Sufficient Conditions on Stable Recovery of Sparse Signals With Partial Support Information

Xiaohan Yu and Seung Jun Baek, *Member, IEEE*

***Abstract*—In this letter, we study signal reconstruction from compressed sensing measurements. We propose new sufficient conditions for stable recovery when partial support information is available. Weighted ℓ_1-minimization is adopted to recover the original signal under three noise models. The proposed approach is to use Ozeki's inequality and shifting inequality in order to bound the errors in the associated weighted ℓ_1-minimization. Our result offers generalized performance bounds on recovery capturing known support information. Improved sufficient conditions for recovery are derived based on our results, even for the cases where the accuracy of prior support information is arbitrarily low.**

***Index Terms*—Compressive sensing (CS), partial support information, weighted ℓ_1-minimization.**

I. Introduction

COMPRESSIVE sensing has recently received much attention in signal and imaging processing [1]–[4]. The idea is to reliably recover high dimensional sparse signals from significantly fewer linear measurements. Consider the following linear model:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ $(m \ll n)$ has unit ℓ_2-norm columns, $\mathbf{x} \in \mathbb{R}^n$ is the original signal to be recovered and $\mathbf{e}$ is a vector of measurement errors. It is well known that ℓ_1-minimization provides effective way to reconstruct sparse signals [1]–[4]. The ℓ_1-minimization problem can be formulated as follows:

$$(P_{\mathfrak{B}}) \quad \min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{subject to } \mathbf{y} - \mathbf{A}\mathbf{u} \in \mathfrak{B} \tag{2}$$

where $\mathfrak{B}$ is a bounded set. In CS, *Restricted Isometry Constant* (RIC) and *Restricted Orthogonality Constant* (ROC) have been widely used to characterize the sensing matrix $\mathbf{A}$ [1], [2].

Definition 1: (RIC). Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $1 \le k \le n$. The *k-restricted isometry constant* δ_k of $\mathbf{A}$ is defined to be the smallest constant such that

$$(1-\delta_k)\|\mathbf{v}\|_2^2 \le \|\mathbf{A}\mathbf{v}\|_2^2 \le (1+\delta_k)\|\mathbf{v}\|_2^2,$$

for every vector $\mathbf{v}$ which is k-sparse.

Definition 2: (ROC). If $k+k' \le n$, the *k, k'-restricted orthogonality constant* $\theta_{k,k'}$ is the smallest constant that satisfies

$$\left|\langle \mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{v}'\rangle\right| \le \theta_{k,k'}\|\mathbf{v}\|_2\|\mathbf{v}'\|_2,$$

for all $\mathbf{v}$ and $\mathbf{v}'$ which are k-sparse and k'-sparse respectively with disjoint supports.

Roughly speaking, RIC and ROC measure how close the submatrices which consist of columns of A are to behaving like an orthonormal system.

Candès and Tao [1] studied the standard ℓ_1-minimization problem where they provided a sufficient condition $\delta_k + \theta_{k,k} + \theta_{k,2k} < 1$ for the stable recovery of k-sparse signals. Cai *et al.* [4] improved it to $\delta_{k+a} + \sqrt{(k/b)}\theta_{k+a,b} < 1$ by using *shifting inequality*. However, they assumed no prior information on the unknown signal which is not practical in many applications. For example, a large estimation problem may contain a CS inner loop which feeds prior information on the signal to the recovery algorithm [5]. In order to exploit the prior information, Friedlander *et al.* [6] introduced the weighted ℓ_1-minimization:

$$\min_{\mathbf{u}} \|\mathbf{u}\|_{1,\mathbf{w}} \quad \text{subject to } \|\mathbf{A}\mathbf{u} - \mathbf{y}\|_2 \le \epsilon \tag{3}$$

where $\|\mathbf{u}\|_{1,\mathbf{w}} = \sum_i w_i|u_i|$ with $w_i \in [0,1]$. The idea is to assign relatively small weights to the entries which belong to the actual support of $\mathbf{x}$ in the objective function. Let us denote the fraction of the accurate portion of the available support information by $\alpha \in (0,1]$. When compared to the case without prior information in [7], their scheme is shown to yield the smaller error bounds when α exceeds 50%.

In this letter we propose a new set of sufficient conditions for stable recovery when prior support information is available. Our main theorem extends and complements the results of [4], [6]. We show that, under certain conditions, the error bounds can be improved irrespective of how small the amount of accurate prior information is, i.e., for arbitrarily small α. Note that, as mentioned above, improved bounds are acquired in [6] only when $\alpha > 0.5$. The key technique in our derivation is to apply *Ozeki's inequality* [8] and *shifting inequality* [4] to weighted ℓ_1-minimization in order to tightly bound the error in ℓ_2-norm. Moreover, the condition for recovery in [6] involves only RIC whereas our condition involves both RIC and ROC. As a result, we are able to propose an improved sufficient condition for stable recovery compared to existing ones. Finally, we extend our result to recovery problems under two different noise models and provide explicit error bounds.

The rest of this letter is organized as follows. In Section II, we introduce basic notations, and prove our main theorem. We compare our results with the existing results, e.g., in [4], [6] in Section III. We present extended results under different noise models in Section IV. Section V concludes the paper.

Manuscript received January 29, 2013; revised March 18, 2013; accepted March 19, 2013. Date of publication March 27, 2013; date of current version April 04, 2013. This work was supported in part by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communications Agency) (KCA-2013-10-911-05-006). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gitta Kutyniok.

The authors are with Department of Computer and Communications Engineering, Korea University, Seoul 136-713, Korea (e-mail: yuxiaohan188@korea.ac.kr; sjbaek@korea.ac.kr).

Digital Object Identifier 10.1109/LSP.2013.2254712

II. Recovery by the Weighted ℓ_1-Minimization

We use boldface letters to denote vectors and matrices. The ith entry of vector $\mathbf{x}$ is denoted by x_i. For a vector $\mathbf{x} \in \mathbb{R}^n$ and a set $B \subseteq \{1, 2, \cdots, n\}$, we shall denote by $\mathbf{x}_B \in \mathbb{R}^n$ the vector whose entries in B are identical to those of $\mathbf{x}$ and the entries not in B are set to zero. We use the standard notation $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ to denote the ℓ_p-norm of a vector $\mathbf{x} \in \mathbb{R}^n$ for $p \geq 1$. A vector $\mathbf{x}$ is said to be k-sparse if the size of support does not exceed k, i.e., $|\mathrm{supp}(\mathbf{x})| \leq k$. A vector $\mathbf{x}$ is said to be *compressible* if $|x_{[i]}| \leq C \cdot i^{-s}$ for some positive C and $s \geq 1$ where $x_{[i]}$ denotes the ith largest (in terms of magnitude) entry of $\mathbf{x}$. Suppose $T_0 \subseteq \{1, 2, \cdots, n\}$ contains the indices of largest-in-magnitude entries of $\mathbf{x}$. We denote prior support information or an estimate of T_0 by $T \subseteq \{1, 2, \cdots, n\}$. Assume that $|T_0| = k$, $|T| = l$ and $|T_0 \cap T| = \alpha l$. Given a known estimate T of $\mathbf{x}$, we will set $w_i = w \in [0, 1]$ whenever $i \in T$, and $w_i = 1$ otherwise.

Theorem 2.1: For a signal $\mathbf{x} \in \mathbb{R}^n$, suppose T_0 with $|T_0| = k$ contains k indices of largest-in-magnitude entries of $\mathbf{x}$ and $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$ with $\mathbf{e}$ satisfying $\|\mathbf{e}\|_2 \leq \epsilon$. Suppose there exist positive integers a, b such that $(1 - \alpha)l \leq a$ and $b \leq 4a$, and the measurement matrix $\mathbf{A}$ satisfies

$$\delta_{k+a} + \gamma\theta_{k+a,b} < 1 \tag{4}$$

where

$$\gamma = \begin{cases} \gamma_1 := \frac{w\sqrt{k}+(1-w)\sqrt{k-\alpha l}}{\sqrt{b}}, & b \leq 2wa, \\ \gamma_2 := \frac{w\sqrt{k}+(1-w)\sqrt{k+l-2\alpha l}}{\sqrt{b}}, & 2wa < b \leq 4a, \end{cases} \tag{5}$$

for some $w \in [0, 1]$. We have that, the solution $\mathbf{x}^*$ of (3) obeys

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_1\epsilon + C_2 E \tag{6}$$

where C_1, C_2 and E are given by

$$\begin{aligned} C_1 &= \frac{2\sqrt{(1+\gamma^2)(1+\delta_{k+a})}}{1-\delta_{k+a}-\gamma\theta_{k+a,b}}, \\ C_2 &= \frac{2}{\sqrt{b}}\left(1 + \frac{(\sqrt{1+\gamma^2})\theta_{k+a,b}}{1-\delta_{k+a}-\gamma\theta_{k+a,b}}\right), \\ E &= w\left\|\mathbf{x}_{T_0^c}\right\|_1 + (1-w)\left\|\mathbf{x}_{T_0^c \cap T^c}\right\|_1. \end{aligned} \tag{7}$$

Note that our theorem is applicable to both sparse and compressible signals. If the original signal is k-sparse, the second term of the RHS of (6) is zero by definition. Hence the exact recovery is guaranteed if we assume that the measurements are acquired without noise. If the original signal is compressible, the second term of the RHS of (6) tends to be small for high dimensional signals, hence the stable recovery is guaranteed.

Proof: Without loss of generality, we assume the first k entries of $\mathbf{x}$ are the largest in magnitude, i.e., $T_0 = \{1, \cdots, k\}$. Denote the error vector $\mathbf{h} = \mathbf{x}^* - \mathbf{x}$. We will partition T_0^c into disjoint subsets of the sizes a and b, similar to the method proposed in [4], as follows. Let $T_* = \{k+1, \cdots, k+a\}$ and $T_i = \{k + a + (i-1)b + 1, \cdots, k + a + ib\}$, $i = 1, 2, \cdots, t$. We append 0 entries to $\mathbf{h}$ so that $|T_t| = b$ if $|T_t| < b$. Let us rearrange the entries of $\mathbf{h}$ such that $|h_{k+1}| \geq |h_{k+2}| \geq \cdots$. We observe that

$$\|\mathbf{h}\|_2^2 = \sum_i \|\mathbf{h}_{T_i}\|_2^2 + \|\mathbf{h}_{T_0 \cup T_*}\|_2^2 \tag{8}$$

$$\leq \left(\sum_i \|\mathbf{h}_{T_i}\|_2\right)^2 + \|\mathbf{h}_{T_0 \cup T_*}\|_2^2. \tag{9}$$

The inequality holds since $\|\mathbf{u}\|_2^2 \leq \|\mathbf{u}\|_1^2$ for some vector $\mathbf{u}$. We begin by introducing Ozeki's inequality [8].

Lemma 2.2: Let $\mathbf{u} = [u_1, \cdots, u_n]$ and $\mathbf{v} = [v_1, \cdots, v_n]$ be two nonnegative n-tuples with $m_u \leq u_i \leq M_u$ and $m_v \leq v_i \leq M_v$ for each $i \in \{1, \cdots, n\}$ and some nonnegative constants m_u, m_v, M_u, M_v. We have that

$$\sum_{i=1}^n u_i^2 \sum_{i=1}^n v_i^2 - \left(\sum_{i=1}^n u_i v_i\right)^2 \leq \frac{n^2}{4}(M_u M_v - m_u m_v)^2.$$

By letting $\mathbf{v} = [1, \cdots, 1]$ and applying $\sqrt{g^2 + h^2} \leq g + h$ for nonnegative g, h to Lemma 2.2, we have that

$$\sqrt{n}\|\mathbf{u}\|_2 \leq \|\mathbf{u}\|_1 + \frac{n}{2}(M_u - m_u). \tag{10}$$

By applying (10) to $\mathbf{h}_{T_i}$ for each $i \in \{1, \cdots, t\}$, we have that

$$\sqrt{b}\,\|\mathbf{h}_{T_i}\|_2 \leq \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2}(M_i - m_i) \tag{11}$$

where M_i (resp. m_i) is the maximum (resp. minimum) entry in magnitude of $\mathbf{h}_{T_i}$. From (11), we have that

$$\begin{aligned} \sqrt{b}\sum_i \|\mathbf{h}_{T_i}\|_2 &\leq \sum_i \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2}\sum_i (M_i - m_i) \\ &= \sum_i \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2}\left[M_1 - \sum_{i=1}^{t-1}(m_i - M_{i+1}) - m_t\right] \\ &\leq \sum_i \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2}M_1 && (12) \\ &\leq \sum_i \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2a}\|\mathbf{h}_{T_*}\|_1 && (13) \\ &= \left\|\mathbf{h}_{T_0^c}\right\|_1 - \left(1 - \frac{b}{2a}\right)\|\mathbf{h}_{T_*}\|_1. && (14) \end{aligned}$$

Note that (12) and (13) hold since we assume $|h_{k+1}| \geq |h_{k+2}| \geq \cdots$, and by the definition of T_*. Next we will bound $\|\mathbf{h}_{T_0^c}\|_1$. Let $T_0 \cap T^c = J_1$ and $T_0^c \cap T = J_2$. Since $\mathbf{x}^*$ is the solution to (3), we have that $\|\mathbf{x}^*\|_{1,\mathbf{w}} \leq \|\mathbf{x}\|_{1,\mathbf{w}}$, or

$$\|\mathbf{x} + \mathbf{h}\|_{1,\mathbf{w}} \leq \|\mathbf{x}\|_{1,\mathbf{w}}. \tag{15}$$

By applying reverse triangle inequality to (15), we obtain (refer to inequality (21) in [6])

$$\left\|\mathbf{h}_{T_0^c}\right\|_1 \leq w\|\mathbf{h}_{T_0}\|_1 + (1-w)\|\mathbf{h}_{J_1 \cup J_2}\|_1 + 2E. \tag{16}$$

We will consider two cases based on how a and b are related.

Case 1) $b \leq 2wa$. Subtracting $(1 - (b/2a))\|\mathbf{h}_{T_*}\|_1$ from both sides of (16), we obtain

$$\begin{aligned} \left\|\mathbf{h}_{T_0^c}\right\|_1 - \left(1 - \frac{b}{2a}\right)\|\mathbf{h}_{T_*}\|_1 &\leq w\|\mathbf{h}_{T_0}\|_1 + (1-w)\|\mathbf{h}_{J_1}\|_1 \\ &\quad + \left((1-w)\|\mathbf{h}_{J_2}\|_1 - \left(1 - \frac{b}{2a}\right)\|\mathbf{h}_{T_*}\|_1\right) + 2E && (17) \\ &\leq w\|\mathbf{h}_{T_0}\|_1 + (1-w)\|\mathbf{h}_{J_1}\|_1 + 2E && (18) \\ &\leq \left(w\sqrt{k} + (1-w)\sqrt{k-\alpha l}\right)\|\mathbf{h}_{T_0 \cup T_*}\|_2 + 2E. && (19) \end{aligned}$$

Note that (18) holds since $b \leq 2wa$ and $(1 - \alpha)l < a$. Inequality (19) holds by Cauchy-Schwarz inequality:

$$\|\mathbf{h}_{T_0}\|_1 \leq \sqrt{k}\,\|\mathbf{h}_{T_0}\|_2 \leq \sqrt{k}\,\|\mathbf{h}_{T_0 \cup T_*}\|_2, \tag{20}$$

$$\|\mathbf{h}_{J_1}\|_1 \leq \sqrt{k - \alpha l}\,\|\mathbf{h}_{J_1}\|_2 \leq \sqrt{k - \alpha l}\,\|\mathbf{h}_{T_0 \cup T_*}\|_2. \tag{21}$$

By applying (19) to (14), we obtain

$$\sum_i \|\mathbf{h}_{T_i}\|_2 \leq \gamma_1\|\mathbf{h}_{T_0 \cup T_*}\|_2 + 2E/\sqrt{b}. \tag{22}$$

Case 2) $2wa < b \leq 4a$. Firstly assume $2wa < b \leq 2a$. We have that, from (13),

$$\sum_i \|\mathbf{h}_{T_i}\|_1 + \frac{b}{2a}\|\mathbf{h}_{T_*}\|_1 \leq \sum_i \|\mathbf{h}_{T_i}\|_1 + \|\mathbf{h}_{T_*}\|_1 = \left\|\mathbf{h}_{T_0^c}\right\|_1,$$

which implies that

$$\sum_i \|\mathbf{h}_{T_i}\|_2 \leq \frac{\left\|\mathbf{h}_{T_0^c}\right\|_1}{\sqrt{b}}. \tag{23}$$

Secondly assume $2a < b \leq 4a$. In this case, by applying shifting inequality [4], one can show that (23) holds as well (refer to the proof of Theorem 2 in [4]). By the assumption $(1-\alpha)l \leq a$ and Cauchy-Schwarz inequality, we have that,

$$\begin{aligned}\|\mathbf{h}_{J_1\cup J_2}\|_1 &\leq \sqrt{k+l-2\alpha l}\,\|\mathbf{h}_{J_1\cup J_2}\|_2\\ &\leq \sqrt{k+l-2\alpha l}\,\|\mathbf{h}_{T_0\cup T_*}\|_2 .\end{aligned} \tag{24}$$

By applying (16), (20) and (24) to (23), we obtain

$$\sum_i \|\mathbf{h}_{T_i}\|_2 \leq \gamma_2 \|\mathbf{h}_{T_0\cup T_*}\|_2 + 2E/\sqrt{b}. \tag{25}$$

Now we observe that $|\langle \mathbf{Ah}, \mathbf{Ah}_{T_0\cup T_*}\rangle|$ is given by

$$\begin{aligned}&\left|\langle \mathbf{Ah}_{T_0\cup T_*}, \mathbf{Ah}_{T_0\cup T_*}\rangle + \sum_i \langle \mathbf{Ah}_{T_i}, \mathbf{Ah}_{T_0\cup T_*}\rangle\right|\\ &\geq \|\mathbf{Ah}_{T_0\cup T_*}\|_2^2 - \sum_i |\langle \mathbf{Ah}_{T_i}, \mathbf{Ah}_{T_0\cup T_*}\rangle|\\ &\geq (1-\delta_{k+a})\,\|\mathbf{h}_{T_0\cup T_*}\|_2^2 - \theta_{k+a,b}\,\|\mathbf{h}_{T_0\cup T_*}\|_2 \sum_i \|\mathbf{h}_{T_i}\|_2 .\end{aligned} \tag{26}$$

By applying (22) and (25) to (26), we have that

$$\begin{aligned}|\langle \mathbf{Ah}, \mathbf{Ah}_{T_0\cup T_*}\rangle| &\geq (1-\delta_{k+a})\,\|\mathbf{h}_{T_0\cup T_*}\|_2^2\\ &\quad -\theta_{k+a,b}\,\|\mathbf{h}_{T_0\cup T_*}\|_2 \left(\gamma \|\mathbf{h}_{T_0\cup T_*}\|_2 + 2E/\sqrt{b}\right).\end{aligned} \tag{27}$$

By the assumption $\|\mathbf{e}\|_2 \leq \epsilon$, we have that $\|\mathbf{Ah}\|_2 \leq \|\mathbf{y} - \mathbf{Ax}^*\|_2 + \|\mathbf{y}-\mathbf{Ax}\|_2 \leq \epsilon$. Consequently,

$$\begin{aligned}|\langle \mathbf{Ah}, \mathbf{Ah}_{T_0\cup T_*}\rangle| &\leq \|\mathbf{Ah}\|_2\,\|\mathbf{Ah}_{T_0\cup T_*}\|_2\\ &\leq 2\epsilon\sqrt{1+\delta_{k+a}}\,\|\mathbf{h}_{T_0\cup T_*}\|_2 .\end{aligned} \tag{28}$$

By combining (27) and (28), we obtain

$$\|\mathbf{h}_{T_0\cup T_*}\|_2 \leq \frac{2\theta_{k+a,b}E/\sqrt{b} + 2\epsilon\sqrt{1+\delta_{k+a}}}{1-\delta_{k+a}-\gamma\theta_{k+a,b}}. \tag{29}$$

In order for (29) to hold, $\delta_{k+a} + \gamma\theta_{k+a,b} < 1$ must be satisfied. Finally, we can bound the error vector as follows.

$$\begin{aligned}\|\mathbf{h}\|_2^2 &\leq \|\mathbf{h}_{T_0\cup T_*}\|_2^2 + \left(\gamma \|\mathbf{h}_{T_0\cup T_*}\|_2 + 2E/\sqrt{b}\right)^2\\ &\leq \left(\sqrt{\gamma^2+1}\,\|\mathbf{h}_{T_0\cup T_*}\|_2 + 2E/\sqrt{b}\right)^2 .\end{aligned} \tag{30}$$

By combining (30) and (29), the theorem is proved. ■

III. Discussion

In this section, we will compare our results to the existing ones. We first observe that $E \leq \|\mathbf{x}_{T_0^c}\|_1$. Suppose $w \in [0, 1)$.

1) If $b \leq 2wa$, the constants C_1 and C_2 are strictly smaller than the analogous constants of Theorem 4 in [4], provided that $\alpha > 0$. This implies that our error bound is strictly smaller than the case without prior support information, for arbitrarily small α. This result is in contrast to that on partial support information in [6] in which error bounds are improved only if $\alpha > 0.5$. Moreover, for any $\alpha > 0$, the following holds:

$$\gamma = \gamma_1 = \sqrt{\frac{k}{b}}\left\{w + (1-w)\sqrt{1-\frac{\alpha l}{k}}\right\} < \sqrt{\frac{k}{b}}.$$

Hence our condition (4) is strictly weaker than $\delta_{k+a} + \sqrt{k/b}\,\theta_{k+a,b} < 1$ as proposed in [4] for any $\alpha > 0$. Note that we extend the results of standard ℓ_1-minimization problem $(w = 1)$ in [4] since we cover the case $b < a$.

2) If $2wa < b \leq 4a$ and $\alpha = 0.5$, the constants C_1, C_2 and the sufficient condition in Theorem 2.1 are identical to those from Theorem 4 in [4].

3) If $2wa < b \leq 4a$ and $\alpha > 0.5$, the constants C_1 and C_2 are smaller than those in [4]. Again, $\gamma < \sqrt{k/b}$ holds, thus our condition (4) is strictly weaker than that in [4].

Below we will derive weaker sufficient conditions than those known in the literature. For the case 1), we can derive $\delta_{2k} + \gamma_1\theta_{k,2k} < 1$ by letting $w = 0.5$ and $a = b = k$. Since $\gamma_1 < 1$ for any $\alpha > 0$, the derived condition is strictly weaker than $\delta_{2k} + \theta_{k,2k} < 1$ as proposed in [9].

For the case 3), let us first introduce a useful lemma.

Lemma 3.1: (Cai [4]) For any $a \geq 1$ and positive integers k, k' such that ak' is an integer, we have that

$$\theta_{k,k'} \leq \delta_{k+k'}, \theta_{k,ak'} \leq \sqrt{a}\theta_{k,k'}.$$

Let $a = pk$, $b = 4pk$ with $p = (\mathbb{Z}_+/k)$ and $\beta = w + (1-w)\sqrt{1+(1-2\alpha)l/k}$. Note that $\gamma = \gamma_2$ holds in this case. If $p \in (0, 1]$, from Lemma 3.1, we have $\theta_{(p+1)k,4pk} < \sqrt{p+1}\theta_{k,4pk} < \sqrt{p+1}\delta_{(4p+1)k}$. Thus our condition reduces to $\delta_{(4p+1)k} < 1/(1+\beta\sqrt{1/4+1/4p})$, or equivalently, $\delta_{(p+1)k} < 1/(1+\beta\sqrt{1/4+1/p})$. If $p > 1$, from Lemma 3.1, we have $\theta_{(p+1)k,4pk} < \delta_{(5p+1)k}$, from which $\delta_{(p+1)k} < 1/(1+\beta\sqrt{5/4p})$ is derived. Thus the proposed condition is

$$\delta_{(p+1)k} < \begin{cases}\delta_1 := \left(1+\beta\sqrt{\frac{4+p}{4p}}\right)^{-1}, & 0 < p \leq 1,\\ \delta_2 := \left(1+\beta\sqrt{\frac{5}{4p}}\right)^{-1}, & p > 1.\end{cases} \tag{31}$$

When $w = 1$ we can obtain different conditions from (31): $\delta_{1.75k} < 0.443$, $\delta_{2k} < 0.472$ and $\delta_{3k} < 0.558$. All of the above conditions are weaker than the existing ones: $\delta_{1.75k} < 0.414$ [3], $\delta_{2k} < 0.414$ [2], and $\delta_{3k} < 0.535$ [4].

The work [6] studied weighted ℓ_1-minimization under partial support information, and introduced the sufficient condition

$$\delta_{pk} + \frac{p}{\beta^2}\delta_{(p+1)k} < \frac{p}{\beta^2} - 1 \tag{32}$$

where $p > 1$. A direct comparison between (4) and (32) is not possible since (32) involves RIC only. Instead, we will consider conditions of the form $\delta_{(p+1)k} < C$ proposed in [6]. When $p = 1$, $\delta_{2k} < (1/1+\beta\sqrt{2})$ was proposed [6], however our condition $\delta_{2k} < (1/1+\beta\sqrt{1.25})$ from (31) is weaker. When $p > 1$, the following was proposed [6]:

$$\delta_{(p+1)k} < \delta_3 := \frac{p-\beta^2}{p+\beta^2}. \tag{33}$$

We compared (31) and (33) in Fig. 1 when $p = 3$ and $l = k$. One can show that our condition (31) is weaker than (33), i.e., $\delta_2 > \delta_3$,

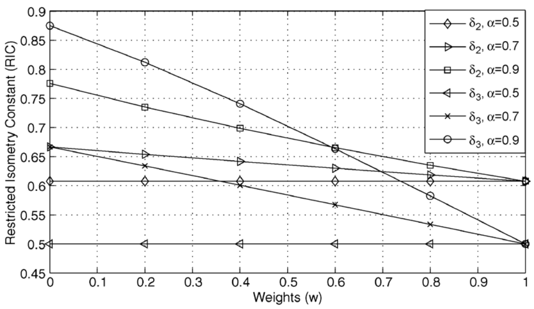

Fig. 1. Comparisons of the RICs δ_2 and δ_3 assuming $l = k$.

when $p \leq 5\beta^2$. Note that (33) is not applicable for $p < 1$. Hence we finally propose the sufficient condition modified from (31) as follows: $\delta_{(p+1)k} < \delta_1$ for $p \in (0, 1]$, and $\delta_{(p+1)k} < \max(\delta_2, \delta_3)$ for $p > 1$.

IV. Extensions

In this section we consider two different noise models: $\|\mathbf{A}^T\mathbf{e}\|_\infty \leq \eta$ and Gaussian noise.

Theorem 4.1: For a signal $\mathbf{x} \in \mathbb{R}^n$, suppose T_0 such that $|T_0| = k$ contains k indices of largest-in-magnitude entries of $\mathbf{x}$ and $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$ with $\mathbf{e}$ satisfying $\|\mathbf{A}^T\mathbf{e}\|_\infty \leq \eta$. Assume the condition (4) holds with a, b and γ defined as previously. We have that, the solution $\mathbf{x}^*$ to the problem

$$\min \|\mathbf{u}\|_{1,\mathbf{w}} \quad s.t. \left\|\mathbf{A}^T(\mathbf{A}\mathbf{u} - \mathbf{y})\right\|_\infty \leq \eta \tag{34}$$

obeys

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_3\eta + C_2E \tag{35}$$

where C_2, E are given in (7) and $C_3 = 2\sqrt{(1+\gamma^2)(k+a)}/1 - \delta_{k+a} - \gamma\theta_{k+a,b}$.

Proof: By the assumption $\|\mathbf{A}^T\mathbf{e}\|_\infty \leq \eta$, we have that

$$\left\|\mathbf{A}_{T_0\cup T_*}^T(\mathbf{A}\mathbf{x}^* - \mathbf{y} + \mathbf{e})\right\|_\infty \leq 2\eta.$$

Consequently,

$$\begin{aligned} |\langle \mathbf{A}\mathbf{h}, \mathbf{A}\mathbf{h}_{T_0\cup T_*}\rangle| &= |\langle \mathbf{A}\mathbf{x}^* - \mathbf{y} + \mathbf{e}, \mathbf{A}_{T_0\cup T_*}\mathbf{h}_{T_0\cup T_*}\rangle| \\ &= \left\|\mathbf{A}_{T_0\cup T_*}^T(\mathbf{A}\mathbf{x}^* - \mathbf{y} + \mathbf{e})\right\|_2 \|\mathbf{h}_{T_0\cup T_*}\|_2 \\ &\leq \sqrt{k+a}\left\|\mathbf{A}_{T_0\cup T_*}^T(\mathbf{A}\mathbf{x}^* - \mathbf{y} + \mathbf{e})\right\|_\infty \|\mathbf{h}_{T_0\cup T_*}\|_2 \\ &\leq 2\eta\sqrt{k+a}\|\mathbf{h}_{T_0\cup T_*}\|_2. \end{aligned} \tag{36}$$

As previously, condition (4) must be satisfied in order to bound $\|\mathbf{h}_{T_0\cup T_*}\|_2$ from above. Thus, by combining (27) and (36),

$$\|\mathbf{h}_{T_0\cup T_*}\|_2 \leq \frac{2\theta_{k+a,b}E/\sqrt{b} + 2\eta\sqrt{k+a}}{1 - \delta_{k+a} - \gamma\theta_{k+a,b}}. \tag{37}$$

By combining (37) and (30), the theorem is proved. ■

Next we consider Gaussian noise case where we let $\mathbf{e} \sim N(0, \sigma^2\mathbf{I}_m)$ for some $\sigma \geq 0$. Let us define the following sets:

$$\mathfrak{B}_1 := \left\{\mathbf{x} : \|\mathbf{x}\|_2 \leq \sigma\sqrt{m + 2\sqrt{m\ln m}}\right\},$$
$$\mathfrak{B}_2 := \left\{\mathbf{x} : \|\mathbf{A}^T\mathbf{x}\|_\infty \leq \sigma\sqrt{2\ln n}\right\}.$$

Lemma 4.2: (Cai [3]) The noise $\mathbf{e} \sim N(0, \sigma^2\mathbf{I}_m)$ satisfies

$$\Pr(\mathbf{e} \in \mathfrak{B}_1) \geq 1 - \frac{1}{m}, \quad \Pr(\mathbf{e} \in \mathfrak{B}_2) \geq 1 - \frac{2}{\sqrt{\pi\ln n}}.$$

Suppose $\mathbf{x}_{\mathfrak{B}_1}^*$ is the solution of

$$\min_{\mathbf{u}} \|\mathbf{u}\|_{1,\mathbf{w}} \quad \text{subject to } \mathbf{y} - \mathbf{A}\mathbf{u} \in \mathfrak{B}_1, \tag{38}$$

and $\mathbf{x}_{\mathfrak{B}_2}^*$ is the solution of

$$\min_{\mathbf{u}} \|\mathbf{u}\|_{1,\mathbf{w}} \quad \text{subject to } \mathbf{y} - \mathbf{A}\mathbf{u} \in \mathfrak{B}_2. \tag{39}$$

A corollary directly follows from the previous theorems.

Corollary 4.3: For a signal $\mathbf{x} \in \mathbb{R}^n$, suppose T_0 such that $|T_0| = k$ contains k indices of largest-in-magnitude entries of $\mathbf{x}$ and $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$ with $\mathbf{e} \sim N(0, \sigma^2\mathbf{I}_n)$. Assume the condition (4) holds with a, b and γ defined as previously. We have that, with probability at least $1 - (1/m)$,

$$\|\mathbf{x} - \mathbf{x}_{\mathfrak{B}_1}^*\|_2 \leq C_1\sigma\sqrt{m + 2\sqrt{m\ln m}} + C_2E. \tag{40}$$

Also with probability at least $1 - (2/\sqrt{\pi\ln n})$,

$$\|\mathbf{x} - \mathbf{x}_{\mathfrak{B}_2}^*\|_2 \leq C_3\sigma\sqrt{2\ln n} + C_2E \tag{41}$$

where C_1, C_2 and E (resp. C_3) are given in Theorem 2.1 (resp. Theorem 4.1).

V. Conclusions

In this letter, we provided a generalized form of sufficient conditions for stable recovery of sparse and/or compressible signals in the presence of noise via weighted ℓ_1-minimization. The key technique lies on the combined use of Ozeki's inequality and shifting inequality. Based on our findings, we proposed improved sufficient conditions for stable recovery.

References

[1] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
[2] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Math.*, vol. 346, no. 9–10, pp. 589–592, 2008.
[3] T. Cai, G. Xu, and J. Zhang, "On recovery of sparse signals via ℓ_1 minimization," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3388–3397, 2009.
[4] T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1300–1308, 2010.
[5] E. Candès, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Applicat.*, vol. 14, no. 5, pp. 877–905, 2008.
[6] M. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1122–1134, 2012.
[7] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
[8] N. Ozeki, "On the estimation of the inequality by the maximum," *J. College Arts, Chiba Univ.*, vol. 5, no. 2, pp. 199–203, 1968.
[9] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, 2007.