Wang and Peng *Journal of Inequalities and Applications* (2017) 2017:313
https://doi.org/10.1186/s13660-017-1590-x

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH Open Access

Analysis of the equivalence relationship between l_0-minimization and l_p-minimization

Changlong Wang and Jigen Peng*

*Correspondence:
jgpengxjtu@126.com
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

Abstract

In signal processing theory, l_0-minimization is an important mathematical model. Unfortunately, l_0-minimization is actually NP-hard. The most widely studied approach to this NP-hard problem is based on solving l_p-minimization ($0 < p \leq 1$). In this paper, we present an analytic expression of $p^*(A, b)$, which is formulated by the dimension of the matrix $A \in \mathbb{R}^{m\times n}$, the eigenvalue of the matrix A^TA, and the vector $b \in \mathbb{R}^m$, such that every k-sparse vector $x \in \mathbb{R}^n$ can be exactly recovered via l_p-minimization whenever $0 < p < p^*(A, b)$, that is, l_p-minimization is equivalent to l_0-minimization whenever $0 < p < p^*(A, b)$. The superiority of our results is that the analytic expression and each its part can be easily calculated. Finally, we give two examples to confirm the validity of our conclusions.

MSC: 60E05; 94A12; 94A20

Keywords: sparse recovery; null space constant; null space property; l_0-minimization; l_p-minimization

1 Introduction

In sparse information theory, a central goal is to get the sparsest solutions of underdetermined linear systems including visual coding [1], matrix completion [2], source localization [3], and face recognition [4]. All these problems are popularly modeled by the following l_0-minimization:

$$\min_{x\in\mathbb{R}^n} \|x\|_0 \quad \text{s.t.} \quad Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ is an underdetermined matrix (*i.e.* $m < n$), and $\|x\|_0$ is the number of nonzero elements of x, which is commonly called l_0-norm although it is not a true vector norm. If $x \in \mathbb{R}^n$ is a unique solution of l_0-minimization, we also say that x can be recovered by l_0-minimization; we adopt these two statements in this paper.

Since A has more columns than rows, the underdetermined linear system $Ax = b$ admits an infinite number of solutions. To find the sparsest one, much excellent theoretical work (see, *e.g.* [5, 6], and [7]) has been devoted to the l_0-minimization. However, Natarajan [8] proved that l_0-minimization is NP-hard. Furthermore, it is combinationally and computationally intractable to solve l_0-minimization directly because of its discrete and discontinuous nature. Therefore, a lot of work put forward some alternative strategies to get the sparsest solution (see, *e.g.* [5, 9–14], and [15]). Among these methods, the most popular one is l_p-minimization with $0 < p \leq 1$ introduced by Gribonval and Nielsen [16],

$$\min_{x\in\mathbb{R}^n} \|x\|_p^p \quad \text{s.t.} \quad Ax = b, \tag{2}$$

where $\|x\|_p^p = \sum_{i=1}^n |x_i|^p$. In the literature, $\|x\|_p$ is still called the p-norm of x though it is only a quasi-norm when $0 < p < 1$ (because in this case it violates the triangle inequality). Due to the fact that $\|x\|_0 = \lim_{p\to 0} \|x\|_p^p$, l_0-minimization and l_p-minimization are collectively called l_p-minimization with $0 \leq p \leq 1$ in this paper.

However, to get the sparsest solution of $Ax = b$ via l_p-minimization, we need certain conditions on A and/or b, for example, the novel restricted isometry property (RIP) of A. A matrix A is said to have restricted isometry property of order k with restricted isometry constant $\delta_k \in (0, 1)$ if δ_k is the smallest constant such that

$$(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2 \tag{3}$$

for all k-sparse vectors x, where a vector x is said to be k-sparse if $\|x\|_0 \leq k$.

There exist a lot of sufficient conditions for the exact recovery by l_1-minimization, such as $\delta_{3k} + 3\delta_{4k} < 2$ in [10], $\delta_{2k} < \sqrt{2} - 1$ in [9], and $\delta_{2k} < 2(3 - \sqrt{2})/7$ in [11]. Cai and Zhang [17] showed that for any given $t \geq \frac{4}{3}$, the condition $\delta_{tk} < \sqrt{\frac{t-1}{t}}$ guarantees recovery of every k-sparse vector by l_1-minimization. From the definition of p-norm it seems to be more natural to consider l_p-minimization with $0 < p < 1$ instead of l_0-minimization. Foucart [11] showed that the condition $\delta_{2k} < 0.4531$ can guarantee exact k-sparse recovery via l_p-minimization for any $0 < p < 1$. Chartrand [18] proved that if $\delta_{2k+1} < 1$, then we can recover a k-sparse vector by l_p-minimization for some $p > 0$ small enough. However, it should be pointed out that the problem of calculating δ_{2k} for a given matrix A is still NP-hard.

Recently, Peng, Yue, and Li [7] have proved that there exists a constant $p(A, b) > 0$ such that every solution of l_p-minimization is also a solution of l_0-minimization whenever $0 < p < p(A, b)$. This result builds a bridge between l_p-minimization and l_0-minimization, and it is important that this conclusion is not limited by the structure of a matrix A. However, the paper [7] does not give an analytic expression of $p(A, b)$. The model of choice of l_p-minimization is still difficult.

As already mentioned, it is NP-hard to calculate δ_{2k} for a given matrix $A \in \mathbb{R}^{m\times n}$ and also to calculate these p. On the other hand, the possibility of recovery of every k-sparse vector by l_0-minimization is just a necessary condition for the existence of such δ_{2k}, and therefore the results based on δ_{2k} lead to limitations of practical application.

We have to emphasize that although l_p-minimization is also difficult due to its nonconvexity and nonsmoothness, a lot of algorithms have been designed to solve l_p-minimization; see *e.g.* [11, 19], and [20]. Moreover, a reasonable range of p in these algorithms is very important. In this paper, we devote ourselves to giving a complete answer to this problem.

Our paper is organized as follows. In Section 2, we present some preliminaries of the l_p-null space property, which plays a core role in the proof of our main theorem. In Section 3, we focus ourselves on proving the main results of this paper: we present an analytic expression of $p^*(A, b)$ such that every k-sparse vector $x \in \mathbb{R}^n$ can be exactly recovered via

l_p-minimization with $0 < p < p^*(A, b)$ as long as x can be recovered via l_0-minimization. Finally, we summarize our findings in the last section.

For convenience, for $x \in \mathbb{R}^n$, its support is defined by support$(x) = \{i : x_i \neq 0\}$, and the cardinality of a set Ω is denoted by $|\Omega|$. Let $\mathrm{Ker}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ be the null space of a matrix A, and denote by $\lambda_{\min^+}(A)$ the minimal nonzero absolute-value eigenvalue of A and by $\lambda_{\max}(A)$ the maximal one. We denote by x_Ω the vector that is equal to x on the index set Ω and zero elsewhere and by A_Ω the submatrix the columns of which are the columns of A that are in the set index Ω. Let Ω^c be the complement of Ω.

2 Preliminaries

To investigate conditions under which both l_0-minimization and l_p-minimization have the same unique solution, it is convenient for us to work with a sufficient and necessary condition of the solutions of l_0-minimization and l_p-minimization. Therefore, in this preliminary section, we focus on introducing such an condition, namely the l_p-null space property.

Definition 1 ([16]) A matrix $A \in \mathbb{R}^{m\times n}$ with $m \leq n$ is said to satisfy the l_p-null space property of order k if

$$\|x_\Omega\|_p < \|x_{\Omega^c}\|_p \tag{4}$$

for every $x \in \mathrm{Ker}(A) \setminus \{\mathbf{0}\}$ and every set $\Omega \subset \{1, 2, \dots, n\}$ with $|\Omega| \leq k$.

In the literature, the null space property usually means the l_1-null space property. We now indicate the relation between the l_p-null space property and exact recovery via l_p-minimization with $0 \leq p \leq 1$.

Theorem 1 ([16, 21]) *Given a matrix $A \in \mathbb{R}^{m\times n}$ with $m \leq n$, every k-sparse vector $x \in \mathbb{R}^n$ can be recovered via l_p-minimization with $0 \leq p \leq 1$ if and only if A satisfies the l_p-null space property of order k.*

Theorem 1 provides a sufficient and necessary condition to judge whether a vector can be recovered by l_p-minimization with $0 \leq p \leq 1$, which is the most important advantage of the l_p-null space property. However, the l_p-null space property is difficult to be checked for a given matrix. To reach our goal, we recall the concept of the null space constant (NSC), which is closely related to the l_p-null space property and offers tremendous help in illustrating the performance of l_0-minimization and l_p-minimization.

Definition 2 ([22]) For any $0 \leq p \leq 1$ and $k > 0$, the null space constant (NSC) $h(p, A, k)$ is the smallest number such that:

$$\sum_{i\in\Omega} |x_i|^p \leq h(p, A, k) \sum_{i\notin\Omega} |x|^p \tag{5}$$

for every index set $\Omega \subset \{1, 2, \dots, n\}$ with $|\Omega| \leq k$ and every $x \in \mathrm{Ker}(A) \setminus \{\mathbf{0}\}$.

Similarly to the l_p-null space property, NSC also can be used for characterizing the performance of l_p-minimization. Combining the definition of NSC and the results in [23] and [22], we can derive the following corollaries.

Corollary 1 *For any* $p \in [0,1]$, $h(p,A,k) < 1$ *is a sufficient and necessary condition for recovery of all* k*-sparse vectors via* l_p*-minimization with* $0 \leq p \leq 1$.

Proof The proof is very easy, and we leave it to the readers. □

Corollary 2 *Given a matrix* $A \in \mathbb{R}^{m\times n}$, *if* $h(0,A,k) < 1$, *we have*:

(a) $\|x\|_0 \geq 2k+1$ *for every* $x \in \operatorname{Ker}(A) \setminus \{\mathbf{0}\}$;

(b) $k \leq \lceil \frac{n-2.5}{2} \rceil + 1$, *where* $\lceil a \rceil$ *represents the integer part of* a.

Proof (a) We assume that there exists a vector $x \in \operatorname{Ker}(A) \setminus \{\mathbf{0}\}$ with $\|x\|_0 \leq 2k$.

Let $\Omega = \operatorname{support}(x)$. If $\|x\|_0 \leq k$, then we get that $\|x_\Omega\|_0 \geq \|x_{\Omega^c}\|_0 = 0$.

If $k < \|x\|_0 \leq 2k$, we consider an arbitrary set $\widetilde{\Omega} \subset \Omega$ with $|\widetilde{\Omega}| = k$. Then we get that $\|x_{\widetilde{\Omega}}\|_0 = k \geq \|x_{\widetilde{\Omega}^c}\|_0$.

According to the definition of $h(p,A,k)$, these two conclusions contradict $h(0,A,k) < 1$, and therefore we have that $\|x\|_0 \geq 2k+1$ for any $x \in \operatorname{Ker}(A) \setminus \{\mathbf{0}\}$.

(b) As has been proved in (a), we get that

$$2k+1 \leq \|x\|_0 \leq n. \tag{6}$$

Due to the integer values of $\|x\|_0$ and k, it is easy to get that

$$k \leq \begin{cases} \frac{n-1}{2} & \text{when } n \text{ is odd,} \\ \frac{n-2}{2} & \text{when } n \text{ is even.} \end{cases}$$

In total, we have $k \leq \lceil \frac{n-2.5}{2} \rceil + 1$. □

Remark 1 In Corollary 2, we obtained a relation of inequality between n and k under the assumption $h(0,A,k) < 1$. Furthermore, Foucart [23, p. 49, Chapter 2] showed another relation of inequality between m and k. If every k-sparse vector $x \in \mathbb{R}^n$ can be recovered via l_0-minimization, then we get that $m \geq 2k$; furthermore, it is easy to get that $k \leq \lceil \frac{m}{2} \rceil$ due to the integer values of k.

Remark 2 Chen and Gu [22] showed some important properties of $h(p,A,k)$. It is shown that $h(p,A,k)$ is a continuous function in $p \in [0,1]$ when $k \leq \operatorname{spark}(A) - 1$, where $\operatorname{spark}(A)$ is the smallest number of columns from A that are linearly dependent. Therefore, if $h(0,A,k) < 1$ for some fixed A and k, then there exists a constant p^* such that $h(p,A,k) < 1$ for $p \in [0,p^*)$, that is, every k-sparse vector can be recovered via both l_0-minimization and l_p-minimization for $p \in (0,p^*)$, which is a corollary of the main theorem in [7].

Theorem 2 ([7]) *There exists a constant* $p(A,b) > 0$ *such that when* $0 < p < p(A,b)$, *every solution to* l_p*-minimization also solves* l_0*-minimization*.

Theorem 2 is the main theorem in [7]. Obviously, this theorem qualitatively proves the effectiveness of solving the original l_0-minimization problem via l_p-minimization. Moreover, the theorem becomes more practical if $p(A,b)$ is computable. At the end of this section, we need to point out a necessary and sufficient condition based on the l_p-null space property, and NSC can provide us the following lemma, which is similar to RIP.

Lemma 1 *Given an underdetermined matrix $A \in \mathbb{R}^{m\times n}$ and an integer k, the inequality $h(0,A,k) < 1$ holds if and only if there exist two constants $0 < u \le w$ with*

$$0 < \lambda_{\min^+}\left(A^T A\right) \le u^2 \le w^2 \le \lambda_{\max}\left(A^T A\right) \tag{7}$$

such that

$$u^2\|x\|_2^2 \le \|Ax\|_2^2 \le w^2\|x\|_2^2 \tag{8}$$

for every $2k$-sparse vector $x \in \mathbb{R}^n$.

Proof *Necessity.* The proof is divided into two steps.

Step 1: Proof of the existence of u.

To prove this result, we just need to prove that the set

$$V = \left\{u : \|Ax\|_2/\|x\|_2 \ge u \text{ for any nonzero } x \text{ with } \|x\|_0 \le 2k\right\}$$

has a nonzero infimum.

If we assume that $\inf V = 0$, then, for any $n \in N^+$, there exists a vector $2k$-sparse vector x_n with $\|x_n\|_2 = 1$ such that $\|Ax_n\|_2 \le n^{-1}$.

Furthermore, it is easy to get a convergent subsequence $\{x_{n_i}\}$ of the bounded sequence $\{x_n\}$, that is, $x_{n_i} \to x_0$, and it is obvious that $Ax_0 = \mathbf{0}$ because the function $y(x) = Ax$ is continuous.

Let $J(x_0) = \{i : (x_0)_i \neq 0\}$. There exists N_i such that $(x_{n_k})_i \neq 0$ when $k \ge N_i$ for any $i \in J(x_0)$.

Let $N = \max_{i\in J(x_0)} N_i$. For any $i \in J(x_0)$, it is easy to get that $(x_{n_k})_i \neq 0$ when $k \ge N$. When $k \ge N$, we get that $\|x_{n_k}\|_0 \ge \|x_0\|_0$ and $\|x_0\|_0 \le 2k$.

However, according to Corollary 2, it is easy to get that $\|x\|_0 \ge 2k+1$ for any $x \in \mathrm{Ker}(A) \setminus \{\mathbf{0}\}$. We notice that $x_0 \in \mathrm{Ker}(A)$, so the result $\|x_0\|_0 \le 2k$ contradicts Corollary 2.

Therefore, there exists a constant $u > 0$ such that $\|Ax\|_2 \ge u\|x\|_2$ for any $x \in \mathbb{R}^n$ with $\|x\|_0 \le 2k$.

Step 2: Proof of $u^2 \ge \lambda_{\min^+}(A^T A)$.

According to the proof above, there exists a vector $\widetilde{x} \in \mathbb{R}^n$ with $\|\widetilde{x}\|_0 \le 2k$ such that $\|A\widetilde{x}\|_2 = u\|\widetilde{x}\|_2$.

Let $V = \mathrm{support}(\widetilde{x})$. It is easy to get that

$$u^2 x^T x \le x^T A_V^T A_V x \tag{9}$$

for all $x \in \mathbb{R}^{|V|}$. Therefore, the smallest eigenvalue of $A_V^T A_V$ is u^2 since $A_V^T A_V \in R^{|V|\times|V|}$ is a symmetric matrix, and we can choose an eigenvector $z \in R^{|V|}$ of eigenvalue u^2.

If $u^2 < \lambda_{\min^+}(A^T A)$, then consider the vector $x' \in \mathbb{R}^n$ with $x_i{}' = z_i$ when $i \in V$ and zero otherwise. Therefore, it is easy to get that $A^T A x' = u^2 x'$, which contradicts the definition of $\lambda_{\min^+}(A^T A)$.

Finally, notice that $A^T A$ is a semipositive definite matrix such that $\|Ax\|_2^2 = x^T A^T A x \le \lambda_{\max}(A^T A)\|x\|_2^2$ for all $x \in \mathbb{R}^n$. So there exists a constant w such that $\|Ax\|_2^2 \le w^2\|x\|_2^2$ for all $\|x\|_0 \le 2k$.

Sufficiency. Let a k-sparse vector x^* be the unique solution of l_0-minimization. For any k-sparse vector x_1, we have that

$$u^2\|x^*-x_1\|_2^2 \le \|A(x^*-x_1)\|_2^2 \le w^2\|x^*-x_1\|_2^2. \tag{10}$$

Therefore, we get that $x^* = x_1$ as long as x_1 is a solution of $Ax = b$, that is, every k-sparse vector can be recovered by l_0-minimization (1), and this is equivalent to $h(0, A, k) < 1$ by Corollary 1. □

3 Main contribution

In this section, we focus ourselves on the proposed problem. By introducing a new technique and utilizing preparations provided in Section 2, we will present an analytic expression of $p^*(A, b)$ such that every k-sparse vector x can be recovered via l_p-minimization with $0 < p < p^*(A, b)$ as long as it can be recovered via l_0-minimization. To this end, we first begin with two lemmas.

Lemma 2 *For any $x \in \mathbb{R}^n$ and $0 < p \le 1$, we have that*

$$\|x\|_p \le \|x\|_0^{\frac{1}{p}-\frac{1}{2}}\|x\|_2.$$

Proof This result can be easily proved by Hölder's inequality. □

Lemma 3 *Given a matrix $A \in \mathbb{R}^{m\times n}$, if $u\|x\|_2 \le \|Ax\|_2 \le w\|x\|_2$ for all $\|x\|_0 \le 2k$, then*

$$\left|\langle Ax_1, Ax_2\rangle\right| \le \frac{w^2-u^2}{2}\|x_1\|_2\|x_2\|_2$$

for all x_1 and x_2 with $\|x_i\|_0 \le k$ $(i = 1, 2)$, and $\operatorname{support}(x_1) \cap \operatorname{support}(x_2) = \varnothing$.

Proof By the assumption on the matrix A, it is easy to get that

$$\begin{aligned}\frac{|\langle Ax_1, Ax_2\rangle|}{\|x_1\|_2\|x_2\|_2} &= \left|\left\langle A\left(\frac{x_1}{\|x_1\|_2}\right), A\left(\frac{x_2}{\|x_2\|_2}\right)\right\rangle\right| \\ &= \frac{1}{4}\left|\left\|A\left(\frac{x_1}{\|x_1\|_2}+\frac{x_2}{\|x_2\|_2}\right)\right\|_2^2 - \left\|A\left(\frac{x_1}{\|x_1\|_2}-\frac{x_2}{\|x_2\|_2}\right)\right\|_2^2\right| \\ &\le \frac{1}{4}\left|w^2\left\|\frac{x_1}{\|x_1\|_2}+\frac{x_2}{\|x_2\|_2}\right\|_2^2 - u^2\left\|\frac{x_1}{\|x_1\|_2}-\frac{x_2}{\|x_2\|_2}\right\|_2^2\right|.\end{aligned} \tag{11}$$

Since $\operatorname{support}(x_1) \cap \operatorname{support}(x_2) = \varnothing$, we have that

$$\left\|\frac{x_1}{\|x_1\|_2}+\frac{x_2}{\|x_2\|_2}\right\|_2^2 = \left\|\frac{x_1}{\|x_1\|_2}-\frac{x_2}{\|x_2\|_2}\right\|_2^2 = 2, \tag{12}$$

from which we get that

$$\left|\langle Ax_1, Ax_2\rangle\right| \le \frac{w^2-u^2}{2}\|x_1\|_2\|x_2\|_2. \tag{13}$$

□

With the above lemmas in hand, we now can prove our main theorems.

Theorem 3 *Given a matrix $A \in \mathbb{R}^{m\times n}$ with $m \leq n$ and $0 < p \leq 1$, if $h(0,A,k) < 1$, then*

$$h(p,A,k) < h^*(p,A,k), \tag{14}$$

where

$$h^*(p,A,k) = \left(\frac{\sqrt{2}+1}{2}\right)^p \left(\frac{k}{k+1}\right)\left[\frac{(\lambda-1)(n-2-k)}{2k} + \left(\lambda + \sqrt{\frac{1}{k+1}}\right)k^{-\frac{1}{2}}\right]^p \tag{15}$$

with

$$\lambda = \frac{\lambda_{\max}(A^TA)}{\lambda_{\min^+}(A^TA)}.$$

Proof According to Theorem 1 and Corollary 2, it is easy to get that $\|x\|_0 \geq 2k+1$ for every $x \in \operatorname{Ker}(A) \setminus \{\mathbf{0}\}$ since $h(0,A,k) < 1$. Furthermore, according to Lemma 1, we can find constants $\lambda_{\min^+}(A^TA) \leq u^2 \leq w^2 \leq \lambda_{\max}(A^TA)$ such that

$$u\|\tilde{x}\|_2 \leq \|A\tilde{x}\|_2 \leq w\|\tilde{x}\|_2 \tag{16}$$

for any $\tilde{x} \in \mathbb{R}^n$ with $\|\tilde{x}\|_0 \leq 2k$.

Now we consider a nonzero vector $x \in \operatorname{Ker}(A) \setminus \{\mathbf{0}\}$ and an arbitrary index set $\Omega_0 \subset \{1,2,\dots,n\}$ with $|\Omega_0| = k$. We partition the complement of Ω_0 as $\Omega_0^c = \bigcup_{i=1}^t \Omega_i$, where

$$\Omega_1 = \{\text{indices of the } k+1 \text{ largest absolute-value components of } x - x_{\Omega_0}\},$$

$$\Omega_2 = \{\text{indices of the } k \text{ largest absolute-value components of } x - x_{\Omega_0} - x_{\Omega_1}\},$$

$$\Omega_3 = \{\text{indices of the } k \text{ largest absolute-value components of } x - x_{\Omega_0} - x_{\Omega_1} - x_{\Omega_2}\},$$

$$\dots$$

$$\Omega_t = \{\text{indices of the remaining components of } x\}.$$

We know that $\|x\|_0 \geq 2k+1$, so both Ω_1 and Ω_0 are not empty, and there are only two cases:

(i) Ω_0 and Ω_i ($i = 2,\dots,t-1$) all have k elements except, possibly, Ω_t.

(ii) Ω_0 has k elements, Ω_1 has less than $k+1$ elements, and Ω_i ($i = 2,\dots,t-1$) are empty.

Furthermore, in both cases, the set Ω_1 can be divided in two parts:

$$\Omega_1^{(1)} = \{\text{indices of the } k \text{ largest absolute-value components of } \Omega_1\},$$

$$\Omega_1^{(2)} = \{\text{indices of the rest components of } \Omega_1\}.$$

It is obvious that $\Omega_1 = \Omega_1^{(1)} \cup \Omega_1^{(2)}$ and the set $\Omega_1^{(2)}$ is not empty since $\|x\|_0 \geq 2k+1$.

Since $u\|\tilde{x}\|_2 \le \|A\tilde{x}\|_2 \le w\|\tilde{x}\|_2$ for any $\|\tilde{x}\|_0 \le 2k$, we have that

$$\begin{aligned}
\|x_{\Omega_0}\|_2^2 + \|x_{\Omega_1}\|_2^2 &= \|x_{\Omega_0}\|_2^2 + \|x_{\Omega_1^{(1)}}\|_2^2 + \|x_{\Omega_1^{(2)}}\|_2^2 \\
&= \|x_{\Omega_0} + x_{\Omega_1^{(1)}}\|_2^2 + \|x_{\Omega_1^{(2)}}\|_2^2 \\
&\le \frac{1}{u^2}\|A(x_{\Omega_0} + x_{\Omega_1^{(1)}})\|_2^2 + \left\|x_{\Omega_1^{(2)}}\right\|_2^2. \qquad (17)
\end{aligned}$$

Since $x = x_{\Omega_0} + x_{\Omega_1^{(1)}} + x_{\Omega_1^{(2)}} + x_{\Omega_2} + \cdots + x_{\Omega_t} \in \mathrm{Ker}(A) \setminus \{\mathbf{0}\}$, we have that

$$\begin{aligned}
\left\|A(x_{\Omega_0} + x_{\Omega_1^{(1)}})\right\|_2^2 &= \left\langle A(-x_{\Omega_0} - x_{\Omega_1^{(1)}}, A(x_{\Omega_1^{(2)}} + x_{\Omega_2} + \cdots + x_{\Omega_t})\right\rangle \\
&= \left\langle A(-x_{\Omega_0} - x_{\Omega_1^{(1)}}), Ax_{\Omega_1^{(2)}}\right\rangle \\
&\quad + \sum_{i=2}^{t}\left(\left\langle A(-x_{\Omega_0}), Ax_{\Omega_i}\right\rangle + \left\langle A(-x_{\Omega_1^{(1)}}), Ax_{\Omega_i}\right\rangle\right). \qquad (18)
\end{aligned}$$

According to Lemma 3, for any $i \in \{2, 3, \ldots, t\}$, we get that

$$\begin{aligned}
\left\langle A(-x_{\Omega_0}), Ax_{\Omega_i}\right\rangle &\le \frac{w^2 - u^2}{2}\|x_{\Omega_0}\|_2\|x_{\Omega_i}\|_2, \\
\left\langle A(-x_{\Omega_1^{(1)}}), Ax_{\Omega_i}\right\rangle &\le \frac{w^2 - u^2}{2}\|x_{\Omega_1^{(1)}}\|_2\|x_{\Omega_i}\|_2.
\end{aligned} \qquad (19)$$

Substituting inequalities (19) into (18), we have

$$\begin{aligned}
\left\|A(x_{\Omega_0} + x_{\Omega_1^{(1)}})\right\|_2^2 &\le \left\|A(-x_{\Omega_0} - x_{\Omega_1^{(1)}})\right\|_2\|Ax_{\Omega_1^{(2)}}\| \\
&\quad + \frac{w^2 - u^2}{2}\left(\sum_{i=2}^{t}\|x_{\Omega_i}\|_2\right)\left(\|x_{\Omega_0}\|_2 + \|x_{\Omega_1^{(1)}}\|_2\right) \\
&\le w^2\left(\|x_{\Omega_0}\|_2 + \|x_{\Omega_1^{(1)}}\|_2\right)\|x_{\Omega_1^{(2)}}\|_2 \\
&\quad + \frac{w^2 - u^2}{2}\left(\sum_{i=2}^{t}\|x_{\Omega_i}\|_2\right)\left(\|x_{\Omega_0}\|_2 + \|x_{\Omega_1^{(1)}}\|_2\right). \qquad (20)
\end{aligned}$$

By the definition of $x_{\Omega_1^{(1)}}$ and x_{Ω_1} it is easy to get that

$$\|x_{\Omega_1^{(1)}}\|_2 \le \|x_{\Omega_1}\|_2 \qquad (21)$$

and

$$\|x_{\Omega_1^{(2)}}\|_2 \le \sqrt{\frac{1}{k+1}}\|x_{\Omega_1}\|_2. \qquad (22)$$

Therefore, we get that

$$\|x_{\Omega_1^{(2)}}\|_2^2 \le \sqrt{\frac{1}{k+1}}\left(\|x_{\Omega_0}\|_2 + \|x_{\Omega_1}\|_2\right)\|x_{\Omega_1^{(2)}}\|_2. \qquad (23)$$

Substituting inequalities (20) and (23) into (17), we have that

$$\begin{aligned}\|x_{\Omega_0}\|_2^2+\|x_{\Omega_1}\|_2^2&\leq\frac{1}{u^2}\big\|A(x_{\Omega_0}+x_{\Omega_1^{(1)}})\big\|_2^2+\|x_{\Omega_1^{(2)}}\|_2^2\\&\leq\big(\|x_{\Omega_0}\|_2+\|x_{\Omega_1}\|_2\big)\Bigg[\frac{w^2-u^2}{2u^2}\sum_{i=2}^{t}\|x_{\Omega_i}\|_2\\&\quad+\left(\frac{w^2}{u^2}+\sqrt{\frac{1}{k+1}}\right)\|x_{\Omega_1^{(2)}}\|_2\Bigg].\end{aligned}\tag{24}$$

For any $2\leq i\leq t$ and any element a of x_{Ω_i}, it is easy to get that $|a|^p\leq\frac{1}{k+1}\|x_{\Omega_1}\|_p^p$, so we have the inequalities:

$$\|x_{\Omega_i}\|_2^2\leq k(k+1)^{-\frac{2}{p}}\|x_{\Omega_1}\|_p^2\tag{25}$$

and

$$\|x_{\Omega_1^{(2)}}\|_2^2\leq(k+1)^{-\frac{2}{p}}\|x_{\Omega_1}\|_p^2.\tag{26}$$

Substituting inequalities (25) and (26) into (24), we derive that

$$\begin{aligned}\|x_{\Omega_0}\|_2^2+\|x_{\Omega_1}\|_2^2&\leq\big(\|x_{\Omega_0}\|_2+\|x_{\Omega_1}\|_2\big)\Bigg[\frac{w^2-u^2}{2u^2}k^{\frac{1}{2}}(k+1)^{-\frac{1}{p}}(t-1)\\&\quad+\left(\frac{w^2}{u^2}+\sqrt{\frac{1}{k+1}}\right)(k+1)^{-\frac{1}{p}}\Bigg]\|x_{\Omega_1}\|_p.\end{aligned}\tag{27}$$

Let $r=\frac{w^2}{u^2}$ and

$$B=\left[\frac{w^2-u^2}{2u^2}k^{\frac{1}{2}}(k+1)^{-\frac{1}{p}}(t-1)+\left(\frac{w^2}{u^2}+\sqrt{\frac{1}{k+1}}\right)(k+1)^{-\frac{1}{p}}\right]\|x_{\Omega_1}\|_p.$$

Then we can rewrite inequality (27) as

$$\|x_{\Omega_0}\|_2^2+\|x_{\Omega_1}\|_2^2\leq B\big(\|x_{\Omega_0}\|_2+\|x_{\Omega_1}\|_2\big),\tag{28}$$

so that

$$\left(\|x_{\Omega_0}\|_2-\frac{B}{2}\right)^2+\left(\|x_{\Omega_1}\|_2-\frac{B}{2}\right)^2\leq\frac{B^2}{2}.$$

Therefore, we get that $\|x_{\Omega_0}\|_2\leq\frac{\sqrt{2}+1}{2}B$.

According to Lemma 2, we have that

$$\|x_{\Omega_0}\|_p\leq k^{\frac{1}{p}-\frac{1}{2}}\|x_{\Omega_0}\|_2\leq k^{\frac{1}{p}-\frac{1}{2}}\frac{\sqrt{2}+1}{2}B.\tag{29}$$

Substituting the expression of B into inequality (29), we obtain that

$$\begin{aligned}\|x_{\Omega_0}\|_p &\le k^{\frac{1}{p}-\frac{1}{2}}\left(\frac{\sqrt{2}+1}{2}\right)\\ &\quad\times\left[\frac{r-1}{2}k^{\frac{1}{2}}(k+1)^{-\frac{1}{p}}(t-1)+\left(r+\sqrt{\frac{1}{k+1}}\right)(k+1)^{-\frac{1}{p}}\right]\|x_{\Omega_1}\|_p\\ &=\left(\frac{k}{k+1}\right)^{\frac{1}{p}}\left(\frac{\sqrt{2}+1}{2}\right)\left[\frac{r-1}{2}(t-1)+\left(r+\sqrt{\frac{1}{k+1}}\right)k^{-\frac{1}{2}}\right]\|x_{\Omega_1}\|_p. \end{aligned} \tag{30}$$

We notice that the sets Ω_0 and Ω_i ($i=2,\dots,t-1$) all have k elements and the set Ω_1 has $k+1$ elements such that $tk+2\le n\le(t+1)k+1$, so we get that $t\le\frac{n-2}{k}$.

According to Lemma 1, we have that $r=\frac{w^2}{u^2}\le\lambda=\frac{\lambda_{\max}(A^TA)}{\lambda_{\min^+}(A^TA)}$. Substituting the inequalities into (30), we obtain

$$\|x_{\Omega_0}\|_p\le\frac{\sqrt{2}+1}{2}\left(\frac{k}{k+1}\right)^{\frac{1}{p}}\left[\frac{(\lambda-1)(n-2-k)}{2k}+\left(\lambda+\sqrt{\frac{1}{k+1}}\right)k^{-\frac{1}{2}}\right]\|x_{\Omega_1}\|_p.$$

It is obvious that $\|x_{\Omega_1}\|_p\le\|x_{\Omega_0^c}\|_p$, and therefore we get that

$$\|x_{\Omega_0}\|_p^p\le h^*(p,A,k)\|x_{\Omega_0^c}\|_p^p,$$

where $h^*(p,A,k)$ is given in (15).

According to the definition of $h(p,A,k)$, we can get that $h(p,A,k)\le h^*(p,A,k)$. □

Theorem 3 presents a result that is very similar to the result in Theorem 1. However, it is worth pointing out that the constant $h^*(p,A,k)$ plays a central role in Theorem 3. In fact, we can treat $h^*(p,A,k)$ as an estimate of $h(p,A,k)$, where the former is calculable, and since the latter is NP-hard, $h^*(p,A,k)$ may be considered as an improvement of $h(p,A,k)$. According to Theorem 1, if we take k as the l_0-norm of the unique solution of l_0-minimization, then we can get the main contribution as soon as the inequality $h^*(p,A,k)^{\frac{1}{p}}<1$ is satisfied.

Theorem 4 *Let $A\in\mathbb{R}^{m\times n}$ be an underdetermined matrix of full rank, and denote $\Omega^*=|\operatorname{support}(A^T(AA^T)^{-1}b)|$. If every k-sparse vector x can be recovered via l_0-minimization, then x also can be recovered via l_p-minimization with $p\in(0,p^*(A,b))$, where*

$$p^*(A,b)=\max\left\{h(\Omega^*),h\left(\left\lceil\frac{n-2.5}{2}\right\rceil+1\right),h\left(\left\lceil\frac{m}{2}\right\rceil\right)\right\} \tag{31}$$

with

$$h(x)=\frac{\ln(x+1)-\ln x}{\ln[(\frac{\sqrt{2}+1}{2})(\frac{(\lambda-1)(n-3)}{2}+\lambda+\sqrt{\frac{1}{2}})]} \tag{32}$$

and

$$\lambda=\frac{\lambda_{\max}(A^TA)}{\lambda_{\min^+}(A^TA)}.$$

Proof Recalling (15), we can get the equivalence between l_0-minimization and l_p-minimization as long as $h^*(p,A,k)^{\frac{1}{p}} < 1$. However, k cannot be calculated directly, and we need to estimate k and change the inequality $h^*(p,A,k)^{\frac{1}{p}} < 1$ into a computable one through inequality technique.

Due to the integer values of $\|x\|_0$, we have that

$$\left(\lambda + \sqrt{\frac{1}{k+1}}\right)k^{-\frac{1}{2}} \leq \lambda + \sqrt{\frac{1}{2}}.$$

Notice that $\lambda > 1$, so we have

$$\frac{(\lambda-1)(n-2-k)}{2k} \leq \frac{(\lambda-1)(n-3)}{2}.$$

Furthermore, according to Corollary 2, we get that $2k+1 \leq n$, and it is easy to get that $(\lambda-1)(n-2-k) \geq 0$, so that

$$h^*(p,A,k)^{\frac{1}{p}} \leq \frac{\sqrt{2}+1}{2}\left(\frac{k}{k+1}\right)^{\frac{1}{p}}\left[\frac{(\lambda-1)(n-3)}{2} + \lambda + \sqrt{\frac{1}{2}}\right].$$

Furthermore, it is obvious that $\frac{(\lambda-1)(n-3)}{2} + \lambda + \sqrt{\frac{1}{2}} > 0$. Therefore, for $x \in (0,+\infty)$ and $p \in (0,1)$, it is easy to prove that the function

$$f(x,p) = \frac{\sqrt{2}+1}{2}\left(\frac{x}{x+1}\right)^{\frac{1}{p}}\left[\frac{(\lambda-1)(n-3)}{2} + \lambda + \sqrt{\frac{1}{2}}\right]$$

increases in x when p is fixed and also increases in p when x is fixed.

According to Corollary 2 and Remark 2, we have that $k \leq \lceil\frac{n-2.5}{2}\rceil + 1$, $k \leq \lceil\frac{m}{2}\rceil$, and $k \leq |\Omega^*|$, where $\Omega^* = |\operatorname{support}(A^T(AA^T)^{-1}b)|$, because it is obvious that $x = A^T(AA^T)^{-1}b$ is a solution of the underdetermined system $Ax = b$.

Therefore, we get that

$$f(k,p) \leq \min\left\{f\left(\left\lceil\frac{n-2.5}{2}\right\rceil + 1,p\right),f\left(\Omega^*,p\right),f\left(\left\lceil\frac{m}{2}\right\rceil,p\right)\right\}. \tag{33}$$

It is obvious that $f(k,p) < 1$ as long as one of three inequalities $f(\lceil\frac{n-2.5}{2}\rceil + 1,p) < 1$, $f(\Omega^*,p) < 1$, and $f(\lceil\frac{m}{2}\rceil,p) < 1$ is satisfied.

Furthermore, the inequality $f(x,p) < 1$ when x fixed is very easily solved, and the range of such p is

$$p < h(x) = \frac{\ln(x+1) - \ln x}{\ln[(\frac{\sqrt{2}+1}{2})(\frac{(\lambda-1)(n-3)}{2} + \lambda + \sqrt{\frac{1}{2}})]}.$$

Hence, for any $0 < p < p^* = \max\{h(\Omega^*), h(\lceil\frac{n-2.5}{2}\rceil + 1), h(\lceil\frac{m}{2}\rceil)\}$, we have that $h^*(p,A,k)^{\frac{1}{p}} \leq f(k,p) < 1$. Therefore, according to Theorem 1, every k-sparse vector $x \in \mathbb{R}^n$ can be recovered via both l_0-minimization and l_p-minimization. □

Combining Theorems 3 and 4, we have reached the major goals of this paper. The most important result in these two theorems is the analytic expression of $p^*(A,b)$, with which the specific range of p can be easily calculated.

Next, we present two examples to demonstrate the validation of Theorem 4. We consider two matrixes of different dimensions of their null spaces and get the unique solution to l_p-minimization to verify whether it is the unique solution to l_0-minimization.

Example 1 We consider an underdetermined system $Ax = b$, where

$$A = \begin{pmatrix} 1 & -1.5 & -0.7 & 0 \\ 0 & -2 & 0.5 & 1 \\ 1 & 0.5 & -1 & 1 \end{pmatrix}$$

and

$$b = \begin{pmatrix} 0.5 \\ 0 \\ 0.5 \end{pmatrix}.$$

It is obvious that the sparsest solution is $x^* = [0.5, 0, 0, 0]^T$ and $\mathrm{Ker}(A)$ is spanned by $[-10, -2, -10, 1]^T$, so the solutions of the equation $Ax = b$ can be expressed in the form

$$x = [0.5, 0, 0, 0]^T + t[-10, -2, -10, 1]^T, \quad \text{where } t \in \mathbb{R}.$$

Therefore, the p-norm of x can be expressed as

$$\|x\|_p^p = |0.5 - 10t|^p + |2t|^p + |10t|^p + |t|^p.$$

Furthermore, it is easy to get that $\lambda_{\max}(A^TA) = 7.2583$, $\lambda_{\min}(A^TA) = 1.1926$, and $\lambda = \frac{\lambda_{\max}(A^TA)}{\lambda_{\min}(A^TA)} = 6.0856$.

We can get that

$$A^T\bigl(AA^T\bigr)^{-1}b = [0.2561, -0.0488, -0.2439, 0.0244]^T,$$

and hence $h(\Omega^*) = 0.0921$ and $h(\lceil \frac{n-2.5}{2} \rceil + 1) = h(\lceil \frac{m}{2} \rceil) = 0.2862$, so $p^*(A,b) = 0.2862$.

As shown in the Figure 1, we can get the solution of l_p-minimization in different cases where $p = 0.2861, 0.2, 0.15$, and 0.1. It is obvious that $l_{0.2861}$-minimization, $l_{0.2}$-minimization, $l_{0.15}$-minimization, and $l_{0.1}$-minimization all reach their minimums at $t = 0$, which corresponds to the sparsest solution $x^* = [0.5, 0, 0, 0]^T$.

Example 2 We consider a more complex situation with $Ax = b$, where

$$A = \begin{pmatrix} 1 & 0 & 3.5 & -3 & -2.7 \\ 0 & 2 & 0 & -1.5 & 4.5 \\ 2 & 2 & -4 & -0.5 & 1.5 \end{pmatrix}$$

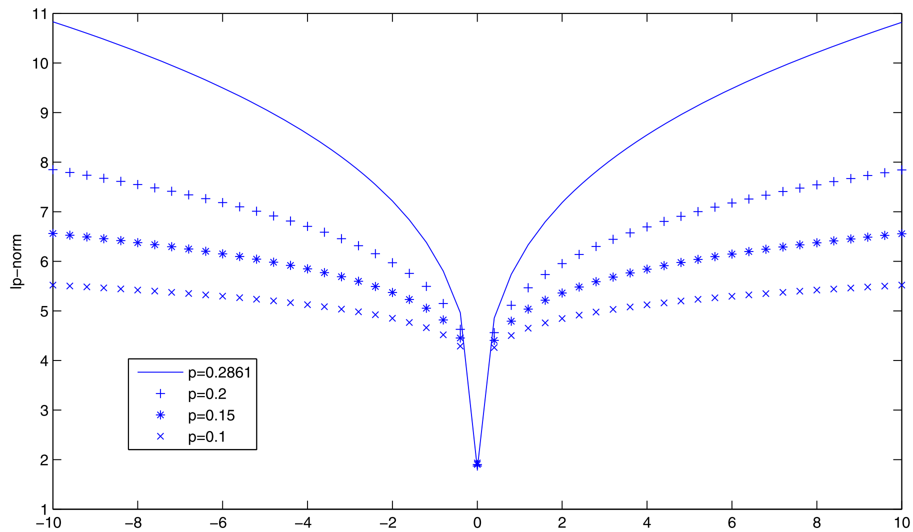

Figure 1 The *p*-norm of the solutions of the different cases in Example 1.

and

$$b = \begin{pmatrix} 0.5 \\ 0 \\ 1 \end{pmatrix}.$$

It is easy to get the sparsest solution $x^* = [0.5, 0, 0, 0, 0]^T$, and the solutions of the underdetermined system $Ax = b$ can be expressed in the parameterized form

$$x = [0.5, 0, 0, 0, 0]^T + s\left[\frac{213}{110}, -\frac{9}{4}, \frac{12}{55}, 0, 1\right]^T + t\left[\frac{17}{22}, \frac{3}{4}, \frac{7}{11}, 1, 0\right]^T, \quad \text{where } s, t \in \mathbb{R}.$$

Therefore

$$\|x\|_p^p = \left|0.5 + \frac{213}{110}s + \frac{17}{22}t\right|^p + \left|-\frac{9}{4}s + \frac{3}{4}t\right|^p + \left|\frac{12}{55}s + \frac{7}{11}t\right|^p + |s|^p + |t|^p.$$

Furthermore, we can get $\lambda = \frac{\lambda_{\max}(A^TA)}{\lambda_{\min}(A^TA)} = 4.1792$. It is easy to get that

$$A^T\left(AA^T\right)^{-1}b = [0.1903, 0.1083, -0.1188, -0.1527, -0.0990]^T.$$

Hence $h(\Omega^*) = 0.0801$, $h(\lceil\frac{n-2.5}{2}\rceil + 1) = 0.1782$, and $h(\lceil\frac{m}{2}\rceil) = 0.3046$, so we take $p^*(A, b) = 0.3046$.

From Figure 2 we can also find the solutions in different cases where $p = 0.3045, 0.3, 0.2$, and 0.1. It is obvious that the minimum is reached at $s = t = 0$, which corresponds to the sparsest solution $x^* = [0.5, 0, 0, 0, 0]^T$. The result can be seen more clearly in Figure 3.

4 Conclusions

In this paper, we have studied the equivalence between l_0-minimization and l_p-minimization. By using the l_p-null space property and a sufficient and necessary condition to recover a sparse vector via these two models, we present an analytic expression

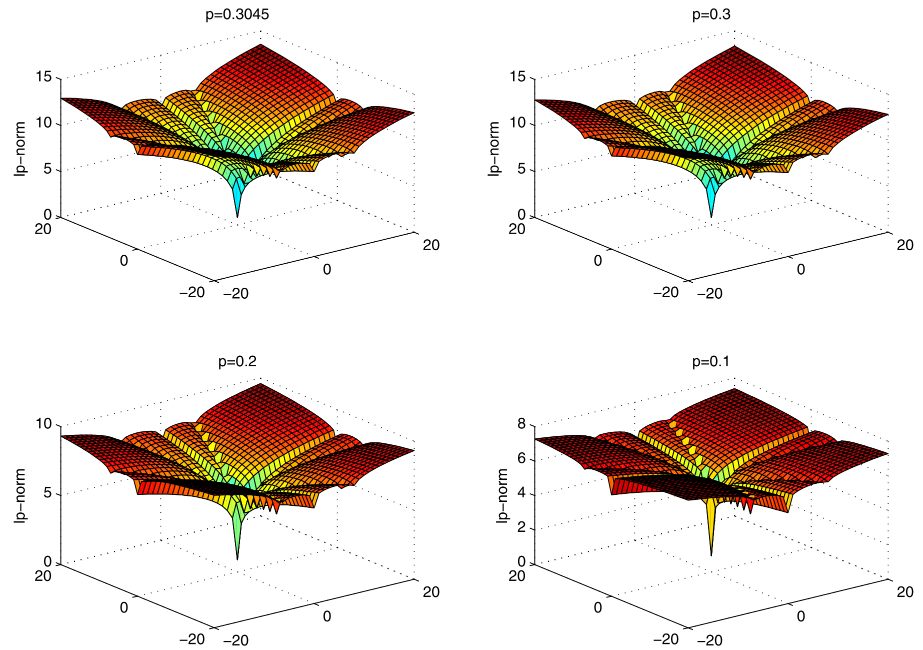

Figure 2 The *p*-norm of the solutions of the different cases in Example 2.

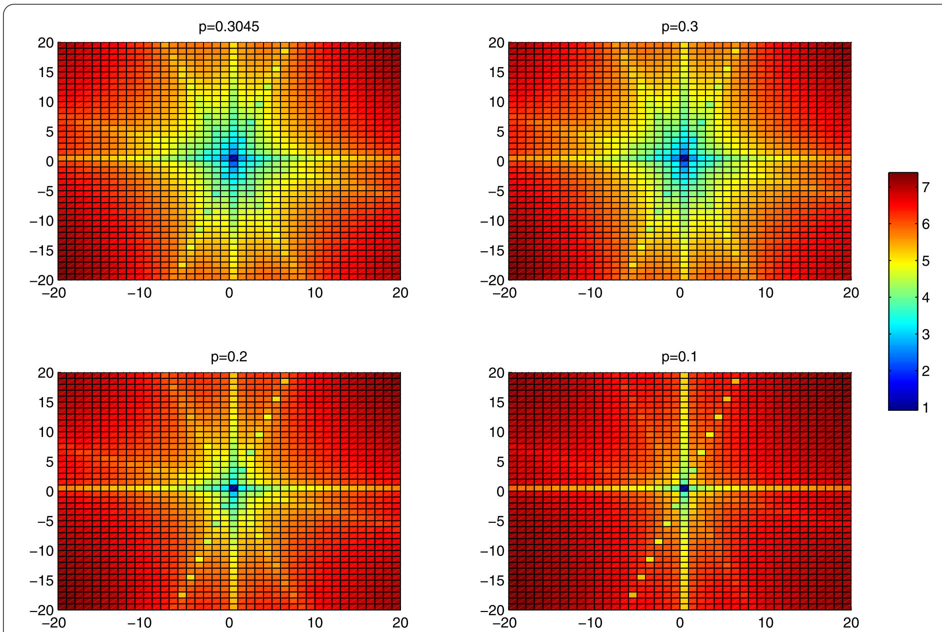

Figure 3 On *s*-*t* plane for the *p*-norm of the solutions to the different cases in Example 2.

of $p^*(A,b)$ such that l_p-minimization is equivalent to l_0-minimization. Although it is NP-hard to find the global optimal solution of l_p-minimization, a local minimizer can be done in polynomial time [24]. Chen [22] proved that $h(p,A,k) < 1$ is a necessary and sufficient condition for the global optimality of l_p-minimization. Therefore, it is confident that we can find the sparse solution with l_p-minimization with $0 < p < p^*(A, b)$ as long as we start with a good initialization.

However, in this paper, we only consider the situation where l_0-minimization has an unique solution. The uniqueness assumption is vital for us to prove the main results. However, from Lemma 1 we see that the uniqueness assumption is equivalent to a certain double-inequality condition, which looks like RIP. The evident difference between them is in that the former possesses the homogeneity rather than the latter. This implies that, unlike RIP, the uniqueness assumption is not in essential conflict with equivalence of all linear systems $\lambda Ax = \lambda x$, $\lambda \in \mathbb{R}$. Therefore, we think that the uniqueness assumption and, equivalently, the resulting double-inequality condition in Lemma 1 can replace the RIP in many cases.

Acknowledgements
The work was supported by the National Natural Science Foundations of China (11131006).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
Both authors contributed equally to this work. Both authors read and approved the final manuscript.

Publisher's Note

Received: 7 August 2017 Accepted: 12 December 2017 Published online: 21 December 2017

References
1. Olshausen, B, Field, DJ: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature **381**(6583), 607 (1996)
2. Candes, EJ, Recht, B: Exact matrix completion via convex optimization. Found. Comput. Math. **9**(6), Article ID 717 (2009)
3. Malioutov, D, Cetin, M, Willsky, AS: A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. Signal Process. **53**(8), 3010-3022 (2005)
4. Wright, J, Ganesh, A, Zhou, Z: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**(2), 210-227 (2009)
5. Candes, EJ, Tao, T: Decoding by linear programming. IEEE Trans. Inf. Theory **51**(12), 4203-4215 (2005)
6. Chen, L, Gu, Y: Local and global optimality of l_p minimization for sparse recovery. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3596-3600 (2015)
7. Peng, J, Yue, S, Li, H: *NP/CMP*, equivalence: a phenomenon hidden among sparsity models, l_0-minimization and l_p-minimization for information processing. IEEE Trans. Inf. Theory **61**(7), 4028-4033 (2015)
8. Natarajan, BK: Sparse approximate solutions to linear systems. SIAM J. Comput. **24**(2), 227-234 (2006)
9. Candes, EJ: The restricted isometry property and its implications for compressed sensing. C. R. Math. **346**(9-10), 589-592 (2008)
10. Candes, EJ, Tao, T: Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inf. Theory **52**(12), 5406-5425 (2006)
11. Foucart, S, Lai, MJ: Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$. Appl. Comput. Harmon. Anal. **26**(3), 395-407 (2009)
12. Donoho, DL, Tanner, J: Sparse nonnegative solution of underdetermined linear equations by linear programming. Proc. Natl. Acad. Sci. USA **102**(27), 9446 (2005)
13. Tropp, JA: Greed is good: algorithmic results for sparse approximation. IEEE Trans. Inf. Theory **50**(10), 2231-2242 (2004)
14. Petukhov, A: Fast implementation of orthogonal greedy algorithm for tight wavelet frames. Signal Process. **86**(3), 471-479 (2008)
15. Gao, Y, Ma, M: A new bound on the block restricted isometry constant in compressed sensing. J. Inequal. Appl. **2017**(1), Article ID 174 (2017)
16. Gribonval, R, Nielsen, M: Sparse representations in unions of bases. IEEE Trans. Inf. Theory **49**(12), 3320-3325 (2004)
17. Cai, TT, Zhang, A: Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. IEEE Trans. Inf. Theory **60**(1), 122-132 (2013)
18. Chartrand, R: Exact reconstruction of sparse signals via nonconvex minimization. IEEE Signal Process. Lett. **14**(10), 707-710 (2007)
19. Sun, Q: Recovery of sparsest signals via l_p-minimization. Appl. Comput. Harmon. Anal. **32**(3), 329-341 (2010)
20. Daubechies, I, Devore, R, Fornasier, M: Iteratively reweighted least squares minimization for sparse recovery. Commun. Pure Appl. Math. **63**(1), 1-38 (2008)
21. Gao, Y, Peng, J, Yue, S, Zhao, Y: On the null space property of l_q-minimization for $0 < q \leq 1$ in compressed sensing. J. Funct. Spaces **2015**, Article ID 579853 (2015)

22. Chen, L, Gu, Y: On the null space constant for l_p-minimization. IEEE Signal Process. Lett. **22**(10), 1600-1603 (2015)
23. Foucart, S, Rauhut, H: A Mathematical Introduction to Compressive Sensing. Birkhäuser, Basel (2013)
24. Ge, D, Jiang, X, Ye, Y: A note on the complexity of l_p-minimization. Math. Program. **129**, 285-299 (2011)