Journal of Mathematical Research with Applications
May, 2023, Vol. 43, No. 3, pp. 363–378
DOI:10.3770/j.issn:2095-2651.2023.03.010
Http://jmre.dlut.edu.cn

Piecewise Sparse Recovery in Union of Bases

Chongjun LI*, **Yijun ZHONG**
School of Mathematical Sciences, Dalian University of Technology, Liaoning 116024*, P. R. China*

Abstract Sparse recovery (or sparse representation) is a widely studied issue in the fields of signal processing, image processing, computer vision, machine learning and so on, since signals such as videos and images, can be sparsely represented under some frames. Most of fast algorithms at present are based on solving l^0 or l^1 minimization problems and they are efficient in sparse recovery. However, the theoretically sufficient conditions on the sparsity of the signal for l^0 or l^1 minimization problems and algorithms are too strict. In some applications, there are signals with structures, i.e., the nonzero entries have some certain distribution. In this paper, we consider the uniqueness and feasible conditions for piecewise sparse recovery. Piecewise sparsity means that the sparse signal $\mathbf{x}$ is a union of several sparse sub-signals $\mathbf{x}_i$ $(i = 1, 2, \ldots, N)$, i.e., $\mathbf{x} = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \ldots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}}$, corresponding to the measurement matrix A which is composed of union of bases $A = [A_1, A_2, \ldots, A_N]$. We introduce the mutual coherence for the sub-matrices A_i $(i = 1, 2, \ldots, N)$ by considering the block structure of A corresponding to piecewise sparse signal $\mathbf{x}$, to study the new upper bounds of $\|\mathbf{x}\|_0$ (number of nonzero entries of signal) recovered by both l^0 and l^1 optimizations. The structured information of measurement matrix A is exploited to improve the sufficient conditions for successfully piecewise sparse recovery and also improve the reliability of l_0 and l_1 optimization models on recovering global sparse vectors.

Keywords piecewise sparse recovery; union of bases; mutual coherence; greedy algorithm; BP method

MR(2020) Subject Classification 94A12

1. Introduction

In this paper, we consider recovering a sparse signal (vector) $\mathbf{x}^* \in \mathbb{R}^n$ from an underdetermined system of linear equation

$$A\mathbf{x}^* = \mathbf{b}, \tag{1.1}$$

where $\mathbf{b} \in \mathbb{R}^m$ is a measurement vector, $A \in \mathbb{R}^{m\times n}$ is a measurement matrix. If the vector $\mathbf{x}^*$ has at most $s \le m < n$ nonzero entries, then it is named as s-sparse vector, the corresponding index set of nonzero entries is called support $\mathbf{S} = \mathrm{supp}(\mathbf{x}^*)$. There are many theories, algorithms and applications on this problem of sparse recovery [1].

Received September 5, 2022; Accepted January 8, 2023
Supported by the National Natural Science Foundation of China (Grant Nos. 11871137; 11572081) and the Fundamental Research Funds for the Central Universities of China (Grant No. QYWKC2018007).
* Corresponding author
E-mail address: chongjun@dlut.edu.cn (Chongjun LI)

One kind of approaches for solving Eq. (1.1) are the greedy algorithms (GA), which approximate the following l^0 minimizing solution, named as P_0 problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \quad \text{s.t.} \quad A\mathbf{x} = \mathbf{b}. \tag{1.2}$$

One of the most popular greedy methods is the orthogonal matching pursuit (OMP) as proposed in [2–4]. It iteratively appends components to the support of the approximation $\mathbf{x}^k$ whose correlation to the current residual is maximal. There are many other greedy methods for sparse recovery, for example, iterative hard thresholding (IHT) [5], stagewise OMP (StOMP) [6], regularized OMP (ROMP) [7, 8], compressive sampling matching pursuit (CoSaMP) [9], subspace pursuit (SP) [10], iterative thresholding with inversion (ITI) [11], hard thresholding pursuit (HTP) [12] and so on.

Another kind of approaches are convex relaxation algorithms which solve a convex program whose minimizer is known to approximate the target signal. Among them, the basis pursuit gains lots of attention which determines the sparsest representation of $\mathbf{x}^*$ by solving the following l^1 minimization problem, named as P_1 problem or Basis Pursuit problem (BP method):

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1, \quad \text{s.t.} \quad A\mathbf{x} = \mathbf{b}. \tag{1.3}$$

Many algorithms have been proposed to complete the optimization, including interior-point methods [13], projected gradient methods [14], and iterative thresholding [15] etc.

There are three fundamental problems concerned on sparse recovery:

(1) Uniqueness of solution of the P_0 problem.

(2) Feasibility of GA for solving the P_0 problem.

(3) Equivalence between the P_1 problem and the P_0 problem, or feasibility of BP method.

There are several tools for dealing with the above three problems, such as mutual coherence [16], the spark [17], the cumulative coherence [18], the exact recovery coefficient (ERC) [18], the restricted isometry property (RIP) conditions and the restricted isometry constants (RICs) [19–21]. It is well-known that the necessary and sufficient condition for the uniqueness of the solution of the P_0 problem (1.2) is [17]

$$\|\mathbf{x}\|_0 < \mathrm{spark}(A)/2, \tag{1.4}$$

or the RIC of the matrix A satisfies $\delta_{2s} < 1$ (see [19]). The equivalence between the P_1 problem and the P_0 problem is guaranteed by $\delta_{2s} < \sqrt{2} - 1$ (see [20, 21]). For a given matrix or dictionary A, however, it is difficult to compute the spark or verify the RIP conditions. By contrast, we can easily compute the mutual coherence of matrix. The general case discussed in [16] showed that one sufficient condition which ensures the uniqueness of the solution of the P_0 problem (1.2) is

$$\|\mathbf{x}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right), \tag{1.5}$$

where $\mu(A)$ is the mutual coherence of measurement matrix A. Furthermore, the condition Eq. (1.5) is also a sufficient condition which ensures the OMP (greedy method) and BP method for recovering the optimal s-sparse solution [18]. However, in applications, OMP or BP method can work well even when the condition (1.5) is not satisfied, i.e., when

$$\frac{1}{2}(1 + 1/\mu(A)) \leq \|\mathbf{x}\|_0 < \text{spark}(A)/2,$$

which means the sufficient condition (1.5) is strict for sparse recovery to some extent, or the "gap" between the optimal upper bound $\frac{1}{2}\text{spark}(A)$ and the upper bound $\frac{1}{2}(1 + 1/\mu(A))$ is big.

Two improved conditions were obtained in the special case where A is in pairs of orthogonal bases. The result in [22] shows that uniqueness of the l^0 minimization (P_0 problem) solution can be achieved for improved condition $\|\mathbf{x}\|_0 < \frac{1}{\mu}$, where $\mu = \mu(A)$. They also showed that the solutions of the P_0 and P_1 problems coincide for $\|\mathbf{x}\|_0 < \frac{\sqrt{2}-0.5}{\mu}$. Further, it was also shown in [18, 22, 23] that if the matrix A is a union of $N(\geq 2)$ orthogonal bases, the sufficient condition for OMP to solve the P_0 problem for N orthogonal bases was improved with

$$\|\mathbf{x}\|_0 < \left(\frac{1}{2} + \frac{1}{2(N-1)}\right)\frac{1}{\mu} \tag{1.6}$$

and for BP to solve P_1 problem was improved with

$$\|\mathbf{x}\|_0 < \left(\sqrt{2} - 1 + \frac{1}{2(N-1)}\right)\frac{1}{\mu}. \tag{1.7}$$

Moreover, if we consider the vector $\mathbf{x}$ as $(s_1, \ldots, s_N)$-piecewise sparse (see Definition 1.1), $\|\mathbf{x}\|_0 = s_1 + \cdots + s_N$, $s_1 \leq \cdots \leq s_N$) or vector $\mathbf{b}$ is a superposition of s_i atoms from the i-th basis, with A in unions of N orthogonal bases, then the exact recovery condition (ERC) is guaranteed by

$$\sum_{j=2}^{N} \frac{\mu s_j}{1 + \mu s_j} < \frac{1}{2(1 + \mu s_1)}. \tag{1.8}$$

Thus, both the OMP and BP can find the sparest solution under condition (1.8) (see [18]). When $n = 2$, Eq. (1.8) is $2\mu^2 s_1 s_2 + \mu s_2 < 1$ $(s_1 \leq s_2)$.

Condition index	Structure of A	The upper bounds of $\|\mathbf{x}\|_0(= s_1 + s_2)$
Condition 1 Eq. (1.5)	General case	$\|\mathbf{x}\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(A)})$
Condition 2 Eq. (1.6)	Pais of orthogonal bases (uniqueness)	$\|\mathbf{x}\|_0 < \frac{1}{\mu(A)}$
Condition 3 Eq. (1.7)	Pais of orthogonal bases (equivalence)	$\|\mathbf{x}\|_0 < \frac{\sqrt{2}-0.5}{\mu(A)}$
Condition 4 Eq. (1.8)	Pais of orthogonal bases (ERC)	$2\mu^2 s_1 s_2 + \mu s_2 < 1$ $(s_1 \leq s_2)$
		$2\mu^2 s_1 s_2 + \mu s_1 < 1$ $(s_1 > s_2)$

Table 1 A list of theoretical upper bounds for sparse recovery

The four Conditions (1.5), (1.6), (1.7) and (1.8) are concluded in Table 1 for the case of A in pairs of orthogonal bases compared with the general case of A. It is observed from Figure 1 (presented in [22]) that the sufficient conditions based on the mutual coherence can be improved by considering the structure of matrix A.

Note that, in many practical applications, the measurement matrix A cannot be always composed of a union of orthogonal bases. Thus, it is necessary to study the sufficient conditions for successful recovery when the measurement matrix A is a union of N general bases or submatrices, i.e., $A = [A_1, A_2, \ldots, A_N]$, where A_i $(i = 1, 2, \ldots, N)$ are not necessary orthogonal bases, each A_i may be a general base or dictionary. Correspondingly, the target signal (vector) $\mathbf{x}$ is also partitioned into N sub-vectors or segments $\mathbf{x} = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \ldots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}}$ according to the structure of matrix A. With considering the piecewise sparsity of vector $\mathbf{x}$ and the structure of matrix A make it possible to deeply study the sufficient conditions for successfully sparse recovery.

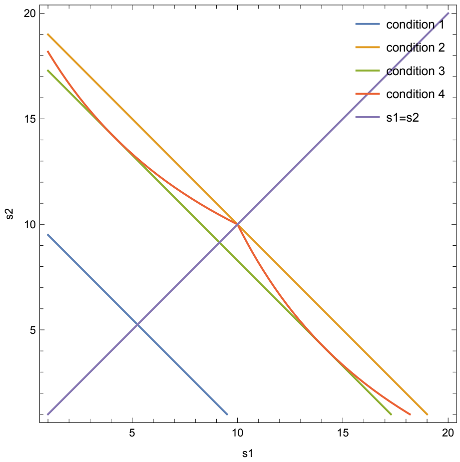

Figure 1 A plot of the upper bound conditions of the sparse vector in Table 1 with $\mu(A) = 0.05$

For a given vector $\mathbf{x} = (\underbrace{x_1, \ldots, x_{d_1}}_{\mathbf{x}_1^{\mathrm{T}}}, \underbrace{x_{d_1+1}, \ldots, x_{d_1+d_2}}_{\mathbf{x}_2^{\mathrm{T}}}, \ldots, \underbrace{x_{n-d_N+1}, \ldots, x_n}_{\mathbf{x}_N^{\mathrm{T}}})^{\mathrm{T}}$, where $n = \sum_{i=1}^N d_i$ and $s_i = \|\mathbf{x}_i\|_0$, $i = 1, 2, \ldots, N$, there are three common types of sparsity of vector $\mathbf{x}$:

(1) Global sparsity. $\mathbf{x}$ is assumed to have $s = \|\mathbf{x}\|_0 = \sum_{i=1}^N \|\mathbf{x}_i\|_0$ nonzero entries.

(2) Block sparsity [24–27]. A block s-sparse vector $\mathbf{x}$ is assumed to have at most s blocks with nonzero entries, i.e., the block l^0 or l^1 norm $\|\mathbf{x}\|_{2,0} = \sum_{i=1}^N I(\|\mathbf{x}_i\|_2)$ or $\|\mathbf{x}\|_{2,1} = \sum_{i=1}^N \|\mathbf{x}_i\|_2$ are minimized.

(3) Piecewise sparsity. As in the following definition.

Definition 1.1 *A vector* $\mathbf{x} = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \ldots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^n$ *is partitioned into* N *segments and it is assumed that each sub-vector* $\mathbf{x}_i \in \mathbb{R}^{n_i}$ *is sparse,* $s_i = \|\mathbf{x}_i\|_0$, $i = 1, 2, \ldots, N$. *We call the vector* $\mathbf{x}$ $(s_1, s_2, \ldots, s_N)$*-piecewise sparse.*

Piecewise sparse vector means that each part of the vector is sparse. The piecewise sparsity is different from the block sparsity. Piecewise sparse recovery is common in applications, such as the problem of the decomposition of texture part and cartoon part of image in [28], i.e., $\mathbf{b} = A_n \mathbf{x}_n + A_t \mathbf{x}_t$ where n and t represent the cartoon and texture. It is assumed that both parts can be represented in some given dictionaries, thus $\mathbf{x}_n$ and $\mathbf{x}_t$ are two sparse vectors. The coefficient vector $\mathbf{x} = (\mathbf{x}_n^{\mathrm{T}}, \mathbf{x}_t^{\mathrm{T}})^{\mathrm{T}}$ is a "piecewise" sparse vector. Another example is the problem of reconstructing a surface from scattered data in approximation space $H = \bigcup_{i=1}^N H_j$, where $H_j \subseteq H_{j+1}$ are principal shift invariant (PSI) spaces generated by a single compactly supported function [29], the fitting surface is $g = \sum_{i=1}^N g_i$, $g_i \in H_i$ with $g_i = \sum_{j=1}^{n_i} c_j^i \phi_j^i$. The coefficients vector $\mathbf{c} = (\mathbf{c}^1, \mathbf{c}^2, \ldots, \mathbf{c}^N)^{\mathrm{T}}$ (with N pieces $\mathbf{c}^i = (c_1^i, \ldots, c_{n_i}^i)^{\mathrm{T}}$, $i = 1, 2, \ldots, N$) is to be determined. Due to the property of PSI spaces, the coefficients to be determined by l^1 minimization in [29] are "piecewise" sparse structure, i.e., each $\mathbf{c}^i \in \mathbb{R}^{n_i}$ is a sparse vector in H_i. In [30], we first try to recover the piecewise sparse vector by the piecewise inverse scale space algorithm with deletion rule.

We study the sparse vectors whose nonzero elements appear in a scattered way, i.e., piecewise sparsity, and the corresponding matrix can be structured in a union of some sub-matrices (orthogonal bases is a special case). It is obvious that piecewise sparsity is more general in applications. By considering the piecewise sparse vector and the structure of its corresponding matrix $A = [A_1, A_2, \dots, A_N]$ we may obtain improved conditions for sparse recovery.

In this paper, we use the mutual coherence and cumulative mutual coherence which can be efficiently calculated for an arbitrary given matrix A to give the new sufficient conditions of piecewise sparse recovery. Since our results are based on the mutual coherence of A and A_i $(i = 1, 2, \dots, N)$, thus the results are applicable to arbitrary structured dictionary $A = [A_1, A_2, \dots, A_N]$. Inspired by the works in [18, 23], which provide improved sufficient conditions for having unique sparse representation of signals in union of orthogonal bases, we study the generalization of the sufficient conditions for having unique sparse representation of piecewise sparse signals corresponding to the unions of general bases (or dictionaries).

2. Preliminaries

In this section, we introduce some necessary notations and definitions as follows.

Notations. We use $\mathbf{x}$ to represent a vector and x to represent a scalar. Define the inner product by $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^{\mathrm{T}}\mathbf{y}$. For a sparse vector $\mathbf{x}^*$, let $\mathbf{S} = \mathrm{supp}(\mathbf{x}^*)$ and T be its complement, i.e., $T = \{i : x_i^* = 0\}$. Denote by $A_{\mathbf{S}}$ the submatrix of A formed by the columns of A in $\mathbf{S}$, which are assumed to be linearly independent. Similarly, define A_T so that $[A_{\mathbf{S}} \;\; A_T] = A$. Denote the number of entries in $\mathbf{S}$ by $s = |\mathbf{S}| = |\mathrm{supp}(\mathbf{x}^*)|$, for a piecewise sparse vector $\mathbf{x}^*$ with piecewise support $\mathbf{S} = S_1 \bigcup S_2 \bigcup \cdots \bigcup S_N$, denote $s_i = |S_i|$, $i = 1, 2, \dots, N$.

2.1. Tools used in sparse recovery

In this part, we introduce the widely used tool for sparse signal recovery: the mutual coherence of a matrix of dictionary $A \in \mathbb{R}^{m \times n}$. Denote $\mathbf{a}_k^i$ by the k-th column in the submatrix A_i, and the matrix $A = [A_1, A_2, \dots, A_N]$ is assumed to have unit l^2 norm for each column, i.e., $\|\mathbf{a}_k^i\|_2 = 1,\ k = 1, 2, \dots, n_i,\ i = 1, 2, \dots, N$.

Definition 2.1 ([16]) *The mutual coherence of A is*

$$\mu = \mu(A) := \max_{k \neq l} |\langle \mathbf{a}_k, \mathbf{a}_l \rangle|,$$

for any two differential columns $\mathbf{a}_k, \mathbf{a}_l \in A$.

Roughly speaking, the coherence measures how much two vectors in the dictionary can look alike. It is obvious that every orthogonal basis has coherence $\mu = 0$. A union of two orthogonal bases has coherence $\mu \geq m^{-1/2}$ (see [16]).

Definition 2.2 ([18]) *The cumulative mutual coherence (Babel function) of A is*

$$\mu_1(s) = \max_{|S|=s} \max_{l\in S^c} \sum_{k\in S} |\langle \mathbf{a}_k, \mathbf{a}_l\rangle|,$$

where $\mathbf{a}_k,\ \mathbf{a}_l \in A$.

A close examination of the formula shows that $\mu_1(1) = \mu$ and that μ_1 is a non-decreasing function of s.

Proposition 2.3 ([18]) *If a dictionary A has coherence μ, then $\mu_1(m) \le m\mu$.*

Theorem 2.4 ([18]) *Exact Recovery Condition (ERC). A sufficient condition for both Orthogonal Matching Pursuit and Basis Pursuit to recover the s-sparse $\mathbf{x}$ with $\mathbf{S} = \operatorname{supp}(\mathbf{x})$ successfully is that*

$$\max_{j\in \mathbf{S}^c} \|(A_{\mathbf{S}}^{\mathrm{T}} A_{\mathbf{S}})^{-1} A_{\mathbf{S}}^{\mathrm{T}} A_j\|_1 < 1,$$

where the 1-norm is the sum of the absolute value of entries of the vector $(A_{\mathbf{S}}^{\mathrm{T}} A_{\mathbf{S}})^{-1} A_{\mathbf{S}}^{\mathrm{T}} A_j$.

Definition 2.5 ([17]) *The spark of A counts the least number of columns which form a linearly dependent set.*

$$\operatorname{spark}(A) = \min_{\mathbf{x}\in \operatorname{Ker}(A),\ \mathbf{x}\neq 0} \|\mathbf{x}\|_0,$$

where the kernel of the dictionary is defined as $\operatorname{Ker}(A) = \{\mathbf{x} : A\mathbf{x} = 0\}$.

2.2. Tools used in piecewise sparse recovery

Assume that $A = [A_1, A_2, \ldots, A_N]$ is a union of N general bases, we generalize the concepts of mutual coherence and cumulative mutual coherence to the piecewise sparse case.

Definition 2.6 *The i-th sub-matrix coherence of A_i is defined as*

$$\mu^{i,i} = \max_{k\neq l} |\langle \mathbf{a}_k^i, \mathbf{a}_l^i\rangle|,$$

where $\mathbf{a}_k^i,\ \mathbf{a}_l^i \in A_i,\quad i = 1, 2, \ldots, N$.

It is clear that the i-th sub-matrix coherence $\mu^{i,i}$ satisfies $0 \le \mu^{i,i} = \alpha_i\mu \le \mu$ with a factor $\alpha_i \in [0, 1]$.

The parameter α_i for i-th block A_i measures the ratio of coherence within A_i compared with the coherence of the whole matrix A. Especially, when A is a union of N orthogonal bases, then $\mu_{i,i} = 0$ and $\alpha_i = 0$, $i = 1, 2, \ldots, N$.

Definition 2.7 *The cumulative coherence between two blocks A_i and A_j is defined as*

$$\mu_1^{i,j}(m) = \max_{|S_i|=m} \max_{l\in\{1,\ldots,n_j\}} \sum_{k\in S_i} |\langle \mathbf{a}_k^i, \mathbf{a}_l^j\rangle|,$$

where $\mathbf{a}_k^i \in A_i,\ \mathbf{a}_l^j \in A_j$, S_i is the index set of m columns in sub-matrix A_i and n_j is the number of columns in A_j.

Remark 2.8 Notice that the cumulative coherence between two blocks A_i and A_j is different from the definition of cumulative block coherence in [24]. The cumulative block coherence $\mu_{1B}(m)$ measures coherence between m blocks, the m represents the number of blocks. In the cumulative coherence between two blocks A_i and A_j, the m represents the number of columns of A_i.

Remark 2.9 The cumulative coherence between two blocks A_i and A_j is bounded by

$$\mu_1^{i,j}(m) \leq m\mu.$$

Definition 2.10 *The cumulative coherence within A_i is defined as*

$$\mu_1^{i,i}(m) = \max_{|S_i|=m} \max_{l \notin S_i} \sum_{k \in S_i} |\langle \mathbf{a}_k^i, \mathbf{a}_l^i \rangle|,$$

where $\mathbf{a}_k^i$, $\mathbf{a}_l^i \in A_i$.

Remark 2.11 Notice that the cumulative coherence within S_i is different from the definition of cumulative mutual coherence in [18]. The cumulative mutual coherence $\mu_1(m)$ measures the maximum overall coherence between one fixed atom (column of A) and a collection of other atoms [18]. The cumulative coherence within A_i can be seen as a cumulative form of the i-th sub-matrix coherence $\mu_{i,i}$, i.e., $\mu_1^{i,i}$ measures how much the atoms in the same block A_i are "speaking the same language".

Remark 2.12 It is clear that

$$\mu_1^{i,i}(m) \leq \alpha_i m\mu.$$

3. Piecewise sparse recovery in union of general bases

In the piecewise sparse setting, system (1.2) is equivalent to the following problem:

$$\begin{aligned} &\min_{\mathbf{x}} \|\mathbf{x}_1\|_0 + \|\mathbf{x}_2\|_0 + \cdots + \|\mathbf{x}_N\|_0 \\ &\text{s.t.} \quad \mathbf{b} = A_1\mathbf{x}_1 + A_2\mathbf{x}_2 + \ldots + A_N\mathbf{x}_N, \end{aligned} \tag{3.1}$$

where $\mathbf{x} = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \ldots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}}$. Denote (3.1) as Piecewise P_0 problem. Notice that the problem (1.2) is indeed equivalent to problem (3.1). We use the form (3.1) to show the piecewise sparse structure of the signal.

3.1. Uniqueness of piecewise sparse recovery via piecewise P_0 problem

By considering the piecewise P_0 problem (3.1) and the sub-block coherence of the measurement matrix, we can improve the uniqueness condition of piecewise sparse recovery as follows.

Theorem 3.1 *Suppose that the measurement matrix $A = [A_1, A_2, \ldots, A_N]$ is a union of N bases with an overall coherence μ and sub–block coherence parameters α_i for $i = 1, 2, \ldots, N$, if $\mathbf{x}$ is a solution of piecewise P_0 problem (3.1) and satisfies*

$$\|\mathbf{x}\|_0 < \frac{N(1 + \alpha_{\max}\mu)}{2(N - 1 + \alpha_{\max})\mu}, \tag{3.2}$$

where $\alpha_{\max} = \max_{i=1,\ldots,N} \alpha_i$, then $\mathbf{x}$ is the unique solution of problem (3.1).

Proof By the necessary and sufficient condition Eq. (1.4) for the P_0 problem, we need find the lower bound of the spark(A) in the piecewise case.

Let $\mathbf{x} \in \mathrm{Ker(A)}$, $\mathbf{x} \neq 0$. Denote the support of $\mathbf{x}$ by $\mathbf{S} = S_1 \bigcup S_2 \bigcup \cdots \bigcup S_N$ and $s_i = |S_i|$, corresponding to the blocks of $A = [A_1, A_2, \ldots, A_N]$. Then

$$\mathrm{spark}(A) = \min_{\mathbf{x}\in\mathrm{Ker(A)},\ \mathbf{x}\neq 0} \|\mathbf{x}\|_0 = \min_{\mathbf{x}\in\mathrm{Ker(A)},\ \mathbf{x}\neq 0} (s_1 + s_2 + \cdots + s_N).$$

Step 1. We start similarly to the proof in [23, Lemma 3]. Let $r_i = \mathrm{rank}(A_i)$, $i = 1, \ldots, N$. Since $A_{\mathbf{S}} = [A_{S_1}, \ldots, A_{S_N}]$, and

$$\mathbf{x}_{\mathbf{S}} = \begin{bmatrix} \mathbf{x}_{S_1} \\ \vdots \\ \mathbf{x}_{S_N} \end{bmatrix} \in \mathrm{Ker}(A).$$

In order to find the minimum of $s_1 + \cdots + s_N$, we can suppose that $s_i \leq r_i$ and A_{S_i} is full column rank for $i = 1, \ldots, N$. Because $\sum_{i=1}^N A_{S_i}\mathbf{x}_{S_i} = 0$, for every i we have $A_{S_i}\mathbf{x}_{S_i} = -\sum_{j\neq i} A_{S_j}\mathbf{x}_{S_j}$, hence $\mathbf{x}_{S_i} = -\sum_{j\neq i}(A_{S_i}^{\mathrm{T}}A_{S_i})^{-1}(A_{S_i}^{\mathrm{T}}A_{S_j})\mathbf{x}_{S_j}$. Then we can deduce that

$$\|\mathbf{x}_{S_i}\|_1 \leq \frac{1}{1-\mu_1^{i,i}(s_i-1)} \sum_{j\neq i} \|A_{S_i}^{\mathrm{T}}A_{S_j}\|_1 \|\mathbf{x}_{S_j}\|_1 \leq \sum_{j\neq i} \frac{\mu_1^{i,j}(s_i)}{1-\mu_1^{i,i}(s_i-1)} \|\mathbf{x}_{S_j}\|_1.$$

Since $\|\mathbf{x}_{\mathbf{S}}\|_1 = \|\mathbf{x}_{S_1}\|_1 + \cdots + \|\mathbf{x}_{S_N}\|_1$, then

$$\Big(1 + \frac{\max_{j\neq i}\mu_1^{i,j}(s_i)}{1-\mu_1^{i,i}(s_i-1)}\Big)\|\mathbf{x}_{S_i}\|_1 \leq \frac{\max_{j\neq i}\mu_1^{i,j}(s_i)}{1-\mu_1^{i,i}(s_i-1)}\|\mathbf{x}_{\mathbf{S}}\|_1,$$

which results in

$$\|\mathbf{x}_{\mathbf{S}}\|_1 \leq \Big(\sum_{i=1}^N \frac{v_1^i}{v_2^i}\Big)\|\mathbf{x}_{\mathbf{S}}\|_1,$$

where $v_1^i = \max_{j\neq i}\mu_1^{i,j}(s_i)/(1-\mu_1^{i,i}(s_i-1))$ and $v_2^i = 1 + v_1^i$. Thus

$$\sum_{i=1}^N \frac{v_1^i}{v_2^i} \geq 1. \tag{3.3}$$

Using the inequalities:

$$\mu_1^{i,i}(s_i-1) = \alpha_i\mu_1(s_i-1) \leq (s_i-1)\alpha_i\mu$$

and

$$\mu_1^{i,j}(s_i) \leq s_i\mu,$$

the inequality (3.3) becomes

$$\sum_{i=1}^N \frac{s_i\mu}{1-(s_i-1)\alpha_{\max}\mu + s_i\mu} \geq \sum_{i=1}^N \frac{s_i\mu}{1-(s_i-1)\alpha_i\mu + s_i\mu} \geq 1,$$

where $\alpha_{\max} = \max_{i=1,\ldots,N}\alpha_i$.

Step 2. In the following we evaluate the spark(A), i.e, when the $s = s_1 + \cdots + s_N$ reaches the minimum.

Denote

$$g_i = \frac{s_i\mu}{1-(s_i-1)\alpha_{\max}\mu + s_i\mu},$$

we consider the following minimization problem

$$\min(s_1+\cdots+s_N), \quad \text{s.t.} \ \sum_{i=1}^{N} g_i - 1 \geq 0.$$

Using the Lagrange function and KKT conditions we obtain that $s=\sum_{i=1}^{N} s_i$ reaches minimum when $s_1=\cdots=s_N$. Then

$$\sum_{i=1}^{N} g_i = \frac{Ns\mu}{N(1+\alpha_{\max}\mu)+(1-\alpha_{\max})s\mu} \geq 1,$$

which results in

$$s \geq \frac{N(1+\alpha_{\max}\mu)}{(N-1+\alpha_{\max})\mu}.$$

By the definition of spark, we have

$$\text{spark}(A) \geq \frac{N(1+\alpha_{\max}\mu)}{(N-1+\alpha_{\max})\mu}.$$

Thus by Eq. (1.4), if

$$\|\mathbf{x}\|_0 < \frac{N(1+\alpha_{\max}\mu)}{2(N-1+\alpha_{\max})\mu},$$

then $\mathbf{x}$ is the unique solution of the piecewise P_0 problem (3.1). □

Remark 3.2 (1) In particular, if $\alpha_{\max}=0$, i.e., A is a union of N orthogonal bases. The result in Theorem 3.1 becomes $\|\mathbf{x}\|_0 < \frac{N}{2(N-1)\mu}$ which corresponds to the upper bound of Eq. (1.6).

(2) When $\alpha_1=\cdots=\alpha_N=1$, i.e., A_i has the same coherence as A. The result in Theorem 3.1 becomes $\|\mathbf{x}\|_0 < \frac{1+\mu}{2\mu}$ which corresponds to the upper bound of Eq. (1.5).

Example 3.3 Consider the case when $N=2$, i.e., $A=[A_1,A_2]$. In this example we set $\mu=0.1$ and $\alpha_{\max}=0.5$, then the sufficient conditions which ensure the uniqueness for P_0 problem and piecewise P_0 problem are listed as follows.

(1) Condition 1 Eq. (1.5): $\|\mathbf{x}\|_0 < \frac{1+\mu}{2\mu}$ (general condition).

(2) Condition 2 Eq. (1.6): $\|\mathbf{x}\|_0 < \frac{1}{\mu}$ (A is union of orthogonal bases).

(3) Condition 5 Eq. (3.2): $\|\mathbf{x}\|_0 < \frac{1+\alpha_{\max}\mu}{(1+\alpha_{\max})\mu}$ (A is union of general bases).

From the observation of Figure 2, in the general case (Condition 1) one can only ensure to recover 4-sparse vector. When it comes to the piecewise sparse recovery, one can recover at least $(5,2)$-piecewise sparse vector with global 7-sparsity by Condition 5. It means that the upper bound in Theorem 3.1 (Condition 5) is more relaxed than the upper bound in Eq. (1.5) (Condition 1). The improved condition also makes a relation between general case and the union of orthogonal bases. Thus the results in Theorem 3.1 enlarge the scope the theoretical guarantees for sparse recovery by considering piecewise sparsity.

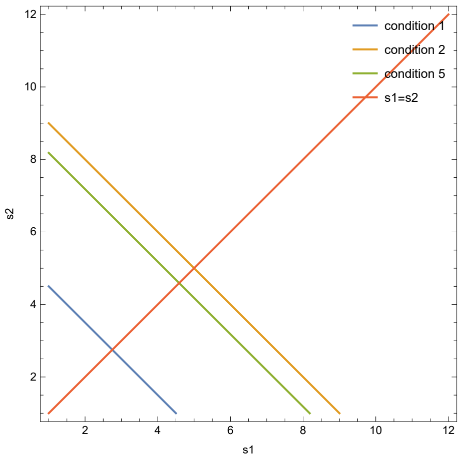

Figure 2 Comparison of the upper bounds for uniqueness in Example 3.3

3.2. Feasible conditions of algorithms for piecewise sparse recovery

Furthermore, combining with the piecewise sparsity of the signal $\mathbf{x} = (\mathbf{x}_1^{\rm T}, \mathbf{x}_2^{\rm T}, \ldots, \mathbf{x}_N^{\rm T})^{\rm T}$, we can also improve the feasible conditions of OMP and BP algorithms for piecewise sparse recovery as follows.

Theorem 3.4 *Suppose that the measurement matrix $A = [A_1, A_2, \ldots, A_N]$ is a union of N bases with an overall coherence μ and sub-block coherence parameters α_i for $i = 1, 2, \ldots, N$, and $\mathbf{x} = (\mathbf{x}_1^{\rm T}, \mathbf{x}_2^{T}, \ldots, \mathbf{x}_N^{\rm T})^{\rm T}$ is $(s_1, s_2, \ldots, s_N)$-piecewise sparse. If*

$$2\sum_{i=1}^{N} \frac{\mu s_i}{1 + \alpha_i\mu + (1-\alpha_i)\mu s_i} \le \frac{1 + \alpha_Z\mu + 2(1-\alpha_Z)\mu s_Z}{1 + \alpha_Z\mu + (1-\alpha_Z)\mu s_Z}, \tag{3.4}$$

where $Z = \{Z : \frac{1+\alpha_Z\mu}{(1-\alpha_Z)s_Z} = \max_{i=1,\ldots,N} \frac{1+\alpha_i\mu}{(1-\alpha_i)s_i}\}$, the exact recovery condition (ERC) holds. In which case both Orthogonal Matching Pursuit and Basis Pursuit recover the sparse representation.

Proof Following the proof in [18, Theorem 3.7] and the notations in the proof of Theorem 3.1, the Grassmannian matrix

$$\begin{aligned}
\Phi_{\mathbf{S}} = A_{\mathbf{S}}^{\rm T} A_{\mathbf{S}} &= \begin{pmatrix} A_{S_1}^{\rm T} \\ \vdots \\ A_{S_N}^{\rm T} \end{pmatrix} \begin{pmatrix} A_{S_1} & \cdots & A_{S_N} \end{pmatrix} \\
&= \begin{pmatrix} A_{S_1}^{\rm T} A_{S_1} & A_{S_1}^{\rm T} A_{S_2} & \cdots & A_{S_1}^{\rm T} A_{S_N} \\ A_{S_2}^{\rm T} A_{S_1} & A_{S_2}^{\rm T} A_{S_2} & \cdots & A_{S_2}^{\rm T} A_{S_N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{S_N}^{\rm T} A_{S_1} & A_{S_N}^{\rm T} A_{S_2} & \cdots & A_{S_N}^{\rm T} A_{S_N} \end{pmatrix} = I_s - G,
\end{aligned}$$

where

$$G = \begin{pmatrix} I_{s_1} - A_{S_1}^{\mathrm{T}} A_{S_1} & -A_{S_1}^{\mathrm{T}} A_{S_2} & \cdots & -A_{S_1}^{\mathrm{T}} A_{S_N} \\ -A_{S_2}^{\mathrm{T}} A_{S_1} & I_{s_2} - A_{S_2}^{\mathrm{T}} A_{S_2} & \cdots & -A_{S_2}^{\mathrm{T}} A_{S_N} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{S_N}^{\mathrm{T}} A_{S_1} & -A_{S_N}^{\mathrm{T}} A_{S_2} & \cdots & I_{s_N} - A_{S_N}^{\mathrm{T}} A_{S_N} \end{pmatrix}$$

with the diag-block matrix $I_{s_i} - A_{S_i}^{\mathrm{T}} A_{S_i}$ of the form

$$\begin{pmatrix} 0 & -A_{i_1}^{\mathrm{T}} A_{i_2} & \cdots & -A_{i_1}^{\mathrm{T}} A_{i_{s_i}} \\ -A_{i_2}^{\mathrm{T}} A_{i_1} & 0 & \cdots & -A_{i_2}^{\mathrm{T}} A_{i_{s_i}} \\ \vdots & \vdots & \ddots & \vdots \\ -A_{i_{s_i}}^{\mathrm{T}} A_{i_1} & -A_{i_{s_i}}^{\mathrm{T}} A_{i_2} & \cdots & 0 \end{pmatrix}.$$

Denote by $|G|$ the entrywise absolute value of the matrix G. Since all the entries in the off-diag blocks of $|G|$ can be bounded by μ, and the diag-block matrix

$$|I_{s_i} - A_{S_i}^{\mathrm{T}} A_{S_i}| \leq \begin{pmatrix} 0 & \mu^{i,i} & \cdots & \mu^{i,i} \\ \mu^{i,i} & 0 & \cdots & \mu^{i,i} \\ \vdots & \vdots & \ddots & \vdots \\ \mu^{i,i} & \mu^{i,i} & \cdots & 0 \end{pmatrix},$$

$i = 1, 2, \ldots, N$, we have $|G| \leq \mu 1_s - \mu B$, where 1_s is the $s \times s$ matrix with unit entries, B is the block matrix

$$B = \begin{pmatrix} B_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_N \end{pmatrix},$$

where $B_i = \alpha_i I_{s_i} + (1 - \alpha_i) 1_{s_i}$ is the matrix with 1 on the diagonal, and all the off-diag entries are $1 - \frac{\mu^{i,i}}{\mu} = 1 - \alpha_i$, $i = 1, 2, \ldots, N$. Hence, we have the entrywise inequality

$$\begin{aligned} |\Phi_{\mathbf{S}}^{-1}| = |(I_s - G)^{-1}| &= \Big| I_s + \sum_{k=1}^{\infty} G^k \Big| \\ &\leq I_s + \sum_{k=1}^{\infty} |G|^k \leq I_s + \sum_{k=1}^{\infty} (\mu 1_s - \mu B)^k \\ &= ((I_s + \mu B) - \mu 1_s)^{-1} \\ &= (I_s - \mu (I_s + \mu B)^{-1} 1_s)^{-1} (I_s + \mu B)^{-1}. \end{aligned}$$

Step 1. Compute

$$(I_s + \mu B)^{-1} = \begin{bmatrix} (I_s + \mu B)_1^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & (I_s + \mu B)_N^{-1} \end{bmatrix},$$

where

$$(I_s + \mu B)_i^{-1} = \frac{1}{1 + \alpha_i \mu} (I_{s_i} - \frac{(1 - \alpha_i)\mu}{1 + \alpha_i \mu + (1 - \alpha_i)\mu s_i} 1_{s_i}).$$

Step 2. Compute

$$(I_s-\mu(I_s+\mu B)^{-1}1_s)^{-1}=I_s+\sum_{k=1}^{\infty}(\mu(I_s+\mu B)^{-1}1_s)^k \tag{3.5}$$

with

$$\mu(I_s+\mu B)^{-1}1_s=\begin{bmatrix}\frac{\mu}{1+\alpha_1\mu+(1-\alpha_1)\mu s_1}\mathbf{1}_{s_1}\\ \vdots\\ \frac{\mu}{1+\alpha_N\mu+(1-\alpha_N)\mu s_N}\mathbf{1}_{s_N}\end{bmatrix}\begin{bmatrix}\mathbf{1}_{s_1}^{\mathrm{T}} & \cdots & \mathbf{1}_{s_N}^{\mathrm{T}}\end{bmatrix}\overset{\text{def}}{=}\mathbf{v}\mathbf{1}_s^{\mathrm{T}}.$$

We use $\mathbf{1}$ to indicate the column vector with unit entries. Moreover, by the inner product

$$\mathbf{1}_s^{\mathrm{T}}\mathbf{v}=\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i},$$

we have the series

$$\begin{aligned}\sum_{k=1}^{\infty}(\mathbf{v}\mathbf{1}_s^{\mathrm{T}})^k&=(\mathbf{v}\mathbf{1}_s^{\mathrm{T}})\sum_{k=1}^{\infty}(\mathbf{1}_s^{\mathrm{T}}\mathbf{v})^{k-1}=(\mathbf{v}\mathbf{1}_s^{\mathrm{T}})\sum_{k=0}^{\infty}(\mathbf{1}_s^{\mathrm{T}}\mathbf{v})^{k}\\&=\frac{1}{1-\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}}(\mathbf{v}\mathbf{1}_s^{\mathrm{T}}).\end{aligned}\tag{3.6}$$

Combined with Eq. (3.5), we have

$$|\Phi_{\mathbf{S}}^{-1}|\le\Big(I_s+\frac{1}{1-\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}}(\mathbf{v}\mathbf{1}_s^{\mathrm{T}})\Big)(I_s+\mu B)^{-1}.\tag{3.7}$$

Step 3. Assume vector A_i is drawn from basis index Z, then

$$|A_{\mathbf{S}}^{\mathrm{T}}A_j|\le\begin{bmatrix}|A_{S_1}^{\mathrm{T}}A_j| & \cdots & |A_{S_N}^{\mathrm{T}}A_j|\end{bmatrix}^{\mathrm{T}}\le\begin{bmatrix}\mu\mathbf{1}_{s_1}^{\mathrm{T}} & \cdots & \alpha_Z\mu\mathbf{1}_{s_Z}^{\mathrm{T}} & \cdots & \mu\mathbf{1}_{s_N}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$$

and

$$(I_s+\mu B)^{-1}|A_{\mathbf{S}}^{\mathrm{T}}A_j|\le\begin{bmatrix}\frac{\mu}{1+\alpha_1\mu+(1-\alpha_1)\mu s_1}\mathbf{1}_{s_1}^{\mathrm{T}}\\ \vdots\\ \frac{\alpha_Z\mu}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}\mathbf{1}_{s_Z}^{\mathrm{T}}\\ \vdots\\ \frac{\mu}{1+\alpha_N\mu+(1-\alpha_N)\mu s_N}\mathbf{1}_{s_N}^{\mathrm{T}}\end{bmatrix}.\tag{3.8}$$

Step 4. Moreover, we calculate the inner product of ERC condition $|(A_{\mathbf{S}}^{\mathrm{T}}A_{\mathbf{S}})^{-1}A_{\mathbf{S}}^{\mathrm{T}}A_j|$ in combination with Eqs. (3.7) and (3.8):

$$|(A_{\mathbf{S}}^{\mathrm{T}}A_{\mathbf{S}})^{-1}A_{\mathbf{S}}^{\mathrm{T}}A_j|=|\Phi_{\mathbf{S}}^{-1}||A_{\mathbf{S}}^{\mathrm{T}}A_j|\le\begin{bmatrix}\frac{\mu}{1+\alpha_1\mu+(1-\alpha_1)\mu s_1}\mathbf{1}_{s_1}\\ \vdots\\ \frac{\alpha_Z\mu}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}\mathbf{1}_{s_Z}\\ \vdots\\ \frac{\mu}{1+\alpha_N\mu+(1-\alpha_N)\mu s_N}\mathbf{1}_{s_N}\end{bmatrix}+$$

$$\frac{\sum_{i\neq Z}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}+\frac{\alpha_Z\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}}{1-\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}}\begin{bmatrix}\frac{\mu}{1+\alpha_1\mu+(1-\alpha_1)\mu s_1}\mathbf{1}_{s_1}\\ \vdots\\ \frac{\mu}{1+\alpha_N\mu+(1-\alpha_N)\mu s_N}\mathbf{1}_{s_N}\end{bmatrix}, \tag{3.9}$$

then apply the l^1 norm to inequality (3.9) to reach

$$\|(A_{\mathbf{S}}^{\mathrm{T}}A_{\mathbf{S}})^{-1}A_{\mathbf{S}}^{\mathrm{T}}A_j\|_1\leq\frac{\sum_{i\neq Z}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}+\frac{\alpha_Z\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}}{1-\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}}. \tag{3.10}$$

Step 5. Since

$$\|A_T^{\mathrm{T}}A_{\mathbf{S}}^{\dagger}\|_\infty=\max_{j\in T}\|(A_{\mathbf{S}}^{\mathrm{T}}A_{\mathbf{S}})^{-1}A_{\mathbf{S}}^{\mathrm{T}}A_j\|_1,$$

we consider the maximum of the right side of Eq. (3.10) and rewrite it as

$$\|(A_{\mathbf{S}}^{\mathrm{T}}A_{\mathbf{S}})^{-1}A_{\mathbf{S}}^{\mathrm{T}}A_j\|_1\leq\frac{\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}-\frac{(1-\alpha_Z)\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}}{1-\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}}. \tag{3.11}$$

The right side of Eq. (3.11) reaches the maximum when

$$f_Z\overset{\text{def}}{=}\frac{(1-\alpha_Z)\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}$$

reaches the minimum,

$$f_Z=\frac{(1-\alpha_Z)s_Z\mu}{(1-\alpha_Z)s_Z\mu+1+\alpha_Z\mu}=\frac{\mu}{\mu+\frac{1+\alpha_Z\mu}{(1-\alpha_Z)s_Z}}.$$

Let

$$Z=\{Z:\frac{1+\alpha_Z\mu}{(1-\alpha_Z)s_Z}=\max_{i=1,\dots,N}\frac{1+\alpha_i\mu}{(1-\alpha_i)s_i}\}$$

and

$$2\sum_{i=1}^{N}\frac{\mu s_i}{1+\alpha_i\mu+(1-\alpha_i)\mu s_i}\leq\frac{1+\alpha_Z\mu+2(1-\alpha_Z)\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}.$$

Then the Exact Recovery Condition holds as $\|A_T^{\mathrm{T}}A_{\mathbf{S}}^{\dagger}\|_\infty<1$, thus we complete the proof. □

In particular, when A is a union of orthogonal bases, i.e., $\alpha_1=\cdots=\alpha_N=0$. Thus, Z is chosen for the minimum s_i, $i=1,\dots,N$, then the condition Eq. (3.4) corresponds to the condition Eq. (1.8) in [18].

Example 3.5 Consider the case where $N=2$, i.e., $A=[A_1,A_2]$ and $\mathbf{x}$ is (s_1,s_2)-piecewise sparse vector. In this example we set overall coherence $\mu=0.1$, $\alpha_1=0.01$, $\alpha_2=0.3$. The following sufficient conditions which ensure the feasibility of OMP and BP algorithms are listed.

(1) Condition 1 Eq. (1.5): $\|\mathbf{x}\|_0<\frac{1}{2}(1+\frac{1}{\mu})$ (general condition).

(2) Condition 3 Eq. (1.7): $\|\mathbf{x}\|_0<\frac{\sqrt{2}-0.5}{\mu}$ (equivalence condition when A is in pairs of orthogonal bases).

(3) Condition 4 Eq. (1.8): $2\mu^2s_1s_2+\mu s_2-1<0$ $(s_1\leq s_2)$, or $2\mu^2s_1s_2+\mu s_1-1<0$ $(s_1>s_2)$ (ERC condition when A is in pairs of orthogonal bases).

(4) Condition 6 Eq. (3.4):

$$2\sum_{i=1}^{2}\frac{\mu s_i}{1+\mu s_i-(s_i-1)\alpha_i\mu}\le\frac{1+\alpha_Z\mu+2(1-\alpha_Z)\mu s_Z}{1+\alpha_Z\mu+(1-\alpha_Z)\mu s_Z}$$

(ERC condition when A is in pairs of general bases).

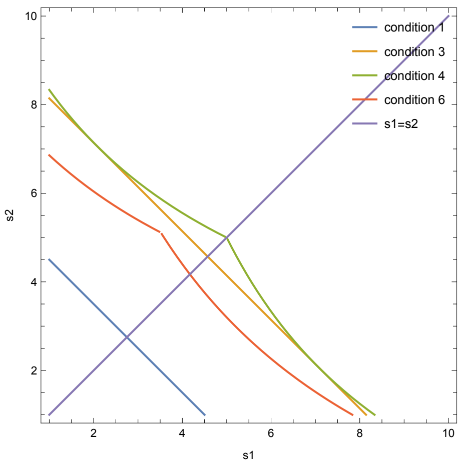

Figure 3 Comparison of upper bounds for feasible conditions of sparse recovery for Example 3.5

The above four conditions are shown in Figure 3, where the curve of Condition 6 lies between the curve of Condition 1 for general case and the curve of Condition 4 for pairs of orthogonal bases (correspongding to $\alpha_1=\alpha_2=0$). It shows that the new sparsity Condition 6 is improved by considering the piecewise sparsity (s_1,s_2) and the sub-block coherence parameters $\alpha_1=0.01$, $\alpha_2=0.3$.

Example 3.6 In this example we show the cases where $\mathbf{x}$ is (s_1,s_2)-piecewise sparse vector with different piecewise sparsities, corresponding to different parameter pairs (α_1,α_2). The upper bounds by the condition 6 Eq. (3.4) are plotted in Figure 4 for the following three cases.

Case 1. $\mu=0.1$, $(\alpha_1,\alpha_2)=(0.7,0.01)$, i.e., the sub-matrix coherence of A_1 differs greatly from that of A_2, α_1 is close to 1.

Case 2. $\mu=0.1$, $(\alpha_1,\alpha_2)=(0.2,0.15)$, i.e., the sub-matrix coherence of A_1 differs slightly from that of A_2, both α_1 and α_2 are small.

Case 3. $\mu=0.1$, $(\alpha_1,\alpha_2)=(0.05,0.02)$, i.e., the sub-matrix coherence of A_1 differs very slightly from that of A_2, both α_1 and α_2 are very small.

Remark 3.7 It is observed from Figure 4 that different (α_1,α_2), i.e., different piecewise sparsities may result in different global sparsity conditions. Especially, when the sub-matrix coherences of A_1 and A_2 are small, some relaxed conditions of the sparsity can be obtained. This phenomenon provides us a guidance on the setting of piecewise structure for a given matrix in order to obtain an optimal piecewise sparsity condition, which is another interesting problem in our future work.

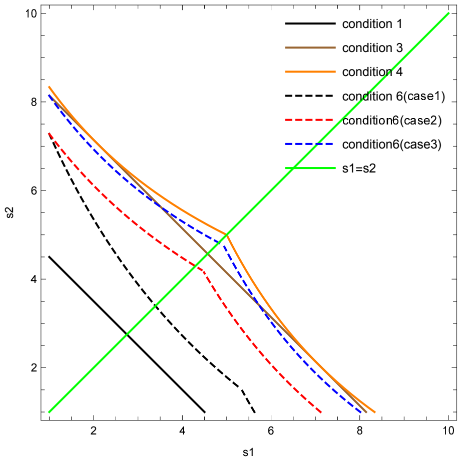

Figure 4 Comparison of upper bounds for feasible conditions of piecewise sparse recovery for Example 3.6

4. Conclusion

In this paper, we introduce the piecewise sparsity of signals and use the mutual coherence for matrix in union of general bases to study the conditions for piecewise sparse recovery. We generalize the results in orthogonal cases to the cases of general bases. We provide the new upper bounds of global sparsity and piecewise sparsity of the signal recovered by both l^0 and l^1 optimizations when the measurement matrix A is a union of general bases. The structured information of the matrix A is exploited to improve the sufficient conditions for successfully piecewise sparse recovery and the reliability of the greedy algorithms and the BP method.

References

[1] M. ELAD. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[2] S. G. MALLAT, Z. ZHANG. *Matching pursuits with time-frequency dictionaries*. IEEE Trans. Signal Process, 1993, **41**(12): 3397–3415.

[3] Y. C. PATI, R. REZAIIFAR, P. S. KRISHNAPRASAD. *Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition*. Proceedings of the 27th Annual Asilomar Conference on Signals, Systems and Computers, 1993, **1/2**: 40–44.

[4] J. A. TROPP, A. C. GILBERT. *Signal recovery from random measurements via orthogonal matching pursuit*. IEEE Trans. Inform. Theory, 2007, **53**(12): 4655–4666.

[5] T. BLUMENSATH, M. E. DAVIES. *Iterative hard thresholding for compressed sensing*. Appl. Comput. Harmon. Anal., 2009, **27**(3): 265–274.

[6] D. L. DONOHO, Y. TSAIG, I. DRORI, et al. *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit (StOMP)*. IEEE Trans. Inform. Theory, 2012, **58**(2): 1094–1121.

[7] D. NEEDELL, R. VERSHYNIN. *Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit*. Found. Comput. Math., 2009, **9**(3): 317–334.

[8] D. NEEDELL, R. VERSHYNIN. *Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit*. IEEE J. Select. Top. Signal Process, 2010, **4**(2): 310–316.

[9] D. NEEDELL, J. A. TROPP. *CoSaMP: iterative signal recovery from incomplete and inaccurate samples*. Appl. Comput. Harmon. Anal., 2009, **26**(3): 301–321.

[10] Wei DAI, O. MILENKOVIC. *Subspace pursuit for compressive sensing signal reconstruction*. IEEE Trans. Inform. Theory, 2009, **55**(5): 2230–2249.

[11] A. MALEKI. *Coherence Analysis of Iterative Thresholding Algorithms*. in: Proceedings of the 47th Annual Allerton Conference on Communication, Control and Computing, IEEE Press, 2009.

[12] S. FOUCART. *Hard thresholding pursuit: an algorithm for compressive sensing*. SIAM J. Numer. Anal., 2011, **49**(6): 2543–2563.

[13] S. J. KIM, K. KOH, M. LUSTIG, et al. *A interior–point method for large–scale l^1-regularized least squares*, IEEE J. Select. Top. Signal Process, 2007, **1**(4): 606–617.

[14] M. A. T. FIGUEIREDO, R. D. NOWAK, S. J. WRIGHT. *Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems*. IEEE J. Select. Top. Signal Process, 2007, **1**(4): 586–598.

[15] I. DAUBECHIES, M. DEFRISE, C. D. MOL. *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*. Comm. Pure Appl. Math., 2004, **57**(11): 1413–1457.

[16] D. L. DONOHO, Xiaoming HUO. *Uncertainty principles and ideal atomic decomposition*. IEEE Trans. Inform. Theory, 2001, **47**(7): 2845–2862.

[17] D. L. DONOHO, M. ELAD. *Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization*. Proc. Natl. Acad. Sci. USA, 2003, **100**(5): 2197–2202.

[18] J. A. TROPP. *Greed is good: algorithmic results for sparse approximation*. IEEE Trans. Inform. Theory, 2004, **50**(10): 2231–2242.

[19] E. J. CANDDES, T. TAO. *Decoding by linear programming*. IEEE Trans. Inform. Theory, 2005, **51**(12): 4203–4215.

[20] E. J. CANDDES, J. K. ROMBERG, T. TAO. *Stable signal recovery from incomplete and inaccurate measurements*. Comm. Pure Appl. Math., 2006, **59**(8): 1207–1223.

[21] E. J. CANDDES. *The restricted isometry property and its implications for compressed sensing*. C. R. Math. Acad. Sci. Paris, 2008, **346**(9): 589–592.

[22] M. ELAD, A. M. BRUCKSTEIN. *A generalized uncertainty principle and sparse representation in pairs of bases*. IEEE Trans. Inform. Theory, 2002, **48**(9): 2558–2567.

[23] R. GRIBONVAL, M. NIELSEN. *Sparse representations in unions of bases*. IEEE Trans. Inform. Theory, 2003, **49**(12): 3320–3325.

[24] L. PEOTTA, P. VANDERGHEYNST. *Matching pursuit with block incoherent dictionaries*. IEEE Trans. Signal Process., 2007, **55**(9): 4549–4557.

[25] Y. C. ELDAR, M. MISHALI. *Block sparsity and sampling over a union of subspaces*. in: International Conference on Digital Signal Processing, IEEE, 2009, 1–8.

[26] Y. C. ELDAR, P. KUPPINGER, H. BOLCSKEI. *Block-sparse signals: uncertainty relations and efficient recovery*. IEEE Trans. Signal Process., 2010, **58**(6): 3042–3054.

[27] E. ELHAMIFAR, R. VIDAL, *Block-sparse recovery via convex optimization*. IEEE Trans. Signal Process., 2011, **60**(8): 4094–4107.

[28] J. L. STARCK, M. ELAD, D. L. DONOHO. *Image decomposition via the combination of sparse representations and a variational approach*. IEEE Trans. Image Process., 2010, **14**(10): 1570–1582.

[29] Yongxia HAO, Chongjun LI, Renhong WANG. *Sparse approximate solution of fitting surface to scattered points by MLASSO model*. Sci. China Math., 2018, **61**(7): 1319–1336.

[30] Yijun ZHONG, Chongjun LI. *Piecewise sparse recovery via piecewise inverse scale space algorithm with deletion rule*. J. Comput. Math., 2020, **38**(2): 375–394.