PIECEWISE SPARSE SIGNAL RECOVERY VIA PIECEWISE ORTHOGONAL MATCHING PURSUIT

Kezhi Li[123], *Cristian R. Rojas*[2], *Tao Yang*[2], *Håkan Hjalmarsson*[2], *Karl H. Johansson*[2], *Shuang Cong*[1*]
[1]Uni. of Sci. & Tech. of China, China; [2]Royal Institute of Technology (KTH), Sweden;
[3]Imperial College London, UK;
{kezhi, crro, taoyang, hjalmars, kallej}@kth.se, *scong@ustc.edu.cn

ABSTRACT

In this paper, we consider the recovery of piecewise sparse signals from incomplete noisy measurements via a greedy algorithm. Here piecewise sparse means that the signal can be approximated in certain domain with known number of nonzero entries in each piece/segment. This paper makes a two-fold contribution to this problem: 1) formulating a piecewise sparse model in the framework of compressed sensing and providing the theoretical analysis of corresponding sensing matrices; 2) developing a greedy algorithm called piecewise orthogonal matching pursuit (POMP) for the recovery of piecewise sparse signals. Experimental simulations verify the effectiveness of the proposed algorithms.

I. INTRODUCTION

Compressed sensing is capable of recovering sparse or compressible signals/coefficients from an underdetermined system of linear equations [1]. In practice, however piecewise sparsity rather than overall sparsity arises more naturally. For instance, suppose that a signal source is broadcasting radio programmes (or microwaves, underwater sound, etc.). In addition to the overall sparsity, the signal frequency is also sparse piecewisely in each spectrum band, because the whole frequency spectrum has to be allocated to L pieces for L different radio stations, and their programme signals are supposed to be sparse in each spectral region, which implies that the overall sparsity lever is $K = \sum_{i=1}^{L} K_i$, where K_i represents the sparsity in the ith piece. In this case a new sparse framework needs to be formulated.

In this paper we provide the analysis tools and an algorithm to solve the piecewise sparse recovery problem specifically. In particular, we address the problem in the following respects:

• Firstly, we formulate a piecewise sparse model in the framework of compressed sensing, and provide the theoretical analysis of the property of the qualified sensing matrix $\mathbf{D}$ for the recovery of a piecewise sparse signal $\mathbf{s}$. Our work is based on the results of the inner product of any two columns of $\mathbf{D}$. The result shows that tighter bounds of coherence and restricted isometry constant (RIC) of $\mathbf{D}$ can be derived theoretically for the piecewise sparse signals compared to the conventional sparse ones.

• Secondly, a greedy algorithm called piecewise orthogonal matching pursuit (POMP) is proposed specially aiming for recovering the piecewise sparse signals from few measurements. This algorithm, inspired by the OMP related algorithms [2]–[5], is capable of choosing the appropriate support of $\mathbf{s}$ and satisfying the constraints of piecewise sparsity $K_i, i \in \{1, \cdots., L\}$ simultaneously.

These two contributions establish an integrated solver for the practical recovery problem for piecewise sparse signals in both theory and algorithm design.

I-A. Relations to Prior Work

Sensing matrices have been studied widely by exploiting RIC and coherences to derive theoretical recovery bounds, such as for Toeplitz matrices [6], convolutional matrices [7] or general structured matrices [8]–[10]. Researchers have also exploited the coherence analyses regarding different sparsity models. The block-sparse model was first proposed in [4], and later extended by [5] [11], in which block orthogonal matching pursuit (BOMP) was developed. Clustered orthogonal matching pursuit (COMP) for clustered sparse model was considered in [3] but without the analysis of sensing matrices. In contrast to some previous works, this paper provides both the theoretical analysis and the POMP algorithm for reconstruction. The differences between general, block and piecewise sparsity are illustrated in Fig. 1.

I-B. Notation

Bold letters are used to denote a vector or a matrix. For vectors, $||\cdot||_1, ||\cdot||_2$ are the l_1 norm $||\mathbf{s}||_1 = \sum_j |s_j|$ and Euclidean norm $||\mathbf{s}||_2^2 = \sqrt{\mathbf{s}^*\mathbf{s}}$, respectively; $||\mathbf{s}||_0$ is the number of nonzero entries. For matrices, $\mathbf{D}^T$ and $\mathbf{D}^*$ denote the transpose, and Hermitian transpose of $\mathbf{D}$, respectively. The spectral norm of $\mathbf{D}$ is defined as $\rho(\mathbf{D}) = \sigma_{\max}(\mathbf{D}) = \sqrt{\lambda_{\max}(\mathbf{D}^*\mathbf{D})}$, where $\sigma_{\max}(\mathbf{D}), \lambda_{\max}(\mathbf{D}^*\mathbf{D})$ denote the largest singular value of $\mathbf{D}$ and the largest eigenvalue of $\mathbf{D}^*\mathbf{D}$. $[n]$ represents the set of $\{1, \cdots, n\}$.

II. PROBLEM SETUP

II-A. Piecewise Sparse

Suppose we need to recover a signal $\mathbf{x} \in \mathbb{R}^N$ that is corrupted by a Gaussian noise $\mathbf{e} \in \mathbb{R}^M \sim \mathcal{N}(0, \sigma^2)$, which

This work was supported by the National Natural Science Foundation of China under Grant No. 61573330, and Swedish Research Council, the Linnaeus Center ACCESS at KTH.

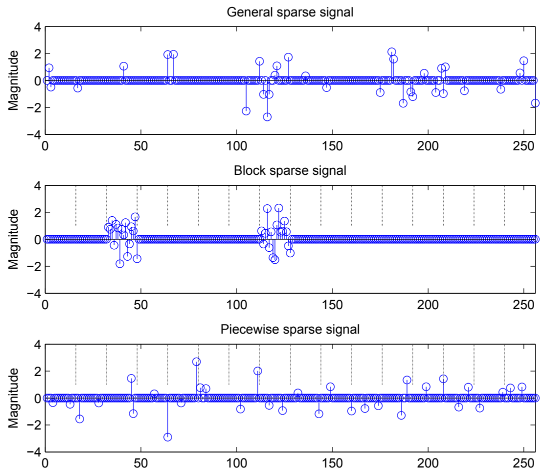

Fig. 1. Comparison of general sparse signal, block sparse signal and piecewise signals. All signals have length $N = 256$ and sparsity $K = 32$. Piecewise sparsity $K_i = 2, i \in \{1, \cdots, 16\}$.

can be written as

$$\mathbf{b} = \mathbf{R}_\Gamma \mathbf{x} + \mathbf{e}, \tag{1}$$

where $\mathbf{b} \in \mathbb{R}^M, M < N$ is the measurements, and $\mathbf{R}_\Gamma$ is a downsampling operator selecting random coefficients of $\mathbf{x}$ indexed by a set $\Gamma \subset [N], |\Gamma| = M$. Conventionally if $\mathbf{x}$ in (1) is K-sparse or can be approximated in the $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$ domain as

$$\mathbf{x} = \mathbf{\Psi}\mathbf{s}, ||\mathbf{s}||_0 \le K, K \ll N, \tag{2}$$

then $\mathbf{b} = \mathbf{D}\mathbf{s} + \mathbf{e}$, $\mathbf{s} \in \mathbb{R}^N, K < M < N$. According to the CS theory, $M \ge O(K\log(N/K))$ measurements are enough to recover $\mathbf{x}$ from (1) by exploiting reconstruction techniques, such as LASSO: $\hat{\mathbf{s}} = \arg_{\mathbf{s}} \min{(1/2)}||\mathbf{D}\mathbf{s}-\mathbf{b}||_2^2 + \lambda||\mathbf{s}||_1, \hat{\mathbf{x}} = \mathbf{\Psi}\hat{\mathbf{s}}$, when the sensing matrix $\mathbf{D} = \mathbf{R}_\Gamma\mathbf{\Psi}$ satisfies the restricted isometry property (RIP) with RIC δ_K defined in Def. 4 (see Appendix), and λ is a regulation parameter [12].

On the other hand, in many scenarios the signals that researchers are interested in are usually sparse in each piece [13], [14]. To address this issue, we consider the case that $\mathbf{x}$ is decomposed piecewise sparsely in the $\mathbf{\Psi}$ domain:

*Definition 1 (***Piecewise Sparse):** In (2), $\mathbf{s} \in \mathbb{R}^N$ is called (N, L, K)-piecewise sparse if $\mathbf{s}$ is a concatenation of L vectors with length $\mathbb{R}^d, d = N/L$, and each vector $\mathbf{s}^T[i], i \in [L]$ is K_i-sparse:

$$\mathbf{s} = [\underbrace{s_1 \cdots s_d}_{\mathbf{s}^T[1]}\underbrace{s_{d+1}\cdots s_{2d}}_{\mathbf{s}^T[2]}\cdots\underbrace{s_{N-d+1}\cdots s_N}_{\mathbf{s}^T[L]}]^T \tag{3}$$

where $\mathbf{s}^T[i]$ denotes the ith piece/block, $\sum_{i=1}^L K_i = K$. This definition is apparently different from conventional sparsity [15] and block-sparsity [5], so the existing unique recovery conditions are not directly applicable. The piecewise compressed sensing problem can be summarized as:

Problem: Assume that $\mathbf{s}$ is (N, L, K)-piecewise sparse. Given the CS model (1,2), how to measure the quality of $\mathbf{D}$, and how to recover $\mathbf{s}$ or $\mathbf{x}$ efficiently.

III. Piecewise Coherence

In CS, the coherence analysis provides an effective method to measure the quality of sensing matrices $\mathbf{D}$. Low coherence implies that the columns of $\mathbf{D}$ are almost mutually orthogonal and that leads to the unique recovery for $\mathbf{s}$ or $\mathbf{x}$. We assume that the l_2 norm of all columns of $\mathbf{D}$, $||\mathbf{d}_i||_2 = 1$ for $i \in [N]$ throughout the paper, where $\mathbf{d}_i$ is the ith column of $\mathbf{D}$. The coherence of $\mathbf{D}$ is defined as:

*Definition 2 (***Coherence [16], [17]):** The coherence of a matrix $\mathbf{D}$ can be defined as

$$\mu = \max_{i,j\neq i} |\mathbf{d}_i^*\mathbf{d}_j|. \tag{4}$$

When the RIC δ_{2K} of a matrix $\mathbf{D}$ satisfies $\delta_{2K} < 1/3$, every K-sparse vector $\mathbf{s} \in \mathbb{R}^N$ can be recovered by l_1 minimization [17], [18]. It has been proved that the RIC δ_K can be bounded by μ, so the unique recovery is guaranteed:

*Proposition 1 (***RIC bound [17]):** If $\mathbf{D}$ has unit-norm columns and coherence μ, then it satisfies the RIP of order K with δ_K

$$\delta_K \le (K-1)\mu \quad \text{for all } K < 1/\mu. \tag{5}$$

The bound in Prop. 1 is for conventional sparse signals. It is natural to seek for a tighter bound of δ_K by exploiting the piecewise sparsity in a piecewise sparse setting. Similar to (3), we represent $\mathbf{D}$ as a concatenation of columns $\mathbf{D}[l]$ of size $M \times d$:

$$\mathbf{D} = [\underbrace{\mathbf{d}_1\cdots\mathbf{d}_d}_{\mathbf{D}[1]}\underbrace{\mathbf{d}_{d+1}\cdots\mathbf{d}_{2d}}_{\mathbf{D}[2]}\cdots\underbrace{\mathbf{d}_{N-d+1}\cdots\mathbf{d}_N}_{\mathbf{D}[L]}], \tag{6}$$

where $\mathbf{D}[i]$ is the ith block matrix. With a slight abuse of notations, we also use $\mathbf{D}[i]$ to represent the columns set $\{\mathbf{d}_{(i-1)\cdot d+1}, \cdots, \mathbf{d}_{i\cdot d}\}, i \in [L]$. Then we define the inner coherence and the piecewise coherence of $\mathbf{D}$ as follows:

*Definition 3 (***Piecewise Coherence):** The (k, r)-inner coherence $\mathbf{v}_k(r)$ for the rth column and the piecewise coherence $\mu_{(K,L)}$ are defined as

$$\mathbf{v}_i(r) = \begin{cases} \max_{j_k\neq r}\sum_{k=1}^{K_i-1}|\mathbf{d}_r^*\mathbf{d}_{j_k}|, \mathbf{d}_{j_k}\in\mathbf{D}[i] & \text{if } \mathbf{d}_r \in \mathbf{D}[i], \\ \max\sum_{k=1}^{K_i}|\mathbf{d}_r^*\mathbf{d}_{j_k}|, \mathbf{d}_{j_k}\in\mathbf{D}[i] & \text{if } \mathbf{d}_r \notin \mathbf{D}[i], \end{cases}$$

$$\mu_{(K,L)} = \frac{1}{K-1}\max_{r\in[N]}\sum_{i=1}^{L}\mathbf{v}_i(r), \tag{7}$$

respectively.

As well as the coherence μ, the inner coherence $\mathbf{v}_k(r)$ and the piecewise coherence $\mu_{(K,L)}$ are also based on the inner product of two columns of $\mathbf{D}$, so $\mathbf{v}_k(r) \in [0, 1], \mu_{(K,L)} \in [0, 1]$. In addition, we derive their relationships through the following propositions.

Proposition 2: For a matrix $\mathbf{D}$ given in (6), set $\mathbf{G}_\Gamma = \mathbf{D}_\Gamma^*\mathbf{D}_\Gamma - \mathbf{I}_K, \mathbf{G}_\Gamma \in \mathbb{R}^{K\times K}$, where $\mathbf{D}_\Gamma$ is the sub-matrix of $\mathbf{D}$ indexed by the piecewise sparse set $\Gamma \subset [N]$, then

$$\mu_{(K,L)} = \frac{1}{K-1} \cdot \max_{\Gamma\subset[N],|\Gamma|\leq K} \max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}|, \qquad (8)$$

where $\mathbf{g}_{\Gamma_r}$ denotes the column of $\mathbf{G}_\Gamma$ corresponding to the rth column of $\mathbf{D}_\Gamma$.

Proof: The Gramian matrix $\mathbf{D}_\Gamma^*\mathbf{D}_\Gamma$ has all ones on its diagonal entries because the columns of $\mathbf{D}$ are normalized. Hence for $\mathbf{G}_\Gamma = \mathbf{D}_\Gamma^*\mathbf{D}_\Gamma - \mathbf{I}_K$ we only need to consider the off-diagonal entries. When Γ is fixed, from the definition of $\mathbf{g}_{\Gamma r}$,

$$\max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}| = \max_{r\in\Gamma} \sum_{j\in\Gamma\backslash\{r\}} |\mathbf{d}_r^*\mathbf{d}_j|. \qquad (9)$$

Taking the maximum operator over all choices of Γ on both sides of the equation,

$$\frac{1}{K-1} \max_{\substack{\Gamma\subset[N]\\|\Gamma|\leq K}} \max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}| = \frac{1}{K-1} \max_{\substack{\Gamma\subset[N]\\|\Gamma|\leq K}} \max_{r\in\Gamma} \sum_{j\in\Gamma\backslash\{r\}} |\mathbf{d}_r^*\mathbf{d}_j|$$
$$= \frac{1}{K-1} \max_{r\in[N]} \max_{\substack{r\notin\Gamma,\Gamma\subset[N]\\|\Gamma\backslash\{r\}|\leq K-1}} \sum_{j\in\Gamma} |\mathbf{d}_r^*\mathbf{d}_j| = \mu_{(K,L)}, \qquad (10)$$

where the last step follows because $\Gamma \subset [N]$ corresponds to the locations of piecewise nonzero coefficients in $\mathbf{s}$, $|\Gamma| = K = \sum_i K_i$, and

$$\max_{\Gamma\subset[N],|\Gamma|\leq K} \max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}| = \max_{\Gamma\subset[N],|\Gamma|=K} \max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}|. \qquad (11)$$

■

Remark 1: Prop. 2 establishes a relationship between the piecewise coherence and the Gramian matrix of sub-matrices of $\mathbf{D}$, which is useful in the proof of the following theorem.

Theorem 1: Denote δ'_K as the restricted isometry constant for the (N, L, K)-piecewise sparse signal $\mathbf{s}$, and $\mu_{(K,L)}, \mu$ as the piecewise coherence and coherence, respectively. Given $K \leq 1/\mu$, for a matrix $\mathbf{D}$ in (6) we have

$$0 \leq \delta'_K \leq (K-1)\mu_{(K,L)} \leq (K-1)\mu < 1. \qquad (12)$$

Proof: (a) To prove that $0 \leq \delta'_K \leq (K-1)\mu_{(K,L)}$. From Def. 4, δ'_K is the smallest constant such that

$$\delta'_K||\mathbf{s}||_2^2 \geq \left|||\mathbf{D}\mathbf{s}||_2^2 - ||\mathbf{s}||_2^2\right| \geq 0 \qquad (13)$$

holds for all (N, L, K)-piecewise sparse $\mathbf{s}$ for the reason in *Remark* 2. Taking the supremum over all possibilities of piecewise sparse $\mathbf{s} \in \mathbb{R}^N$ with its nonzero indexes $\Gamma \subset [N], |\Gamma| = K = \sum_i K_i$ and $||\mathbf{s}||_2^2 = 1$ normalized,

$$\begin{aligned}\delta'_K &= \sup_{\mathbf{s}} |((\mathbf{D}^*\mathbf{D} - \mathbf{I})\mathbf{s})^* \cdot \mathbf{s}| \\ &= \sup_{\mathbf{s}_\Gamma} |(\mathbf{G}_\Gamma\mathbf{s}_\Gamma)^* \cdot \mathbf{s}_\Gamma| \leq \sup(\rho(\mathbf{G}_\Gamma)) \\ &\leq \max_{\Gamma\subset[N],|\Gamma|\leq K} \max_{r\in\Gamma} |\mathbf{g}_{\Gamma r}| = (K-1)\mu_{(K,L)},\end{aligned} \qquad (14)$$

where $\mathbf{G}_\Gamma$ has the expression in Prop 2. The third step follows because $\mathbf{G}_\Gamma$ is Hermitian, $||\mathbf{s}||_2 = 1$. The last two steps are due to the Gershgorin's disk theorem [19] and Prop. 2.

(b) To prove that $(K-1)\mu_{(K,L)} \leq (K-1)\mu \leq 1$.

$$\begin{aligned}(K-1)\mu_{(K,L)} &= \max_{r\in[N]} \sum_{k=1}^{L} \mathbf{v}_k(r) \\ &\leq \left(\sum_i^L K_i - 1\right)\mu = (K-1)\mu < 1,\end{aligned} \qquad (15)$$

for $K < 1/\mu$ due to equation (4) and (7). The combination of (a) and (b) completes the proof. ■

Remark 2: 1) The piecewise restricted constant δ'_K is a special case of the general δ_K for piecewise sparse signals $\mathbf{s}$. Given the piecewise sparsity of $\mathbf{s}$ as prior knowledge, a sensing operator $\mathbf{D}_1$ with $\delta'_K(\mathbf{D}_1)$ inherits the properties of $\mathbf{D}_2$ with $\delta_K(\mathbf{D}_2)$ (such as the uniqueness of recovery) when $\delta'_K(\mathbf{D}_1) = \delta_K(\mathbf{D}_2)$, but the choice of $\mathbf{D}_1$ is more flexible than $\mathbf{D}_2$. From another perspective, for a matrix $\mathbf{D}$ the bound of $\delta'_K(\mathbf{D})$ is tighter than the bound of $\delta_K(\mathbf{D})$ in terms of the coherence, so a reconstruction with better accuracy is expected to be feasible for piecewise sparse signals if a proper algorithm is adopted.
2) It is possible to prove a tighter bound for δ, like $\delta \leq \max_{r\in[N]} \max_{\substack{\Gamma\subseteq[N]\backslash\{r\}\\|\Gamma|\leq K-1}} \sum_{j\in\Gamma} |\mathbf{d}_j^*\mathbf{d}_r| \leq (K-1)\mu$. However it is also easy to see that $(K-1)\mu_{(K,L)} \leq \max_{r\in[N]} \max_{\substack{\Gamma\subseteq[N]\backslash\{r\}\\|\Gamma|\leq K-1}} \sum_{j\in\Gamma} |\mathbf{d}_j^*\mathbf{d}_r|$, so δ' still has a tighter bound than δ.

IV. PIECEWISE ORTHOGONAL MATCHING PURSUIT

In this section, we develop a novel algorithm called piecewise-OMP (POMP) that can be used for reconstructing piecewise sparse signals efficiently. The algorithm uses the OMP framework that increases the support set by one in each iteration, and utilizes the information whose upper bound of number of nonzero entries is at most K_i in the piece $i \in [L]$. POMP behaves like regular OMP when the support number of each piece $\mathbf{s}_i$ is less than K_i for all $i \in [L]$. The main difference between POMP and OMP appears in how to avoid the circumstance that the number of columns selected from each block exceeds the sparsity constraints K_i during the support size increases. Specifically, when the support number in the ith piece $\mathbf{s}[i], i \in [L]$ increases gradually and reaches the limit K_i, POMP absorbs the final column index in the ith piece and disables other candidates of columns in $\mathbf{D}[i]$, which lets K_i stop at the upper bound of $||\mathbf{s}_i||_0$. This step is described by the {if, $\cdots$, end} paragraph below from line 6 to 8. The details of the steps for recovering piecewise sparse signals are demonstrated in Algorithm 1. Please also note that the algorithm cannot be implemented by applying OMP L times independently, because only one comprehensive measurement vector $\mathbf{b}$ is received.

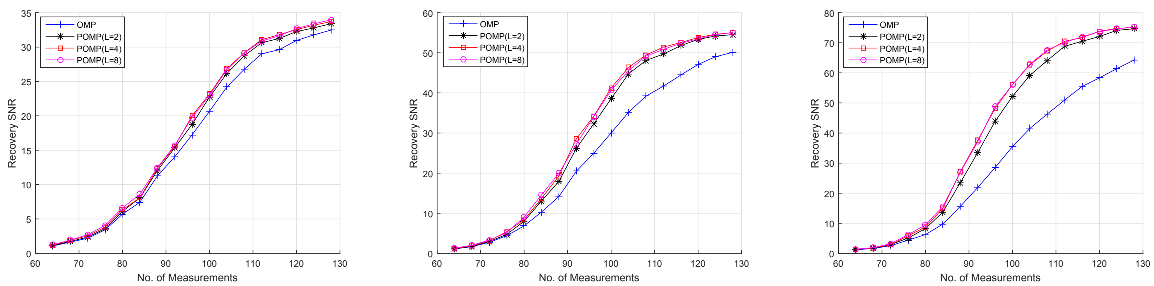

Fig. 2. Recovery performance of piecewise sparse signals with 10dB, 30dB, 50dB noise versus different number of measurements M using POMP in terms of SNR. The testing signal $\mathbf{s} \in \mathbb{R}^N$, $N = 256$, is generated randomly with the piecewise sparsity $\{K_i = i, i \in \{1, \cdots, 8\}\}$, $K = 36$.

Algorithm 1: Piecewise Orthogonal Matching Pursuit

Input: matrix $\mathbf{D} \in \mathbb{R}^{M\times N}$, measurements $\mathbf{b} \in \mathbb{R}^M$, sparsity bound K, piecewise sparsity $K_i, i \in [L]$.
Set: residual bound ξ.
Initialize: support set $\Gamma_0 = \varnothing$, abandon set $\Gamma_0' = \varnothing$, universe set $\Lambda = \{1, \cdots, N\}$, residual $\mathbf{r}_0 = \mathbf{b}$, solution $\mathbf{s}_0 = \mathbf{0}$, max number of iteration $k_{max} \le K$.
Iteration: 1. for $k = 1, \cdots, k_{max}$ do
2. $\mathbf{D}_k = \mathbf{D}_{\Lambda/\Gamma'_{k-1}}$ \\ update the available columns in $\mathbf{D}$
3. $\mathbf{h}_k = \mathbf{D}_k^T \mathbf{r}_{k-1}$ \\ form the residual signal estimate
4. $\Omega_k = \arg\max_j |\mathbf{h}_k(j)|$ \\ find the best support
5. $\Gamma_k = \Gamma_{k-1} \cup \Omega_k$ \\ update the support set
\\ abandon redundant columns
6. if $|\Gamma_{[\Omega_k]}| = K_i$ \\ identify the sparsity in the block,
7. $\Gamma'_k = \Gamma'_{k-1} \cup \{[\Omega_k]/(\Gamma_k \cap [\Omega_k])\}$ \\ add other \\ indexes in the block into the abandoned set
8. end \\ end if
9. $\mathbf{s}_k = \arg\min_{\mathbf{s}:\text{supp}(\mathbf{s})\subseteq\Gamma_k} ||\mathbf{b} - \mathbf{D}_k\mathbf{s}||_2$ \\ where supp($\mathbf{s}$) is the estimate nonzero support
10. $\mathbf{r}_k = \mathbf{b} - \mathbf{D}\mathbf{s}_k$ \\ calculate the residual by given $\mathbf{s}_k$
11. If $||\mathbf{r}_k||_2^2 \le \xi$ or $|\Gamma_k| = K$, break
12. end \\ end for
Output: $\hat{\mathbf{x}} = \mathbf{\Psi}\mathbf{s}_k$

In Algorithm 1, $[\Omega_k]$ represents the set of indexes in the ith block that contains Ω_k. $\{[\Omega_k]/(\Gamma_k \cap [\Omega_k])\}$ is the set of all column indexes in the block containing Ω_k that do not exist in the support set Γ_k. Therefore once the sparsity in a block i reaches its upper bound K_i, the algorithm will abandon other redundant columns in this block and move them into the abandoned column set Γ'. The algorithm continues until the stop criteria is satisfied. The choice of threshold ξ depends on the problems we are dealing with. Normally we set $\xi = 0.01||\mathbf{y}||_2^2$. We could also implement the subspace pursuit [20] instead of the matching pursuit to define the iterations aiming at accelerating the convergence of the algorithm.

V. Numerical Simulations

The aim of this section is to examine the proposed algorithms in recovering signals by taking piecewise sparsity explicitly into account. The sensing matrix $\mathbf{D} \in \mathbb{R}^{M\times N}$ whose entries are drawn from a random i.i.d. Gaussian distribution, is normalized so that the resulting columns have norm 1. The true piecewise sparse signal $\mathbf{s} \in \mathbb{R}^N$ is generated with length $N = 256$. More specifically, $\mathbf{s}$ is divided into $L = 8$ consecutive pieces with each size $N/L = 32$, and the piecewise sparsity of the ith piece is a varying value $K_i = i, i \in \{1, \cdots, 8\}$. Hence the overall sparsity of $\mathbf{s}$ is $K = \sum_{i=1}^{8} K_i = 36$. $\mathbf{s}$ has entries with i.i.d. Gaussian magnitudes on a randomly chosen support set following a uniform distribution in each piece with constraints on sparsity K_i. When $L = 2$, we set $K_1 = \sum_{i=1}^{4} i = 10$ and $K_2 = \sum_{i=5}^{8} i = 26$. Similar settings are implemented when $L = 4$. The recovery results in terms of signal to noise ratio (SNR) versus different number of measurements are shown in Fig. 2 for input noises SNR$=$ $10, 30, 50$dB, respectively. Each data point is based on the average of 1000 iterations. We can see that when the number of measurements increases, the performance of recovery becomes better in terms of SNR, and the recovery results of POMP outperform OMP by about $2, 5, 10$dB in SNR in different circumstances of noises. In addition the performance of knowing more information about the sparsity (with larger L) gives a little better reconstruction.

VI. Conclusions

This paper considered the recovery of piecewise sparse signals from noisy incomplete measurements, using the techniques of compressed sensing. We provided a whole set of analysis tools with tighter coherence and a novel greedy algorithm to solve the piecewise sparse recovery problem specifically. The simulations showed the superiority of the proposed POMP over conventional OMP clearly in terms of reconstruction SNR.

VII. Appendix

Definition 4 (**Restricted Isometry Constant (RIC)**): The restricted isometry constant δ_K of a $\mathbf{D} \in \mathbb{R}^{M\times N}$ is defined as the smallest $\delta_K > 0$ such that

$$(1 - \delta_K)||\mathbf{s}||_2^2 \le ||\mathbf{D}\mathbf{s}||_2^2 \le (1 + \delta_K)||\mathbf{s}||_2^2 \qquad (16)$$

for all K-sparse $\mathbf{s} \in \mathbb{R}^N$. If a matrix $\mathbf{D}$ has a RIC δ_K, it is equivalent to say that $\mathbf{D}$ satisfies the restricted isometry property of order K with δ_K.

VIII. REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, pp. 1289–1306, Jul. 2006.

[2] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2231–2242, Oct. 2004.

[3] B. Sanandaji, T. Vincent, and M. Wakin, "Compressive topology identification of interconnected dynamic systems via clustered orthogonal matching pursuit," in *Proceedings. 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Dec. 2011, pp. 174–180.

[4] Y. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Processing*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[5] Y. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[6] J. Haupt, W. Bajwa, G. Raz, and R. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5862–5875, Nov. 2010.

[7] K. Li, L. Gan, and C. Ling, "Convolutional compressed sensing using deterministic sequences," *IEEE Trans. Signal Processing*, vol. 61, no. 3, pp. 740–752, 2013.

[8] H. Rauhut, "Compressive sensing and structured random matrices," *Radon Series Comp. Appl. Math.*, Theoretical Foundations and Numercial Methods for Sparse Recovery, Publisher: DeGruyter, vol. 9, pp. 1-92, 2010.

[9] N. Shervashidze and F. Bach, "Learning the structure for structured sparsity," *IEEE Trans. Signal Processing*, vol. 63, no. 18, pp. 4894–4902, Sep. 2015.

[10] K. Li and S. Cong, "State of the Art and Prospects of Structured Sensing Matrices in Compressed Sensing," *Frontier of Computer Science*, vol. 9, no. 5, pp. 665-677, 2015.

[11] S. Jiang and Y. Gu, "Block-sparsity-induced adaptive filter for multi-clustering system identification," *IEEE Trans. Signal Processing*, vol. 63, no. 20, pp. 5318–5330, Oct. 2015.

[12] E. Candès and J. Romberg, "Quantitative robust uncertainty principles and optimally sparse decompositions," *Found. Comput. Math.*, vol. 6, no. 8, pp. 227–254, 2006.

[13] M. Mishali and Y. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[14] D. Angelosante and G. Giannakis, "Group lassoing change-points in piecewise-stationary AR signals," in *17th International Conference on Digital Signal Processing (DSP),*, Jul. 2011, pp. 1–8.

[15] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, pp. 969–985, Jun. 2007.

[16] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[17] H. Boche, R. Calderbank, G. Kutyniok, and J. Vybiral, "A survey of compressed sensing," available online.

[18] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure and Applied Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[19] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[20] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity," *IEEE Trans. Inform. Theory*, vol. 55, pp. 2230–2249, May 2009.