Sparse Signal Recovery via Multipath Matching Pursuit

Suhyuk Kwon, Jian Wang, and Byonghyo Shim
Information System Laboratory
School of Information and Communications, Korea University
email: {shkwon, jwang, bshim}@isl.korea.ac.kr

***Abstract*—In this paper, we propose a sparse recovery algorithm, termed multiple path matching pursuit (MMP), that improves the recovery performance of sparse signals. By investigating the multiple paths and then choosing the most promising path in the final moment, the MMP algorithm improves the chance of finding the true support and therefore enhances the recovery performance. From the restricted isometry property (RIP) analysis, we show that the MMP algorithm can perfectly reconstruct any K-sparse ($K > 1$) signals, provided that the sensing matrix satisfies RIP with $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+3\sqrt{L}}$. We demonstrate by empirical simulations that the MMP algorithm is very competitive in both noisy and noiseless scenarios.**

I. Introduction

In recent years, compressed sensing (CS) has received much attention as a means to reconstruct sparse signals from compressed measurements [1]–[7]. Basic premise of the CS is that the sparse signals $\mathbf{x} \in \mathbb{R}^n$ can be reconstructed from the compressed measurements $\mathbf{y} = \mathbf{\Phi x} \in \mathbb{R}^m$ 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).

Early works on this topic focus on the reconstruction of original sparse signals using ℓ_1-norm minimization approach [1], [2]. In [2], Candes, Romberg, and Tao show that the method so called Basis Pursuit (BP) recovers the original sparse signal with high probability if the number of samples is high enough compared to the sparsity level K. Another line of research, designed to alleviate the computational complexity of the BP, is greedy search approaches. The greedy search algorithms include orthogonal matching pursuit (OMP) [3] and its variants such as stagewise OMP (StOMP) [8], subspace pursuit (SP) [9], compressive sampling matching pursuit (CoSaMP) [10], and generalized orthogonal matching pursuit (gOMP) [11]. In a nutshell, greedy algorithms search the support (index set of nonzero elements) of the sparse vector $\mathbf{x}$ in an iterative fashion, generating a series of locally optimal updates.

In this work, we propose an algorithm termed multiple path matching pursuit (MMP) that improves the reconstruction performance of the sparse signals. While most of the conventional greedy algorithms maintain a *single path* (candidate), the MMP algorithm searches *multiple promising paths* and then chooses the most promising path minimizing the ℓ_2-norm cost function $J(\mathbf{x}) = \|\mathbf{y} - \mathbf{\Phi x}\|_2$ in the final moment. Owing to the investigation of the multiple promising paths together with deliberately designed search strategy, the chance of finding the true support set T improves considerably with only marginal overhead in the computational complexity. From the recovery condition analysis (expressed in terms of restricted isometry property (RIP)) and also empirical simulations, we demonstrate that the proposed MMP algorithm is competitive for both noiseless and noisy conditions.

TABLE I
MMP ALGORITHM

```
Input: Measurement y, sensing matrix Φ, sparsity K,
       number of path L
Output: Estimated signal x̂
Initialization: k := 0, r^0 := y, S̃^0 := {∅}
while k < K do
  k := k + 1, u := 0, rpow_min := ∞, S̃^k := ∅
  for i := 1 → |S̃^{k-1}| do
    t := arg max_{|t|=L} ||(Φ' r_i^{k-1})_t||_2^2  /*choose L largest indices*/
    for j := 1 → L do
      s_j^k := s_i^{k-1} ∪ {t(j)}  /*support update*/
      /*check the duplicated path*/
      if s_j^k ∉ S̃^k then
        u := u + 1
        S̃^k := S̃^k ∪ {s_j^k}
        x̂_u^k := Φ^†_{s_j^k} y  /*estimate u-th path in k-iteration*/
        r_u^k := y − Φ_{s_j^k} x̂_u^k  /*residual update*/
        if ||r_u^k||_2^2 < rpow_min then
          x̂* := x̂_u^k
          rpow_min := ||r_u^k||_2^2
        end if
      end if
    end for
  end for
end while
return x̂*
```

The rest of this paper is organized as follows. In Section II, we introduce the proposed MMP algorithm. In Section III, we provide the RIP based recovery condition of the MMP ensuring the perfect reconstruction of sparse signals. In Section IV, we provide empirical simulation results and conclude the paper in Section V.

II. MMP Algorithm

A. Overview of MMP Algorithm

The ℓ_0-norm minimization problem to find out the sparsest solution of underdetermined system is given by

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

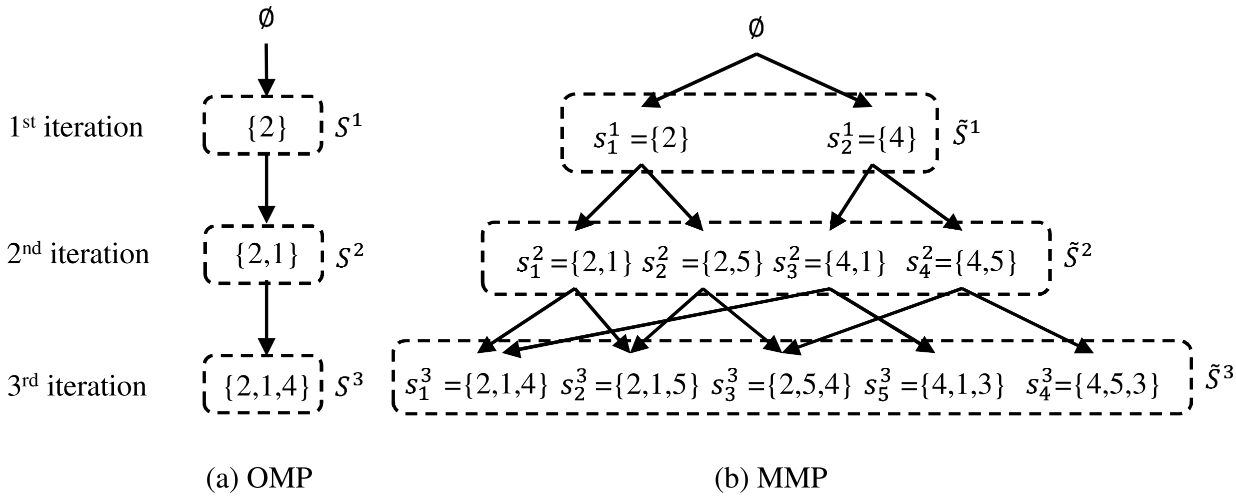

Fig. 1. Comparison between the OMP and the MMP algorithm ($L = 2$ and $K = 3$).

In finding out the solution of this problem, all combinations of columns satisfying the constraint should be tested. In particular, if the sparsity level is set to K, then $\binom{n}{K}$ combinations of columns should be investigated, which is obviously prohibitive for large n and nontrivial K. In the OMP algorithm, however, only one path is being searched using a sequence of greedy operations [3], [12]. Therefore, when an incorrect index is chosen in the iteration loop of the OMP algorithm, the output of the OMP algorithm will be simply incorrect. As a way to overcome this drawback, number of approaches investigating multiple indices has been proposed. For example, in the StOMP algorithm, indices whose magnitude of correlation exceeds a deliberately designed threshold are chosen [8]. Whereas, the CoSaMP and gOMP algorithm choose more than one indices in each iteration [10], [11].

While these approaches exploit *multiple indices* to improve the chance of finding the optimal path, the proposed MMP algorithm searches *multiple paths* and then chooses one minimizing $J(\mathbf{x}) = \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_2$ in the final moment. We compare the operations of the OMP and MMP algorithm in Fig. 1. In contrast to the OMP algorithm, each path s_i^k in a layer k generates L child paths in the MMP algorithm. Note that newly added support element in the path corresponds to the index of columns whose correlation with the residual is maximum (i.e., $t = \{t(1), \cdots, t(L)\} = \arg\max_{|t|=L} \|(\mathbf{\Phi}'\mathbf{r}_i^{k-1})_t\|_2^2$). At a glance, it seems that the number of paths increases by the factor of L for each layer so that the number of path becomes L^K after K iterations. In practice, however, the number of paths increases moderately since large number of paths are overlapping during the search. For example, as illustrated in the Fig. 1(b), $\{2, 5, 4\}$ is the child of $\{2, 5\}$ and $\{4, 5\}$, and $\{2, 1, 4\}$ is the child of $\{2, 1\}$ and $\{4, 1\}$ so that the number of paths in the 2nd iteration is 4 while that in the 3rd iteration becomes only 5. In fact, as shown in Fig. 2, the number of paths of the MMP is much smaller than that of whole enumeration scheme.

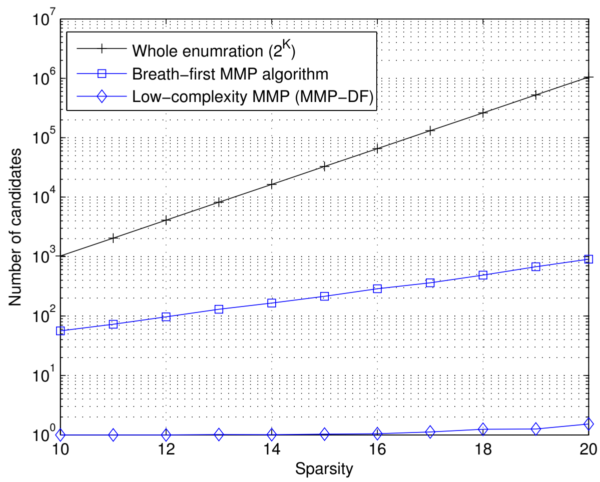

Fig. 2. The required number of paths of the MMP algorithm vs. whole enumeration.

B. Depth-First MMP (MMP-DF)

The MMP algorithm described in the previous subsection can be classified as a breadth first search algorithm (henceforth refer to it as MMP-BF) that requires parallel search of multiple paths. Although large number of paths are merged in the middle of the search, computational overhead is still burdensome and also computational complexity required to obtain the final output varies considerably among realizations of $\mathbf{\Phi}$ and $\mathbf{x}$.

To address this issue, we propose a modification of the MMP algorithm referred to as the depth first MMP (MMP-DF). In the MMP-DF, the path search is performed in a sequential manner and the search will be finished if the magnitude of the residual satisfies suitably chosen termination condition (e.g., $\|\mathbf{r}_\ell\|^2 = 0$ in the noiseless condition) or the

TABLE II
THE SEARCH ORDER AND ITERATION ORDER FOR $L = 4$ AND $K = 3$

Search order ℓ of s_ℓ^K	Layer order (c_1, c_2, c_3) of s_ℓ^K
1	$(1,1,1)$
2	$(2,1,1)$
3	$(3,1,1)$
4	$(4,1,1)$
5	$(1,2,1)$
6	$(2,2,1)$
$\vdots$	$\vdots$
64	$(4,4,4)$

number of paths reaches the predefined maximum value.

In searching the optimal path, there exists $L^K!$ ways in the worst case, and an important question is how to design the best possible search strategy to find out the optimal solution as early as possible. Clearly, an algorithm should have a lower computational cost than the breadth first search yet still have a high chance of finding the solution (even when the number of search is limited). To meet this somewhat complicated goal, we propose a search strategy so called *modulo strategy*. The key motivation of the modulo strategy is to avoid the microsearch around the local optimum and facilitate the search over the wide space. To be specific, ℓ-th path s_ℓ^K (i.e., path being searched at the order ℓ) is determined by

$$\ell = 1 + \sum_{k=1}^{K} (c_k - 1) L^{k-1} \tag{2}$$

where c_k is the order at iteration k. One can easily show that there exists one-to-one correspondence between the search order ℓ and layer order $(c_1, \ldots, c_k)$. Let $L = 4$ and $K = 3$, then the path of the first path s_1^K is expressed as $(c_1, c_2, c_3) = (1, 1, 1)$ so that the best index is added to s_1^K for all iterations (i.e., $s_1^K = \{i_1, i_2, i_3\}$ where $i_k = \arg\max_i |\langle \varphi_i, \mathbf{r}^k \rangle|$ for $k = 1, 2,$ and 3). Whereas, for the 6th path, the path is expressed as $(c_1, c_2, c_3) = (2, 2, 1)$ so that the second best index is chosen in the first and second iterations and the best index is chosen in the last iteration. In Table II, we plot the mapping between the search order and iteration order for $L = 4$ and $K = 3$.

As seen from Fig. 2, the number of path visited of MMP-DF is far smaller than (about two order of magnitude smaller) that of MMP-BF, resulting in significant savings in computational cost.

III. RECOVERY CONDITION ANALYSIS

In this section, we provide the restricted isometry property (RIP) based analysis under which MMP can perfectly recover K-sparse signals. We skip the basic definition of RIP and restricted isometry constant δ_k (see [13] for details).

A. Noiseless Scenario

In this scenario, our analysis is divided into two parts. In the first part, we consider the condition guaranteeing the success in the initial iteration ($k = 1$). In the second part, we investigate the condition guaranteeing the success in the non-initial iteration ($k > 1$). By success we means that the true support index is selected in the iteration. By combining two conditions (choosing a strict condition between two), we obtain a condition guaranteeing the perfect recovery of K-sparse signals.

Recall that the MMP computes the correlation between residual $\mathbf{r}^k$ and each column ϕ_i of $\mathbf{\Phi}$ and then selects L indices of columns having largest correlation in magnitude in each iteration. Following theorem provides the condition for which at least one correct index belonging to T is selected in the first iteration.

Theorem 1: Suppose $\mathbf{x} \in \mathbb{R}^n$ is a K-sparse signal, then among L paths at least one contains correct index in the first iteration of the MMP algorithm if the sensing matrix $\mathbf{\Phi}$ satisfies the RIP constant with

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

Proof: Let Λ^1 denote the set of L indices chosen in the first iteration. Then, elements of $\mathbf{\Phi}'_{\Lambda^1}\mathbf{y}$ are L significant elements in $\mathbf{\Phi}'\mathbf{y}$ and thus

$$\|\mathbf{\Phi}'_{\Lambda^1}\mathbf{y}\|_2 = \max_{|I|=L} \sqrt{\sum_{i \in I} |\langle \phi_i, \mathbf{y} \rangle|^2} \tag{4}$$

where ϕ_i denotes the i-th column in $\mathbf{\Phi}$. One can show that

$$\|\mathbf{\Phi}'_{\Lambda^1}\mathbf{y}\|_2 \geq \sqrt{\frac{L}{K}} \|\mathbf{\Phi}'_T \mathbf{\Phi}_T \mathbf{x}_T\|_2 \geq \sqrt{\frac{L}{K}} (1 - \delta_K) \|\mathbf{x}\|_2 . \tag{5}$$

On the other hand, when no correct index is chosen in the first iteration (i.e., $\Lambda^1 \cap T = \emptyset$), we have

$$\|\mathbf{\Phi}'_{\Lambda^1}\mathbf{y}\|_2 = \|\mathbf{\Phi}'_{\Lambda^1}\mathbf{\Phi}_T \mathbf{x}_T\|_2 \leq \delta_{K+L} \|\mathbf{x}\|_2 . \tag{6}$$

The last inequality contradicts (5) if

$$\delta_{K+L} \|\mathbf{x}\|_2 < \sqrt{\frac{L}{K}} (1 - \delta_K) \|\mathbf{x}\|_2 . \tag{7}$$

This implies that at least one correct index is chosen in the first iteration under (7).

Since $\delta_K \leq \delta_{K+N}$ by the monotonicity of the isometry constant, (7) holds true when

$$\delta_{K+L} \|\mathbf{x}\|_2 < \sqrt{\frac{L}{K}} (1 - \delta_{K+L}) \|\mathbf{x}\|_2 . \tag{8}$$

Equivalently, $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$. In summary, if $\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+\sqrt{L}}$, then Λ^1 contains at least one element of T in the first iteration of the MMP. ∎

In the non-initial iteration , we focus only on the path s_i^{k-1} whose elements are those of true support T (i.e., $s_i^{k-1} \subset T$).

Theorem 2: Suppose the path s_i^{k-1} consists only of true supports, then among L child paths at least one contains correct index under

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K} + 3\sqrt{L}}. \tag{9}$$

Before we proceed, we provide definitions useful in our analysis. First, 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} ($\mathbf{r}^{k-1} = \mathbf{y} - \mathbf{\Phi}_{s_i^{k-1}} \hat{\mathbf{x}}_{s_i^{k-1}}$) and columns indexed by incorrect indices. That is,

$$\alpha_j^k = \left| \langle \phi_{\varphi(j)}, \mathbf{r}^{k-1} \rangle \right| \tag{10}$$

where $\varphi(j) = \arg \max_{j \in T^C \setminus \{\varphi(1), \ldots, \varphi(j-1)\}} \left| \langle \phi_j, \mathbf{r}^{k-1} \rangle \right|$. Note that α_j^k are ordered in magnitude ($\alpha_1^k \geq \alpha_2^k \geq \cdots$). Next, let β_1^k be the largest correlation in magnitude between $\mathbf{r}^{k-1}$ and columns whose indices belong to $T - T_i^{k-1}$. That is,

$$\beta_j^k = \left| \langle \phi_{\varphi(j)}, \mathbf{r}^{k-1} \rangle \right| \tag{11}$$

where $\varphi(j) = \arg \max_{j \in (T - T^{k-1}) \setminus \{\varphi(1), \ldots, \varphi(j-1)\}} \left| \langle \phi_j, \mathbf{r}^{k-1} \rangle \right|$. 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, respectively (see [14]).

Lemma 3: α_L^k satisfies

$$\alpha_L^k \leq \left(\delta_{L+K-k+1} + \frac{\delta_{L+k-1}\delta_K}{1-\delta_{k-1}} \right) \frac{\left\| \mathbf{x}_{T-s_j^{k-1}} \right\|_2}{\sqrt{L}}. \tag{12}$$

Lemma 4: β_1^k satisfies

$$\beta_1^k \geq \left(1 - \delta_{K-k+1} - \frac{\sqrt{1+\delta_{K-k+1}}\sqrt{1+\delta_{k-1}}\delta_K}{1-\delta_{k-1}} \right) \cdot \frac{\left\| \mathbf{x}_{T-s_j^{k-1}} \right\|_2}{\sqrt{K-k+1}} \tag{13}$$

We provide a sketch of proof due to the page limitation. Readers are referred to [14] for details. First, one can easily show that the true support is contained among the L largest correlated supports if

$$\alpha_L^k < \beta_1^k. \tag{14}$$

In order to identify the condition satisfying (14), we find out 1) the upper bound α_L^k and 2) the lower bound β_1^k. Using Lemma 3 and 4, one can further have

$$\alpha_L^k \leq \frac{\delta_{L+K}}{1-\delta_{L+K}} \frac{\left\| \mathbf{x}_{T-s_j^{k-1}} \right\|_2}{\sqrt{L}} \tag{15}$$

and

$$\beta_1^k \geq \frac{1-3\delta_{L+K}}{1-\delta_{L+K}} \frac{\left\| \mathbf{x}_{T-s_j^{k-1}} \right\|_2}{\sqrt{K-k+1}}. \tag{16}$$

After some manipulations, we finally have

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+3\sqrt{L}}. \tag{17}$$

In Theorem 1 and 2, we investigated conditions guaranteeing the success of the MMP algorithm in the initial iteration and non-initial iterations. The overall sufficient condition,

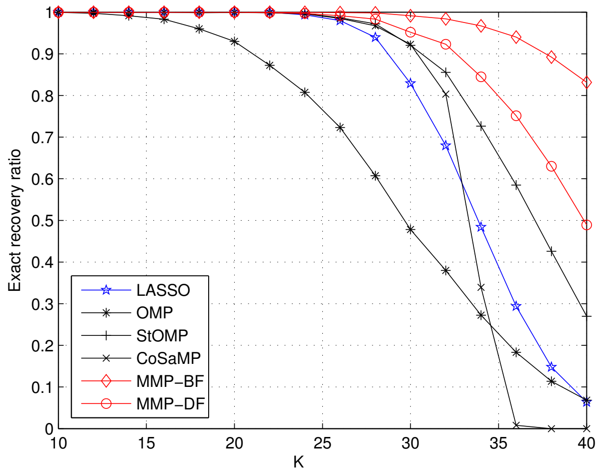

Fig. 3. Reconstruction performance for K-sparse Gaussian signal vector as a function of sparsity K.

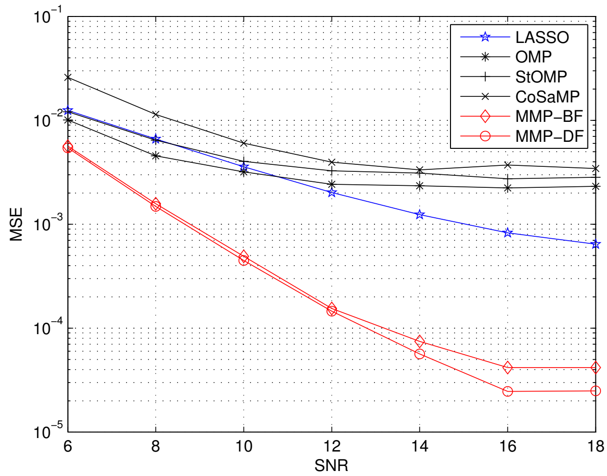

Fig. 4. MSE performance for K-sparse Gaussian signal vector as a function of SNR.

which should be the strict condition between (3) and (9), becomes

$$\delta_{K+L} < \frac{\sqrt{L}}{\sqrt{K}+3\sqrt{L}}. \tag{18}$$

Corollary 5: It is worth mentioning that the perfect recovery condition of the MMP becomes $\delta_{2K} < 0.25$ if $L = K$. When compared to the perfect recovery condition of the CoSaMP algorithm ($\delta_{4K} < 0.1$) [10], the SP algorithm ($\delta_{3K} < 0.165$) [9], and the gOMP algorithm $\left(\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K}+3\sqrt{N}}\right)$ [11], the MMP provides better (more relaxed) recovery condition, which in turn implies that the set of sensing matrices for which exact recovery of the sparse signals is possible gets larger.

IV. Simulations and Discussions

In this section, we observe the performance of the MMP algorithm. As a performance measure, we use the exact recovery

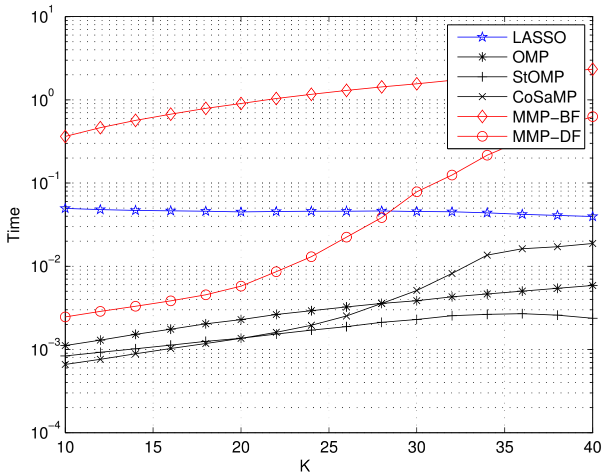

Fig. 5. Running time for K-sparse Gaussian signal vector as a function of sparsity K.

ratio for noiseless scenario and the mean square error (MSE) for noisy scenario. The exact recovery ratio performance is evaluated by checking the maximal sparsity level (often called *critical sparsity*) at which the perfect recovery is ensured. The MSE performance is evaluated by computing $E\|\mathbf{x} - \hat{\mathbf{x}}\|_2$ in various signal-to-noise ratio (SNR) levels. In our simulations, following algorithms are considered:

1) LASSO algorithm
2) OMP algorithm
3) StOMP with false alarm control (FAC)
4) CoSaMP algorithm
5) MMP-BF and MMP-DF. [1]

In each trial, we construct $m \times n$ ($m = 100$ and $n = 256$) sensing matrix $\mathbf{\Phi}$ with entries drawn independently from Gaussian distribution $\mathcal{N}(0, 1/m)$. In addition, we generate K-sparse signal vector $\mathbf{x}$ whose support is chosen at random. In each recovery algorithm, we perform 5,000 independent trials and plot the empirical frequency of exact reconstruction and MSE. In Fig. 3, we provide the recovery performance as a function of the sparsity level K. The simulation results reveal that the critical sparsity of the MMP algorithm is larger than that of the StOMP, CoSaMP, and OMP algorithms. In Fig. 4, we provide the MSE performance as a function of the SNR. In the simulations, K is set to 30. Overall, we observe that MMP outperforms the existing algorithms by a large margin, in particular for high SNR regimes. These results demonstrate that MMP is very competitive in recovering sparse signals for both noiseless and noisy scenarios. Fig. 5 shows the running time as a function of the sparsity level K. We see that the running time of MMP-DF is much smaller than that of MMP-BF and also comparable to the competing algorithms for wide range of sparsity.

[1] Both algorithms expand 6 child paths in every iteration ($L = 6$). The MMP-BF maintains maximally 50 paths in each iteration ($L_{\max} = 50$) and the maximum number of paths for MMP-DF is set to 100 ($l_{\max} = 100$).

V. Conclusion

In this paper, we proposed the sparse signal recovery scheme referred to as the MMP algorithm that improves recovery performance of sparse signals. From the RIP analysis and simulation results, we could observe that the proposed MMP algorithm outperforms conventional sparse recovery algorithms. In particular, we observed that the MMP algorithm is very effective in reconstructing sparse signals in noisy scenarios.

Acknowledgment

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013 (KCA-12-911-01-110), and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A2A2A01047510)

References

[1] D. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845 – 2862, Nov. 2001.

[2] E. 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] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655 –4666, Dec. 2007.

[4] M. Lustig, D. Donoho, J. Santos, and J. Pauly, "Compressed sensing mri," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 72 –82, march 2008.

[5] R. Ward, "Compressed sensing with cross validation," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5773 –5782, Dec. 2009.

[6] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 12, pp. 3728 –3738, Dec. 2010.

[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] D. 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.

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

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

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

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

[14] S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," in preparation.