

# 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  $\mathbf{y} = \Phi\mathbf{x} \in \mathbb{R}^m$  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. $L_0$ minimization

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

#### 2. $L_1$ minimization

- Since  $\ell_0$ -minimization problem is NP-hard and hence not so practical, early works focused on the reconstruction of sparse signals using  $\ell_1$ -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  $\mathbf{x}$  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  $\Phi$  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

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

### 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.



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

## RIP

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

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

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

- The minimum of all constants  $\delta$  satisfying (2) is called the restricted isometry constant  $\delta_K$ .

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

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

$$(1-\delta_{|I|})\|\mathbf{x}\|_2 \leq \|\Phi_I' \Phi_I \mathbf{x}\|_2 \leq (1+\delta_{|I|})\|\mathbf{x}\|_2 \quad (3)$$

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

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

$$\|\Phi_{I_1}' \Phi_{I_2} \mathbf{x}\|_2 \leq \delta_{|I_1|+|I_2|} \|\mathbf{x}\|_2 \quad (5)$$

holds for any  $\mathbf{x}$ .

*Lemma 3.4:* For  $m \times n$  matrix  $\Phi$ ,  $\|\Phi\|_2$  satisfies

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

### A. Success Condition in Initial Iteration

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

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

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  $\Phi$  satisfies the RIP with

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + \sqrt{L}}. \quad (8)$$

*Proof:* From (7), we have

$$\frac{1}{\sqrt{L}} \|\Phi'_{\Lambda}\mathbf{y}\|_2 = \frac{1}{\sqrt{L}} \max_{|I|=L} \sqrt{\sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2} \quad (9)$$

$$= \max_{|I|=L} \sqrt{\frac{1}{|I|} \sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2} \quad (10)$$

$$\geq \sqrt{\frac{1}{|T|} \sum_{i \in T} |\langle \phi_i, \mathbf{y} \rangle|^2} \quad (11)$$

$$= \frac{1}{\sqrt{K}} \|\Phi'_{T}\mathbf{y}\|_2 \quad (12)$$

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

$$\|\Phi'_{\Lambda}\mathbf{y}\|_2 \geq \sqrt{\frac{L}{K}} \|\Phi'_{T}\Phi_T \mathbf{x}_T\|_2 \quad (13)$$

$$\geq \sqrt{\frac{L}{K}} (1 - \delta_K) \|\mathbf{x}\|_2 \quad (14)$$

where (14) is due to Lemma 3.2.

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

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

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

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

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

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

Equivalently,

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + \sqrt{L}}. \quad (18)$$

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

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + 3\sqrt{L}}. \quad (19)$$

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,  $f_j = \arg \max_{j \in T^c \setminus \{f_1, \dots, f_{j-1}\}} |\langle \phi_j, \mathbf{r}^{k-1} \rangle|$ . Let  $F_L$  be the set of these indices ( $F_L = \{f_1, f_2, \dots, f_L\}$ ). Also, let  $\alpha_j^k$  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,

$$\alpha_j^k = |\langle \phi_{f_j}, \mathbf{r}^{k-1} \rangle|. \quad (20)$$

Note that  $\alpha_j^k$  are ordered in magnitude ( $\alpha_1^k \geq \alpha_2^k \geq \dots$ ). Finally, let  $\beta_j^k$  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,

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

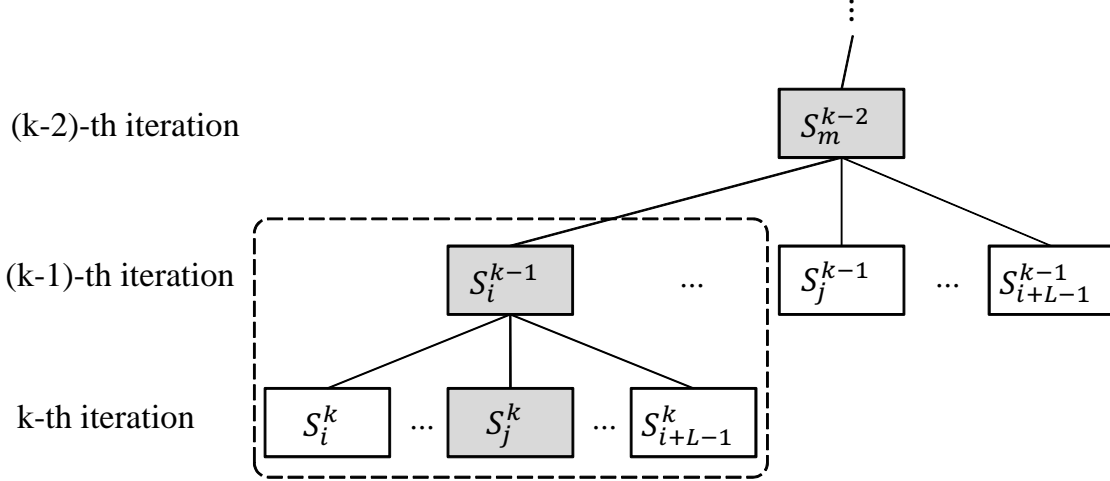


Fig. 3. Relationship between the candidates in  $(k-1)$ -th iteration and those in  $k$ -th iteration. Candidates inside the gray box contain elements of true support  $T$  only.

where  $\varphi(j) = \arg \max_{j \in (T-T^{k-1}) \setminus \{\varphi(1), \dots, \varphi(j-1)\}} |\langle \phi_j, \mathbf{r}^{k-1} \rangle|$ . Similar to  $\alpha_j^k$ ,  $\beta_j^k$  are ordered in magnitude ( $\beta_1^k \geq \beta_2^k \geq \dots$ ). In the following lemmas, we provide the upper bound of  $\alpha_L^k$  and lower bound of  $\beta_1^k$ .

*Lemma 3.7:*  $\alpha_L^k$  satisfies

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

*Proof:* See Appendix A. ■

*Lemma 3.8:*  $\beta_1^k$  satisfies

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

*Proof:* See Appendix B. ■

**Proof of Theorem 3.6:** From the definitions of  $\alpha_j^k$  and  $\beta_j^k$ , 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

$$\alpha_L^k < \beta_1^k \quad (24)$$

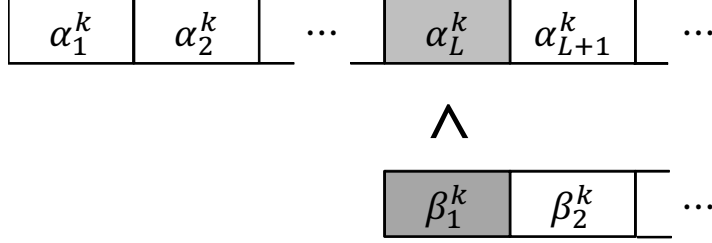


Fig. 4. Comparison between  $\alpha_N^k$  and  $\beta_1^k$ . If  $\beta_1^k > \alpha_N^k$ , then among  $L$  indices chosen in  $K$ -iteration, at least one is from the true support  $T$ .

First, from Lemma 3.1 and 3.7, we have

$$\alpha_L^k \leq \left( \delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}} \right) \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{L}} \quad (25)$$

$$\leq \left( \delta_{L+K} + \frac{\delta_{L+K}\delta_{L+K}}{1-\delta_{L+K}} \right) \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{L}} \quad (26)$$

$$= \frac{\delta_{L+K}}{1-\delta_{L+K}} \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{L}}. \quad (27)$$

Also, from Lemma 3.1 and 3.8, we have

$$\beta_1^k \geq \left( 1 - \delta_{K-k+1} - \frac{\sqrt{1+\delta_{K-k+1}}\sqrt{1+\delta_{k-1}}\delta_K}{1-\delta_{k-1}} \right) \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{K-k+1}} \quad (28)$$

$$\geq \left( 1 - \delta_{L+K} - \frac{(1+\delta_{L+K})\delta_{L+K}}{(1-\delta_{L+K})} \right) \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{K-k+1}} \quad (29)$$

$$= \frac{1-3\delta_{L+K}}{1-\delta_{L+K}} \frac{\|\mathbf{x}_{T-s_j^{k-1}}\|_2}{\sqrt{K-k+1}}. \quad (30)$$

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

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

From (31), we further have

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

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.  $\blacksquare$



APPENDIX A  
PROOF OF LEMMA 3.7

*Proof:* The  $\ell_2$ -norm of the correlation  $\Phi'_{F_L} \mathbf{r}^{k-1}$  is expressed as

$$\left\| \Phi'_{F_L} \mathbf{r}^{k-1} \right\|_2 = \left\| \Phi'_{F_L} \mathbf{P}_{s_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (102)$$

$$= \left\| \Phi'_{F_L} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-s_j^{k-1}} - \Phi'_{F_L} \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (103)$$

$$\leq \left\| \Phi'_{F_L} \Phi_{T-s_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 + \left\| \Phi'_{F_L} \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2. \quad (104)$$

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,

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

Using this together with Lemma 3.3,

$$\left\| \Phi'_{F_L} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-s_j^{k-1}} \right\|_2 \leq \delta_{L+K-k+1} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2. \quad (106)$$

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

$$\left\| \Phi'_{F_L} \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \leq \delta_{L+k-1} \left\| \Phi_{T_j^{k-1}}^\dagger \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (107)$$

where

$$\left\| \Phi_{T_j^{k-1}}^\dagger \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 = \left\| \left( \Phi'_{T_j^{k-1}} \Phi_{T_j^{k-1}} \right)^{-1} \Phi'_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (108)$$

$$\leq \frac{1}{1 - \delta_{k-1}} \left\| \Phi'_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (109)$$

$$\leq \frac{\delta_{(k-1)+K-(k-1)}}{1 - \delta_{k-1}} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (110)$$

$$= \frac{\delta_K}{1 - \delta_{k-1}} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (111)$$

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

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

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

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

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

$$\|\Phi'_{F_L} \mathbf{r}^{k-1}\|_2 \geq \frac{1}{\sqrt{L}} \sum_{i=1}^L \alpha_i^k \quad (114)$$

$$\geq \|\Phi'_{F_L} \mathbf{r}^{k-1}\|_2 \quad (115)$$

$$\geq \frac{1}{\sqrt{L}} L \alpha_L^k = \sqrt{L} \alpha_L^k \quad (116)$$

where  $\alpha_j^k$  is  $|\langle \phi_{f_j}, \mathbf{r}^{k-1} \rangle|$ <sup>5</sup> and  $\alpha_1^k \geq \alpha_2^k \geq \dots \geq \alpha_L^k$ . Combining (113) and (116), we have

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

and hence

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

■

## APPENDIX B

### PROOF OF LEMMA 3.8

*Proof:* Since  $\beta_1^k$  is the largest correlation in magnitude between  $\mathbf{r}^{k-1}$  and  $\{\phi_j\}_{j \in T-T_j^{k-1}}$  ( $|\langle \phi_{\varphi(j)}, \mathbf{r}^{k-1} \rangle|$ )<sup>6</sup>, it is clear that

$$\beta_1^k \geq |\langle \phi_j, \mathbf{r}^{k-1} \rangle| \quad (119)$$

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

$$\beta_1^k \geq \frac{1}{\sqrt{K - (k-1)}} \|\Phi'_{T-T_j^{k-1}} \mathbf{r}^{k-1}\| \quad (120)$$

$$= \frac{1}{\sqrt{K - k + 1}} \|\Phi'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}}^\perp \Phi \mathbf{x}\| \quad (121)$$

<sup>5</sup>  $f_j = \arg \max_{j \in T^C \setminus \{f_1, \dots, f_{(j-1)}\}} |\langle \phi_j, \mathbf{r}^{k-1} \rangle|$

<sup>6</sup>  $\varphi(j) = \arg \max_{j \in (T-T^{k-1}) \setminus \{\varphi(1), \dots, \varphi(j-1)\}} |\langle \phi_j, \mathbf{r}^{k-1} \rangle|$

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

$$\beta_1^k \geq \frac{1}{\sqrt{K-k+1}} \left\| \Phi'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}}^\perp \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (122)$$

$$\geq \frac{\left\| \Phi'_{T-T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 - \left\| \Phi'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2}{\sqrt{K-k+1}}. \quad (123)$$

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

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

and also

$$\left\| \Phi'_{T-T_j^{k-1}} \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \leq \left\| \Phi'_{T-T_j^{k-1}} \right\|_2 \left\| \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (125)$$

$$\leq \sqrt{1 + \delta_{K-k+1}} \left\| \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (126)$$

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

$$\left\| \mathbf{P}_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (127)$$

$$= \left\| \Phi_{T_j^{k-1}} \left( \Phi'_{T_j^{k-1}} \Phi_{T_j^{k-1}} \right)^{-1} \Phi'_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (128)$$

$$\leq \sqrt{1 + \delta_{k-1}} \left\| \left( \Phi'_{T_j^{k-1}} \Phi_{T_j^{k-1}} \right)^{-1} \Phi'_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (129)$$

$$\leq \frac{\sqrt{1 + \delta_{k-1}}}{1 - \delta_{k-1}} \left\| \Phi'_{T_j^{k-1}} \Phi_{T-T_j^{k-1}} \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (130)$$

$$\leq \frac{\delta_{(k-1)+K-(k-1)} \sqrt{1 + \delta_{k-1}}}{1 - \delta_{k-1}} \left\| \mathbf{x}_{T-T_j^{k-1}} \right\|_2 \quad (131)$$

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

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

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

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

■

## 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. Verdú, *Multiuser Detection*, Cambridge University Press, 1998.
- [19] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathématique*, 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.