SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2018, Vol. 61 102303:1–102303:14
https://doi.org/10.1007/s11432-017-9289-2

Improved RIP-based performance guarantees for multipath matching pursuit

Juan ZHAO[1,2], Xia BAI[1,2*] & Ran TAO[1,2]

[1]*School of Information and Electronics, Beijing Institute of Technology, Beijing* 100081*, China;*
[2]*Beijing Key Laboratory of Fractional Signals and Systems, Beijing Institute of Technology, Beijing* 100081*, China*

Received 10 July 2017/Revised 13 September 2017/Accepted 20 October 2017/Published online 22 May 2018

Abstract The multipath matching pursuit (MMP) is a generalization of the orthogonal matching pursuit (OMP), which generates multiple child paths for every candidate in each iteration and selects the candidate having the minimal residual as the final support set when iteration ends. In this paper we analyze its performance in both noiseless and noisy cases. The restricted isometry property (RIP)-based condition of MMP that ensures accurate recovery of sparse signals in the noiseless case is derived by using a simple technique. The performance guarantees of the MMP for support recovery in noisy cases are also discussed. It is shown that under certain conditions on the RIP and minimum magnitude of nonzero components of the sparse signal, the MMP will exactly recover the true support of the sparse signal in cases of bounded noises and recover the true support with a high probability in the case of Gaussian noise. Our bounds on the RIP improve the existing results.

Keywords compressed sensing (CS), sparse signal reconstruction, multipath matching pursuit (MMP), support recovery, restricted isometry property (RIP)

Citation Zhao J, Bai X, Tao R. Improved RIP-based performance guarantees for multipath matching pursuit. Sci China Inf Sci, 2018, 61(10): 102303, https://doi.org/10.1007/s11432-017-9289-2

1 Introduction

Compressed sensing (CS) is a novel sampling paradigm in signal processing that guarantees the reconstruction of a sparse or compressible signal with a lower sampling rate than the Nyquist rate [1]. It can significantly reduce the amount of data that must be stored and transmitted, and has been successfully applied to radar, communications, and other fields [2–5].

Suppose that $\boldsymbol{x}$ is a N-length signal and is said to be K-sparse if $\boldsymbol{x}$ has only K nonzero entries or significant coefficients with $K \ll N$. The support of $\boldsymbol{x}$ is denoted by $T = \mathrm{supp}(\boldsymbol{x}) = \{i|\, x_i \neq 0, i = 1, 2, \ldots, N\}$, where x_i is the i-th element of $\boldsymbol{x}$. The compressed linear measurements are expressed as $\boldsymbol{y} = \boldsymbol{\Phi x}$, where $\boldsymbol{y} \in \mathbb{R}^M$ is the measurement signal and $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ $(M < N)$ is the sensing matrix. Let $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_N]$, where ϕ_i is column i of $\boldsymbol{\Phi}$. Because $M < N$, reconstruction of $\boldsymbol{x}$ from $\boldsymbol{y}$ is generally ill-posed. According to CS theory, such a sparse signal $\boldsymbol{x}$ can be accurately reconstructed via properly designed recovery algorithms by solving a l_0 norm optimization problem $\min_{\boldsymbol{x}} ||\boldsymbol{x}||_0$, s.t. $\boldsymbol{y} = \boldsymbol{\Phi x}$, where $||\cdot||_0$ is a measure of sparsity counting the number of nonzero entries. To guarantee the success of the recovery algorithms, the restricted isometry property (RIP) condition of $\boldsymbol{\Phi}$ is required [6]. The sensing matrix $\boldsymbol{\Phi}$ is said to satisfy the RIP of order K if there exists a smallest constant $\delta_K \in (0, 1)$ such that

$$(1-\delta_K)\,\|\boldsymbol{x}\|_2^2 \leqslant \|\boldsymbol{\Phi x}\|_2^2 \leqslant (1+\delta_K)\,\|\boldsymbol{x}\|_2^2 \tag{1}$$

* Corresponding author (email: bai@bit.edu.cn)

holds for any K-sparse $\boldsymbol{x}$. The constant δ_K is called the order-k restricted isometry constant (RIC).

Greedy recovery algorithms have received significant attention because of their good performance and simplicity. They recover the supports of sparse signals in an iterative fashion. The orthogonal matching pursuit (OMP) [7–9] is the classical greedy algorithm. In each iteration, it chooses one column of $\boldsymbol{\Phi}$ (also called atom) that has the largest correlation with the current residual. To improve the performance of the OMP, many greedy algorithms have been proposed, including regularized OMP (ROMP) [10], stagewise OMP (StOMP) [11], compressive sampling matching pursuit (CoSaMP) [12], subspace pursuit (SP) [13], generalized OMP (GOMP) [14–16], blind sparsity weak SP (BSWSP) [17]. In contrast to the above greedy algorithms that output a single candidate for the support set, some tree-search-based greedy algorithms have been proposed, which use tree-search structures to produce multiple promising candidates for the support set to improve the chance of selecting the true support [18–21]. Among these algorithms, a modified OMP called multipath matching pursuit (MMP) [21] exploits the structure of a combinatoric tree to generate L child paths for every candidate in each iteration (see Figure 1). The MMP selects the candidate having the minimal residual as the final support set when iteration ends. It has shown that the MMP can effectively reconstruct sparse signals and it has been used for sparse detection in communications and large MIMO systems [22, 23].

The RIP-based performance guarantees for the MMP have been discussed in [21] and it is proved that the MMP can exactly recover K-sparse signals if $\boldsymbol{\Phi}$ satisfies the RIP with the RIC $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+2\sqrt{L}}$. However, this bound is overly restrictive and is not consistent (for $L = 1$) with the upper bound $\frac{1}{\sqrt{K+1}}$ of δ_{K+1} in the OMP [8]. In this paper, an improved bound $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$ is derived by using a simpler technique than that proposed in [21]. Furthermore, in the presence of l_2 bounded noise, by more straightforward analysis, we derive a looser RIP-based sufficient condition on the minimum magnitude of the nonzero components of the sparse signal to guarantee exact support recovery of the sparse signal via MMP. The performances of the MMP in cases of Gaussian noise and l_∞ bounded noise are also considered.

The remainder of this paper is organized as follows. In Section 2, the MMP algorithm is described in detail and some properties of the RIP are presented. In Section 3, we derive the RIP-based condition of the MMP that ensures the accurate recovery of sparse signals in the noiseless case and then derive the RIP-based condition under which the support of sparse signals can be exactly recovered by the MMP in noisy cases. Finally, our conclusion is summarized in Section 4.

2 The MMP algorithm and some properties of RIP

2.1 Review of the MMP algorithm

The MMP algorithm is a generalization of the OMP algorithm, which produces multiple promising candidates by constructing the combinatoric tree structure illustrated in Figure 1, where each candidate will generate L child candidates in each iteration. For the sake of uniformity, we follow the same notations as used in [21]. Let $S^k = \left\{s_1^k, s_2^k, \ldots, s_u^k\right\}$ be the set of candidates in the k-th iteration and s_i^k denote the i-th candidate in the k-th iteration. At the $(k-1)$-th iteration, for the i-th path (candidate s_i^{k-1}), $\boldsymbol{r}_i^{k-1}$ denotes the residual corresponding to s_i^{k-1}. L indices $\tilde{\pi}_1, \tilde{\pi}_2, \ldots, \tilde{\pi}_L$ whose columns are mostly correlated with the residual $\boldsymbol{r}_i^{k-1}$ will be selected and added to the candidate s_i^{k-1}. L child candidates will then be generated. If the child candidate s_u^k does not overlap with the existing candidates, the estimated vector $\hat{\boldsymbol{x}}_u^k$ and residual $\boldsymbol{r}_u^k$ will be calculated. After K iterations, the candidate having the minimum residual will be chosen (i.e., $\hat{s} = s_{\bar{u}}^K$, where $\bar{u} = \arg\min_u \|\boldsymbol{r}_u^K\|_2^2$). The details of the MMP algorithm are presented in Algorithm 1.

In the MMP algorithm, $|\cdot|$ denotes cardinality and $*$ refers to the transpose. $\boldsymbol{\Phi}_\Lambda \in \mathbb{R}^{M\times|\Lambda|}$ represents the submatrix of $\boldsymbol{\Phi}$ with columns indexed by Λ. $\boldsymbol{\Phi}_\Lambda^\dagger$ denotes the Moore-Penrose pseudo-inverse of $\boldsymbol{\Phi}_\Lambda$ and $\boldsymbol{\Phi}_\Lambda^\dagger = (\boldsymbol{\Phi}_\Lambda^*\boldsymbol{\Phi}_\Lambda)^{-1}\boldsymbol{\Phi}_\Lambda^*$ if $\boldsymbol{\Phi}_\Lambda$ has full column rank. Similarly, $\boldsymbol{x}_\Lambda$ represents the subvector of $\boldsymbol{x}$ containing the entries indexed by Λ.

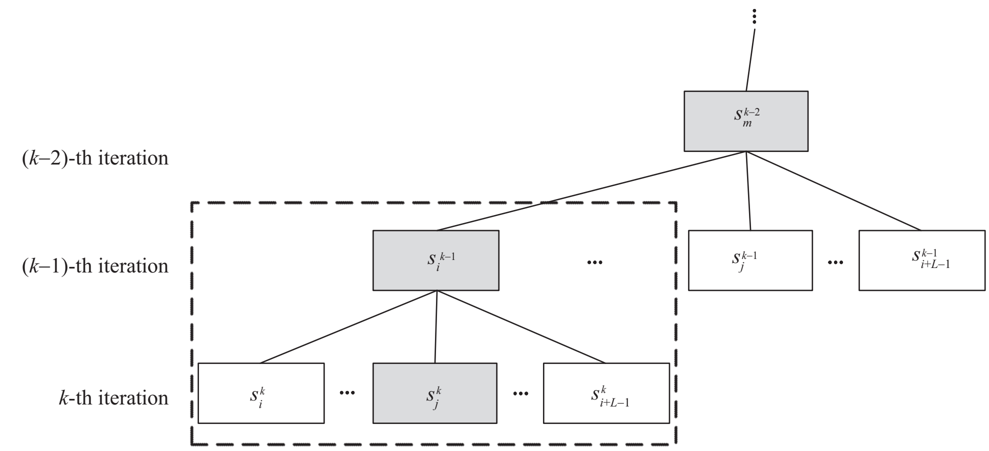

Figure 1 Relationship between the candidates in different iterations of the MMP.

Algorithm 1 MMP algorithm [21]

Input: measurement $\boldsymbol{y}$, sensing matrix $\boldsymbol{\Phi}$, sparsity K, number of path L;
  Initialization: $k=0$ (iteration index), $\boldsymbol{r}^0=\boldsymbol{y}$ (initial residual), $S^0=\{\emptyset\}$;
  while $k<K$ **do**
    $k=k+1$, $u=0$, $S^k=\emptyset$;
    for $i=1$ to $|S^{k-1}|$ **do**
      $\{\tilde{\pi}_1,\ldots,\tilde{\pi}_L\}=\arg\max_{|\pi|=L}\|(\boldsymbol{\Phi}^*\boldsymbol{r}_i^{k-1})_\pi\|_2^2$; (choose L best indices)
      for $j=1$ to L **do**
        $s_{\text{temp}}=s_i^{k-1}\bigcup\{\tilde{\pi}_j\}$; (construct a temporary path)
        if $s_{\text{temp}}\notin S^k$ **then** (check if the path already exists)
          $u=u+1$; (candidate index update)
          $s_u^k=s_{\text{temp}}$; (path update)
          $S^k=S^k\bigcup\{s_u^k\}$; (update the set of paths)
          $\hat{\boldsymbol{x}}_u^k=\boldsymbol{\Phi}_{s_u^k}^{\dagger}\boldsymbol{y}$; (perform estimation)
          $\boldsymbol{r}_u^k=\boldsymbol{y}-\boldsymbol{\Phi}_{s_u^k}\hat{\boldsymbol{x}}_u^k$; (residual update)
        endif
      end for
    end for
  end while
  $\bar{u}=\arg\min_u\|\boldsymbol{r}_u^K\|_2^2$; (find index of the best candidate)
Output: estimated support set $\hat{s}=s_{\bar{u}}^K$ and estimated signal

$$\hat{\boldsymbol{x}}=\begin{cases}\boldsymbol{\Phi}_{\hat{s}}^{\dagger}\boldsymbol{y}, & \text{on the support set } \hat{s},\\ \boldsymbol{0}, & \text{elsewhere.}\end{cases}$$

The MMP algorithm generally chooses the candidate having the minimum residual as the estimated support set after K iterations in cases with non-noise, l_2 bounded noise, and Gaussian noise. In the case of l_∞ bounded noise (i.e., $\boldsymbol{y}=\boldsymbol{\Phi}\boldsymbol{x}+\boldsymbol{v}$, where $\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty\leqslant\varepsilon$), the MMP will choose the candidate $\hat{s}=s_{\tilde{u}}^K$ that satisfies $\tilde{u}=\arg\min_u\|\boldsymbol{\Phi}^*\boldsymbol{r}_u^K\|_\infty$ as the estimated support set when iteration ends.

2.2 Properties of RIP

Some notations used in this subsection are presented below. Let $I_1-I_2=I_1\backslash(I_1\bigcap I_2)$ denote the set with elements contained in I_1, but not in I_2. $\Omega=\{1,2,\ldots,N\}$ represents the column indices of matrix $\boldsymbol{\Phi}$. $\boldsymbol{P}_\Lambda=\boldsymbol{\Phi}_\Lambda\boldsymbol{\Phi}_\Lambda^{\dagger}$ is the projection onto span$(\boldsymbol{\Phi}_\Lambda)$ and $\boldsymbol{P}_\Lambda^{\perp}=\boldsymbol{I}-\boldsymbol{P}_\Lambda$ is the orthogonal complement of span$(\boldsymbol{\Phi}_\Lambda)$.

The following lemmas related to the RIP properties will be used in Section 3.

Lemma 1 ([6])**.** Suppose that the sensing matrix $\boldsymbol{\Phi}$ satisfies the RIP with orders K_1 and K_2. If $K_1\leqslant K_2$, then $\delta_{K_1}\leqslant\delta_{K_2}$.

Lemma 2 ([6])**.** For $I \subset \Omega$, if $\delta_{|I|} < 1$, then for any $\boldsymbol{x} \in \mathbb{R}^{|I|}$,

$$(1-\delta_{|I|})\|\boldsymbol{x}\|_2 \leqslant \|\boldsymbol{\Phi}_I^*\boldsymbol{\Phi}_I\boldsymbol{x}\|_2 \leqslant (1+\delta_{|I|})\|\boldsymbol{x}\|_2, \tag{2}$$

$$\frac{1}{1+\delta_{|I|}}\|\boldsymbol{x}\|_2 \leqslant \left\|(\boldsymbol{\Phi}_I^*\boldsymbol{\Phi}_I)^{-1}\boldsymbol{x}\right\|_2 \leqslant \frac{1}{1-\delta_{|I|}}\|\boldsymbol{x}\|_2. \tag{3}$$

Lemma 3 ([24])**.** Let $I_1, I_2 \subset \Omega$ be two disjoint sets ($I_1 \bigcap I_2 = \emptyset$). If $\delta_{|I_1|+|I_2|} < 1$, then

$$\left\|\boldsymbol{\Phi}_{I_1}^*\boldsymbol{\Phi}_{I_2}\boldsymbol{x}\right\|_2 \leqslant \delta_{|I_1|+|I_2|}\|\boldsymbol{x}\|_2. \tag{4}$$

Lemma 4 ([21])**.** For an $M \times N$ matrix $\boldsymbol{\Phi}$, $\|\boldsymbol{\Phi}\|_2$ satisfies

$$||\boldsymbol{\Phi}||_2 = \sqrt{\lambda_{\max}(\boldsymbol{\Phi}^*\boldsymbol{\Phi})} \leqslant \sqrt{1+\delta_{\min(M,N)}}, \tag{5}$$

where $\lambda_{\max}$ denotes the maximum eigenvalue.

Lemma 5 (Lemma 1 in [16])**.** For any $\boldsymbol{x} \in \mathbb{R}^{|\Gamma-\Lambda|}$, the following inequalities hold:

$$(1-\delta_{|\Gamma\cup\Lambda|})\|\boldsymbol{x}\|_2^2 \leqslant \left\|\boldsymbol{P}_\Lambda^\perp\boldsymbol{\Phi}_{\Gamma-\Lambda}\boldsymbol{x}\right\|_2^2 \leqslant (1+\delta_{|\Gamma\bigcup\Lambda|})\|\boldsymbol{x}\|_2^2. \tag{6}$$

3 Main results

3.1 Exact signal recovery condition for the noiseless case

Consider the model $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$, it has been proven that the MMP can exactly recover any K-sparse signal if the RIC satisfies $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+2\sqrt{L}}$ [21]. Here, we derive an improved condition of the MMP for exact signal recovery with a simpler technique than that proposed in [21].

Theorem 1. For the measurements $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x}$, where $\boldsymbol{x}$ is a K-sparse vector, if the sensing matrix $\boldsymbol{\Phi}$ satisfies the RIP with

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}, \tag{7}$$

then the MMP will perfectly recover the K-sparse signal $\boldsymbol{x}$.

Proof. The K-sparse signal $\boldsymbol{x}$ can be accurately recovered if the support of the sparse signal is recovered exactly via MMP. To guarantee that the MMP can find the true support in K iterations, it is required that at least one candidate chooses the correct index in the true support set T in every iteration, which can be proved by mathematical induction.

Suppose that the MMP performs $k-1$ iterations successfully, that is, there exists a candidate s_i^{k-1} containing $k-1$ correct indices, $s_i^{k-1} \subset T$ and $|s_i^{k-1}| = k-1$ (see Figure 1). This assumption naturally holds when $k=1$. We need to find the condition under which at least one of L indices is chosen from T in the k-th iteration.

Let β_1^k denote the largest magnitude of correlation between $\boldsymbol{r}_i^{k-1}$ and the columns whose indices are in $T - s_i^{k-1}$ (the set of remaining correct indices), and α_L^k denote the L-th largest magnitude of correlation between $\boldsymbol{r}_i^{k-1}$ and the columns whose indices are in the complement set $T^C = \{i|\, x_i = 0, i = 1,2,\ldots,N\}$ (the set of incorrect indices).

According to the definitions of β_1^k and α_L^k, it is clear that in the k-th iteration, at least one of L indices is chosen from T if the following condition holds:

$$\beta_1^k > \alpha_L^k. \tag{8}$$

Now it needs to derive a lower bound of β_1^k and an upper bound of α_L^k.

From $\boldsymbol{y} = \boldsymbol{\Phi x}$, similar to [15], the residual $\boldsymbol{r}_i^{k-1}$ can be written as

$$\begin{aligned}\boldsymbol{r}_i^{k-1} &= \boldsymbol{y} - \boldsymbol{\Phi}_{s_i^{k-1}}\hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}_T\boldsymbol{x}_T - \boldsymbol{\Phi}_{s_i^{k-1}}\hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}_{s_i^{k-1}}\boldsymbol{x}_{s_i^{k-1}} + \boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}} - \boldsymbol{\Phi}_{s_i^{k-1}}\hat{\boldsymbol{x}}_{s_i^{k-1}} \\ &= \boldsymbol{\Phi}_{s_i^{k-1}}\left(\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}}\right) + \boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}} = \left[\boldsymbol{\Phi}_{s_i^{k-1}}\boldsymbol{\Phi}_{T-s_i^{k-1}}\right]\begin{pmatrix}\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} \\ \boldsymbol{x}_{T-s_i^{k-1}}\end{pmatrix} \\ &= \boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}, \end{aligned} \tag{9}$$

where

$$\tilde{\boldsymbol{x}}_i^{k-1} = \begin{pmatrix}\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} \\ \boldsymbol{x}_{T-s_i^{k-1}}\end{pmatrix},$$

and $\hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}^\dagger_{s_i^{k-1}}\boldsymbol{y} = (\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}})^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{y}$.

Noting that $|s_i^{k-1}| = k-1$, then $|T - s_i^{k-1}| = K-(k-1)$ and we have

$$\begin{aligned}\beta_1^k &= \max_{j\in T-s_i^{k-1}} \left|\boldsymbol{\phi}_j^*\boldsymbol{r}_i^{k-1}\right| \overset{(a)}{\geqslant} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}^*_{T-s_i^{k-1}}\boldsymbol{r}_i^{k-1}\right\|_2 \overset{(b)}{=} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}_T^*\boldsymbol{r}_i^{k-1}\right\|_2 \\ &\overset{(c)}{=} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 \overset{(d)}{\geqslant} \frac{1}{\sqrt{K-(k-1)}}(1-\delta_K)\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 \\ &\overset{(e)}{\geqslant} \frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2, \end{aligned} \tag{10}$$

where (a) follows from $\|\boldsymbol{z}\|_\infty \geqslant \frac{1}{\sqrt{\|\boldsymbol{z}\|_0}}\|\boldsymbol{z}\|_2$, (b) follows from $\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{r}_i^{k-1} = \boldsymbol{0}$, and (c), (d), and (e) follow from (9), (2), and Lemma 1, respectively.

Let α_j^k denote the j-th largest magnitude of correlation between $\boldsymbol{r}_i^{k-1}$ and $\{\boldsymbol{\phi}_j, j \in T^C\}$. That is, $\alpha_j^k = |\langle\boldsymbol{\phi}_{f_j}, \boldsymbol{r}_i^{k-1}\rangle|$, where the corresponding index $f_j = \arg\max_{u\in T^C\setminus\{f_1,\ldots,f_{j-1}\}}|\langle\boldsymbol{\phi}_u, \boldsymbol{r}_i^{k-1}\rangle|$. Then, we have $\alpha_1^k \geqslant \alpha_2^k \geqslant \cdots \geqslant \alpha_L^k$. By defining $W = \{f_1, f_2, \ldots, f_L\}$, we can obtain the upper bound of α_L^k as follows:

$$\begin{aligned}\alpha_L^k &\leqslant \frac{1}{L}\sum_{j=1}^L \alpha_j^k = \frac{1}{L}\sum_{j=1}^L\left|\langle\boldsymbol{\phi}_{f_j}, \boldsymbol{r}_i^{k-1}\rangle\right| = \frac{1}{L}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{r}_i^{k-1}\right\|_1 \\ &\overset{(f)}{\leqslant} \frac{1}{\sqrt{L}}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{r}_i^{k-1}\right\|_2 \overset{(g)}{=} \frac{1}{\sqrt{L}}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 \overset{(h)}{\leqslant} \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2, \end{aligned} \tag{11}$$

where (f) follows from the norm inequality $\|\boldsymbol{z}\|_1 \leqslant \sqrt{\|\boldsymbol{z}\|_0}\|\boldsymbol{z}\|_2$, and (g) and (h) follow from (9) and (4), respectively.

According to (10) and (11), the sufficient condition for (8) is

$$\frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 > \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2,$$

which can be simplified as

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K-(k-1)}+\sqrt{L}}.$$

Because $\sqrt{K-k+1} < \sqrt{K}$ for $k > 1$, the above inequality holds under (7). Thus, we can obtain that (8) will hold under the condition given by (7), which guarantees that in the k-th iteration, at least one of L indices generated from s_i^{k-1} is an index in T. That is, $s_i^{k-1} \subseteq T$ will hold and the sparse signal can be perfectly recovered after K iterations. Theorem 1 is proved.

Remark 1. According to the above analysis, the key point of the proof is to derive the bounds of β_1^k and α_L^k. From Lemma 3.8 and (22) in [21], the bound of β_1^k becomes

$$\beta_1^k \geqslant \frac{1-2\delta_{K+L}}{\sqrt{K-k+1}}\frac{\|\boldsymbol{x}_{T-s_i^{k-1}}\|_2}{1-\delta_{K+L}}. \tag{12}$$

Similarly, from Lemma 3.7 and (21) in [21], it can be seen that the bound of α_L^k is changed as

$$\alpha_L^k \leqslant \frac{\delta_{K+L}}{\sqrt{L}} \frac{\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2}{1-\delta_{K+L}}. \tag{13}$$

The bounds derived in (10) and (11) are tighter than those shown in (12) and (13), respectively (refer to Appendix A), which is useful for obtaining the looser condition shown in (7).

Remark 2. From Theorem 1, it can be seen that when $L = 1$, the MMP becomes the OMP and the sufficient condition (7) is consistent with the upper bound on δ_{K+1} in the OMP algorithm [8].

3.2 Exact support recovery conditions for noisy cases

In this subsection, we first analyze the performance of the MMP in cases of l_2 bounded noise and Gaussian noise. We then proceed with l_∞ bounded noise.

3.2.1 l_2 *bounded noise*

In practice, we often encounter noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where $\boldsymbol{v}$ is a noise vector and $\|\boldsymbol{v}\|_2 \leqslant \varepsilon$. In contrast to the noiseless case, it has been pointed out in [21] that we not only need to derive the condition under which at least one candidate of the MMP is the true support set, but also require a condition ensuring that the candidate with the minimal residual is the true support. Combining these conditions, it can guarantee that the output candidate of the MMP having the minimal residual is the true support,that is, the MMP can recover the support of the sparse signal exactly.

First, consider the condition under which at least one candidate of the MMP is the true support. Let $x_{\min}$ denote the minimum magnitude of the nonzero elements of $\boldsymbol{x}$, that is, $x_{\min} = \min\{|x_j|, |x_j| \neq 0\}$. Theorem 2 shows that at least one candidate of the MMP is the true support under a looser condition than that in [21].

Theorem 2. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where $\boldsymbol{x}$ is a K-sparse signal and $\boldsymbol{v}$ denotes the noise. Suppose that the sensing matrix $\boldsymbol{\Phi}$ satisfies $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$, then at least one candidate of the MMP will recover the support set T of the sparse signal $\boldsymbol{x}$ if

$$x_{\min} > \frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+1)}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}} \|\boldsymbol{v}\|_2. \tag{14}$$

Proof. To guarantee that at least one candidate of the MMP is the true support in K iterations, it is required that at least one correct index is selected in each iteration, which can be proved by mathematical induction.

Suppose that the MMP performs $k-1$ iterations successfully, that is, there exists a candidate s_i^{k-1} containing $k-1$ correct indices, $s_i^{k-1} \subset T$ and $|s_i^{k-1}| = k-1$. This assumption naturally holds when $k = 1$. We need to prove that at least one of L indices is chosen from T under the condition (14). At the k-th iteration, Eq. (8) is still required to guarantee choosing at least one correct index. According to $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, similar to (9), the residual $\boldsymbol{r}_i^{k-1}$ can be written as

$$\begin{aligned}
\boldsymbol{r}_i^{k-1} &= \boldsymbol{y} - \boldsymbol{\Phi}_{s_i^{k-1}} \hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}_T \boldsymbol{x}_T + \boldsymbol{v} - \boldsymbol{\Phi}_{s_i^{k-1}} \hat{\boldsymbol{x}}_{s_i^{k-1}} \\
&= \boldsymbol{\Phi}_{s_i^{k-1}} \boldsymbol{x}_{s_i^{k-1}} + \boldsymbol{\Phi}_{T-s_i^{k-1}} \boldsymbol{x}_{T-s_i^{k-1}} - \boldsymbol{\Phi}_{s_i^{k-1}} \hat{\boldsymbol{x}}_{s_i^{k-1}} + \boldsymbol{v} \\
&= \boldsymbol{\Phi}_{s_i^{k-1}} \left(\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}}\right) + \boldsymbol{\Phi}_{T-s_i^{k-1}} \boldsymbol{x}_{T-s_i^{k-1}} + \boldsymbol{v} \\
&= \left[\boldsymbol{\Phi}_{s_i^{k-1}} \boldsymbol{\Phi}_{T-s_i^{k-1}}\right] \begin{pmatrix} \boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} \\ \boldsymbol{x}_{T-s_i^{k-1}} \end{pmatrix} + \boldsymbol{v} = \boldsymbol{\Phi}_T \tilde{\boldsymbol{x}}_i^{k-1} + \boldsymbol{v},
\end{aligned} \tag{15}$$

where

$$\tilde{\boldsymbol{x}}_i^{k-1} = \begin{pmatrix} \boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} \\ \boldsymbol{x}_{T-s_i^{k-1}} \end{pmatrix},$$

and $\hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}^{\dagger}_{s_i^{k-1}}\boldsymbol{y} = (\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}})^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{y}$. From similar analysis to that for the noiseless case, it needs to give a lower bound of β_1^k and an upper bound of α_L^k.

Note that $|s_i^{k-1}| = k-1$, then $|T - s_i^{k-1}| = K-(k-1)$ and we have

$$\begin{aligned}
\beta_1^k &= \max_{j\in T-s_i^{k-1}} \left|\boldsymbol{\phi}_j^*\boldsymbol{r}_i^{k-1}\right| \overset{\text{(i)}}{\geqslant} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}^*_{T-s_i^{k-1}}\boldsymbol{r}_i^{k-1}\right\|_2 \\
&\overset{\text{(j)}}{=} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}_T^*\boldsymbol{r}_i^{k-1}\right\|_2 \overset{\text{(k)}}{=} \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1} + \boldsymbol{\Phi}_T^*\boldsymbol{v}\right\|_2 \\
&\geqslant \frac{1}{\sqrt{K-(k-1)}}\left\|\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2 \\
&\overset{\text{(l)}}{\geqslant} \frac{1}{\sqrt{K-(k-1)}}(1-\delta_K)\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2,
\end{aligned} \tag{16}$$

where (i) follows from $\|\boldsymbol{z}\|_\infty \geqslant \frac{1}{\sqrt{\|\boldsymbol{z}\|_0}}\|\boldsymbol{z}\|_2$, (j) follows from $\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{r}_i^{k-1} = \boldsymbol{0}$, and (k) and (l) follow from (15) and (2), respectively. Additionally, from Lemma 4, we get

$$\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2 \leqslant \|\boldsymbol{\Phi}_T^*\|_2\|\boldsymbol{v}\|_2 \leqslant \sqrt{1+\delta_K}\|\boldsymbol{v}\|_2. \tag{17}$$

Thus, using (16), (17), and Lemma 1, a lower bound of β_1^k can be obtained as

$$\begin{aligned}
\beta_1^k &\geqslant \frac{1}{\sqrt{K-(k-1)}}(1-\delta_K)\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\sqrt{1+\delta_K}\|\boldsymbol{v}\|_2 \\
&\geqslant \frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\sqrt{1+\delta_{K+L}}\|\boldsymbol{v}\|_2.
\end{aligned} \tag{18}$$

Similar to (11), we can obtain an upper bound of α_L^k as follows:

$$\begin{aligned}
\alpha_L^k &\leqslant \frac{1}{L}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{r}_i^{k-1}\right\|_1 \overset{\text{(m)}}{\leqslant} \frac{1}{\sqrt{L}}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{r}_i^{k-1}\right\|_2 \overset{\text{(n)}}{=} \frac{1}{\sqrt{L}}\left\|\boldsymbol{\Phi}_W^*(\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}+\boldsymbol{v})\right\|_2 \\
&\leqslant \frac{1}{\sqrt{L}}\left\|\boldsymbol{\Phi}_W^*\boldsymbol{\Phi}_T\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \overset{\text{(o)}}{\leqslant} \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \\
&\overset{\text{(p)}}{\leqslant} \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\sqrt{1+\delta_L}\|\boldsymbol{v}\|_2 \leqslant \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\sqrt{1+\delta_{K+L}}\|\boldsymbol{v}\|_2,
\end{aligned} \tag{19}$$

where (m) follows from the norm inequality $\|\boldsymbol{z}\|_1 \leqslant \sqrt{\|\boldsymbol{z}\|_0}\|\boldsymbol{z}\|_2$, (n) follows from (15), and (o) and (p) follow from (4) and (17), respectively. The last inequality in (19) follows from Lemma 1.

Combining (18) and (19), to guarantee that (8) holds, we require

$$\begin{aligned}
&\frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\sqrt{1+\delta_{K+L}}\|\boldsymbol{v}\|_2 \\
&> \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\sqrt{1+\delta_{K+L}}\|\boldsymbol{v}\|_2,
\end{aligned}$$

which can be simplified as

$$\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 > \frac{\sqrt{\frac{1+\delta_{K+L}}{L}} + \sqrt{\frac{1+\delta_{K+L}}{K-(k-1)}}}{\frac{1-\delta_{K+L}}{\sqrt{K-(k-1)}} - \frac{\delta_{K+L}}{\sqrt{L}}}\|\boldsymbol{v}\|_2. \tag{20}$$

Because $\|\tilde{\boldsymbol{x}}_i^{k-1}\|_2 \geqslant \|\boldsymbol{x}_{T-s_i^{k-1}}\|_2 \geqslant \sqrt{K-(k-1)}x_{\min}$, we obtain the sufficient condition for (20) as

$$x_{\min} > \frac{\sqrt{\frac{1+\delta_{K+L}}{L}} + \sqrt{\frac{1+\delta_{K+L}}{K-(k-1)}}}{1-\delta_{K+L} - \frac{\delta_{K+L}}{\sqrt{L}}\sqrt{K-(k-1)}}\|\boldsymbol{v}\|_2.$$

Note that $k \leqslant K$, so we can obtain that (8) will hold if

$$x_{\min} > \frac{\sqrt{\frac{1+\delta_{K+L}}{L}} + \sqrt{1+\delta_{K+L}}}{1-\delta_{K+L} - \frac{\delta_{K+L}}{\sqrt{L}}\sqrt{K}} ||\boldsymbol{v}||_2,$$

which can be reduced to (14). That is, at least one correct index is chosen in the k-th iteration under the condition given by (14) (i.e., $s_j^k \subseteq T$ will hold). Theorem 2 is proved.

Remark 3. Let $\frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+1)}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}} = \zeta$, then (14) becomes $x_{\min} > \zeta\|\boldsymbol{v}\|_2$. It can be verified that

$$\mu = \frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+\sqrt{K})(1-\delta_{K+L})}{\sqrt{L}-(2\sqrt{L}+\sqrt{K})\delta_{K+L}} > \frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+\sqrt{K})}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}} \geqslant \zeta, \tag{21}$$

and

$$\zeta \geqslant \frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+\sqrt{K})}{\sqrt{LK}-(\sqrt{LK}+K)\delta_{K+L}} = \gamma. \tag{22}$$

It is shown that at least one candidate of the MMP is the true support if $x_{\min} > \max(\mu,\gamma)\|\boldsymbol{v}\|_2$ (from Theorems 4.1 and 4.2 in [21]), which can be reduced to $x_{\min} > \mu\|\boldsymbol{v}\|_2$ from (21) and (22). Thus, it can be seen that the bound given in (14) is looser than that given in [21].

Remark 4. When $L=1$, we have $|W|=1$. If the columns of $\boldsymbol{\Phi}$ have a unit l_2 norm, then $\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \leqslant \|\boldsymbol{\Phi}_W^*\|_2\|\boldsymbol{v}\|_2 \leqslant \|\boldsymbol{v}\|_2$. Thus, Eq. (19) can be relaxed to

$$\alpha_L^k \leqslant \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \leqslant \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\|\boldsymbol{v}\|_2,$$

and (14) will become

$$x_{\min} > \frac{1+\sqrt{1+\delta_{K+1}}}{1-(\sqrt{K}+1)\delta_{K+1}} ||\boldsymbol{v}||_2,$$

which is consistent with the bound on $x_{\min}$ in the OMP algorithm (see Theorem 1 in [25]).

Second, to guarantee selection of the true support, we must derive the condition ensuring that the candidate having the minimal residual is the true support. The condition given in (48) in [21] is not correct because the first inequalities in (90) and (91) in [21] do not always hold. Theorem 3 provides a simpler and looser condition in a different way.

Theorem 3. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where $\boldsymbol{v}$ denotes the noise. If $\boldsymbol{\Phi}$ satisfies the RIP condition and the K-sparse vector $\boldsymbol{x}$ satisfies

$$x_{\min} \geqslant \frac{2}{\sqrt{1-\delta_{2K}}}\|\boldsymbol{v}\|_2, \tag{23}$$

then

$$\|\boldsymbol{r}_T\|_2 \leqslant \min_{\Gamma\in\Omega^K} \|\boldsymbol{r}_\Gamma\|_2, \tag{24}$$

where Ω^K denotes the set of all combinations of K columns in $\boldsymbol{\Phi}$.

Proof. First, the upper bound of $\|\boldsymbol{r}_T\|_2$ is obtained as

$$\|\boldsymbol{r}_T\|_2 = \|\boldsymbol{y} - \boldsymbol{\Phi}_T\hat{\boldsymbol{x}}_T\|_2 = \min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{\Phi}_T\boldsymbol{x}\|_2 \leqslant \|\boldsymbol{y} - \boldsymbol{\Phi}_T\boldsymbol{x}_T\|_2 = \|\boldsymbol{v}\|_2, \tag{25}$$

because $\hat{\boldsymbol{x}}_T = \boldsymbol{\Phi}_T^\dagger\boldsymbol{y} = (\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T)^{-1}\boldsymbol{\Phi}_T^*\boldsymbol{y} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{\Phi}_T\boldsymbol{x}\|_2$.

Noting that $\boldsymbol{y} = \boldsymbol{\Phi}_T\boldsymbol{x}_T + \boldsymbol{v} = \boldsymbol{\Phi}_\Gamma\boldsymbol{x}_\Gamma + \boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma} + \boldsymbol{v}$ and $\boldsymbol{P}_\Gamma^\perp = \boldsymbol{I} - \boldsymbol{P}_\Gamma = \boldsymbol{I} - \boldsymbol{\Phi}_\Gamma\boldsymbol{\Phi}_\Gamma^\dagger$, for any set $\Gamma \in \Omega^K$, we have

$$\begin{aligned} \boldsymbol{r}_\Gamma &= \boldsymbol{y} - \boldsymbol{\Phi}_\Gamma\hat{\boldsymbol{x}}_\Gamma = \boldsymbol{y} - \boldsymbol{\Phi}_\Gamma\boldsymbol{\Phi}_\Gamma^\dagger\boldsymbol{y} = (\boldsymbol{I} - \boldsymbol{\Phi}_\Gamma\boldsymbol{\Phi}_\Gamma^\dagger)\boldsymbol{y} \\ &= \boldsymbol{P}_\Gamma^\perp(\boldsymbol{\Phi}_\Gamma\boldsymbol{x}_\Gamma + \boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma} + \boldsymbol{v}) = \boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma} + \boldsymbol{P}_\Gamma^\perp\boldsymbol{v}, \end{aligned} \tag{26}$$

because $\boldsymbol{P}_\Gamma^\perp \boldsymbol{\Phi}_\Gamma = \boldsymbol{0}$. Using $\left\|\boldsymbol{P}_\Gamma^\perp \boldsymbol{v}\right\|_2 \leqslant \|\boldsymbol{v}\|_2$ and Lemma 5, we can obtain

$$\begin{aligned}\|\boldsymbol{r}_\Gamma\|_2 &= \left\|\boldsymbol{P}_\Gamma^\perp \boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma} + \boldsymbol{P}_\Gamma^\perp \boldsymbol{v}\right\|_2 \geqslant \left\|\boldsymbol{P}_\Gamma^\perp \boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_2 - \left\|\boldsymbol{P}_\Gamma^\perp \boldsymbol{v}\right\|_2 \\ &\geqslant \left\|\boldsymbol{P}_\Gamma^\perp \boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_2 - \|\boldsymbol{v}\|_2 \geqslant \sqrt{\left(1-\delta_{|\Gamma\bigcup T|}\right)}\|\boldsymbol{x}_{T-\Gamma}\|_2 - \|\boldsymbol{v}\|_2. \end{aligned} \tag{27}$$

Thus, combining (25) and (27), the sufficient condition for (24) is

$$\sqrt{\left(1-\delta_{|\Gamma\bigcup T|}\right)}\|\boldsymbol{x}_{T-\Gamma}\|_2 - \|\boldsymbol{v}\|_2 \geqslant \|\boldsymbol{v}\|_2,$$

which can be simplified as

$$\|\boldsymbol{x}_{T-\Gamma}\|_2 \geqslant \frac{2}{\sqrt{1-\delta_{|T\bigcup\Gamma|}}}\|\boldsymbol{v}\|_2. \tag{28}$$

Noting that $\|\boldsymbol{x}_{T-\Gamma}\|_2 \geqslant \sqrt{|T-\Gamma|}x_{\min}$ and $1 \leqslant |T-\Gamma| \leqslant K$, $|\Gamma\bigcup T| \leqslant 2K$, it can be seen that (28) is guaranteed under (23), which implies that (24) holds under (23). Theorem 3 is proved.

By using Theorems 2 and 3, the condition under which the MMP will exactly recover the true support in the case of l_2 bounded noise can be obtained as follows.

Theorem 4. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where the noise $\boldsymbol{v}$ satisfies $\|\boldsymbol{v}\|_2 \leqslant \varepsilon$. If $\boldsymbol{\Phi}$ satisfies the condition $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$ and the K-sparse vector $\boldsymbol{x}$ satisfies

$$x_{\min} > c \cdot \varepsilon, \tag{29}$$

where

$$c = \max\left\{\frac{\sqrt{1+\delta_{K+L}}(\sqrt{L}+1)}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}}, \frac{2}{\sqrt{1-\delta_{2K}}}\right\}, \tag{30}$$

then the MMP will recover the support of the sparse signal exactly.

3.2.2 *Gaussian noise*

In practice, Gaussian noise is commonly encountered random noise. If $\boldsymbol{e} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I}_M)$, it can be shown that [26]

$$p\left(||\boldsymbol{e}||_2 \leqslant \sigma\sqrt{M+2\sqrt{M\log M}}\right) \geqslant 1-1/M. \tag{31}$$

By using (31) combined with Theorem 4, we can directly obtain Theorem 5.

Theorem 5. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{e}$, where $\boldsymbol{e} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I}_M)$. If $\boldsymbol{\Phi}$ satisfies the condition $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$ and the K-sparse vector $\boldsymbol{x}$ satisfies

$$x_{\min} > c \cdot \sigma\sqrt{M+2\sqrt{M\log M}}, \tag{32}$$

then the MMP will recover the support of the sparse signal exactly with a probability of at least $1-1/M$.

3.2.3 l_∞ *bounded noise*

For l_∞bounded noise, the MMP chooses the candidate $\hat{s} = s_{\tilde{u}}^K$ that satisfies $\tilde{u} = \arg\min_u \|\boldsymbol{\Phi}^*\boldsymbol{r}_u^K\|_\infty$ as the estimated support set when iteration ends. Similar to l_2 bounded noise, it needs to derive two conditions: one ensuring that at least one candidate of the MMP is the true support set and another guaranteeing that the candidate whose residual has the minimal correlation with the columns of the sensing matrix is the true support.

First, Theorem 6 gives the condition under which at least one candidate of the MMP is the true support.

Theorem 6. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where $\boldsymbol{x}$ is a K-sparse signal and $\boldsymbol{v}$ denotes l_∞ bounded noise (i.e., $\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty \leqslant \varepsilon$). Suppose that the sensing matrix $\boldsymbol{\Phi}$ satisfies $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$, then at least one candidate of the MMP will recover the support set T of the sparse signal $\boldsymbol{x}$ if

$$x_{\min} > \frac{\sqrt{L}(\sqrt{K}+1)}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}}\varepsilon. \tag{33}$$

Proof. Similar to Theorem 1, it needs to prove that under condition (33), at least one correct index is selected in each iteration. Suppose that the MMP performs $k-1$ iterations successfully, that is, there exists a candidate s_i^{k-1} that contains $k-1$ correct indices, $s_i^{k-1} \subset T$ and $|s_i^{k-1}| = k-1$. This assumption naturally holds when $k = 1$. In the k-th iteration, to guarantee choosing at least one correct index, we need to prove that (8) holds under condition (33). Because $|T| = K$, we have

$$\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2 \leqslant \sqrt{K}\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_\infty \leqslant \sqrt{K}\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty \leqslant \sqrt{K}\varepsilon. \tag{34}$$

Then, from (16), we can obtain the lower bound of β_1^k as follows:

$$\begin{aligned}\beta_1^k &\geqslant \frac{1}{\sqrt{K-(k-1)}}(1-\delta_K)\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{1}{\sqrt{K-(k-1)}}\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2 \\ &\geqslant \frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{\sqrt{K}}{\sqrt{K-(k-1)}}\varepsilon. \end{aligned} \tag{35}$$

Additionally, because $|W| = L$, similar to (34), we have $\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \leqslant \sqrt{L}\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_\infty \leqslant \sqrt{L}\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty \leqslant \sqrt{L}\varepsilon$. Thus, using (19), we can also obtain an upper bound of α_L^k as follows:

$$\alpha_L^k \leqslant \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \frac{1}{\sqrt{L}}\|\boldsymbol{\Phi}_W^*\boldsymbol{v}\|_2 \leqslant \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \varepsilon. \tag{36}$$

Combining (35) and (36), to guarantee that (8) holds, we require

$$\frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 - \frac{\sqrt{K}}{\sqrt{K-(k-1)}}\varepsilon > \frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 + \varepsilon.$$

By simplification, the above inequality becomes

$$\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 > \frac{1+\sqrt{\frac{K}{K-(k-1)}}}{\frac{1-\delta_{K+L}}{\sqrt{K-(k-1)}} - \frac{\delta_{K+L}}{\sqrt{L}}}\varepsilon. \tag{37}$$

Because $\|\tilde{\boldsymbol{x}}_i^{k-1}\|_2 \geqslant \|\boldsymbol{x}_{T-s_i^{k-1}}\|_2 \geqslant \sqrt{K-(k-1)}x_{\min}$, we can obtain the sufficient condition for (37) as follows:

$$x_{\min} > \frac{1+\sqrt{\frac{K}{K-(k-1)}}}{1-\delta_{K+L}-\frac{\delta_{K+L}}{\sqrt{L}}\sqrt{K-(k-1)}}\varepsilon.$$

Noting that $k \leqslant K$, we can obtain that (8) will hold if

$$x_{\min} > \frac{1+\sqrt{K}}{1-\delta_{K+L}-\frac{\delta_{K+L}}{\sqrt{L}}\sqrt{K}}\varepsilon,$$

which can be reduced to (33). That is, at least one correct index is chosen in the k-th iteration under the condition given by (33). Theorem 6 is proved.

Remark 5. When $L = 1$, the MMP becomes the OMP and (33) will reduce to

$$x_{\min} > \frac{1+\sqrt{K}}{1-(\sqrt{K}+1)\delta_{K+1}}\varepsilon. \tag{38}$$

In [25], it has been shown that the OMP can exactly recover the support of any K-sparse signal if

$$x_{\min} > \frac{(1+\sqrt{1+\delta_{K+1}})\sqrt{K}}{1-(\sqrt{K}+1)\delta_{K+1}}\varepsilon. \tag{39}$$

It is easy to verify that the above bound, (38), is looser than the bound (39) in the OMP algorithm [25].

Second, to guarantee selecting the true support, Theorem 7 provides the condition ensuring that the candidate whose residual has the minimal correlation with the columns of the sensing matrix is the true support.

Theorem 7. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi x} + \boldsymbol{v}$, where $\boldsymbol{v}$ denotes the l_∞ bounded noise (i.e., $\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty \leqslant \varepsilon$). If $\boldsymbol{\Phi}$ satisfies the RIP condition and the K-sparse vector $\boldsymbol{x}$ satisfies

$$x_{\min} \geqslant \frac{2}{1-\delta_{2K}}\left(1+\frac{\delta_{K+1}}{1-\delta_K}\sqrt{K}\right)\varepsilon, \tag{40}$$

then

$$\|\boldsymbol{\Phi}^*\boldsymbol{r}_T\|_\infty \leqslant \min_{\Gamma\in\Omega^K}\|\boldsymbol{\Phi}^*\boldsymbol{r}_\Gamma\|_\infty, \tag{41}$$

where Ω^K denotes the set of all combinations of K columns in $\boldsymbol{\Phi}$.

Proof. First, because $\boldsymbol{\Phi}_T^*\boldsymbol{r}_T = \boldsymbol{0}$, there exists a $j_0 \in T^C$ such that $\|\boldsymbol{\Phi}^*\boldsymbol{r}_T\|_\infty = |\boldsymbol{\phi}_{j_0}^*\boldsymbol{r}_T|$. The residual $\boldsymbol{r}_T$ can be written as

$$\boldsymbol{r}_T = \boldsymbol{y} - \boldsymbol{\Phi}_T\hat{\boldsymbol{x}}_T = \boldsymbol{\Phi}_T\boldsymbol{x}_T + \boldsymbol{v} - \boldsymbol{\Phi}_T\hat{\boldsymbol{x}}_T = \boldsymbol{\Phi}_T(\boldsymbol{x}_T - \hat{\boldsymbol{x}}_T) + \boldsymbol{v}.$$

Thus, the upper bound of $\|\boldsymbol{\Phi}^*\boldsymbol{r}_T\|_\infty$ is obtained as

$$\begin{aligned}\|\boldsymbol{\Phi}^*\boldsymbol{r}_T\|_\infty &= \left|\boldsymbol{\phi}_{j_0}^*\boldsymbol{r}_T\right| = \left|\boldsymbol{\phi}_{j_0}^*\boldsymbol{\Phi}_T(\boldsymbol{x}_T-\hat{\boldsymbol{x}}_T) + \boldsymbol{\phi}_{j_0}^*\boldsymbol{v}\right| \leqslant \left|\boldsymbol{\phi}_{j_0}^*\boldsymbol{\Phi}_T(\boldsymbol{x}_T-\hat{\boldsymbol{x}}_T)\right| + \left|\boldsymbol{\phi}_{j_0}^*\boldsymbol{v}\right| \\ &\leqslant \left\|\boldsymbol{\phi}_{j_0}^*\boldsymbol{\Phi}_T(\boldsymbol{x}_T-\hat{\boldsymbol{x}}_T)\right\|_2 + \varepsilon \overset{(\mathrm{q})}{\leqslant} \delta_{K+1}\|\boldsymbol{x}_T-\hat{\boldsymbol{x}}_T\|_2 + \varepsilon,\end{aligned} \tag{42}$$

where (q) follows from (4). Additionally, we have

$$\boldsymbol{x}_T - \hat{\boldsymbol{x}}_T = \boldsymbol{x}_T - (\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T)^{-1}\boldsymbol{\Phi}_T^*\boldsymbol{y} = \boldsymbol{x}_T - (\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T)^{-1}\boldsymbol{\Phi}_T^*(\boldsymbol{\Phi}_T\boldsymbol{x}_T + \boldsymbol{v}) = -(\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T)^{-1}\boldsymbol{\Phi}_T^*\boldsymbol{v},$$

and we can obtain

$$\|\boldsymbol{x}_T - \hat{\boldsymbol{x}}_T\|_2 = \left\|-(\boldsymbol{\Phi}_T^*\boldsymbol{\Phi}_T)^{-1}\boldsymbol{\Phi}_T^*\boldsymbol{v}\right\|_2 \overset{(\mathrm{r})}{\leqslant} \frac{1}{1-\delta_K}\|\boldsymbol{\Phi}_T^*\boldsymbol{v}\|_2 \overset{(\mathrm{s})}{\leqslant} \frac{\sqrt{K}\varepsilon}{1-\delta_K}, \tag{43}$$

where (r) and (s) follow from (3) and (34), respectively. Using (42) and (43), we have

$$\|\boldsymbol{\Phi}^*\boldsymbol{r}_T\|_\infty \leqslant \delta_{K+1}\|\boldsymbol{x}_T - \hat{\boldsymbol{x}}_T\|_2 + \varepsilon \leqslant \frac{\sqrt{K}\varepsilon}{1-\delta_K}\delta_{K+1} + \varepsilon. \tag{44}$$

For any set $\Gamma \in \Omega^K$, using (26) we have

$$\begin{aligned}\|\boldsymbol{\Phi}^*\boldsymbol{r}_\Gamma\|_\infty &\geqslant \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{r}_\Gamma\right\|_\infty = \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma} + \boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\right\|_\infty \\ &\geqslant \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_\infty - \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\right\|_\infty.\end{aligned} \tag{45}$$

Thus, a lower bound for $\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\|_\infty$ is

$$\begin{aligned}&\left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_\infty \\ &\overset{(\mathrm{t})}{\geqslant} \frac{1}{\sqrt{|T-\Gamma|}}\left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_2 = \frac{\|\boldsymbol{x}_{T-\Gamma}\|_2}{\sqrt{|T-\Gamma|}\|\boldsymbol{x}_{T-\Gamma}\|_2}\left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_2 \\ &\overset{(\mathrm{u})}{\geqslant} \frac{|\boldsymbol{x}_{T-\Gamma}^*\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}|}{\sqrt{|T-\Gamma|}\|\boldsymbol{x}_{T-\Gamma}\|_2} \overset{(\mathrm{v})}{=} \frac{\left\|\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_2^2}{\sqrt{|T-\Gamma|}\|\boldsymbol{x}_{T-\Gamma}\|_2}\end{aligned}$$

$$\overset{\text{(w)}}{\geqslant} \frac{(1-\delta_{|T\cup\Gamma|})\,\|\boldsymbol{x}_{T-\Gamma}\|_2^2}{\sqrt{|T-\Gamma|}\|\boldsymbol{x}_{T-\Gamma}\|_2} = \frac{(1-\delta_{|T\cup\Gamma|})\,\|\boldsymbol{x}_{T-\Gamma}\|_2}{\sqrt{|T-\Gamma|}}$$
$$\geqslant (1-\delta_{2K})x_{\min}, \tag{46}$$

where (t) follows from $\|\boldsymbol{z}\|_\infty \geqslant \frac{1}{\sqrt{\|\boldsymbol{z}\|_0}}\|\boldsymbol{z}\|_2$, (u) follows from the Cauchy-Schwarz inequality, (v) follows from $\boldsymbol{P}_\Gamma^\perp = \boldsymbol{P}_\Gamma^{\perp *}\boldsymbol{P}_\Gamma^\perp$, (w) follows from Lemma 5, and the last inequality holds because $\|\boldsymbol{x}_{T-\Gamma}\|_2 \geqslant \sqrt{|T-\Gamma|}x_{\min}$ and $|\Gamma \bigcup T| \leqslant 2K$. Then, let $i_0 \in T-\Gamma$ such that $\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\|_\infty = |\boldsymbol{\phi}_{i_0}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}|$. An upper bound for $\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\|_\infty$ can be obtained as

$$\begin{aligned}\left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\right\|_\infty &= \left|\boldsymbol{\phi}_{i_0}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\right| = \left|\boldsymbol{\phi}_{i_0}^*(\boldsymbol{I}-\boldsymbol{P}_\Gamma)\boldsymbol{v}\right| = \left|\boldsymbol{\phi}_{i_0}^*\boldsymbol{v} - \boldsymbol{\phi}_{i_0}^*\boldsymbol{P}_\Gamma\boldsymbol{v}\right| \\ &\leqslant \left|\boldsymbol{\phi}_{j_0}^*\boldsymbol{v}\right| + \left|\boldsymbol{\phi}_{i_0}^*\boldsymbol{P}_\Gamma\boldsymbol{v}\right| \leqslant \varepsilon + \left\|\boldsymbol{\phi}_{j_0}^*\boldsymbol{P}_\Gamma\boldsymbol{v}\right\|_2 = \varepsilon + \left\|\boldsymbol{\phi}_{j_0}^*\boldsymbol{\Phi}_\Gamma(\boldsymbol{\Phi}_\Gamma^*\boldsymbol{\Phi}_\Gamma)^{-1}\boldsymbol{\Phi}_\Gamma^*\boldsymbol{v}\right\|_2 \\ &\overset{\text{(x)}}{\leqslant} \varepsilon + \delta_{K+1}\left\|(\boldsymbol{\Phi}_\Gamma^*\boldsymbol{\Phi}_\Gamma)^{-1}\boldsymbol{\Phi}_\Gamma^*\boldsymbol{v}\right\|_2 \overset{\text{(y)}}{\leqslant} \varepsilon + \frac{\delta_{K+1}}{1-\delta_K}\|\boldsymbol{\Phi}_\Gamma^*\boldsymbol{v}\|_2 \leqslant \left(1+\delta_{K+1}\frac{\sqrt{K}}{1-\delta_K}\right)\varepsilon,\end{aligned} \tag{47}$$

where (x) and (y) follow from (4) and (3), respectively. Using (45), (46), and (47), we have

$$\|\boldsymbol{\Phi}^*\boldsymbol{r}_\Gamma\|_\infty \geqslant \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{\Phi}_{T-\Gamma}\boldsymbol{x}_{T-\Gamma}\right\|_\infty - \left\|\boldsymbol{\Phi}_{T-\Gamma}^*\boldsymbol{P}_\Gamma^\perp\boldsymbol{v}\right\|_\infty \geqslant (1-\delta_{2K})x_{\min} - \left(1+\delta_{K+1}\frac{\sqrt{K}}{1-\delta_K}\right)\varepsilon. \tag{48}$$

Thus, combining (44) and (48), the sufficient condition for (41) is

$$(1-\delta_{2K})x_{\min} - \left(1+\delta_{K+1}\frac{\sqrt{K}}{1-\delta_K}\right)\varepsilon \geqslant \frac{\sqrt{K}\varepsilon}{1-\delta_K}\delta_{K+1} + \varepsilon,$$

which can be simplified as (40). Theorem 7 is proved.

By using Theorems 6 and 7, the condition under which the MMP will exactly recover the true support in the case of l_∞ bounded noise can be obtained as follows.

Theorem 8. Consider the noisy measurements $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{v}$, where $\boldsymbol{v}$ denotes the l_∞ bounded noise (i.e., $\|\boldsymbol{\Phi}^*\boldsymbol{v}\|_\infty \leqslant \varepsilon$). If $\boldsymbol{\Phi}$ satisfies the condition $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$ and the K-sparse vector $\boldsymbol{x}$ satisfies

$$x_{\min} > \lambda \cdot \varepsilon, \tag{49}$$

where

$$\lambda = \max\left\{\frac{\sqrt{L}(\sqrt{K}+1)}{\sqrt{L}-(\sqrt{L}+\sqrt{K})\delta_{K+L}}, \frac{2}{1-\delta_{2K}}\left(1+\frac{\delta_{K+1}}{1-\delta_K}\sqrt{K}\right)\right\}, \tag{50}$$

then the MMP will recover the support of the sparse signal exactly.

4 Conclusion

In this paper, we have analyzed the RIP performance of the MMP algorithm. For the noiseless case, we have presented an improved bound for the RIC of the sensing matrix to guarantee exact recovery of sparse signals via MMP. More importantly, in the presence of l_2 bounded noise, the requirement for the minimum magnitude of the nonzero elements of the sparse signal has been obtained through exploitation of the RIP properties and is more relaxed when compared to existing results. The RIP conditions of the MMP in the cases of Gaussian noise and l_∞ bounded noise have been also derived. Recently, a sharp RIC bound for the OMP has been proposed in [9]. Motivated by this result for the OMP, our future work will focus on further improving the bounds of the MMP and other tree-based greedy algorithms with pruning strategies.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 61421001, 61331021, 61501029, 61671063, 61771046).

References

1 Candes E J, Wakin M B. An introduction to compressive sampling. IEEE Signal Process Mag, 2008, 25: 21–30
2 Bu H X, Bai X, Tao R. Compressed sensing SAR imaging based on sparse representation in fractional Fourier domain. Sci China Inf Sci, 2012, 55: 1789–1800
3 Zhang B C, Zhang Z, Jiang C L, et al. System design and first airborne experiment of sparse microwave imaging radar: initial results. Sci China Inf Sci, 2015, 58: 062306
4 Hou Q K, Liu Y, Fan L J, et al. Compressed sensing digital receiver and orthogonal reconstructing algorithm for wideband ISAR radar. Sci China Inf Sci, 2015, 58: 020302
5 Yu H, Shu F, You Y, et al. Compressed sensing-based time-domain channel estimator for full-duplex OFDM systems with IQ-imbalances. Sci China Inf Sci, 2017, 60: 082303
6 Candes E J, Tao T. Decoding by linear programming. IEEE Trans Inform Theory, 2005, 51: 4203–4215
7 Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory, 2007, 53: 4655–4666
8 Wang J, Shim B. On the recovery limit of sparse signals using orthogonal matching pursuit. IEEE Trans Signal Process, 2012, 60: 4973–4976
9 Wen J M, Zhou Z C, Wang J, et al. A sharp condition for exact support recovery with orthogonal matching pursuit. IEEE Trans Signal Process, 2017, 65: 1370–1382
10 Needell D, Vershynin R. Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. IEEE J Sel Top Signal Process, 2010, 4: 310–316
11 Donoho D L, Tsaig Y, Drori I, et al. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. IEEE Trans Inf Theory, 2012, 58: 1094–1121
12 Needell D, Tropp J A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples. Appl Comput Harmonic Anal, 2009, 26: 301–321
13 Dai W, Milenkovic O. Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans Inf Theory, 2009, 55: 2230–2249
14 Wang J, Kwon S K, Shim B. Generalized orthogonal matching pursuit. IEEE Trans Signal Process, 2012, 60: 6202–6216
15 Satpathi S, Das R L, Chakraborty M. Improving the bound on the RIP constant in generalized orthogonal matching pursuit. IEEE Signal Process Lett, 2013, 20: 1074–1077
16 Shen Y, Pan W L, Li J, et al. Analysis of generalised orthogonal matching pursuit using restricted isometry constant. Electron Lett, 2014, 50: 1020–1022
17 Rui G, Tian W. Blind sparsity weak subspace pursuit for compressed sensing. Electron Lett, 2013, 49: 369–370
18 Cotter S F, Rao B D. Application of tree-based searches to matching pursuit. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Salt Lake City, 2001. 3933–3936
19 Karabulut G Z, Moura L, Panario D, et al. Integrating flexible tree searches to orthogonal matching pursuit algorithm. IEE Proc Vis Image Process, 2006, 153: 538–548
20 Karahanoglu N B, Erdogan H. A* orthogonal matching pursuit: best-first search for compressed sensing signal recovery. Digit Signal Process, 2012, 22: 555–568
21 Wang J, Kwon S, Shim B. Multipath matching pursuit. IEEE Trans Inf Theory, 2014, 60: 2986–3001
22 Shim B, Kwon S, Song B. Sparse detection with integer constraint using multipath matching pursuit. IEEE Commun Lett, 2014, 18: 1851–1854
23 Sah A K, Chaturvedi A K. An MMP-based approach for detection in large MIMO systems using sphere decoding. IEEE Wirel Commun Lett, 2017, 6: 158–161
24 Candés E J. The restricted isometry property and its implications for compressed sensing. Comput Rendus Math, 2008, 346: 589–592
25 Wu R, Huang W, Chen D R. The exact support recovery of sparse signals with noise via orthogonal matching pursuit. IEEE Signal Process Lett, 2013, 20: 403–406
26 Cai T T, Wang L. Orthogonal matching pursuit for sparse signal recovery with noise. IEEE Trans Inf Theory, 2011, 57: 4680–4688

Appendix A

First, we show that the bound (10) is tighter than that shown in (12), i.e.,

$$\frac{1}{\sqrt{K-(k-1)}}(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 > \frac{1-2\delta_{K+L}}{\sqrt{K-k+1}}\frac{\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2}{1-\delta_{K+L}}. \tag{A1}$$

Because $(1-\delta_{K+L})^2 > 1-2\delta_{K+L}$ and

$$\tilde{\boldsymbol{x}}_i^{k-1} = \begin{pmatrix} \boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} \\ \boldsymbol{x}_{T-s_i^{k-1}} \end{pmatrix},$$

it can be easily verified that

$$(1-\delta_{K+L})\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 \geqslant (1-\delta_{K+L})\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2 > \frac{1-2\delta_{K+L}}{1-\delta_{K+L}}\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2.$$

Thus, (A1) holds.

Second, we demonstrate that the bound (11) is tighter than the bound in (13), i.e.,

$$\frac{1}{\sqrt{L}}\delta_{K+L}\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 < \frac{\delta_{K+L}}{\sqrt{L}}\frac{\|\boldsymbol{x}_{T-s_i^{k-1}}\|_2}{1-\delta_{K+L}}. \tag{A2}$$

Because $\hat{\boldsymbol{x}}_{s_i^{k-1}} = \boldsymbol{\Phi}^{\dagger}_{s_i^{k-1}}\boldsymbol{y} = (\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}})^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{y}$ and $\boldsymbol{y} = \boldsymbol{\Phi}_T\boldsymbol{x}_T = \boldsymbol{\Phi}_{s_i^{k-1}}\boldsymbol{x}_{s_i^{k-1}} + \boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}}$, we have

$$\begin{aligned}
\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}} &= \boldsymbol{x}_{s_i^{k-1}} - \left(\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}}\right)^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{y} \\
&= \boldsymbol{x}_{s_i^{k-1}} - \left(\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}}\right)^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\left(\boldsymbol{\Phi}_{s_i^{k-1}}\boldsymbol{x}_{s_i^{k-1}} + \boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}}\right) \\
&= -\left(\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}}\right)^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}}.
\end{aligned}$$

Noting that $|s_i^{k-1}| = k-1$, using (3), (4), and Lemma 1, we can obtain

$$\begin{aligned}
\left\|\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}}\right\|_2 &= \left\|\left(\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{s_i^{k-1}}\right)^{-1}\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2 \\
&\leqslant \frac{1}{1-\delta_{k-1}}\left\|\boldsymbol{\Phi}^*_{s_i^{k-1}}\boldsymbol{\Phi}_{T-s_i^{k-1}}\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2 \\
&\leqslant \frac{\delta_K}{1-\delta_{k-1}}\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2 \leqslant \frac{\delta_K}{1-\delta_{K-1}}\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2.
\end{aligned}$$

Thus, we get

$$\begin{aligned}
\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2^2 &= \left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2^2 + \left\|\boldsymbol{x}_{s_i^{k-1}} - \hat{\boldsymbol{x}}_{s_i^{k-1}}\right\|_2^2 \\
&\leqslant \left[1 + \left(\frac{\delta_K}{1-\delta_{K-1}}\right)^2\right]\cdot\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2^2 \\
&\leqslant \left[1 + \left(\frac{\delta_{K+L}}{1-\delta_{K+L}}\right)^2\right]\cdot\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2^2.
\end{aligned}$$

Because $\sqrt{1+(\frac{\delta_{K+L}}{1-\delta_{K+L}})^2} < \frac{1}{1-\delta_{K+L}}$ (which can be easily verified, due to that $\delta_{K+L} < 1$), it can be shown that

$$\left\|\tilde{\boldsymbol{x}}_i^{k-1}\right\|_2 \leqslant \sqrt{1+\left(\frac{\delta_{K+L}}{1-\delta_{K+L}}\right)^2}\cdot\left\|\boldsymbol{x}_{T-s_i^{k-1}}\right\|_2 < \frac{\|\boldsymbol{x}_{T-s_i^{k-1}}\|_2}{1-\delta_{K+L}}.$$

From the above inequality, it can be seen that (A2) holds.