \alpha_j, P_{H_{\Omega_s}}\left(\sum_{i=1}^{s} z_i\alpha_i\right)\right\rangle = \operatorname{Re}\left\langle \alpha_j, \sum_{i=1}^{s} z_i\alpha_i\right\rangle$$
$$= \frac{1}{4t}\left(\left\|\sum_{i=1}^{s} z_i\alpha_i + t\alpha_j\right\|_2^2 - \left\|\sum_{i=1}^{s} z_i\alpha_i - t\alpha_j\right\|_2^2\right)$$
$$\leq \frac{1}{4t}\left((1+\delta_{K+1})\left(\|Z\|_2^2 + t^2\right) - (1-\delta_{K+1})\left(\|Z\|_2^2 + t^2\right)\right)$$
$$= \frac{\delta_{K+1}}{2t}\left(\|Z\|_2^2 + t^2\right) = \frac{\delta_{K+1}}{2}\left(\frac{1}{t}\|Z\|_2^2 + t\right). \quad (41)$$

(40) and (41) together yield

$$(1-\delta_{K+1})\|Z\|_2^2 \leq \min_{t>0}\left\{\frac{\delta_{K+1}}{2}\left(\frac{1}{t}\|Z\|_2^2 + t\right)\right\} = \delta_{K+1}\|Z\|_2,$$

i.e., $\|Z\|_2 \leq \frac{\delta_{K+1}}{1-\delta_{K+1}}$. Notice that $\left\|A^* v\right\|_\infty = \max_{1\leq i\leq N}\left|\left\langle \alpha_i, v\right\rangle\right|$. Thus from (p4) and the fact that $\left\|A^* v\right\|_\infty \leq \varepsilon$, we have

$$\left|\left\langle P_{H_{\Omega_s}}^{\perp}\alpha_j, v\right\rangle\right| = \left|\left\langle \alpha_j - P_{H_{\Omega_s}}\alpha_j, v\right\rangle\right| \leq \left|\left\langle \alpha_j, v\right\rangle\right| + \left|\left\langle P_{H_{\Omega_s}}\alpha_j, v\right\rangle\right|$$
$$\leq \varepsilon + \left|\left\langle \sum_{i=1}^{m} z_i\alpha_i, v\right\rangle\right| \leq \varepsilon + \|Z\|_1\varepsilon \leq \varepsilon + \sqrt{s}\|Z\|_2\varepsilon$$
$$\leq \left(1 + \frac{\sqrt{s}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon.$$

So, the lemma holds.

C. Proof of Lemma 3

Mark $P_{H_{\Omega_s}} v = \sum_{i=1}^{s} v_i\alpha_i = A_{\Omega_s}V$ with $V = (v_1, v_2, \cdots, v_s)^T$. Notice that

$$\sum_{i=1}^{s}\hat{x}_i\alpha_i = P_{H_{\Omega_s}} y = P_{H_{\Omega_s}}(Ax+v) = Ax + P_{H_{\Omega_s}} v = \sum_{i=1}^{s}(x_i+v_i)\alpha_i.$$

Thus we have $\hat{x}_i = x_i + v_i$ for any $1 \leq i \leq s$. It is obvious that (10) and (11) hold for $V = 0$. In the following, we prove that (10) and (11) still hold for $V \neq 0$.

First, we prove (10). From (p2), we have

$$\left\|P_{H_{\Omega_s}} v\right\|_2^2 = \left\|A_{\Omega_s}V\right\|_2^2 = V^* A_{\Omega_s}^* A_{\Omega_s} V \geq (1-\delta_{K+1})\|V\|_2^2. \quad (42)$$

For $\|v\|_2 \leq \varepsilon$, it follows from (p1) that

$$\left\|P_{H_{\Omega_s}} v\right\|_2 \leq \varepsilon. \quad (43)$$

By combining (p4), (42) and (43), we have

$$\|V\|_\infty \leq \|V\|_2 \leq \frac{\varepsilon}{\sqrt{1-\delta_{K+1}}}. \quad (44)$$

Then it follows from (44) that

$$\min_{1\leq i\leq s}|\hat{x}_i| = \min_{1\leq i\leq s}|x_i + v_i| \geq \min_{1\leq i\leq s}|x_i| - \|V\|_\infty \geq \min_{1\leq i\leq s}|x_i| - \frac{\varepsilon}{\sqrt{1-\delta_{K+1}}},$$

i.e., (10) holds.

Next, we prove (11). Notice that $\left\|A^* v\right\|_\infty = \max_{1\leq i\leq N}\left|\left\langle \alpha_i, v\right\rangle\right|$. Then from $\left\|A^* v\right\|_\infty \leq \varepsilon$ we have

$$\left\|P_{H_{\Omega_s}} v\right\|_2^2 = \left\langle P_{H_{\Omega_s}} v, P_{H_{\Omega_s}} v\right\rangle = \left\langle P_{H_{\Omega_s}} v, \sum_{i=1}^{s} v_i\alpha_i\right\rangle$$
$$= \left\langle v, P_{H_{\Omega_s}}\left(\sum_{i=1}^{s} v_i\alpha_i\right)\right\rangle = \left|\left\langle v, \sum_{i=1}^{s} v_i\alpha_i\right\rangle\right| \leq \|V\|_1\varepsilon. \quad (45)$$

By combining (p4), (42) and (45), we obtain

$$\|V\|_\infty \leq \|V\|_2 \leq \frac{\varepsilon}{1-\delta_{K+1}}\frac{\|V\|_1}{\|V\|_2} \leq \frac{\sqrt{s}\varepsilon}{1-\delta_{K+1}} \leq \frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}}. \quad (46)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2017.2711543, IEEE Transactions on Signal Processing

It follows from (46) that

$$\min_{1\le i\le s}|\hat{x}_i| = \min_{1\le i\le s}|x_i+v_i| \ge \min_{1\le i\le s}|x_i| - \|V\|_\infty \ge \min_{1\le i\le s}|x_i| - \frac{\sqrt{K}\varepsilon}{1-\delta_{K+1}},$$

i.e., (11) holds. Therefore, the lemma holds.

D. Proof of Lemma 4

First we prove (12). For $\|v\|_2 \le \varepsilon$ and $Ax = A_{\Omega_s}x_{\Omega_s}$, we have $\left\|P^\perp_{H_{\Omega_j}} v\right\|_2 \le \varepsilon$ with $0 \le j \le N$. Since $H_{\Omega_0} \subset H_{\Omega_1} \cdots \subset H_{\Omega_{s-1}}$, it is clear that $H^\perp_{\Omega_0} \supset H^\perp_{\Omega_1} \cdots \supset H^\perp_{\Omega_{s-1}}$. Thus we have $\left\|P^\perp_{H_{\Omega_0}} y\right\|_2 \ge \left\|P^\perp_{H_{\Omega_1}} y\right\|_2 \ge \left\|P^\perp_{H_{\Omega_{s-1}}} y\right\|_2$. Notice that

$$\begin{aligned}\left\|P^\perp_{H_{\Omega_{s-1}}} y\right\|_2 &= \left\|P^\perp_{H_{\Omega_{s-1}}}\left(A_{\Omega_s}x_{\Omega_s}+v\right)\right\|_2 \ge \left\|P^\perp_{H_{\Omega_{s-1}}}A_{\Omega_s}x_{\Omega_s}\right\|_2 - \left\|P^\perp_{H_{\Omega_{s-1}}}v\right\|_2 \\ &\ge \left\|P^\perp_{H_{\Omega_{s-1}}}A_{\Omega_s}x_{\Omega_s}\right\|_2 - \varepsilon. \qquad (47)\end{aligned}$$

It is easy to check that

$$P^\perp_{H_{\Omega_{s-1}}}A_{\Omega_s}x_{\Omega_s} = P^\perp_{H_{\Omega_{s-1}}}A_{\Omega_s\setminus\Omega_{s-1}}x_{\Omega_s\setminus\Omega_{s-1}} = x_s P^\perp_{H_{\Omega_{s-1}}}\alpha_s \qquad (48)$$

and

$$\begin{aligned}C_{s-1}\left(A^*_{\Omega_s}A_{\Omega_s}\right) &= \alpha_s^*\alpha_s - \alpha_s^*A_{\Omega_{s-1}}\left(A^*_{\Omega_{s-1}}A_{\Omega_{s-1}}\right)^{-1}A^*_{\Omega_{s-1}}\alpha_s \\ &= \alpha_s^* P^\perp_{H_{\Omega_{s-1}}}\alpha_s. \qquad (49)\end{aligned}$$

By combining (p2), (p3), (48) and (49) we have

$$\begin{aligned}\left\|P^\perp_{H_{\Omega_{s-1}}}A_{\Omega_s}x_{\Omega_s}\right\|_2^2 &= |x_s|^2\alpha_s^*P^\perp_{H_{\Omega_{s-1}}}\alpha_s = |x_s|^2 C_{s-1}\left(A^*_{\Omega_s}A_{\Omega_s}\right) \\ &\ge \left(1-\delta_{K+1}\right)\left(\min_{1\le i\le s}|x_i|\right)^2. \qquad (50)\end{aligned}$$

Then (47) and (50) together yield (12).

Next we prove (13) and (14). For $\left\|A^*v\right\|_\infty \le \varepsilon$ and $Ax = A_{\Omega_s}x_{\Omega_s}$, it follows from (9) that

$$\left\|A^*P^\perp_{H_{\Omega_s}}y\right\|_\infty = \max_{s<j\le N}\left|\left\langle \alpha_j, P^\perp_{H_{\Omega_s}}y\right\rangle\right| \le \left(1+\frac{\sqrt{K}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon$$

i.e., (13) holds.

Notice that for $k < s$,

$$\begin{aligned}\left\|A^*P^\perp_{H_{\Omega_k}}y\right\|_\infty &= \left\|A^*P^\perp_{H_{\Omega_k}}\left(Ax+v\right)\right\|_\infty \\ &\ge \left\|A^*P^\perp_{H_{\Omega_k}}Ax\right\|_\infty - \left\|A^*P^\perp_{H_{\Omega_k}}v\right\|_\infty. \qquad (51)\end{aligned}$$

From (p2) and (p3), we have

$$\begin{aligned}&\left\|P^\perp_{H_{\Omega_k}}Ax\right\|_2^2 \\ &= \left\|P^\perp_{H_{\Omega_k}}A_{\Omega_s}x_{\Omega_s}\right\|_2^2 = \left\|P^\perp_{H_{\Omega_k}}A_{\Omega_s\setminus\Omega_k}x_{\Omega_s\setminus\Omega_k}\right\|_2^2 \\ &= x^*_{\Omega_s\setminus\Omega_k}A^*_{\Omega_s\setminus\Omega_k}P^\perp_{H_{\Omega_k}}A_{\Omega_s\setminus\Omega_k}x_{\Omega_s\setminus\Omega_k} \\ &= x^*_{\Omega_s\setminus\Omega_k}A^*_{\Omega_s\setminus\Omega_k}\left(I_n - A_{\Omega_k}\left(A^*_{\Omega_k}A_{\Omega_k}\right)^{-1}A^*_{\Omega_k}\right)A_{\Omega_s\setminus\Omega_k}x_{\Omega_s\setminus\Omega_k} \\ &= x^*_{\Omega_s\setminus\Omega_k}C_k\left(A^*_{\Omega_s}A_{\Omega_s}\right)x_{\Omega_s\setminus\Omega_k} \ge \left(1-\delta_{K+1}\right)\left\|x_{\Omega_s\setminus\Omega_k}\right\|_2^2 \qquad (52)\end{aligned}$$

Meanwhile, it holds that

$$\begin{aligned}&\left\|P^\perp_{H_{\Omega_k}}Ax\right\|_2^2 \\ &= \left|\left\langle P^\perp_{H_{\Omega_k}}Ax, P^\perp_{H_{\Omega_k}}\sum_{i=1}^{s}x_i\alpha_i\right\rangle\right| = \left|\left\langle P^\perp_{H_{\Omega_k}}Ax, \sum_{i=k+1}^{s}x_i\alpha_i\right\rangle\right| \\ &= \left|\sum_{i=k+1}^{s}\overline{x}_i\left\langle P^\perp_{H_{\Omega_k}}Ax, \alpha_i\right\rangle\right| \le \left\|x_{\Omega_s\setminus\Omega_k}\right\|_1 \max_{1\le i\le s}\left|\left\langle P^\perp_{H_{\Omega_k}}Ax,\alpha_i\right\rangle\right| \\ &= \left\|x_{\Omega_s\setminus\Omega_k}\right\|_1\left\|A^*_{\Omega_s}P^\perp_{H_{\Omega_k}}Ax\right\|_\infty \le \left\|x_{\Omega_s\setminus\Omega_k}\right\|_1\left\|A^*P^\perp_{H_{\Omega_k}}Ax\right\|_\infty \qquad (53)\end{aligned}$$

Then, from (p4), (52) and (53) we obtain

$$\begin{aligned}&\left\|A^*P^\perp_{H_{\Omega_k}}Ax\right\|_\infty \\ &\ge \left(1-\delta_{K+1}\right)\frac{\left\|x_{\Omega_s\setminus\Omega_k}\right\|_2^2}{\left\|x_{\Omega_s\setminus\Omega_k}\right\|_1} \ge \frac{1-\delta_{K+1}}{\sqrt{s-k}}\left\|x_{\Omega_s\setminus\Omega_k}\right\|_2 \\ &\ge \frac{1-\delta_{K+1}}{\sqrt{s-k}}\sqrt{s-k}\min_{1\le i\le s}|x_i| = \left(1-\delta_{K+1}\right)\min_{1\le i\le s}|x_i| \qquad (54)\end{aligned}$$

It follows from Lemma 2 that

$$\begin{aligned}\left\|A^*P^\perp_{H_{\Omega_k}}v\right\|_\infty &= \max_{1\le i\le N}\left|\left\langle\alpha_i, P^\perp_{H_{\Omega_k}}v\right\rangle\right| \\ &= \max_{1\le i\le N}\left|\left\langle P^\perp_{H_{\Omega_k}}\alpha_i, v\right\rangle\right| \le \left(1+\frac{\sqrt{k}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \\ &\le \left(1+\frac{\sqrt{s-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \le \left(1+\frac{\sqrt{K-1}\delta_{K+1}}{1-\delta_{K+1}}\right)\varepsilon \qquad (55)\end{aligned}$$

Then (51), (54) and (55) together yield (14). Thus, the lemma holds.

Appendix II

Here, we verify that $R(x) \le \sqrt{g_K(\alpha)}$ if x is an α-strongly-decaying K-sparse signal by exploiting the theory of majorization [20]. First, we need to introduce the following basic notions.

Let $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$. Let $x^\downarrow$ be the vectors obtained by rearranging the coordinates of x in the decreasing orders. Thus, if $x^\downarrow = \left(x_1^\downarrow, x_2^\downarrow, \cdots, x_n^\downarrow\right)^T$, then $x_1^\downarrow \ge x_2^\downarrow \ge \cdots \ge x_n^\downarrow$. Let $x, y \in \mathbb{R}^n$. We say that x is *majorised* by y, in symbols $x \prec y$, if

$$\sum_{j=1}^{n}x_j^\downarrow = \sum_{j=1}^{n}y_j^\downarrow,\quad \sum_{j=1}^{k}x_j^\downarrow \le \sum_{j=1}^{k}y_j^\downarrow,\quad 1\le k\le n-1.$$

A real valued function φ on $\mathbb{R}^n$ is called *Schur-convex* if

$$x \prec y \Rightarrow \varphi(x) \le \varphi(y).$$

The proof can be divided into the following two steps.

Step 1: Let $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$ where $x_i \ge \alpha x_{i+1} > 0$ with $\alpha > 1$ and $i = 1, 2, \cdots, n-1$. Then we need to prove the following majorisation relation.

$$\left(\frac{1}{\alpha\sum_{j=1}^{n}\alpha^{-j}},\frac{1}{\alpha^2\sum_{j=1}^{n}\alpha^{-j}},\cdots,\frac{1}{\alpha^n\sum_{j=1}^{n}\alpha^{-j}}\right)^T \prec \left(\frac{x_1}{\|x\|_1},\frac{x_2}{\|x\|_1},\cdots,\frac{x_n}{\|x\|_1}\right)^T. \quad (56)$$

Obviously,

$$\sum_{i=1}^{n}\frac{1}{\alpha^i\sum_{j=1}^{n}\alpha^{-j}}=\sum_{j=1}^{n}\frac{x_j}{\|x\|_1}=1.$$

Thus it suffices to verify that

$$\sum_{i=1}^{k}\frac{1}{\alpha^i\sum_{j=1}^{n}\alpha^{-j}}\le\sum_{j=1}^{k}\frac{x_j}{\|x\|_1}$$

for $1\le k\le n-1$. Notice that

$$\sum_{j=1}^{k}x_j\ge\left(\sum_{j=1}^{k}\alpha^{k-j}\right)x_k,\ \sum_{j=k+1}^{n}x_j\le\left(\sum_{j=k+1}^{n}\alpha^{k-j}\right)x_k$$

Thus we have

$$\frac{\|x\|_1}{\sum_{j=1}^{k}x_j}=\frac{\sum_{j=1}^{n}x_j}{\sum_{j=1}^{k}x_j}=1+\frac{\sum_{j=k+1}^{n}x_j}{\sum_{j=1}^{k}x_j}\le 1+\frac{\sum_{j=k+1}^{n}\alpha^{k-j}}{\sum_{j=1}^{k}\alpha^{k-j}}=\frac{\sum_{j=1}^{n}\alpha^{-j}}{\sum_{j=1}^{k}\alpha^{-j}}.$$

Then,

$$\sum_{i=1}^{k}\frac{1}{\alpha^i\sum_{j=1}^{n}\alpha^{-j}}=\frac{\sum_{j=1}^{k}\alpha^{-j}}{\sum_{j=1}^{n}\alpha^{-j}}\le\sum_{j=1}^{k}\frac{x_j}{\|x\|_1}.$$

i.e., (56) holds.

Step 2: Suppose that x is an α-strongly-decaying *K*-sparse signal. Without loss of generality, we assume that $\mathrm{supp}(x)=\Omega_s$ with $s\le K$ and that all the nonzero entries of x are in the decreasing orders, i.e., $|x_1|\ge|x_2|\ge\cdots\ge|x_s|>0$. It is easy to check that

$$R(x)=\max\left\{\frac{\|x_\Lambda\|_1}{\|x_\Lambda\|_2}\middle|\Lambda\subset\Omega_N,\Lambda\ne\varnothing\right\}$$
$$=\max\left\{\frac{\|x_\Lambda\|_1}{\|x_\Lambda\|_2}\middle|\Lambda\subset\Omega_s,\Lambda\ne\varnothing\right\}. \quad (57)$$

To show $R(x)\le\sqrt{g_K(\alpha)}$, we only need to show

$$\frac{\|x_\Lambda\|_1}{\|x_\Lambda\|_2}\le\sqrt{g_k(\alpha)}$$

for any $\Lambda=\{\lambda_1,\ \lambda_2,\cdots,\ \lambda_m\}\subset\Omega_s$ with $1\le\lambda_1<\lambda_2<\cdots<\lambda_m\le s$. Obviously,

$$\left|x_{\lambda_i}\right|\ge\alpha\left|x_{\lambda_{i+1}}\right|,i=1,2,\cdots,m-1. \quad (58)$$

One can verify that $\varphi(x)=\sum_{i=1}^{m}x_i^2$ is Schur-convex on $\mathbb{R}^m$ (see [20, Theorem II.3.14]). From (56) and (58) we have

$$\frac{\|x_\Lambda\|_2^2}{\|x_\Lambda\|_1^2}=\varphi\left(\frac{|x_{\lambda_1}|}{\|x_\Lambda\|_1},\frac{|x_{\lambda_2}|}{\|x_\Lambda\|_1},\cdots,\frac{|x_{\lambda_m}|}{\|x_\Lambda\|_1}\right)$$
$$\ge\varphi\left(\frac{1}{\alpha\sum_{j=1}^{m}\alpha^{-j}},\frac{1}{\alpha^2\sum_{j=1}^{m}\alpha^{-j}},\cdots,\frac{1}{\alpha^m\sum_{j=1}^{m}\alpha^{-j}}\right)$$
$$=\frac{\sum_{j=1}^{m}\alpha^{-2j}}{\left(\sum_{j=1}^{m}\alpha^{-j}\right)^2}=\frac{(\alpha^m+1)(\alpha-1)}{(\alpha^m-1)(\alpha+1)}. \quad (59)$$

Notice that $\frac{\alpha^m-1}{\alpha^m+1}\le\frac{\alpha^K-1}{\alpha^K+1}$ if $m\le K$. Thus it follows from (59) that

$$\frac{\|x_\Lambda\|_1}{\|x_\Lambda\|_2}\le\sqrt{\frac{(\alpha^m-1)(\alpha+1)}{(\alpha^m+1)(\alpha-1)}}\le\sqrt{\frac{(\alpha^K-1)(\alpha+1)}{(\alpha^K+1)(\alpha-1)}}=\sqrt{g_K(\alpha)}. \quad (60)$$

Finally, from (57) and (60) we obtain $R(x)\le\sqrt{g_K(\alpha)}$.

ACKNOWLEDGE

The authors thank the reviewers for their feedback, which has greatly improved the quality of this article.

REFERENCES

[1] 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.

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

[3] M. A. Davenport, M. Duarte, Y. C. Eldar, and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[4] G. Tang and A. Nehorai, "Performance analysis for sparse support recovery," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1383–1399, Mar. 2010.

[5] R. Wu, W. Huang, and D.-R. Chen, "The exact support recovery of sparse signals with noise via orthogonal matching pursuit," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 403–406, Apr. 2013.

[6] L.-H. Chang and J.-Y. Wu, "An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 707–710, 2014.

[7] Y.Shen and S.Li, "Sparse signals recovery from noisy measurements by orthogonal matching pursuit," *Inverse problems and Imaging*, vol. 9, no. 1, pp. 231-238, 2015.

[8] J. Wen, Z. Zhou, J. Wang, X. Tang, Q. Mo, "A Sharp Condition for Exact Support Recovery of Sparse Signals with Orthogonal Matching Pursuit," 2016, *arXiv: 1512.07248*, [Online]. Available: http://arxiv.org/pdf/1512.07248.pdf

[9] 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.

[10] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[11] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci. Paris*, Ser. I, vol. 346, no. 11, pp. 589–592, 2008.
[12] T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
[13] 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, 2010.
[14] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser, 2013.
[15] Q. Mo and S. Yi, "A remark on the restricted isometry property in orthogonal matching pursuit," IEEE Trans. Inf. Theory, vol. 58, no. 6, pp. 3654–3656, 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, 2012.
[17] Q. Mo, "A sharp restricted isometry constant bound of orthogonal matching pursuit," 2015, *arXiv: 1501.01708*, [Online]. Available: http://arxiv.org/pdf/1501.01708v1.pdf
[18] J. Wen, Z. Zhou, Z. Liu, et al. "Sharp Sufficient Conditions for Stable Recovery of Block Sparse Signals by Block Orthogonal Matching Pursuit," 2016, *arXiv: 1605.02894*, [Online]. Available: http://arxiv.org/pdf/1605.02894.pdf
[19] R. A. Horn and C. R. Johnson, "Positive definite and semidefinite matrices," in *Matrix Analysis*, 2nd ed., New York: Cambridge Univ. Press, 2013, pp. 425-515.
[20] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, New York: Springer-Verlag, 1996.

Chang Liu received the B.S. degree in School of Communication Engineering from Nanjing University of Information Science & Technology, Nanjing, China, in 2015. Now he is a M.S. candidate in School of Communication and Information Engineering of Shanghai University, Shanghai, China. His current research interest is the communication signal processing.

Yong Fang received the Ph.D. degree in Electronic Engineering from City University of Hong Kong, Hong Kong, in 1999. He is now a professor in the School of Communication and Information Engineering, Shanghai University, Shanghai, China. His research interests include communication signal processing, blind signal processing, and adaptive information system.

Jianzhong Liu received the B.S. degree in mathematics from Northwest University, Xi'an, China, in 1983. He is now a professor in the School of Mathematics and Physics, Jiangsu University of Technology, Changzhou, China. His research interests include matrix theory and multilinear algebra.