J. Korean Math. Soc. **56** (2019), No. 3, pp. 689–701
https://doi.org/10.4134/JKMS.j180321
pISSN: 0304-9914 / eISSN: 2234-3008

STABLE AND ROBUST ℓ_p-CONSTRAINED COMPRESSIVE SENSING RECOVERY VIA ROBUST WIDTH PROPERTY

Jun Yu and Zhiyong Zhou

Abstract. We study the recovery results of ℓ_p-constrained compressive sensing (CS) with $p \geq 1$ via robust width property and determine conditions on the number of measurements for standard Gaussian matrices under which the property holds with high probability. Our paper extends the existing results in Cahill and Mixon [2] from ℓ_2-constrained CS to ℓ_p-constrained case with $p \geq 1$ and complements the recovery analysis for robust CS with ℓ_p loss function.

1. Introduction

There has been enormous research interest in the field of compressive sensing (CS) since its advent in [4, 8]. CS aims to recover an unknown signal with underdetermined linear measurements (see [11, 12] for a comprehensive view). Specifically, let $x^\natural$ be some unknown signal, which lies in a finite-dimensional Hilbert space $\mathcal{H}$, and let $\Phi : \mathcal{H} \to \mathbb{F}^m$ be the underdetermined measurement linear operator, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. We denote the usual ℓ_p norm in $\mathbb{F}^m$ as $\|u\|_{\ell_p} = (\sum_{i=1}^m |u_i|^p)^{1/p}$ for $1 \leq p < \infty$ and $\|u\|_\infty = \max_{1\leq i\leq m} |u_i|$. Then, given the measurements $y = \Phi x^\natural + e$ for some unknown error $e \in \mathbb{F}^m$ with $\|e\|_{\ell_2} \leq \varepsilon$ and $x^\natural$ is close to a particular subset $\mathcal{A} \subseteq \mathcal{H}$ that consists of members with some specific structure, there is a norm $\|\cdot\|_\sharp$ corresponding to $\mathcal{A}$ such that

$$\triangle_{\sharp,2}(y) := \arg\min_x \|x\|_\sharp \text{ subject to } \|\Phi x - y\|_{\ell_2} \leq \varepsilon \tag{1}$$

is a good estimate of the original signal $x^\natural$, provided that the measurement matrix Φ satisfies certain properties (e.g., robust null space property [6,7,15,19], restricted isometry property [3,4]). However, these conditions are not known to be necessary. Recently, for a broad class of triples $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$, [2] completely characterized Φ for which

$$\|\triangle_{\sharp,2}(y) - x^\natural\|_2 \leq C_0\|x^\natural - a\|_\sharp + C_1\varepsilon \ \forall\ x^\natural \in \mathcal{H}, e \in \mathbb{F}^m, \|e\|_{\ell_2} \leq \varepsilon, a \in \mathcal{A} \tag{2}$$

Received May 9, 2018; Revised September 13, 2018; Accepted December 6, 2018.
2010 *Mathematics Subject Classification.* 94A12, 94A20.
Key words and phrases. compressive sensing, robust width property, robust null space property, restricted isometry property.

©2019 Korean Mathematical Society

holds for any given constants C_0 and C_1 via robust width property (RWP). Here the $\|\cdot\|_2$ on the left hand side above denotes the norm induced by the inner product over $\mathcal{H}$. Meanwhile, it was shown in [24] that the RWP can also be perfectly used to establish recovery results for the renowned Lasso [22] and Dantzig selector [5] models.

But in some certain cases, it is of special interest to use the ℓ_p-norms with $1 \leq p \leq \infty$ which is different from 2, as the loss function for measuring the noise level (see [1, 7, 14, 20, 23]). Hereafter, we will focus on the solutions of the corresponding ℓ_p-constrained CS problem:

$$\triangle_{\sharp,p}(y) := \arg\min_x \|x\|_\sharp \text{ subject to } \|\Phi x - y\|_{\ell_p} \leq \varepsilon. \tag{3}$$

The case of taking $\mathcal{H}$ to be either $\mathbb{R}^N$ or $\mathbb{C}^N$ and $\|\cdot\|_\sharp = \|\cdot\|_{\ell_1}$, and the case of taking $\mathcal{H}$ to be the Hilbert space of real (or complex) matrices with inner product $\langle X, Y\rangle = Tr(XY^*)$ and considering the nuclear norm $\|\cdot\|_\sharp = \|\cdot\|_*$, have been studied via the corresponding robust null space properties in [7] and [15], respectively. In this paper, we complement these recovery analysis results via RWP, a kind of weaker property compared to the robust null space property (NSP) and restricted isometry property (RIP). And from the other point of view, we simultaneously extend the existing results of RWP for ℓ_2-constrained CS in [2] to a much more general ℓ_p-constrained case with $p \geq 1$.

The remainder of the paper is organized as follows. In Section 2, we introduce the RWP and give the equivalent condition for stable and robust recovery results. We discuss the relationship between RWP and robust NSP in Section 3, and the relationship between RWP and RIP in Section 4. In Section 5, we obtain the optimal minimum required number of measurements for standard Gaussian matrices via RWP. We finish in Section 6 with a conclusion.

2. Robust width property

Firstly, we present the definition of RWP which acts as our basic recovery analysis tool. As will be given below, our newly defined RWP involves a new parameter p and it is obvious that the original RWP in [2] is a special case of ours with $p = 2$.

Definition 1. For any $p \geq 1$, we say a linear operator $\Phi : \mathcal{H} \to \mathbb{F}^m$ satisfies the (p, ρ, α)-RWP over the unit $\sharp$-ball $B_\sharp$ if

$$\|x\|_2 \leq \rho \|x\|_\sharp$$

for every $x \in \mathcal{H}$ such that $\|\Phi x\|_{\ell_p} < \alpha \|x\|_2$.

Remark 1. Obviously, this statement is equivalent to say that a linear operator $\Phi : \mathcal{H} \to \mathbb{F}^m$ satisfies the (p, ρ, α)-RWP over the unit $\sharp$-ball $B_\sharp$ if $\|\Phi x\|_{\ell_p} \geq \alpha \|x\|_2$ for every $x \in \mathcal{H}$ such that $\|x\|_2 > \rho \|x\|_\sharp$. In the special case that $\mathcal{H} = \mathbb{R}^N$, $\|\cdot\|_\sharp = \|\cdot\|_{\ell_1}$ and $p = 2$, with the ℓ_1-constrained minimal singular

value (CMSV) of Φ being defined as

$$r_s(\Phi) = \min_{x\neq 0, \|x\|_{\ell_1}^2/\|x\|_{\ell_2}^2 \leq s} \frac{\|\Phi x\|_{\ell_2}}{\|x\|_{\ell_2}}, \tag{4}$$

see [21], then the corresponding RWP reduces to the ℓ_1-CMSV property $r_{\frac{1}{p^2}}(\Phi) \geq \alpha$. The main benefit of adopting this kind of incoherence measure is that it is computable since it's hard to verify NSP and RIP for an arbitrary measurement matrix, see [21] for more arguments. Other related studies include the geometrical aspect of ℓ_1-CMSV discussed in [25], and the general extension to q-ratio CMSV with some $1 < q \leq \infty$ established recently in [26], in which setting the ℓ_1-CMSV corresponds to the special case of q-ratio CMSV with $q = 2$.

In order to apply the results to various instances of CS, we adopt the definition of CS spaces from [2] here.

Definition 2 (Definition 1 in [2])**.** A CS space $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ with bound L consists of a finite-dimensional Hilbert space $\mathcal{H}$, a subset $\mathcal{A} \subseteq \mathcal{H}$, and a norm $\|\cdot\|_\sharp$ with the following properties:

(i) $0 \in \mathcal{A}$.

(ii) For every $a \in \mathcal{A}$ and $v \in \mathcal{H}$, there exists a decomposition $v = v_1 + v_2$ such that

$$\|a + v_1\|_\sharp = \|a\|_\sharp + \|v_1\|_\sharp, \ \|v_2\|_\sharp \leq L\|v\|_2.$$

Remark 2. As presented in Section 4 of [2], a variety of examples of CS framework satisfy this framework, including weighted sparsity [13,16,19], block sparsity [9,10], gradient sparsity [17,18] and low-rank matrices [15].

Next, we give the main result that completely characterizes the measurement operators Φ for which the stable and robust recovery results for ℓ_p-constrained CS are achieved.

Theorem 1. *For any $p \geq 1$, any CS space $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ with bound L and any linear operator $\Phi : \mathcal{H} \to \mathbb{F}^m$, the following statements are equivalent up to constants*:

(a) *Φ satisfies the (p, ρ, α)-RWP over $B_\sharp$.*

(b) *For every $x^\natural \in \mathcal{H}$, $y = \Phi x^\natural + e$ with $e \in \mathbb{F}^m$ and $\|e\|_{\ell_p} \leq \varepsilon$, any solution $x^\star$ to*

$$\min_x \|x\|_\sharp \ \textit{subject to} \ \|\Phi x - y\|_{\ell_p} \leq \varepsilon \tag{5}$$

satisfies $\|x^\star - x^\natural\|_2 \leq C_0\|x^\natural - a\|_\sharp + C_1\varepsilon$ for every $a \in \mathcal{A}$.

In particular, (a) *implies* (b) *with $C_0 = 4\rho$ and $C_1 = \frac{2}{\alpha}$ provided $\rho \leq 1/(4L)$. Also,* (b) *implies* (a) *with $\rho = 2C_0$ and $\alpha = \frac{1}{2C_1}$.*

When taking $p = 2$, this theorem goes to Theorem 3 in [2]. Its proof follows from Theorem 3 in [2] with minor modification. For the sake of completeness, we present the proof here.

Proof. (a)⇒(b): Pick $a \in \mathcal{A}$ and decompose $v = x^\star - x^\natural = v_1 + v_2$ according to (ii) in Definition 2 such that $\|a + v_1\|_\sharp = \|a\|_\sharp + \|v_1\|_\sharp$ and $\|v_2\|_\sharp \le L\|v\|_2$. Then

$$\begin{aligned}
\|a\|_\sharp + \|x^\natural - a\|_\sharp &\ge \|x^\natural\|_\sharp \\
&\ge \|x^\star\|_\sharp \\
&= \|x^\natural + (x^\star - x^\natural)\|_\sharp \\
&= \|a + (x^\natural - a) + v_1 + v_2\|_\sharp \\
&\ge \|a + v_1\|_\sharp - \|(x^\natural - a) + v_2\|_\sharp \\
&\ge \|a + v_1\|_\sharp - \|x^\natural - a\|_\sharp - \|v_2\|_\sharp \\
&= \|a\|_\sharp + \|v_1\|_\sharp - \|x^\natural - a\|_\sharp - \|v_2\|_\sharp.
\end{aligned}$$

Thus $\|v_1\|_\sharp \le 2\|x^\natural - a\|_\sharp + \|v_2\|_\sharp$, which implies that

$$\begin{aligned}
\|x^\star - x^\natural\|_\sharp &\le \|v_1\|_\sharp + \|v_2\|_\sharp \\
&\le 2\|x^\natural - a\|_\sharp + 2\|v_2\|_\sharp \\
&\le 2\|x^\natural - a\|_\sharp + 2L\|x^\star - x^\natural\|_2.
\end{aligned} \tag{6}$$

Assume that $\|x^\star - x^\natural\|_2 > C_1\varepsilon$, since otherwise the desired result holds. Then, we have

$$\|\Phi x^\star - \Phi x^\natural\|_{\ell_p} \le \|\Phi x^\star - (\Phi x^\natural + e)\|_{\ell_p} + \|e\|_{\ell_p} \le 2\varepsilon < \frac{2}{C_1}\|x^\star - x^\natural\|_2 = \alpha\|x^\star - x^\natural\|_2$$

by taking $C_1 = \frac{2}{\alpha}$. Therefore, by (a), it holds that

$$\|x^\star - x^\natural\|_2 \le \rho\|x^\star - x^\natural\|_\sharp. \tag{7}$$

Therefore, by substituting (7) into (6), we have

$$\|x^\star - x^\natural\|_\sharp \le \frac{2}{1 - 2\rho L}\|x^\natural - a\|_\sharp \le 4\|x^\natural - a\|_\sharp$$

provided that $\rho \le \frac{1}{4L}$. Finally, applying (7) again implies that

$$\|x^\star - x^\natural\|_2 \le \rho\|x^\star - x^\natural\|_\sharp \le 4\rho\|x^\natural - a\|_\sharp = C_0\|x^\natural - a\|_\sharp \le C_0\|x^\natural - a\|_\sharp + C_1\varepsilon$$

with $C_0 = 4\rho$.

(b)⇒(a): Pick $x^\natural$ such that $\|\Phi x^\natural\|_{\ell_p} < \alpha\|x^\natural\|_2$. Then by setting $\varepsilon = \alpha\|x^\natural\|_2$ and $e = 0$, we observe that $x^\star = 0$ is a feasible solution of (5), and so with $a = 0$ we obtain

$$\|x^\natural\|_2 = \|x^\star - x^\natural\|_2 \le C_0\|x^\natural\|_\sharp + C_1\varepsilon = C_0\|x^\natural\|_\sharp + \alpha C_1\|x^\natural\|_2.$$

Thus, we have

$$\|x^\natural\|_2 \le \frac{C_0}{1 - \alpha C_1}\|x^\natural\|_\sharp = \rho\|x^\natural\|_\sharp$$

by taking $\alpha = \frac{1}{2C_1}$ and $\rho = 2C_0$. □

3. Robust NSP implies RWP

In this section, we propose a general version of robust NSP for CS spaces $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ and establish its relationship with RWP.

Definition 3. For CS space $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ with bound L and $p \geq 1$, we say a linear operator $\Phi : \mathcal{H} \to \mathbb{F}^m$ satisfies the robust NSP with respect to ℓ_p with constants $\phi > 0$ and $\tau > 0$ if for all $v \in \mathcal{H}$ and $a \in \mathcal{A}$

$$\|v\|_2 \leq \phi\|v - a\|_\sharp + \tau\|\Phi v\|_{\ell_p}. \tag{8}$$

In fact, this kind of definition of robust NSP is a weaker version compared to the traditional ones [7,12,15,19] we used before. Several examples are listed here.

Example 1. Let $[N]$ denote the set $\{1, 2, \ldots, N\}$, $\mathrm{card}(S)$ denote the cardinality of a set S and S^c denote the complement $[N] \setminus S$ of a set S in $[N]$. Take $\mathcal{H}$ to be $\mathbb{C}^N$, $\mathcal{A} = \Sigma_s := \{x : \|x\|_{\ell_0} := \mathrm{card}(\{j \in [N] : x_j \neq 0\}) \leq s\}$ and $\|\cdot\|_\sharp = \|\cdot\|_{\ell_1}$. Then the traditional robust NSP with respect to ℓ_p norm of order s with constants $0 < \psi < 1$ and $\tau > 0$ [7,12] is defined as for any set $S \subset [N]$ with $|S| \leq s$ and any $v \in \mathbb{C}^N$,

$$\|v_S\|_{\ell_2} \leq \frac{\psi}{\sqrt{s}}\|v_{S^c}\|_{\ell_1} + \tau\|\Phi v\|_{\ell_p}.$$

If we take S to be the index set of s largest (in modulus) entries of v, then for any $v \in \mathbb{C}^N$ we have

$$\begin{aligned}
\|v\|_{\ell_2} &\leq \|v_S\|_{\ell_2} + \|v_{S^c}\|_{\ell_2} \\
&\leq \|v_S\|_{\ell_2} + \|v_{S^c}\|_{\ell_1} \\
&\leq \frac{\psi}{\sqrt{s}}\|v_{S^c}\|_{\ell_1} + \tau\|\Phi v\|_{\ell_p} + \|v_{S^c}\|_{\ell_1} \\
&= (\frac{\psi}{\sqrt{s}} + 1)\|v_{S^c}\|_{\ell_1} + \tau\|\Phi v\|_{\ell_p} \\
&\leq (\frac{\psi}{\sqrt{s}} + 1)\|v - a\|_{\ell_1} + \tau\|\Phi v\|_{\ell_p}.
\end{aligned}$$

Therefore, the robust NSP given in Definition 3 holds with $\phi = (\frac{\psi}{\sqrt{s}}+1)$. Similar conclusion holds in the weighted sparsity case. Specifically, for every $v \in \mathbb{C}^N$ and $0 < q \leq 2$, we define $\|v\|_{w,q} = \left(\sum_{j=1}^N w_j^{2-q}|v_j|^q\right)^{1/q}$ with $w = (w_j)_{j\in[N]}$ being a vector of weights $w_j \geq 1$. Set $\|v\|_\sharp = \|v\|_{w,1}$ and $\mathcal{A} = \Sigma_{w,s} := \left\{x : \sum_{\{j:x_j\neq 0\}} w_j^2 \leq s\right\}$. Then the weighted robust NSP with respective to ℓ_p norm of order s with constants $\psi \in (0, 1)$ and $\tau > 0$ (the case $p = 2$ see the Definition 4.1 in [19]):

$$\|v_S\|_{\ell_2} \leq \frac{\psi}{\sqrt{s}}\|v_{S^c}\|_{w,1} + \tau\|\Phi v\|_{\ell_p}$$

for all $v \in \mathbb{C}^N$ and any set $S \subset [N]$ with $w(S) := \sum_{j\in S} w_j^2 \leq s$, will also imply the robust NSP given in Definition 3.

Example 2. Take $\mathcal{H}$ to be the Hilbert space of $C^{n_1\times n_2}$ with inner product $\langle X, Y\rangle = Tr(XY^*)$, and consider the nuclear norm $\|X\|_\sharp = \|X\|_*$. Let $\mathcal{A}$ denote the sets of matrices with rank at most r. With the Schatten q-norm of $X \in C^{n_1\times n_2}$ given by $\|X\|_q = \left(\sum_{j=1}^n \sigma_j(X)^q\right)^{1/q} (q \geq 1)$ where $\sigma_j(X), j = 1, \ldots, n$ denote the singular values of X, the nuclear norm $\|\cdot\|_*$ and the Frobenius norm $\|\cdot\|_F$ correspond to the cases that $q = 1$ and $q = 2$, respectively. We write $X = X_r + X_c$, where X_r is the best rank r approximation of X with respect to any Schatten q-norm of X. Then, for $p \geq 1$, we say that $\Phi : \mathbb{C}^{n_1\times n_2} \to \mathbb{C}^m$ satisfies the Frobenius-robust rank NSP with respect to ℓ_p of order r with constants $0 < \psi < 1$ and $\tau > 0$ [15] if for all $M \in \mathbb{C}^{n_1\times n_2}$, the singular values of M satisfy

$$\|M_r\|_2 \leq \frac{\psi}{\sqrt{r}}\|M_c\|_1 + \tau\|\Phi M\|_{\ell_p}.$$

Then, for all M, it is obvious that

$$\begin{aligned}
\|M_r\|_2 &\leq \|M_r\|_2 + \|M_c\|_2 \\
&\leq \|M_r\|_2 + \|M_c\|_1 \\
&\leq \frac{\psi}{\sqrt{r}}\|M_c\|_1 + \tau\|\Phi M\|_{\ell_p} + \|M_c\|_1 \\
&\leq (\frac{\psi}{\sqrt{r}} + 1)\|M_c\|_1 + \tau\|\Phi M\|_{\ell_p} \\
&\leq (\frac{\psi}{\sqrt{r}} + 1)\|M - a\|_1 + \tau\|\Phi M\|_{\ell_p}
\end{aligned}$$

holds for all $a \in \mathcal{A}$. Therefore, the robust NSP given in Definition 3 holds in this specific CS space.

Next, we show that compared to the robust NSP given above, RWP is actually a weaker property.

Theorem 2. *For any CS space $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ with bound L, if a linear operator $\Phi : \mathcal{H} \to \mathbb{F}^m$ satisfies the robust NSP with respect to ℓ_p with constants $\phi > 0$ and $\tau > 0$, then this linear operator Φ satisfies the (p, ρ, α)-RWP over $B_\sharp$ with $\rho = 2\phi$ and $\alpha = \frac{1}{2\tau}$.*

Proof. If for any $x \in \mathcal{H}$ such that $\|\Phi x\|_{\ell_p} < \alpha\|x\|_2$, then by applying the robust NSP with respect to ℓ_p of Φ, we have

$$\|x\|_2 \leq \phi\|x - a\|_\sharp + \tau\|\Phi x\|_{\ell_p} \leq \phi\|x - a\|_\sharp + \alpha\tau\|x\|_2$$

holds for all $a \in \mathcal{A}$. Let $a = 0$ and rearrange the inequality, we obtain that

$$\|x\|_2 \leq \frac{\phi}{1 - \alpha\tau}\|x\|_\sharp.$$

Thus, $\|x\|_2 \le \rho\|x\|_\sharp$ holds by assuming that $\rho = 2\phi$ and $\alpha = \frac{1}{2\tau}$. □

4. A direct proof of $RIP_{p,2}$ implies RWP

Basically, it's often inconvenient to show the robust NSP directly for a given matrix. Thus, one usually works instead with bit stronger properties, i.e., various versions of RIP. And it's well known that RIP implies part (b) in Theorem 1. Then, one could come into conclusion that RIP implies the RWP based on the equivalence of RWP to the stable and robust recovery result. For the sake of completeness, we will give a direct proof of RIP implies RWP in this section. In addition, we merely focus on the traditional sparsity case for CS space here instead of different versions of RIP for different CS spaces. The similar proofs hold for other cases. To handle the ℓ_p-constrained CS, the following $RIP_{p,2}$ [14] is used.

Definition 4. We say Φ satisfies the $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$ if

$$\mu_{p,2}(1-\delta)\|x\|_{\ell_2} \le \|\Phi x\|_{\ell_p} \le \mu_{p,2}(1+\delta)\|x\|_{\ell_2} \tag{9}$$

for every s-sparse vector x.

It is known that the stable and robust recovery result for the ℓ_p-constrained CS is guaranteed if Φ satisfies the $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$ with some small δ in the case of $2 \le p < \infty$, see Theorem 2 in [14] for detailed arguments. But this RIP condition is strictly stronger than the RWP condition. Next, we are going to present the fact and proof of that the RIP with a small constant δ does imply the RWP. Before that, the following lemma which will be used in this proof is provided. It is a direct extension of Lemma 12 in [2] from $p = 2$ to $p \ge 1$.

Lemma 1. *Suppose that Φ satisfies the right-hand inequality of $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$. Then for every x such that $\|x\|_{\ell_2} > \rho\|x\|_{\ell_1}$, we have*

$$\|x - x_S\|_{\ell_2} < \frac{1}{\rho\sqrt{s}}\|x\|_{\ell_2}, \ \ \|\Phi(x - x_S)\|_{\ell_p} < \frac{(1+\delta)\mu_{p,2}}{\rho\sqrt{s}}\|x\|_{\ell_2},$$

where S denotes the indices of the s largest (in modulus) entries of x.

Proof. Let $S_0 = S$, and for $j \ge 1$, S_j denote the indices of the s largest (in modulus) entries of x not covered by S_i for $i < j$. Then, for all $j \ge 0$, it holds that

$$\|x_{S_{j+1}}\|_{\ell_2} \le \sqrt{s}\max_{i\in S_{j+1}}|x_i| \le \sqrt{s}\min_{i\in S_j}|x_i| \le \frac{1}{\sqrt{s}}\|x_{S_j}\|_{\ell_1}.$$

Therefore, for every x such that $\|x\|_{\ell_2} > \rho\|x\|_{\ell_1}$,

$$\|x - x_S\|_{\ell_2} \le \sum_{j\ge1}\|x_{S_j}\|_{\ell_2} \le \frac{1}{\sqrt{s}}\sum_{j\ge0}\|x_{S_j}\|_{\ell_1} = \frac{1}{\sqrt{s}}\|x\|_{\ell_1} < \frac{1}{\rho\sqrt{s}}\|x\|_{\ell_2}.$$

Similarly, it holds that

$$\|\Phi(x - x_S)\|_{\ell_p} \le \sum_{j\ge1}\|\Phi x_{S_j}\|_{\ell_p} \le (1+\delta)\mu_{p,2}\sum_{j\ge1}\|x_{S_j}\|_{\ell_2} < \frac{(1+\delta)\mu_{p,2}}{\rho\sqrt{s}}\|x\|_{\ell_2}.$$

□

Theorem 3. *Suppose Φ satisfies $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$ with $\delta < \frac{1}{3}$. Then Φ satisfies the (p, ρ, α)-RWP over B_{ℓ_1} with*

$$\rho = \frac{3}{\sqrt{s}}, \ \ \alpha = (\frac{1}{3} - \delta)\mu_{p,2}.$$

Proof. The proof procedure follows from the proof of Theorem 11 in [2] with minor modification. It will be proved by contradiction. We assume Φ satisfies the right-hand inequality of $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$, but violates the (p, ρ, α)-RWP, then the proof of the theorem is completed if we can show that Φ necessarily violates the left-hand inequality of $(\mu_{p,2}, s, \delta)$-$RIP_{p,2}$.

In fact, if we pick x such that $\|\Phi x\|_{\ell_p} < \alpha\|x\|_{\ell_2}$ but $\|x\|_{\ell_2} > \rho\|x\|_{\ell_1}$, then we have

$$\begin{aligned} \|\Phi x_S\|_{\ell_p} &\le \|\Phi x\|_{\ell_p} + \|\Phi(x - x_S)\|_{\ell_p} \\ &< \alpha\|x\|_{\ell_2} + \frac{(1+\delta)\mu_{p,2}}{\rho\sqrt{s}}\|x\|_{\ell_2} \\ &= \frac{2}{3}(1-\delta)\mu_{p,2}\|x\|_{\ell_2} \end{aligned} \tag{10}$$

by using Lemma 1 and our specific choices of ρ and α. By adopting Lemma 1 again, it holds that

$$\|x\|_{\ell_2} \le \|x_S\|_{\ell_2} + \|x - x_S\|_{\ell_2} < \|x_S\|_{\ell_2} + \frac{1}{3}\|x\|_{\ell_2},$$

which implies $\|x\|_{\ell_2} < \frac{3}{2}\|x_S\|_{\ell_2}$. Therefore, we have

$$\|\Phi x_S\|_{\ell_p} < \frac{2}{3}(1-\delta)\mu_{p,2}\|x\|_{\ell_2} = (1-\delta)\mu_{p,2}\|x_S\|_{\ell_2}.$$

This is the desired result and completes the proof. □

5. Recovery results via RWP

In this section, we will provide the recovery results for ℓ_p-constrained CS directly via RWP and find the required minimum number of measurements for standard Gaussian matrices. We take $\mathcal{H} = \mathbb{R}^N$ and $\mathbb{F} = \mathbb{R}$ at the remainder of this section. To obtain the main result, it will be much more convenient to adopt the contrapositive statement of (p, ρ, α)-RWP: for any $p \ge 1$, a linear operator $\Phi : \mathbb{R}^N \to \mathbb{R}^m$ satisfies the (p, ρ, α)-RWP over B_{ℓ_1} if and only if $\|\Phi x\|_{\ell_p} \ge \alpha\|x\|_{\ell_2}$ for every $x \in \mathbb{R}^N$ such that $\|x\|_{\ell_2} > \rho\|x\|_{\ell_1}$. Then, by scaling, (p, ρ, α)-RWP is further equivalent to having $\|\Phi x\|_{\ell_p} \ge \alpha$ for every $x \in \mathbb{R}^N$ such that $\|x\|_{\ell_2} = 1$ and $\|x\|_{\ell_1} < \rho^{-1}$. Taking $T = \rho^{-1}B_{\ell_1} \cap \mathbb{S}^{N-1}$ where $\mathbb{S}^{N-1}$ denotes the unit sphere in $\mathbb{R}^N$, we have the following lemma:

Lemma 2. *A linear operator $\Phi : \mathbb{R}^N \to \mathbb{R}^m$ satisfies (p, ρ, α)-RWP over B_{ℓ_1} if and only if $\|\Phi x\|_{\ell_p} \ge \alpha$ for every $x \in T$.*

Remark 3. This lemma acts as a crucial result which characterizes the geometric feature of the RWP. A similar argument for the robust NSP was used in [7] (see its statement (7) in Proof of Theorem III.3), but with a different constraint set from T. Roughly speaking, comparing to the RIP, the upper bounds on the restricted eigenvalues are not required for the performance analysis of CS recovery.

Let $X_1, \ldots, X_m$ be i.i.d. copies of a random vector X in $\mathbb{R}^N$. Then we have the following key lemma concerning a lower bound for empirical processes:

Lemma 3 (Lemma III.1. in [7])**.** *Fix $1 \le p < \infty$, let $\mathcal{F}$ be a class of functions from $\mathbb{R}^N$ into $\mathbb{R}$. Consider $Q_{\mathcal{F}}(u) = \inf_{f\in\mathcal{F}} P(|f(X)| \ge u)$ and $R_m(\mathcal{F}) = E\sup_{f\in\mathcal{F}} |\frac{1}{m}\sum_{i=1}^m \varepsilon_i f(X_i)|$, where $\{\varepsilon_i\}_{i\ge1}$ is a Rademacher sequence. Let $u > 0$ and $t > 0$, then, with probability at least $1 - 2e^{-2t^2}$,*

$$\inf_{f\in\mathcal{F}} \frac{1}{m}\sum_{i=1}^m |f(X_i)|^p \ge u^p \left(Q_{\mathcal{F}}(2u) - \frac{4}{u}R_m(\mathcal{F}) - \frac{t}{\sqrt{m}} \right).$$

With the above two lemmas being applied, the main recovery result for standard Gaussian matrices is provided as follows:

Theorem 4. *Take $\rho > 0$ and let Φ be an $m \times N$ standard Gaussian matrix. Define the Gaussian width of T as $w(T) = E\sup_{x\in T}\langle g, x\rangle$, where g has i.i.d. $N(0,1)$ entries. For any $p \ge 1$, set $\alpha = c_1 m^{1/p}$ and if*

$$m \ge c_0 w(T)^2,$$

then Φ satisfies the (p, ρ, α)-RWP over B_{ℓ_1} with probability exceeding $1-2e^{-c_2 m}$.

Proof. According to Lemma 2, it suffices to show that

$$P\left(\inf_{x\in T} \|\Phi x\|_{\ell_p} \ge \alpha \right) \ge 1 - 2e^{-c_2 m}.$$

We write

$$\inf_{x\in T} \|\Phi x\|_{\ell_p} = \inf_{x\in T} \left(\sum_{i=1}^m |\langle X_i, x\rangle|^p \right)^{1/p},$$

where X_i denotes the i-th row of Φ.

Firstly, we suppose $1 \le p < \infty$, then applying Lemma 3 with $\mathcal{F} = \{\langle \cdot, x\rangle : x \in T\}$, we obtain that

$$\begin{aligned} \inf_{x\in T} \|\Phi x\|_{\ell_p} &= \inf_{x\in T} \left(\sum_{i=1}^m |\langle X_i, x\rangle|^p \right)^{1/p} \\ &\ge m^{1/p} u \left(Q_{\mathcal{F}}(2u) - \frac{4}{u}R_m(\mathcal{F}) - \frac{t}{\sqrt{m}} \right)^{1/p} \end{aligned} \tag{11}$$

holds with probability at least $1-2e^{-2t^2}$. Next, we estimate the small ball probability $Q_{\mathcal{F}}(2u)$ and the expected Rademacher supremum $R_m(\mathcal{F})$ respectively. Since for any $x \in \mathbb{S}^{N-1}$, $\|x\|_{\ell_2} = 1$, therefore

$$P(|\langle X_i, x\rangle| \geq u) = P(|G| \geq u),$$

where G is a standard Gaussian real-valued random variable. Hence, it holds that

$$Q_{\mathcal{F}}(2u) \geq P(|G| \geq 2u). \tag{12}$$

In addition, let $V = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \varepsilon_i X_i$, then V is a standard Gaussian vector. Then we have

$$R_m(\mathcal{F}) = \frac{1}{\sqrt{m}} E \sup_{x \in T} \langle V, x\rangle = \frac{1}{\sqrt{m}} w(T).$$

Now pick $u_\star$ small enough such that the right side of (12) is bigger than $\frac{1}{2}$ and $t = \sqrt{\frac{c_2 m}{2}}$ with $c_2 = 2\left(\frac{1}{2} - \frac{4}{u_\star\sqrt{c_0}} - (\frac{c_1}{u_\star})^p\right)^2$, then with probability exceeding $1 - 2e^{-c_2 m}$ it holds that

$$\begin{aligned}
\inf_{x \in T} \|\Phi x\|_{\ell_p} &\geq m^{1/p} u_\star \left(\frac{1}{2} - \frac{4}{u_\star} \frac{1}{\sqrt{m}} w(T) - \frac{t}{\sqrt{m}}\right)^{1/p} \\
&\geq m^{1/p} u_\star \left(\frac{1}{2} - \frac{4}{u_\star\sqrt{c_0}} - \sqrt{\frac{c_2}{2}}\right)^{1/p} \\
&=: c_1 m^{1/p} = \alpha.
\end{aligned}$$

Finally, when $p = \infty$, since $\|\Phi x\|_{\ell_{\log m}} \leq e\|\Phi x\|_{\ell_\infty}$, hence

$$P(\inf_{x \in T} \|\Phi x\|_{\ell_\infty} \geq c_1) \geq P(\inf_{x \in T} \|\Phi x\|_{\ell_{\log m}} \geq c_1 e).$$

Thus, the result follows from the already proved case that $p = \log m$ and the proof is completed. □

To further compute the lower bound of the required number of measurements m, we need the following upper bound for the Gaussian width of set T:

Lemma 4 (Lemma 16 in [2])**.** *There exists an absolute constant c such that*

$$w(\sqrt{s} B_{\ell_1} \cap \mathbb{S}^{N-1}) \leq c\sqrt{s \log(eN/s)} \tag{13}$$

for every positive integer s.

Then, the order of the minimum required number of measurements m for standard Gaussian matrices to guarantee the stable and robust recovery result with high probability is established here. A similar result is obtained in Theorem III.3 of [7] via robust NSP. Since it's shown in Section 3 that RWP is weaker than robust NSP, thus we can weaken the condition of the required minimum m up to constants although it seems that the obtained orders are the same and optimal via both methods.

Corollary 1. *Let Φ be an $m \times N$ standard Gaussian matrix. Fix $1 \le p \le \infty$ and $0 < \eta < 1$. There exists the set of constants $\{C, D, C', D'\}$ such that if*

$$m \ge Cs \log(eN/s) + D \log(2\eta^{-1}),$$

then with probability exceeding $1 - \eta$, for every $x^\natural \in \mathbb{R}^N$, $y = \Phi x^\natural + e$ with $e \in \mathbb{R}^m$ and $\|e\|_{\ell_p} \le \varepsilon$, any solution $x^\star$ to

$$\min_x \|x\|_{\ell_1} \quad \text{subject to} \quad \|\Phi x - y\|_{\ell_p} \le \varepsilon \tag{14}$$

satisfies $\|x^\star - x^\natural\|_{\ell_2} \le \frac{C'}{\sqrt{s}} \sigma_s(x^\natural)_1 + \frac{D'}{m^{1/p}} \varepsilon$, where $\sigma_s(x^\natural)_1 = \inf_{\|z\|_{\ell_0} \le s} \|x^\natural - z\|_{\ell_1}$.

Proof. Take the CS space $(\mathcal{H}, \mathcal{A}, \|\cdot\|_\sharp)$ with bound L to be the traditional sparse space with $\mathcal{H} = \mathbb{R}^N$, $\mathcal{A} = \Sigma_s$ and $\|\cdot\|_\sharp = \|\cdot\|_{\ell_1}$. Then as given in [2], we have $L = \frac{1}{\sqrt{s}}$. Set $\rho = \frac{C'}{4\sqrt{s}} \le \frac{1}{4\sqrt{s}}$ with $0 < C' \le 1$ and $\alpha = \frac{2}{D'} m^{1/p}$, then the desired result follows from Theorem 1 if the (p, ρ, α)-RWP over B_{ℓ_1} of Φ is verified. In fact, by adopting Theorem 4 and Lemma 4, this holds when $m \ge Cs \log(eN/s) + D \log(2\eta^{-1})$ with C and D being appropriately chosen. □

6. Conclusion

In this paper, we studied the equivalent RWP condition to guarantee the stable and robust recovery result for ℓ_p-constrained CS procedure. We proposed a general version of robust NSP and verified that both robust NSP and RIP imply RWP. Finally, the specific recovery condition for standard Gaussian matrices was established via RWP and the optimal order of the minimum required number of measurements was obtained. The present paper extends the CS recovery analysis results via RWP in [2] from ℓ_2-constrained to ℓ_p-constrained case with $p \ge 1$, and also largely complements the existing theoretical results for robust CS with ℓ_p loss function, see [1, 7, 14, 15, 20, 23], and among others.

Acknowledgements. We would like to thank the reviewer, the Associate Editor, and the Editor for their detailed and insightful comments and suggestions that help to improve the quality of this paper. This work is supported by the Swedish Research Council grant (Reg.No. 340-2013-5342).

References

[1] Z. Allen-Zhu, R. Gelashvili, and I. Razenshteyn, *Restricted isometry property for general p-norms*, IEEE Trans. Inform. Theory **62** (2016), no. 10, 5839–5854.

[2] J. Cahill and D. G. Mixon, *Robust width: A characterization of uniformly stable and robust compressed sensing*, arXiv preprint arXiv:1408.4409, 2014.

[3] E. J. Candès, *The restricted isometry property and its implications for compressed sensing*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 9-10, 589–592.

[4] E. J. Candès, J. K. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59** (2006), no. 8, 1207–1223.

[5] E. Candes and T. Tao, *The Dantzig selector: statistical estimation when p is much larger than n*, Ann. Statist. **35** (2007), no. 6, 2313–2351.

[6] A. Cohen, W. Dahmen, and R. DeVore, *Compressed sensing and best k-term approximation*, J. Amer. Math. Soc. **22** (2009), no. 1, 211–231.

[7] S. Dirksen, G. Lecué, and H. Rauhut, *On the gap between restricted isometry properties and sparse recovery conditions*, IEEE Trans. Inform. Theory **64** (2018), no. 8, 5478–5487.
[8] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52** (2006), no. 4, 1289–1306.
[9] M. F. Duarte and Y. C. Eldar, *Structured compressed sensing: from theory to applications*, IEEE Trans. Signal Process. **59** (2011), no. 9, 4053–4085.
[10] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, *Block-sparse signals: uncertainty relations and efficient recovery*, IEEE Trans. Signal Process. **58** (2010), no. 6, 3042–3054.
[11] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.
[12] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Applied and Numerical Harmonic Analysis, Birkhäuser/Springer, New York, 2013.
[13] M. P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, *Recovering compressively sampled signals using partial support information*, IEEE Trans. Inform. Theory **58** (2012), no. 2, 1122–1134.
[14] L. Jacques, D. K. Hammond, and J. M. Fadili, *Dequantizing compressed sensing: when oversampling and non-Gaussian constraints combine*, IEEE Trans. Inform. Theory **57** (2011), no. 1, 559–571.
[15] M. Kabanava, R. Kueng, H. Rauhut, and U. Terstiege, *Stable low-rank matrix recovery via null space properties*, Inf. Inference **5** (2016), no. 4, 405–441.
[16] H. Mansour and R. Saab, *Recovery analysis for weighted ℓ_1-minimization using the null space property*, Appl. Comput. Harmon. Anal. **43** (2017), no. 1, 23–38.
[17] D. Needell and R. Ward, *Near-optimal compressed sensing guarantees for total variation minimization*, IEEE Trans. Image Process. **22** (2013), no. 10, 3941–3949.
[18] ———, *Stable image reconstruction using total variation minimization*, SIAM J. Imaging Sci. **6** (2013), no. 2, 1035–1058.
[19] H. Rauhut and R. Ward, *Interpolation via weighted ℓ_1 minimization*, Appl. Comput. Harmon. Anal. **40** (2016), no. 2, 321–351.
[20] A. A. Saleh, F. Alajaji, and W. Y. Chan, *Compressed sensing with non-Gaussian noise and partial support information*, IEEE Signal Processing Letters **22** (2015), no. 10, 1703–1707.
[21] G. Tang and A. Nehorai, *Performance analysis of sparse recovery based on constrained minimal singular values*, IEEE Trans. Signal Process. **59** (2011), no. 12, 5734–5745.
[22] R. Tibshirani, *Regression shrinkage and selection via the lasso*, J. Roy. Statist. Soc. Ser. B **58** (1996), no. 1, 267–288.
[23] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, *Robust sparse recovery in impulsive noise via ℓ_p-ℓ_1 optimization*, IEEE Trans. Signal Process. **65** (2017), no. 1, 105–118.
[24] H. Zhang, *On robust width property for Lasso and Dantzig selector*, Commun. Math. Sci. **15** (2017), no. 8, 2387–2393.
[25] H. Zhang and L. Cheng, *On the constrained minimal singular values for sparse signal recovery*, IEEE Signal Processing Letters **19** (2012), no. 8, 499–502.
[26] Z. Zhou and J. Yu, *Sparse recovery based on q-ratio constrained minimal singular values*, Signal Processing **155** (2019), 247–258.

Jun Yu
Department of Mathematics and Mathematical Statistics
Umeå University
Umeå, 901 87, Sweden
Email address: jun.yu@umu.se

Zhiyong Zhou
Department of Mathematics and Mathematical Statistics
Umeå University
Umeå, 901 87, Sweden
Email address: zhiyong.zhou@umu.se