Multipath Matching Pursuit

Submitted to IEEE trans. on Information theory

Authors: S. Kwon, J. Wang, and B. Shim

Presenter: Hwanchol Jang

- Multipath is investigated rather than a single path for a greedy type of search
- In the final moment, the most promising path is chosen.
- They propose "breadth-first search" and "depth-first search" for greedy algorithm.
- They provide analysis for the performance of MMP with RIP

I. Introduction

CS

- The sparse signals $\mathbf{x} \in \mathbb{R}^n$ can be reconstructed from the compressed measurements $\mathbf{y} = \mathbf{\Phi x} \in \mathbb{R}^n$ even when the system representation is underdetermined ($m<n$), as long as the signal to be recovered is sparse (i.e., number of nonzero elements in the vector is small).

Reconstruction

1. L_0 mimimization

- K-sparse signal $\mathbf{x}$ can be accurately reconstructed using $m=2K$ measurements in a noiseless scenario [2].

2. L1 minimization

- Since ℓ_0-minimization problem is NP-hard and hence not so practical, early works focused on the reconstruction of sparse signals using ℓ_1-norm minimization technique (e.g., basis pursuit [2]).

3. Greedy search

- the greedy search approach is designed to further reduce the computational complexity of the basis pursuit.

- In a nutshell, greedy algorithms identify the support (index set of nonzero elements) of the sparse vector x in an iterative fashion, generating a series of locally optimal updates.

OMP

- In the orthogonal matching pursuit (OMP) algorithm, the index of column that maximizes the magnitude of correlation between columns of Φ and the modified measurements (often called residual) is chosen as a new support element in each iteration.
- If at least one incorrect index is chosen in the middle of the search, the output of OMP will be simply incorrect.

II. MMP algorithm

L0 minimization

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{\Phi x} = \mathbf{y}. \ (1)$$

OMP

- OMP is simple to implement and also computationally efficient
- Due to the choice of the single candidate it is very sensitive to the selection of index.
- The output of OMP will be simply wrong if an incorrect index is chosen in the middle of the search.

Multiple indices

- StOMP algorithm identifying more than one indices in each iteration was proposed. In this approach, indices whose magnitude of correlation exceeds a deliberately designed threshold are chosen [9].
- CoSaMP and SP algorithms maintaining K supports in each iteration were introduced.
- In [12], generalized OMP (gOMP), was proposed. By choosing multiple indices corresponding to $N\,(> 1)$ largest correlation in magnitude in each iteration, gOMP reduces the misdetection probability at the expense of increase in the false alarm probability.

MMP

- The MMP algorithm searches *multiple promising candidates* and then chooses one minimizing the residual in the final moment.
- Due to the investigation of multiple full-blown candidates instead of partial ones, MMP improves the chance of selecting the true support.
- The effect of the random noise vector cannot be accurately judged by just looking at the partial candidate, and more importantly, incorrect decision affects subsequent decision in many greedy algorithms.
- MMP is effective in noisy scenario.

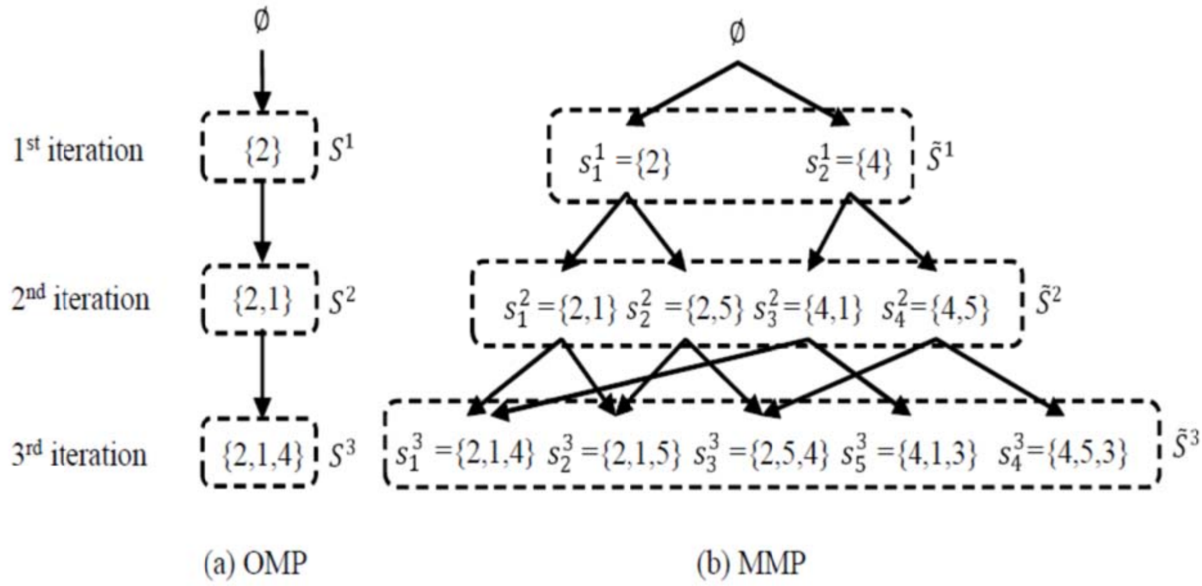

Fig. 1. Comparison between the OMP and the MMP algorithm ($L = 2$ and $K = 3$).

III. Perfect Recovery Condition for MMP

- A recovery condition under which MMP can accurately recover K-sparse signals in the noiseless scenario.
- two parts:
  - A condition ensuring the successful recovery in the initial iteration ($k = 1$).
  - A condition guaranteeing the success in the non-initial iteration ($k > 1$).

- By success we mean that an index of the true support T is chosen in the iteration.

RIP

- A sensing matrix $\mathbf{\Phi}$ is said to satisfy the RIP of order K if there exists a constant $\delta \in (0,1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{\Phi x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \quad (2)$$

for any K-sparse vector $\mathbf{x}$.

- The minimum of all constants δ satisfying (2) is called the restricted isometry constant δ_K .

Lemma 3.1 (Monotonicity of the restricted isometry constant [1]): If the sensing matrix $\mathbf{\Phi}$ satisfies the RIP of both orders K_1 and K_2, then $\delta_{K_1} \le \delta_{K_2}$ for any $K_1 \le K_2$.

Lemma 3.2 (Consequences of RIP [1]): For $I \subset \Omega$, if $\delta_{|I|} < 1$ then for any $\mathbf{x} \in \mathbb{R}^{|I|}$,

$$(1-\delta_{|I|})\|\mathbf{x}\|_2 \le \|\mathbf{\Phi}_I{}'\mathbf{\Phi}_I\mathbf{x}\|_2 \le (1+\delta_{|I|})\|\mathbf{x}\|_2 \quad (3)$$

$$\frac{1}{1+\delta_{|I|}}\|\mathbf{x}\|_2 \le \|(\mathbf{\Phi}_I{}'\mathbf{\Phi}_I)^{-1}\mathbf{x}\|_2 \le \frac{1}{1-\delta_{|I|}}\|\mathbf{x}\|_2 \quad (4)$$

Lemma 3.3 (Lemma 2.1 in [19]): Let $I_1, I_2 \subset \Omega$ be two disjoint sets ($I_1 \cap I_2 = \varnothing$). If $\delta_{|I_1|+|I_2|} < 1$, then

$$\|\mathbf{\Phi}_{I_1}{}'\mathbf{\Phi}_{I_2}\mathbf{x}\|_2 \le \delta_{|I_1|+|I_2|}\|\mathbf{x}\|_2 \quad (5)$$

holds for any $\mathbf{x}$.

Lemma 3.4: For $m \times n$ matrix $\mathbf{\Phi}$, $\|\mathbf{\Phi}\|_2$ satisfies

$$\|\mathbf{\Phi}\|_2 = \sqrt{\lambda_{\max}(\mathbf{\Phi}'\mathbf{\Phi})} \le \sqrt{1+\delta_{\min(m,n)}} \quad (6)$$

A. Success Condition in Initial Iteration

In the first iteration, MMP computes the correlation between measurements $\mathbf{y}$ and each column ϕ_i of $\mathbf{\Phi}$ and then selects L indices whose column has largest correlation in magnitude. Let Λ be the set of L indices chosen in the first iteration, then

$$\|\mathbf{\Phi}'_\Lambda \mathbf{y}\|_2 = \max_{|I|=L} \sqrt{\sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2}. \tag{7}$$

Following theorem provides a condition under which at least one correct index belonging to T is chosen in the first iteration.

Theorem 3.5: Suppose $\mathbf{x} \in \mathbb{R}^n$ is K-sparse signal, then among L candidates at least one contains the correct index in the first iteration of the MMP algorithm if the sensing matrix $\mathbf{\Phi}$ satisfies the RIP with

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + \sqrt{L}}. \tag{8}$$

Proof: From (7), we have

$$\begin{aligned}
\frac{1}{\sqrt{L}} \|\mathbf{\Phi}'_\Lambda \mathbf{y}\|_2 &= \frac{1}{\sqrt{L}} \max_{|I|=L} \sqrt{\sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2} && (9) \\
&= \max_{|I|=L} \sqrt{\frac{1}{|I|} \sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2} && (10) \\
&\geq \sqrt{\frac{1}{|T|} \sum_{i \in T} |\langle \phi_i, \mathbf{y} \rangle|^2} && (11) \\
&= \frac{1}{\sqrt{K}} \|\mathbf{\Phi}'_T \mathbf{y}\|_2 && (12)
\end{aligned}$$

where $|T| = K$. Since $\mathbf{y} = \mathbf{\Phi}_T \mathbf{x}_T$, we further have

$$\begin{aligned}
\|\mathbf{\Phi}'_\Lambda \mathbf{y}\|_2 &\geq \sqrt{\frac{L}{K}} \|\mathbf{\Phi}'_T \mathbf{\Phi}_T \mathbf{x}_T\|_2 && (13) \\
&\geq \sqrt{\frac{L}{K}} (1 - \delta_K) \|\mathbf{x}\|_2 && (14)
\end{aligned}$$

where (14) is due to Lemma 3.2.

On the other hand, when an incorrect index is chosen in the first iteration (i.e., $\Lambda \cap T = \emptyset$),

$$\|\mathbf{\Phi}'_\Lambda \mathbf{y}\|_2 = \|\mathbf{\Phi}'_\Lambda \mathbf{\Phi}_T \mathbf{x}_T\|_2 \leq \delta_{K+L} \|\mathbf{x}\|_2, \tag{15}$$

where the inequality follows from Lemma 3.3. This inequality contradicts (14) if

$$\delta_{K+L} \left\| \mathbf{x} \right\|_2 < \sqrt{\frac{L}{K}} \left(1 - \delta_K \right) \left\| \mathbf{x} \right\|_2 . \tag{16}$$

In other words, under (16) at least one correct index should be chosen in the first iteration ($T_i^1 \in \Lambda$). Further, since $\delta_K \leq \delta_{K+N}$ by Lemma 3.1, (16) holds true if

$$\delta_{K+L} \left\| \mathbf{x} \right\|_2 < \sqrt{\frac{L}{K}} \left(1 - \delta_{K+L} \right) \left\| \mathbf{x} \right\|_2 . \tag{17}$$

Equivalently,

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + \sqrt{L}}. \tag{18}$$

In summary, if $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$, then among L indices at least one belongs to T in the first iteration of MMP. ■

B. Success Condition in Non-initial Iterations

Now we turn to the analysis of the success condition for non-initial iterations. In the k-th iteration ($k > 1$), we focus on the candidate s_i^{k-1} whose elements are exclusively from the true support T (see Fig. 3). In short, our key finding is that at least one of L indices chosen by s_i^{k-1} is from T under $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+3\sqrt{L}}$. Formal description of our finding is as follows.

Theorem 3.6: Suppose a candidate s_i^{k-1} includes indices only in T, then among L children generated from s_i^{k-1} at least one candidate chooses an index in T under

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + 3\sqrt{L}}. \tag{19}$$

Before we proceed, we provide definitions and lemmas useful in our analysis. Let f_i be the i-th largest correlated index in magnitude between $\mathbf{r}^{k-1}$ and $\{\phi_j\}_{j \in T^C}$. That is, $f_j = \arg \max\limits_{j \in T^C \setminus \{f_1, \ldots, f_{(j-1)}\}} \left| \left\langle \phi_j, \mathbf{r}^{k-1} \right\rangle \right|$. Let F_L be the set of these indices ($F_L = \{f_1, f_2, \cdots, f_L\}$). Also, let α_j^k be the j-th largest correlation in magnitude between the residual $\mathbf{r}^{k-1}$ associated with s_i^{k-1} and columns indexed by incorrect indices. That is,

$$\alpha_j^k = \left| \left\langle \phi_{f_j}, \mathbf{r}^{k-1} \right\rangle \right| . \tag{20}$$

Note that α_j^k are ordered in magnitude ($\alpha_1^k \geq \alpha_2^k \geq \cdots$). Finally, let β_j^k be the j-th largest correlation in magnitude between $\mathbf{r}^{k-1}$ and columns whose indices belong to $T - T_i^{k-1}$ (the set of remaining true indices). That is,

$$\beta_j^k = \left| \left\langle \phi_{\varphi(j)}, \mathbf{r}^{k-1} \right\rangle \right| \tag{21}$$

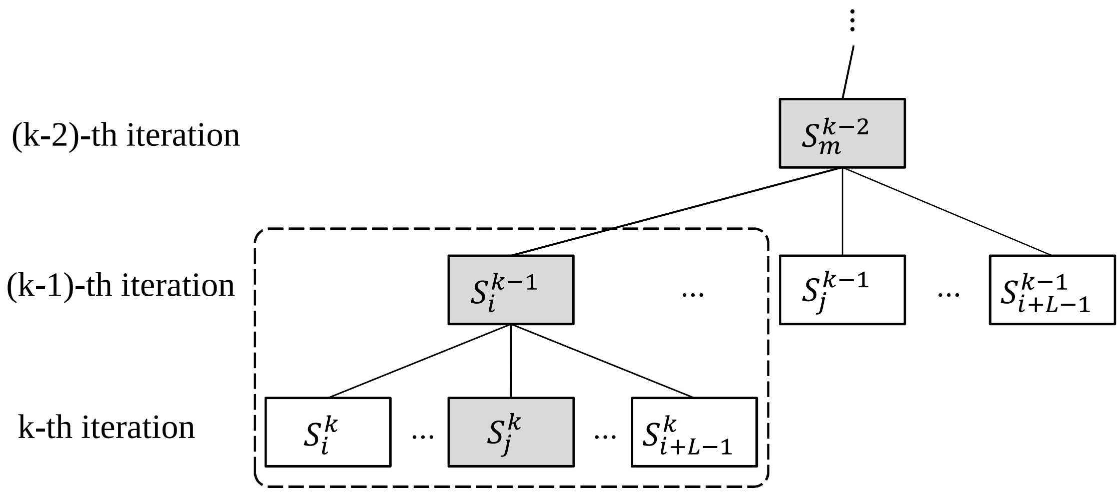

Fig. 3. Relationship between the candidates in $(k-1)$-th iteration and those in k-th iteration. Candidates inside the gray box contain elements of true support T only.

where $\varphi(j) = \arg \max_{j \in \left(T - T^{k-1}\right) \backslash \{\varphi(1),\ldots,\varphi(j-1)\}} \left|\left\langle \phi_j, \mathbf{r}^{k-1} \right\rangle\right|$. Similar to α_j^k, β_j^k are ordered in magnitude ($\beta_1^k \geq \beta_2^k \geq \cdots$). In the following lemmas, we provide the upper bound of α_L^k and lower bound of β_1^k.

Lemma 3.7: α_L^k satisfies

$$\alpha_L^k \leq \left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}}\right) \frac{\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2}{\sqrt{L}}. \tag{22}$$

Proof: See Appendix A. ■

Lemma 3.8: β_1^k satisfies

$$\beta_1^k \geq \left(1 - \delta_{K-k+1} - \frac{\sqrt{1+\delta_{K-k+1}}\sqrt{1+\delta_{k-1}}\delta_K}{1-\delta_{k-1}}\right) \frac{\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2}{\sqrt{K-k+1}}. \tag{23}$$

Proof: See Appendix B. ■

Proof of Theorem 3.6*:* From the definitions of α_j^k and β_j^k, it is clear that a (sufficient) condition under which at least one out of L indices is true in k-th iteration of MMP is

$$\alpha_L^k < \beta_1^k \tag{24}$$

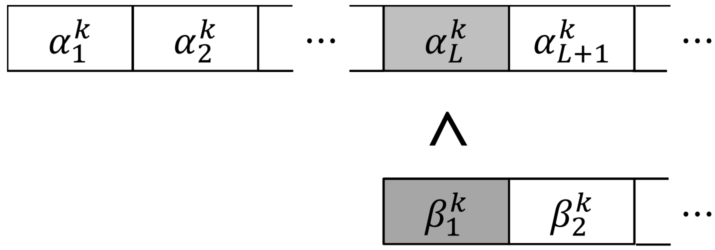

Fig. 4. Comparison between α_N^k and β_1^k. If $\beta_1^k > \alpha_N^k$, then among L indices chosen in K-iteration, at least one is from the true support T.

First, from Lemma 3.1 and 3.7, we have

$$\begin{aligned}
\alpha_L^k &\leq \left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}}\right)\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{L}} && (25)\\
&\leq \left(\delta_{L+K} + \frac{\delta_{L+K}\delta_{L+K}}{1-\delta_{L+K}}\right)\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{L}} && (26)\\
&= \frac{\delta_{L+K}}{1-\delta_{L+K}}\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{L}}. && (27)
\end{aligned}$$

Also, from Lemma 3.1 and 3.8, we have

$$\begin{aligned}
\beta_1^k &\geq \left(1-\delta_{K-k+1} - \frac{\sqrt{1+\delta_{K-k+1}}\sqrt{1+\delta_{k-1}}\delta_K}{1-\delta_{k-1}}\right)\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{K-k+1}} && (28)\\
&\geq \left(1-\delta_{L+K} - \frac{(1+\delta_{L+K})\,\delta_{L+K}}{(1-\delta_{L+K})}\right)\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{K-k+1}} && (29)\\
&= \frac{1-3\delta_{L+K}}{1-\delta_{L+K}}\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{K-k+1}}. && (30)
\end{aligned}$$

Using (24), (27), and (30), we can obtain the sufficient condition of (24) as

$$\frac{1-3\delta_{L+K}}{1-\delta_{L+K}}\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{K-k+1}} > \frac{\delta_{L+K}}{1-\delta_{L+K}}\frac{\left\|\mathbf{x}_{T-s_j^{k-1}}\right\|_2}{\sqrt{L}}. \tag{31}$$

From (31), we further have

$$\delta_{L+K} < \frac{\sqrt{L}}{\sqrt{K-k+1}+3\sqrt{L}}. \tag{32}$$

Since $\sqrt{K-k+1} < \sqrt{K}$ for $k > 1$, (32) holds under $\delta_{L+K} < \frac{\sqrt{L}}{\sqrt{K}+3\sqrt{L}}$, which completes the proof. ■

APPENDIX A
PROOF OF LEMMA 3.7

Proof: The ℓ_2-norm of the correlation $\mathbf{\Phi}'_{F_L}\mathbf{r}^{k-1}$ is expressed as

$$\begin{aligned}\left\|\mathbf{\Phi}'_{F_L}\mathbf{r}^{k-1}\right\|_2 &= \left\|\mathbf{\Phi}'_{F_L}\mathbf{P}^{\perp}_{s_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(102)\\ &= \left\|\mathbf{\Phi}'_{F_L}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-s_j^{k-1}} - \mathbf{\Phi}'_{F_L}\mathbf{P}_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(103)\\ &\le \left\|\mathbf{\Phi}'_{F_L}\mathbf{\Phi}_{T-s_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 + \left\|\mathbf{\Phi}'_{F_L}\mathbf{P}_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2. &(104)\end{aligned}$$

Since F_L and $T-s_j^{k-1}$ are disjoint ($F_L\cap(T-s_j^{k-1})=\emptyset$) and also noting that the number of correct indices in s_j^k is k by the hypothesis,

$$|F_L| + |T-s_j^{k-1}| = L + K - (k-1). \tag{105}$$

Using this together with Lemma 3.3,

$$\left\|\mathbf{\Phi}'_{F_L}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-s_j^{k-1}}\right\|_2 \le \delta_{L+K-k+1}\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2. \tag{106}$$

Similarly, noting that $F_L\cap T_j^{k-1}=\emptyset$ and $|F_L|+|s_j^{k-1}|=L+k-1$, we have

$$\left\|\mathbf{\Phi}'_{F_L}\mathbf{P}_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 \le \delta_{L+k-1}\left\|\mathbf{\Phi}^{\dagger}_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 \tag{107}$$

where

$$\begin{aligned}\left\|\mathbf{\Phi}^{\dagger}_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &= \left\|\left(\mathbf{\Phi}'_{T_j^{k-1}}\mathbf{\Phi}_{T_j^{k-1}}\right)^{-1}\mathbf{\Phi}'_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(108)\\ &\le \frac{1}{1-\delta_{k-1}}\left\|\mathbf{\Phi}'_{T_j^{k-1}}\mathbf{\Phi}_{T-T_j^{k-1}}\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(109)\\ &\le \frac{\delta_{(k-1)+K-(k-1)}}{1-\delta_{k-1}}\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(110)\\ &= \frac{\delta_K}{1-\delta_{k-1}}\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2 &(111)\end{aligned}$$

where (109) and (110) follow from Lemma 3.2 and 3.3, respectively. Since T_j^{k-1} and $T-T_j^{k-1}$ are disjoint, if the number of correct indices in T_j^{k-1} is $k-1$, then

$$\left|T_j^{k-1}\cup\left(T-T_j^{k-1}\right)\right| = (k-1) + K - (k-1). \tag{112}$$

Using (104), (106), (107), and (111), we have

$$\left\|\mathbf{\Phi}'_{F_L}\mathbf{r}^{k-1}\right\|_2 \leq \left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}}\right)\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2. \tag{113}$$

Using the norm inequality $\left(\|\mathbf{z}\|_1 \leq \sqrt{\|\mathbf{z}\|_0}\,\|\mathbf{z}\|_2\right)$, we further have

$$\begin{aligned}
\left\|\mathbf{\Phi}'_{F_L}\mathbf{r}^{k-1}\right\|_2 &\geq \frac{1}{\sqrt{L}}\sum_{i=1}^{L}\alpha_i^k && (114)\\
&\geq \left\|\mathbf{\Phi}'_{F_L}\mathbf{r}^{k-1}\right\|_2 && (115)\\
&\geq \frac{1}{\sqrt{L}}L\alpha_L^k = \sqrt{L}\alpha_L^k && (116)
\end{aligned}$$

where α_j^k is $\left|\left\langle \phi_{f_j}, \mathbf{r}^{k-1}\right\rangle\right|$ [5] and $\alpha_1^k \geq \alpha_2^k \geq \cdots \geq \alpha_L^k$. Combining (113) and (116), we have

$$\left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}}\right)\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2 \geq \sqrt{L}\alpha_L^k, \tag{117}$$

and hence

$$\alpha_L^k \leq \left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}}\right)\frac{\left\|\mathbf{x}_{T-T_j^{k-1}}\right\|_2}{\sqrt{L}}. \tag{118}$$

■

APPENDIX B

PROOF OF LEMMA 3.8

Proof: Since β_1^k is the largest correlation in magnitude between $\mathbf{r}^{k-1}$ and $\{\phi_j\}_{j\in T-T_j^{k-1}}$ $\left(\left|\left\langle \phi_{\varphi(j)}, \mathbf{r}^{k-1}\right\rangle\right|\right)$[6], it is clear that

$$\beta_1^k \geq \left|\left\langle \phi_j, \mathbf{r}^{k-1}\right\rangle\right| \tag{119}$$

for all $j \in T - T_j^{k-1}$, and hence

$$\begin{aligned}
\beta_1^k &\geq \frac{1}{\sqrt{K-(k-1)}}\left\|\mathbf{\Phi}'_{T-T_j^{k-1}}\mathbf{r}^{k-1}\right\| && (120)\\
&= \frac{1}{\sqrt{K-k+1}}\left\|\mathbf{\Phi}'_{T-T_j^{k-1}}\mathbf{P}^{\perp}_{T_j^{k-1}}\mathbf{\Phi}\mathbf{x}\right\| && (121)
\end{aligned}$$

[5] $f_j = \arg\max_{j\in T^C\setminus\{f_1,\ldots,f_{(j-1)}\}} \left|\left\langle \phi_j, \mathbf{r}^{k-1}\right\rangle\right|$

[6] $\varphi(j) = \arg\max_{j\in\left(T-T^{k-1}\right)\setminus\{\varphi(1),\ldots,\varphi(j-1)\}} \left|\left\langle \phi_j, \mathbf{r}^{k-1}\right\rangle\right|$

where (121) follows from $\mathbf{r}^{k-1} = \mathbf{y} - \mathbf{\Phi}_{T_j^{k-1}} \mathbf{\Phi}^{\dagger}_{T_j^{k-1}} \mathbf{y} = \mathbf{P}^{\perp}_{T_j^{k-1}} \mathbf{y}$. Using the triangle inequality,

$$\begin{aligned} \beta_1^k &\geq \frac{1}{\sqrt{K-k+1}} \left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{P}^{\perp}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (122) \\ &\geq \frac{\left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 - \left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2}{\sqrt{K-k+1}}. && (123) \end{aligned}$$

Since $\left| T - T_j^{k-1} \right| = K - (k-1)$,

$$\left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \geq (1 - \delta_{K-k+1}) \left\| \mathbf{x}_{T-T_j^{k-1}} \right\| \tag{124}$$

and also

$$\begin{aligned} \left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 &\leq \left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \right\|_2 \left\| \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (125) \\ &\leq \sqrt{1 + \delta_{K-k+1}} \left\| \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (126) \end{aligned}$$

where (126) follows from Lemma 3.4. Further, we have

$$\begin{aligned} & \left\| \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (127) \\ &= \left\| \mathbf{\Phi}_{T_j^{k-1}} \left(\mathbf{\Phi}'_{T_j^{k-1}} \mathbf{\Phi}_{T_j^{k-1}} \right)^{-1} \mathbf{\Phi}'_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (128) \\ &\leq \sqrt{1 + \delta_{k-1}} \left\| \left(\mathbf{\Phi}'_{T_j^{k-1}} \mathbf{\Phi}_{T_j^{k-1}} \right)^{-1} \mathbf{\Phi}'_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (129) \\ &\leq \frac{\sqrt{1 + \delta_{k-1}}}{1 - \delta_{k-1}} \left\| \mathbf{\Phi}'_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (130) \\ &\leq \frac{\delta_{(k-1)+K-(k-1)} \sqrt{1 + \delta_{k-1}}}{1 - \delta_{k-1}} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2 && (131) \end{aligned}$$

where (129) and (130) are from the definition of RIP and Lemma 3.2. (131) follows from Lemma 3.3 and $\left| T_j^{k-1} \cup \left(T - T_j^{k-1} \right) \right| = (k-1) + K - (k-1)$ since T_j^{k-1} and $T - T_j^{k-1}$ are disjoint sets. Using (126) and (131), we obtain

$$\left\| \mathbf{\Phi}'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}} \mathbf{\Phi}_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \leq \frac{\sqrt{1 + \delta_{K-k+1}} \sqrt{1 + \delta_{k-1}} \delta_K}{1 - \delta_{k-1}} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2. \tag{132}$$

Finally, by combining (123), (124) and (132), we have

$$\beta_1^k \geq \left(1 - \delta_{K-k+1} - \frac{\sqrt{1 + \delta_{K-k+1}} \sqrt{1 + \delta_{k-1}} \delta_K}{1 - \delta_{k-1}} \right) \frac{\left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2}{\sqrt{K-k+1}}. \tag{133}$$

■

REFERENCES

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

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

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

[4] 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

[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, Sept. 2010.

[6] 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, July 2011.

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

[8] R. G. Baraniuk, M. A. Davenport, R. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, pp. 253–263, Dec. 2008.

[9] 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.

[10] D. Needell and J. A. Tropp, "Cosamp: iterative signal recovery from incomplete and inaccurate samples," *Commun. ACM*, vol. 53, no. 12, pp. 93–100, Dec. 2010.

[11] 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.

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

[13] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, April 1967.

[14] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, July 1999.

[15] B. Shim and I. Kang, "Sphere decoding with a probabilistic tree pruning," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4867–4878, Oct. 2008.

[16] B. M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, March 2003.

[17] W. Chen, M. R. D. Rodrigues, and I. J. Wassell, "Projection design for statistical compressive sensing: A tight frame based approach," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2016–2029, 2013.

[18] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998.

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

[20] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Foundations of Computational Mathematics*, vol. 9, pp. 317–334, June 2009.