A Two Stage Generalized Block Orthogonal Matching Pursuit (TSGBOMP) Algorithm

Samrat Mukhopadhyay, *Member, IEEE*, and Mrityunjoy Chakraborty, *Senior Member, IEEE*

***Abstract*—Recovery of an unknown sparse signal from a few of its projections is the key objective of compressed sensing. Often one comes across signals that are not ordinarily sparse but are sparse blockwise. Existing block sparse recovery algorithms like BOMP make the assumption of uniform block size and known block boundaries, which are, however, not very practical in many applications. This paper addresses this problem and proposes a two step procedure, where the first stage is a coarse block location identification stage while the second stage carries out finer localization of a non-zero cluster within the window selected in the first stage. A detailed convergence analysis of the proposed algorithm is carried out by first defining a so-called pseudoblock-interleaved block RIP for the given generalized block sparse signal and then imposing upper bounds on the corresponding RIC. Simulation results confirm significantly improved performance of the proposed algorithm as compared to BOMP.**

***Index Terms*—Compressed Sensing (CS), block sparse, restricted isometry property (RIP).**

I. Introduction

Though last two decades have seen a great flurry of activities in the domain of compressed sensing (CS) ([1], [2]), in particular towards development of *greedy* algorithms for recovering an unknown vector that is *sparse* in some domain, almost all these treatments considered signals that are ordinarily sparse and not *block sparse* where the nonzero elements tend to occur in clusters ([3], [4]), even though block sparse signals are encountered in many applications ([5]–[8]).

In [3], the CS recovery problem with block sparse structure was studied in detail and a l_1-norm minimization based recovery algorithm was proposed which was analyzed by introducing the notion of the so-called block restricted isometry property (BRIP) of the sensing matrix. In [7], a Block OMP (BOMP) algorithm was proposed as an extension of OMP [9] for block-sparse recovery from compressed measurements, which was analyzed using the concept of block-coherence. Recently, a BRIP based analysis of the BOMP algorithm has been presented [8] that derives recovery conditions in terms of the block restricted isometry constant (BRIC) of the sensing matrix. Both the block sparse recovery algorithms ([3], [7]), however, assume that the overall support of the signal vector can be *uniformly* partitioned into blocks of identical length and of known block boundaries. These assumptions, however, do not hold good in many applications, e.g., atomic decomposition of audio signals [10], where the exact block partitions of the unknown vector are not known beforehand and a non-zero cluster can overlap partially or fully into two or more adjacent blocks. Although a few algorithms have been proposed to address the recovery of this kind of signals ([11], [12]), all of them use the Bayesian learning framework, which impose prior distributional assumptions on the unknown vector.

In this paper, we propose and study a new non-Bayesian algorithm called the two stage generalized block OMP (TSGBOMP), which has a structure similar to the BOMP algorithm except that the block identification is performed in two stages. The first stage is a coarse block location identification stage where, similar to the BOMP algorithm, from a prespecified set of *windows* (i.e. sets of consecutive columns that the whole set of columns is partitioned into), a window of columns having the maximum correlation (with some prior residual) is selected. In the second stage, the algorithm conducts a finer search for a block by calculating the correlations (with some prior residual) corresponding to all overlapping consecutive cluster of columns throughout the window selected, and then finding the cluster having the largest absolute correlation[1]. Presence of the two stages, however, makes analysis of the proposed TSGBOMP algorithm very challenging, as, success at each iteration requires first selection of a window that overlaps with at least one true cluster and then in the second stage, identification of that true cluster. For this, we introduce a new RIP tailored to the particular structure of the unknown vector, termed as pseudoblock-interleaved block RIP (PIBRIP) and study the properties of the corresponding pseudoblock-interleaved block RIC (PIBRIC). The PIBRIP is motivated by the model-RIP introduced in [14] for analyzing signals with union of subspace structure. We carry out a detailed PIBRIP based analysis of the TSGBOMP algorithm and determine conditions on the PIBRIC that ensures exact recovery using TSGBOMP. Finally, we use numerical simulations to exhibit the superior probability of recovery performances of the TSGBOMP

Manuscript received August 20, 2020; revised April 14, 2021 and August 24, 2021; accepted September 13, 2021. Date of publication September 24, 2021; date of current version November 4, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xue Jiang. *(Corresponding author: Mrityunjoy Chakraborty.)*

Samrat Mukhopadhyay is with the Department of Electronics Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India (e-mail: samrat@iitism.ac.in).

Mrityunjoy Chakraborty is with the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur 721302, India (e-mail: mrityun@ece.iitkgp.ac.in).

This article has supplementary downloadable material available at https://doi.org/10.1109/TSP.2021.3114977, provided by the authors.

Digital Object Identifier 10.1109/TSP.2021.3114977

[1]This philosophy of block selection in the second stage is inspired from a recent paper [13] which searches for a block by calculating absolute correlations corresponding to all possible overlapping clusters of columns in the matrix with certain residual vector and then selecting the one having the highest value.

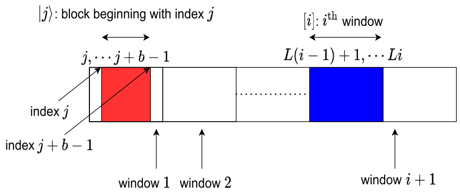


Fig. 1. Explanation of the notations $[i]$ and $|i\rangle$.

algorithm with respect to the BOMP algorithm for recovering the signals with generalized block sparse structure.

II. Notations

The following notations have been used throughout the paper : 't' in superscript indicates matrix / vector transpose, $\mathcal{H}$ denotes the set of indices $\{1, 2, \ldots, n\}$. For any vector $\boldsymbol{x} \in \mathbb{R}^n$, the entries are denoted by $x_1, \ldots, x_n$ and the *support* of $\boldsymbol{x}$, denoted by $\texttt{supp}(\boldsymbol{x})$, is defined as the set of indices corresponding to the nonzero values of $\boldsymbol{x}$, i.e., $\texttt{supp}(\boldsymbol{x}) = \{i \in \mathcal{H} | x_i \neq 0\}$[2]. The symbol $\boldsymbol{\phi}_i$ denotes the i th column of $\boldsymbol{\Phi}$, $i \in \mathcal{H}$ and all the columns of $\boldsymbol{\Phi}$ are assumed to have unit l_2 norm, i.e., $\|\boldsymbol{\phi}_i\|_2 = 1$, which is a common assumption in literature [16]. For any two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^t \boldsymbol{v}$. For any $S \subseteq \mathcal{H}$, $\boldsymbol{x}_S$ denotes a vector comprising those entries of $\boldsymbol{x}$ that are indexed by numbers belonging to S. Similarly, $\boldsymbol{\Phi}_S$ denotes the submatrix of $\boldsymbol{\Phi}$ formed with the columns of $\boldsymbol{\Phi}$ having column numbers given by the index set S. We denote by $[i]$, the set of indices $\{(i-1)L+1, \ldots, Li\}$, for any $i = 1, 2, \ldots, \frac{n}{L}$ (throughout the paper, we assume n to be divisible by L), which is hereafter referred to as the i^{th} *window* or the window with index i. For a set $S \subseteq \mathcal{H}$, we denote by $[S]$ the set union of the windows indexed by the elements of S, i.e., $[S] = \cup_{i \in S}[i]$. A set of $b(\geq 1)$ consecutive indices is called a *block*. We denote by $|i\rangle$ the set of indices $\{i, i+1, \ldots, i+b-1\}$, which is hereafter referred to as the block with starting index i. Analogous to the windows, for any set $S \subseteq \mathcal{H}$, we define $|S\rangle = \cup_{i \in S} |i\rangle$. Fig. 1 provides a pictorial description of the above notations. In this paper, we consider blocks that are non-overlapping though they can be adjacent. It is assumed that the unknown vector $\boldsymbol{x}$ contains at most K nonzero blocks[3] with starting indices $t_1, t_2, \ldots, t_K$, where $t_1 = \min\{j \geq 1 | x_j \neq 0\}$, and $t_{k+1} = \min\{j \geq t_k + b | x_j \neq 0\}$, $1 \leq k \leq K-1$. Using these indices, we define $T = \{t_1, t_2, \ldots, t_K\}$.

If $\boldsymbol{\Phi}_{|S\rangle}$ has full column rank of $|S| \times b$ ($|S| \times b < m$), then the Moore-Penrose pseudo-inverse of $\boldsymbol{\Phi}_{|S\rangle}$ is given by $\boldsymbol{\Phi}^{\dagger}_{|S\rangle} = (\boldsymbol{\Phi}^t_{|S\rangle}\boldsymbol{\Phi}_{|S\rangle})^{-1}\boldsymbol{\Phi}^t_{|S\rangle}$. The matrices $\mathbf{P}_S = \boldsymbol{\Phi}_{|S\rangle}\boldsymbol{\Phi}^{\dagger}_{|S\rangle}$ and $\mathbf{P}^{\perp}_S = \boldsymbol{I} - \mathbf{P}_S$ respectively denote the orthogonal projection operators associated with $\texttt{span}(\boldsymbol{\Phi}_{|S\rangle})$ and the orthogonal complement of $\texttt{span}(\boldsymbol{\Phi}_{|S\rangle})$. Finally, for any matrix $\boldsymbol{A}$, we denote by $\|\boldsymbol{A}\|_{2\to 2}$ the operator norm of $\boldsymbol{A}$ defined as $\|\boldsymbol{A}\|_{2\to 2} = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$, and can be shown to be equivalent to $\max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{|\boldsymbol{x}^t\boldsymbol{A}\boldsymbol{x}|}{\boldsymbol{x}^t\boldsymbol{x}}$ when $\boldsymbol{A}$ is symmetric [2, pp. 519]. We use the abbreviation w.l.o.g. for without loss of generality.

[2]The treatment in this paper is restricted to the real case only. For extension to the complex case, some additional steps and one more lemma are required. Details of this can be found in [15].

[3]By nonzero block of $\boldsymbol{x}$ is meant a block over which $\boldsymbol{x}$ has non-zero values.

TABLE I
Proposed TSGBOMP Algorithm

Input: Measurement vector $\boldsymbol{y} \in \mathbb{R}^m$, sensing matrix $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$; sparsity level K; prespecified residual threshold ϵ; window size ($L(\geq bp)$); maximum cluster size $B = pb,\ p \geq 1$.
Initialize: Counter $k = 0$, residue $\boldsymbol{r}^0 = \boldsymbol{y}$, estimated support set, $T^0 = \emptyset$.
While ($\|\boldsymbol{r}^k\|_2 \geq \epsilon$ and $k < K$)
1) $k = k+1$,
2) $w^k = \arg\max_{1\leq l \leq n/L} \left\|\boldsymbol{\Phi}^t_{[l]}\boldsymbol{r}^{k-1}\right\|_2$ and $b^k = \{\max\{1, L(w^k-1)+1-(B-1)\}, \cdots, Lw^k\}$,
3) $i^k = \arg\max_{i\in b^k} \left\|\boldsymbol{\Phi}^t_{|S\rangle}\boldsymbol{r}^{k-1}\right\|_2$, where $S = \{i,\ i+b, \cdots,\ i+(p-1)b\}$;
4) $T^k = T^{k-1} \cup h^k$, where $h^k = \{i^k,\ i^k + b, \cdots,\ i^k + (p-1)b\}$;
5) $\boldsymbol{x}^k_{|T^k\rangle} = \arg\min_{\boldsymbol{u}} \left\|\boldsymbol{y} - \boldsymbol{\Phi}_{|T^k\rangle}\boldsymbol{u}\right\|_2$,
6) $\boldsymbol{r}^k = \boldsymbol{y} - \boldsymbol{\Phi}_{|T^k\rangle}\boldsymbol{x}^k_{|T^k\rangle}$.
End While

Output: Estimated support set $\{T^K\}$ and estimated vector $\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{u}:\texttt{supp}(\boldsymbol{u})=\{T^K\}} \left\|\boldsymbol{y} - \boldsymbol{\Phi}_{|T^K\rangle}\boldsymbol{u}\right\|_2$.

III. Proposed Algorithm

The proposed TSGBOMP algorithm aims to recover an unknown vector $\boldsymbol{x}$, and more specifically, to recover its support $|T\rangle$ that has a maximum of K nonzero blocks each of size b. However, unlike conventional approaches like the BOMP [7], it does not assume the exact block locations to be known a priori. It is assumed that there can be at most p adjacent non-zero blocks in $\boldsymbol{x}$, forming a nonzero *cluster* of maximum size $B(= pb)$.[4] The nonzero clusters are not contiguous (i.e., they are separated by zeros), and if there are a total of r such nonzero clusters, with the s^{th} cluster having size $j_s b$, $1 \leq j_s \leq p$, $s = 1, 2, \ldots, r$ (i.e., it has j_s contiguous blocks of size b each), then $\sum_{s=1}^r j_s = K$. The whole signal range is divided into n/L windows of size L, with B, L satisfying $L \geq B$. It is also assumed that any two consecutive nonzero clusters of $\boldsymbol{x}$ are *well-separated* by a zone of at least $L' = L + 2B - b$ zeros. Although, in principle, such a constraint is not necessary for the execution of the algorithm, it ensures that the range of indices, associated to an window identified by the first stage of TSGBOMP, can contain only one nonzero true cluster. This makes the analysis of the algorithm less complicated. Also, we assume that the signal length n is large enough to accommodate any arrangement of nonzero clusters with K blocks in the signal (of size b) such that any two consecutive clusters are separated by at least L' zeros.

The proposed TSGBOMP algorithm, given in Table I, employs a two stage search procedure, of which the first one is similar to the BOMP. At any iteration $k(\geq 1)$ of the algorithm, it assumes that a residual vector $\boldsymbol{r}^{k-1}$ and a partially

[4]In this paper, we use the notion of cluster in most cases rather than block, as the former is more general (a block is a cluster with $p = 1$).

constructed set T^{k-1} are already available from step $k-1$ ($\boldsymbol{r}^0 = \boldsymbol{y}$, $T^0 = \emptyset$). Then, following the BOMP procedure, it carries out a *window-wise scanning* and identifies a window of length L and index w^k from the range 1 to n/L, for which the l_2 norm of the correlation vector $\boldsymbol{\Phi}^t_{[w^k]}\boldsymbol{r}^{k-1}$ is maximum. Next, it carries out a *pointwise scanning* over the set of indices $b^k = \{\max\{1, L(w^k-1)+1-(B-1)\}, \ldots, Lw^k\}$ and identifies a cluster of size B that has non-empty overlap with the chosen window and for which, the correlation vector $\boldsymbol{\Phi}^t_{|h^k\rangle}\boldsymbol{r}^{k-1}$ has maximum l_2 norm, where h^k denotes the set of the first indices of the elementary blocks (of size b) contained in the cluster, i.e., $h^k = \{i^k,\, i^k+b, \ldots,\, i^k+(p-1)b\}$, with $i^k \in b^k$. The set h^k is then appended to T^{k-1} to construct T^k, and the residual vector is updated to $\boldsymbol{r}^k$ by computing $\mathbf{P}^{\perp}_{T^k}\boldsymbol{y}$.

IV. Signal Recovery Using TSGBOMP

In this section, we derive sufficient recovery condition for TSGBOMP to successfully reconstruct the correct support of an unknown vector $\boldsymbol{x} \in \mathbb{R}^n$ from a set of noisy measurements, given by $\boldsymbol{y}(\in \mathbb{R}^m) = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{e}$, where $\boldsymbol{e}$ is a measurement noise vector that is assumed to be l_2-bounded by a positive number ϵ, i.e., $\|\boldsymbol{e}\|_2 \le \epsilon$. For this, we follow the approach of induction, i.e., at any k-th step of iteration ($k \ge 1$), we assume that in each of the previous $(k-1)$ steps, at least one true cluster of $\boldsymbol{x}$ has been selected. We then first find a condition ensuring that TSGBOMP identifies a window which has a non-empty intersection with one of the true (as yet unidentified) clusters of $\boldsymbol{x}$. Then we find a condition that ensures that a correct cluster (of size between b and pb) is identified among all possible consecutive clusters having nonzero overlap with the identified window. Finally we combine these two conditions to find a uniform recovery condition under which TSGBOMP is successful in identifying a correct cluster at step k. For the analysis of finding a correct window, our proof closely follows the arguments of Theorem 1 of [8] which is widely used for the analysis of BOMP using block sparse structure with predefined block boundaries and uniform block lengths. However, as the TSGBOMP does not assume fixed block boundaries and uniform block length, certain steps in the analysis of Wen *et al.* [8] cannot be directly extended to TSGBOMP, and instead, certain novel structures are required to be defined to continue adopting the analysis of [8]. These significantly change the final conditions for successful recovery.

A. Condition for Identifying a Correct Window At Step $k(K \ge 1)$

To ensure success at iteration k, we need to ensure that the window selected in iteration k has a nonempty overlap with at least one of the true clusters of $\boldsymbol{x}$, having blocks with first indices in the set $T \setminus T^{k-1}$. To express this mathematically, let us define, for any subset $S \subseteq \mathcal{H}$, $O_S = \{i : [i] \cap |S\rangle \neq \emptyset\}$, which is the set containing the first indices of the windows that have nonempty overlap with the blocks beginning with the indices in the set S. Then the necessary condition for selecting a correct window at iteration k is given by

$$\max_{i \in O_{T\setminus T^{k-1}}} \left\|\boldsymbol{\Phi}^t_{[i]}\boldsymbol{r}^{k-1}\right\|_2 > \max_{j \in O^C_{T\setminus T^{k-1}}} \left\|\boldsymbol{\Phi}^t_{[j]}\boldsymbol{r}^{k-1}\right\|_2,$$

$$\Leftrightarrow \left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\boldsymbol{r}^{k-1}\right\|_{2,\infty} > \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\boldsymbol{r}^{k-1}\right\|_{2,\infty}, \tag{1}$$

where $S^{k-1} = T \setminus T^{k-1}$. To find a condition to ensure (1), we first observe that using steps similar to the ones used to derive Eq (4.16) of [8], one obtains

$$\begin{aligned}
\boldsymbol{r}^{k-1} &= \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{y} \\
&= \mathbf{P}^{\perp}_{T^{k-1}}(\boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{e}) \\
&= \mathbf{P}^{\perp}_{T^{k-1}}\left(\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle} + \boldsymbol{e}\right) \\
&= \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle} + \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}.
\end{aligned} \tag{2}$$

Plugging in the expression of $\boldsymbol{r}^{k-1}$ from (2), it is easy to see (using triangle and reverse triangle inequalities respectively) that the condition (1) is satisfied if the following is ensured:

$$\begin{aligned}
&\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&- \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&> \left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty} + \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty}.
\end{aligned} \tag{3}$$

Now we will find a lower bound of the left hand side (LHS) and upper bound of the right hand side (RHS), of the inequality (3) and compare them to come up with a sufficient condition to ensure (3).

In order to proceed further, we first note that one can write

$$\begin{aligned}
&\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&= \frac{\sum_{i\in O_{S^{k-1}}}\left\|\boldsymbol{x}_{[i]}\right\|_2}{\sum_{i\in O_{S^{k-1}}}\left\|\boldsymbol{x}_{[i]}\right\|_2}\cdot \\
&\quad\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&\overset{(a)}{\ge} \frac{\sum_{i\in O_{S^{k-1}}}\left|\boldsymbol{x}^t_{[i]}\boldsymbol{\Phi}^t_{[i]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right|}{\sum_{i\in O_{S^{k-1}}}\left\|\boldsymbol{x}_{[i]}\right\|_2} \\
&\overset{(b)}{\ge} \frac{\left|\left\langle\boldsymbol{\Phi}_{[O_{S^{k-1}}]}\boldsymbol{x}_{[O_{S^{k-1}}]}, \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\rangle\right|}{\sum_{i\in O_{S^{k-1}}}\left\|\boldsymbol{x}_{[i]}\right\|_2} \\
&\overset{(c)}{=} \frac{\left|\left\langle\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}, \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\rangle\right|}{\sum_{i\in O_{S^{k-1}}}\left\|\boldsymbol{x}_{[i]}\right\|_2} \\
&\overset{(d)}{\ge} \frac{\left|\left\langle\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}, \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\rangle\right|}{\sqrt{d_k}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \\
&= \frac{\left\|\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2}{\sqrt{d_k}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2},
\end{aligned} \tag{4}$$

where $d_k = |O_{S^{k-1}}|$. Here step (a) uses the Hölder's-$(1,\infty)$ inequality followed by Cauchy-Schwartz inequality, while (b) uses the triangle inequality. Step (c) uses the observation that

$|S^{k-1}\rangle \subset [O_{S^{k-1}}]$ along with the fact that a window can contain at most one nonzero cluster (of size jb, $1 \le j \le p$), since such nonzero clusters are separated by at least L' zeros. Finally, step (d) uses Cauchy-Schwartz inequality and the observation that $\|\boldsymbol{x}_{[O_{S^{k-1}}]}\|_2 = \|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2$.

We now proceed to find an upper bound of $\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\|_{2,\infty}$. In order to do that, following the proof of Lemma 4.1 of [8], we will instead find an upper bound of $\|\boldsymbol{\Phi}^t_{[j]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\|_2$ for an arbitrary $j \in O^C_{S^{k-1}}$ which will hold uniformly for all $j \in O^C_{S^{k-1}}$. Fix any $j \in O^C_{S^{k-1}}$ and let

$$\boldsymbol{q}^{k-1} = \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}. \tag{5}$$

Following the proof of Lemma 4.1 of [8], we now define a list of quantities for expressing $\|\boldsymbol{\Phi}^t_{[j]}\boldsymbol{q}^{k-1}\|_2$ in a convenient way. Let $\theta > 0$ be an arbitrary positive number. Define,

$$\mu = -\frac{\sqrt{\theta+1}-1}{\sqrt{\theta}}, \tag{6}$$

$$h_l = \frac{\left(\boldsymbol{\Phi}_{[j]}\right)^t_l \boldsymbol{q}^{k-1}}{\left\|\boldsymbol{\Phi}^t_{[j]}\boldsymbol{q}^{k-1}\right\|_2},\ 1 \le l \le L, \tag{7}$$

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{x}_{|S^{k-1}\rangle} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{||S^{k-1}\rangle|+L}, \tag{8}$$

$$\boldsymbol{w} = \mu \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 \cdot \begin{bmatrix} \mathbf{0} \\ \boldsymbol{h} \end{bmatrix} \in \mathbb{R}^{||S^{k-1}\rangle|+L}, \tag{9}$$

$$\boldsymbol{B} = \mathbf{P}^{\perp}_{T^{k-1}}\left[\boldsymbol{\Phi}_{|S^{k-1}\rangle} \quad \boldsymbol{\Phi}_{[j]}\right]. \tag{10}$$

Using these definitions, it is straightforward to verify that $\boldsymbol{B}\boldsymbol{u} = \boldsymbol{q}^{k-1}$ and $\boldsymbol{B}\boldsymbol{w} = \mu\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{[j]}\boldsymbol{h}$, so that

$$\begin{aligned}&\boldsymbol{w}^t\boldsymbol{B}^t\boldsymbol{B}\boldsymbol{u}\\ &= \mu\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 \boldsymbol{h}^t\boldsymbol{\Phi}^t_{[j]}(\mathbf{P}^{\perp}_{T^{k-1}})^t\boldsymbol{q}^{k-1}\\ &= \mu\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 \boldsymbol{h}^t\boldsymbol{\Phi}^t_{[j]}\boldsymbol{q}^{k-1}\\ &(\because (\mathbf{P}^{\perp}_{T^{k-1}})^t\mathbf{P}^{\perp}_{T^{k-1}} = \mathbf{P}^{\perp}_{T^{k-1}})\\ &= \mu\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2\left\|\boldsymbol{\Phi}^t_{[j]}\boldsymbol{q}^{k-1}\right\|_2.\end{aligned}$$

Consequently, it can be verified that the following identity holds (see the proof of Lemma 4.1 of [8] for details):

$$\begin{aligned}&\|\boldsymbol{B}(\boldsymbol{u}+\boldsymbol{w})\|_2^2 - \left\|\boldsymbol{B}(\mu^2\boldsymbol{u}-\boldsymbol{w})\right\|_2^2\\ &= (1-\mu^4)\Big(\|\boldsymbol{B}\boldsymbol{u}\|_2^2 -\\ &\sqrt{\theta}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2\left\|\boldsymbol{\Phi}^t_{[j]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2\Big).\end{aligned} \tag{11}$$

Clearly, if a lower bound of the LHS of (11) can be found, one can find an upper bound of $\|\boldsymbol{\Phi}^t_{[j]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\|_2$. Consider the two terms of the LHS of (11), i.e., $\|\boldsymbol{B}(\boldsymbol{u}+\boldsymbol{w})\|_2^2 \equiv \|\mathbf{P}^{\perp}_{T^{k-1}}[\boldsymbol{\Phi}_{|S^{k-1}\rangle} \quad \boldsymbol{\Phi}_{[j]}](\boldsymbol{u}+\boldsymbol{w})\|_2^2$, and $\|\boldsymbol{B}(\mu^2\boldsymbol{u}-\boldsymbol{w})\|_2^2 \equiv \|\mathbf{P}^{\perp}_{T^{k-1}}[\boldsymbol{\Phi}_{|S^{k-1}\rangle} \quad \boldsymbol{\Phi}_{[j]}](\mu^2\boldsymbol{u}-\boldsymbol{w})\|_2^2$. In the case of BOMP, both the window and block are the same entity, of length, say, L, meaning, $\boldsymbol{\Phi}_{|S^{k-1}\rangle}$ has an integral multiple of L (i.e., $|S^{k-1}|L$) number of columns, while $\boldsymbol{\Phi}_{[j]}$ has L columns.

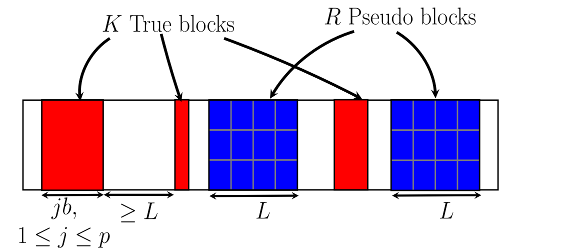


Fig. 2. Illustration of a PIBS vector: there are $k (\le K)$ true clusters (shown in solid), each of size of the form jb, where $1 \le j \le p$, and any two consecutive true clusters are separated by at least L' zeros; these gaps (as well as the gaps between the signal edges and the first and last true clusters) contain $r(\le R)$ another type of nonzero blocks (shown in grids), called pseudo blocks, which have fixed size l $(0 \le l \le L')$.

Together, $\boldsymbol{\Phi}_{|S^{k-1}\rangle}$ along with $\boldsymbol{\Phi}_{[j]}$ constitute the columns of $\boldsymbol{\Phi}$ corresponding to the support of a conventional block sparse vector of block sparsity $|S^{k-1}|+1$ and block size L. As a result, in the analysis of BOMP by Wen *et al* [8], the properties of block RIP (See Sections 1 and 2 of [8]) could be leveraged to find an upper bound of the LHS of (11). However, in the case of TSG-BOMP, the columns of $\boldsymbol{\Phi}_{|S^{k-1}\rangle}$ correspond to the indices of the as yet unidentified true non-zero blocks of $\boldsymbol{x}$ (totalling $|S^{k-1}|b$ columns) and the columns of $\boldsymbol{\Phi}_{[j]}$ correspond to the columns of the window with index j and size L. Since the unknown vector $\boldsymbol{x}$ has a special non-uniform block structure with *unspecified boundaries* (as described in Sec. III), together, these columns do not correspond to a conventional block sparse vector. This necessitates definition of a new block sparse structure as given below.

1) The Pseudoblock Interleaved Block Sparse Structure (PIBS): The proposed PIBS structure is a more general form of support set than given by the columns of the matrices $\boldsymbol{\Phi}_{|S^{k-1}\rangle}$ and $\boldsymbol{\Phi}_{[j]}$. In this, clusters of consecutive non-zero blocks (called *true* clusters from hereafter) with sizes given by integer multiples of a constant b, and having unspecified boundaries (analogous to S^{k-1}) are well separated by several indices, and these gaps may contain a second type of non-zero blocks that we call *pseudo* blocks (analogous to the window $[j]$ in $\boldsymbol{\Phi}_{[j]}$). However, unlike above that has only one pseudo block $\boldsymbol{\Phi}_{[j]}$, we consider the more general case of r pseudo blocks, $0 \le r \le R$. Further, we do not constrain each pseudo block to remain confined to any specific window. Instead, they can be positioned anywhere in the cluster space such that they do not overlap with any of the true clusters. Positions not occupied by either the true clusters or the pseudo blocks are filled with zeros. The specific structure of a PIBS vector of length n is described by a $6-$tuple (b, p, l, L', K, R) that is explained below and is also illustrated in Fig 2:

1) There are k true clusters, with $0 \le k \le K$, having lengths $j_1 b, \dots, j_k b$, where $1 \le j_s \le p$, $1 \le s \le k$, such that $\sum_{s=1}^k j_s = K$.

2) Any two consecutive true clusters are separated by at least L' zeros. We will always assume that $L' = L + 2bp - b$, where L is the length of a window, defined in Section II.

3) There are r pseudo blocks, $0 \le r \le R$, which are a second type of nonzero blocks, that are of length l each, $0 \le l \le L'$, and are positioned anywhere but in a way so that they do not overlap with any of the k true clusters defined above.

Moreover, the signal length n is assumed to be sufficiently large so that all the R pseudo blocks, as well as $0 \leq k \leq K$ true clusters can be accommodated within the signal without any overlap between the true clusters and the pseudo blocks.

Using this definition of the PIBS structure it can be easily seen that the columns of $\Phi_{|S^{k-1}\rangle}$, along with the columns of $\Phi_{[j]}$ correspond to the support of a PIBS vector with parameters $(b, p, L, L', c_k, 1)$ ($c_k = |S^{k-1}|$), where the pseudo block is given by the window $[j]$, and the true clusters are given by $|S^{k-1}\rangle$.

Next we define a restricted isometry property for this PIBS signal structure, and state and prove a few Lemmas associated to it which will be required in the subsequent analysis of the TSGBOMP algorithm.

2) Useful Definitions and Lemmas Related to the PIBS Structure: We first define an analog of the celebrated restricted isometry property (RIP) [17] for the PIBS vectors. For this, it may be useful to recall the structure of a PIBS vector that is specified by the 6-tuple (b, p, l, L', k, r) (apart from the length n). In this, b is the length of a true block and a true cluster has length $bj,\ 1 \leq j \leq p$; l is the number of pseudo blocks; L' is the minimum separation of indices between two consecutive true clusters; k is the total number of true clusters while the total number of true blocks in them is K, and r is the total number of pseudo blocks such that $0 \leq r \leq R$.

Definition 4.1: The pseudoblock interleaved block restricted isometry constant (PIBRIC) of order (K, R) with parameters b, p, l, L' of a matrix $\Phi \in \mathbb{R}^{m \times n}$ is defined as

$$\delta_{b,p,l,L'}(K, R) = \max_{S \subset \mathcal{H}: S \in \cup_{r=0}^{R} \cup_{k=0}^{K} \Sigma_{b,p,l,L'}(k,r)} \left\| \Phi_S^t \Phi_S - I_{|S| \times |S|} \right\|_{2 \to 2}, \tag{12}$$

where $\Sigma_{b,p,l,L'}(k, r)$ is the collection of all subsets $S \subset \mathcal{H}$ such that a PIBS vector with parameters (b, p, l, L', k, r) can be supported on S.

The following lemma shows that the above definition is equivalent to the following interpretation of PIBRIC which is frequently used in the definition of RIP in literature.

Lemma 4.1: A matrix $\Phi \in \mathbb{R}^{m \times n}$ having the PIBRIC $\delta_{b,p,l,L'}(K, R)$ of order (K, R) with parameters b, p, l, L' satisfies the following inequality for every PIBS vector $x \in \mathbb{R}^n$ with parameters $(b, p, l, L', k, r),\ 0 \leq k \leq K,\ 0 \leq r \leq R$:

$$(1 - \delta) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta) \|x\|_2^2, \tag{13}$$

for all $\delta \geq \delta_{b,p,l,L'}(K, R)$.

Proof: The proof is supplied in Appendix A. ■

Conventionally, in the literature the RIC has been required to be bounded above by a constant strictly smaller than 1 to ensure success of compressed sensing algorithms, which is why we also aim to derive such an upper bound on a PIBRIC of certain order to ensure success of the TSGBOMP algorithm. In the sequel, we adopt the convention that a matrix $\Phi \in \mathbb{R}^{m \times n}$ is said to satisfy pseudo block interleaved restricted isometry property (PIBRIP) of order (K, R) with parameters b, p, l, L' if $\delta_{b,p,l,L'}(K, R) \in (0, 1)$.

Note that the PIBS vector with parameters (b, p, l, L', K, R) is at most $Kb + Rl$ sparse in the conventional sense. From definition 4.1, the computation of $\delta_{b,p,l,L'}(K, R)$ requires one to search for the maximum eigenvalue of $\|\Phi_S^t \Phi_S - I_S\|_{2 \to 2}$ only over $\cup_{r=0}^{R} \cup_{k=0}^{K} \Sigma_{b,p,l,L'}(k, r)$, whereas the computation of δ_{Kb+Rl} requires finding the maximum eigenvalue of a similar submatrix over *all* subsets S of size $\leq (Kb + Rl)$, meaning, the PIBRIC $\delta_{b,p,l,L'}(K, R)$ is smaller than the conventional RIC δ_{Kb+Rl}. Similar observations were also made for the block sparse vectors in [3]. Such smaller restricted isometry constants are inherent in signals with specialized signal structure, which belong to a class of general signal structures referred to as model sparse signals [14].

Now we state a few lemmas which are useful for analysing the TSGBOMP algorithm. Similar lemmas have already appeared in the literature in the context of the RIP for sparse [18]–[20] and block sparse vectors [8]. We modify them to be applicable to the PIBS structure and state them below. The proofs of these Lemmas are sketched in Appendix A, with details given in the accompanying supplementary material.

Lemma 4.2 (Block number monotonicity): Let the matrix Φ satisfy PIBRIP with parameters $(b, p, l, L', K_i, R_j),\ 1 \leq i,\ j \leq 2$, and $K_1 \leq K_2,\ R_1 \leq R_2$. Then, $\delta_{b,p,l,L'}(K_1, R_1) \leq \delta_{b,p,l,L'}(K_2, R_1) \leq \delta_{b,p,l,L'}(K_2, R_2)$, and $\delta_{b,p,l,L'}(K_1, R_1) \leq \delta_{b,p,l,L'}(K_1, R_2) \leq \delta_{b,p,l,L'}(K_2, R_2)$.

Lemma 4.3: For fixed b, p, K, if $0 \leq l < L'$, then $\delta_{b,p,l,L'}(K, R) \leq \delta_{b,p,L',L'}(K, R)$ for $R = 0, 1, 2$.

Lemma 4.4: For any $K \geq 1$ and $L' = L + 2bp - b > L \geq bp$, $\delta_{b,p,L',L'}(K, 1) \leq \delta_{b,p,L',L'}(K - 1, 2)$.

Lemma 4.5 (Projected matrix PIBRIP): Let $\Phi \in \mathbb{R}^{m \times n}$ be a matrix such that $\delta_{b,p,l,L'}(K, R) < 1$. Let S be the support of a PIBS vector of length n with parameters $(b, p, l, L', k, r),\ 0 \leq k \leq K,\ 0 \leq r \leq R$, i.e., S contains the indices of the true clusters as well as the pseudo blocks of such a vector. Let $S_1,\ S_2 \subset S$ be such that $S = \{S_1\} \cup S_2$ (S_1 contains the starting indices of some of the true blocks in S), and let $x \in \mathbb{R}^n$ such that $\mathtt{supp}(x) = S_2$. Then,

$$(1 - \delta_{b,p,l,L'}(K, R)) \|x\|_2^2 \leq \left\| \mathbf{P}_{S_1}^{\perp} \Phi_{S_2} x_{S_2} \right\|_2^2 \leq (1 + \delta_{b,p,l,L'}(K, R)) \|x\|_2^2. \tag{14}$$

Lemma 4.6: Let the matrix Φ satisfy $\delta_{b,p,l,L'}(K, R) < 1$. Let $S_1, S_2, S_3 \subset \mathcal{H}$ be such that the sets $S_2,\ S_3$ are disjoint and $|S_1\rangle \cup S_2 \cup S_3$ is the support of a PIBS vector with parameters $(b, p, l, L', k, r),\ 0 \leq k \leq K,\ 0 \leq r \leq R$. Then, for any two vectors u, v that are supported on the sets S_2, S_3 respectively,

$$\left| \left\langle \mathbf{P}_{S_1}^{\perp} \Phi u, \Phi v \right\rangle \right| \leq \delta_{b,p,l,L'}(K, R) \|u\|_2 \|v\|_2. \tag{15}$$

Lemma 4.7: Given that the matrix Φ has columns with unit norm and that it satisfies $\delta_{b,p,1,L'}(K, 1) < 1$. Let $S \subset \mathcal{H}$ be such that $|S\rangle$ is the support of the true clusters of a PIBS vector with parameters $(b, p, l, L', k, r),\ 0 \leq k \leq K,\ 0 \leq r \leq R$, and let j be an index such that $j \notin |S\rangle$. Then,

$$\left\| \mathbf{P}_S^{\perp} \phi_j \right\|_2 \geq \sqrt{1 - \delta_{b,p,1,L'}^2(K, 1)}. \tag{16}$$

B. Recovery Guarantee

Equipped with the definition of PIBS structure in Sec. IV-A1, it can be easily seen that the unknown vector $\boldsymbol{x}$, described in Sec. III, has the structure of a PIBS vector of with parameters $(b, p, L, L', K, 0)$. We now state a sufficient condition that the measurement matrix $\boldsymbol{\Phi}$ as well as the unknown vector $\boldsymbol{x}$ need to satisfy so that the TSGBOMP algorithm can exactly recover the support of $\boldsymbol{x}$ within K iterations.

Theorem 4.1: Let $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ and the measurement matrix $\boldsymbol{\Phi}$ as well as the unknown vector $\boldsymbol{x}$, with parameters $(b, p, L, L', K, 0)$ satisfy the conditions:

$$\delta < \frac{1}{\sqrt{2K+1}}, \tag{17}$$

$$x_{\min} > \frac{x_{\max}\delta\sqrt{B'b}}{(1+\delta)}\left[\frac{K}{4B'}(1+\delta)+\sqrt{K+1}+1\right] + \frac{\sqrt{2(1+B')(1+\delta)}\epsilon}{(1-\delta\sqrt{2K+1})}, \tag{18}$$

where $B' = pb - b + 1$, $\delta := \delta_{b,p,L',L'}(K-1, 2), x_{\min} = \min\{|x_j| : j \in \mathrm{supp}(\boldsymbol{x})\}$, $x_{\max} = \max\{|x_j| : j \in \mathrm{supp}(\boldsymbol{x})\}$. Then the TSGBOMP algorithm can recover the support of $\boldsymbol{x}$ exactly within K iterations.

Proof: The derivations are detailed in Sections IV-A, IV-C, IV-D, and IV-E. ■

C. Back to Finding a Condition for Identifying a Correct Window at Step $k (K \geq 1)$

We now return to complete the analysis of Sec. IV-A and proceed toward finding the sufficient condition for success of TSGBOMP at any step as provided by Theorem 4.1. Throughout the analysis we maintain that all the PIBS vectors considered will have length n where n is chosen large enough to accommodate all possible configurations of the different PIBS vectors emerging in the analysis.[5] We will now find an upper bound of the left hand side (LHS) of (11) by finding the PIBRIC of the associated PIBRIP of $\boldsymbol{B} = \mathbf{P}^{\perp}_{T^{k-1}}[\boldsymbol{\Phi}_{|S^{k-1}\rangle}\ \boldsymbol{\Phi}_{[j]}]$.

First, observe that it can be now easily verified that both the vectors $\boldsymbol{u}+\boldsymbol{w}$ as well as $\mu^2\boldsymbol{u}-\boldsymbol{w}$ have the common support (for $\mu \neq 0$) $S' = |S^{k-1}\rangle \cup [j]$, which corresponds to the support of 1 pseudo block of length L and $c_k (= |S^{k-1}|)$ true clusters of a PIBS vector with parameters $(b, p, L, L', c_k, 1)$. Now, let us denote $\widetilde{S} = |T^{k-1}\rangle \cup S'$. Note that as $j \in O^C_{S^{k-1}}$ the window j does not have any overlap with $|S^{k-1}\rangle$, but might have an overlap with $|T^{k-1}\rangle$. Consider first the case that the window j does not have any overlap with $|T^{k-1}\rangle$. Then $\widetilde{S}$ corresponds to the support of a PIBS vector with parameters $(b, p, L, L', K, 1)$. Consequently, according to Lemma 4.5, the matrix $\boldsymbol{B} = \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{S'}$ satisfies PIBRIP with PIBRIC given by $\delta_{b,p,L,L'}(K, 1)$, which is upper bounded by $\delta_{b,p,L',L'}(K, 1)$ by Lemma 4.3. On the other hand, if there is a nonempty overlap of the window j with $|T^{k-1}\rangle$, there can be overlap with at most one of the true clusters in $|T^{k-1}\rangle$. Call that true cluster (i.e. the set of its indices) $\mathcal{C}'$. Now, there are two cases to consider. In one case if $\mathcal{C}'$ is a proper subset of the window j, the set $\widetilde{S}$ corresponds to the support of a PIBS vector with parameters $(b, p, L, L', K-k', 1)$, where k' is the number of true blocks in the cluster $\mathcal{C}'$. In this case, by Lemma 4.5 the matrix $\boldsymbol{B}$ satisfies PIBRIP with PIBRIC $\delta_{b,p,L,L'}(K-k', 1)$, which is upper bounded by $\delta_{b,p,L',L'}(K, 1)$ by Lemmas 4.2 and 4.3. On the other hand, if $\mathcal{C}'$ has only partial overlap with the window j, with say l' being the size of the overlap (clearly, $1 \leq l' < pb \leq L$), the set $\widetilde{S}$ corresponds to the support of a PIBS vector with parameters $(b, p, L-l', L', K, 1)$. In this case, by Lemma 4.5 the matrix $\boldsymbol{B}$ satisfies PIBRIP with PIBRIC given by $\delta_{b,p,L-l',L'}(K, 1)$, which in turn is upper bounded by $\delta_{b,p,L',L'}(K, 1)$ by Lemma 4.3. Therefore, using the expressions of $\boldsymbol{u}, \boldsymbol{w}$ from Eqs. (8) and (9) and the fact that $\boldsymbol{u}$ and $\boldsymbol{w}$ are orthogonal, we obtain,

$$\begin{aligned}
&\|\boldsymbol{B}(\boldsymbol{u}+\boldsymbol{w})\|_2^2 - \left\|\boldsymbol{B}(\mu^2\boldsymbol{u}-\boldsymbol{w})\right\|_2^2 \\
&\geq (1-\delta_{b,p,L',L'}(K,1))\,\|\boldsymbol{u}+\boldsymbol{w}\|_2^2 \\
&- (1+\delta_{b,p,L',L'}(K,1))\left\|\mu^2\boldsymbol{u}-\boldsymbol{w}\right\|_2^2 \\
&= \left[(1-\delta_{b,p,L',L'}(K,1)) - \mu^4(1+\delta_{b,p,L',L'}(K,1))\right]\|\boldsymbol{u}\|_2^2 \\
&- 2\delta_{b,p,L',L'}(K,1)\,\|\boldsymbol{w}\|_2^2, \\
&= (1-\mu^4)\left[1-\frac{1+\mu^4}{1-\mu^4}\delta_{b,p,L',L'}(K,1)\right]\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 \\
&- 2\mu^2\delta_{b,p,L',L'}(K,1)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 \\
&= (1-\mu^4)\left[1-\frac{1+\mu^2}{1-\mu^2}\delta_{b,p,L',L'}(K,1)\right]\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2.
\end{aligned} \tag{19}$$

Now, using the expression of μ from (6), we obtain,

$$\frac{1+\mu^2}{1-\mu^2} = \frac{\theta+\left(\sqrt{\theta+1}-1\right)^2}{\theta-\left(\sqrt{\theta+1}-1\right)^2} = \frac{2\theta+2-2\sqrt{\theta+1}}{2\sqrt{\theta+1}-2} = \sqrt{\theta+1}. \tag{20}$$

Therefore, one can find a lower bound of the RHS of (11) as the following:

$$\begin{aligned}
&\|\boldsymbol{B}(\boldsymbol{u}+\boldsymbol{w})\|_2^2 - \left\|\boldsymbol{B}(\mu^2\boldsymbol{u}-\boldsymbol{w})\right\|_2^2 \\
&\geq (1-\mu^4)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2\left(1-\sqrt{\theta+1}\delta_{b,p,L',L'}(K,1)\right).
\end{aligned} \tag{21}$$

Since, it is easy to verify that $|\mu| < 1$, and since the identity (11) and the inequality (21) hold true for all $j \in O^C_{S^k}$, it follows that

$$\begin{aligned}
&\|\boldsymbol{B}\boldsymbol{u}\|_2^2 \\
&\quad - \sqrt{\theta}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 \cdot \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&\geq \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2\left(1-\sqrt{\theta+1}\delta_{b,p,L',L'}(K,1)\right).
\end{aligned} \tag{22}$$

Therefore, using (4), (22), and $\boldsymbol{B}\boldsymbol{u} = \mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}$, it follows that

$$\begin{aligned}
&\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} \\
&- \sqrt{\frac{\theta}{d_k}}\left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty}
\end{aligned}$$

[5]It can be shown that by taking $n \geq Kb + 2L' + K + 1$ this property is satisfied.

$$\geq \frac{\left(1-\sqrt{\theta+1}\delta_{b,p,L',L'}(K,1)\right)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{d_k}}. \quad (23)$$

Clearly, putting $\theta = d_k$ we recover the LHS of (3) from the LHS of (23). Consequently, we obtain,

$$\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|T\setminus T^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty} - \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_{2,\infty}$$
$$\geq \frac{\left(1-\sqrt{d_k+1}\delta_{b,p,L',L'}(K,1)\right)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{d_k}}. \quad (24)$$

To find an upper bound on the RHS of inequality (3), we follow the derivation of the inequality (4.22) in [8]. First we note that there exists windows indexed by $i_0 \in O_{S^{k-1}}$, and $j_0 \in O^C_{S^{k-1}}$, such that $\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\|_{2,\infty} = \|\boldsymbol{\Phi}^t_{[i_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\|_2$, and $\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\|_{2,\infty} = \|\boldsymbol{\Phi}^t_{[j_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\|_2$. Therefore,

$$\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty} + \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty}$$
$$= \left\|\boldsymbol{\Phi}^t_{[i_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_2 + \left\|\boldsymbol{\Phi}^t_{[j_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_2$$
$$\leq \sqrt{2}\left\|\boldsymbol{\Phi}^t_{[i_0\cup j_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_2$$
$$\leq \sqrt{2}\sigma_{\max}\left(\boldsymbol{\Phi}^t_{[i_0\cup j_0]}\mathbf{P}^{\perp}_{T^{k-1}}\right)\|\boldsymbol{e}\|_2$$
$$= \sqrt{2\lambda_{\max}\left(\boldsymbol{\Phi}^t_{[i_0\cup j_0]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{[i_0\cup j_0]}\right)}\|\boldsymbol{e}\|_2, \quad (25)$$

where in the last three steps, for any matrix $\boldsymbol{A}$, $\sigma_{\max}(\boldsymbol{A})$ denotes the maximum singular value of $\boldsymbol{A}$ and for any symmetric matrix $\boldsymbol{B}$, $\lambda_{\max}(\boldsymbol{B})$ denotes the largest eigenvalue of $\boldsymbol{B}$.

Now, observe that although the window i_0 cannot have an overlap with the set $|T^{k-1}\rangle$, there might be overlap of the latter with window j_0. If there is no overlap, then the set $S'' = |T^{k-1}\rangle \cup [i_0 \cup j_0]$ corresponds to the support of a PIBS vector with parameters $(b,p,L,L',|T^{k-1}|,2)$. If there is overlap, let $\mathcal{C}''$ be the true cluster in $|T^{k-1}\rangle$ which has a nonempty overlap with the window j_0. If $\mathcal{C}''$ is a subset of window j_0, the set S'' corresponds to the support of a PIBS vector with parameters $(b,p,L,L',|T^{k-1}|-k,''2)$, where k'' is the number of true blocks in the cluster $\mathcal{C}''$ (note that $1 \leq k'' \leq \min\{p,|T^{k-1}|\}$). On the other hand, if $\mathcal{C}''$ has only partial overlap with window j_0, then, assuming that the length of overlap between $\mathcal{C}''$ and the window j_0 is $l''(1 \leq l'' \leq k''b-1)$, the set S'' can be covered by another set S''' which consists of the unions of the windows i_0, j_0, the set $|T^{k-1}\rangle \setminus \mathcal{C}''$ and another true cluster of size $k''b$ obtained by prefixing or suffixing (as the case may be) l'' indices to the non-overlapping side of $\mathcal{C}'' \setminus [j_0]$. Clearly, The set S''' corresponds to the support of a PIBS vector with parameters $(b,p,L,L',|T^{k-1}|,2)$.

Therefore, using the Lemmas 4.5, 4.2 as well as 4.3, it can be seen that the matrix $\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{[i_0\cup j_0]}$ satisfies PIB-RIP with PIBRIC given by $\delta_{b,p,L',L'}(|T^{k-1}|,2)$. Consequently, $\lambda_{\max}(\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{\Phi}_{[i_0\cup j_0]}) \leq (1+\delta_{b,p,L',L'}(|T^{k-1}|,2))$. Therefore,

$$\left\|\boldsymbol{\Phi}^t_{[O_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty} + \left\|\boldsymbol{\Phi}^t_{[O^C_{S^{k-1}}]}\mathbf{P}^{\perp}_{T^{k-1}}\boldsymbol{e}\right\|_{2,\infty}$$
$$\leq \sqrt{2(1+\delta_{b,p,L',L'}(|T^{k-1}|,2))}\epsilon, \quad (26)$$

where we have assumed that $\|\boldsymbol{e}\|_2 \leq \epsilon$. Thus, using inequalities (3), (24) and (26) the following sufficient condition is derived to ensure a correct window selection at step $k(k \geq 1)$:

$$\frac{\left(1-\sqrt{d_k+1}\delta_{b,p,L',L'}(K,1)\right)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{d_k}} > \sqrt{2(1+\delta_{b,p,L',L'}(|T^{k-1}|,2))}\epsilon. \quad (27)$$

D. Condition for True Cluster Selection At Step $k(K \geq 1)$

We assume that a correct window, indexed by w^k (that is the w^kth window), has already been selected at step k, i.e. set of the columns of $\boldsymbol{\Phi}_{[w^k]}$ has a nonempty intersection with the set of columns of $\boldsymbol{\Phi}_{|S^{k-1}\rangle}$. We now find a condition to ensure that a true cluster from $|S^{k-1}\rangle$, having a nonempty overlap with the window $[w^k]$, is selected at step k.

Let the set of indices corresponding to the true cluster having a non-empty overlap with window $[w^k]$ be denoted by C^k and let t $(1 \leq t \leq p)$ be the number of true blocks in C^k. Clearly, $C^k \subset \beta^k$, where $\beta^k = \{L(w^k-1)+1-(B-1), L(w^k-1)-(B-1)+2,\ldots,Lw^k+B-1\}$. Note that the assumption that any two consecutive true clusters are separated by at least $L' = L+2bp-b$ zeros, ensures that the set β^k does not have *any* overlap with with a true cluster from $|S^{k-1}\rangle$ other than C^k, that is, it ensures that $|S^{k-1}\rangle \cap \beta^k \setminus C^k = \emptyset$.

Let $\mathcal{W}$ be the collection of all sets S of size B such that $S \subset \beta^k$ and such that S covers C^k, i.e., $C^k \subseteq S$. On the other hand, let $\mathcal{W}'$ be the collection of all sets S' of size B such that $S' \subset \beta^k$ and that S' does not cover C^k, i.e., $C^k \not\subseteq S'$. Then, a set $S \in \mathcal{W}$ is selected at step $k(\geq 1)$ if and only if

$$\max_{S\in\mathcal{W}}\left\|\boldsymbol{\Phi}^t_S\boldsymbol{r}^{k-1}\right\|_2 > \max_{S'\in\mathcal{W}'}\left\|\boldsymbol{\Phi}^t_{S'}\boldsymbol{r}^{k-1}\right\|_2. \quad (28)$$

Now, let $S_0 = \arg\max_{S\in\mathcal{W}}\|\boldsymbol{\Phi}^t_S\boldsymbol{r}^{k-1}\|_2$. Consider any $S' \in \mathcal{W}'$. Let $s' = |S' \cap C^k|$, $t' = |S' \cap S_0 \setminus C^k|$. Note that

$$\left\|\boldsymbol{\Phi}^t_{S_0\setminus S'}\boldsymbol{r}^{k-1}\right\|_2 \geq \left\|\boldsymbol{\Phi}^t_{C_k\setminus S'}\boldsymbol{r}^{k-1}\right\|_2$$
$$\geq \sqrt{tb-s'}\min_{j\in C^k}\left|\left\langle\boldsymbol{\phi}_j,\boldsymbol{r}^{k-1}\right\rangle\right|, \quad (29)$$

where we have used $|C_k \setminus S'| = |C_k| - |C_k \cap S'| = tb - s'$. On the other hand,

$$\left\|\boldsymbol{\Phi}^t_{S'\setminus S_0}\boldsymbol{r}^{k-1}\right\|_2 \leq \sqrt{pb-s'-t'}\max_{l\in\beta^k\setminus C^k}\left|\left\langle\boldsymbol{\phi}_l,\boldsymbol{r}^{k-1}\right\rangle\right|, \quad (30)$$

where we have used $|S' \setminus S_0| = |S'| - |S' \cap S_0| = |S'| - |S' \cap S_0 \setminus C^k| - |S' \cap C^k| = pb - t' - s'$. Therefore, the inequality (28) is satisfied if

$$\min_{j\in C^k}\left|\left\langle\boldsymbol{\phi}_j,\boldsymbol{r}^{k-1}\right\rangle\right| \geq \sqrt{\frac{pb-t'-s'}{tb-s'}}\max_{l\in\beta^k\setminus C^k}\left|\left\langle\boldsymbol{\phi}_l,\boldsymbol{r}^{k-1}\right\rangle\right|, \quad (31)$$

where in the above we have used the fact that $0 \leq s' \leq tb-1$ since, by definition, S' cannot fully cover C^k.

Now, note that $\frac{pb-t'-s'}{tb-s'} = \frac{(p-t)b-t'}{tb-s'} + 1$ is an increasing function of s' since $0 \leq t' = |S' \cap S_0 \setminus C^k| \leq |S_0 \setminus C^k| = (p-t)b$. Since $s' \leq tb-1$, we obtain that $\frac{pb-t'-s'}{tb-s'} \leq (p-t)b - t' +$

$1 \leq pb - b + 1$, where we have used the facts $t \geq 1,\ t' \geq 0$. Therefore, the inequality (31) is satisfied if the following is satisfied[6]:

$$\min_{j \in C^k} \left|\langle \boldsymbol{\phi}_j, \boldsymbol{r}^{k-1} \rangle\right| \geq \sqrt{B'} \max_{l \in \beta^k \setminus C^k} \left|\langle \boldsymbol{\phi}_l, \boldsymbol{r}^{k-1} \rangle\right|, \quad (32)$$

where $B' = pb - b + 1$.

Fix any $j_1 \in C^k,\ j_2 \in \beta^k \setminus C^k$. Now, we obtain

$$\begin{aligned} &\left|\langle \boldsymbol{\phi}_{j_1}, \boldsymbol{r}^{k-1} \rangle\right| - \sqrt{B'} \left|\langle \boldsymbol{\phi}_{j_2}, \boldsymbol{r}^{k-1} \rangle\right| \\ &\overset{(d)}{\geq} \underbrace{\left|\langle \boldsymbol{\phi}_{j_1}, \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle} \rangle\right|}_{F_1} \\ &\quad - \sqrt{B'} \underbrace{\left|\langle \boldsymbol{\phi}_{j_2}, \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle} \rangle\right|}_{F_2} \\ &\quad - \underbrace{\left(\left|\langle \boldsymbol{\phi}_{j_1}, \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{e} \rangle\right| + \sqrt{B'} \left|\langle \boldsymbol{\phi}_{j_2}, \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{e} \rangle\right|\right)}_{F_3}. \quad (33) \end{aligned}$$

Here, step (d) uses the expression for $\boldsymbol{r}^{k-1}$ from (2) and the reverse triangle inequality and triangle inequality, respectively. We now proceed to find upper bounds of F_2, F_3 and a lower bound of F_1.

First consider F_2. To find its upper bound, a procedure exactly similar to the one used via (11), (21) and (22) to calculate an upper bound of a similar quantity $\|\boldsymbol{\Phi}^t_{[j]} \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle}\|_2$ can be used. Since, like the window $[j]$ which is disjoint to $|S^{k-1}\rangle$, the column $j_2 \notin |S^{k-1}\rangle$, this will imply simply replacing the window $[j]$ of length L by a window of length 1 consisting of the column j_2 only. We make corresponding changes in the definitions of $\mu, \boldsymbol{h}, \boldsymbol{u}, \boldsymbol{w}, \boldsymbol{B}$ as given by (6)–(10) respectively, by replacing θ by some positive number α, L by 1, and $\boldsymbol{\Phi}_{[j]}$ by $\boldsymbol{\phi}_{j_2}$. To describe the structure of the PIBS vector that emerges as a result of a similar analysis, note that since $j_2 \in \beta^k \setminus C^k$, we always have $j_2 \notin |T^{k-1}\rangle$. As a result, we have a PIBS vector with parameters $(b, p, 1, L', K, 1)$ (corresponding PIBRIC : $\delta_{b,p,1,L'}(K, 1)$). Consequently, following the steps of (11), (21) and (22), we obtain the following inequality (for an arbitrary positive number α):

$$\begin{aligned} &\left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 - \sqrt{\alpha} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 F_2 \\ &\geq \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 \left(1 - \sqrt{\alpha + 1}\,\delta_{b,p,1,L'}(K, 1)\right). \quad (34) \end{aligned}$$

To find an upper bound of F_3, we derive:

$$\begin{aligned} F_3 &\leq \sqrt{1 + B'} \left\|\boldsymbol{\Phi}^t_U \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{e}\right\|_2 \\ &\leq \sqrt{1 + B'} \sqrt{\lambda_{\max}\left(\boldsymbol{\Phi}^t_U \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_U\right)} \|\boldsymbol{e}\|_2 \\ &\leq \sqrt{(B' + 1)(1 + \delta_{b,p,1,L'}(\left|T^{k-1}\right|, 2))}\epsilon, \quad (35) \end{aligned}$$

where $U = \{j_1, j_2\}$, and the last step follows from Lemma 4.5 which uses the observation that $|T^{k-1}\rangle \cup U$ corresponds to the support of a PIBS vector with parameters $(b, p, 1, L', |T^{k-1}|, 2)$ (which is true because $j_1, j_2 \notin |T^{k-1}\rangle$ and $j_1 \neq j_2$).

[6]Similar sufficient condition was derived in [13] in the context of a coherence based analysis of the sliding-block type algorithm proposed therein.

Now, we proceed to find a lower bound of F_1. For this, first we define, for any $u \in \mathbb{R}$, $\mathtt{sgn}(u) = 1$ if $u \geq 0$ and $\mathtt{sgn}(u) = -1$ if $u < 0$. Then recalling that $\boldsymbol{q}^{k-1} = \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle}$, we have $F_1 = |\langle \boldsymbol{\phi}_{j_1}, \boldsymbol{q}^{k-1} \rangle| = \langle s_{j_1} \boldsymbol{\phi}_{j_1}, \boldsymbol{q}^{k-1} \rangle \equiv \langle s_{j_1} \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\phi}_{j_1}, \boldsymbol{q}^{k-1} \rangle$, where, $s_{j_1} = \mathtt{sgn}(\langle \boldsymbol{\phi}_{j_1}, \boldsymbol{q}^{k-1} \rangle)$. Also, for two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we can write $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ as $\frac{1}{2}[\|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2 - \|\boldsymbol{x} - \boldsymbol{y}\|_2^2]$. Then, defining $\alpha' = \alpha / B'$, one can write,

$$\begin{aligned} &F_1 - \frac{\left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2}{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \\ &\overset{(e)}{=} \frac{1}{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \left(\frac{\alpha' \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 \left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\phi}_{j_1}\right\|_2^2}{4} \right. \\ &\quad \left. - \left\|\frac{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2 \mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\phi}_{j_1} s_{j_1}}{2} - \boldsymbol{q}^{k-1}\right\|_2^2\right) \\ &= \frac{1}{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \left(\frac{\alpha' \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 \left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\phi}_{j_1}\right\|_2^2}{4} \right. \\ &\quad \left. - \left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{v}_{|S^{k-1}\rangle}\right\|_2^2\right), \end{aligned}$$

where the vector $\boldsymbol{v} \in \mathbb{R}^n$ is defined as follows : $v_r = 0\ \forall r \notin |S^{k-1}\rangle$ and $\boldsymbol{v}_{|S^{k-1}\rangle} = \boldsymbol{x}_{|S^{k-1}\rangle}$, except for the index j_1, for which, $v_{j_1} = x_{j_1} - \frac{s_{j_1} \sqrt{\alpha'} \|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2}{2}$. Then, as $j_1 \notin |T^{k-1}\rangle$, using Lemma 4.7 we obtain

$$\left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\phi}_{j_1}\right\|_2^2 \geq (1 - \delta^2_{b,p,1,L'}(\left|T^{k-1}\right|, 1)). \quad (36)$$

On the other hand, as the support of $\boldsymbol{v}$ is $|S^{k-1}\rangle$ and since $|T^{k-1}\rangle \cup |S^{k-1}\rangle$ is the support of a PIBS vector with parameters $(b, p, 1, L', K, 0)$, using Lemma 4.5, we obtain

$$\left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{v}_{|S^{k-1}\rangle}\right\|_2^2 \leq (1 + \delta_{b,p,1,L'}(K, 0)) \|\boldsymbol{v}\|_2^2. \quad (37)$$

Therefore, we obtain,

$$\begin{aligned} &F_1 - \frac{\left\|\mathbf{P}^{\perp}_{T^{k-1}} \boldsymbol{\Phi}_{|S^{k-1}\rangle} \boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2}{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \\ &\geq \frac{1}{\sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2} \left(\alpha'/4 \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 (1 - \delta^2_{b,p,1,L'}(\left|T^{k-1}\right|, 1)) \right. \\ &\quad \left. - (1 + \delta_{b,p,1,L'}(K, 0)) \|\boldsymbol{v}\|_2^2\right). \quad (38) \end{aligned}$$

Now observe that

$$\begin{aligned} \|\boldsymbol{v}\|_2^2 &= \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 - x_{j_1}^2 + \left(x_{j_1} - \frac{s_{j_1} \sqrt{\alpha'} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{2}\right)^2 \\ &= \left(1 + \frac{\alpha'}{4}\right) \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 - \sqrt{\alpha'} s_{j_1} x_{j_1} \left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2. \end{aligned} \quad (39)$$

Therefore, from the inequalities (33)–(35), (38), and (39), one can deduce that,

$$\begin{aligned} &\left|\langle \boldsymbol{\phi}_{j_1}, \boldsymbol{r}^{k-1} \rangle\right| - \left|\langle \boldsymbol{\phi}_{j_2}, \boldsymbol{r}^{k-1} \rangle\right| \\ &\geq \frac{\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{\alpha'}} (1 - \sqrt{\alpha + 1}\,\delta_{b,p,1,L'}(K, 1)) \end{aligned}$$

$$+ \frac{1}{\sqrt{\alpha'}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}\left[\frac{\alpha'}{4}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 (1-\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1)) - (1+\delta_{b,p,1,L'}(K,0))\left\{\left(1+\frac{\alpha'}{4}\right)\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2^2 - \sqrt{\alpha'}s_{j_1}x_{j_1}\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2\right\}\right] - \sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}\epsilon$$
$$\geq -\frac{\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{\alpha'}}\left[\frac{\alpha'}{4}\left(\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1) + \delta_{b,p,1,L'}(K,0)\right) + \sqrt{\alpha+1}\delta_{b,p,1,L'}(K,1) + \delta_{b,p,1,L'}(K,0)\right] + \min_{j_1\in C^k} s_{j_1}x_{j_1}(1+\delta_{b,p,1,L'}(K,0)) - \sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}\epsilon. \tag{40}$$

Clearly, the condition (32) is satisfied if the right hand side of the inequality (40) is non-negative. However, observe that, this is not possible unless $\min_{j_1\in C^k} s_{j_1}x_{j_1} > 0$. Now we have,

$$s_{j_1} = \mathtt{sgn}\left(\langle\boldsymbol{\phi}_{j_1}, \mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\Phi}_{|S^{k-1}\rangle}\boldsymbol{x}_{|S^{k-1}\rangle}\rangle\right) = \mathtt{sgn}\left(\left\|\mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\phi}_{j_1}\right\|_2^2 x_{j_1} + \left\langle\boldsymbol{\phi}_{j_1}, \mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\Phi}_{V_{j_1}^{k-1}}\boldsymbol{x}_{V_{j_1}^{k-1}}\right\rangle\right), \tag{41}$$

where the set $|S^{k-1}\rangle$ is partitioned into the index j_1 and the set $V_{j_1}^{k-1}$. Now, it is easy to check that for any $a, b \in \mathbb{R}$, $\mathtt{sgn}(a+b) = \mathtt{sgn}(a)$ if $|a| > |b|$. Therefore, a sufficient condition for $s_{j_1}x_{j_1} > 0, \forall j_1 \in C^k$ to be satisfied is the following:

$$\left\|\mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\phi}_{j_1}\right\|_2^2 |x_{j_1}| > \left|\left\langle\boldsymbol{\phi}_{j_1}, \mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\Phi}_{V_{j_1}^{k-1}}\boldsymbol{x}_{V_{j_1}^{k-1}}\right\rangle\right|, \tag{42}$$

$\forall j_1 \in C^k$. Clearly, under this condition, one has $s_{j_1}x_{j_1} = |x_{j_1}|$ We have already found during the derivation of inequality (38) that $\|\mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\phi}_{j_1}\|_2^2 \geq (1-\delta_{b,p,1,L'}^2(|T^{k-1}|,1))$. Moreover, to find an upper bound on $|\langle\boldsymbol{\phi}_{j_1}, \mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\Phi}_{V_{j_1}^{k-1}}\boldsymbol{x}_{V_{j_1}^{k-1}}\rangle|$, first note that the set $V_{j_1}^{k-1}$ is disjoint to the index j_1, and note that the union of the sets $|T^{k-1}\rangle$ and $V_{j_1}^{k-1}$ and the index j_1 corresponds to the union of the sets $|T^{k-1}\rangle$ and $|S^{k-1}\rangle$, which corresponds to the support of a PIBS vector with parameters $(b, p, 1, L', K, 0)$. Therefore, using Lemma 4.6 we obtain $|\langle\boldsymbol{\phi}_{j_1}, \mathbf{P}_{T^{k-1}}^{\perp}\boldsymbol{\Phi}_{V_{j_1}^{k-1}}\boldsymbol{x}_{V_{j_1}^{k-1}}\rangle| \leq \delta_{b,p,1,L'}(K,0)\|\boldsymbol{x}_{V_{j_1}^{k-1}}\|_2$. Hence, if for all $j_1 \in C^k$, $(1-\delta_{b,p,1,L'}^2(|T^{k-1}|,1))|x_{j_1}| > \delta_{b,p,1,L'}(K,0)\|\boldsymbol{x}_{V_{j_1}^{k-1}}\|_2$ holds, then $s_{j_1}x_{j_1} > 0,\ \forall j_1 \in C^k$. Since $\|\boldsymbol{x}_{V_{j_1}^{k-1}}\|_2 = \sqrt{\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2^2 - |x_{j_1}|^2}$, we further deduce that $s_{j_1}x_{j_1} > 0$ if $(1-\delta_{b,p,1,L'}^2(|T^{k-1}|,1))|x_{j_1}| > \delta_{b,p,1,L'}(K,0)\sqrt{\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2^2 - |x_{j_1}|^2}$, which is equivalent to

$$\frac{|x_{j_1}|}{\frac{\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{c_k b}}} > \frac{\delta_{b,p,1,L'}(K,0)\sqrt{c_k b}}{\sqrt{(1-\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1))^2 + \delta_{b,p,1,L'}^2(K,0)}}, \tag{43}$$

for all $j_1 \in C^k$. Now, let $x_{\min,k-1} = \min\{|x_j| : j \in |S^{k-1}\rangle\} \leq |x_{j_1}|$, and $x_{\max,k-1} := \max\{|x_j| : j \in |S^{k-1}\rangle\} \geq \frac{\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2}{\sqrt{c_k b}}$. Thus, to ensure that $\Re(s_{j_1}^* x_{j_1}) > 0$ for all $j_1 \in C^k$, the following serves as a sufficient condition:

$$\frac{x_{\min,k-1}}{x_{\max,k-1}} > \frac{\delta_{b,p,1,L'}(K,0)\sqrt{c_k b}}{\sqrt{(1-\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1))^2 + \delta_{b,p,1,L'}^2(K,0)}}. \tag{44}$$

Consequently, condition (32) is satisfied if the condition (44) is satisfied and if the right hand side of inequality (40) is made greater than or equal to 0, which is ensured by the following sufficient condition:

$$x_{\min,k-1} \geq \frac{\left\|\boldsymbol{x}_{|S^{k-1}\rangle}\right\|_2}{\sqrt{\alpha'}(1+\delta_{b,p,1,L'}(K,0))}\left[\frac{\alpha'}{4}\left(\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1) + \delta_{b,p,1,L'}(K,0)\right) + \sqrt{\alpha+1}\delta_{b,p,1,L'}(K,1) + \delta_{b,p,1,L'}(K,0)\right] + \frac{\sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}\epsilon}{(1+\delta_{b,p,1,L'}(K,0))}. \tag{45}$$

The above in turn is ensured by the following sufficient condition:

$$x_{\min,k-1} \geq \frac{x_{\max,k-1}\sqrt{c_k B' b}}{\sqrt{\alpha}(1+\delta_{b,p,1,L'}(K,0))}\left[\frac{\alpha'}{4}\left(\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1) + \delta_{b,p,1,L'}(K,0)\right) + \sqrt{\alpha+1}\delta_{b,p,1,L'}(K,1) + \delta_{b,p,1,L'}(K,0)\right] + \frac{\sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}\epsilon}{(1+\delta_{b,p,1,L'}(K,0))}. \tag{46}$$

As α can be any arbitrary positive number, choosing $\alpha = c_k$ and defining $c_k' = c_k/B$, we arrive at the following sufficient condition:

$$x_{\min,k-1} \geq \frac{x_{\max,k-1}\sqrt{B' b}}{(1+\delta_{b,p,1,L'}(K,0))}\left[\frac{c_k'}{4}\left(\delta_{b,p,1,L'}^2(\left|T^{k-1}\right|,1) + \delta_{b,p,1,L'}(K,0)\right) + \sqrt{c_k+1}\delta_{b,p,1,L'}(K,1) + \delta_{b,p,1,L'}(K,0)\right] + \frac{\sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}\epsilon}{(1+\delta_{b,p,1,L'}(K,0))}. \tag{47}$$

E. Condition for Overall Success

In this section we claim that the conditions (17), and (18) stated in Theorem 4.1 simultaneously satisfy the inequalities (27), (44), and (47) for all iterations $1 \leq k \leq K$. In the following δ will be used to denote $\delta_{b,p,L',L'}(K-1,2)$ unless otherwise specified.

We first verify that the inequality (27) is satisfied for all $1 \leq k \leq K$ under the conditions (17), and (18). Indeed, using Lemmas 4.4 and 4.2 we obtain $\delta_{b,p,L',L'}(K,1) \leq \delta_{b,p,L',L'}(K-1,2)$, and $\delta_{b,p,L',L'}(|T^{k-1}|,2) \leq \delta_{b,p,L',L'}(K-1,2)$, respectively. Moreover, for all $1 \leq k \leq K$, $d_k = |O_{S^{k-1}}| \leq 2|S^{k-1}| = 2c_k \leq 2|T| = 2K$, since each cluster in S^{k-1}

might have non-zero overlap with at most two windows in $O_{S^{k-1}}$. Therefore, $\frac{\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2}{\sqrt{d_k}} \geq \frac{\|\boldsymbol{x}_{|S^{k-1}\rangle}\|_2}{\sqrt{2c_k}} \geq \frac{\sqrt{b}x_{\min}}{\sqrt{2}} \geq \frac{x_{\min}}{\sqrt{2}} > \frac{\sqrt{(1+B')(1+\delta_{b,p,L',L'}(|T^{k-1}|,2))}\epsilon}{(1-\delta_{b,p,L',L'}(K,1)\sqrt{d_k+1})} \geq \frac{\sqrt{2(1+\delta_{b,p,L',L'}(|T^{k-1}|,2))}\epsilon}{(1-\delta_{b,p,L',L'}(K,1)\sqrt{d_k+1})}$, where the last two inequalities follow from condition (18) (by considering only the second term of the RHS of (18)) and the fact that $B' \geq 1$, respectively. After rearrangement, this results in the condition (27).

Now, to show that (17) and (18) imply (44), we first observe that

$$\begin{aligned}
&\frac{K}{4\sqrt{B'}} + \frac{\sqrt{B'}(\sqrt{K+1}+1)}{1+\delta} - \frac{\sqrt{K}}{\sqrt{\delta^2+(1-\delta^2)^2}} \\
&\overset{(g)}{\geq} \frac{K}{4\sqrt{B'}} + \frac{\sqrt{2B'}(\sqrt{K+1}+1)}{\sqrt{2}+1} - \sqrt{\frac{4K}{3}} \\
&= \frac{1}{\sqrt{B'}}\left(\frac{\sqrt{K}}{2} - \frac{2\sqrt{B'}}{\sqrt{3}}\right)^2 \\
&\quad + \frac{\sqrt{2B'}}{\sqrt{2}+1}\left(\sqrt{K+1} - \frac{2\sqrt{2}+1}{3}\right) \overset{(h)}{>} 0,
\end{aligned}$$

where step (g) uses $\delta < \frac{1}{\sqrt{2}}$ (which follows from (17) since $K \geq 1$), and the fact that the function $\delta^2 + (1-\delta^2)^2$ is monotonically decreasing for $\delta \in [0, \frac{1}{\sqrt{2}}]$ with the minimum at $\frac{1}{\sqrt{2}}$, and step (h) uses the simple observation that $\frac{2\sqrt{2}+1}{3} < \sqrt{2} \leq \sqrt{K+1}$ since $K \geq 1$. Then, using (18) and (17) it follows that

$$\begin{aligned}
x_{\min,k-1} &\geq x_{\min} > \frac{x_{\max}\delta\sqrt{Kb}}{\sqrt{\delta^2+(1-\delta^2)^2}} \\
&\geq \frac{x_{\max,k-1}\delta_{b,p,1,L'}(K,0)\sqrt{c_k b}}{\sqrt{(1-\delta^2_{b,p,1,L'}(|T^{k-1}|,1))^2 + \delta^2_{b,p,1,L'}(K,0)}}. \quad (48)
\end{aligned}$$

In the above we use the fact that the function $\frac{\delta}{\sqrt{(1-\delta^2)^2+\delta^2}}$ is monotonically increasing for $\delta \in [0,1]$, so that, using $x_{\max} \geq x_{\max,k-1}$, $c_k \leq K$, and the Lemmas 4.2, 4.3 and 4.4, we obtain, $\frac{\delta}{\sqrt{(1-\delta^2)^2+\delta^2}} \geq \frac{\delta_{b,p,1,L'}(K,1)}{\sqrt{(1-\delta^2_{b,p,1,L'}(K,1))^2+\delta^2_{b,p,1,L'}(K,1)}} = \frac{1}{\sqrt{1+(\frac{1-\delta^2_{b,p,1,L'}(K,1)}{\delta_{b,p,1,L'}(K,1)})^2}}$. Finally, using $\delta_{b,p,1,L'}(K,1) \geq \delta_{b,p,1,L'}(|T^{k-1}|,1)$, and $\delta_{b,p,1,L'}(K,1) \geq \delta_{b,p,1,L'}(K,0)$, we obtain the above inequality.

To show that (47) follows from (18) and (17), first observe that the condition (18) followed by (17) imply that

$$\begin{aligned}
&x_{\min} \\
&> x_{\max}\delta\sqrt{B'b}\left[\frac{K}{4B'}(1+\delta) + \sqrt{K+1}+1\right] - x_{\min}\delta \\
&+ \epsilon\sqrt{(1+B')(1+\delta)} \\
&\geq \left[x_{\max,k-1}\sqrt{B'b}\left(\frac{c_k}{4B'}+1\right) - x_{\min,k-1}\right]\delta \\
&\quad + x_{\max,k-1}\sqrt{B'b}\left[\frac{c_k}{4B'}\delta^2 + \delta\sqrt{c_k+1}\right] \\
&\quad + \epsilon\sqrt{(1+B')(1+\delta)}, \quad (49)
\end{aligned}$$

since $x_{\max} \geq x_{\max,k-1} \geq x_{\min,k-1} \geq x_{\min}$ and $K \geq c_k$. Note that the factor multiplied with δ in the first term in the RHS above is non-negative. Therefore, applying Lemmas 4.2, 4.4 and 4.3 it follows that,

$$\begin{aligned}
&x_{\min,k-1} \\
&> \left[x_{\max,k-1}\sqrt{B'b}\left(\frac{c_k}{4B'}+1\right) - x_{\min,k-1}\right]\delta_{b,p,1,L'}(K,0) \\
&+ x_{\max,k-1}\sqrt{B'b}\left[\frac{c_k}{4B'}\delta^2_{b,p,1,L'}(\left|T^{k-1}\right|,1)\right. \\
&\left.+\delta_{b,p,1,L'}(K,1)\sqrt{c_k+1}\right] \\
&+ \epsilon\sqrt{(1+B')(1+\delta_{b,p,1,L'}(\left|T^{k-1}\right|,2))}, \quad (50)
\end{aligned}$$

which, after a rearrangement, is identical to the condition (47). In the above we have used the following inequalities to arrive at the final inequality: $\delta = \delta_{b,p,L',L'}(K-1,2) \geq \delta_{b,p,L',L'}(K,1)$(Lemma 4.4), $\delta_{b,p,L',L'}(K,1) \geq \delta_{b,p,1,L'}(K,0)$(Lemmas 4.3 and 4.2), $\delta_{b,p,L',L'}(K,1) \geq \delta_{b,p,L',L'}(|T^{k-1}|,1) \geq \delta_{b,p,1,L'}(|T^{k-1}|,1)$ (by Lemma 4.2 followed by Lemma 4.3), and finally, $\delta_{b,p,L',L'}(K-1,2) \geq \delta_{b,p,1,L'}(K-1,2) \geq \delta_{b,p,1,L'}(|T^{k-1}|,2)$(by Lemma 4.3 followed by Lemma 4.2).

V. Simulation Results

In this section, we compare the recovery success performance of TSGBOMP with BOMP, over a range of sparsity values for different sensing matrix sizes. In all the simulation exercises, we generate random matrices of size $m \times n$, having i.i.d. Gaussian entries with zero mean and unit variance, and columns normalized to have unit norm. For the unknown signal, we generate a PIBS vector with parameters $(b, p, L, L', K, 0)$. Here the parameters n, b, p, L, K are chosen such that all the true clusters of such a vector can be accommodated within a signal length of n, and $1 \leq p \leq K$. Once the support of such a vector is generated, the entries are filled randomly with values ± 10. The results are averaged over 1000 such trials. All the experiments are carried out on MATLAB 2017a running on a Core i-5 Laptop with 8 GB RAM, 64 b processor and 1.60 GHz processor speed.

In the experiments below, we study the recovery probability of TSGBOMP in terms of recovery of the true support of the unknown vector. For our experiments, we take $n = 200,\ b = 4$ and $p = 1, 2$. Furthermore, we take $L = 8$ and maintain the well-separation between consecutive cluster locations, that is, we maintain at least $L' = L + 2B - b$ zeros between consecutive clusters where $B = bp$, and the true clusters are placed randomly within the vector subject to this.

1) Probability of Recovery Vs. Sparsity: In our first experiment, we fix the number of measurements to $m = 160$ and vary the sparsity K over the range 1 to 16and compare the frequencies of exact support recovery of TSGBOMP and BOMP. The resulting plots are shown in Figs. 3(a) and 3(b). It is easy to see from Fig. 3 that while TSGBOMP exhibits very good recovery probability performance, the BOMP algorithm, on the other hand, performs quite poorly in recovering the unknown vector. This is, however, to be expected, as the BOMP algorithm

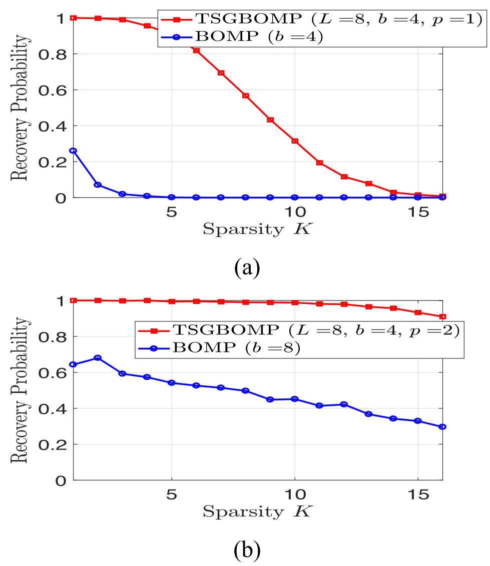


Fig. 3. Probability of recovery vs sparsity for $n = 200,\ m = 160$.

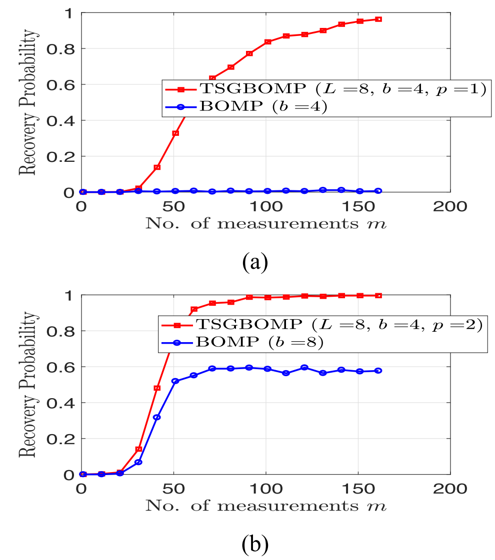


Fig. 4. Probability of recovery vs no of measurements for $b = 4, L = 8, K = 4$. In Fig. (a), $p = 1$, and in Fig. (b), $p = 2$.

was not designed to consider recovery of the general block structured vectors considered in this paper which might not exactly fit into prespecified block boundaries, and instead, might have partial overlaps in consecutive windows. We also observe from Figs. 3(a) and 3(b) that the recovery performance of TSGBOMP improves significantly when p is increased. This improvement can be explained by the fact that increasing p produces PIBS vectors with larger true clusters having sizes varying from b to bp, while keeping the total size of the sum of the nonzero blocks to Kb. This implies that once a true cluster is found, usually more than just one block of size b are detected, which reduces the chances of false detection, thereby increasing the probability of true signal recovery.

2) Probability of Recovery Vs. Number of Measurements: In our next simulation experiment, we compare the recovery probabilities of TSGBOMP and BOMP at various values of the number of measurements m. We choose $K = 4$ for the experiments and vary m from 1 to 200. It can be seen from the plots in Figs. 4(a) and 4(b) that TSGBOMP enjoys far superior probability recovery performance compared to BOMP for the whole range of number of measurements used in the experiment. When $p = 1$, BOMP can hardly recover the support of the true clusters whereas TSGBOMP enjoys much higher probability of recovery. On the other hand, for $p = 2$, the effective maximum cluster length can be 8 and in this case the performance of BOMP is somewhat improved while the performance of TSGBOMP is even better than the $p = 1$ case. As the value of p increases, the BOMP partitions the signal length into larger block lengths and thus the chances of finding true clusters increase. However, BOMP might fail to recover those clusters that lie at the intersection of two consecutive blocks of BOMP. Since TSGBOMP is tailor-made for such signal structures, its recovery is much better and it can detect even those clusters lying at the intersection of consecutive windows.

3) An Application - ECG Signal Recovery: We next apply the TSGBOMP and BOMP algorithms to the recovery of electrocardiogram (ECG) signals, as a test case. The ECG signal we have considered is a noisy version of an ideal ECG signal and is shown in Fig. 5[7], having length $n = 250$. The structure of an ideal ECG signal segment consists of three prominent clusters in the signal, namely the QRS complex in the middle, the P cluster at the left and the T cluster at the right. The sizes of these clusters are different in general, although the separation between two consecutive clusters, for example P and QRS or QRS and T might not be too large, compared to the size of the largest among the P, QRS and T clusters. Nevertheless, we still apply the TSGBOMP to recover this signal from a few compressed measurements obtained using a 125×250 random matrix with zero mean, unity variance i.i.d. Gaussian entries. We compare our recovered signal with that obtained by applying BOMP. Since the ECG signal of Fig. 5 is a noisy signal, we measure the normalized mean squared deviation (NMSD) $\|\hat{\boldsymbol{x}} - \boldsymbol{x}/\|_2 \|\boldsymbol{x}\|_2$, where $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ are the original and the recovered signals respectively. As for the parameters, we choose $b = 20,\ L = bp$ and $K = 5, 4$ with $p = 1$ for $K = 5$ and $p = 2$ for $K = 4$. The recovered signals along with their respective NMSD are shown in Figs. 6(a) and 6(b). From these plots, it can be seen that the NMSD of TSGBOMP is quite lower than that of BOMP. Furthermore,

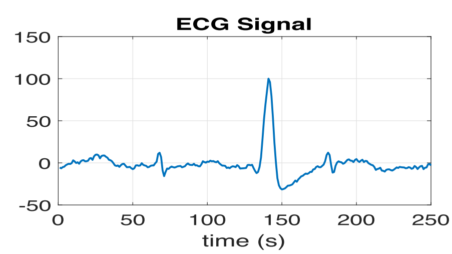


Fig. 5. A real world ECG signal [11].

[7]The data can be found in https://sites.google.com/site/researchbyzhang/bsbl [11]

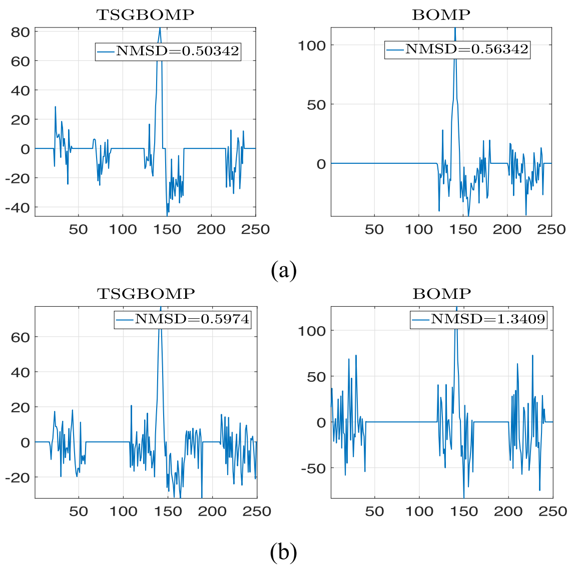


Fig. 6. ECG signal recovery using TSGBOMP and BOMP. Here $b = 20$ and $L = bp$, Fig. (a) uses $p = 1, K = 5$, and Fig. (b) uses $p = 2, K = 4$.

TSGBOMP can reconstruct distinct clusters which relate to the clusters in the real world signal, while BOMP may not always succeed in doing so.

VI. Conclusions

In this paper, a novel greedy algorithm called TSGBOMP is proposed to recover a block sparse vector from small number of measurements, where the cluster (of blocks) sizes are nonuniform and their locations are not known a priori. The algorithm uses a two stage procedure. In the first stage, it selects a window which provides a coarse estimate of the location of a nonzero cluster of the unknown vector, while in the second stage, it tries to make a fine estimate of the exact location of the cluster within the selected window. For analyzing the algorithm, a new block structure called the pseudo-block interleaved block structure (PIBS) is defined and the concept of RIP is extended to obtain a PIBS analog of block RIP (PIBRIP). A recovery analysis of TSGBOMP is carried out and conditions for recovery of the exact support in terms of PIBRIP are presented.

Appendix A

Proofs of the Lemmas in Section IV-A2

1) Proof of Lemma 4.1: The result is a straightforward implication of the definition 4.1 of PIBRIC. For details, please refer to the accompanying supplementary material.

2) Proof of Lemma 4.2: The proof follows immediately after writing down the expressions for $\delta_{b,p,l,L'}(K_i, R_j),\ 1 \le i, j \le 2$ using Eq (12) from definition 4.1 and subsequently noting that if $K_i(\text{resp. } R_i) \le K_j(\text{resp. } R_j)$ then the collection of sets over which the search for $\delta_{b,p,l,L'}(K_j, R_j)$ is conducted is a superset of the collection of sets over which $\delta_{b,p,l,L'}(K_i, R_i)$ is searched.

3) Proof of Lemma 4.3: Choose some k, r such that $0 \le k \le K,\ \ 0 \le r \le R,\ R = 0, 1, 2$. Let $S \subset \mathcal{H}$ be an arbitrary set $S \in \Sigma_{b,p,l,L'}(k, r)$. Then, the Lemma 4.3 will be proved if we can prove that $\|\boldsymbol{\Phi}_S^t \boldsymbol{\Phi}_S - \boldsymbol{I}_S\|_{2\to 2} \le \delta_{b,p,L',L'}(K, R)$, i.e., if we can show that there exists a set $S' \subset \mathcal{H}$ such that $S' \in \cup_{k=0}^{K} \cup_{r=0}^{R} \Sigma_{b,p,L',L'}(k, r)$ and $\|\boldsymbol{\Phi}_S^t \boldsymbol{\Phi}_S - \boldsymbol{I}_S\|_{2\to 2} \le \|\boldsymbol{\Phi}_{S'}^t \boldsymbol{\Phi}_{S'} - \boldsymbol{I}_{S'}\|_{2\to 2}$, which in turn, will be proved if one can show that $S \subset S'$ for some $S' \in \Sigma_{b,p,L',L'}(k', r')$ for some $0 \le k' \le K$, and $0 \le r' \le R$. The rest of the proof is devoted to constructing such S' for any given S for different values of $0 \le r \le R$, for $R = 0, 1, 2$. Although such constructions require simple combinatorial arguments, they are rather long and tedious. The details can be found in the accompanying supplementary material.

4) Proof of Lemma 4.4: Consider a PIBS vector with parameters $(b, p, L', L', K, 1),\ K \ge 1$ and let S denote its support set. The solitary pseudo block may lie either in between two consecutive true clusters or beside just one true cluster. Since we are dealing with both $\delta_{b,p,L',L'}(K, 1)$ as well $\delta_{b,p,L',L'}(K-1, 2)$, it is implicit that the size of the index set n is large enough to accommodate both K true blocks together with one pseudo block of size L', as well as $K-1$ true blocks together with two pseudo blocks of size L' each. It then easily follows that b consecutive indices (i.e., one true block) either from the front of the true cluster to the left or from the rear of the true cluster to the right can be covered by part of a pseudo block of size L', with the remaining part of the pseudo block of size $(L' - b)$ supported in the inter-cluster space. This results in a support set S' with $S \subset S'$ and thus, $\delta_{b,p,L',L'}(K, 1) \le \delta_{b,p,L',L'}(K-1, 2)$.

5) Proof of Lemma 4.5: The proof follows by combining the arguments of the proof of Lemma 5 of [21] with necessary modifications for the PIBS structure. The details of the proof can be found in the accompanying supplementary material.

6) Proof of Lemma 4.6: The proof follows the same steps as used to prove the inequality (3.7) of [22], using Lemma 4.5 in place of the corresponding result for conventional RIC. The details can be found in the enclosed supplementary material.

7) Proof of Lemma 4.7: First, note that $\|\mathbf{P}_S^{\perp} \boldsymbol{\phi}_j\|_2^2 = 1 - \|\mathbf{P}_S \boldsymbol{\phi}_j\|_2^2$. One can express $\mathbf{P}_S \boldsymbol{\phi}_j$ as $\boldsymbol{\Phi}_{|S\rangle} \boldsymbol{z}_{|S\rangle}$, for some vector $\boldsymbol{z} \in \mathbb{R}^n$ supported on $|S\rangle$. Define $S' = |S\rangle \cup \{j\}$. Then, $\|\mathbf{P}_S \boldsymbol{\phi}_j\|_2^2 = \langle \mathbf{P}_S \boldsymbol{\phi}_j, \mathbf{P}_S \boldsymbol{\phi}_j \rangle = \langle \mathbf{P}_S \boldsymbol{\phi}_j, \boldsymbol{\phi}_j \rangle = \langle \boldsymbol{\Phi}_{S'} \boldsymbol{z}_{S'}, \boldsymbol{\Phi}_{S'} \boldsymbol{e}_{j,S'} \rangle$, where $\boldsymbol{e}_j \in \mathbb{R}^n$, is the standard j^{th} basis vector[8]. Note that $\langle \boldsymbol{z}, \boldsymbol{e}_j \rangle = 0$, since $j \notin |S\rangle$. By assumption, S' can be associated with a PIBS vector with parameters $(b, p, 1, L', k, 1),\ 0 \le k \le K$. Defining $\lambda = \frac{\lambda_{\max}(\boldsymbol{\Phi}_{S'}^t \boldsymbol{\Phi}_{S'})}{\lambda_{\min}(\boldsymbol{\Phi}_{S'}^t \boldsymbol{\Phi}_{S'})}$ and $\delta = \frac{1+\delta_{b,p,1,L'}(K,1)}{1-\delta_{b,p,1,L'}(K,1)}$, we have, using the PIBRIP, $\lambda \le \delta$, as $\delta_{b,p,1,L'}(K, 1) < 1$. From this and the fact that $f(\lambda) = \frac{\lambda - 1}{\lambda + 1}$ is an increasing function in λ, from Weilandt'stheorem [23, Theorem 7.4.34][9], we have, $\|\mathbf{P}_S \boldsymbol{\phi}_j\|_2^2 \le \delta_{b,p,1,L'}(K, 1) \|\boldsymbol{\Phi}_{S'} \boldsymbol{z}_{S'}\|_2 \|\boldsymbol{\phi}_j\|_2 = \delta_{b,p,1,L'}(K, 1) \|\mathbf{P}_S \boldsymbol{\phi}_j\|_2 \Rightarrow \|\mathbf{P}_S \boldsymbol{\phi}_j\|_2 \le \delta_{b,p,1,L'}(K, 1)$, where we have used the fact that $\|\boldsymbol{\phi}_j\|_2 = 1$, and $\boldsymbol{\Phi}_{S'} \boldsymbol{z} = \mathbf{P}_S \boldsymbol{\phi}_j$. Hence the result follows.

[8] $\boldsymbol{e}_j$ has all its entries 0 except for the j^{th} entry, which is 1.

[9] Wielandt's theorem states that for any two vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$, that are orthogonal, i.e., $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, and for any symmetric positive definite matrix $\boldsymbol{B} \in \mathbb{R}^{n \times n}$, one has $|\boldsymbol{u}^t \boldsymbol{B} \boldsymbol{v}|^2 \le \left(\frac{\lambda_{\max}(\boldsymbol{B}) - \lambda_{\min}(\boldsymbol{B})}{\lambda_{\max}(\boldsymbol{B}) + \lambda_{\min}(\boldsymbol{B})}\right)^2 (\boldsymbol{u}^t \boldsymbol{B} \boldsymbol{u})(\boldsymbol{v}^t \boldsymbol{B} \boldsymbol{v})$

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York, NY, USA: Springer, 2013.

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

[4] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi, "Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 3, pp. 275–285, Jun. 2008.

[5] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Jan. 2009.

[6] J. Chen and X. Huo, "Sparse representations for multiple measurement vectors (MMV) in an over-complete dictionary," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, Mar. 2005, pp. iv257–iv260.

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

[8] J. Wen, Z. Zhou, Z. Liu, M.-J. Lai, and X. Tang, "Sharp sufficient conditions for stable recovery of block sparse signals by block orthogonal matching pursuit," *Appl. Comput. Harmon. Anal.*, vol. 47, no. 3, pp. 948–974, Nov. 2019.

[9] 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, Jun. 2012.

[10] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 101–111, Jan. 2003.

[11] Z. Zhang and B. D. Rao, "Extension of SBL algorithms for the recovery of block sparse signals with intra-block correlation," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 2009–2015, Jan. 2013.

[12] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 360–372, Nov. 2015.

[13] M. A. and A. P. Kannu, "Sinusoid signal estimation using generalized block orthogonal matching pursuit algorithm," in *Proc. Int. Conf. Signal Process. Commun.*, Bangalore, India, Jul. 2018, pp. 60–64.

[14] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[15] S. Mukhopadhyay and M. Chakraborty, "A two stage generalized block orthogonal matching pursuit (TSGBOMP) algorithm," Aug. 2020, *arXiv:2008.08031v1* [*cs.IT*].

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

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

[18] 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.

[19] J. Wen, Z. Zhou, D. Li, and X. Tang, "A novel sufficient condition for generalized orthogonal matching pursuit," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 805–808, Dec. 2017.

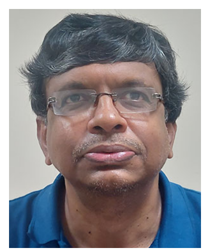
[20] J. Wen, J. Wang, and Q. Zhang, "Nearly optimal bounds for orthogonal least squares," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5347–5356, Jul. 2017.

[21] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jun. 2011.

[22] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, Dec. 2011.

[23] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix Analysis*. Cambridge Univ. Press, 1990.

Samrat Mukhopadhyay (Member, IEEE) received the B.E. degree from the Indian Institute of Engineering Science and Technology, Shibpur, the M.E. degree from the Indian Institute of Science, Bangalore, and the Ph.D. degree from the Indian Institute of Technology, Kharagpur, in 2012, 2014, and 2021, respectively. He is currently an Assistant Professor with the Electronics Engineering Department of Indian Institute of Technology (Indian School of Mines), Dhanbad, India. His research interests include compressed sensing, adaptive filtering, high dimensional statistics, and machine learning and optimization.



Mrityunjoy Chakraborty (Senior Member, IEEE) received the B.E. degree in electronics and telecommunication engineering from Jadavpur University, Kolkata, in 1983, the M.Tech. degree in 1985 and the Ph.D. degree in 1994 from Indian Institute of Technology at Kanpur and Delhi respectively, both in electrical engineering. He joined the Department of Electronics and Electrical Communication Engineering at the Indian Institute of Technology, Kharagpur as a Lecturer in 1994 and is currently a Professor in the same department, with teaching and research interests in digital and adaptive signal processing, VLSI signal processing, linear algebra and compressive sensing.

He has been a Senior Editorial Board (SEB) member of the IEEE SIGNAL PROCESSING MAGAZINE from 2017 to 2020. Earlier, he had been an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, Part I (2004–2007, 2010–2012) and Part II (2008–2009), apart from being a SEB member of the IEEE JOURNAL ON EMERGING TECHNIQUES IN CIRCUITS AND SYSTEMS (2016–2017). He served as a Chair of the DSP Technical Committee (TC) of the IEEE Circuits and Systems Society during 2016–2018. He has also been a Guest Editor of the EURASIP JASP and also special issues of TCAS-II, DSP track Co-Chair of ISCAS during 2015–2021, TPC Co-Chair of IEEE SIPS-2018, Special Session Co-Chair of DSP-18 and a Gabor Track Chair of DSP-15. He is a co-founder of the Asia Pacific Signal and Information Processing Association (APSIPA), has been a member of the APSIPA BOG (2013–2016) and also, served as Chair of the APSIPA TC on Signal and Information Processing Theory and Methods. He has also been the General Chair of the National Conference on Communications - 2012, 2020.

Prof. Chakraborty is a Fellow of the National Academy of Sciences, India, and also of the Indian National Academy of Engineering (INAE). Recently he received the prestigious Chair Professorship of the INAE. During 2012-2013, he was selected as a Distinguished Lecturer of the APSIPA.