Uniqueness Conditions for A Class of ℓ_0-Minimization Problems

Chunlei Xu* and Yun-Bin Zhao†

October, 2013, Revised January 2014

Abstract. We consider a class of ℓ_0-minimization problems, which is to search for the partial sparsest solution to an underdetermined linear system with additional constraints. We introduce several concepts, including l_p-induced quasi-norm $(0 < p < 1)$, maximal scaled spark and scaled mutual coherence, to develop several new uniqueness conditions for the partial sparsest solution to this class of ℓ_0-minimization problems. A further improvement of some of these uniqueness criteria has been also achieved through the so-called concepts such as maximal scaled (sub)coherence rank.

Key words. ℓ_0-minimization, uniqueness condition, l_p-induced quasi-norm, scaled spark, scaled mutual (sub)coherence, maximal scaled coherence rank.

1 Introduction

Sparse representation, using only a few elementary atoms from a dictionary to represent data (signals, images, etc.), has been widely used in engineering and applied sciences recently (see, e.g., [15, 17, 14, 11, 4, 9, 8, 10, 22, 20, 21] and the references therein). For a vector x, let $\|x\|_0$ denote the 'ℓ_0-norm' of x, namely, the number of nonzero components of x. In this paper, we consider the following model for the sparse representation of the vector $b \in R^m$:

$$\min \left\{ \|x\|_0 : \ M \begin{pmatrix} x \\ y \end{pmatrix} = b, \ y \in C \right\}, \tag{1}$$

where $M = [A_1, A_2] \in R^{m\times(n_1+n_2)}$, $m \leq n_1$, is a concatenation of $A_1 \in R^{m\times n_1}$ and $A_2 \in R^{m\times n_2}$, and C is a convex set in R^{n_2} which can be interpreted as certain constraints on the variable $y \in R^{n_2}$. Throughout the paper, we assume that the null space of A_2^T is nonzero, namely, $\mathcal{N}(A_2^T) \neq \{0\}$. The solution to the system

$$M \begin{pmatrix} x \\ y \end{pmatrix} = b, \ y \in C, \tag{2}$$

*School of Mathematics, University of Birmingham, Edgbaston, B15 2TT, United Kingdom (email: cxx028@bham.ac.uk).

†School of Mathematics, University of Birmingham, Edgbaston, B15 2TT, United Kingdom (email: y.zhao.2@bham.ac.uk). This author was supported by the Engineering and Physical Sciences Research Council (EPSRC) under the grant #EP/K00946X/1, and partially supported by the National Natural Science Foundation of China (NSFC) under the grant #11301016.

includes two parts: $x \in R^{n_1}$ and $y \in R^{n_2}$. The ℓ_0-minimization problem (1) is to seek a solution $z = (x, y)$ to the system (2) such that the x-part is the sparsest one, but there is no requirement on the sparsity of the y-part of the solution. Such a sparsest solution x can be called the *sparsest x-part solution* to the system (2). The ℓ_0-minimization problem (1) is NP-hard (see Natarajan [16]), and can be called a partial ℓ_0-minimization problem, or partial sparsity-seeking problem.

Problem (1) is closely related to partial sparsity recovery theory (see, e.g., Bandeira et al.[1], and Jacques [13]), and partial imaging reconstruction (Vaswani and Lu [18]), and the sparse Hessian recovery (Bandeira et al. [2]). Problem (1) is general enough to include some importance sparsity seeking problems as special cases. For instance, the normal ℓ_0-minimization

$$\min\{\|x\|_0 : Ax = b\} \tag{3}$$

is an important special case of (1). In fact, Problem (1) is reduced to (3) when $A_2 = 0$.

The uniqueness of the standard ℓ_0-minimization (3) has been widely investigated, and has been established by using the so-called spark of a matrix A (see Donoho and Elad [6]), denoted by Spark(A), which is the smallest number of columns of a matrix that are linearly dependent. It was shown in [6, 5, 4] that for a given linear system $Ax = b$, if there exists a solution x satisfying $\|x\|_0 < \frac{1}{2}\text{Spark}(A)$, then x is necessarily the unique sparsest solution to (3).

Since the computation of spark is generally intractable, some other verifiable conditions have been developed in the literature, for instance, by the mutual coherence [7], the largest absolute value of inner products between different normalized columns of A, i.e.,

$$\mu(A) = \max_{1 \le i \ne j \le n} \frac{|\langle a_i, a_j \rangle|}{\|a_i\|_2 \cdot \|a_j\|_2},$$

where a_i is the i-th column of A, $i = 1, ..., n$. The mutual coherence gives a computable lower bound for the spark [6], i.e., Spark$(A) \ge 1 + \frac{1}{\mu(A)}$, which yields following uniqueness condition (see, e.g., [6, 5, 4]): For a given linear system $Ax = b$, if there exists a solution x satisfying $\|x\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(A)})$, then x is necessarily the unique sparsest solution to (3). However, the mutual coherence condition might be very restrictive in some situations, and fails to provide a good lower bound for the spark. In order to improve the lower bound for spark, Zhao [19] has introduced the concept of coherence rank, submutual coherence and scaled mutual coherence, and has developed several new and improved uniqueness sufficient conditions for the solution to ℓ_0-minimization (3).

So far, the uniqueness of the sparsest x-part solution to the general sparsity module (1) has not well developed. The main purpose of this paper is to study such uniqueness and to establish some criteria under which the problem (1) has a unique sparsest x-part solution. These results will be established through some new concepts such as the l_p-induced quasi-norm, the (maximal) scaled spark, coherence, and coherence rank associated with a pair of matrices $(A_1 \in R^{m \times n_1}, A_2 \in R^{m \times n_2})$. These concepts can be seen as a generalization of those in [19].

This paper is organized as follows. In Section 2, we develop sufficient conditions for the uniqueness of x-part solutions to the ℓ_0-minimization problem (1) in terms of l_p- induced quasi-norm, and such concepts as maximal scaled spark, and minimal or maximal scaled mutual coherence. A further improvement of these conditions is provided in Section 3.

2 Uniqueness criteria for the ℓ_0-minimization problem (1)

The uniqueness of the sparsest x-part solution to the system (2) can be developed through different concepts and properties of matrices. One of such important concept is *spark* together with its variants, which provides a connection between the null space of a matrix and the sparsest solution to linear equations. In this section, we show that the method used for developing uniqueness claims for the ℓ_0-minimization (3) can be used for the development of similar claims to the system (2), while the extra variable y in the system $M\begin{pmatrix} x \\ y \end{pmatrix} = b$ increases the complexity of the problem (1). Our first sufficient uniqueness condition for the sparsest x-part solution to (1) can be developed by using the so-called l_p-induced quasi-norm, as shown in the following subsection.

2.1 An l_p-induced quasi-norm-based uniqueness condition

For any $0 < p < \infty$ and a vector $x \in R^n$, let $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. When $p \in (0,1)$, $\|x\|_p$ is called the l_p quasi-norm of x. We now introduce the following concept.

Definition 2.1 *For any given matrix $A \in R^{m\times n}$, when $0 < p < 1$, the l_p-induced quasi-norm of A, denoted by $\psi_p(A)$, is defined by*

$$\psi_p(A) = \sup_{0\neq z\in R^n} \frac{\|Az\|_p^p}{\|z\|_p^p} = \sup_{\|z\|_p^p \leq 1} \|Az\|_p^p. \quad (4)$$

Clearly, for a fixed $p \in (0,1)$, $\psi_p(A)$ satisfies the following properties: $\psi_p(A) \geq 0$, $\psi_p(A) > 0$ for any $A \neq 0$, and $\psi_p(A+B) \leq \psi_p(A) + \psi_p(B)$ for any matrices A, B with same dimensions. It is worth mentioning that the triangle inequality above follows from the property: $\|x+y\|_p^p \leq \|x\|_p^p + \|y\|_p^p$ (see, e.g., [12]). We see that for $\alpha > 0$, $\psi_p(\alpha A) \neq \alpha\psi_p(A)$ in general. So $\psi_p(A)$ is a quasi-norm of A. Note that, for every entry z_i, as p tends to zero, $|z_i|^p$ approaches to 1 for $z_i \neq 0$ and 0 for $z_i = 0$. Thus for any given $z \in R^n$, we have

$$\lim_{p\to 0} \|z\|_p^p = \lim_{p\to 0} \sum_{i=1}^n |z_i|^p = \|z\|_0, \quad (5)$$

which indicates that the 'ℓ_0-norm' $\|z\|_0$ can be approximated by $\|x\|_p^p$ with sufficiently small $p \in (0,1)$. Note that for a given matrix A, $\psi_p(A)$ is continuous with respect to $p \in (0,1)$. Thus there might exists a positive number η such that $\eta = \lim_{p\to 0^+} \psi_p(A)$. We assume that the following property holds for the matrix $M = [A_1, A_2]$ when p tends to 0.

Assumption 2.2 *Assume that matrices A_1, A_2 satisfy the following properties: (i) $A_2^T A_2$ is a nonsingular matrix, and (ii) there exists a positive constant, denoted by $\psi_0(A_2^\dagger A_1)$, such that*

$$\psi_0(A_2^\dagger A_1) = \lim_{p\to 0^+} \psi_p(A_2^\dagger A_1),$$

where $A_2^\dagger = (A_2^T A_2)^{-1} A_2^T$, the pseudo-inverse of A_2.

Under Assumption 2.2 and by (4) and (5), we immediately have the following inequality:

$$\|(A_2^\dagger A_1)z\|_0 = \lim_{p\to 0_+} \|(A_2^\dagger A_1)z\|_p^p \le \lim_{p\to 0_+} \psi_p(A_2^\dagger A_1)\|z\|_p^p = \psi_0(A_2^\dagger A_1)\|z\|_0 \tag{6}$$

for any $z \in R^n$. We now state a uniqueness condition for Problem (1) under Assumption 2.2.

Theorem 2.3 *Consider the system (2) with $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$, and $m < n_1$. Let Assumption 2.2 be satisfied. Then if there exists a solution (x, y) to the system (2) satisfying that*

$$\|x\|_0 < \frac{1}{2}\frac{Spark(M)}{(1+\psi_0(A_2^\dagger A_1))}, \tag{7}$$

x must be the unique sparsest x-part solution to the system (2).

Proof. Assume the contrary that there is another solution $(x^{(1)}, y^{(1)})$ to the system (2) such that $x^{(1)}$ is the sparsest x-part and $x^{(1)} \neq x$ and $\|x^{(1)}\|_0 \le \|x\|_0 < \frac{1}{2}\frac{\text{Spark}(M)}{(1+\psi_0(A_2^\dagger A_1))}$. Since both (x, y) and $(x^{(1)}, y^{(1)})$ are solutions to the linear system $M\begin{pmatrix} x \\ y \end{pmatrix} = b$, we have

$$A_1(x - x^{(1)}) + A_2(y - y^{(1)}) = 0. \tag{8}$$

Since $A_2^T A_2$ is nonsingular, $y - y^{(1)}$ can be uniquely determined by $x - x^{(1)}$, i.e.,

$$y^{(1)} - y = A_2^\dagger A_1(x - x^{(1)}), \tag{9}$$

where $A_2^\dagger$ is the pseudo-inverse of A_2 given by $A_2^\dagger = (A_2^T A_2)^{-1}A_2^T$. From (8), we know that $\begin{pmatrix} x - x^{(1)} \\ y - y^{(1)} \end{pmatrix}$ is in the null space of the matrix $M = [A_1, A_2]$. This implies that the Spark(M) is a lower bound for $\left\|\begin{pmatrix} x - x^{(1)} \\ y - y^{(1)} \end{pmatrix}\right\|_0$, i.e.,

$$\|x - x^{(1)}\|_0 + \|y - y^{(1)}\|_0 = \left\|\begin{pmatrix} x - x^{(1)} \\ y - y^{(1)} \end{pmatrix}\right\|_0 \ge \text{Spark}(M). \tag{10}$$

Substituting (9) into (10) leads to

$$\|x - x^{(1)}\|_0 + \|A_2^\dagger A_1(x - x^{(1)})\|_0 \ge \text{Spark}(M). \tag{11}$$

Under Assumption 2.2, one has

$$\|A_2^\dagger A_1(x^{(1)} - x)\|_0 \le \psi_0(A_2^\dagger A_1) \cdot \|x - x^{(1)}\|_0.$$

Merging (11) and the inequality above leads to

$$(1 + \psi_0(A_2^\dagger A_1))\|x - x^{(1)}\|_0 \ge \text{Spark}(M),$$

Therefore,

$$2\|x\|_0 \ge \|x^{(1)}\|_0 + \|x\|_0 \ge \|x - x^{(1)}\|_0 \ge \frac{\text{Spark}(M)}{1 + \psi_0(A_2^\dagger A_1)}.$$

Thus $\|x\|_0 \geq \frac{1}{2}\frac{\text{Spark}(M)}{(1+\psi_0(A_2^{\dagger}A_1))}$, contradicting with (7). Therefore x must be the unique sparsest x-part solution to system (2).

The above result provides a new uniqueness criteria for the problem (1) by using l_p-induced quasi-norm. However, the above analysis relies on the nonsingularity of $A_2^T A_2$ which might not be satisfied in more general situations. Thus we need to develop some other uniqueness criteria for the problem from other perspectives.

2.2 Uniqueness based on scaled spark and scaled mutual coherence

In this section, we develop uniqueness conditions for problem (1) by using the so-called scaled spark and scaled mutual coherence.

Lemma 2.4 ([4]) *For any matrix M and any scaling matrix W, one has*

$$Spark(WM) \geq 1 + \frac{1}{\mu(WM)}.$$

We use $\mathcal{N}(\cdot)$ to denote the null space of a matrix. Our first uniqueness criterion based on scaled spark is given as follows.

Theorem 2.5 *Consider the system (2) where $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$ and $m < n_1$. If there exists a solution (x, y) to the system (2) satisfying*

$$\|x\|_0 < \frac{1}{2} Spark(B^T A_1), \tag{12}$$

where B is a basis of $\mathcal{N}(A_2^T)$, then x is the unique sparsest x-part solution to the system (2).

Proof. Assume that $(x^{(1)}, y^{(1)}) \neq (x, y)$ is a solution to the system (2) satisfying that $x^{(1)} \neq x$, and $\|x^{(1)}\|_0 \leq \|x\|_0 < \frac{1}{2}\text{Spark}(B^T A_1)$, where B is a basis of $\mathcal{N}(A_2^T)$. Note that $\begin{pmatrix} x^{(1)} - x \\ y^{(1)} - y \end{pmatrix}$ is in the null space of $M = [A_1, A_2]$, so

$$A_1(x^{(1)} - x_2) = -A_2(y^{(1)} - y). \tag{13}$$

Note that the range space of A_2 is orthogonal to the null space of A_2^T, namely,

$$\mathcal{R}(A_2) = \mathcal{N}(A_2^T)^{\perp}.$$

Let B be an arbitrary basis of $\mathcal{N}(A_2^T)$. Since the right-hand side of (13) is in $\mathcal{R}(A_2)$, by multiplying both sides of the equation (13) by B^T, we get

$$B^T A_1(x^{(1)} - x) = 0,$$

which implies that

$$\|x^{(1)} - x\|_0 \geq \text{Spark}(B^T A_1). \tag{14}$$

Therefore,

$$2\|x|_0 \geq \|x^{(1)}\|_0 + \|x\|_0 \geq \text{Spark}(B^T A_1).$$

i.e., $\|x\|_0 \geq \frac{1}{2}\text{Spark}(B^T A_1)$, leading to a contradiction. Therefore, the system (2) has a unique sparsest x-part solution.

Let F be a set of all bases of $\mathcal{N}(A_2^T)$, namely,

$$F = \{B \in R^{m\times q}: \; B \text{ is a basis of } \; \mathcal{N}(A_2^T)\},$$

where q is the dimension of $\mathcal{N}(A_2^T)$.

From the definition of the spark, we know that $\text{Spark}(B^T A_1)$ is bounded. Hence, there exists the supremum of $\text{Spark}(B^T A_1)$, defined as follows.

Definition 2.6 *For any matrix $A_1 \in R^{m\times n_1}$ with $m < n_1$, let*

$$\mathit{Spark}^*_{A_2}(A_1) = \sup_{B\in F} \mathit{Spark}(B^T A_1). \tag{15}$$

$\mathit{Spark}^*_{A_2}(A_1)$ is called the maximal scaled spark of A_1 over F (the set of bases of $\mathcal{N}(A_2^T)$).

The inequality (14) in the proof of Theorem 2.5 holds for all bases B of $\mathcal{N}(A_2^T)$. Therefore, the spark condition (12) can be further enhanced by using $\text{Spark}^*_{A_2}(A_1)$.

Theorem 2.7 *Consider the system (2) where $A_1 \in R^{m\times n_1}$ and $A_2 \in R^{m\times n_2}$ and $m < n_1$. If there exists a solution (x, y) to the system (2) satisfying*

$$\|x\|_0 < \frac{1}{2}\mathit{Spark}^*_{A_2}(A_1), \tag{16}$$

where $\mathit{Spark}^*_{A_2}(A_1)$ is given by (15), then x is the unique sparsest x-part solution to the system (2).

From Lemma 2.4, the scaled mutual coherence may provide a lower bound for the scaled spark. An immediate consequence of Theorem 2.5 is the corollary below.

Corollary 2.8 *For a given system (2) where $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$ and $m < n_1$. If there exists a solution (x, y) to the system (2) satisfying*

$$\|x\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(B^T A_1)}\right), \tag{17}$$

where B is a basis of $\mathcal{N}(A_2^T)$, then x is the unique sparsest x-part solution to the system (2).

Note that Corollary 2.8 holds for any bases B of $\mathcal{N}(A_2^T)$. So it makes sense to further enhance the bound (17) by introducing the following definition.

Definition 2.9 *For any matrix $A_1 \in R^{m\times n_1}$ $(m < n_1)$ and $A_2 \in R^{m\times n_2}$, let*

$$\mu^*_{A_2}(A_1) = \inf_{B\in F} \mu(B^T A_1), \quad \mu^{**}_{A_2}(A_1) = \sup_{B\in F} \mu(B^T A_1). \tag{18}$$

$\mu^*_{A_2}(A_1)$ is called the minimal scaled coherence of A_1 over F, and $\mu^{**}_{A_2}(A_1)$ is called the maximal scaled coherence of A_1 over F.

Based on Lemma 2.4 and the above definition, we have the following result.

Lemma 2.10 *For any basis B of $\mathcal{N}(A_2^T)$, we have*

$$1+\frac{1}{\mu(B^TA_1)} \leq 1+\frac{1}{\mu^*_{A_2}(A_1)} \leq Spark^*_{A_2}(A_1), \tag{19}$$

Proof. The first inequality holds by the definition of $\mu^*_{A_2}(A_1)$. From Lemma 2.4, we have

$$1+\frac{1}{\mu(B^TA_1)} \leq \text{Spark}(B^TA_1) \quad \text{for every basis } B \text{ of } \mathcal{N}(A_2^T).$$

By (15), we see that $\text{Spark}(B^TA_1) \leq \text{Spark}^*_{A_2}(A_1)$, thus

$$1+\frac{1}{\mu(B^TA_1)} \leq \text{Spark}^*_{A_2}(A_1) \quad \text{for all } B \in F.$$

Since the right-hand side of the above is fixed, which is an upper bound for the left-hand side for any $B \in F$, we conclude that

$$\text{Spark}^*_{A_2}(A_1) \geq \sup_{B\in F}\left\{1+\frac{1}{\mu(B^TA_1)}\right\} = 1+\frac{1}{\inf_{B\in F}\{\mu(B^TA_1)\}} = 1+\frac{1}{\mu^*_{A_2}(A_1)}.$$

By Theorem 2.7 and Lemma 2.10, we have the next enhanced uniqueness claim.

Theorem 2.11 *For a given system (2) with $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$ and $m < n_1$. If there exists a solution $(x, y)^T$ satisfying*

$$\|x\|_0 < \frac{1}{2}\left(1+\frac{1}{\mu^*_{A_2}(A_1)}\right), \tag{20}$$

where $\mu^*_{A_2}(A_1)$ is the minimal scaled coherence of A_1 over F, then x is the unique sparsest x-part solution to the system (2).

Remark. The uniqueness criteria established in this section can be seen as certain generalization of that of sparsest solutions to systems of linear equations. For instance, when $A_2 = 0$, the null space of A_2^T is the whole space R^m. Hence, by letting $B = I$, the corresponding scaled mutual coherence and scaled spark become

$$\mu(B^TA_1) = \mu(A_1), \quad \text{spark}(B^TA_1) = \text{spark}(A_1).$$

The results in this section are reduced to the existing ones [6, 5, 4]. It is worth noting that the spark type uniqueness conditions are derived from the property of null spaces. It is worth mentioning that the null space based analysis is not the unique way to derive uniqueness criteria for sparsest solutions. Some other approaches such as the so-called range space property (see, e.g., [19, 20, 21]) and orthogonal projection from $R^{n_1+n_2}$ to $\mathcal{N}(A_2^T)$ [1] can be also used to develop uniqueness criteria.

3 Further Improvement of some uniqueness conditions

Since spark conditions are difficult to verify, the mutual coherence conditions play an important role in the uniqueness theory for the ℓ_0-minimization problem (1). As shown in Lemma 2.10, $1+\frac{1}{\mu^*_{A_2}(A_1)}$ is a good lower bound for $\mathrm{Spark}^*_{A_2}(A_1)$ which is an improved version of the bound (17). In this section, we aim to further enhance the uniqueness claim (20) by further improving the lower bound of $\mathrm{Spark}^*_{A_2}(A_1)$ under some situations. Following the discussions in [19], we introduce the so-called scaled coherence rank, scaled sub-coherence and scaled sub-coherence rank to achieve certain improvement on uniqueness conditions developed in section 2.

3.1 Maximal (sub) coherence and rank

Let us first recall several concepts which were introduced by Zhao [19]. For a given matrix $A \in R^{m\times n}$ with columns $a_i, i = 1, ..., n$, consider the index set

$$S_i(A) := \left\{ j: \ j \neq i, \ \frac{|a_i^T a_j|}{\|a_i\|_2 \cdot \|a_j\|_2} = \mu(A) \right\}, \ i = 1, ..., m.$$

Let $\alpha_i(A)$ be the cardinality of $S_i(A)$, and $\alpha(A)$ be the largest one among $\alpha_i(A)$'s, i.e.,

$$\alpha(A) = \max_{1\le i\le m} \alpha_i(A) = \max_{1\le i\le m} |S_i(A)|.$$

$\alpha(A)$ is called the coherence rank of A. Let i_0 be an index such that $\alpha(A) = \alpha_{i_0}(A) = |S_{i_0}(A)|$. Define

$$\beta(A) = \max_{1\le i\le m, \ i\neq i_0} \alpha_i(A) = \max_{1\le i\le m, \ i\neq i_0} |S_i(A)|,$$

which is called the sub-coherence rank of A. Also we define by

$$\mu^{(2)}(A) = \max_{i\neq j} \left\{ \frac{|a_i^T a_j|}{\|a_i\|_2 \cdot \|a_j\|_2} : \ \frac{|a_i^T a_j|}{\|a_i\|_2 \cdot \|a_j\|_2} < \mu(A) \right\},$$

the second largest absolute value of the inner product between two normalized columns of A. $\mu^{(2)}(A)$ is called the sub-mutual coherence of A.

Consider the sub-mutual coherence $\mu^{(2)}(B^T A_1)$ with a scaling matrix $B \in F$. We introduce the following new concept.

Definition 3.1 *Let $A_1 \in R^{m\times n_1}$ ($m < n_1$) and $A_2 \in R^{m\times n_2}$ be two matrices, and F is the set of bases of $\mathcal{N}(A_2^T)$.*

(i) The maximal scaled sub-mutual coherence of A_1 on F, denoted by $\mu^{**(2)}_{A_2}(A_1)$, is defined as

$$\mu^{**(2)}_{A_2}(A_1) = \sup_{B\in F} \mu^{(2)}(B^T A_1). \tag{21}$$

(ii) The maximal scaled coherence rank of A_1 on F, denoted by $\alpha^*_{A_2}(A_1)$, is defined as

$$\alpha^*_{A_2}(A_1) = \sup_{B\in F} \{\alpha(B^T A_1)\}. \tag{22}$$

(iii) The maximal scaled sub-coherence rank on F, denoted by $\beta^*_{A_2}(A_1)$, is defined as

$$\beta^*_{A_2}(A_1) = \sup_{B\in F}\{\beta(B^TA_1)\}. \tag{23}$$

It is easy to see the following relationship between $\alpha(B^TA_1)$, $\beta(B^TA_1)$, $\alpha^*_{A_2}(A_1)$ and $\beta^*_{A_2}(A_1)$: For every basis B of $\mathcal{N}(A_2^T)$, we have

$$1 \le \beta(B^TA_1) \le \alpha(B^TA_1) \le \alpha^*_{A_2}(A_1),\ 1 \le \beta(B^TA_1) \le \beta^*_{A_2}(A_1) \le \alpha^*_{A_2}(A_1). \tag{24}$$

3.2 Improved lower bounds of $Spark^*_{A_2}(A_1)$

Following the method used to improve the lower bound of Spark(A) in [19], we can find an enhanced lower bound of $\text{Spark}^*_{A_2}(A_1)$ via the concepts introduced in Section 3.1. We will make use of the following two lemmas.

Lemma 3.2 *(Brauer [3]) For any matrix $A \in R^{n\times n}$ with $n \ge 2$, if λ is an eigenvalue of A, there is a pair (i, j) of positive integers with $i \ne j$ $(1 \le i, j \le n)$ such that*

$$|\lambda - a_{ii}| \cdot |\lambda - a_{jj}| \le \Delta_i\Delta_j,$$

where $\Delta_i := \sum_{j=1,j\ne i}^{n} |a_{ij}|$ for $1 \le i \le n$.

Merging Theorem 2.5 and Proposition 2.6 in [19] yields the following result.

Lemma 3.3 *(Zhao [19]) Let $A \in R^{m\times n}$, and let $\alpha(A)$ and $\beta(A)$ be the coherence rank and sub-coherence rank of A, respectively. Suppose that one of the following conditions holds: (i) $\alpha(A) < \frac{1}{\mu(A)}$; (ii) $\alpha(A) \le \frac{1}{\mu(A)}$ and $\beta(A) < \alpha(A)$. Then $\mu^{(2)}(A) > 0$ and*

$$\begin{aligned} Spark(A) &\ge 1 + \frac{2[1-\alpha(A)\beta(A)\bar{\mu}(A)^2]}{\mu^{(2)}(A)\{\bar{\mu}(A)(\alpha(A)+\beta(A)) + \sqrt{\bar{\mu}(A)^2(\alpha(A)-\beta(A))^2+4}\}} \\ &> 1 + \frac{1}{\mu(A)} \end{aligned}$$

where $\bar{\mu}(A) = \mu(A) - \mu^{(2)}(A)$ and $\mu^{(2)}(A)$ is the subcoherence of A.

Based on Lemma 3.3, we can construct an enhanced lower bound of $\text{Spark}^*_{A_2}(A_1)$ under some conditions, in terms of the scaled coherence rank and scaled sub-coherence rank.

Theorem 3.4 *Consider the system (2) where $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$ and $m < n_1$. Suppose that one of the following conditions holds: (i) $\alpha(B^TA_1) < \frac{1}{\mu(B^TA_1)}$ for all $B \in F$; (ii) $\alpha(B^TA_1) \le \frac{1}{\mu(B^TA_1)}$ and $\beta(B^TA_1) < \alpha(B^TA_1)$ for all $B \in F$. Then for any $B \in F$, we have that $\mu^{(2)}(B^TA_1) > 0$ and*

$$\begin{aligned} Spark^*_{A_2}(A_1) &\ge \sup_{B\in F}\left\{1 + \frac{2[1-\alpha(B^TA_1)\beta(B^TA_1)\bar{\mu}(B^TA_1)^2]}{\mu^{(2)}(B^TA_1)\{\bar{\mu}(B^TA_1)(\alpha(B^TA_1)+\beta(B^TA_1)) + \sqrt{\Delta}\}}\right\} \\ &\ge 1 + \frac{1}{\mu^*_{A_2}(A_1)}. \end{aligned}$$

where $\bar{\mu}(B^TA_1) = \mu(B^TA_1) - \mu^{(2)}(B^TA_1)$ and $\Delta = [\bar{\mu}(B^TA_1)]^2(\alpha(B^TA_1) - \beta(B^TA_1))^2 + 4$.

Proof. Under conditions (i) and (ii), by Lemma 3.3, for any $B \in F$ we have that $\mu^{(2)}(B^T A_1) > 0$ and

$$\begin{aligned} \text{Spark}(B^T A_1) &\geq \varphi(B^T A_1) \\ &=: 1 + \frac{2[1 - \alpha(B^T A_1)\beta(B^T A_1)\bar{\mu}(B^T A_1)^2]}{\mu^{(2)}(B^T A_1)\{\bar{\mu}(B^T A_1)(\alpha(B^T A_1) + \beta(B^T A_1)) + \sqrt{\Delta}\}}, \end{aligned} \tag{25}$$

where $\bar{\mu}(B^T A_1) = \mu(B^T A_1) - \mu^{(2)}(B^T A_1)$ and $\Delta = [\bar{\mu}(B^T A_1)]^2(\alpha(B^T A_1) - \beta(B^T A_1))^2 + 4$. The above inequality holds for any basis $B \in F$. By the definition of $\text{Spark}^*_{A_2}(A_1)$, we have

$$\text{Spark}^*_{A_2}(A_1) \geq \text{Spark}(B^T A_1) \text{ for any } B \in F.$$

Thus it follows from (25) that

$$\text{Spark}^*_{A_2}(A_1) \geq \varphi(B^T A_1) \quad \text{for all } B \in F \tag{26}$$

Inequality (26) implies that the value of $\varphi(B^T A_1)$ is bounded by the constant $\text{Spark}^*_{A_2}(A_1)$. Hence, the supremum of $\varphi(B^T A_1)$ over F should be bounded by $\text{Spark}^*_{A_2}(A_1)$, namely,

$$\text{Spark}^*_{A_2}(A_1) \geq \sup_{B \in F} \varphi(B^T A_1).$$

By Lemma 3.3 again, under conditions (i) and (ii), we see that $\varphi(B^T A_1) > 1 + \frac{1}{\mu(B^T A_1)}$. Therefore, the superimum of $\varphi(B^T A_1)$ should be greater than the value of $1 + \frac{1}{\mu(B^T A_1)}$ for any basis $B \in F$, i.e.,

$$\sup_{B \in F} \varphi(B^T A_1) > 1 + \frac{1}{\mu(B^T A_1)} \text{ for any } B \in F.$$

This in turn implies that

$$\sup_{B \in F}\{\varphi(B^T A_1)\} \geq \sup_{B \in F}\left\{1 + \frac{1}{\mu(B^T A_1)}\right\} = 1 + \frac{1}{\mu^*_{A_2}(A_1)},$$

where the last equality follows from the definition of $\mu^*_{A_2}(A_1)$. Therefore, under conditions (i) and (ii), we conclude that

$$\text{Spark}^*_{A_2}(A_1) \geq \sup_{B \in F}\{\varphi(B^T A_1)\} \geq 1 + \frac{1}{\mu^*_{A_2}(A_1)},$$

as claimed.

Conditions (i) and (ii) in Theorem 3.4 rely on $B \in F$. A similar condition without relying on B can be also established as shown by the next result.

Theorem 3.5 *Consider the system (2) with $A_1 \in R^{m \times n_1}$, $A_2 \in R^{m \times n_2}$ and $m < n_1$. Let $\mu^{**}_{A_2}(A_1)$, and $\mu^{**(2)}_{A_2}(A_1), \alpha^*_{A_2}(A_1), \beta^*_{A_2}(A_1)$ be four constants defined by (18), (21)-(23), respectively. Suppose that one of the following conditions holds: (i) $\alpha^*_{A_2}(A_1) < \frac{1}{\mu^{**}_{A_2}(A_1)}$; (ii) $\alpha^*_{A_2}(A_1) \leq \frac{1}{\mu^{**}_{A_2}(A_1)}$ and $\beta^*_{A_2}(A_1) < \alpha^*_{A_2}(A_1)$. Then $\mu^{**(2)}_{A_2}(A_1) > 0$ and*

$$\text{Spark}^*_{A_2}(A_1) \geq \varphi^* = 1 + \frac{\sqrt{\rho} - (\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^*}{2\mu^{**(2)}_{A_2}(A_1)},$$

where $\bar{\mu}^* = \mu^{**}_{A_2}(A_1) - \mu^{**(2)}_{A_2}(A_1)$ and $\rho = \left(\alpha^*_{A_2}(A_1) - \beta^*_{A_2}(A_1)\right)^2 (\bar{\mu}^*)^2 + 4$.

Proof. Note that $\alpha(B^T A_1) \in \{1, ..., n_1 - 1\}$ for any $B \in F$. By the definition of $\alpha^*_{A_2}(A_1)$ which is the maximum value of $\alpha(B^T A_1)$ over F, this maximum is attainable, that is, there exists a $\widehat{B} \in F$ such that

$$\alpha^*_{A_2}(A_1) = \alpha(\widehat{B}^T A_1).$$

For such a basis $\widehat{B} \in F$, without loss of generality, we assume that all columns of $\widehat{B}^T A_1$ are normalized in the sense that the l_2-norm of every column of $\widehat{B}^T A_1$ is 1. Note also that the spark, mutual coherence, sub-coherence, coherence rank, and sub-coherence rank are invariant under normalization.

Let $p = \text{Spark}(\widehat{B}^T A_1)$ and $\{c_1, \cdots, c_p\}$ be the set of p columns from $\widehat{B}^T A_1$ that are linearly dependent. Denote C_p the submatrix consisting of these p columns. Then the Gram matrix of C_p, $G_{pp} = C_p^T C_p \in R^{p\times p}$, is singular. Since all diagonal entries of G_{pp} are 1's, and the absolute value of off-diagonal entries are less than or equal to $\mu(\widehat{B}^T A_1)$. Under either condition (i) or (ii) of the theorem, we have

$$\alpha^*_{A_2}(A_1) \le \frac{1}{\mu^{**}_{A_2}(A_1)} \le \frac{1}{\mu(B^T A_1)} \text{ for any } B \in F.$$

In particular, we have

$$\alpha^*_{A_2}(A_1) \le \frac{1}{\mu(\widehat{B}^T A_1)} \le \text{Spark}(\widehat{B}^T A_1) - 1 = p - 1. \tag{27}$$

Since G_{pp} is a $p \times p$ matrix, in each row of G_{pp}, there are at most $\alpha^*_{A_2}(A_1) = \alpha(\widehat{B}^T A_1)$ entries whose absolute values are equal to $\mu(\widehat{B}^T A_1)$, and the absolute values of the remaining $(p - 1 - \alpha^*_{A_2}(A_1))$ entries are less than or equal to $\mu^{(2)}(\widehat{B}^T A_1)$. By the singularity of G_{pp}, we know that $\lambda = 0$ is an eigenvalue of G_{pp}. By Lemma 3.2, there exist two rows of G_{pp}, say, the ith row and the jth row ($i \neq j$), satisfying that

$$|0 - G_{ii}| \cdot |0 - G_{jj}| \le \Delta_i \cdot \Delta_j = \sum_{t=1, t\neq i}^{p} |c_i^T c_t| \cdot \sum_{t=1, t\neq j}^{p} |c_j^T c_t|. \tag{28}$$

By the definitions of coherence rank and sub-coherence rank, if there are $\alpha^*_{A_2}(A_1)(= \alpha(\widehat{B}^T A_1))$ entries whose absolute values are $\mu(\widehat{B}^T A_1)$ in the ith row, then for the jth row, there are at most $\beta(\widehat{B}^T A_1)$ entries whose absolute values are $\mu(\widehat{B}^T A_1)$. And the absolute values of the remaining entries in either row are less than or equal to $\mu^{(2)}(\widehat{B}^T A_1)$. Therefore, from (28), we have that

$$\begin{aligned} 1 \ \le \ & [\alpha^*_{A_2}(A_1)\mu(\widehat{B}^T A_1) + (p - 1 - \alpha^*_{A_2}(A_1))\mu^{(2)}(\widehat{B}^T A_1)] \cdot \\ & [\beta(\widehat{B}^T A_1)\mu(\widehat{B}^T A_1) + (p - 1 - \beta(\widehat{B}^T A_1))\mu^{(2)}(\widehat{B}^T A_1)]. \end{aligned} \tag{29}$$

Let $p^* = \text{Spark}^*_{A_2}(A_1)$. Since $\text{Spark}^*_{A_2}(A_1)$ is the supremum of $\text{Spark}(B^T A_1)$ over F, we have $p \le p^*$. Thus it follows from (29) that

$$\begin{aligned} 1 \ \le \ & [\alpha^*_{A_2}(A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \alpha^*_{A_2}(A_1))\mu^{(2)}(\widehat{B}^T A_1)] \cdot \\ & [\beta(\widehat{B}^T A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \beta(\widehat{B}^T A_1))\mu^{(2)}(\widehat{B}^T A_1)]. \end{aligned} \tag{30}$$

By the definition of $\beta^*_{A_2}(A_1)$, we have $\beta(\widehat{B}^T A_1) \leq \beta^*_{A_2}(A_1)$. This, together with $\mu(\widehat{B}^T A) \geq \mu^{(2)}(\widehat{B}^T A_1)$, implies that

$$\begin{aligned} & \beta(\widehat{B}^T A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \beta(\widehat{B}^T A_1))\mu^{(2)}(\widehat{B}^T A_1) \\ \leq \ & \beta^*_{A_2}(A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \beta^*_{A_2}(A_1))\mu^{(2)}(\widehat{B}^T A_1). \end{aligned}$$

Combining (30) with the inequality above yields

$$\begin{aligned} 1 \ \leq \ & [\alpha^*_{A_2}(A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \alpha^*_{A_2}(A_1))\mu^{(2)}(\widehat{B}^T A_1)] \cdot \\ & [\beta^*_{A_2}(A_1)\mu(\widehat{B}^T A_1) + (p^* - 1 - \beta^*_{A_2}(A_1))\mu^{(2)}(\widehat{B}^T A_1)]. \end{aligned} \tag{31}$$

Note that

$$\beta^*_{A_2}(A_1) \leq \alpha^*_{A_2}(A_1) \leq p - 1 \leq p^* - 1, \ \mu(\widehat{B}^T A_1) \leq \mu^{**}_{A_2}(A_1), \ \mu^{(2)}(\widehat{B}^T A_1) \leq \mu^{**(2)}_{A_2}(A_1).$$

So from (31), we obtain

$$\begin{aligned} 1 \ \leq \ & [\alpha^*_{A_2}(A_1)\mu^{**}_{A_2}(A_1) + (p^* - 1 - \alpha^*_{A_2}(A_1))\mu^{**(2)}_{A_2}(A_1)] \cdot \\ & [\beta^*_{A_2}(A_1)\mu^{**}_{A_2}(A_1) + (p^* - 1 - \beta^*_{A_2}(A_1))\mu^{**(2)}_{A_2}(A_1)]. \end{aligned}$$

Denote by $\bar{\mu}^* := \mu^{**}_{A_2}(A_1) - \mu^{**(2)}_{A_2}(A_1)$. The above inequality can be written as

$$\left[(p^* - 1)\mu^{**(2)}_{A_2}(A_1)\right]^2 + (p^* - 1)(\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^* \mu^{**(2)}_{A_2}(A_1) + \alpha^*_{A_2}(A_1)\beta^*_{A_2}(A_1)(\bar{\mu}^*)^2 \geq 1. \tag{32}$$

By the definition of $\mu^{**(2)}_{A_2}(A_1)$, we know that $\mu^{**(2)}_{A_2}(A_1) \geq 0$. We now prove that $\mu^{**(2)}_{A_2}(A_1) > 0$. In fact, if $\mu^{**(2)}_{A_2}(A_1) = 0$, then the quadratic inequality (32) becomes

$$\alpha^*_{A_2}(A_1)\beta^*_{A_2}(A_1)\left(\mu^{**}_{A_2}(A_1)\right)^2 \geq 1,$$

which contradicts to either condition (i) or condition (ii) of the theorem. Thus $\mu^{**(2)}_{A_2}(A_1)$ must be positive. Consider the following quadratic equation in variable t :

$$h(t) := t^2 + t(\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^* + \alpha^*_{A_2}(A_1)\beta^*_{A_2}(A_1)(\bar{\mu}^*)^2 - 1 = 0$$

which has only one positive root under conditions (i) and (ii). This positive root is given by

$$t^* = \frac{-(\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^* + \sqrt{\rho}}{2},$$

where $\rho = (\alpha^*_{A_2}(A_1) - \beta^*_{A_2}(A_1))^2(\bar{\mu}^*)^2 + 4$. Let $\gamma = (p^* - 1)\mu^{**(2)}_{A_2}(A_1)$. The inequality (32) shows that $h(\gamma) \geq 0$. Thus $\gamma \geq t^*$, that is,

$$(p^* - 1)\mu^{**(2)}_{A_2}(A_1) \geq \frac{-(\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^* + \sqrt{\rho}}{2}.$$

Therefore,

$$\mathrm{Spark}^*_{A_2}(A_1) = p^* \geq 1 + \frac{\sqrt{\rho} - (\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^*}{2\mu^{**(2)}_{A_2}(A_1)},$$

as desired.

By Theorem 2.7 and Theorem 3.5, we immediately have the next uniqueness condition.

Corollary 3.6 *Consider the system (2) where* $A_1 \in R^{m\times n_1}$, $A_2 \in R^{m\times n_2}$, *and* $m < n_1$. *Under the same condition of Theorem 3.5. If there exists a solution* (x, y) *to the system (2) satisfying that*

$$\|x\|_0 < \frac{1}{2}\varphi^* =: \frac{1}{2}\left(1 + \frac{\sqrt{\rho} - (\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^*}{2\mu^{**(2)}_{A_2}(A_1)}\right),$$

then x *is the unique sparsest* x*-part solution to the system (2).*

The above corollary may also provide a tighter lower bound of $\text{Spark}^*_{A_2}(A_1)$ than Theorems 2.11 under some conditions, as indicated by the following proposition.

Proposition 3.7 *Let* φ^* *be a lower bound of* $\text{Spark}^*_{A_2}(A_1)$ *given in Theorem 3.5. Assume that* $\alpha^*_{A_2}(A_1) = 1$ *and* $\alpha^*_{A_2}(A_1) < \frac{1}{\mu^{**}_{A_2}(A_1)}$. *If* $\mu^{**(2)}_{A_2}(A_1) < \mu^*_{A_2}(A_1)(1 - \bar{\mu}^*)$ *where* $\bar{\mu}^* = \mu^{**}_{A_2}(A_1) - \mu^{**(2)}_{A_2}(A_1)$, *we have* $\varphi^* > 1 + \frac{1}{\mu^*_{A_2}(A_1)}$.

Proof. Under condition $\alpha^*_{A_2}(A_1) < \frac{1}{\mu^{**}_{A_2}(A_1)}$, by Theorem (3.5) we get the following lower bond of $\text{Spark}^*_{A_2}(A_1)$:

$$\varphi^* = 1 + \frac{\sqrt{\rho} - (\alpha^*_{A_2}(A_1) + \beta^*_{A_2}(A_1))\bar{\mu}^*}{2\mu^{**(2)}_{A_2}(A_1)}. \tag{33}$$

By (24), we see that $\alpha^*_{A_2}(A_1) = 1$ implies that $\beta^*_{A_2}(A_1) = 1$. Thus (33) is reduced to $\varphi^* - 1 = \frac{1-\bar{\mu}^*}{\mu^{**(2)}_{A_2}(A_1)}$. Note that

$$\begin{aligned}\frac{1-\bar{\mu}^*}{\mu^{**(2)}_{A_2}(A_1)} &= \frac{1}{\mu^*_{A_2}(A_1)} + \left(\frac{1-\bar{\mu}^*}{\mu^{**(2)}_{A_2}(A_1)} - \frac{1}{\mu^*_{A_2}(A_1)}\right) \\ &= \frac{1}{\mu^*_{A_2}(A_1)} + \frac{\mu^*_{A_2}(A_1)(1-\bar{\mu}^*) - \mu^{**(2)}_{A_2}(A_1)}{\mu^{**(2)}_{A_2}(A_1)\mu^*_{A_2}(A_1)},\end{aligned}$$

Thus if $\mu^{**(2)}_{A_2}(A_1) < \mu^*_{A_2}(A_1)(1 - \bar{\mu}^*)$, we must have $\varphi^* > 1 + \frac{1}{\mu^*_{A_2}(A_1)}$.

The discussion in this section demonstrates that the concepts introduced in this section such as maximal scaled coherence rank and sub-coherence rank, and minimal/maximal scaled mutual coherence are quite useful in the development of uniqueness criteria for the ℓ_0-minimization problem (1).

4 Conclusion

In this paper, we have established several uniqueness conditions for the solution to a class of ℓ_0-minimization problems which seek sparsity only for part of the variables of the problem. This problem includes several important sparsity-seeking models as special cases. To obtain uniqueness conditions, several concepts such as maximal/minimal scaled coherence, maximal scaled coherence rank, and maximal scaled spark have been introduced in this paper. Also the l_p-induced quasi-norm has been defined and used to establish a sufficient condition for the uniqueness of the solution to the underlying ℓ_0-minimization problems as well.

References

[1] A. S. Bandeira and K. Scheinberg and L. N. Vicente, On partial sparse recovery, Submitted to *IEEE Signal Processing letters*, 2013

[2] A. S. Bandeira and K. Scheinberg and L. N. Vicente, Computation of sparse low degree interpolating polynomials and their application to derivative-free optimization, *Mathematical Programming*, 134 (2012), 223-257.

[3] A. Brauer, Limits for the characteristic roots of a matrix, *Duke Math. J.*, 13 (1946), 387-395.

[4] A. Bruckstein and M.Elad and D. Donoho, From sparse solutions of systems of equations to sparse modeling of signals and images, *SIAM Review*, 51(2009), 34-81.

[5] M. Davenport and M. Duarte and Y. Eldar and G. Kutyniok, *Introduction to Compressive Sensing*, in Compressive Sensing: Theory and Applications (Eldar and Kutyniok eds.), University of Cambridge, 2012.

[6] D. Donoho and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via L^1 minimization, *Proc. Nat. Acad. Sci.*, 100 (2003), 2197-2202.

[7] D. Donoho and X. Huo, Uncertainty principles and ideal atomic decompositions, *IEEE Transactions on Information Theory*, 47 (2001), 2845-2862.

[8] A. Doostan and H. Owhadi and A. Lashgari and G. Iaccarino, Non-adapted sparse approximation of PDEs with stochastic inputs, *Journal of Computational Physics*, 230 (2011), 3015-3034.

[9] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, New York, 2010.

[10] Y. Eldar and G. Kutyniok, *Compressive Sensing: Theory and Applications*, Cambridge University Press, 2012.

[11] M. Figueiredo and R. Nowak and S. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, *IEEE Journal of Selected Topics in Signal Processing*, 1 (2007), 586-597.

[12] S. Foucart and A. Pajor and H. Rauhut and T. Ullrich, The Gelfand widths of l_p-balls for $0 < p < 1$, *Journal of Complexity*, 26 (2010), 629-640.

[13] L. Jacques, A short note on compressed sensing with partially known signal support, *Signal Processing*, 90 (2010), 3308-3312.

[14] M. Lustig and D. Donoho and J. Pauly, Sparse MRI: the application of compressed sensing for rapid MR imaging, *Magnetic Resonance in Medicine* 58 (2007), 11821195.

[15] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[16] B. Natarajan, Sparse approximate solutions to linear systems, *SIAM Journal on Computing*, 24 (1995), 227-234.

[17] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic, 2001.

[18] N. Vaswani and W. Lu, Modified-CS: Modifying compressive sensing for problems with partially known support, *IEEE Transactions on Signal Processing*, 58 (2010), 4595-4607.

[19] Y.B. Zhao, New and improved conditions for uniqueness of sparsest solutions of underdetermined linear systems, *Applied Mathematics and Computation*, 224 (2013), 58-73.

[20] Y.B. Zhao, RSP-Based analysis for sparsest and least l_1-norm solutions to underdetermined linear systems, *IEEE Transactions on Signal Processing*, 61 (2013), no. 22, 5777-5788.

[21] Y.B. Zhao, Equivalence and strong equivalence between sparsest and least ℓ_1-norm nonnegative solutions of linear systems and their application, Technical Report, Univ of Birmingham, June 2012.

[22] Y.B. Zhao and D. Li, Reweighted ℓ_1-minimization for sparse solutions to underdetermined linear system, *SIAM Journal on Optimization*, 22 (2012), no. 3, 1065-1088.