On the Null Space Constant for ℓ_p Minimization

Laming Chen, *Student Member, IEEE*, and Yuantao Gu, *Member, IEEE*

***Abstract*—The literature on sparse recovery often adopts the ℓ_p "norm" ($p \in [0,1]$) as the penalty to induce sparsity of the signal satisfying an underdetermined linear system. The performance of the corresponding ℓ_p minimization problem can be characterized by its null space constant. In spite of the NP-hardness of computing the constant, its properties can still help in illustrating the performance of ℓ_p minimization. In this letter, we show the strict increase of the null space constant in the sparsity level k and its continuity in the exponent p. We also indicate that the constant is strictly increasing in p with probability 1 when the sensing matrix $\mathbf{A}$ is randomly generated. Finally, we show how these properties can help in demonstrating the performance of ℓ_p minimization, mainly in the relationship between the the exponent p and the sparsity level k.**

***Index Terms*—Continuity, ℓ_p minimization , monotonicity, null space constant, sparse recovery.**

I. Introduction

An IMPORTANT problem that often arises in signal processing, machine learning, and statistics is sparse recovery [1]–[3]. It is in general formulated in the standard form

$$\arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{1}$$

where the sensing matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ has more columns than rows and the ℓ_0 "norm" $\|\mathbf{x}\|_0$ denotes the number of nonzero entries of the vector $\mathbf{x}$. The combinatorial optimization (1) is NP-hard and therefore cannot be solved efficiently [4]. A standard method to solve this problem is by relaxing the non-convex discontinuous ℓ_0 "norm" to the convex ℓ_1 norm [5], i.e.,

$$\arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}. \tag{2}$$

It is theoretically proved that under some certain conditions [5], [6], the optimum solution of (2) is identical to that of (1).

Some works try to bridge the gap between ℓ_0 "norm" and ℓ_1 norm by non-convex but continuous ℓ_p "norm" ($0 < p < 1$) [7]–[10], and consider the ℓ_p minimization problem

$$\arg\min_{\mathbf{x}} \|\mathbf{x}\|_p^p \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \tag{3}$$

where $\|\mathbf{x}\|_p^p = \sum_{i=1}^N |x_i|^p$. Though finding the global optimal solution of ℓ_p minimization is still NP-hard, computing a local minimizer can be done in polynomial time [11]. The global optimality of (3) has been studied and various conditions have been derived, for example, those based on restricted isometry property [7]–[9], [12] and null space property [10], [13]. Among them, a necessary and sufficient condition is based on the null space property and its constant [10], [13], [14].

Definition 1: For any $0 \le p \le 1$, define null space constant $\gamma(\ell_p, \mathbf{A}, k)$ as the smallest quantity such that

$$\sum_{i\in S} |z_i|^p \le \gamma(\ell_p, \mathbf{A}, k) \sum_{i\notin S} |z_i|^p \tag{4}$$

holds for any set $S \subset \{1,2,\dots,N\}$ with $\#S \le k$ and for any vector $\mathbf{z} \in \mathcal{N}(\mathbf{A})$ which denotes the null space of $\mathbf{A}$.

It has been shown that for any $p \in [0,1]$, $\gamma(\ell_p, \mathbf{A}, k) < 1$ is a necessary and sufficient condition such that for any k-sparse $\mathbf{x}^*$ and $\mathbf{y} = \mathbf{A}\mathbf{x}^*$, $\mathbf{x}^*$ is the unique solution of ℓ_p minimization [10]. Therefore, $\gamma(\ell_p, \mathbf{A}, k)$ is a tight quantity in indicating the performance of ℓ_p minimization ($0 \le p \le 1$) in sparse recovery. However, it has been shown that calculating $\gamma(\ell_p, \mathbf{A}, k)$ is in general NP-hard [15], which makes it difficult to check whether the condition is satisfied or violated. Despite this, properties of $\gamma(\ell_p, \mathbf{A}, k)$ are of tremendous help in illustrating the performance of ℓ_p minimization, e.g., non-decrease of $\gamma(\ell_p, \mathbf{A}, k)$ in $p \in [0,1]$ shows that if ℓ_p minimization guarantees successful recovery of all k-sparse signal and $0 \le q \le p$, then ℓ_q minimization also does [10].

In this letter, we give some new properties of the null space constant $\gamma(\ell_p, \mathbf{A}, k)$. Specifically, we prove that $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in k and is continuous in p. For random sensing matrix $\mathbf{A}$, the non-decrease of $\gamma(\ell_p, \mathbf{A}, k)$ in p can be improved to strict increase with probability 1. Based on them, the performance of ℓ_p minimization can be intuitively demonstrated and understood.

II. Main Contribution

This section introduces some properties of null space constant $\gamma(\ell_p, \mathbf{A}, k)(0 \le p \le 1)$. We begin with a lemma about $\gamma(\ell_p, \mathbf{A}, k)$ which will play a central role in the theoretical analysis. The spark of a matrix $\mathbf{A}$, denoted as $\mathrm{Spark}(\mathbf{A})$[16], is the smallest number of columns from $\mathbf{A}$ that are linearly dependent.

Lemma 1: Suppose $\mathrm{Spark}(\mathbf{A}) = L+1$. For $p \in [0,1]$,

1) $\gamma(\ell_p, \mathbf{A}, k)$ is finite if and only if $k \le L$;
2) For $k \le L$, there exist $S' \subset \{1,2,\dots,N\}$ with $\#S' \le k$ and $\mathbf{z}' \in \mathcal{N}(\mathbf{A})\backslash\{\mathbf{0}\}$ such that

$$\sum_{i\in S'} |z_i'|^p = \gamma(\ell_p, \mathbf{A}, k) \sum_{i\notin S'} |z_i'|^p \tag{5}$$

Proof: See Section III-A. ■

First, we show the strict increase of $\gamma(\ell_p, \mathbf{A}, k)$ in k.

Theorem 1: Suppose $\mathrm{Spark}(\mathbf{A}) = L+1$. Then for $p \in [0,1]$, $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in k when $k \le L$.

Proof: See Section III-B. ■

Manuscript received December 15, 2014; revised March 02, 2015; accepted March 18, 2015. Date of publication March 23, 2015; date of current version March 30, 2015. This work was supported by National 973 Program of China under Grant 2013CB329203, and by the National Natural Science Foundation of China under Grant 61371137. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gitta Kutyniok.

The authors are with Department of Electronic Engineering and Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China (e-mail:chen-lm06@mails.tsinghua.edu.cn; gyt@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2015.2416003

Remark 1: For any $p \in [0,1]$, we can define a set $\mathcal{K}_p(\mathbf{A})$ of all positive integers k that every k–sparse $\mathbf{x}^*$ can be recovered as the unique solution of ℓ_p minimization (3) with $\mathbf{y} = \mathbf{A}\mathbf{x}^*$. According to Theorem 1, $\mathcal{K}_p(\mathbf{A})$ contains successive integers starting from 1 to some integer $k_p^*(\mathbf{A})$ and is possibly empty.

Remark 2: If $\text{Spark}(\mathbf{A}) = L+1$, then $k_0^*(\mathbf{A}) = \lfloor L/2 \rfloor$ [16]. Therefore, if $L \geq 2$, $k_0^*(\mathbf{A}) \geq 1$.

Remark 3: For $\mathbf{A}$ with identical column norms, if $\text{Spark}(\mathbf{A}) = L+1$ and $L \geq 2$, then $k_1^*(\mathbf{A}) \geq 1$. To show this, we only need to prove that $\gamma(\ell_1, \mathbf{A}, 1) < 1$. First, for any $1 \leq i \leq N$ and $\mathbf{z} \in \mathcal{N}(\mathbf{A})\backslash\{\mathbf{0}\}$, since $\mathbf{A}\mathbf{z} = 0$, $z_i\mathbf{a}_i = -\sum_{j\neq i} z_j\mathbf{a}_j$ where $\mathbf{a}_i$ is the ith column of $\mathbf{A}$. Since

$$|z_i| \cdot \|\mathbf{a}_i\|_2 = \|z_i\mathbf{a}_i\|_2 = \|\sum_{j\neq i} z_j\mathbf{a}_j\|_2 \leq \sum_{j\neq i} |z_j| \cdot \|\mathbf{a}_j\|_2$$

with equality holds only when $z_j\mathbf{a}_j (j \neq i)$ are all on the same ray, which cannot be true since $\text{Spark}(\mathbf{A}) = L+1 \geq 3$. Since $\mathbf{A}$ has identical column norms, $|z_i| < \sum_{j\neq i} |z_j|$ holds, which leads to $\gamma(\ell_1, \mathbf{A}, 1) < 1$ because of Lemma 1.2).

Now we turn to the properties of $\gamma(\ell_p, \mathbf{A}, k)$ as a function of p. The following result reveals the continuity of $\gamma(\ell_p, \mathbf{A}, k)$ in p.

Theorem 2: Suppose $\text{Spark}(\mathbf{A}) = L+1$. Then for $k \leq L$, $\gamma(\ell_p, \mathbf{A}, k)$ is a continuous function in $p \in [0,1]$.

Proof: See Section III-C. ■

Remark 4: Some works have discussed the equivalence of ℓ_0 and ℓ_p minimizations. In [17], it is shown that the sufficient condition for the equivalence of these two minimization problems approaches the necessary and sufficient condition for the uniqueness of solutions of ℓ_0 minimization. In [7], it is shown that for any k-sparse $\mathbf{x}^*$ and $\mathbf{y} = \mathbf{A}\mathbf{x}^*$, if $\delta_{2k+1} < 1$, then there is $p > 0$ such that $\mathbf{x}^*$ is the unique solution of ℓ_p minimization. This result is improved to $\delta_{2k} < 1$ which is optimal since it is exactly the necessary and sufficient condition for $\mathbf{x}^*$ being the unique solution of ℓ_0 minimization [12]. [18] shows the equivalence of the ℓ_0- and the ℓ_p-norm minimization problem for sufficiently small p. According to Theorem 2, we can also justify this result: For any k-sparse $\mathbf{x}^*$ and $\mathbf{y} = \mathbf{A}\mathbf{x}^*$, if $\gamma(\ell_0, \mathbf{A}, k) < 1$, then there is $p > 0$ such that $\gamma(\ell_p, \mathbf{A}, k) < 1$ and $\mathbf{x}^*$ is the unique solution of ℓ_p minimization.

Remark 5: In [10], the author defines a set $\mathcal{P}_k(\mathbf{A})$ of reconstruction exponents, that is the set of all exponents $0 < p \leq 1$ for which every k-sparse $\mathbf{x}^*$ is recovered as the unique solution of ℓ_p minimization with $\mathbf{y} = \mathbf{A}\mathbf{x}^*$. It is shown that $\mathcal{P}_k(\mathbf{A})$ is a (possibly empty) open interval $(0, p_k^*(\mathbf{A}))$[10]. This result can be easily shown by Theorem 2. Since $\gamma(\ell_p, \mathbf{A}, k)$ is a non-decreasing [13] continuous function in $p \in [0,1]$, the inverse image of the open interval $(-\infty, 1)$ is also an open interval of $[0,1]$. Therefore, the requirement that $\gamma(\ell_p, \mathbf{A}, k) < 1$ is equivalent to $p \in [0, p_k^*(\mathbf{A}))$.

Remark 6: For any $\mathbf{A}$, we can plot $k_p^*(\mathbf{A})$ as a function of p, as shown in Fig. 1. For concision, we omit the argument $\mathbf{A}$ in the figure. It is obvious that $k_p^*(\mathbf{A})$ is a step function decreasing from $k_0^*(\mathbf{A})$ to $k_1^*(\mathbf{A})$. Three facts needs to be pointed out. First, $k_p^*(\mathbf{A})$ is right-continuous, which is an easy consequence of Theorem 2. Second, the points (p_0, k_0) corresponding to the hollow circles in Fig. 1 satisfy $\gamma(\ell_{p_0}, \mathbf{A}, k_0) = 1$. Third, for the p-axis p_0 of the points of discontinuity, the one-sided limits satisfy $\lim_{p\to p_0^-} k_p^*(\mathbf{A}) - \lim_{p\to p_0^+} k_p^*(\mathbf{A}) = 1$. This can be proved by Theorem 1 that if $\gamma(\ell_{p_0}, \mathbf{A}, k_0) = 1$, then $\gamma(\ell_{p_0}, \mathbf{A}, k_0 - 1) < 1$.

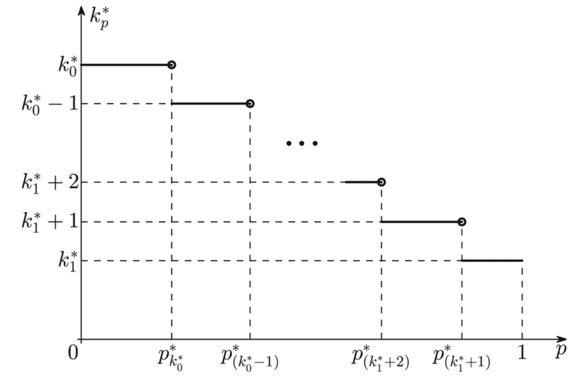

Fig. 1. The figure shows $k_p^*(\mathbf{A})$ as a function of p, where the argument $\mathbf{A}$ is omitted for concision.

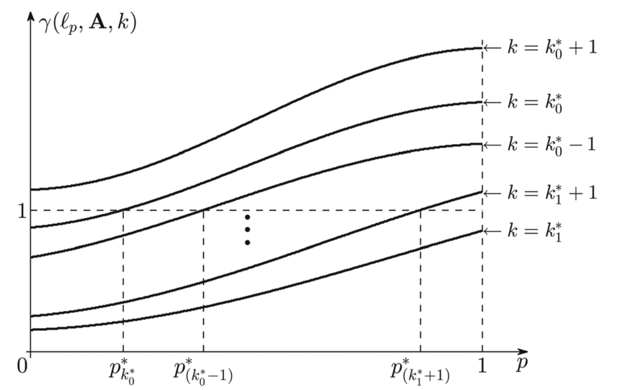

Fig. 2. This figure shows a diagrammatic sketch of $\gamma(\ell_p, \mathbf{A}, k)$ as a function of p for different k when $\mathbf{A}$ is a random matrix.

Finally, we introduce an important property of $\gamma(\ell_p, \mathbf{A}, k)$ as a function of p with regard to random matrix $\mathbf{A}$.

Theorem 3: Suppose the entries of $\mathbf{A} \in \mathbb{R}^{M\times N}$ are i.i.d. and satisfy a continuous probability distribution. Then for $k \leq M$, $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in $p \in [0,1]$ with probability one.

Proof: See Section III-D. ■

Remark 7: It needs to be noted that there exists $\mathbf{A}$ such that $\gamma(\ell_p, \mathbf{A}, k)$ is a constant number for all $p \in [0,1]$. For example, for

$$\mathbf{A} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \tag{6}$$

$\text{Spark}(\mathbf{A}) = 2$. Since $\mathcal{N}(\mathbf{A}) = \text{span}([1, -1]^T)$, it is easy to check that for all $p \in [0,1]$, $\gamma(\ell_p, \mathbf{A}, 1) = 1$.

Remark 8: To sum up, we can schematically show $\gamma(\ell_p, \mathbf{A}, k)$ as a function of p for different k in Fig. 2. According to Theorem 1, these curves are strictly in order without intersections. Theorem 2 reveals that $\gamma(\ell_p, \mathbf{A}, k)$ is continuous in p. For a random matrix $\mathbf{A}$ with i.i.d. entries satisfying a continuous probability distribution, $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in p with probability 1 by Theorem 3. According to the definition of $k_p^*(\mathbf{A})$, the curves intersecting $\gamma(\ell_p, \mathbf{A}, k) = 1 (0 \leq p \leq 1)$ are those with $k_1^*(\mathbf{A}) + 1 \leq k \leq k_0^*(\mathbf{A})$. According to the definition of $p_k^*(\mathbf{A})$, the p-axis of these intersections are $p_{k_0^*}^*$, $p_{k_0^*-1}^*$, $\ldots$, $p_{k_1^*+1}^*$ from left to right. Therefore, it is easy to derive Fig. 1 based on Fig. 2 when $\mathbf{A}$ is a random matrix.

III. Proofs

A. Proof of Lemma 1

Proof: 1) Since $\mathrm{Spark}(\mathbf{A}) = L+1$, $\mathcal{N}(\mathbf{A})$ contains an $(L+1)$-sparse signal, and it is easy to show that for any $k \geq L+1$, $\gamma(\ell_p, \mathbf{A}, k) = +\infty$ according to Definition 1. Next we prove that for $k \leq L$, $\gamma(\ell_p, \mathbf{A}, k)$ is finite. Define

$$\theta(p, \mathbf{z}, S) = \frac{\sum_{i \in S} |z_i|^p}{\sum_{i \notin S} |z_i|^p} \tag{7}$$

and $\mathcal{N}_1(\mathbf{A}) = \mathcal{N}(\mathbf{A}) \cap \{\mathbf{z} : \|\mathbf{z}\|_2 = 1\}$ which is a compact set. Then it is easy to see that the definition of null space constant is equivalent to

$$\gamma(\ell_p, \mathbf{A}, k) = \max_{\#S \leq k} \sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(p, \mathbf{z}, S). \tag{8}$$

If $\gamma(\ell_p, \mathbf{A}, k)$ is not finite, then there exists S' with $\#S' \leq k$ such that $\sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(p, \mathbf{z}, S')$ is not finite. Therefore, for any $n \in \mathbb{N}^+$, there exists $\mathbf{z}^{(n)} \in \mathcal{N}_1(\mathbf{A})$ such that

$$\theta(p, \mathbf{z}^{(n)}, S') \geq n. \tag{9}$$

If $p = 0$, since $\mathbf{z}^{(n)}$ is at least $(L+1)$-sparse, it is easy to see that $\theta(0, \mathbf{z}^{(n)}, S') \leq k$ holds for any $n \in \mathbb{N}^+$. This contradicts (9) when $n > k$. If $p \in (0, 1]$, according to Lemma 4.5 in [10], $\|\mathbf{z}^{(n)}\|_p \leq N^{\frac{1}{p} - \frac{1}{2}} \|\mathbf{z}^{(n)}\|_2 = N^{\frac{1}{p} - \frac{1}{2}}$, and (9) implies

$$\sum_{i \notin S'} |z_i^{(n)}|^p \leq \frac{N^{1 - \frac{p}{2}}}{n+1}. \tag{10}$$

Due to the compactness of $\mathcal{N}_1(\mathbf{A})$, the sequence $\{\mathbf{z}^{(n)}\}_n$ has a convergent subsequence $\{\mathbf{z}^{(n_m)}\}_m$, and its limit $\mathbf{z}'$ also lies in $\mathcal{N}_1(\mathbf{A})$. Then (10) implies $z'_i = 0$ for $i \notin S'$, i.e., $\mathcal{N}_1(\mathbf{A})$ contains a k-sparse element $\mathbf{z}'$. This contradicts the assumption that $\mathrm{Spark}(\mathbf{A}) = L+1 > k$.

2) If $p = 0$, for any S with $\#S \leq k$ and any $\mathbf{z} \in \mathcal{N}(\mathbf{A}) \backslash \{\mathbf{0}\}$, it holds that

$$\theta(0, \mathbf{z}, S) \leq \frac{k}{L+1-k}. \tag{11}$$

On the other hand, since $\mathrm{Spark}(\mathbf{A}) = L+1$, $\mathcal{N}(\mathbf{A})$ contains an $(L+1)$-sparse signal $\mathbf{z}'$ with T as its support set. For any $S' \subset T$ with $\#S' = k$, $\theta(0, \mathbf{z}', S') = k/(L+1-k)$, and therefore (5) holds.

If $p \in (0, 1]$, recalling the equivalent definition (8), there exists S' with $\#S' \leq k$ such that

$$\gamma(\ell_p, \mathbf{A}, k) = \sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(p, \mathbf{z}, S'). \tag{12}$$

Since $\mathcal{N}_1(\mathbf{A})$ is compact and the function $\theta(p, \mathbf{z}, S')$ is continuous in $\mathbf{z}$ on $\mathcal{N}_1(\mathbf{A})$, it is easy to show that there exists $\mathbf{z}' \in \mathcal{N}_1(\mathbf{A})$ such that $\gamma(\ell_p, \mathbf{A}, k) = \theta(p, \mathbf{z}', S')$. ∎

B. Proof of Theorem 1

Proof: We prove that when $p \in [0, 1]$ and $2 \leq k \leq L$,

$$\gamma(\ell_p, \mathbf{A}, k-1) < \gamma(\ell_p, \mathbf{A}, k). \tag{13}$$

According to Lemma 1.2), there exist S' with $\#S' \leq k-1$ and $\mathbf{z}' \in \mathcal{N}_1(\mathbf{A})$ such that

$$\gamma(\ell_p, \mathbf{A}, k-1) = \theta(p, \mathbf{z}', S'). \tag{14}$$

Since $\mathbf{z}'$ is at least $(L+1)$-sparse, there exists an index $s' \in \{1, 2, \ldots, N\} \backslash S'$ such that $z'_{s'} \neq 0$. Let $S'' = S' \cup \{s'\}$, then

$$\sum_{i \in S'} |z'_i|^p < \sum_{i \in S''} |z'_i|^p, \sum_{i \notin S'} |z'_i|^p > \sum_{i \notin S''} |z'_i|^p > 0 \tag{15}$$

and hence

$$\theta(p, \mathbf{z}', S') < \theta(p, \mathbf{z}', S''). \tag{16}$$

Recalling (14) and the equivalent definition (8), we can get (13) and complete the proof. ∎

C. Proof of Theorem 2

Proof: According to Theorem 5 in [13], $\gamma(\ell_p, \mathbf{A}, k)$ is nondecreasing in $p \in [0, 1]$ and therefore can only have jump discontinuities. We show this is impossible by two steps.

First, for any $p \in (0, 1]$, we prove the one-sided limit from the negative direction satisfies

$$L^- := \lim_{q \to p^-} \gamma(\ell_q, \mathbf{A}, k) = \gamma(\ell_p, \mathbf{A}, k). \tag{17}$$

According to Lemma 1.2), there exist S' with $\#S' \leq k$ and $\mathbf{z}' \in \mathcal{N}_1(\mathbf{A})$ satisfying

$$\gamma(\ell_p, \mathbf{A}, k) = \theta(p, \mathbf{z}', S'). \tag{18}$$

According to the definition of $\theta(p, \mathbf{z}, S)$, it is easy to show that

$$\lim_{q \to p^-} \theta(q, \mathbf{z}', S') = \theta(p, \mathbf{z}', S'), \tag{19}$$

and then (17) holds obviously.

Second, for any $p \in [0, 1)$, we prove the one-sided limit from the positive direction satisfies

$$L^+ := \lim_{q \to p^+} \gamma(\ell_q, \mathbf{A}, k) = \gamma(\ell_p, \mathbf{A}, k). \tag{20}$$

Since $p < 1$, there exists $N_0 \in \mathbb{N}^+$ such that $p + N_0^{-1} \leq 1$. Then for $n \geq N_0$, Lemma 1.2) reveals that there exist $S^{(n)}$ with $\#S^{(n)} \leq k$ and $\mathbf{z}^{(n)} \in \mathcal{N}_1(\mathbf{A})$ such that

$$\gamma(\ell_{p+n^{-1}}, \mathbf{A}, k) = \theta(p + n^{-1}, \mathbf{z}^{(n)}, S^{(n)}). \tag{21}$$

Since there are only finite different S satisfying $\#S \leq k$, there exists S' with $\#S' \leq k$ such that an infinite subsequence of $\{\mathbf{z}^{(n)}\}_n$ is associated with S'. Due to the compactness of $\mathcal{N}_1(\mathbf{A})$, this subsequence has a convergent subsequence $\{\mathbf{z}^{(n_m)}\}_m$, and its limit $\mathbf{z}'$ also lies in $\mathcal{N}_1(\mathbf{A})$. According to the definition of $\theta(p, \mathbf{z}, S)$ and (21),

$$\theta(p, \mathbf{z}', S') = \lim_{m \to +\infty} \theta(p + n_m^{-1}, \mathbf{z}^{(n_m)}, S') = L^+, \tag{22}$$

and consequently $\gamma(\ell_p, \mathbf{A}, k) \geq L^+$. Since $\gamma(\ell_p, \mathbf{A}, k)$ is nondecreasing in p, $\gamma(\ell_p, \mathbf{A}, k) \leq L^+$ and (20) is proved. ∎

D. Proof of Theorem 3

Proof: First, we show that $\mathrm{Spark}(\mathbf{A}) = M+1$ with probability 1. Let $\mathcal{M}(M)$ denote the M^2-dimensional vector space of $M \times M$ real matrices. For any $0 \leq k \leq M$, let $\mathcal{M}_k(M)$ denote the subset of $\mathcal{M}(M)$ consisting of matrices of rank k. It can be proved that $\mathcal{M}_k(M)$ is an embedded submanifold of dimension $k(2M-k)$ in $\mathcal{M}(M)$[19]. Consequently, for $M \times M$ matrices with i.i.d. entries drawn from a continuous distribution, the M^2-dimensional volume of the set of singular matrices $\bigcup_{k=0}^{M-1} \mathcal{M}_k(M)$ is zero. In other words, any M, or fewer, random vectors in $\mathbb{R}^M$ with i.i.d. entries drawn from a continuous distribution are linearly independent with probability 1. On the other hand, more than M vectors in $\mathbb{R}^M$ are always linearly dependent. Therefore, $\text{Spark}(\mathbf{A}) = M + 1$ with probability 1.

Next, with the equivalent definition (8), we prove that for $k \le M$ and $0 \le p < q \le 1$,

$$\max_{\#S \le k} \sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(p, \mathbf{z}, S) < \max_{\#S \le k} \sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(q, \mathbf{z}, S) \tag{23}$$

holds with probability 1. According to Lemma 1.2), there exist S' with $\#S' \le k$ and $\mathbf{z}' \in \mathcal{N}_1(\mathbf{A})$ such that

$$\theta(p, \mathbf{z}', S') = \max_{\#S \le k} \sup_{\mathbf{z} \in \mathcal{N}_1(\mathbf{A})} \theta(p, \mathbf{z}, S). \tag{24}$$

Suppose $\mathbf{z}'$ has N_* nonzero entries with T as its support set, then $N_* \ge M + 1$ with probability 1. It is obvious that $S' \subset T$, and for any $i \in S'$ and any $l \in T \backslash S'$, $|z_i'| \ge |z_l'| > 0$. Since $p < q$, $|z_i'|^{q-p} \ge |z_l'|^{q-p}$ and therefore

$$|z_i'|^q |z_l'|^p \ge |z_i'|^p |z_l'|^q. \tag{25}$$

Summing (25) with i in S' and l in $T \backslash S'$, we can obtain

$$\sum_{i \in S'} |z_i'|^q \sum_{l \in T \backslash S'} |z_l'|^p \ge \sum_{i \in S'} |z_i'|^p \sum_{l \in T \backslash S'} |z_l'|^q \tag{26}$$

which is equivalent to

$$\theta(p, \mathbf{z}', S') \le \theta(q, \mathbf{z}', S'). \tag{27}$$

Since $p < q$, it is easy to check that the equality in (27) holds only when $|z_i'| = |z_l'|$ for all $i \in S'$ and all $l \in T \backslash S'$, i.e., the nonzero entries of $\mathbf{z}'$ have the same magnitude. We prove that $\mathcal{N}_1(\mathbf{A})$ contains such $\mathbf{z}'$ with probability 0, which together with (24) imply that

$$\gamma(\ell_p, \mathbf{A}, k) = \theta(p, \mathbf{z}', S') < \theta(q, \mathbf{z}', S') \le \gamma(\ell_q, \mathbf{A}, k) \tag{28}$$

holds with probability 1.

To this end, let $\mathcal{M}(M, N)$ denote the MN-dimensional vector space of $M \times N$ real matrices. For fixed $\mathbf{z} \in \mathbb{R}^N$ with $\|\mathbf{z}\|_2 = 1$, it can be easily shown that the subset

$$\mathcal{M}_{\mathbf{z}}(M, N) = \{\mathbf{A} \in \mathcal{M}(M, N) : \mathbf{A}\mathbf{z} = \mathbf{0}\} \tag{29}$$

is an $M(N-1)$-dimensional subspace in $\mathcal{M}(M, N)$. Therefore, for $\mathbf{A} \in \mathcal{M}(M, N)$ with i.i.d. entries drawn from a continuous probability distribution, $\mathcal{N}_1(\mathbf{A})$ contains $\mathbf{z}$ with probability 0. In $\{\mathbf{z} \in \mathbb{R}^N : \|\mathbf{z}\|_2 = 1\}$, the number of vectors whose nonzero entries have the same magnitude is

$$\sum_{i=1}^{N} \binom{N}{i} 2^i = 3^N - 1 \tag{30}$$

which is a finite number. Therefore, with probability 0, $\mathcal{N}_1(\mathbf{A})$ contains a vector $\mathbf{z}'$ which makes the equality in (27) hold. That is, $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in $p \in [0, 1]$ with probability 1. ■

IV. Conclusion

In characterizing the performance of ℓ_p minimization in sparse recovery, null space constant $\gamma(\ell_p, \mathbf{A}, k)$ can be served as a necessary and sufficient condition for the perfect recovery of all k-sparse signals. This letter derives some basic properties of $\gamma(\ell_p, \mathbf{A}, k)$ in k and p. In particular, we show that $\gamma(\ell_p, \mathbf{A}, k)$ is strictly increasing in k and is continuous in p, meanwhile for random $\mathbf{A}$, the constant is strictly increasing in p with probability 1. Possible future works include the properties of $\gamma(\ell_p, \mathbf{A}, k)$ in $\mathbf{A}$, for example, the requirement of number of measurements M to guarantee $\gamma(\ell_p, \mathbf{A}, k) < 1$ with high probability when $\mathbf{A}$ is randomly generated.

References

[1] R. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
[2] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[3] N. Meinshausen and B. Yu, "Lasso-type recovery of sparse representations for high-dimensional data," *Ann. Statist.*, vol. 37, no. 1, pp. 246–270, Feb. 2009.
[4] B. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, Apr. 1995.
[5] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
[6] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[7] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.
[8] R. Saab, R. Chartrand, and O. Yilmaz, "Stable sparse approximations via nonconvex optimization," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Mar. 2008, pp. 3885–3888.
[9] S. Foucart and M. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \le 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, May 2009.
[10] S. Foucart, *Notes on compressed sensing for Math 394*. Berlin, Germany: Springer, 2009.
[11] D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of L_p minimiation," *Math. Program.*, vol. 129, no. 2, pp. 285–299, Jun. 2011.
[12] Q. Sun, "Recovery of sparsest signals via ℓ^q-minimization," *Appl. Comput. Harmon. Anal.*, vol. 32, no. 3, pp. 329–341, May 2012.
[13] R. Gribonval and M. Nielsen, "Highly sparse representations from dictionaries are unique and independent of the sparseness measure," *Appl. Comput. Harmon. Anal.*, vol. 22, no. 3, pp. 335–355, May 2007.
[14] L. Chen and Y. Gu, "The convergence guarantees of a non-convex approach for sparse recovery," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3754–3767, Aug. 2014.
[15] A. Tillmann and M. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1248–1259, Feb. 2014.
[16] D. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," in *Proc. Nat. Acad. Sci.*, Mar. 2003, vol. 100, no. 5, pp. 2197–2202.
[17] D. Malioutov, M. Cetin, and A. Willsky, "Optimal sparse representations in general overcomplete bases," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2004, vol. 2, pp. 793–796.
[18] G. Fung and O. Mangasarian, "Equivalence of minimal ℓ_0- and ℓ_p-norm solutions of linear equalities, inequalities and linear programs for sufficiently small p," *J. Optim. Theory Applicat.*, vol. 151, no. 1, pp. 1–10, 2011.
[19] J. Lee, *Introduction to Smooth Manifolds*. Berlin, Germany: Springer, 2012, vol. 218.