An Improved RIP-Based Performance Guarantee for Sparse Signal Recovery via Orthogonal Matching Pursuit

Ling-Hua Chang and Jwo-Yuh Wu

***Abstract*— A sufficient condition reported very recently for perfect recovery of a K-sparse vector via orthogonal matching pursuit (OMP) in K iterations (when there is no noise) is that the restricted isometry constant (RIC) of the sensing matrix satisfies $\delta_{K+1} < (1/\sqrt{K}+1)$. In the noisy case, this RIC upper bound along with a requirement on the minimal signal entry magnitude is known to guarantee exact support identification. In this paper, we show that, in the presence of noise, a relaxed RIC upper bound $\delta_{K+1} < (\sqrt{4K+1}-1/2K)$ together with a relaxed requirement on the minimal signal entry magnitude suffices to achieve perfect support identification using OMP. In the noiseless case, our result asserts that such a relaxed RIC upper bound can ensure exact support recovery in K iterations: this narrows the gap between the so far best known bound $\delta_{K+1} < (1/\sqrt{K}+1)$ and the ultimate performance guarantee $\delta_{K+1} = (1/\sqrt{K})$. Our approach relies on a newly established near orthogonality condition, characterized via the achievable angles between two orthogonal sparse vectors upon compression, and, thus, better exploits the knowledge about the geometry of the compressed space. The proposed near orthogonality condition can be also exploited to derive less restricted sufficient conditions for signal reconstruction in two other compressive sensing problems, namely, compressive domain interference cancellation and support identification via the subspace pursuit algorithm.**

***Index Terms*— Compressive sensing, interference cancellation, orthogonal matching pursuit, restricted isometry property (RIP), restricted isometry constant (RIC), subspace pursuit.**

I. Introduction

A. Overview

ORTHOGONAL matching pursuit (OMP) [1]–[4] is a popular greedy algorithm capable of recovering a K-sparse signal $\mathbf{x} \in \mathrm{R}^N$ based on a set of incomplete measurements $\mathbf{y} \in \mathrm{R}^M$ obeying the linear model

$$\mathbf{y} = \mathbf{\Phi x} + \mathbf{w}, \tag{1.1}$$

where $\mathbf{\Phi} \in \mathrm{R}^{M\times N}$ is the sensing matrix with $N \gg M(\gg K)$, and $\mathbf{w} \in \mathrm{R}^M$ is the measurement noise. Basically, OMP (outlined in Table I) is an iterative algorithm, whereby in each iteration an index of the signal support is identified. When noise is absent, i.e., $\mathbf{w} = \mathbf{0}$ in (1.1), the study of sufficient conditions for perfect signal recovery using OMP recently received considerable attention in the area of compressive sensing [2]–[4]. Various reports on the performance guarantee of OMP have been available, all of which are specified in terms of either the restricted isometry property (RIP) or the mutual coherence of the sensing matrix $\mathbf{\Phi}$ [4]–[11]. RIP-based analysis of an OMP-like algorithm, in which the square-error metric for support identification in each iteration is replaced by a general convex objective function, is considered in [12]. Notably, by allowing more than K iterations in OMP, the RIP-based recovery condition reported in [12] is similar to that in the ℓ_1-minimization algorithm [3] for both the noiseless and noisy cases; in addition, the recovery condition under noise does not rely on the magnitude of the signal coefficients. Reconstruction of a class of structured sparse signals modeled by trigonometric polynomials via OMP is addressed in [13] and [14]. A comprehensive review of greedy

TABLE I

OMP Algorithm. In Step 3.2, $\mathbf{e}_i$ Denotes the ith Unit Standard Vector of a Suitable Dimension; in Step 3.4, the Entries of $\mathbf{q}_{\mathbf{\Omega}^j} \in \mathrm{R}^j$ Are Those of the Updated $\mathbf{q}$ Corresponding to the Index Set $\mathbf{\Omega}^j$

1. Input: $\mathbf{y}$, $\mathbf{\Phi}$
2. Initialize: $j=0$, $\mathbf{r}^0 := \mathbf{y} = \mathbf{\Phi x}$, $\mathbf{\Omega}^0 := [\ ]$ and $\mathbf{\Phi}_{\mathbf{\Omega}^0} = [\ ]$
3. While $\left\|\mathbf{r}^j\right\|_2 > \varepsilon_1$ ($\left\|\mathbf{\Phi}^*\mathbf{r}^j\right\|_\infty > \varepsilon_2$ for l_∞ -bounded noise case) 3.1 $j = j+1$ 3.2 $\rho^j = \arg\max_{i=1,\dots,N} \left|\left\langle \mathbf{\Phi e}_i, \mathbf{r}^{j-1}\right\rangle\right|$ and $\mathbf{\Omega}^j = \mathbf{\Omega}^{j-1} \cup \rho^j$ 3.3 $\mathbf{q} = [0\ 0\ \cdots\ 0]^* \in \mathbb{R}^N$ 3.4 $\mathbf{q}_{\Omega^j} = \arg\min_{\mathbf{b}} \left\|\mathbf{\Phi}_{\Omega^j}\mathbf{b} - \mathbf{y}\right\|_2 = (\mathbf{\Phi}^*_{\Omega^j}\mathbf{\Phi}_{\Omega^j})^{-1}\mathbf{\Phi}^*_{\Omega^j}\mathbf{y}$ 3.5 $\mathbf{r}^j = \mathbf{y} - \mathbf{\Phi}_{\Omega^j}\mathbf{q}_{\Omega^j}$ end while
4. Output: $\hat{\mathbf{x}} = \mathbf{q}$

Manuscript received March 20, 2013; revised April 20, 2014; accepted June 13, 2014. Date of publication July 11, 2014; date of current version August 14, 2014. This work was supported by the National Science Council of Taiwan under Grant NSC-102-2221-E-009-019-MY3, Grant NSC-100-2221-E-009-104-MY3, and Grant MOST 103-2221-E-009-050-MY3, in part by the Ministry of Education of Taiwan through the ATU Program, and in part by the Telecommunication Laboratories, Chunghwa Telecom Company, Ltd., Taipei, Taiwan, under Grant TL-99-G107. This paper was presented at the 6th International Symposium on Communications, Control, and Signal Processing, in 2014, and the 8th IEEE Sensor Array and Multichannel Signal Processing Workshop in 2014.

The authors are with the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: iamjaung@gmail.com; jywu@cc.nctu.edu.tw).

Communicated by O. Milenkovic, Associate Editor for Coding Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2014.2338314

algorithms for sparse signal recovery, as well as the related analytical performance guarantees, can be found in [3, Ch. 8].

B. RIP-Based Performance Guarantee: Existing Results

The sensing matrix $\mathbf{\Phi}$ is said to satisfy RIP of order K [15]–[17] if there exists $0 < \delta_K < 1$ such that

$$(1-\delta_K)\,\|\mathbf{s}\|_2^2 \le \|\mathbf{\Phi s}\|_2^2 \le (1+\delta_K)\,\|\mathbf{s}\|_2^2 \tag{1.2}$$

holds for all K-sparse $\mathbf{s}$. The constant δ_K is the so-called restricted isometry constant (RIC) of the sensing matrix $\mathbf{\Phi}$. Under the noiseless environment, i.e. $\mathbf{w} = \mathbf{0}$ in (1.1), Davenport and Wakin [5] showed that the OMP algorithm can exactly identify the support of a K-sparse signal in K iterations if $\mathbf{\Phi}$ satisfies RIP of order $K+1$ with RIC $\delta_{K+1} < 1/(3\sqrt{K})$. Hung and Zhu [6] then derived the less restricted sufficient condition $\delta_{K+1} < 1/(1+\sqrt{2K})$. As the latest report, Mo and Shen [7], and Wang and Shim [8], both proved that the upper bound on δ_{K+1} can be further relaxed to

$$\delta_{K+1} < \frac{1}{\sqrt{K}+1}. \tag{1.3}$$

In [7] and [8], examples are also given to illustrate the failure of exact support identification in K iterations in case that

$$\delta_{K+1} = \frac{1}{\sqrt{K}}. \tag{1.4}$$

Such results also verified the conjecture made by Dai and Milenkovic in [9], viz., values of δ_{K+1} no less than $1/\sqrt{K}$ may result in the failure of perfect support recovery. RIP-based performance analysis for OMP in the noisy case has also been considered in [18] and [19]. Specifically, to guarantee exact support identification under measurement noise, sufficient conditions specified in terms of the RIC bound (1.3) along with certain requirements on the minimal signal entry magnitude have been derived in [18].

C. Paper Contribution

In this paper, we derive improved performance guarantees for perfect support recovery via OMP. Under either the ℓ_2 - or ℓ_∞ -bounded noise model, we show that exact support identification is guaranteed if the RIC δ_{K+1} satisfies

$$\delta_{K+1} < \frac{\sqrt{4K+1}-1}{2K}, \tag{1.5}$$

and the minimal amplitude of the signal entries exceeds a certain lower bound. Since $\frac{1}{\sqrt{K}+1} < \frac{\sqrt{4K+1}-1}{2K}$ and the derived lower bound for the signal entry amplitudes for either noise model is smaller than that given in [18], our sufficient conditions are less conservative. In the noiseless case, our result asserts that OMP can perfectly identify the support in K iterations if the RIC δ_{K+1} satisfies (1.5). Since $\frac{1}{\sqrt{K}+1} < \frac{\sqrt{4K+1}-1}{2K} < \frac{1}{\sqrt{K}}$, the proposed bound (1.5) narrows the gap between the so far best known bound (1.3) and the ultimate performance guarantee (1.4). Our proofs exploit a newly established near-orthogonality property, which is characterized via achievable angles between two orthogonal sparse vectors upon compression. Hence, the improved performance guarantee benefits from more explicit geometric insights of compressed sparse vectors under the RIP of the sensing matrix. The developed near-orthogonality condition as well as the proof techniques has a far-reaching impact: it finds applications in RIP-based performance analyses in two other signal reconstruction problems in compressive sensing. Specifically,

- *Compressive-Domain Interference Cancellation via Orthogonal Projection [20]–[24]:* In this problem, a central issue regarding the study of the signal reconstruction performance upon interference removal is to specify the RIC of the effective sensing matrix, which is a product of an orthogonal projection matrix and a random sensing matrix [20]–[24]. Based on the developed analysis techniques, we derive a more accurate estimate of the RIC of the effective sensing matrix as compared to the previous works [20] and [23].
- *Support Identification via Subspace Pursuit (SP) [9]:* SP is another popular greedy algorithm for sparse signal recovery in compressive sensing [9], and RIP-based performance guarantee for SP has been investigated in [9] and [10]. By using the proposed approach (in particular, the approximate orthogonality condition), we show in this paper that, to guarantee perfect/stable signal reconstruction via SP, the requirement on the RIC of the sensing matrix $\mathbf{\Phi}$ can be relaxed even further. Specifically, assuming that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$, it is shown that $\delta_{3K} \le 0.2412$ suffices to guarantee exact (stable, respectively) support identification via SP in the noiseless (noisy, respectively) case. Our bound thus improves the results in [9] and [10]: in the absence of noise, the requirement on RIC reported in [9] is $\delta_{3K} < 0.165$; when noise is present, the required bound shown in [9] is $\delta_{3K} < 0.083$, and in [10] is then pushed to $\delta_{3K} < 0.139$.

The organization of this paper is as follows. Section II derives improved performance guarantees for OMP in both noisy and noiseless cases. Section III further investigates the applications of the proposed proof techniques in the study of two other signal reconstruction problems, namely, compressive-domain interference cancellation and signal recovery via SP. For the former, a more accurate estimate of the RIC of the effective sensing matrix upon interference removal is derived; for the later, less conservative sufficient conditions for signal reconstruction are developed. Some concluding remarks are then drawn in Section IV. To ease reading, most of the technical proof is relegated to the appendix.

Notation List: For $S \subset \{1,\ldots,N\}$ with cardinality $|S|$, we use $\mathbf{\Phi}_S \in \mathrm{R}^{M\times|S|}$ to denote the matrix obtained from $\mathbf{\Phi} \in \mathrm{R}^{M\times N}$ by retaining its columns indexed by the subset S. For $\mathbf{u} \in \mathrm{R}^N$ with $(\mathbf{u})_i$ as the ith entry, $\mathbf{u}_S \in \mathrm{R}^{|S|}$ denotes the vector whose entries consist of those of $\mathbf{u}$ indexed by S; we write $\tilde{\mathbf{u}}_S \in \mathrm{R}^N$ to be the zero-padded version of $\mathbf{u}_S$ such that $(\tilde{\mathbf{u}}_S)_i = (\mathbf{u})_i$ for $i \in S$ and $(\tilde{\mathbf{u}}_S)_i = 0$ otherwise (thus, $\tilde{\mathbf{u}}_S$ is $|S|$-sparse with support S). Throughout the paper, $\mathbf{e}_i \in \mathrm{R}^N$ denotes the ith unit standard vector, $\mathbf{I}$ denotes the identity matrix of a proper dimension, $\mathbf{0}$ represents the zero vector of a proper dimension, and $()^*$ stands for the transpose operation.

$\mathcal{R}(\mathbf{M})$ denotes the column space of the matrix $\mathbf{M}$. $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denote, respectively, the vector two-norm and sup-norm [25]; $\langle \mathbf{u}, \mathbf{v}\rangle$ represents the standard inner product between the two vectors $\mathbf{u}$ and $\mathbf{v}$.

II. Improved Performance Guarantee for OMP

This section presents the derived improved RIP-based performance guarantees for OMP in both the noiseless and noisy cases. In the sequel, we denote the support of the desired K-sparse vector $\mathbf{x}$ by T, with cardinality $|T| = K$. In addition, as in various previous works [8], [10], and [11], it is assumed that all columns of the sensing matrix $\mathbf{\Phi}$ are normalized to be of unit-norm. Section II-A first shows the near-orthogonality property, which is the foundation for our analyses. Section II-B then presents the main results. Finally, some discussions are given in Section II-C.

A. Near Orthogonality Property

The development of the improved performance guarantees relies crucially on the next lemma, which characterizes the achievable angle between two compressed orthogonal sparse vectors in terms of the RIC of the sensing matrix.

Lemma 2.1 (Near-Orthogonality Property): Let $\mathbf{u}$ and $\mathbf{v}$ be two orthogonal sparse vectors with supports T_u and T_v fulfilling $|T_u \cup T_v| \le K$. Suppose that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order K with RIC δ_K. Then we have

$$|\cos \measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})| \le \delta_K, \tag{2.1}$$

where $\measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})$ denotes the angle between $\mathbf{\Phi u}$ and $\mathbf{\Phi v}$.

Proof: See Appendix A. □

Lemma 2.1 asserts (in terms of the achievable angle) that, for a small δ_K, the compressed vectors $\mathbf{\Phi u}$ and $\mathbf{\Phi v}$ are nearly orthogonal as long as $\mathbf{u}$ and $\mathbf{v}$ are orthogonal. Notably, under the same assumptions made as in Lemma 2.1 and based on a plane-geometry analysis, the following upper bound on $|\cos \measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})|$ has been derived in [24, Corollary 5.3]:

$$|\cos \measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})| \le \frac{\delta_K}{\sqrt{1-\delta_K^2}}. \tag{2.2}$$

It can be seen that the proposed bound (2.1), which exploits a geometric interpretation of the two-norm condition number (details referred to Appendix A), improves the result (2.2). An alternative characterization of orthogonality is via the inner product between $\mathbf{\Phi u}$ and $\mathbf{\Phi v}$ [17]; more precisely, for $\mathbf{u}$ and $\mathbf{v}$ with non-overlapping supports (thus, $\mathbf{u}$ and $\mathbf{v}$ are orthogonal), it has been shown in [17] that

$$|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle| \le \delta_K \|\mathbf{u}\|_2 \cdot \|\mathbf{v}\|_2 . \tag{2.3}$$

It is worthy of noting that (2.3) in conjunction with the RIP condition (1.2) can be directly used to derive an upper bound of $|\cos \measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})|$, as can be seen by

$$\begin{aligned} |\cos \measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})| &= \frac{|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle|}{\|\mathbf{\Phi u}\|_2 \cdot \|\mathbf{\Phi v}\|_2} \\ &\overset{(a)}{\le} \frac{\delta_K \|\mathbf{u}\|_2 \cdot \|\mathbf{v}\|_2}{\|\mathbf{\Phi u}\|_2 \cdot \|\mathbf{\Phi v}\|_2} \overset{(b)}{\le} \frac{\delta_K \|\mathbf{u}\|_2 \cdot \|\mathbf{v}\|_2}{(1-\delta_K)\|\mathbf{u}\|_2 \cdot \|\mathbf{v}\|_2} \\ &= \frac{\delta_K}{1-\delta_K}, \end{aligned} \tag{2.4}$$

where (a) follows from (2.3) and (b) holds due to the RIP (1.2). The upper bound (2.4) derived by using the simple algebraic approach shown above is even worse than (2.2); this is not unexpected since algebraic analyses using RIP are known to yield the worst-case estimate [26], [27]. To summarize, Lemma 2.1 asserts that, when $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, the measure of orthogonality between $\mathbf{\Phi u}$ and $\mathbf{\Phi v}$ in terms of the achievable angle $\measuredangle (\mathbf{\Phi u}, \mathbf{\Phi v})$ can be improved as compared to the previous results given by (2.2) and (2.4). We would like to mention that the "near orthogonality" condition (either in the form of (2.1) or (2.3)) plays a fundamental role in the study of the signal reconstruction performance in compressive sensing [17], [21], [24]. Thanks to the improved bound (2.1), less restricted sufficient conditions for signal reconstruction via OMP can be obtained in the next subsection; the impacts of (2.1) on other two signal reconstruction problems will be studied in Section III.

B. Main Results

To guarantee exact support identification via OMP in the presence of noise,[1] sufficient conditions specified in terms of the RIC δ_{K+1} and certain lower bounds for the signal entry amplitudes have been reported in [18], and are summarized in the next proposition. Throughout this section, $\mathbf{r}^j$ denotes the residual vector in the j-th iteration of the OMP algorithm (cf. Step 3.5 in Table I).

Proposition 2.2 [18]: Consider the signal model (1.1). Then the following results hold.

(1). (ℓ_2-bounded noise) Under $\|\mathbf{w}\|_2 \le \varepsilon_1$, OMP with the stopping criterion $\left\|\mathbf{r}^j\right\|_2 \le \varepsilon_1$ can exactly identify the support T of the K-sparse signal $\mathbf{x}$ if $\delta_{K+1} < 1/(\sqrt{K}+1)$ and the minimum magnitude of the nonzero entries of $\mathbf{x}$ satisfies

$$\min_{i\in T} |(\mathbf{x})_i| > \frac{\left(\sqrt{1+\delta_{K+1}}+1\right)\varepsilon_1}{1-\delta_{K+1}-\sqrt{K}\delta_{K+1}}. \tag{2.5}$$

(2). (ℓ_∞-bounded noise) Under $\left\|\mathbf{\Phi}^*\mathbf{w}\right\|_\infty \le \varepsilon_2$, OMP with the stopping criterion $\left\|\mathbf{\Phi}^*\mathbf{r}^j\right\|_\infty \le \varepsilon_2$ can exactly identify the support T of the K-sparse signal $\mathbf{x}$ if $\delta_{K+1} < 1/(\sqrt{K}+1)$ and the minimum magnitude of the nonzero entries of $\mathbf{x}$ satisfies

$$\min_{i\in T} |(\mathbf{x})_i| > \frac{\left(\sqrt{K}+\sqrt{K}\sqrt{1+\delta_{K+1}}\right)\varepsilon_2}{1-\delta_{K+1}-\sqrt{K}\delta_{K+1}}. \tag{2.6}$$

□

By exploiting the near-orthogonality condition established in Lemma 2.1, less conservative sufficient conditions for exact support identification via OMP with noise corruption are obtained below. For this, we need the following two technical lemmas.

Lemma 2.3: Under the same assumptions as in Lemma 2.1, we have

$$|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle| \le \delta_K \|\mathbf{\Phi u}\|_2 \cdot \|\mathbf{\Phi v}\|_2 . \tag{2.7}$$

[1] In our study, we focus on OMP with a residual based stopping criterion, as in [18]. Notably, there have been several different implementations of OMP, e.g., by setting a fixed number of iterations to halt the algorithm, see [28] and [29].

Proof: The result follows from $|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle| = \|\mathbf{\Phi u}\|_2 \cdot \|\mathbf{\Phi v}\|_2 \cdot |\cos(\mathbf{\Phi u}, \mathbf{\Phi v})|$ and using (2.1). □

Lemma 2.4: Let $\mathbf{x}$ be a K-sparse vector with support T. Assume that the sensing matrix $\mathbf{\Phi} \in \mathrm{R}^{M\times N}$ satisfies RIP of order $K+1$ with RIC δ_{K+1}. Then $\left\|\mathbf{\Phi}_T^* \mathbf{\Phi x}\right\|_2 \geq \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi x}\|_2 \geq (1-\delta_{K+1})\,\|\mathbf{x}\|_2$, or equivalently,

$$\left\|\mathbf{\Phi}_T^* \mathbf{\Phi}_T \mathbf{x}_T\right\|_2 \geq \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}_T \mathbf{x}_T\|_2 \geq (1-\delta_{K+1})\,\|\mathbf{x}_T\|_2\,. \tag{2.8}$$

Proof: See Appendix B. □

We note that the inequality $\left\|\mathbf{\Phi}_T^* \mathbf{\Phi}_T \mathbf{x}_T\right\|_2 \geq (1-\delta_{K+1})\,\|\mathbf{x}_T\|_2$, which provides an *ambient-domain* lower bound for $\left\|\mathbf{\Phi}_T^* \mathbf{\Phi}_T \mathbf{x}_T\right\|_2$, has been derived in, see [30, Proposition 3.1]. The significance of Lemma 2.4 lies in that a tighter lower bound for $\left\|\mathbf{\Phi}_T^* \mathbf{\Phi}_T \mathbf{x}_T\right\|_2$ is provably to be $\sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}_T \mathbf{x}_T\|_2$ (*a compressed-domain bound*). Based on Lemmas 2.3 and 2.4, the main results of this section are given in the next theorem.

Theorem 2.5: Consider the signal model (1.1). Then the following results hold.

(1). (ℓ_2-bounded noise) Under $\|\mathbf{w}\|_2 \leq \varepsilon_1$, OMP with the stopping criterion $\left\|\mathbf{r}^j\right\|_2 \leq \varepsilon_1$ can exactly identify the support T of the K-sparse signal $\mathbf{x}$ if the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $K+1$ with RIC δ_{K+1} fulfilling

$$\delta_{K+1} < \frac{\sqrt{4K+1}-1}{2K}, \tag{2.9}$$

and the minimum magnitude of the nonzero entries of $\mathbf{x}$ satisfies

$$\min_{i\in T} |(\mathbf{x})_i| > \frac{(\sqrt{1+\delta_{K+1}}+1)\varepsilon_1}{1-\delta_{K+1}-\sqrt{1-\delta_{K+1}}\sqrt{K}\delta_{K+1}}. \tag{2.10}$$

(2). (ℓ_∞-bounded noise) Under $\left\|\mathbf{\Phi}^*\mathbf{w}\right\|_\infty \leq \varepsilon_2$, OMP with the stopping criterion $\left\|\mathbf{\Phi}^*\mathbf{r}^j\right\|_\infty \leq \varepsilon_2$ can exactly identify the support T of the K-sparse signal $\mathbf{x}$ if $\mathbf{\Phi}$ satisfies RIP of order $K+1$ with RIC δ_{K+1} fulfilling (2.9), and the minimum magnitude of the nonzero elements of $\mathbf{x}$ satisfies

$$\min_{i\in T} |(\mathbf{x})_i| > \frac{(\sqrt{K}+1)\varepsilon_2}{1-\delta_{K+1}-\sqrt{1-\delta_{K+1}}\sqrt{K}\delta_{K+1}}. \tag{2.11}$$

Proof: See Appendix C. □

For either noise model, it can be readily checked that the derived lower bound (i.e., (2.10) or (2.11)) is smaller than that shown in [18] (i.e., (2.5) or (2.6)). Our results thus assert that, to guarantee exact support identification via OMP in the noisy environment, the requirements on the RIC and the strength of the signal components can be further relaxed as compared to that reported in [18]. It is worthy of noting that, in the noiseless case, Theorem 2.5 leads to the following corollary.

Corollary 2.6: Consider the noiseless case, i.e., $\mathbf{w} = \mathbf{0}$ in (1.1). Assume that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $K+1$ with RIC fulfilling (2.9). Then, for any K-sparse $\mathbf{x}$, OMP can perfectly identify the support of $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x}$ in K iterations.

Proof: See also Appendix C. □

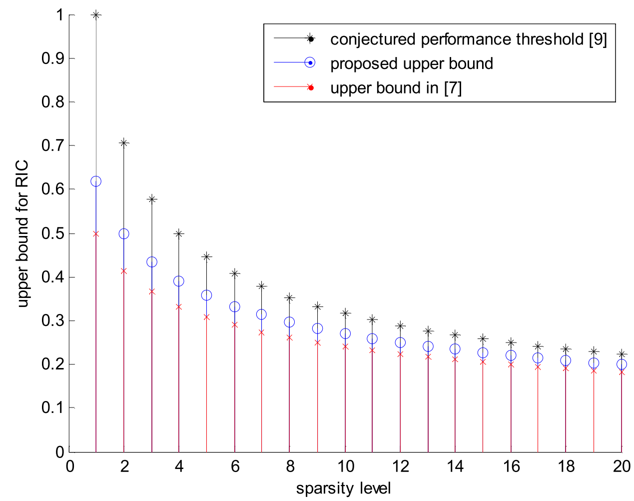

Fig. 1. Comparison of three bounds (i.e., (1.3), (1.4), and (2.9)) on δ_{K+1} for different sparsity levels K.

C. Discussions

Some discussions regarding the improved performance guarantees derived in the previous subsection are given below.

1. It is noted that most existing RIP-based analyses for OMP rely on analyzing $\left|\langle \mathbf{\Phi e}_i, \mathbf{r}^j\rangle\right|$, namely, the magnitude of the correlation between columns of the sensing matrix $\mathbf{\Phi}$ and the residual vector $\mathbf{r}^j$ (see [7], [8], [18], and our paper). A backbone of such analyses consists in an upper bound of $|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle|$, where $\mathbf{u}$ and $\mathbf{v}$ are sparse with disjoint supports (see [7, Proof of Lemma 2.1], [18, Lemma 3 and eq. (16)], and eq. (C.5) in Appendix C of our paper). It is worthy of further noting that both $\left|\langle \mathbf{\Phi e}_i, \mathbf{r}^j\rangle\right|$ and $|\langle \mathbf{\Phi u}, \mathbf{\Phi v}\rangle|$ are entirely about inner products in the *compressed domain*, and this to a large extent reflects the underlying fact: Since measurement processing for signal reconstruction is done in the compressed domain, the geometry of the compressed space is important. With all the above in mind, a fundamental difference between the technical approaches used in the previous works ([7], [8], [18]) and our study can be seen as follows. In [7], [8], and [18], the upper bound given in (2.3), developed by Candes [17], is employed. Notably, the bound in (2.3) is an ambient-domain bound, since it is specified in terms of $\|\mathbf{u}\|_2$ and $\|\mathbf{v}\|_2$, the norms of $\mathbf{u}$ and $\mathbf{v}$ in the ambient signal domain. By contrast, the proposed approach relies on the compressed-domain bound (2.7) (it is completely specified in terms of $\|\mathbf{\Phi u}\|_2$ and $\|\mathbf{\Phi v}\|_2$, the norms of $\mathbf{u}$ and $\mathbf{v}$ upon compression). With the aid of such compressed-domain bounds ((2.7) as well as (2.8)), our approach can better exploit the geometry of the compressed space: this overall leads to the improved performance guarantees derived in Theorem 2.5.
2. Consider the noiseless case. Since $\frac{1}{\sqrt{K}+1} < \frac{\sqrt{4K+1}-1}{2K} < \frac{1}{\sqrt{K}}$, the proposed RIC bound (2.9) thus improves the best-known result $\frac{1}{\sqrt{K}+1}$ reported in [7] and [8]. Figure 1 shows the three bounds (namely, (1.3), (1.4), and (2.9))

for various sparsity levels K. As can be seen from the figure, the improvement over $\frac{1}{\sqrt{K+1}}$ is slight when K is large. This is not unexpected since, for large K, the gap between $\frac{1}{\sqrt{K+1}}$ and the ultimate performance guarantee $\frac{1}{\sqrt{K}}$ is pretty small, and it is therefore rather difficult to achieve a substantial improvement.

3. For the Gaussian noise case, $\|\mathbf{w}\|_2$ can be bounded from above with a sufficiently high probability. Hence, under an additional probability constraint, a relaxed lower bound for the minimal signal entry magnitude for the Gaussian case can also be obtained by following essentially the same procedures as in the ℓ_2-bounded noise case.

III. Impacts of the Approximate Orthogonality Condition (2.1)

As mentioned above, the approximate orthogonality condition (2.1) (measured in terms of achievable angles) not only enjoys its own technical novelty but also has a wide spectrum of applications: it can be used for developing improved RIP-based performance characterizations for other sparse signal reconstruction schemes. Below we discuss two such applications in details.

A. Sparse Signal Recovery Against Sparse Interference via Orthogonal Projection

Consider the following compressive sensing system under sparse interference corruption [20], [23]

$$\mathbf{y} = \mathbf{\Phi}(\mathbf{x} + \mathbf{d}) = \mathbf{\Phi}\mathbf{x} + \mathbf{\Phi}\mathbf{d}, \tag{3.1}$$

where $\mathbf{d} \in \mathrm{R}^N$ is a sparse interference/disturbance with support T_d. As in various previous works [20]–[24], it is assumed that T_d is known and does not overlap with the signal support T. To exploit the knowledge of T_d for interference removal, one typical approach is via orthogonal projection. More specifically, the measurement $\mathbf{y}$ is projected onto the orthogonal complement of the interference subspace $\mathcal{R}(\mathbf{\Phi}_{T_d})$ to obtain [20]–[24]

$$\bar{\mathbf{y}} \triangleq \mathbf{Q}\mathbf{y} = \mathbf{Q}\mathbf{\Phi}\mathbf{x} + \mathbf{Q}\mathbf{\Phi}\mathbf{d} = \mathbf{Q}\mathbf{\Phi}\mathbf{x}, \tag{3.2}$$

where the projection matrix $\mathbf{Q} \triangleq \mathbf{I} - \mathbf{\Phi}_{T_d}(\mathbf{\Phi}_{T_d}^*\mathbf{\Phi}_{T_d})^{-1}\mathbf{\Phi}_{T_d}^*$ removes all the components of $\mathbf{y}$ lying in $\mathcal{R}(\mathbf{\Phi}_{T_d})$. Upon interference removal, the effective sensing matrix in (3.2) is $\mathbf{Q}\mathbf{\Phi}$, which is a product of an orthogonal projection matrix $\mathbf{Q}$ and the original random sensing matrix $\mathbf{\Phi}$. The performance of sparse signal reconstruction based on (3.2) depends crucially on the RIC of $\mathbf{Q}\mathbf{\Phi}$. The RIP of $\mathbf{Q}\mathbf{\Phi}$ as well as the achievable RIC was first studied in [20]; the results are summarized in the next proposition.

Proposition 3.1 [20]: Consider the system (3.2). Assume that $\mathbf{\Phi}$ satisfies the RIP of order K with RIC given by δ_K, and that the interference support T_d satisfies $|T_d| < K$. The following inequality holds for all $(K - |T_d|)$-sparse $\mathbf{x}$ whose support does not overlap with T_d:

$$\left(1 - \bar{\delta}_A\right)\|\mathbf{x}\|_2^2 \le \|\mathbf{Q}\mathbf{\Phi}\mathbf{x}\|_2^2 \le (1 + \delta_K)\|\mathbf{x}\|_2^2, \tag{3.3}$$

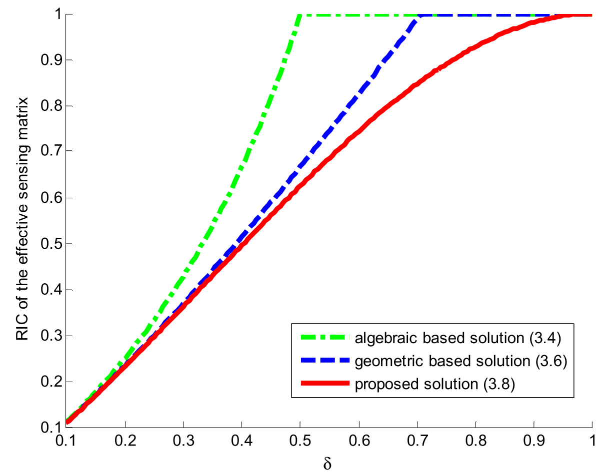

Fig. 2. RIC of $\mathbf{Q}\mathbf{\Phi}$ for different values of δ, the RIC of $\mathbf{\Phi}$.

where

$$\bar{\delta}_A \triangleq \min\left\{1, \frac{\delta_K}{1 - \delta_K}\right\}. \tag{3.4}$$

□

Proposition 3.1 asserts that $\mathbf{Q}\mathbf{\Phi}$ still enjoys RIP, but is with an RIC $\bar{\delta}_A$ larger than δ_K. In [23] and [24], an improved estimate of RIC of $\mathbf{Q}\mathbf{\Phi}$ was obtained by means of certain plane geometry analyses, as asserted in the next proposition.

Proposition 3.2 [23], [24]: Under the same assumptions as in Proposition 3.1, the following inequality holds for all $(K - |T_d|)$-sparse $\mathbf{x}$ whose support does not overlap with T_d:

$$\left(1 - \bar{\delta}_G\right)\|\mathbf{x}\|_2^2 \le \|\mathbf{Q}\mathbf{\Phi}\mathbf{x}\|_2^2 \le (1 + \delta_K)\|\mathbf{x}\|_2^2, \tag{3.5}$$

where

$$\bar{\delta}_G \min\left\{1, \delta_K + \frac{\delta_K^2}{1 + \delta_K}\right\}. \tag{3.6}$$

□

By leveraging Lemma 2.1, the following theorem shows that the estimate of the RIC of $\mathbf{Q}\mathbf{\Phi}$ can be improved even further.

Theorem 3.3: Under the same assumptions as in Proposition 3.1, the following inequality holds for all $(K - |T_d|)$-sparse $\mathbf{x}$ whose support does not overlap with T_d:

$$\left(1 - \bar{\delta}\right)\|\mathbf{x}\|_2^2 \le \|\mathbf{Q}\mathbf{\Phi}\mathbf{x}\|_2^2 \le (1 + \delta_K)\|\mathbf{x}\|_2^2, \tag{3.7}$$

where

$$\bar{\delta} \triangleq \min\left\{1, \delta_K + \delta_K^2(1 - \delta_K)\right\}. \tag{3.8}$$

Proof: See Appendix D. □

It is easy to verify that $\bar{\delta} < \bar{\delta}_G < \bar{\delta}_A$, viz., the proposed solution (3.8) improves the previous estimates in [20], [23], and [24] (this is also confirmed by Figure 2, which plots the three estimated RIC of the effective sensing matrix $\mathbf{Q}\mathbf{\Phi}$ with respect to different values of δ_K, the RIC of the random sensing matrix $\mathbf{\Phi}$). Since a smaller RIC results in a better signal reconstruction performance [3], our result implies that the achievable performance of sparse signal recovery with interference-nulling is actually better than as predicted by the previous works [20], [23], and [24].

TABLE II
THE SP ALGORITHM

1. Input: $\mathbf{y}$, $\mathbf{\Phi}$
Initialize:
2.1 $\mathbf{r}^0 = \mathbf{y}$
2.2 $\mathbf{\Omega}^1 = \{K$ indexes yielding the largest magnitude entries of $\mathbf{\Phi}^*\mathbf{r}^0\}$
2.3 $\mathbf{q} = [0\ 0\ \cdots\ 0]^* \in \mathbb{R}^N$
2.4 $\mathbf{q}_{\Omega^1} = \arg\min_{\mathbf{b}} \left\|\mathbf{\Phi}_{\Omega^1}\mathbf{b} - \mathbf{y}\right\|_2 = (\mathbf{\Phi}^*_{\Omega^1}\mathbf{\Phi}_{\Omega^1})^{-1}\mathbf{\Phi}^*_{\Omega^1}\mathbf{y}$
2.5 $\mathbf{r}^1 := \mathbf{y} - \mathbf{\Phi}_{\Omega^1}\mathbf{q}_{\Omega^1}$
2.6 $j = 1$
3. While $\left\|\mathbf{r}^{j-1}\right\|_2 < \left\|\mathbf{r}^j\right\|_2$
3.0 $j = j + 1$
3.1 $\mathbf{\Omega}^j_\Delta = \{K$ indexes yielding the largest magnitude entries of $\mathbf{\Phi}^*\mathbf{r}^{j-1}\}$
3.2 $\breve{\mathbf{\Omega}}^j \triangleq \mathbf{\Omega}^j_\Delta \cup \mathbf{\Omega}^{j-1}$
3.3 $\breve{\mathbf{q}} = [0\ 0\ \cdots\ 0]^* \in \mathbb{R}^N$
3.4 $\breve{\mathbf{q}}_{\breve{\Omega}^j} = \arg\min_{\mathbf{b}} \left\|\mathbf{\Phi}_{\breve{\Omega}^j}\mathbf{b} - \mathbf{y}\right\|_2 = (\mathbf{\Phi}^*_{\breve{\Omega}^j}\mathbf{\Phi}_{\breve{\Omega}^j})^{-1}\mathbf{\Phi}^*_{\breve{\Omega}^j}\mathbf{y}$
3.5 $\mathbf{\Omega}^j = \{K$ indexes yielding the largest magnitude entries of $\breve{\mathbf{q}}\}$
3.6 $\mathbf{q} = [0\ 0\ \cdots\ 0]^* \in \mathbb{R}^N$
3.7 $\mathbf{q}_{\Omega^j} = \arg\min_{\mathbf{b}} \left\|\mathbf{\Phi}_{\Omega^j}\mathbf{b} - \mathbf{y}\right\|_2 = (\mathbf{\Phi}^*_{\Omega^j}\mathbf{\Phi}_{\Omega^j})^{-1}\mathbf{\Phi}^*_{\Omega^j}\mathbf{y}$
3.8 $\mathbf{r}^j \triangleq \mathbf{y} - \mathbf{\Phi}_{\Omega^j}\mathbf{q}_{\Omega^j} = \mathbf{y} - \mathbf{P}_{\Omega^j}\mathbf{y}$, where $\mathbf{P}_{\Omega^j} \triangleq \mathbf{\Phi}_{\Omega^j}(\mathbf{\Phi}^*_{\Omega^j}\mathbf{\Phi}_{\Omega^j})^{-1}\mathbf{\Phi}^*_{\Omega^j}$
end while
4. Output: $\hat{\mathbf{x}} = \mathbf{q}$

B. Support Identification via SP

SP is another popular greedy algorithm for sparse signal recovery in the area of compressive sensing [9]; see Table II for an outline of the algorithm. In each iteration, SP tries to keep track of an estimated support consisting of K elements by adding and then removing certain elements to and from the candidate set. RIP-based performance guarantees for SP, in both noiseless and noisy cases, have been reported in [9] and [10]. The following proposition summarizes the result in [9] when noise is absent. In the sequel, $\mathbf{r}^j$ denotes the residual vector in the j-th iteration of the SP algorithm (cf. Step 3.8 in Table II).

Proposition 3.4 [9]: Assume that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$ with RIC

$$\delta_{3K} < 0.165. \tag{3.9}$$

Then, for any K-sparse $\mathbf{x}$, the SP algorithm with stopping criterion $\left\|\mathbf{r}^j\right\|_2 \geq \left\|\mathbf{r}^{j-1}\right\|_2$ can perfectly identify the support of $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x}$ via a finite number of iterations. □

By means of Lemma 2.1, an improved result (in terms of a less strict requirement on the RIC of the sensing matrix) is derived in the next theorem.

Theorem 3.5: Assume that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$ with RIC

$$\delta_{3K} \leq 0.2412. \tag{3.10}$$

Then, for any K-sparse $\mathbf{x}$, the SP algorithm with stopping criterion $\left\|\mathbf{r}^j\right\|_2 \geq \left\|\mathbf{r}^{j-1}\right\|_2$ can perfectly identify the support of $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x}$ via a finite number of iterations.

Proof: See Appendix E. □

When noise is present, SP is capable of achieving stable signal reconstruction, in the sense that, if the sensing matrix satisfies RIP with a small RIC, the reconstruction error is bounded and the size does not exceed a constant times the noise level. The following proposition, which was established in [9], makes this point precise.

Proposition 3.6 [9]: Assume that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$ with RIC

$$\delta_{3K} < 0.083. \tag{3.11}$$

Then, the SP algorithm reconstructs the K-sparse vector $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x} + \mathbf{w}$ with the reconstruction error bounded as

$$\left\|\mathbf{x} - \hat{\mathbf{x}}\right\|_2 \leq c'_K \left\|\mathbf{w}\right\|_2, \text{ with } c'_K \triangleq \frac{1 + \delta_{3K} + \delta_{3K}^2}{\delta_{3K}(1 - \delta_{3K})}, \tag{3.12}$$

where $\hat{\mathbf{x}}$ is the estimated sparse signal vector. □

By using the variation of the proof of [9, Th. 10], an improved performance guarantee has been derived in [10], and is given in the next proposition.

Proposition 3.7 [10]: Assume that the sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$ with RIC

$$\delta_{3K} < 0.139. \tag{3.13}$$

Then, the SP algorithm reconstructs the K-sparse vector $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x} + \mathbf{w}$ with the reconstruction error bounded as

$$\left\|\mathbf{x} - \hat{\mathbf{x}}\right\|_2 \leq \bar{c}_K \left\|\mathbf{\Phi}^*_{T_e}\mathbf{w}\right\|_2, \text{ with } \bar{c}_K \triangleq 2\left(\frac{7 - 9\delta_{3K} + 7\delta_{3K}^2 - \delta_{3K}^3}{(1 - \delta_{3K})^4}\right), \tag{3.14}$$

where $\hat{\mathbf{x}}$ is the estimated sparse signal vector and $T_e \triangleq \arg\max_{S \text{ with } |S| = K} \left\|\mathbf{\Phi}^*_S\mathbf{w}\right\|_2$. □

By exploiting the approximate orthogonality property shown in Lemma 2.1, we can obtain a less conservative sufficient condition for guaranteeing stable signal reconstruction as well as a tighter reconstruction error bound. Specifically, we have the following theorem.

Theorem 3.8: Assume that sensing matrix $\mathbf{\Phi}$ satisfies RIP of order $3K$ with RIC

$$\delta_{3K} \leq 0.2412. \tag{3.15}$$

Then, the SP algorithm reconstructs the K-sparse vector $\mathbf{x}$ from the measurement $\mathbf{y} = \mathbf{\Phi x} + \mathbf{w}$ with the reconstruction error bounded as

$$\left\|\mathbf{x} - \hat{\mathbf{x}}\right\|_2 \leq c_K \left\|\mathbf{w}\right\|_2,$$

with

$$c_K \triangleq \left(1 + \frac{\delta_{3K}\sqrt{1 + \delta_{3K}}}{\sqrt{1 - \delta_{3K}}}\right)\left(\frac{2 + \sqrt{1 + \delta_{3K}}\beta}{\sqrt{1 - \delta_{3K}} - \sqrt{1 + \delta_{3K}}\alpha}\right) + \frac{1}{\sqrt{1 - \delta_{3K}}}, \tag{3.16}$$

where $\hat{\mathbf{x}}$ is the estimated sparse signal vector,

$$\alpha \triangleq \left(\frac{2\delta_{3K}}{1 - \delta_{3K}}\right)\sqrt{1 + \delta_{3K}^2\frac{1 + \delta_{3K}}{1 - \delta_{3K}}}\sqrt{1 + 4\delta_{3K}^2\frac{1 + \delta_{3K}}{1 - \delta_{3K}}}, \tag{3.17}$$

and

$$\beta \triangleq \left(\frac{2\sqrt{1+\delta_{3K}}}{1-\delta_{3K}}\right)\sqrt{1+\frac{4\delta_{3K}^2(1+\delta_{3K})}{1-\delta_{3K}}}+\frac{2}{\sqrt{1-\delta_{3K}}}. \quad (3.18)$$

Proof: See Appendix F. □

In summary, for either the noiseless or the noisy case, we have pushed the bound for δ_{3K} in the sufficient conditions to 0.2412. In addition, when noise is present, it can be shown through some algebra that the proposed reconstruction error upper bound (3.16) is smaller than (3.12) (under a fixed RIC δ_{3K}). To further compare our error bound (3.16) with the result in [10], we first use (3.14) to obtain the following bound independent of $\mathbf{\Phi}_{T_e}^*$:

$$\left\|\mathbf{x}-\hat{\mathbf{x}}\right\|_2 \le \bar{c}_K \left\|\mathbf{\Phi}_{T_e}^*\mathbf{w}\right\|_2 \le \bar{c}_K\sqrt{1+\delta_{3K}}\,\|\mathbf{w}\|_2, \quad (3.19)$$

where the second inequality can be obtained by using Lemma A.1 in Appendix A. Based on (3.19) and by invoking the definition of $\bar{c}_K$ in (3.14), it can be shown that $c_K < \bar{c}_K\sqrt{1+\delta_{3K}}$, viz., the proposed bound (3.16) is also smaller than (3.19). Hence, our analysis shows that the reconstruction performance of SP is actually better than as asserted in [9] and [10].

IV. Concluding Remarks

In this paper, we derive improved RIP-based performance guarantees for perfect support identification via OMP. In the noisy case, less restricted sufficient conditions for exact support recovery via OMP is first obtained. Compared to the most recent work [18], our result shows that relaxed requirements on the RIC and the smallest signal entry magnitude can ensure exact support identification. In the noiseless case, our result narrows the gap between the so-far best known bound on the RIC of the sensing matrix and the ultimate performance guarantee. The proposed approach exploits a newly established approximate orthogonality condition, which is characterized via the achievable angles between two compressed orthogonal sparse vectors under RIP. Such a near-orthogonality property in conjunction with the developed analysis techniques evidenced a wider spectrum of applications. Indeed, for the problem of compressive-domain interference cancellation, we derive a more accurate estimate of the RIC of the effective sensing matrix (in comparison to the results in [20] and [23]). Also, we study support identification via SP in both noiseless and noisy settings. By means of the approximate orthogonality condition, it is shown that, compared to [9] and [10], the requirement on the RIC of the sensing matrix for guaranteeing exact/stable signal recovery can be further relaxed; in addition, when noise is present, the reconstruction error upper bound is provably to be smaller. Improved RIP-based performance analysis of other greedy algorithms, such as CoSaMP [30], by using the developed analysis techniques in this paper is currently under investigation.

Appendix

A. Proof of Lemma 2.1

To prove Lemma 2.1, we need the following two lemmas.

Lemma A.1 [9]: Assume that $\mathbf{\Phi} \in \mathrm{R}^{M\times N}$ satisfies RIP of order K with RIC δ_K. Then, for $S \subset \{1,\ldots,N\}$ with $|S| \le K$, the $M\times|S|$ sub-matrix $\mathbf{\Phi}_S$ is of full column rank with singular values $\sigma_1(\mathbf{\Phi}_S) \ge \sigma_2(\mathbf{\Phi}_S) \ge \cdots \ge \sigma_{|S|}(\mathbf{\Phi}_S) > 0$ satisfying

$$\sqrt{1-\delta_K} \le \sigma_j(\mathbf{\Phi}_S) \le \sqrt{1+\delta_K}, \quad j=1,\ldots,|S|. \quad (A.1)$$

□

Lemma A.2: Let $\mathbf{A} \in \mathrm{R}^{m\times n}$, with $m \ge n$, be of full column rank, and $\sigma_1(\mathbf{A}) \ge \sigma_2(\mathbf{A}) \ge \cdots \ge \sigma_n(\mathbf{A}) > 0$ be the singular values of $\mathbf{A}$. Let $\kappa(\mathbf{A}) = \sigma_1(\mathbf{A})/\sigma_n(\mathbf{A})$ be the condition number of $\mathbf{A}$. Then we have

$$\kappa(\mathbf{A}) = \cot(\theta(\mathbf{A})/2) \quad (A.2)$$

where $0 < \theta(\mathbf{A}) \le \pi/2$ is given by

$$\theta(\mathbf{A}) = \min_{\langle\mathbf{x},\mathbf{y}\rangle=0} \measuredangle(\mathbf{Ax},\mathbf{Ay}). \quad (A.3)$$

Proof: The assertion can be shown by directly following the proof in [25, Example 7.4.26]. □

Proof of Lemma 2.1: Let $T_{uv} = T_u \cup T_v$. As $\mathbf{\Phi u} = \mathbf{\Phi}_{T_{uv}}\mathbf{u}_{T_{uv}}$ and $\mathbf{\Phi v} = \mathbf{\Phi}_{T_{uv}}\mathbf{v}_{T_{uv}}$, we have

$$|\cos\measuredangle(\mathbf{\Phi u},\mathbf{\Phi v})| = \left|\cos\left(\mathbf{\Phi}_{T_{uv}}\mathbf{u}_{T_{uv}},\mathbf{\Phi}_{T_{uv}}\mathbf{v}_{T_{uv}}\right)\right|. \quad (A.4)$$

Since $\mathbf{u}$ and $\mathbf{v}$, supported on, respectively, T_u and T_v, are orthogonal, so are the pairs $(\mathbf{u}_{T_{uv}},\mathbf{v}_{T_{uv}})$ and $(-\mathbf{u}_{T_{uv}},\mathbf{v}_{T_{uv}})$. Also, since $|T_u \cup T_v| \le K$ and arbitrary K columns of $\mathbf{\Phi}$ are linearly independent, $\mathbf{\Phi}_{T_{uv}}$ is of full column rank. By Lemma A.2, it follows immediately that

$$\measuredangle(\pm\mathbf{\Phi}_{T_{uv}}\mathbf{u}_{T_{uv}},\mathbf{\Phi}_{T_{uv}}\mathbf{v}_{T_{uv}}) \ge \theta(\mathbf{\Phi}_{T_{uv}}). \quad (A.5)$$

From (A.5), we also have

$$\measuredangle(\mp\mathbf{\Phi}_{T_{uv}}\mathbf{u}_{T_{uv}},\mathbf{\Phi}_{T_{uv}}\mathbf{v}_{T_{uv}}) \le \pi-\theta(\mathbf{\Phi}_{T_{uv}}). \quad (A.6)$$

With (A.5) and (A.6), it then follows that

$$\left|\cos\measuredangle\left(\mathbf{\Phi}_{T_{uv}}\mathbf{u}_{T_{uv}},\mathbf{\Phi}_{T_{uv}}\mathbf{v}_{T_{uv}}\right)\right| \le \cos(\theta(\mathbf{\Phi}_{T_{uv}})). \quad (A.7)$$

It thus remains to show

$$\cos(\theta(\mathbf{\Phi}_{T_{uv}})) \le \delta_K. \quad (A.8)$$

The assertion of Lemma 2.1 then follows from (A.4), (A.7) and (A.8). To prove (A.8), it is noted that, since $\kappa(\mathbf{\Phi}_{T_{uv}}) = \cot(\theta(\mathbf{\Phi}_{T_{uv}})/2)$ (cf. (A.2)), we have

$$\cos(\theta(\mathbf{\Phi}_{T_{uv}})) = \frac{\cot^2\left(\theta(\mathbf{\Phi}_{T_{uv}})/2\right)-1}{\cot^2\left(\theta(\mathbf{\Phi}_{T_{uv}})/2\right)+1} = \frac{\kappa^2(\mathbf{\Phi}_{T_{uv}})-1}{\kappa^2(\mathbf{\Phi}_{T_{uv}})+1} = 1-\frac{2}{\kappa^2(\mathbf{\Phi}_{T_{uv}})+1}. \quad (A.9)$$

By definition, $\kappa(\mathbf{\Phi}_{T_{uv}}) = \bar{\sigma}(\mathbf{\Phi}_{T_{uv}})/\sigma(\mathbf{\Phi}_{T_{uv}})$, where $\bar{\sigma}(\mathbf{\Phi}_{T_{uv}})$ and $\bar{\sigma}(\mathbf{\Phi}_{T_{uv}})$ are, respectively, the maximal and minimal singular values of $\mathbf{\Phi}_{T_{uv}}$. Hence, with Lemma A.1 it follows that

$$\kappa(\mathbf{\Phi}_{T_{uv}}) = \frac{\bar{\sigma}(\mathbf{\Phi}_{T_{uv}})}{\sigma(\mathbf{\Phi}_{T_{uv}})} \le \frac{\sqrt{1+\delta_K}}{\sigma(\mathbf{\Phi}_{T_{uv}})} \le \frac{\sqrt{1+\delta_K}}{\sqrt{1-\delta_K}}. \quad (A.10)$$

Inequality (A.8) can be directly obtained from (A.9) and (A.10). □

B. Proof of Lemma 2.4

Let $\mathbf{\Phi}_T = \mathbf{U}\Sigma\mathbf{V}^*$ be a singular value decomposition of $\mathbf{\Phi}_T \in \mathrm{R}^{M\times|T|}$, where $\mathbf{U} \in \mathrm{R}^{M\times M}$ and $\mathbf{V} \in \mathrm{R}^{|T|\times|T|}$ are orthogonal, and $\Sigma = \begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix} \in \mathrm{R}^{M\times|T|}$ with $\Sigma_1 \in \mathrm{R}^{|T|\times|T|}$ being diagonal with positive diagonal entries. The assertion of Lemma 3.2 follows from the following set of relations:

$$\begin{aligned}
\|\mathbf{\Phi}_T^*\mathbf{\Phi}\mathbf{x}\|_2 &= \|\mathbf{\Phi}_T^*\mathbf{\Phi}_T\mathbf{x}_T\|_2 = \|\mathbf{V}\Sigma^*\mathbf{U}^*\mathbf{U}\Sigma\mathbf{V}^*\mathbf{x}_T\|_2 \\
&\overset{(a)}{=} \left\|\mathbf{V}[\Sigma_1^*\mathbf{0}]\begin{bmatrix}\mathbf{U}_1^* \\ \mathbf{U}_2^*\end{bmatrix}[\mathbf{U}_1\mathbf{U}_2]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&= \left\|\mathbf{V}[\Sigma_1^*\mathbf{0}]\begin{bmatrix}\mathbf{U}_1^* \\ \mathbf{U}_2^*\end{bmatrix}[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&\overset{(b)}{=} \left\|[\Sigma_1^*\mathbf{0}]\begin{bmatrix}\mathbf{U}_1^* \\ \mathbf{U}_2^*\end{bmatrix}[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&= \left\|\Sigma_1^*\mathbf{U}_1^*[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&\overset{(c)}{\geq} \sqrt{1-\delta_{K+1}}\left\|\mathbf{U}_1^*[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&= \sqrt{1-\delta_{K+1}}\left\|\begin{bmatrix}\mathbf{U}_1^* \\ \mathbf{0}\end{bmatrix}[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&\overset{(d)}{=} \sqrt{1-\delta_{K+1}}\left\|\begin{bmatrix}\mathbf{U}_1^* \\ \mathbf{U}_2^*\end{bmatrix}[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&\overset{(e)}{=} \sqrt{1-\delta_{K+1}}\left\|[\mathbf{U}_1\mathbf{0}]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&= \sqrt{1-\delta_{K+1}}\left\|[\mathbf{U}_1\mathbf{U}_2]\begin{bmatrix}\Sigma_1 \\ \mathbf{0}\end{bmatrix}\mathbf{V}^*\mathbf{x}_T\right\|_2 \\
&= \sqrt{1-\delta_{K+1}}\,\|\mathbf{U}\Sigma\mathbf{V}^*\mathbf{x}_T\|_2 \\
&= \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}_T\mathbf{x}_T\|_2 \\
&= \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}\mathbf{x}\|_2 \\
&\overset{(f)}{\geq} (1-\delta_{K+1})\,\|\mathbf{x}\|_2,
\end{aligned}$$

where in (a) we partition $\mathbf{U}$ into $\mathbf{U} = [\mathbf{U}_1\mathbf{U}_2]$, where $\mathbf{U}_1 \in \mathrm{R}^{M\times|T|}$ and $\mathbf{U}_2 \in \mathrm{R}^{M\times(M-|T|)}$, (b) holds since $\mathbf{V}$ is orthogonal, (c) is true since Lemma A.1 asserts that the singular values of $\mathbf{\Phi}_T$ (appearing as the diagonal entries of Σ_1^*) are no less than $\sqrt{1-\delta_{K+1}}$, (d) is true because $\mathbf{U}_2^*\mathbf{U}_1 = \mathbf{0}$, and (e) follows since $[\mathbf{U}_1\mathbf{U}_2]^* = \mathbf{U}^*$ is orthogonal, and (f) is due to RIP. This thus proves the lemma. □

C. Proof of Theorem 2.5 and Corollary 2.6

We present first the proof of performance guarantees via OMP with l_2- and l_∞-bounded noise (case (1) and case (2) in Theorem 2.5). The proof for noiseless case (Corollary 2.6) is similar and is shown next.

Proof of Case (1): Note that, in the j-th iteration, the index ρ^j determined yields the maximal $\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right|$ (see Step 3.2 in Table I). We first claim that, if δ_{K+1} satisfies (2.9) and the requirement (2.10) holds, such ρ^j's, where $j = 1, \ldots, K$, belong to the support T. Also, according to the orthogonality property of OMP [8, Lemma 7], all the selected indexes ρ^j's, $j = 1, \ldots, K$, are distinct. We will then prove that OMP with the stopping criterion $\left\|\mathbf{r}^j\right\|_2 \leq \varepsilon_1$ halts exactly after K iterations. The assertion of the theorem then follows.

To prove the claim, it suffices to show that the following conditions hold for all $j = 1, \ldots, K$: in the j-th iteration, there exists some K-sparse $\mathbf{z}^j \in \mathrm{R}^N$ with support T such that $\mathbf{r}^{j-1}$ in Step 3.2 reads $\mathbf{r}^{j-1} = \mathbf{\Phi}\mathbf{z}^j + \mathbf{w}$,

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right| \leq \left\|\mathbf{\Phi}\mathbf{z}^j\right\|_2 \delta_{K+1} + \varepsilon_1 \text{ for all } i \notin T, \tag{C.1}$$

and

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right| > \left\|\mathbf{\Phi}\mathbf{z}^j\right\|_2 \delta_{K+1} + \varepsilon_1 \text{ for some } i \in T. \tag{C.2}$$

Hence, we have $\rho^j \in T$ for all $j = 1, \ldots, K$. The proof of (C.1) and (C.2) is done by induction. In the first iteration ($j = 1$), $\mathbf{r}^{j-1}$ needed to compute the inner product in Step 3.2 is

$$\mathbf{r}^0 = \mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{w}, \tag{C.3}$$

and hence

$$\begin{aligned}
\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right| &= \left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{0}\right\rangle\right| = |\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{\Phi}\mathbf{x}\rangle + \langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{w}\rangle| \\
&\leq |\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{\Phi}\mathbf{x}\rangle| + |\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{w}\rangle| \\
&\leq |\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{\Phi}\mathbf{x}\rangle| + \varepsilon_1,
\end{aligned} \tag{C.4}$$

where the last inequality follows since each column of $\mathbf{\Phi}$ is of unit-norm and, thus, $|\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{w}\rangle| \leq \|\mathbf{\Phi}\mathbf{e}_i\|_2\,\|\mathbf{w}\|_2 \leq \varepsilon_1$. Note that, for $i \notin T$, we have $\langle \mathbf{e}_i, \mathbf{x}\rangle = 0$ and $|\{i\} \cup T| = K + 1$. From Lemma 2.3 and since each column of $\mathbf{\Phi}$ is of unit-norm, it follows immediately that

$$\begin{aligned}
|\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{\Phi}\mathbf{x}\rangle| &\leq \|\mathbf{\Phi}\mathbf{e}_i\|_2\,\|\mathbf{\Phi}\mathbf{x}\|_2\,\delta_{K+1} \\
&= \|\mathbf{\Phi}\mathbf{x}\|_2\,\delta_{K+1} \text{ for all } i \notin T.
\end{aligned} \tag{C.5}$$

From (C.4) and (C.5), we have

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{0}\right\rangle\right| \leq \|\mathbf{\Phi}\mathbf{x}\|_2\,\delta_{K+1} + \varepsilon_1 \text{ for all } i \notin T, \tag{C.6}$$

i.e., (C.1) holds when $j = 1$ with $\mathbf{z}^1 = \mathbf{x}$. We then go on to show by contradiction that (C.2) is also true when $j = 1$ with $\mathbf{z}^1 = \mathbf{x}$. Assume otherwise that

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{0}\right\rangle\right| \leq \|\mathbf{\Phi}\mathbf{x}\|_2\,\delta_{K+1} + \varepsilon_1 \textit{ for all } i \in T. \tag{C.7}$$

Then, it follows from (C.7) that

$$\begin{aligned}
\left\|\mathbf{\Phi}_T^*\mathbf{r}^0\right\|_2 &= \sqrt{\sum_{i\in T}\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{0}\right\rangle\right|^2} \leq \sqrt{K}\,(\|\mathbf{\Phi}\mathbf{x}\|_2\,\delta_{K+1} + \varepsilon_1) \\
&= \sqrt{K}\delta_{K+1}\,\|\mathbf{\Phi}\mathbf{x}\|_2 + \sqrt{K}\varepsilon_1.
\end{aligned} \tag{C.8}$$

Also,

$$\begin{aligned}
\left\|\mathbf{\Phi}_T^*\mathbf{r}^0\right\|_2 &= \left\|\mathbf{\Phi}_T^*(\mathbf{\Phi}\mathbf{x} + \mathbf{w})\right\|_2 \geq \left\|\mathbf{\Phi}_T^*\mathbf{\Phi}\mathbf{x}\right\|_2 - \left\|\mathbf{\Phi}_T^*\mathbf{w}\right\|_2 \\
&\overset{(a)}{\geq} \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}\mathbf{x}\|_2 - \left\|\mathbf{\Phi}_T^*\mathbf{w}\right\|_2 \\
&\overset{(b)}{\geq} \sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}\mathbf{x}\|_2 - \sqrt{1+\delta_{K+1}}\varepsilon_1,
\end{aligned} \tag{C.9}$$

where (a) follows from Lemma 2.4 and (b) is true due to Lemma A.1 and $\|\mathbf{w}\|_2 \leq \varepsilon_1$. Under assumption (2.9), it can be shown that

$$\begin{aligned}
\sqrt{1-\delta_{K+1}}\,\|\mathbf{\Phi}\mathbf{x}\|_2 - \sqrt{1+\delta_{K+1}}\varepsilon_1 &> \sqrt{K}\delta_{K+1}\,\|\mathbf{\Phi}\mathbf{x}\|_2 \\
&\quad + \sqrt{K}\varepsilon_1.
\end{aligned} \tag{C.10}$$

Then, combining (C.9) and (C.10) yields

$$\left\|\mathbf{\Phi}_T^*\mathbf{r}^0\right\|_2 > \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1, \tag{C.11}$$

which contradicts with (C.8). This then implies (C.2) is true for $j = 1$ with $\mathbf{z}^1 = \mathbf{x}$. To verify (C.10), we first write

$$\begin{aligned}\sqrt{1-\delta_{K+1}}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 - \sqrt{1+\delta_{K+1}}\varepsilon_1 &= \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1 \\ &\quad + \underbrace{\left(\sqrt{1-\delta_{K+1}} - \sqrt{K}\delta_{K+1}\right)}_{\gamma}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 \\ &\quad - \left(\sqrt{1+\delta_{K+1}} + \sqrt{K}\right)\varepsilon_1. \end{aligned} \tag{C.12}$$

Under assumption (2.9), one can check that

$$\gamma = \sqrt{1-\delta_{K+1}} - \sqrt{K}\delta_{K+1} \geq 0. \tag{C.13}$$

Thus, it follows from (C.12) and (C.13) that

$$\begin{aligned}&\sqrt{1-\delta_{K+1}}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 - \sqrt{1+\delta_{K+1}}\varepsilon_1 \\ &\overset{(a)}{\geq} \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1 \\ &\quad + \underbrace{\left(\sqrt{1-\delta_{K+1}} - \sqrt{K}\delta_{K+1}\right)}_{=\gamma}\sqrt{1-\delta_{K+1}}\left\|\mathbf{x}\right\|_2 \\ &\quad - \left(\sqrt{1+\delta_{K+1}} + \sqrt{K}\right)\varepsilon_1 \\ &\overset{(b)}{>} \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1 \\ &\quad + \sqrt{K}(\sqrt{1+\delta_{K+1}}+1)\varepsilon_1 - \left(\sqrt{1+\delta_{K+1}} + \sqrt{K}\right)\varepsilon_1 \\ &= \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1 + \underbrace{\varepsilon_1(\sqrt{K}-1)\sqrt{1+\delta_{K+1}}}_{\geq 0} \\ &\geq \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \sqrt{K}\varepsilon_1, \end{aligned} \tag{C.14}$$

where (a) follows since $\mathbf{\Phi}$ satisfies RIP and $\mathbf{x}$ is a K-sparse vector, (b) follows since $\mathbf{x}$ is K-sparse (thus, with at most K nonzero entries), and, using assumption (2.10) together with straightforward manipulations.

We move on to the next iteration. Since the index selected in the first iteration belongs to T, in the second iteration ($j = 2$), $\mathbf{r}^{j-1}$ needed in Step 3.2 is

$$\begin{aligned}\mathbf{r}^{j-1} = \mathbf{r}^1 &= \mathbf{y} - \mathbf{\Phi}_{\mathbf{\Omega}^1}\mathbf{q}_{\mathbf{\Omega}^1} = \mathbf{\Phi}\mathbf{x} - \mathbf{\Phi}_{\mathbf{\Omega}^1}\mathbf{q}_{\mathbf{\Omega}^1} + \mathbf{w} \\ &= \mathbf{\Phi}\mathbf{x} - \mathbf{\Phi}\tilde{\mathbf{q}}_{\mathbf{\Omega}^1} + \mathbf{w} = \mathbf{\Phi}\underbrace{(\mathbf{x} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^1})}_{\triangleq \mathbf{z}^2}\mathbf{z}^2 + \mathbf{w}, \end{aligned} \tag{C.15}$$

where $\tilde{\mathbf{q}}_{\mathbf{\Omega}^1} \in \mathrm{R}^N$ is a 1-sparse vector with all entries equal to zero except the one indexed by $\mathbf{\Omega}^1 \subset T$. Then $\mathbf{z}^2 = \mathbf{x} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^1}$ is a K-sparse vector with support T. By following essentially the same procedures as in the first iteration, (C.1) and (C.2) can be shown to be true for $j = 2$ with the K-sparse vector $\mathbf{z}^2$. By repeating such procedures, one can inductively show that the first K selected indexes all belong to the support T. Also, since the selected indexes are distinct, the columns of $\mathbf{\Phi}_{\mathbf{\Omega}^K} \in \mathrm{R}^{M\times K}$ are those of $\mathbf{\Phi}$ indexed by T.

Now we turn to show that, with the stopping criterion $\left\|\mathbf{r}^j\right\|_2 \leq \varepsilon_1$, OMP stops exactly after K iterations. This can be done by showing that

$$\left\|\mathbf{r}^j\right\|_2 > \varepsilon_1 \quad \text{for all } j = 0, \ldots, K-1, \quad \text{and } \left\|\mathbf{r}^j\right\|_2 \leq \varepsilon_1 \text{ for } j = K. \tag{C.16}$$

The condition (C.16) can be proved by following essentially the same procedures as in the analyses between eq. (24)∼(26) in [18]; the details are thus omitted (the interested readers are referred to [31]).

Proof of Case (2): Similar to case (1), it suffices to show that the indexes selected in the first K iterations all belong to T by the following claim: in the j -th iteration, $j = 1, \ldots, K$, there exists some K-sparse $\mathbf{z}^j \in \mathrm{R}^N$ with support T such that $\mathbf{r}^{j-1}$ in Step 3.2 reads $\mathbf{r}^{j-1} = \mathbf{\Phi}\mathbf{z}^j + \mathbf{w}$,

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right| \leq \left\|\mathbf{\Phi}\mathbf{z}^j\right\|_2\delta_{K+1} + \varepsilon_2 \text{ for all } i \notin T, \tag{C.17}$$

and

$$\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^{j-1}\right\rangle\right| > \left\|\mathbf{\Phi}\mathbf{z}^j\right\|_2\delta_{K+1} + \varepsilon_2 \text{ for some } i \in T. \tag{C.18}$$

By going through essentially the same induction-and-contradiction procedures as in Case (1), it can be verified that (C.17) and (C.18) are true. Also, the indices thus identified are distinct, according to [8]. It then remains to show that OMP with the stopping criterion $\max_{i=1,\ldots,N}\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^j\right\rangle\right| \leq \varepsilon_2$ will not halt during the first K iterations. This can be done by proving that

$$\max_{i=1,\ldots,N}\left|\left\langle \mathbf{\Phi}\mathbf{e}_i, \mathbf{r}^j\right\rangle\right| > \varepsilon_2 \text{ for all } j = 0, \ldots, K-1. \tag{C.19}$$

Still, the condition (C.19) can be proved by following essentially the same procedures in the proof of [18, Th. 2], and the details are thus omitted (the interested readers are referred to [31]).

Proof of Corollary 2.6 (Noiseless Case): Since the selected indexes are distinct [8, Lemma 7], it suffices to show that the first K identified indexes belong to the support T. The proof is essentially the same as that of the l_2–bounded noise case. In particular, we show that (C.1) and (C.2) hold for $j = 1, \ldots, K$ with $\varepsilon_1 = 0$. For $j = 1$, the assertion can be easily shown by first going through the proof from (C.3) to (C.13). Note that, with $\varepsilon_1 = 0$, (C.12) becomes $\sqrt{1-\delta_{K+1}}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 = \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 + \left(\sqrt{1-\delta_{K+1}} - \sqrt{K}\delta_{K+1}\right)\left\|\mathbf{\Phi}\mathbf{x}\right\|_2$. This equality, together with (C.13), implies

$$\sqrt{1-\delta_{K+1}}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2 \leq \sqrt{K}\delta_{K+1}\left\|\mathbf{\Phi}\mathbf{x}\right\|_2. \tag{C.20}$$

Then, replace (C.14) by (C.20). Thus we can conclude that (C.1) and (C.2) with $\varepsilon_1 = 0$ are true for $j = 1$. By induction, one can show that (C.1) and (C.2) with $\varepsilon_1 = 0$ holds for $j = 2, \ldots, K$. □

D. Proof of Theorem 3.3

In the sequel, for $S \subset \{1, \ldots, N\}$ with $|S| < K$, $\mathbf{P}_S \triangleq \mathbf{\Phi}_S(\mathbf{\Phi}_S^*\mathbf{\Phi}_S)^{-1}\mathbf{\Phi}_S^*$ represents the orthogonal projection onto $\mathcal{R}(\mathbf{\Phi}_S)$. We need the next lemma to complete the proof.

Lemma A.3: Assume that $\mathbf{\Phi}$ satisfies RIP of order K with RIC δ_K. Let $S \subset \{1, \ldots, N\}$ with $|S| < K$, and $\mathbf{P}_S$ be the

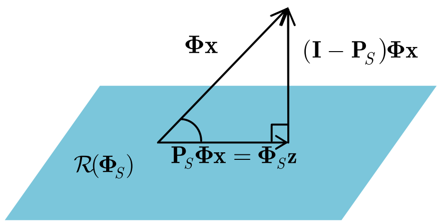

Fig. 3. Schematic description of orthogonal projection of Φx onto the column space of Φ_S.

orthogonal projection onto $\mathcal{R}(\Phi_S)$. Then for all $(K - |S|)$-sparse $\mathbf{x}$ whose support does not overlap with S, we have

$$\|\mathbf{P}_S\Phi\mathbf{x}\|_2^2 \le \delta_K^2 \|\Phi\mathbf{x}\|_2^2 \le (1+\delta_K)\delta_K^2 \|\mathbf{x}\|_2^2, \quad \text{(D.1)}$$

and

$$\begin{aligned}\|(\mathbf{I}-\mathbf{P}_S)\Phi\mathbf{x}\|_2^2 &\ge (1-\delta_K^2)\|\Phi\mathbf{x}\|_2^2 \\ &\ge (1-\delta_K)(1-\delta_K^2)\|\mathbf{x}\|_2^2. \quad \text{(D.2)}\end{aligned}$$

Proof of Lemma A.3: Since $\mathbf{P}_S = \Phi_S(\Phi_S^*\Phi_S)^{-1}\Phi_S^*$, we have $\mathbf{P}_S\Phi\mathbf{x} = \Phi_S(\Phi_S^*\Phi_S)^{-1}\Phi_S^*\Phi\mathbf{x} = \Phi_S\mathbf{z}$ with

$$\mathbf{z} = (\Phi_S^*\Phi_S)^{-1}\Phi_S^*\Phi\mathbf{x}. \quad \text{(D.3)}$$

It follows that (see Figure 3)

$$\|\mathbf{P}_S\Phi\mathbf{x}\|_2^2 = \|\Phi\mathbf{x}\|_2^2 \left|\cos\measuredangle(\Phi\mathbf{x}, \Phi_S\mathbf{z})\right|^2$$

and

$$\|(\mathbf{I}-\mathbf{P}_S)\Phi\mathbf{x}\|_2^2 = \|\Phi\mathbf{x}\|_2^2 \left|\sin\measuredangle(\Phi\mathbf{x}, \Phi_S\mathbf{z})\right|^2. \quad \text{(D.4)}$$

Since $\Phi_S\mathbf{z}$ can be written as

$$\Phi_S\mathbf{z} = \Phi\tilde{\mathbf{z}}, \quad \text{(D.5)}$$

where $\tilde{\mathbf{z}}$ is obtained by padding zeros to $\mathbf{z}$ (the support of $\tilde{\mathbf{z}}$ is thus S), we obtain

$$\measuredangle(\Phi\mathbf{x}, \Phi_S\mathbf{z}) = \measuredangle(\Phi\mathbf{x}, \Phi\tilde{\mathbf{z}}). \quad \text{(D.6)}$$

Note that $\mathbf{x}$ and $\tilde{\mathbf{z}}$ are orthogonal since the supports of $\mathbf{x}$ and $\tilde{\mathbf{z}}$ do not overlap. According to Lemma 2.1, it follows

$$\left|\cos\measuredangle(\Phi\mathbf{x}, \Phi\tilde{\mathbf{z}})\right|^2 \le \delta_K^2 \quad \text{(D.7)}$$

and thus

$$\begin{aligned}\left|\sin\measuredangle(\Phi\mathbf{x}, \Phi\tilde{\mathbf{z}})\right|^2 &= 1 - \left|\cos\measuredangle(\Phi\mathbf{x}, \Phi\tilde{\mathbf{z}})\right|^2 \\ &\ge 1 - \delta_K^2. \quad \text{(D.8)}\end{aligned}$$

Combining (D.4), (D.7) and (D.8) yields $\|\mathbf{P}_S\Phi\mathbf{x}\|_2^2 \le \delta_K^2\|\Phi\mathbf{x}\|_2^2$ and $\|(\mathbf{I}-\mathbf{P}_S)\Phi\mathbf{x}\|_2^2 \ge (1-\delta_K^2)\|\Phi\mathbf{x}\|_2^2$. Also, since Φ satisfies RIP, (D.1) and (D.2) directly follow. □

Proof of Theorem 3.3: By definition, $\mathbf{Q}$ is the orthogonal projection onto the orthogonal complement of $\mathcal{R}(\Phi_S)$. Under the assumptions that Φ satisfies the RIP of order K with RIC δ_K, and that the support of the $(K - |S|)$-sparse $\mathbf{x}$ does not overlap with T_d, it then follows from Lemma A.3 that

$$\|\mathbf{Q}\Phi\mathbf{x}\|_2^2 \ge (1-\delta_K)(1-\delta_K^2)\|\mathbf{x}\|_2^2. \quad \text{(D.9)}$$

Also, we have

$$\|\mathbf{Q}\Phi\mathbf{x}\|_2^2 \le \|\Phi\mathbf{x}\|_2^2 \le (1+\delta_K)\|\mathbf{x}\|_2^2. \quad \text{(D.10)}$$

Combining (D.9) with (D.10) immediately gives

$$(1-\delta_K)(1-\delta_K^2)\|\mathbf{x}\|_2^2 \le \|\mathbf{Q}\Phi\mathbf{x}\|_2^2 \le (1+\delta_K)\|\mathbf{x}\|_2^2. \quad \text{(D.11)}$$

The assertion thus follows from (D.11). □

E. Proof of Theorem 3.5

The proof is based on that of [9, Th. 1]; in particular, it suffices to show

$$\|\mathbf{r}^j\|_2 < \|\mathbf{r}^{j-1}\|_2 \text{ if } \|\mathbf{r}^{j-1}\|_2 > 0, \quad \text{(E.1)}$$

that is, the norm of the residual is reduced iteration by iteration provided that $\left\|\mathbf{r}^{j-1}\right\|_2 > 0$. If (E.1) is true, the stopping criterion $\left\|\mathbf{r}^{j}\right\|_2 \ge \left\|\mathbf{r}^{j-1}\right\|_2$ will be satisfied only when $\left\|\mathbf{r}^{j_0-1}\right\|_2 = 0$ for some $j_0 - 1$[2]. Indeed, if $\left\|\mathbf{r}^{j_0-1}\right\|_2 = 0$, SP will halt at the j_0-th iteration since $\left\|\mathbf{r}^{j_0}\right\|_2 \ge \left\|\mathbf{r}^{j_0-1}\right\|_2 = 0$. It then remains to show that, for such a j_0, we also have $\left\|\mathbf{r}^{j_0-1}\right\|_2 = 0$, implying that the estimated sparse vector after j_0 iterations is exactly $\mathbf{x}$. Hence, the proof of theorem 3.5 is completed.

Notably, in [9], the proof first relies on Theorems 3 and 4 to obtain [9, eq. (6)], which then leads to (E.1). Our proof basically follows the similar flow; in particular, the improved performance guarantee shown in Theorem 3.5 is obtained by using the newly derived approximate orthogonality condition (2.1) to tighten the performance bounds in [9, Th. 3 and 4] and [9, eq. (16)]. First of all, based on the newly derived approximate orthogonality condition (2.1) we obtain the following two lemmas, which improve the results in, respectively, [9, Th. 3 and 4].

Lemma A.4: Let Ω_Δ^j be the index set, with $|\Omega_\Delta^j| = K$, such that $\left\{\left|(\Phi^*\mathbf{r}^{j-1})_i\right|\right\}_{i\in\Omega_\Delta^j}$ consists of the K largest elements of $\left\{\left|(\Phi^*\mathbf{r}^{j-1})_1\right|, \ldots, \left|(\Phi^*\mathbf{r}^{j-1})_N\right|\right\}$, where $\mathbf{r}^{j-1}$ is the residual vector in the $(j-1)$-th iteration. Then, for $\breve{\Omega}^j = \Omega^{j-1} \cup \Omega_\Delta^j$,

$$\left\|\mathbf{x}_{T\setminus\breve{\Omega}^j}\right\|_2 \le \frac{2\delta_{3K}}{1-\delta_{3K}}\sqrt{1+\frac{\delta_{3K}^2(1+\delta_{3K})}{1-\delta_{3K}}}\left\|\mathbf{x}_{T\setminus\Omega^{j-1}}\right\|_2. \quad \text{(E.2)}$$

Proof: The proof is placed at the end of this appendix. □

Lemma A.5: Under the same assumptions made as in Lemma A.4, we have the following result:

$$\left\|\mathbf{x}_{T\setminus\Omega^j}\right\|_2 \le \sqrt{1+\frac{4\delta_{3K}^2(1+\delta_{3K})}{1-\delta_{3K}}}\left\|\mathbf{x}_{T\setminus\breve{\Omega}^j}\right\|_2. \quad \text{(E.3)}$$

Proof: The proof is placed at the end of this appendix. □

Based on Lemma A.4 and Lemma A.5, we have the following lemma, which improves the bound provided in [9, eq. (6)].

Lemma A.6: Assume that the sensing matrix Φ satisfies RIP of order $3K$ with RIC $\delta_{3K} \le 0.2412$. Let Ω^j be the estimated support in the j-th iteration of the SP algorithm. Then we have

$$\left\|\mathbf{x}_{T\setminus\Omega^j}\right\|_2 \le \alpha\left\|\mathbf{x}_{T\setminus\Omega^{j-1}}\right\|_2, \quad \text{(E.4)}$$

[2] An upper bound for $j_0 - 1$ can be found in [9].

where α is defined in (3.17).

Proof: Inequality (E.4) can be directly obtained from (E.2), (E.3), together with some straightforward manipulations. □

To claim (E.1), we first recall from Steps 3.7-3.8 in Table II that the residual vector is

$$\mathbf{r}^j = \mathbf{y} - \mathbf{P}_{\mathbf{\Omega}^j}\mathbf{y} = [\mathbf{I} - \mathbf{P}_{\mathbf{\Omega}^j}]\underbrace{(\mathbf{\Phi}_{\Omega^j}\mathbf{x}_{\mathbf{\Omega}^j} + \mathbf{\Phi}_{T\backslash\mathbf{\Omega}^j}\mathbf{x}_{T\backslash\mathbf{\Omega}^j})}_{\mathbf{y}=\mathbf{\Phi x}}$$
$$\overset{(a)}{=} (\mathbf{I} - \mathbf{P}_{\mathbf{\Omega}^j})\mathbf{\Phi}_{T\backslash\mathbf{\Omega}^j}\mathbf{x}_{T\backslash\mathbf{\Omega}^j}, \tag{E.5}$$

where (a) holds since $\mathbf{\Phi}_{\mathbf{\Omega}^j}\mathbf{x}_{\mathbf{\Omega}^j} \in \mathcal{R}(\mathbf{\Phi}_{\mathbf{\Omega}^j})$ and $\mathbf{P}_{\mathbf{\Omega}^j}$ is the orthogonal projection onto $\mathcal{R}(\mathbf{\Phi}_{\mathbf{\Omega}^j})$. With (E.5), we have

$$\begin{aligned}\|\mathbf{r}^j\|_2 &= \|(\mathbf{I} - \mathbf{P}_{\mathbf{\Omega}^j})\mathbf{\Phi}_{T\backslash\mathbf{\Omega}^j}\mathbf{x}_{T\backslash\mathbf{\Omega}^j}\|_2 \le \|\mathbf{\Phi}_{T\backslash\mathbf{\Omega}^j}\mathbf{x}_{T\backslash\mathbf{\Omega}^j}\|_2\\ &\overset{(a)}{\le} \sqrt{1+\delta_{3K}}\|\mathbf{x}_{T\backslash\mathbf{\Omega}^j}\|_2\\ &\overset{(b)}{\le} \sqrt{1+\delta_{3K}}\alpha\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2,\end{aligned} \tag{E.6}$$

where (a) follows due to Lemma A.1 and (b) holds from (E.4). Now, based on Step 3.8, $\mathbf{r}^j$ can also be written as[3]

$$\begin{aligned}\mathbf{r}^j &= \mathbf{y} - \mathbf{\Phi}_{\mathbf{\Omega}^j}\mathbf{q}_{\mathbf{\Omega}^j} = \mathbf{\Phi}_{T\backslash\mathbf{\Omega}^j}\mathbf{x}_{T\backslash\mathbf{\Omega}^j} + \mathbf{\Phi}_{\mathbf{\Omega}^j}\mathbf{x}_{\Omega^j} - \mathbf{\Phi}_{\mathbf{\Omega}^j}\mathbf{q}_{\mathbf{\Omega}^j}\\ &= \mathbf{\Phi}(\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^j} + \tilde{\mathbf{x}}_{\mathbf{\Omega}^j} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^j})\end{aligned}$$

thus we have

$$\begin{aligned}\|\mathbf{r}^{j-1}\|_2 &= \|\mathbf{\Phi}(\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}} + \tilde{\mathbf{x}}_{\Omega^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}})\|_2\\ &\overset{(a)}{\ge} \sqrt{1-\delta_{3K}}\|\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}} + \tilde{\mathbf{x}}_{\mathbf{\Omega}^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}}\|_2\\ &\overset{(b)}{=} \sqrt{1-\delta_{3K}}\sqrt{\|\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2^2 + \|\tilde{\mathbf{x}}_{\Omega^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}}\|_2^2}\\ &\ge \sqrt{1-\delta_{3K}}\|\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2\\ &= \sqrt{1-\delta_{3K}}\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2,\end{aligned} \tag{E.7}$$

where (a) follows from the RIP, and (b) is true since the supports of $\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}}$ and $\tilde{\mathbf{x}}_{\mathbf{\Omega}^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}}$ are disjoint. One can verify that if $\delta_{3K} \le 0.2412$, the following inequality holds:

$$\sqrt{1-\delta_{3K}} > \sqrt{1+\delta_{3K}}\alpha. \tag{E.8}$$

The assertion (E.1) thus follows by combining (E.6), (E.7), and (E.8).

Now we turn to show that if $\|\mathbf{r}^{j_0-1}\|_2 = 0$ for some $j_0 - 1$, we also have $\|\mathbf{r}^{j_0}\|_2 = 0$. Since $\mathbf{r}^{j_0-1} = \mathbf{\Phi x} - \mathbf{\Phi}_{\mathbf{\Omega}^{j_0-1}}\mathbf{q}_{\mathbf{\Omega}^{j_0-1}} = \mathbf{\Phi}(\mathbf{x} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j_0-1}})$ (see Step 3.8 in Table II), we have $0 = \|\mathbf{r}^{j_0-1}\|_2 = \|\mathbf{\Phi}(\mathbf{x} - \tilde{\mathbf{q}}_{\Omega^{j_0-1}})\|_2 \ge \sqrt{1-\delta_{3K}}\|\mathbf{x} - \tilde{\mathbf{q}}_{\Omega^{j_0-1}}\|_2$, immediately leading to $\mathbf{x} = \tilde{\mathbf{q}}_{\Omega^{j_0-1}}$. It then implies the supports of $\mathbf{x}$ and $\tilde{\mathbf{q}}_{\mathbf{\Omega}^{j_0-1}}$, i.e., T and Ω^{j_0-1}, are the same, thereby $\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j_0-1}}\|_2 = 0$. Finally, it follows from (E.4) that $\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j_0}}\|_2 = 0$ as well, namely, the estimated support at the j_0-th iteration is still $\mathbf{\Omega}^{j_0} = T$. Then, $\mathbf{q}$ obtained by Steps 3.6-3.7 is exactly $\mathbf{x}$.

[3]We remind the readers of the following notation usage: For $\mathbf{u} \in \mathrm{R}^N$, $\mathbf{u}_S \in \mathrm{R}^{|S|}$ denotes the vector whose entries consist of those of $\mathbf{u}$ indexed by the subset $S \subset \{1,\ldots,N\}$, and $\tilde{\mathbf{u}}_S \in \mathrm{R}^N$ is the zero-padded version of $\mathbf{u}_S$ such that $(\tilde{\mathbf{u}}_S)_i = (\mathbf{u})_i$ for $i \in S$ and $(\tilde{\mathbf{u}}_S)_i = 0$ otherwise (thus, $\tilde{\mathbf{u}}_S$ is $|S|$-sparse with support S).

[Proof of Lemma A.4]: Write

$$\mathbf{P}_{\mathbf{\Omega}^{j-1}}\mathbf{\Phi}_{T\backslash\mathbf{\Omega}^{j-1}}\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}} = \mathbf{\Phi z} \quad \text{for some } \mathbf{z} \text{ supported on } \mathbf{\Omega}_{j-1}. \tag{E.9}$$

Then, the residual $\mathbf{r}^{j-1}$ in (E.5) can be expressed as

$$\mathbf{r}^{j-1} = \mathbf{\Phi}\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}} - \mathbf{\Phi z} = \mathbf{\Phi h}^{j-1}, \quad \text{where } \mathbf{h}^{j-1} \triangleq \tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}} - \mathbf{z}. \tag{E.10}$$

According to Step 3.1 of the SP algorithm (see Table II), we must have $\|\mathbf{\Phi}_S^*\mathbf{r}^{j-1}\|_2 \le \|\mathbf{\Phi}^*_{\Omega^j_\Delta}\mathbf{r}^{j-1}\|_2$ for any subset $S \subset \{1,\ldots,N\}$ of K elements, and thus $\|\mathbf{\Phi}_T^*\mathbf{r}^{j-1}\|_2^2 \le \|\mathbf{\Phi}^*_{\Omega^j_\Delta}\mathbf{r}^{j-1}\|_2^2$. Since $\|\mathbf{\Phi}_T^*\mathbf{r}^{j-1}\|_2^2 = \|\mathbf{\Phi}^*_{T\backslash\mathbf{\Omega}^j_\Delta}\mathbf{r}^{j-1}\|_2^2 + \|\mathbf{\Phi}^*_{T\cap\Omega^j_\Delta}\mathbf{r}^{j-1}\|_2^2$ and $\|\mathbf{\Phi}^*_{\Omega^j_\Delta}\mathbf{r}^{j-1}\|_2^2 = \|\mathbf{\Phi}^*_{\Omega^j_\Delta\backslash T}\mathbf{r}^{j-1}\|_2^2 + \|\mathbf{\Phi}^*_{\Omega^j_\Delta\cap T}\mathbf{r}^{j-1}\|_2^2$, it then implies

$$\left\|\mathbf{\Phi}^*_{T\backslash\mathbf{\Omega}^j_\Delta}\mathbf{r}^{j-1}\right\|_2^2 \le \left\|\mathbf{\Phi}^*_{\Omega^j_\Delta\backslash T}\mathbf{r}^{j-1}\right\|_2^2. \tag{E.11}$$

From [9, eq. (28) and (29)], an upper bound for the right-hand-side (RHS) of (E.11) is

$$\left\|\mathbf{\Phi}^*_{\Omega^j_\Delta\backslash T}\mathbf{r}^{j-1}\right\|_2 \le \delta_{3K}\left\|\mathbf{h}^{j-1}\right\|_2, \tag{E.12}$$

and a lower bound for the left-hand-side (LHS) of (E.11) is

$$\left\|\mathbf{\Phi}^*_{T\backslash\mathbf{\Omega}^j_\Delta}\mathbf{r}^{j-1}\right\|_2 \ge \left(1-\delta_{3K}\right)\left\|\mathbf{x}_{T\backslash\breve{\mathbf{\Omega}}^j}\right\|_2 - \delta_{3K}\left\|\mathbf{h}^{j-1}\right\|_2. \tag{E.13}$$

It then follows

$$\begin{aligned}\left\|\mathbf{x}_{T\backslash\breve{\mathbf{\Omega}}^j}\right\|_2 &\overset{(a)}{\le} \frac{2\delta_{3K}}{1-\delta_{3K}}\left\|\mathbf{h}^{j-1}\right\|_2\\ &\overset{(b)}{=} \frac{2\delta_{3K}}{1-\delta_{3K}}\sqrt{\left\|\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}}\right\|_2^2 + \left\|\mathbf{z}\right\|_2^2},\end{aligned} \tag{E.14}$$

where (a) is obtained by combining (E.11)~(E.13), and (b) is obtained by using the definition of $\mathbf{h}^{j-1}$ in (E.10) and the fact that the supports of $\tilde{\mathbf{x}}_{T\backslash\mathbf{\Omega}^{j-1}}$ and $\mathbf{z}$ are disjoint. We also note that

$$\begin{aligned}\sqrt{1-\delta_{3K}}\|\mathbf{z}\|_2 &\overset{(a)}{\le} \|\mathbf{\Phi z}\|_2 \overset{(b)}{=} \|\mathbf{P}_{\mathbf{\Omega}^{j-1}}\mathbf{\Phi}_{T\backslash\mathbf{\Omega}^{j-1}}\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2\\ &\overset{(c)}{\le} \delta_{3K}\sqrt{1+\delta_{3K}}\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2,\end{aligned} \tag{E.15}$$

where (a) is due to RIP, (b) holds with (E.9), and (c) is obtained based on Lemma A.3. It follows directly from (E.15) that

$$\|\mathbf{z}\|_2 \le \delta_{3K}\frac{\sqrt{1+\delta_{3K}}}{\sqrt{1-\delta_{3K}}}\|\mathbf{x}_{T\backslash\mathbf{\Omega}^{j-1}}\|_2. \tag{E.16}$$

The assertion of Lemma A.4 follows immediately by combining (E.14) and (E.16). □

Proof of Lemma A.5: According to the SP algorithm (see Step 3.5 in Table II), the K elements of the estimated support $\Omega^j \subset \breve{\mathbf{\Omega}}^j$ are the K indexes corresponding to the K largest magnitudes of $\check{\mathbf{q}}$. Since the entries of $\check{\mathbf{q}}$ are all zeros except those $2K$ elements indexed by $\breve{\mathbf{\Omega}}^j$ (see Steps 3.3-3.4 in Table II), the entries of $\check{\mathbf{q}}$ indexed by $\bar{\mathbf{\Omega}}^j \triangleq \breve{\mathbf{\Omega}}^j\backslash\mathbf{\Omega}^j$ are thus those yielding the K smallest nonzero magnitudes; more precisely, we have $\|\check{\mathbf{q}}_{\bar{\mathbf{\Omega}}^j}\|_2 \le \|\check{\mathbf{q}}_{S_2}\|_2$ for all $S_2 \subset \breve{\mathbf{\Omega}}^j$ of K (or more) elements. Since $(\breve{\boldsymbol{\Omega}}^j \backslash T) \subset \breve{\boldsymbol{\Omega}}^j$ consists of K or more elements, we have

$$\|\breve{\mathbf{q}}_{\bar{\boldsymbol{\Omega}}^j}\|_2 \leq \|\breve{\mathbf{q}}_{\breve{\boldsymbol{\Omega}}^j \backslash T}\|_2. \tag{E.17}$$

As in [9], our proof is done by deriving an upper bound (lower bound, respectively,) for the RHS (LHS, respectively) of (E.17). Thanks to the improved orthogonality condition (2.1), our bounds are tighter than those given in [9], thereby leading to the better performance guarantee (3.10). We claim that

$$\|\breve{\mathbf{q}}_{\breve{\boldsymbol{\Omega}}^j \backslash T}\|_2 \leq \frac{\delta_{3K}\sqrt{1+\delta_{3K}}}{\sqrt{1-\delta_{3K}}}\|\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2, \tag{E.18}$$

and

$$\|\breve{\mathbf{q}}_{\bar{\boldsymbol{\Omega}}^j}\|_2 \geq \sqrt{\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^j}\|_2^2 - \|\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2^2} - \frac{\delta_{3K}\sqrt{1+\delta_{3K}}}{\sqrt{1-\delta_{3K}}}\|\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2. \tag{E.19}$$

Then, inequality (E.3) can be obtained based on (E.17)$\sim$(E.19) together with some straightforward manipulations.

We first prove (E.18). Since $\mathbf{P}_{\breve{\boldsymbol{\Omega}}^j}\boldsymbol{\Phi}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j} = \boldsymbol{\Phi}\mathbf{z}$ for some $\mathbf{z}$ supported on $\breve{\boldsymbol{\Omega}}^j$, it follows that

$$\begin{aligned}\mathbf{P}_{\breve{\boldsymbol{\Omega}}^j}\mathbf{y} &= \mathbf{P}_{\breve{\boldsymbol{\Omega}}^j}(\boldsymbol{\Phi}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j} + \boldsymbol{\Phi}_{\breve{\boldsymbol{\Omega}}^j}\mathbf{x}_{\breve{\boldsymbol{\Omega}}^j}) \\ &\overset{(a)}{=} \boldsymbol{\Phi}\mathbf{z} + \boldsymbol{\Phi}\tilde{\mathbf{x}}_{\breve{\boldsymbol{\Omega}}^j} = \boldsymbol{\Phi}\underbrace{(\mathbf{z}+\tilde{\mathbf{x}}_{\breve{\boldsymbol{\Omega}}^j})}_{\breve{\mathbf{q}} \text{ in Step 3.5}};\end{aligned} \tag{E.20}$$

where (a) holds because $\mathbf{P}_{\breve{\boldsymbol{\Omega}}^j}\boldsymbol{\Phi}_{\breve{\boldsymbol{\Omega}}^j}\mathbf{x}_{\breve{\boldsymbol{\Omega}}^j} = \boldsymbol{\Phi}_{\breve{\boldsymbol{\Omega}}^j}\mathbf{x}_{\breve{\boldsymbol{\Omega}}^j} = \boldsymbol{\Phi}\tilde{\mathbf{x}}_{\breve{\boldsymbol{\Omega}}^j}$. Then, based on (E.20), we obtain

$$\begin{aligned}\|\breve{\mathbf{q}}_{\breve{\boldsymbol{\Omega}}^j \backslash T}\|_2 &= \|(\mathbf{z}+\tilde{\mathbf{x}}_{\breve{\boldsymbol{\Omega}}^j})_{\breve{\boldsymbol{\Omega}}^j \backslash T}\|_2 \\ &\overset{(a)}{=} \|\mathbf{z}_{\breve{\boldsymbol{\Omega}}^j \backslash T}\|_2 \leq \|\mathbf{z}\|_2,\end{aligned} \tag{E.21}$$

where (a) follows since only those entries of $\mathbf{x}$ indexed by T are nonzero. Also, by using the same technique in deriving (E.15)$\sim$(E.16), we have

$$\|\mathbf{z}\|_2 \leq \frac{\delta_{3K}\sqrt{1+\delta_{3K}}}{\sqrt{1-\delta_{3K}}}\|\mathbf{x}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2, \tag{E.22}$$

Inequality (E.18) directly follows from (E.21) and (E.22).

Now we turn to prove (E.19). From [9, eq. (37)], we have

$$\|\breve{\mathbf{q}}_{\bar{\boldsymbol{\Omega}}^j}\|_2 \geq \|\tilde{\mathbf{x}}_{\bar{\boldsymbol{\Omega}}^j}\|_2 - \|\tilde{\mathbf{z}}\|_2. \tag{E.23}$$

Also, since $T\backslash \boldsymbol{\Omega}^j = T\backslash(\breve{\boldsymbol{\Omega}}^j \backslash \bar{\boldsymbol{\Omega}}^j) = (T\backslash \breve{\boldsymbol{\Omega}}^j)\cup(T\cap\bar{\boldsymbol{\Omega}}^j)$, where $T\backslash \breve{\boldsymbol{\Omega}}^j$ and $T\cap\bar{\boldsymbol{\Omega}}^j$ are disjoint, we obtain $\tilde{\mathbf{x}}_{T\backslash \boldsymbol{\Omega}^j} = \tilde{\mathbf{x}}_{T\backslash \breve{\boldsymbol{\Omega}}^j} + \tilde{\mathbf{x}}_{T\cap\bar{\boldsymbol{\Omega}}^j}$, and

$$\begin{aligned}\|\tilde{\mathbf{x}}_{T\backslash \boldsymbol{\Omega}^j}\|_2^2 &= \|\tilde{\mathbf{x}}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2^2 + \|\tilde{\mathbf{x}}_{T\cap\bar{\boldsymbol{\Omega}}^j}\|_2^2 \\ &\overset{(a)}{=} \|\tilde{\mathbf{x}}_{T\backslash \breve{\boldsymbol{\Omega}}^j}\|_2^2 + \|\tilde{\mathbf{x}}_{\bar{\boldsymbol{\Omega}}^j}\|_2^2,\end{aligned} \tag{E.24}$$

where (a) follows because $\mathbf{x}$ is supported on T. The assertion (E.19) thus holds by combining (E.22), (E.23) and (E.24). □

F. Proof of Theorem 3.8

In what follows, assume that the SP algorithm terminates after l iterations.[4] We first note that

$$\begin{aligned}\mathbf{P}_{\boldsymbol{\Omega}^l}\mathbf{y} &= \mathbf{P}_{\boldsymbol{\Omega}^l}[\boldsymbol{\Phi}(\tilde{\mathbf{x}}_{\boldsymbol{\Omega}^l} + \tilde{\mathbf{x}}_{T\backslash \boldsymbol{\Omega}^l}) + \mathbf{w}] \\ &= \boldsymbol{\Phi}_{\boldsymbol{\Omega}^l}\mathbf{x}_{\boldsymbol{\Omega}^l} + \underbrace{\mathbf{P}_{\boldsymbol{\Omega}^l}\boldsymbol{\Phi}_{T\backslash \boldsymbol{\Omega}^l}\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}}_{\triangleq \boldsymbol{\Phi}_{\boldsymbol{\Omega}^l}\mathbf{f}} + \underbrace{\mathbf{P}_{\boldsymbol{\Omega}^l}\mathbf{w}}_{\triangleq \boldsymbol{\Phi}_{\boldsymbol{\Omega}^l}\mathbf{g}} \\ &= \boldsymbol{\Phi}_{\boldsymbol{\Omega}^l}\underbrace{(\mathbf{x}_{\boldsymbol{\Omega}^l} + \mathbf{f} + \mathbf{g})}_{=\mathbf{q}_{\boldsymbol{\Omega}^l} \text{ in step 3.7}}.\end{aligned} \tag{F.1}$$

Since $\hat{\mathbf{x}} = \mathbf{q}$ is supported by $\boldsymbol{\Omega}^l$ (see Step 4 of SP), the reconstruction error is

$$\begin{aligned}\|\mathbf{x} - \widehat{\mathbf{x}}\|_2 &= \Bigg\|\underbrace{\tilde{\mathbf{x}}_{T\backslash \boldsymbol{\Omega}^l} + \tilde{\mathbf{x}}_{\boldsymbol{\Omega}^l}}_{=\mathbf{x}} - \tilde{\mathbf{q}}_{\boldsymbol{\Omega}^l}\Bigg\|_2 \\ &\leq \|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\|_2 + \|\mathbf{x}_{\Omega^l} - \mathbf{q}_{\boldsymbol{\Omega}^l}\|_2 \\ &\overset{(a)}{\leq} \|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\|_2 + \|\mathbf{f}\|_2 + \|\mathbf{g}\|_2,\end{aligned} \tag{F.2}$$

where (a) follows since $\mathbf{q}_{\Omega^l} - \mathbf{x}_{\boldsymbol{\Omega}^l} = \mathbf{f} + \mathbf{g}$ (cf. (F.1)) and using the triangular inequality. As in the proof of [9, Th. 9], our proof relies on deriving upper bounds of $\left\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\right\|_2$, $\|\mathbf{f}\|_2$, and $\|\mathbf{g}\|_2$ in terms of $\|\mathbf{w}\|_2$ under the given RIC requirement (3.15). By exploiting the near-orthogonality condition (2.1), our derived bounds are tighter, and thus improve the achievable performance results. By means of Lemma A.1, an upper bound for $\|\mathbf{g}\|_2$ can be immediately obtained based on the following:

$$\sqrt{1-\delta_{3K}}\,\|\mathbf{g}\|_2 \leq \left\|\boldsymbol{\Phi}_{\boldsymbol{\Omega}^l}\mathbf{g}\right\|_2 = \left\|\mathbf{P}_{\boldsymbol{\Omega}^l}\mathbf{w}\right\|_2 \leq \|\mathbf{w}\|_2. \tag{F.3}$$

Also, using the same technique in deriving (E.15)$\sim$(E.16), it follows that

$$\|\mathbf{f}\|_2 \leq \frac{\delta_{3K}\sqrt{1+\delta_{3K}}}{\sqrt{1-\delta_{3K}}}\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\|_2. \tag{F.4}$$

It then remains to derive an upper bound of $\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\|_2$ to complete the proof. Toward this end, we need the following lemma, which improves the bound provided in eq. (17) in [9, Th. 10].

Lemma A.7: For α and β defined in, respectively, (3.17) and (3.18), we have

$$\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^j}\|_2 \leq \alpha\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^{j-1}}\|_2 + \beta\|\mathbf{w}\|_2. \tag{F.5}$$

Proof: The proof is relegated to the end of this appendix. □

Based on Lemma A.7, an upper bound on $\|\mathbf{x}_{T\backslash \boldsymbol{\Omega}^l}\|_2$ is given in the next lemma.

[4]Since there are totally $C_K^N = \frac{N!}{K!(N-K)!}$ candidate supports, the SP algorithm with the stopping criterion $\|\mathbf{r}^j\|_2 \geq \|\mathbf{r}^{j-1}\|_2$ halts after at most $C_K^N + 1$ iterations. This is because $\Omega^{C_K^N+1}$, the support identified in the (C_K^N+1)-th iteration, must be the same as Ω^{j_0}, the support identified in the j_0-th iteration for some $1 \leq j_0 \leq C_K^N$. As a result, in the worst case we have $\|\mathbf{r}^{C_K^N+1}\|_2 = \|\mathbf{r}^{j_0}\|_2 > \|\mathbf{r}^{j_0+1}\|_2 > \cdots > \|\mathbf{r}^{C_K^N}\|_2$.

Lemma A.8: Under the assumptions as in Theorem 3.8, it follows that

$$\|\mathbf{x}_{T\setminus\mathbf{\Omega}^l}\|_2 \leq \frac{(2+\sqrt{1+\delta_{3K}}\beta)\|\mathbf{w}\|_2}{\sqrt{1-\delta_{3K}}-\sqrt{1+\delta_{3K}}\alpha}, \tag{F.6}$$

where α and β are defined in (3.17) and (3.18).

Proof: The proof is relegated to the end of this appendix. □

Using Lemma A.8, (F.2), (F.3) and (F.4) together with some straightforward manipulations, the assertion in Theorem 3.8 thus follows. □

Proof of Lemma A.7: The proof procedures are similar to the proof of [9, Th. 10], which utilizes [9, eq. (15) and (16)] for obtaining the reconstruction error given in [9, eq. (17)]. By exploiting the developed approximate orthogonality condition (2.1), we can obtain the following two upper bounds (which improves the bounds given in [9, eq. (15) and (16)]:

$$\|\mathbf{x}_{T\setminus\breve{\mathbf{\Omega}}^j}\|_2 \leq \frac{2\delta_{3K}}{1-\delta_{3K}}\sqrt{1+\delta_{3K}^2\frac{1+\delta_{3K}}{1-\delta_{3K}}}\|\mathbf{x}_{T\setminus\mathbf{\Omega}^{j-1}}\|_2 + \frac{2\sqrt{1+\delta_{3K}}}{1-\delta_{3K}}\|\mathbf{w}\|_2, \tag{F.7}$$

and

$$\|\mathbf{x}_{T\setminus\mathbf{\Omega}^j}\|_2 \leq \sqrt{1+\frac{4\delta_{3K}^2(1+\delta_{3K})}{1-\delta_{3K}}}\|\mathbf{x}_{T\setminus\breve{\mathbf{\Omega}}^j}\|_2 + \frac{2}{\sqrt{1-\delta_{3K}}}\|\mathbf{w}\|_2. \tag{F.8}$$

Inequality (F.5) follows immediately by substituting the upper bound on $\left\|\mathbf{x}_{T\setminus\breve{\mathbf{\Omega}}^j}\right\|_2$ given in (F.7) into (F.8) together with some straightforward manipulations. The derivations of (F.7) and (F.8) are similar to, respectively, the derivations of (E.2) in Lemma A.4 and (E.3) in Lemma A.5, except that the noise effect is taken into account. The details are thus omitted, and are referred to the supplementary result [31]. □

Proof of Lemma A.8: We have

$$\begin{aligned}\|\mathbf{r}^j\|_2 &\overset{(a)}{\leq} \sqrt{1+\delta_{3K}}\|\mathbf{x}_{T\setminus\mathbf{\Omega}^j}\|_2 + \|\mathbf{w}\|_2 \\ &\overset{(b)}{\leq} \sqrt{1+\delta_{3K}}\alpha\|\mathbf{x}_{T\setminus\mathbf{\Omega}^{j-1}}\|_2 \\ &\quad + (1+\sqrt{1+\delta_{3K}}\beta)\|\mathbf{w}\|_2,\end{aligned} \tag{F.9}$$

where (a) is from [9, eq. (19)] and (b) follows from Lemma A.7. Note that $\mathbf{r}^j$ can be expressed as $\mathbf{r}^j = \mathbf{y} - \mathbf{\Phi}_{\mathbf{\Omega}^j}\mathbf{q}_{\mathbf{\Omega}^j} = \mathbf{\Phi}\mathbf{x} + \mathbf{w} - \mathbf{\Phi}\tilde{\mathbf{q}}_{\mathbf{\Omega}^j}$ (see Step 3.8 of SP). We then obtain

$$\begin{aligned}\|\mathbf{r}^{j-1}\|_2 &= \|\mathbf{\Phi}(\tilde{\mathbf{x}}_{T\setminus\mathbf{\Omega}^{j-1}} + \tilde{\mathbf{x}}_{\mathbf{\Omega}^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}}) + \mathbf{w}\|_2 \\ &\geq \|\mathbf{\Phi}(\tilde{\mathbf{x}}_{T\setminus\mathbf{\Omega}^{j-1}} + \tilde{\mathbf{x}}_{\mathbf{\Omega}^{j-1}} - \tilde{\mathbf{q}}_{\mathbf{\Omega}^{j-1}})\|_2 - \|\mathbf{w}\|_2 \\ &\overset{(a)}{\geq} \sqrt{1-\delta_{3K}}\|\mathbf{x}_{T\setminus\mathbf{\Omega}^{j-1}}\|_2 - \|\mathbf{w}\|_2,\end{aligned} \tag{F.10}$$

where (a) follows using the same technique in deriving (E.7). By assumption, the SP algorithm terminates after l iterations, i.e., $\|\mathbf{r}^l\|_2 \geq \|\mathbf{r}^{l-1}\|_2$, which together with (F.9) and (F.10) implies

$$(\sqrt{1-\delta_{3K}} - \sqrt{1+\delta_{3K}}\alpha)\|\mathbf{x}_{T\setminus\mathbf{\Omega}^l}\|_2 \leq (2+\sqrt{1+\delta_{3K}}\beta)\|\mathbf{w}\|_2. \tag{F.11}$$

Since $(\sqrt{1-\delta_{3K}} - \sqrt{1+\delta_{3K}}\alpha) > 0$ as $\delta_{3K} \leq 0.2412$ (see (E.8)), the assertion of Lemma A.8 directly follows from (F.11). □

Acknowledgment

The authors thank the two reviewers, whose constructive suggestions and comments significantly improve this paper.

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

[4] M. Elad, *Sparse and Redundant Representation: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer-Verlag, 2010.

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

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

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

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

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

[10] R. Giryes and M. Elad, "RIP-based near-oracle performance guarantees for SP, CoSaMP, and IHT," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1465–1568, Mar. 2012.

[11] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030–5043, Oct. 2010.

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

[13] H. Rauhut, "On the impossibility of uniform sparse reconstruction using greedy methods," *Sampling Theory Signal Image Process.*, vol. 7, no. 2, pp. 197–215, 2008.

[14] H. Rahut, "Stability results for random sampling of sparse trigonometric polynomials," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5661–5670, Dec. 2008.

[15] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–124, Jul. 2007.

[16] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[17] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comp. Rendus Math.*, vol. 346, nos. 9–10, pp. 589–592, May 2008.

[18] 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.

[19] T. 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, Jul. 2011.

[20] M. A. Davenport, P. T. Boufounos, and R. G. Baraniuk, "Compressive domain interference cancellation," in *Proc. Workshop Signal Process. Adapt. Sparse Struct. Represent. (SPARS)*, Saint-Malo, France, 2009.

[21] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[22] C. Stuber, P. Kuppinger, G. Pope, and H. Bolcskei, "Recovery of sparsely corrupted signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3115–3130, May 2012.

[23] L.-H. Chang and J.-Y. Wu, "Compressive-domain interference cancellation via orthogonal projection: How small the restricted isometry constant of the effective sensing matrix can be?" in *Proc. IEEE Wireless Commun. Netw. Conf. (IEEE WCNC)*, Paris, France, Apr. 2012, pp. 256–261.

[24] L.-H. Chang and J.-Y. Wu, "Achievable angles between two compressed sparse vectors under norm/distance constraints imposed by the restricted isometry property: A plane geometry approach," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2059–2081, Apr. 2013.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[26] M. E. Davies and R. Gribonval, "Restricted isometry constants where ℓ^p sparse recovery can fail for $0 \ll p \leq 1$," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2203–2214, May 2009.

[27] M. Elad, "Sparse and redundant representation modeling—What next?" *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 922–928, Dec. 2012.

[28] 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, Jan. 2006.

[29] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.

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

[31] L. H. Chang and J. Y. Wu. "An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit." [Online]. Available: http://arxiv.org/abs/1401.0578

Ling-Hua Chang received the B. S. degree from the Department of Electrical and Computer Engineering, National Chiao Tung University, Taiwan, in 2010. She is currently a Ph. D. student in the Institute of Communications Engineering, National Chiao Tung University, Taiwan. Her research interests include signal processing and linear algebra.

Jwo-Yuh Wu received the B. S. degree in 1996, the M. S. degree in 1998, and the Ph. D. degree in 2002, all in Electrical and Control Engineering, National Chiao Tung University, Taiwan. During 2003 and 2007, he was a post doctor research fellow in the Department of Communications Engineering, National Chiao Tung University, Taiwan. Starting from 2008, he has been a faculty member in the Department of Electrical and Computer Engineering, and the Institute of Communications Engineering, National Chiao Tung University, Taiwan, where he has been an associate professor since 2011. His general research interests are in signal processing, wireless communications, control systems, linear algebra, and applied functional analysis. Dr. Wu received the 2013 Y. Z. Hsu Scientific Paper Award (Communication Technology Category), and 2013 Best Paper Award for Young Scholars by IEEE Information Theory Society Taipei Chapter and IEEE Communications Society Taipei/Tainan Chapter.