Mobile Netw Appl (2017) 22:377–382
DOI 10.1007/s11036-016-0773-x

Oblique Projection Matching Pursuit

Jian Wang[1] · Feng Wang[2] · Yunquan Dong[3] · Byonghyo Shim[1]

Published online: 22 November 2016

Abstract Recent theory of compressed sensing (CS) tells us that sparse signals can be reconstructed from a small number of random samples. In reconstruction of sparse signals, greedy algorithms, such as the orthogonal matching pursuit (OMP), have been shown to be computationally efficient. In this paper, the performance of OMP is shown to be dependent on how well information of the underlying signals is preserved in the residual vector. Further, to improve the information preservation, we present a modification of OMP, called oblique projection matching pursuit (ObMP), which updates the residual in a oblique projection manor. Using the restricted isometric property (RIP), we build a solid yet very intuitive grasp of the more accurate phenomenon of ObMP. We also show from empirical experiments that the ObMP achieves improved reconstruction performance over the conventional OMP algorithm in terms of support detection ratio and mean squared error (MSE).

Keywords Compressed sensing (CS) · Sparse recovery · Orthogonal matching pursuit (OMP) · Restricted isometry property (RIP)

✉ Feng Wang
wangfenglucy@gmail.com

Jian Wang
wangjianeee@gmail.com

Yunquan Dong
yunquandong@nuist.edu.cn

Byonghyo Shim
bshim@snu.ac.kr

1 Department of Electrical and Computer Engineering, Seoul National University, Seoul, Korea

2 School of Industrial Management Engineering, Korea University, Seoul, Korea

3 School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

1 Introduction

Compressed sensing (CS) has recently attracted considerable attention in many fields such as electrical engineering, statistics, and applied mathematics [1–3]. The central problem of CS is to estimate a K-sparse vector $\mathbf{x} \in \mathbb{R}^n$ from a small number of linear samples

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x}, \tag{1}$$

where $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ is called the sensing matrix. As $\mathbf{y}$ may belong to one of $\binom{n}{K}$ subspaces spanned by K columns of $\mathbf{\Phi}$, an approach seeking the simplest explanation fitting the data, known as ℓ_0-optimization, has been suggested [1],

$$\min_{x \in \mathbb{R}^n} \|\mathbf{x}\|_0 \qquad \text{subject to} \quad \mathbf{\Phi}\mathbf{x} = \mathbf{y}. \tag{2}$$

This task, however, is known as intractable in practice because the number of cases to be tested grows exponentially with n. Thus much attention has been paid to computationally tractable methods. One popular approach pursuing the relaxation to convex optimization is called ℓ_1-optimization [2],

$$\min_{x \in \mathbb{R}^n} \|\mathbf{x}\|_1 \qquad \text{subject to} \quad \mathbf{\Phi}\mathbf{x} = \mathbf{y}. \tag{3}$$

Alternative approach towards approximation of the sparse solution relies on greedy search. Representative algorithms in this group include matching pursuit (MP) [4], orthogonal matching pursuit (OMP) [5–11]. OMP has been popularly used due to its computational simplicity and competitive recovery performance [12]. For each iteration OMP adds to the list one column that is most strongly correlated with the residual vector. The component of $\mathbf{y}$ associated with the selected list is then cancelled with an orthogonal projection, yielding the updated residual for the next iteration. Since $\mathbf{x}$ is K-sparse, only K iterations are needed for OMP to obtain the sparse solution.

As CS recovery is the inverse process to the sensing process. To ensure reliable recovery, sufficient information should be captured in the sensing process. Therefore, it is important to know how much information captured in $\mathbf{y}$ suffices. This arises the question of how to quantitatively measure the captured information. To the end, a property of the sensing matrix, called restricted isometry property (RIP), has been widely used in the literature. A matrix $\mathbf{\Phi}$ satisfies the RIP of order K if there exists a constant $\delta \in (0, 1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{\Phi x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \tag{4}$$

for all K-sparse $\mathbf{x}$. The minimum of all such δ is called the isometry constant and denoted as δ_K. RIP is a strong descriptor of the ability of $\mathbf{\Phi}$ in preserving information (length) of K-sparse vector $\mathbf{x}$ [2] and has been popularly used to build recovery conditions for variety of CS recovery algorithms. For example, the ℓ_1-optimization achieves exact recovery of any K-sparse $\mathbf{x}$ under $\delta_{2K} < \sqrt{2} - 1$ [13].

In this paper, we present a recovery method, referred to as oblique projection matching pursuit (ObMP), which differs to the conventional OMP algorithm in the residual updating step. Unlike the conventional OMP algorithm that cancels components of selected columns with an orthogonal projection, the ObMP algorithm updates the residual using a oblique projection. The oblique projection incorporates "prior" information of partial support of $\mathbf{x}$ obtained by matched filtering, which is shown to have a better information preservation of the underlying signal to be recovered. Numerical experiments show that ObMP provides improvement over the MP and OMP algorithms in terms of support detection probability as well as mean square error (MSE).

The rest of this paper is organized as follows: In Section 2 we present brief overview of the OMP algorithm and then introduce the ObMP algorithm; We analyze the ability of ObMP in preserving information in Section 3; In Section 4 we present experimental comparison of MP, OMP and ObMP in terms of support detection probability and MSE; Concluding remarks are given in Section 5.

2 The ObMP algorithm

2.1 Notation

We briefly summarize notations in this paper. Let $T = \text{supp}(\mathbf{x}) = \{i | i \in \{1, \cdots, n\}, x_i \neq 0\}$ denote the support set of vector $\mathbf{x}$. For given set $S \subseteq \{1, \cdots, n\}$, $|S|$ is the cardinality of S. $T \setminus S$ is the set of all elements contained in T but not in S. $\mathbf{x}_S \in \mathbb{R}^{|S|}$ is the restriction of the vector $\mathbf{x}$ to the elements with indices in S. $\mathbf{\Phi}_S \in \mathbb{R}^{m \times |S|}$ is a submatrix of $\mathbf{\Phi}$ that only contains columns indexed by S. If $\mathbf{\Phi}_S$ has full column rank, then $\mathbf{\Phi}_S^\dagger = (\mathbf{\Phi}'_S \mathbf{\Phi}_S)^{-1} \mathbf{\Phi}'_S$ is the pseudoinverse of $\mathbf{\Phi}_S$ where $\mathbf{\Phi}'_S$ denotes the transpose of $\mathbf{\Phi}_S$. span$(\mathbf{\Phi}_S)$ is the span of columns in $\mathbf{\Phi}_S$. $\mathbf{P}_S = \mathbf{\Phi}_S \mathbf{\Phi}_S^\dagger$ is the projection onto span$(\mathbf{\Phi}_S)$. $\mathbf{P}_S^\perp = \mathbf{I} - \mathbf{P}_S$ is the projection onto the orthogonal complement of span$(\mathbf{\Phi}_S)$ where $\mathbf{I}$ is the identity matrix.

2.2 Algorithm

Let us begin with a brief overview of the OMP algorithm. At the k-th iteration ($1 \le k < K$), OMP selects an index corresponding to the column of $\mathbf{\Phi}$ that is most strongly correlated with current residual, i.e.,

$$t^k = \arg\max_j |\langle \mathbf{r}^{k-1}, \boldsymbol{\phi}_j \rangle|, \tag{5}$$

and then adds the index to the list:

$$T^k = T^{k-1} \cup \{t^k\}. \tag{6}$$

Based on T^k, a k-sparse signal is estimated as

$$\mathbf{x}^k = \arg \min_{\mathbf{u}:\text{supp}(\mathbf{u}) = T^k} \|\mathbf{y} - \mathbf{\Phi u}\|_2. \tag{7}$$

Finally, the residual is updated by filtering out the component of $\mathbf{y}$ in span$(\mathbf{\Phi}_{T^k})$. That is,

$$\mathbf{r}^k = \mathbf{P}_{T^k}^\perp \mathbf{y}. \tag{8}$$

Now we take some observations on the residual updating step in Eq. 8. Suppose OMP selects k true indices in previous k iterations ($T^k \subset T$). Then, $\mathbf{y}$ admits a parallelogram-wise decomposition:

$$\mathbf{y} = \mathbf{\Phi}_{T^k} \mathbf{x}_{T^k} + \mathbf{\Phi}_{T \setminus T^k} \mathbf{x}_{T \setminus T^k}. \tag{9}$$

See Fig. 1 for a geometric illustration of Eq. 9. Ideally, we would hope that the orthogonal complement operator $\mathbf{P}_{T^k}^\perp$ in $\mathbf{r}^k = \mathbf{P}_{T^k}^\perp \mathbf{y}$ cancels the term $\mathbf{\Phi}_{T^k} \mathbf{x}_{T^k}$ completely but does not affect $\mathbf{\Phi}_{T \setminus T^k} \mathbf{x}_{T \setminus T^k}$. In other words, we hope for an ideal residual

$$\mathbf{r}_0^k = \mathbf{\Phi}_{T \setminus T^k} \mathbf{x}_{T \setminus T^k}. \tag{10}$$

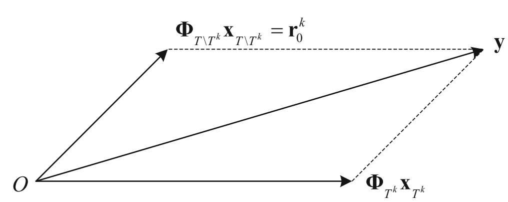

Fig. 1 Sketch of the OMP and the improved OMP algorithm

This, however, is not possible because span($\mathbf{\Phi}_{T^k}$) is not orthogonal to span($\mathbf{\Phi}_{T\setminus T^k}$) in general. In fact, from Eq. 8 and Eq. 9, $\mathbf{r}^k$ is given by

$$\mathbf{r}^k = \mathbf{P}_{T^k}^{\perp}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}, \tag{11}$$

which can be viewed as the "measurements" of vector $\mathbf{x}_{T\setminus T^k}$ using sensing matrix $\mathbf{P}_{T^k}^{\perp}\mathbf{\Phi}_{T\setminus T^k}$. For subsequent iterations, OMP essentially estimates a $(K-k)$-sparse signal (i.e., the vector with $\mathbf{x}_{T\setminus T^k}$ being its nonzero part and zero otherwise) from its "measurements".

As mentioned, the quality of CS recovery is fundamentally determined by the ability of sensing matrix preserving information (typically, the length) of the underlying signals to be recovered. Therefore, the estimate of $\mathbf{x}_{T\setminus T^k}$ relies heavily on how well information of $\mathbf{x}_{T\setminus T^k}$ is preserved in "measurements" $\mathbf{r}^k$.

From Eq. 10 and the definition of RIP, the ideal residual preserves the length of $\mathbf{x}_{T\setminus T^k}$ as

$$(1-\delta_{K-k})\|\mathbf{x}_{T\setminus T^k}\|_2^2 \le \|\mathbf{r}_0^k\|_2^2 \le (1-\delta_{K-k})\|\mathbf{x}_{T\setminus T^k}\|_2^2. \tag{12}$$

In contrast, by Eq. 11 and [14, Lemma 1],

$$(1-\delta_K)\|\mathbf{x}_{T\setminus T^k}\|_2^2 \le \|\mathbf{r}^k\|_2^2 \le (1+\delta_K)\|\mathbf{x}_{T\setminus T^k}\|_2^2. \tag{13}$$

This motivate us to design a new residual update, say $\mathbf{r}_1^k$, such that

$$\|\mathbf{r}^k\|_2^2 < \|\mathbf{r}_1^k\|_2^2 \le \|\mathbf{r}_0^k\|_2^2. \tag{14}$$

In other words, $\mathbf{r}_1^k$ is closer to the ideal residual and hence better preserving the energy of $\mathbf{x}_{T\setminus T^k}$ than $\mathbf{r}^k$. To the end, we propose to use oblique projection to update the residual. Specifically, we first estimate a partial support of $\mathbf{x}$ as

$$\Lambda_\alpha = \{i \mid |(\mathbf{\Phi}'\mathbf{y})_i| \ge \lambda_\alpha\}, \tag{15}$$

where λ_α is the $\lceil \alpha K \rceil$-th largest element (in magnitude) in $\mathbf{\Phi}'\mathbf{y}$ and $0 < \alpha \le 1$. This information can be obtained at the first iteration. Then, by employing the estimated partial support, we estimate $\mathbf{x}^k$ as

$$\mathbf{x}^k = \arg\min_{\mathbf{u}:supp(\mathbf{u})=T^k\cup\Lambda_\alpha} \|\mathbf{y}-\mathbf{\Phi}\mathbf{u}\|_2. \tag{16}$$

Finally, we update the residual as

$$\mathbf{r}_1^k = \mathbf{y} - \mathbf{\Phi}_{T^k}\mathbf{x}_{T^k}^k. \tag{17}$$

Geometrically, $\mathbf{\Phi}_{T^k}\mathbf{x}_{T^k}^k$ is the oblique projection of $\mathbf{y}$ onto span($\mathbf{\Phi}_{T^k}$) against span($\mathbf{\Phi}_{\Lambda_\alpha}$) (see Fig. 2). We henceforth refer to this algorithm as oblique projection matching pursuit (ObMP). A detailed description of ObMP is given in Table 1.

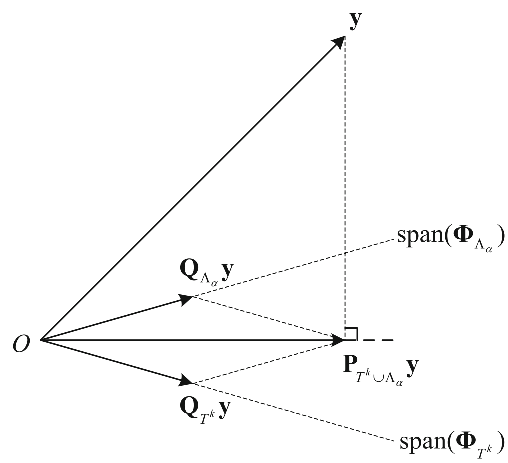

Fig. 2 Sketch of oblique projection of $\mathbf{y}$ onto span($\mathbf{\Phi}_{T^k}$) against span($\mathbf{\Phi}_{\Lambda_\alpha}$)

Proposition 1 *At the k-th iteration* ($1 \le k < K$)*, ObMP satisfies*

$$\mathbf{x}^k = \left(\mathbf{P}_{\Lambda_\alpha}^{\perp}\mathbf{\Phi}_{T^k}\right)^{\dagger}\mathbf{y}, \tag{18}$$

$$\mathbf{r}_1^k = \left(\mathbf{I} - \mathbf{\Phi}_{T^k}\left(\mathbf{P}_{\Lambda_\alpha}^{\perp}\mathbf{\Phi}_{T^k}\right)^{\dagger}\right)\mathbf{y}. \tag{19}$$

Proof See Appendix A □

Let us define

$$\mathbf{Q}_{T^k} = \mathbf{\Phi}_{T^k}\left(\mathbf{P}_{\Lambda_\alpha}^{\perp}\mathbf{\Phi}_{T^k}\right)^{\dagger}, \tag{20}$$

$$\mathbf{Q}_{T^k}^{\perp} = \mathbf{I} - \mathbf{\Phi}_{T^k}\left(\mathbf{P}_{\Lambda_\alpha}^{\perp}\mathbf{\Phi}_{T^k}\right)^{\dagger}, \tag{21}$$

$$\mathbf{Q}_{\Lambda_\alpha} = \mathbf{\Phi}_{\Lambda_\alpha}\left(\mathbf{P}_{T^k}^{\perp}\mathbf{\Phi}_{\Lambda_\alpha}\right)^{\dagger}. \tag{22}$$

Table 1 The ObMP algorithm

Input	$\mathbf{\Phi}$, $\mathbf{y}$, K, and Λ^α.
Initialize	iteration counter $k = 0$,
	estimated support $T^0 = \emptyset$,
	and residual vector $\mathbf{r}^0 = \mathbf{y}$.
While	$k < K$, **do**
	$k = k + 1$.
	Identify $t^k = \arg\max_{i\in\{1,\cdots,n\}\setminus T^{k-1}} \lvert\langle \phi_i, \mathbf{r}^{k-1}\rangle\rvert$.
	Enlarge $T^k = T^{k-1} \cup t^k$.
	Estimate $\mathbf{x}^k = \arg\min_{\mathbf{u}:supp(\mathbf{u})=T^k\cup\Lambda_\alpha} \|\mathbf{y}-\mathbf{\Phi}\mathbf{u}\|_2$.
	Update $\mathbf{r}_1^k = \mathbf{y} - \mathbf{\Phi}_{T^k}\mathbf{x}_{T^k}^k$.
End	
Output	estimated support $\hat{T} = T^K$ and estimated signal $\hat{\mathbf{x}}$
	satisfying $\hat{\mathbf{x}}_{T^K} = \mathbf{x}_{T^K}^{T^K}$ and $\hat{\mathbf{x}}_{\{1,\cdots,n\}\setminus T^K} = \mathbf{0}$.

Then,

$$\mathbf{r}_1^k = \mathbf{Q}_{T^k}^{\perp}\mathbf{y} = \mathbf{Q}_{T^k}^{\perp}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}, \tag{23}$$

which can be viewed as the "measurements" of vector $\mathbf{x}_{T\setminus T^k}$ using sensing matrix $\mathbf{Q}_{T^k}^{\perp}\mathbf{\Phi}_{T\setminus T^k}$.

3 Analysis

3.1 Main results

In this section, we show that at the k-th iteration, the residual of ObMP better preserves the length of $\mathbf{x}_{T\setminus T^k}$ than that of OMP. Our result is formally described in the following theorem.

Theorem 1 *Let $\mathbf{x} \in \mathbb{R}^n$ be any K-sparse vector and $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ be the measurement matrix. Also, let $\mathbf{r}^k$, $\mathbf{r}_1^k$, and $\mathbf{r}_0^k$ be the residual of OMP, ObMP and the ideal case, respectively (see definitions in* Eqs. 10*,* 11 *and* 23*). Then,*

$$\|\mathbf{r}^k\|_2^2 < \|\mathbf{r}_1^k\|_2^2 \le \|\mathbf{r}_0^k\|_2^2. \tag{24}$$

provided that $\mathbf{\Phi}$ satisfies the RIP with isometry constant

$$\delta_{2K} < \frac{1}{3}. \tag{25}$$

We would like to remark that Eq. 25 only imposes a mild constraint to the sensing matrix. Indeed, many random measurement matrices satisfy the RIP with overwhelming probability when the number of measurements scales linearly with the sparsity. For example, a random matrix $\mathbf{\Phi} \in \mathbb{R}^{m\times n}$ with entries drawn i.i.d. from Gaussian distribution $\mathcal{N}(0, \frac{1}{m})$ obeys the RIP with $\delta_K = \varepsilon \in (0, 1)$ with overwhelming probability if [2]

$$m \ge \frac{cK \log \frac{n}{K}}{\varepsilon^2} \tag{26}$$

for some constant $c > 0$.

Before we proceed to the proof for Theorem 1, we provide some useful lemmas.

Lemma 1 (Monotonicity [2]) *If a matrix satisfies the RIP of both orders K_1 and K_2 where $K_1 \le K_2$, then $\delta_{K_1} \le \delta_{K_2}$.*

This property is commonly known as monotonicity of the isometry constant.

Lemma 2 (Consequences of RIP [15, 16]) *Let $S \subseteq \Omega$. If $\delta_{|S|} < 1$, then for any vector $\mathbf{u} \in \mathcal{R}^{|S|}$,*

$$(1-\delta_{|S|})\,\|\mathbf{u}\|_2 \le \left\|\mathbf{\Phi}_S'\mathbf{\Phi}_S\mathbf{u}\right\|_2 \le (1+\delta_{|S|})\,\|\mathbf{u}\|_2\,,$$
$$(1-\delta_{|S|})\|(\mathbf{\Phi}_S'\mathbf{\Phi}_S)^{-1}\mathbf{u}\|_2 \le \|\mathbf{u}\|_2 \le (1+\delta_{|S|})\|(\mathbf{\Phi}_S'\mathbf{\Phi}_S)^{-1}\mathbf{u}\|_2.$$

Lemma 3 (Lemma 1 in [14]) *Let $S_1, S_2 \subseteq \{1, \cdots, n\}$ and $S_1 \cap S_2 = \emptyset$. If $\mathbf{\Phi}$ satisfies the RIP of order $|S_1 \cup S_2|$, then for any $|S_1|$-sparse vector $\mathbf{u} \in \mathbb{R}^n$,*

$$(1-\delta_{|S_1\cup S_2|})\|\mathbf{u}\|_2^2 \le \|\mathbf{P}_{S_2}^{\perp}\Phi\mathbf{u}\|_2^2 \le (1+\delta_{|S_1\cup S_2|})\|\mathbf{u}\|_2^2. \tag{27}$$

Lemma 4 (Lemma 2 in [14]) *For disjoint sets $S_1, S_2, S_3 \subset \{1, \cdots, n\}$. If $\delta_{|S_1\cup S_2\cup S_3|} < 1$, then for any vector $\mathbf{u} \in \mathbb{R}^{|S_3|}$,*

$$\|\mathbf{\Phi}_{S_1}'\mathbf{P}_{S_2}^{\perp}\mathbf{\Phi}_{S_3}\mathbf{u}\|_2^2 \le \delta_{|S_1\cup S_2\cup S_3|}\|\mathbf{u}\|_2 \tag{28}$$

Lemma 5 (Proposition 3.1 in [16]) *Let $S \subseteq \Omega$. If $\delta_{|S|} < 1$, then for any $\mathbf{u} \in \mathcal{R}^{|S|}$,*

$$\|\mathbf{u}\|_2 \le \sqrt{1+\delta_{|S|}}\|(\mathbf{\Phi}_S^{\dagger})'\mathbf{u}\|_2. \tag{29}$$

3.2 Proof of Theorem 1

The proof of Theorem 1 consists of two two steps. First, we derive a lower bound for $\|\mathbf{r}^k - \mathbf{r}_0^k\|_2$ and an upper bound for $\|\mathbf{r}_1^k - \mathbf{r}_0^k||_2$. By relating the two bounds, we obtain an condition ensuring

$$\|\mathbf{r}_1^k - \mathbf{r}_0^k\|_2 < \|\mathbf{r}^k - \mathbf{r}_0^k\|_2. \tag{30}$$

In the second step, we show that Eq. 24 holds true under Eq. 30.

Now we consider the first step of proof.

Proposition 2 *We have*

$$\|\mathbf{r}_1^k - \mathbf{r}_0^k\|_2 \le \frac{\delta_{2K}\sqrt{1+\delta_K}}{1-\delta_K}\|\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2, \tag{31}$$

$$\|\mathbf{r}^k - \mathbf{r}_0^k\|_2 \ge \frac{1-\delta_K}{\sqrt{1+\delta_K}}\|\mathbf{x}_{T\setminus T^k}\|_2. \tag{32}$$

The proof of Proposition 2 is relegated to Appendix B. From Proposition 2, it is clear that Eq. 30 is guaranteed whenever

$$\frac{\delta_K\sqrt{1+\delta_K}}{1-\delta_K}\|\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2 < \frac{1-\delta_K}{\sqrt{1+\delta_K}}\|\mathbf{x}_{T\setminus T^k}\|_2. \tag{33}$$

Noting that

$$\|\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2 \le |\mathbf{x}_{T\setminus T^k}\|_2,$$

it is clear that Eq. 33 is ensured if

$$\delta_{2K}(1+\delta_K) < (1-\delta_K)^2, \tag{34}$$

In fact, by Lemma 1, Eq. 34 holds true under

$$\delta_{2K} < \frac{1}{3}. \tag{35}$$

Next, we proceed to the second step of proof. Since

$$\|\mathbf{r}_1^k - \mathbf{r}_0^k\|_2^2 < \|\mathbf{r}^k - \mathbf{r}_0^k\|_2^2, \tag{36}$$

we have

$$\|\mathbf{r}_1^k - \mathbf{r}^k\|_2^2 > 0. \tag{37}$$

Also, since $\mathbf{r}^k - \mathbf{r}_1^k$ is orthogonal to $\mathbf{r}^k$, by Pythagorean law we have

$$\|\mathbf{r}_1^k\|_2^2 = \|\mathbf{r}_1^k - \mathbf{r}^k\|_2^2 + \|\mathbf{r}^k\|_2^2 > \|\mathbf{r}^k\|_2^2. \tag{38}$$

On the other hand, noting that $\mathbf{r}^k - \mathbf{r}_0^k$ is orthogonal to $\mathbf{r}^k$,

$$\begin{aligned}\|\mathbf{r}_0^k\|_2^2 &= \|\mathbf{r}^k\|_2^2 + \|\mathbf{r}^k - \mathbf{r}_0^k\|_2^2 \\ &\geq \|\mathbf{r}^k\|_2^2 + \|\mathbf{r}^k - \mathbf{r}_1^k\|_2^2 \\ &= \|\mathbf{r}_1^k\|_2^2,\end{aligned} \tag{39}$$

which completes the proof.

4 Experimental results

In this section, we conduct numerical experiments to show the effectiveness of the proposed ObMP algorithm. Let us first introduce our experiment setup. We choose signal $\mathbf{x}$ of length $n = 256$ with K nonzero spike positions randomly generated. We set $m = 50$ and draw a random $m \times n$ matrix $\mathbf{\Phi}$ with i.i.d. Gaussian entries from $\mathcal{N}(0, 1)$. From the CS theory, $\mathbf{\Phi}$ satisfy the RIP with high probability. For comparative purpose, we consider MP, OMP and ObMP algorithms in our experiments. For each algorithm, we run 2,000 independent trials of reconstruction with various sparsity K. Figure 3 depicts the support detection ratio as a function of K. Overall, we observe that support detection ratio decreases as K increases. Compared to MP and OMP, the ObMP algorithm performs better. Figure 4 depicts a plot of the MSE performance as a function of K. We observe that for all recovery algorithms under test, the MSE increases in K. The proposed ObMP exhibits the lowest MSE, demonstrating the performance gain of using oblique projection.

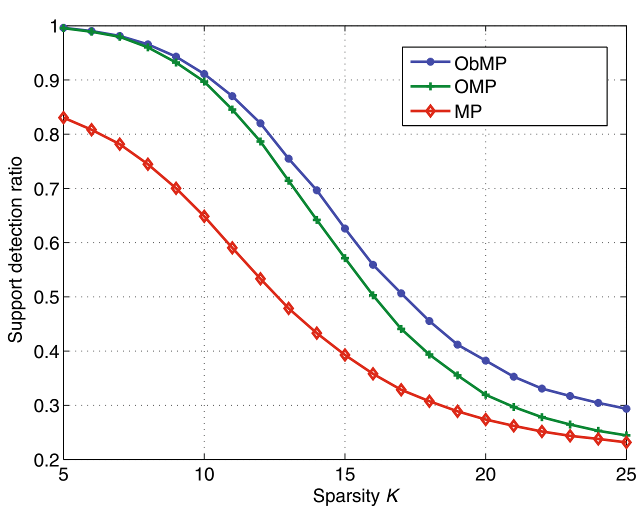

Fig. 3 Support recovery from random measurement with $m = 50$ and $n = 256$ using MP, OMP, and ObMP

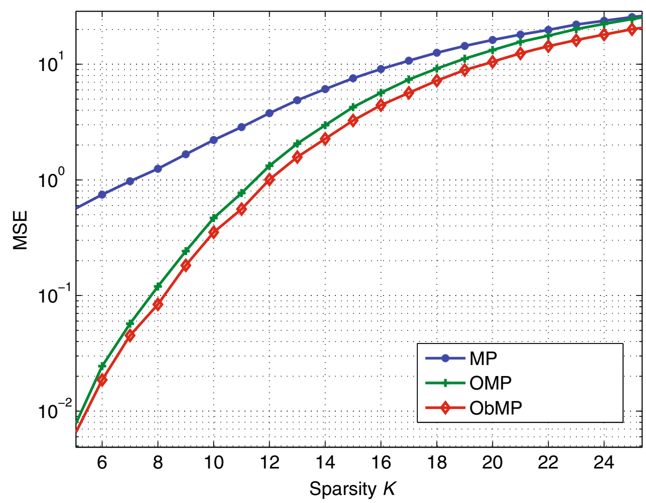

Fig. 4 MSE performance of the signal recovery from random measurement with $m = 50$ and $n = 256$ using MP, OMP, and ObMP

5 Conclusion

In this paper, we have studied the a greedy algorithm called ObMP, which is a modification of OMP with distinct residual updating step. The ObMP algorithm is built on the observations 1) that the recovery performance of OMP depends on how well information about the signals of interest is preserved in the measurements, and 2) that a better information preservation can be achieved by employing an oblique projection that incorporates "prior" knowledge of partial support of input signals. From RIP analysis, we have shown that under a mild constraint on the sensing matrix, the proposed ObMP algorithm is guaranteed to offer better information preservation. In addition, we have demonstrated from numerical experiments that ObMP has very competitive reconstruction performance compared to MP and OMP.

Acknowledgments This work was supported in part by the institute for Information and Communications Technology Promotion Grant Funded by the Korean Government (MSIP) under Grant B0126-16-1017 and in part by the research professor program of Korea University, Korea.

Appendix A: Proof of Proposition 1

Proof Let $\Omega = T^k \cup \Lambda_\alpha$. Then, the projection of $\mathbf{y}$ onto span$(\mathbf{\Phi}_\Omega)$ can be given by

$$\mathbf{P}_\Omega \mathbf{y} = \mathbf{\Phi}_\Omega \mathbf{\Phi}_\Omega^\dagger \mathbf{y}. \tag{40}$$

Following the argument in [17], we can show that

$$\mathbf{P}_\Omega \mathbf{y} = \mathbf{\Phi}_{T^k} (\mathbf{P}_{\Lambda_\alpha}^\perp \mathbf{\Phi}_{T^k})^\dagger \mathbf{y} + \mathbf{\Phi}_{\Lambda_\alpha} (\mathbf{P}_{T^k}^\perp \mathbf{\Phi}_{\Lambda_\alpha})^\dagger \mathbf{y}. \tag{41}$$

Since $\mathbf{P}_{\Omega}\mathbf{y}$ lies in the column span of $\mathbf{\Phi}_{\Omega} = [\mathbf{\Phi}_{T^k}, \mathbf{\Phi}_{\Lambda_\alpha}]$ where $\mathbf{\Phi}_{T^k}$ and $\mathbf{\Phi}_{\Lambda_\alpha}$ are disjoint, it has a unique linear representation with $\mathbf{\Phi}_{T^k}$ and $\mathbf{\Phi}_{\Lambda_\alpha}$. That is,

$$\begin{aligned} \mathbf{P}_{\Omega}\mathbf{y} &= [\mathbf{\Phi}_{T^k}, \mathbf{\Phi}_{\Lambda_\alpha}] \begin{bmatrix} \widehat{\mathbf{x}}_{T^{(k)}} \\ \widehat{\mathbf{x}}_{\Lambda} \end{bmatrix} \\ &= \mathbf{\Phi}_{T^k}\mathbf{x}^k_{T^k} + \mathbf{\Phi}_{\Lambda_\alpha}\mathbf{x}^k_{\Lambda_\alpha}. \end{aligned} \tag{42}$$

From Eqs. 41 and 42, we have

$$\mathbf{x}^k_{T^k} = (\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T^k})^{\dagger})\mathbf{y}. \tag{43}$$

Furthermore, from Table 1, the residual vector of ObMP becomes

$$\begin{aligned} \mathbf{r}^k_1 &= \mathbf{y} - \mathbf{\Phi}_{T^k}\mathbf{x}^k_{T^k} \\ &= \mathbf{y} - \mathbf{\Phi}_{T^k}(\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T^k})^{\dagger})\mathbf{y} \\ &= (\mathbf{I} - \mathbf{\Phi}_{T^k}(\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T^k})^{\dagger})\mathbf{y}, \end{aligned} \tag{44}$$

which completes the proof. □

Appendix B: Proof of Proposition 2

Proof First, for the OMP algorithm, $\|\mathbf{r}^k - \mathbf{r}^k_0\|_2$ can be lower bounded as

$$\begin{aligned} \|\mathbf{r}^k - \mathbf{r}^k_0\|_2 &= \|\mathbf{P}_{T^k}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &= \|\mathbf{\Phi}_{T^k}(\mathbf{\Phi}'_{T^k}\mathbf{\Phi}_{T^k})^{-1}\mathbf{\Phi}'_{T^k}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &= \|(\mathbf{\Phi}^{\dagger}_{T^k})'\mathbf{\Phi}'_{T^k}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &\overset{(a)}{\geq} \frac{\|\mathbf{\Phi}'_{T^k}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2}{\sqrt{1+\delta_K}} \\ &\geq \frac{\sigma_{\min}(\mathbf{\Phi}_{T^k})\|\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2}{\sqrt{1+\delta_K}} \\ &\geq \sqrt{\frac{1-\delta_K}{1+\delta_K}}\|\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &\overset{\text{RIP}}{\geq} \frac{1-\delta_K}{\sqrt{1+\delta_K}}\|\mathbf{x}_{T\setminus T^k}\|_2. \end{aligned} \tag{45}$$

where (a) is from Lemma 5.

We next consider the ObMP algorithm. $\|\mathbf{r}^k_1 - \mathbf{r}^k_0\|_2$ can be upper bounded as follows:

$$\begin{aligned} \|\mathbf{r}^k_1 - \mathbf{r}^k_0\|_2 &= \|\mathbf{Q}_{T^k}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &\overset{(20)}{=} \|\mathbf{\Phi}_{T^k}(\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T^k})^{\dagger}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &\overset{\text{RIP}}{\leq} \sqrt{1+\delta_K}\|(\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T^k})^{\dagger}\mathbf{\Phi}_{T\setminus T^k}\mathbf{x}_{T\setminus T^k}\|_2 \\ &\overset{(a)}{\leq} \frac{\sqrt{1+\delta_K}}{1-\delta_K}\|\mathbf{\Phi}'_{T^k\setminus\Lambda_\alpha}\mathbf{P}^{\perp}_{\Lambda_\alpha}\mathbf{\Phi}_{T\setminus(T^k\cup\Lambda_\alpha)}\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2 \\ &\overset{(b)}{\leq} \frac{\delta_{K+\lceil\alpha K\rceil}\sqrt{1+\delta_K}}{1-\delta_K}\|\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2 \\ &\overset{(c)}{\leq} \frac{\delta_{2K}\sqrt{1+\delta_K}}{1-\delta_K}\|\mathbf{x}_{T\setminus(T^k\cup\Lambda_\alpha)}\|_2, \end{aligned} \tag{46}$$

where (a) is from Lemma 3, (b) is due to Lemma 4, and (c) is because $\alpha \in (0, 1]$.

The proof is now complete. □

References

1. Donoho DL (2006) Compressed sensing. IEEE Trans Inf Theory 52(4):1289–1306
2. Candès EJ, Tao T (2005) Decoding by linear programming. IEEE Trans Inf Theory 51(12):4203–4215
3. Candès EJ, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inf Theory 52(2):489–509
4. Mallat SG, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process 41(12):3397–3415
5. Pati YC, Rezaiifar R, Krishnaprasad PS (1993) Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In: Proceedings of 27th annual Asilomar conference signals, systems, and computers, vol 1. IEEE, Pacific Grove, pp 40–44
6. Wang J, Shim B (2012) On the recovery limit of sparse signals using orthogonal matching pursuit. IEEE Trans Signal Process 60(9):4973–4976
7. Wang J, Kwon S, Shim B (2012) Generalized orthogonal matching pursuit. IEEE Trans Signal Process 60(12):6202–6216
8. Wen J, Zhu X, Li D (2013) Improved bounds on restricted isometry constant for orthogonal matching pursuit. Electron Lett 49(23):1487–1489
9. Wang J (2015) Support recovery with orthogonal matching pursuit in the presence of noise. IEEE Trans Signal Process 63(21):5868–5877
10. Wang J, Kwon S, Li P, Shim B (2016) Recovery of sparse signals via generalized orthogonal matching pursuit: a new analysis. IEEE Trans Signal Process 64(4):1076–1089
11. Wang J, Shim B (2016) Exact recovery of sparse signals using orthogonal matching pursuit: how many iterations do we need? IEEE Trans Signal Process 64(16):4194–4202
12. Tropp JA (2004) Greed is good: algorithmic results for sparse approximation. IEEE Trans Inf Theory 50(10):2231–2242
13. Candès EJ (2008) The restricted isometry property and its implications for compressed sensing. Comptes Rendus Mathematique 346(9–10):589–592
14. Li B, Shen Y, Wu Z, Li J (2015) Sufficient conditions for generalized orthogonal matching pursuit in noisy case. Signal Process 108:111–123
15. Wen J, Li D, Zhu F (2015) Stable recovery of sparse signals via ℓ_p-minimization. Appl Comput Harmon Anal 38(1):161–176
16. Needell D, Tropp JA (2009) CoSaMP: iterative signal recovery from incomplete and inaccurate samples. Appl Comp Harmonic Anal 26(3):301–321
17. Behrens R, Scharf L (1994) Signal processing applications of oblique projection operators. IEEE Trans Signal Process 42(6):1413–1424