SPARSE SYMBOL DETECTION BY A GREEDY TREE SEARCH

Jaeseok Lee and Byonghyo Shim

Information System Laboratory
School of Electrical and Computer Engineering, Seoul National University
email : jslee@islab.snu.ac.kr, bshim@snu.ac.kr

ABSTRACT

In this paper, we consider a detection problem of the underdetermined system when the input vector is sparse and its elements are chosen from a set of finite alphabets. We propose a greedy sparse recovery algorithm dubbed as the sparse detection matching pursuit (SDMP) that is effective in recovering the sparse signals with integer constraint. In our performance guarantee analysis and empirical simulations, we show that SDMP is effective in recovering sparse signals in both noiseless and noisy scenarios.

Index Terms— Sparse detection, tree search, compressed sensing, greedy algorithm, underdetermined system

1. INTRODUCTION

The relationship between a transmit signal $\mathbf{x}$ and a received signal vector $\mathbf{y}$ in many wireless communication systems can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{1}$$

where $\mathbf{H} \in \mathcal{C}^{M\times N}$ is the channel matrix, $\mathbf{v} \sim \mathcal{CN}(0, \sigma_v^2\mathbf{I})$, $\mathbf{x}$ is the transmit signal whose entries are from the finite symbol constellation set $\mathbb{X}$. In this work, we are concerned with the scenario where 1) the input signal $\mathbf{x}$ is sparse (i.e., number of nonzero elements in a signal vector is small) and 2) the dimension of observation vector $\mathbf{y}$ is smaller than that of the input vector ($M < N$). Conventional way of detecting the input signals is to use the estimation technique such as linear minimum mean square error (LMMSE) estimation [1] followed by the symbol slicing. For instance, let $\tilde{\mathbf{x}}$ and $\hat{\mathbf{x}}$ be the output of the LMMSE estimator and the symbol detector, respectively, then

$$\begin{aligned}\tilde{\mathbf{x}} &= (\mathbf{H}'\mathbf{H} + \sigma^2\mathbf{I})^{-1}\mathbf{H}'\mathbf{y} && (2)\\ \hat{\mathbf{x}} &= Q_{\mathbb{X}}(\tilde{\mathbf{x}}) && (3)\end{aligned}$$

where $Q_{\mathbb{X}}(\cdot)$ is a function mapping the input to the closest symbol in $\mathbb{X}$ (i.e., $Q_{\mathbb{X}}(x) = \arg\min_{\xi\in\mathbb{X}} \|x-\xi\|_2$). Since the system is underdetermined, this approach, which in essence tries to find the solution by inverting the covariance matrix of the received vector, does not provide satisfactory result in general.

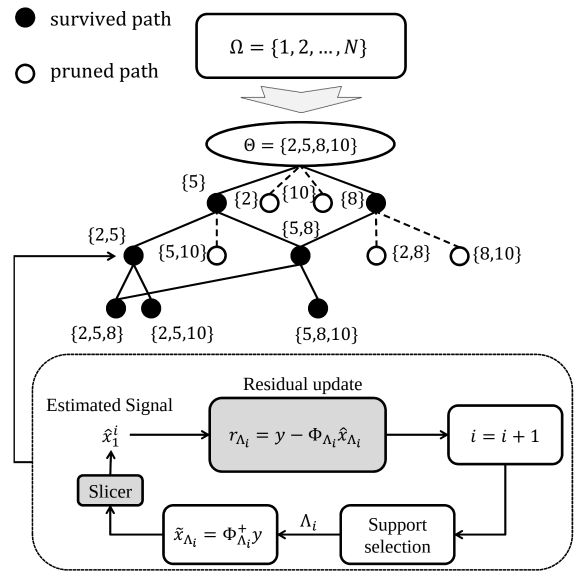

Fig. 1. Illustration of proposed SDMP algorithm.

A better way to exploit the given underdetermined constraint is to use the sparse recovery algorithm. Overall, there are two distinct classes of sparse signal recovery algorithms: ℓ_1-norm minimization technique [2] and greedy approach. In our work, we are mainly interested in the greedy sparse recovery algorithm. Greedy algorithm attempts to find the support (index set of nonzero entries of the original signal vector) in an iterative fashion, returning a sequence of estimates of the sparse input vector. In the orthogonal matching pursuit (OMP) algorithm [3], for example, an index of column in $\mathbf{H}$ maximally correlated to the observation $\mathbf{y}$ is selected in each iteration. That is,

$$i^* = \arg\max_{j\in\{1,\cdots,N\}} |\mathbf{h}_j'\mathbf{y}| \tag{4}$$

This work was partly supported by the ICT R&D program of MSIP/IITP (No. 1291101110-130010100) and National Research Foundation of Korea(NRF) grant funded by the Korea government (MEST) (No. 2014049151).

Table 1. SDMP

Input: measurement $\mathbf{y}$, sensing matrix $\mathbf{\Phi}$, sparsity K, initial threshold ϵ_1
Output: Estimated signal $\hat{\mathbf{x}}$
Initialization: $i := 0$, $S^0 := \emptyset$

$\Theta = f(\mathbf{y}, \mathbf{\Phi}, p)$ (*pre-screening*)
while $i < K$ **do**
  $i := i+1$, $S^i := \emptyset$, $\epsilon_{i+1} := \epsilon_i$
  for $l = 1$ **to** $|S^{i-1}|$ **do**
   $\theta := \Theta \setminus \hat{s}_1^{i-1}(l)$
   for $j = 1$ **to** $|\theta|$ **do**
    $\hat{s}_1^i := \hat{s}_1^{i-1}(l) \cup \{s_i(j)\}$ (*update j-th path*)
    if $\hat{s}_1^i \notin S^i$ **then** (*check the duplicated path*)
     $\tilde{s}_{i+1}^K = \arg\max_{\substack{s \subset \Omega, \\ |s| = K-i}} \|\mathbf{\Phi}'_s \mathbf{r}_{\hat{s}_1^i}\|_2$ (*support estimation*)
     $\bar{s}_1^K = \hat{s}_1^i \cup \tilde{s}_{i+1}^K$, $\hat{\mathbf{x}}_{\bar{s}_1^K} = Q_{\mathbb{X}}(\mathbf{\Phi}^{\dagger}_{\bar{s}_1^K}\mathbf{y})$
     $\mathbf{r}_{\bar{s}_1^K} = \mathbf{y} - \mathbf{\Phi}_{\bar{s}_1^K}\hat{\mathbf{x}}_{\bar{s}_1^K}$
     if $\|\mathbf{r}_{\bar{s}_1^K}\|_2 \le \epsilon_i$ **then** (*pruning decision*)
      $S^i := S^i \cup \hat{s}_1^i$, $I^* := \bar{s}_1^K$
      if $\|\mathbf{r}_{I^*}\|_2 \le \epsilon_{i+1}$ **then**
       $\epsilon_{i+1} := \|\mathbf{r}_{I^*}\|_2$ (*update pruning threshold*)
      end if
     end if
    end if
   end for
  end for
end while
return $\hat{\mathbf{x}}^* = Q_{\mathbb{X}}\left(\mathbf{\Phi}^{\dagger}_{I^*}\mathbf{y}\right)$ (*signal reconstruction*)

where $\mathbf{h}_j$ is the j-th column of $\mathbf{H}$. The chosen index i^* is added to Λ^i (the set of selected indices) and then the estimate $\tilde{\mathbf{x}}_{\Lambda^i} = \mathbf{H}^{\dagger}_{\Lambda^i}\mathbf{y}$ of $\mathbf{x}_{\Lambda^i}$ is generated. Finally, the residual is updated as

$$\mathbf{r}_{\Lambda^i} = \mathbf{y} - \mathbf{H}_{\Lambda^i}\tilde{\mathbf{x}}_{\Lambda^i} = \mathbf{y} - \mathbf{H}_{\Lambda^i}\mathbf{H}^{\dagger}_{\Lambda^i}\mathbf{y}. \tag{5}$$

Well-known drawback of greedy approaches is that the final candidate is often not the optimal solution due to the greedy mechanism of the support selection process. Furthermore, greedy algorithm in itself does not exploit the property that the nonzero elements of the input signal are from finite symbol constellation.

In this paper, we propose a new sparse signal recovery algorithm dubbed as the sparse detection matching pursuit (SDMP). SDMP is effective in recovering the integer-based sparse signals while imposing relatively small computational cost. SDMP accomplishes the mission by 1) the cost-effective tree search based on tree pruning and 2) residual update using integer sliced estimate. In the tree search, SDMP attempts to minimize the cost function by selecting the candidate (set of indices) with minimal residual power[1]. That is,

$$\Lambda^* = \arg\min_{|\Lambda| = K} \|\mathbf{y} - \mathbf{H}_{\Lambda}\hat{\mathbf{x}}_{\Lambda}\|_2. \tag{6}$$

Note that SDMP uses the sliced version of the estimated signal $\hat{\mathbf{x}}_{\Lambda^i} = Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\Lambda^i})$ in the generation of the residual:

$$\mathbf{r}_{\Lambda^i} = \mathbf{y} - \mathbf{H}_{\Lambda^i}Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\Lambda^i}) = \mathbf{y} - \mathbf{H}_{\Lambda^i}\hat{\mathbf{x}}_{\Lambda^i}. \tag{7}$$

Since multiple candidates are investigated in the tree search, SDMP improves the chance of selecting true indices (indices in the support) significantly. Furthermore, even in the presence of noise, exact reconstruction of sparse signals can be achieved due to the symbol slicing in each layer.

2. SPARSE SYMBOL DETECTION VIA GREEDY TREE SEARCH

Major components of SDMP are 1) *pre-screening* to put a limitation on columns of the channel matrix and then 2) *pruning-based tree search* to eliminate unpromising paths from the tree. In the first stage of SDMP, indices that are highly likely to be the elements of support T are selected. In other words, we do our best guess to find the columns of channel matrix that are associated with nonzero elements of the sparse vector. If we denote the set of *roughly chosen* indices as Θ, then the search set is reduced from $\Omega = \{1, 2, \cdots, N\}$ to Θ, a small subset of Ω (i.e., $\Theta \subset \Omega$). In generating Θ, any sparse recovery algorithm can be used. In this work, we use the generalized OMP (gOMP) algorithm [4]. Note that gOMP performs K iterations and chooses L indices per iteration. Since multiple indices are chosen per iteration, misdetection probability of true indices decreases at the increase of the false alarm rate.

In the second stage of SDMP, pruning-based tree search is performed to find a full-blown path Λ with cardinality K ($|\Lambda| = K$) minimizing the cost function $J(\Lambda)$. Since computing cost function of full-blown path is not possible in the middle of the search, we combine the already selected indices (*causal path*) and roughly estimated indices (*noncausal set*). For example, in i-th layer of the tree ($i < K$), the noncausal set $\tilde{s}_{i+1}^K$, temporarily needed for each causal path $\hat{s}_1^i$, is generated by choosing $K-i$ indices of columns in $\Omega \setminus \hat{s}_1^i$. That is,

$$\tilde{s}_{i+1}^K = \arg\max_{\substack{s \subset \Omega \setminus \hat{s}_1^i, \\ |s| = K-i}} \left\|\mathbf{\Phi}'_s \mathbf{r}_{\hat{s}_1^i}\right\|_2. \tag{8}$$

Note that the residual is updated using the sliced estimate as $\mathbf{r}_{\hat{s}_1^i} = \mathbf{y} - \mathbf{H}_{\hat{s}_1^i}Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\hat{s}_1^i}) = \mathbf{y} - \mathbf{H}_{\hat{s}_1^i}\hat{\mathbf{x}}_{\hat{s}_1^i}$. After obtaining the noncausal set $\tilde{s}_{i+1}^K$, we combine this with the causal path $\hat{s}_1^i$

[1]Since the cost function corresponds to the ℓ_2-norm of the residual, minimizing the residual in magnitude is equivalent to minimizing the cost function.

to make a full-blown path $\bar{s}_1^K = \hat{s}_1^i \cup \tilde{s}_{i+1}^K$ and then compute cost function of $\bar{s}_1^K$. If the cost function associated with $\bar{s}_1^K$ is greater than the the pruning threshold ϵ (i.e., $J(\bar{s}_1^K) > \epsilon$), we decide that further investigation is hopeless and prune the path from the tree. It is worth noting that by using the sliced estimate $Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\hat{s}_1^i})$, the estimation performance of the true path can be improved substantially. For example, if $x_k = 1$ and $\tilde{x}_k = 0.4$, then $Q_{\mathbb{X}}(\tilde{x}_k) = 1$ for $\mathbb{X} = \{-1, +1\}$, and hence the estimation error before and after the slicing are $\|x_k - \tilde{x}_k\|_2 = 0.6$ and $\|x_k - \hat{x}_k\|_2 = 0$, respectively. Clearly, slicing entirely removes the residual estimation error of the true path.

3. PERFORMANCE GUARANTEE ANALYSIS

In this section, we provide the sufficient condition under which SDMP accurately reconstructs the sparse signal whose nonzero entries are from finite symbol constellation. Due to the page limitation, we skip the details. The conditions ensuring that SDMP accurately recovers the sparse signals are as follows:

1. At least one true index (index contained in the support T) is selected by the pre-screening ($\Theta \cap T \neq \emptyset$).
2. A true path $\hat{s}_1^i \subset T$ should be survived from the pruning process.
3. For any true path $\hat{s}_1^i \subset T$, the sliced version of $\tilde{\mathbf{x}}_{\hat{s}_1^i}$ should be true symbol ($Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\hat{s}_1^i}) = \mathbf{x}_{\hat{s}_1^i}$).

Since one or more true path should be investigated during the tree search, at least one true index should be selected in the pre-screening stage. In the tree search, the support T should not be removed by the pruning strategy. Even if the path $\hat{s}_1^i$ contains only true indices, exact reconstruction would be achieved only when the sliced version of estimated signal equals the true symbols. In other words, $Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\hat{s}_1^i}) = \mathbf{x}_{\hat{s}_1^i}$ should be satisfied for any $\hat{s}_1^i \subset T$. In order to meet this requirement, the estimation error should be smaller than $\frac{\Delta}{2}$ ($\|\mathbf{x}_{\hat{s}_1^i} - \tilde{\mathbf{x}}_{\hat{s}_1^i}\| < \frac{\Delta}{2}$) where Δ is the minimum distance between symbols.

In our analysis, we use the gOMP algorithm as the pre-screening algorithm [4]. The following theorem provides the condition under which at least one support index is chosen by the pre-screening.

Theorem 1 *The gOMP algorithm identifies at least one support element if the nonzero entries of* $\mathbf{x}$ *satisfy*

$$\min_{j \in T} |x_j| > \eta \|\mathbf{v}\|_2 \tag{9}$$

where $\eta = \frac{(\sqrt{K}+\sqrt{L})\sqrt{1+\delta_{L+K}}}{\sqrt{L}(1-\delta_K)-\sqrt{K}\delta_{L+K}}$.

Next theorem provides the condition under which the path containing true indices is found exclusively in the tree search process.

Theorem 2 *A true path is survived from the pruning process if*

$$\min_{j \in T} |x_j| > \max\{\mu, \nu\} \|\mathbf{v}\|_2 \tag{10}$$

where $\mu = \frac{2\sqrt{1-\delta_M}}{\sqrt{1-\delta_M}\sqrt{1-\delta_K}-\delta_{K+1}}$ *and* $\nu = \frac{2(1-\delta_K)}{1-3\delta_{2K}}$.

The third condition describes the condition that the sliced version is identical to the original signal when the signal is estimated by any true path $\hat{s}_1^i \subset T$. In order to meet this requirement, the estimation error $\|\mathbf{x}_{\hat{s}_1^i} - \tilde{\mathbf{x}}_{\hat{s}_1^i}\|_2$ should be smaller than $\frac{\Delta}{2}$ for any $\hat{s}_1^i \subset T$.

Theorem 3 *For any true path* $\hat{s}_1^i \subset T$*, the estimated signal is detected to true symbols (*$Q_{\mathbb{X}}(\tilde{\mathbf{x}}_{\hat{s}_1^i}) = \mathbf{x}_{\hat{s}_1^i}$*) if*

$$\|\mathbf{v}\|_2 < \frac{(1-\delta_K)^{\frac{3}{2}}}{1+\delta_K}\frac{\Delta}{2} - \frac{\|\mathbf{x}\|_2}{\sqrt{1-\delta_K}} \tag{11}$$

for any $1 \leq i \leq K-1$.

Finally, SDMP exactly reconstructs the sparse signals if Theorem 1, 2, and 3 are jointly satisfied. Next theorem provides the overall sufficient condition ensuring exact reconstruction of sparse signals by SDMP.

Theorem 4 *When* (11) *is satisfied, SDMP accurately reconstructs the sparse signals whose nonzero elements are from finite set of symbol constellation if*

$$\min_{j \in T} |x_j| > \frac{\gamma(1-\delta_K)^2 \Delta}{2(1+\delta_K)(\gamma + \sqrt{1-\delta_K})} \tag{12}$$

where $\gamma = \max\{\eta, \mu, \nu\}$.

It is worth noting that when the condition (11) and (12) are jointly satisfied, the support is accurately identified and the sparse coefficients are also reconstructed accurately. Interestingly, when the signal power is sufficiently large, SDMP exactly reconstructs the original sparse signal even in the presence of noise. To be specific, under Theorem 4, any true path $\hat{s}_1^i \subset T$ is survived from the tree search and the signal is accurately reconstructed (i.e., $\hat{\mathbf{x}}_{\hat{s}_1^i} = \mathbf{x}_{\hat{s}_1^i}$) for each layer. Eventually, SDMP obtains the overdetermined system model $\mathbf{y} = \mathbf{H}_T \mathbf{x}_T + \mathbf{v}$, and the final output is identical to the original signal (i.e., $\hat{\mathbf{x}}_T = Q_{\mathbb{X}}(\tilde{\mathbf{x}}_T) = \mathbf{x}_T$).

4. SIMULATION AND DISCUSSION

4.1. Simulation Setup

In this section, we provide the empirical performance of sparse recovery algorithms including the proposed SDMP algorithm. The simulation is based on the channel matrix $\mathbf{H}$ of size 24×48 whose entries are from the independent Gaussian distribution $\mathcal{CN}(0, \frac{1}{M})$. We generate the K-sparse vector $\mathbf{x}$

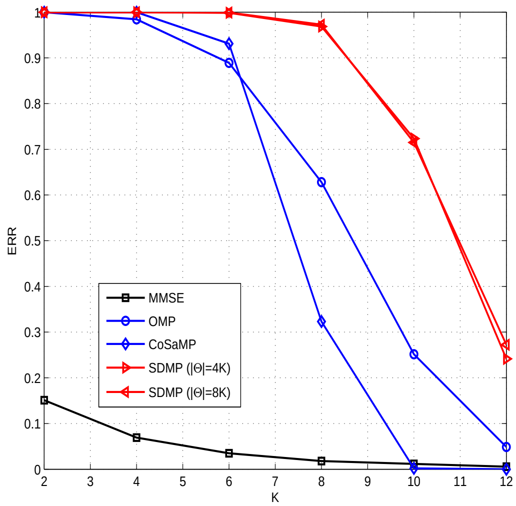

Fig. 2. ERR performance of sparse detection at SNR = 25 (dB).

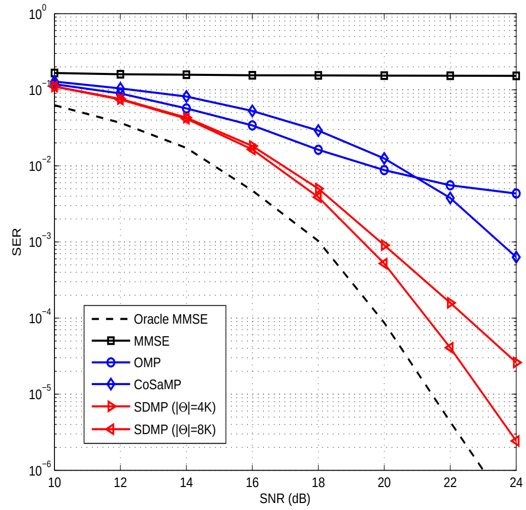

Fig. 3. SER performance of sparse detection when $K = 5$.

whose nonzero positions are randomly chosen. Elements of nonzero positions are chosen from 16-QAM symbol constellation. Two performance measures are used to evaluate the effectiveness of SDMP: 1) exact recovery ratio (ERR) and 2) symbol error rate (SER). In our simulations, we observe the performance of SDMP when $|\Theta| = 4K$ and $8K$. Other than SDMP, we test OMP, CoSaMP [5], LMMSE estimation, and Oracle LMMSE estimation[2]. Note that the Oracle estimator has the prior knowledge on the support and hence solves the problem using the overdetermined system model $\mathbf{y} = \mathbf{H}_T\mathbf{x}_T + \mathbf{v}$.

4.2. Simulation Results and Discussion

In Fig. 2, we plot the ERR performance when signal-to-noise ratio (SNR) is set to 25 dB. Since the system is underdetermined, we observe that the performance of LMMSE estimator exploiting whole channel matrix to estimate the signal vector is not working well. We also observe that the ERR of SDMP is higher than other approaches.

Fig. 3 provides the SER performance when $K = 5$ (10% of original signal is nonzero) as a function of SNR. Since multiple candidates are investigated, SDMP shows the best performance among all sparse recovery algorithms under test. Also, we observe that while performance improvement of conventional algorithms diminishes with SNR, performance of the proposed SDMP improves with SNR and maintains constant gap (around 1.5 dB when $|\Theta| = 8K$) from Oracle estimator.

[2] In all greedy algorithms under test, we added the slicing after each iteration.

5. REFERENCES

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Number V. 1 in Fundamentals of Statistical Signal Processing. Prentice-Hall PTR, 1998.

[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] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, Asilomar Conference on*, Nov. 1993, pp. 40–44.

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

[5] 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.