Stable Recovery of Signals from Highly Corrupted Measurements

Ningning Li¹, Wengu Chen² and Peng Li³†

Abstract—In this paper, we consider the stable recovery of sparse or proximally sparse signals \( x \in \mathbb{R}^n \) from highly corrupted linear measurements \( b = Ax + f + e \), where \( f \in \mathbb{R}^m \) is a sparse error vector whose nonzero entries may be arbitrarily large and \( e \in \mathbb{R}^m \) is a stochastic noise. We propose an extended Dantzig selector model which considers sparsity of both \( x \) and \( f \). We establish sufficient conditions under the restricted isometry property, which guarantee the signal stable recovery from extended Dantzig selector model and extended Lasso model respectively.

Index Terms—Corrupted compressed sensing, Dantzig selector, Lasso, Restricted isometry property.

I. INTRODUCTION

In the last dozen years, sparse signal recovery has attracted much attention in applied mathematics, statistics and electrical engineering. Increasing efforts have been devoted to initiating and developing compressed sensing (CS) theory. The CS theme is to recover a high dimensional sparse signal \( x \in \mathbb{R}^n \) from a small number of linear measurements

\[
b = Ax + e,
\]

where \( A \in \mathbb{R}^{m \times n}(m < n) \) is the sensing matrix, \( e \in \mathbb{R}^m \) denotes the measurement noise and \( b \in \mathbb{R}^m \) is the observed data. The goal is to reconstruct the signal \( x \) exactly or stably from the sensing matrix \( A \) and observed data \( b \). This is now well established by the following optimization problem

\[
\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad b - Ax \in \mathcal{B},
\]

where \( \mathcal{B} \) is a set determined by the error structure. In the noiseless case \( \mathcal{B} = \{0\} \), (2) becomes

\[
\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad Ax = b,
\]

which was called basis pursuit (BP) [11].

In the presence of noise, Candés, Romberg and Tao [2] and Donoho, Elad and Temlyakov [12] proposed the following method

\[
\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad \|b - Ax\|_2 \leq \eta
\]

for some constant \( \eta > 0 \), which is called quadratically constrained basis pursuit (QCBP). It has been well known that if \( A \) satisfies restricted isometry property (RIP) [7], then the linear program (3) can faithfully recover \( x \). A matrix \( A \) is said to satisfy RIP when there is some \( \delta \in [0, 1) \) such that for any \( x \) with \( |\text{supp}(x)| \leq s \),

\[
(1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2.
\]

Another type of \( \ell_1 \)-minimization method is the Dantzig selector [8], i.e.,

\[
\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad \|A^T (Ax - b)\|_{\infty} \leq \eta,
\]

There are plenty of works about models (4) and (5) and readers can refer to [3, 4, 5, 6, 31, 16]. The problem (4) can be equivalently converted into an unconstrained formulation as follows

\[
\min_{x \in \mathbb{R}^n} \rho \|x\|_1 + \frac{1}{2}\|Ax - b\|_2^2,
\]

where \( \rho > 0 \) is a regularization parameter. This Lasso model was introduced in [24]. There are many works about this model. Readers can refer to [12, 13, 1, 10, 25, 19, 21, 28, 30, 16].

But there are some limitations in the previous models. Let \( \tilde{x} \) be the original signal of (1). It has been proved that the solution of (4) is a stable solution of (1) when the sensing matrix \( A \) satisfies the RIP and \( \eta \) is relatively small [2, 12]. In practical settings, actually, noise \( e \) could be generated by measurement hardware, or the signal was contaminated during transmission, etc. Therefore the noise may not be ignored. The authors of [2] proved that as the noise energy gets larger, the solution of (4) might be very different from \( \tilde{x} \). To overcome this limitation, Wright et al [26, 27] proposed the following extended \( \ell_1 \)-minimization problem

\[
\min_{x \in \mathbb{R}^n, f \in \mathbb{R}^m} \|x\|_1 + \|f\|_1 \quad \text{subject to} \quad Ax + f = b,
\]

where \( f \) is the noise that can not be ignored. When there exists another bounded noise \( e \), the observation \( b \) can be obtained by

\[
b = Ax + f + e = [A, I] \begin{bmatrix} x \\ f \end{bmatrix} + e.
\]

The problem of recovering sparse or compressible \( \{x, f\} \) from (8) is called as corrupted compressed sensing. For corrupted CS, several works have been done in the past several years, for example Li [17] considered to recover \( x \) from (8) by \( \ell_1 \)-minimization method as follows

\[
\min_{x \in \mathbb{R}^n, f \in \mathbb{R}^m} \|x\|_1 + \lambda \|f\|_1 \quad \text{subject to} \quad \|b - Ax - f\|_2 \leq \eta,
\]

where \( \lambda \geq 0 \) is a balance parameter. And Li [17] established the sufficient condition of stable recovery under generalized restricted isometry property (see Definition 2). At the same
time, Nguyen and Tran [20] extended the classic Lasso (6) to corrupted CS as follows

$$\min_{x \in \mathbb{R}^n, f \in \mathbb{R}^m} \|x\|_1 + \lambda \|f\|_1 + \frac{1}{2\rho} \|Ax + f - b\|_2^2,$$

where $\rho > 0$ is a regularization parameter and $\lambda \geq 0$ is a balance parameter. It was called extended Lasso [20]. Recovery guarantees based on an extended restricted eigenvalue condition of the matrix $A$ and bounds for the parameters $\lambda$ and $\rho$ are studied in [20]. Readers can refer to [15, 22, 20, 17, 23, 14, 29] to see more works or a survey [9] on corrupted CS.

It is well known that the Dantzig selector (5) relates closely to Lasso (6). In some sense, Lasso estimator and Dantzig selector does not force the residual $Ax + f - b$ to behave like Gaussian noise. Therefore, the extended Dantzig selector has a wide range of potential applications, especially in statistics. We will establish the generalized restricted isometry property characterization of the extended Dantzig selector model and the extended Lasso model.

The contributions of this paper are summarized as follows:

1. We propose an extended Dantzig selector model which considers the sparsity of both $x$ and $f$.
2. Our results provide the recovery guarantee for proximately sparse signals $x$, and Li [17] only considered the sparse signal recovery. And if we take $c_1 = c_2 = 2$ in Theorem 1, then our condition $\delta_{2(1 + \epsilon)} < 1/(2c_1c_2 + 1) = 1/9$ coincides with that of [17, Lemma 2.3] in some sense.
3. In Theorem 2, we are always able to take $c_1 \geq 2, c_2 \geq 1$ such that $1 \in \left[\frac{1}{c_1^2\sqrt{1} + c_2^2\sqrt{1}}, \frac{1}{c_2^2\sqrt{1}}\right]$. Although the condition $\delta_{2(1 + \epsilon)} < 1/12c_1c_2$ is stronger than the sufficient condition $\delta_{2(1 + \epsilon)} < 0.2$ [21, Theorem 3], our result is more general in some sense.

The organization of this paper is arranged as follows. In Section II, we introduce some definitions and give some supporting lemmas. In Section III, we establish the sufficient condition to recover signals from extended Dantzig selector model (11). In Section IV, we give the sufficient condition to recover signals from extended Lasso model (10).

Throughout the article, we use the following basic notations. If $x \in \mathbb{R}^n$, let $x_S$ be a vector of $\mathbb{R}^n$, the $i$-th component of $x_S$ equals to $x_i$ for $i \in S$ and zero otherwise. For $x \in \mathbb{R}^n$, denote $x_{\text{max}}(s)$ as the vector $x$ with all but the last $s$ entries in absolute value set to zero, and $x_{\text{max}}(s) = x - x_{\text{max}}(s)$. Let $A^T$ denote the transpose of matrix $A$. And $I$ denotes an identity matrix. We use boldfaced letter to denote matrix or vector.

II. DEFINITIONS AND SUPPORTING LEMMAS

In this section, we first recall some basic definitions.

**Definition 1.** The support of a vector $x \in \mathbb{R}^n$ is the index set of its nonzero entries, i.e.,

$$\text{supp}(x) := \{ j \in [n] : x_j \neq 0 \}.$$ 

The vector $x \in \mathbb{R}^n$ is called $s$-sparse if at most $s$ of its entries are nonzero, i.e.,

$$\|x\|_0 := |\text{supp}(x)| \leq s.$$ 

To solve the system (8), the generalized RIP was introduced in [17].

**Definition 2.** For any matrix $\Phi \in \mathbb{R}^{m \times (n + m)}$, the $(s, t)$- RIP-constant $\delta_{s, t}$ is defined as the infimum value of $\delta$ such that

$$(1 - \delta) \left( \|x\|_2^2 + \|f\|_2^2 \right) \leq \left\| \Phi x f \right\|_2^2 \leq (1 + \delta) \left( \|x\|_2^2 + \|f\|_2^2 \right),$$

holds for any $x \in \mathbb{R}^n$ with $|\text{supp}(x)| \leq s$ and $f \in \mathbb{R}^m$ with $|\text{supp}(f)| \leq t$.

**A. Auxiliary Lemmas**

In this subsection, we will give some auxiliary lemmas. The first one states that the generalized restricted isometry property can be satisfied by some matrices $\Phi$, which was introduced in [15].

**Lemma 1.** Let matrix $A \in \mathbb{R}^{m \times n}$ with elements $A_{i,j}$ drawn according to $\mathcal{N}(0, 1/m)$ and let matrix $\Omega \in \mathbb{R}^{m \times m}$ with orthonormal columns. If

$$m \geq K_1(s + t) \log \left( \frac{n + m}{s + t} \right)$$

then $[A, \Omega]$ satisfies the $(s, t)$- RIP with probability exceeding $1 - 3e^{-K_2m}$, where $K_1$ and $K_2$ are constants that depends only on the desired RIP constant $\delta$.

The second one provides a way to estimate the inner product $|\Phi w, \Phi y|$ by restricted isometry property, which comes from [17].

**Lemma 2.** For any $x, y, z \in \mathbb{R}^n$ and $f, g \in \mathbb{R}^m$, if $\text{supp}(x) \cap \text{supp}(y) = \emptyset$, $\|x\|_0 + \|y\|_0 \leq s$, and $\text{supp}(f) \cap \text{supp}(g) = \emptyset$, $\|f\|_0 + \|g\|_0 \leq t$, then

$$\left| \Phi \begin{bmatrix} x \\ f \end{bmatrix}, \Phi \begin{bmatrix} y \\ g \end{bmatrix} \right| \leq \delta_{s,t} \sqrt{\|x\|_2^2 + \|f\|_2^2} \sqrt{\|y\|_2^2 + \|g\|_2^2}. \tag{13}$$

To deal with the extended Dantzig selector model (11), we need the following cone constrained inequality.

**Lemma 3.** For any $x, \hat{x} \in \mathbb{R}^n$, $f, \hat{f} \in \mathbb{R}^m$, $z = \hat{x} - x$, $h = f - \hat{f}$, if $(\hat{x}, \hat{f})$ satisfies $\|\hat{x}\|_1 + \lambda \|\hat{f}\|_1 \leq \|x\|_1 + \lambda \|f\|_1$, then for any positive integer $s \leq n$, $t \leq m$,

$$\|z_{\text{max}}(s)\|_1 + \lambda \|h_{\text{max}}(t)\|_1 \leq \|x\|_1 + \lambda \|f\|_1.$$
Therefore, we have
\[
\parallel \hat{x} - x \parallel_1 = 1,
\]
and
\[
\parallel x + z_{max}(s) \parallel_1 + \parallel \hat{x} - x \parallel_1 - 2 \parallel z_{max}(s) \parallel_1 - 2 \parallel \hat{x} - x \parallel_1.
\]
(14)
Similarly, one has
\[
\parallel f \parallel_1 - \parallel \hat{f} \parallel_1 \geq \parallel h_{max(t)} \parallel_1 - \parallel h_{max(t)} \parallel_1 - 2 \parallel f_{-max(t)} \parallel_1.
\]
(15)
Therefore, we have
\[
(\parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1) - (\parallel x \parallel_1 + \lambda \parallel f \parallel_1) \geq \lambda \parallel h_{max(t)} \parallel_1 - 2 \lambda \parallel f_{-max(t)} \parallel_1 + \parallel z_{max}(s) \parallel_1 - 2 \parallel \hat{x} - x \parallel_1.
\]
(16)
where (1) from (14) and (15). Note that \(\parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1 \leq \parallel x \parallel_1 + \lambda \parallel f \parallel_1,\) which implies that
\[
(\parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1) - (\parallel x \parallel_1 + \lambda \parallel f \parallel_1) \leq 0.
\]
(17)
Combining (16) with (17), we have
\[
\parallel z_{max(s)} \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1 \geq \parallel z_{max(s)} \parallel_1 + 2 \parallel \hat{x} - x \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1 + 2 \parallel z_{max(s)} \parallel_1 - 2 \parallel \hat{x} - x \parallel_1.
\]
which finishes the proof.

In order to solve the corresponding problem for extended Lasso model (10), we establish a lemma as follows.

Lemma 4. For \(A \in \mathbb{R}^{m \times n}, \) let \(e = Ax + f - b, \) \(\parallel A^T e \parallel_\infty \leq \frac{\rho}{2} \) and \(\parallel e \parallel_\infty \leq \frac{\rho}{2}. \)If \((\hat{x}, \hat{f})\) is the solution of (10), then
\[
\frac{1}{\rho} \parallel Az + h \parallel_2^2 + \parallel z_{max(s)} \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1 \leq \frac{1}{\rho} \parallel Az + h \parallel_2^2 + \parallel z_{max(s)} \parallel_1 + 2 \parallel \hat{x} - x \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1 + 4 \parallel f_{-max(t)} \parallel_1,
\]
where \(z = \hat{x} - x, \) \(h = \hat{f} - f.\)

Proof: By direct computation, one has
\[
\frac{1}{2\rho} \parallel Az + f - b \parallel_2^2 - \parallel Ax + f - b \parallel_2^2
\]
\[
= \frac{1}{2\rho} \parallel Az + h \parallel_2^2 + \parallel z_{max(s)} \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1
\]
\[
\geq \frac{1}{2\rho} \parallel Az + h \parallel_2^2 - \frac{1}{\rho} \parallel z \parallel_1 \parallel A^T e \parallel_\infty - \frac{1}{\rho} \parallel h \parallel_1 \parallel e \parallel_\infty
\]
\[
\geq \frac{1}{2\rho} \parallel Az + h \parallel_2^2 - \frac{1}{2} (\parallel z_{max(s)} \parallel_1 + \parallel z_{max(s)} \parallel_1)
\]
\[
- \frac{\lambda}{2} (\parallel h_{max(t)} \parallel_1 + \parallel h_{max(t)} \parallel_1),
\]
(18)
where (1) is due to \(Ax + f - b = e, \) (2) follows from \(\langle Az, e \rangle = \langle z, A^T e \rangle \geq -\parallel z \parallel_1 \parallel A^T e \parallel_\infty \) and \((h, e) \geq -\parallel h \parallel_1 \parallel e \parallel_\infty, \) and (3) is from \(\parallel z \parallel_1 = \parallel z_{max(s)} \parallel_1 + \parallel z_{max(s)} \parallel_1 \) and \(\parallel h \parallel_1 = \parallel h_{max(t)} \parallel_1 + \parallel h_{max(t)} \parallel_1.\)

By \((\hat{x}, \hat{f})\) is the solution of (10), one has
\[
0 \geq \parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1 + \frac{1}{2\rho} \parallel Az + f - b \parallel_2^2
\]
\[
- \parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1 + \frac{1}{2\rho} \parallel Az + f - b \parallel_2^2
\]
\[
= (\parallel \hat{x} \parallel_1 - \parallel x \parallel_1) + \lambda (\parallel f \parallel_1 - \parallel f \parallel_1)
\]
\[
+ \frac{1}{2\rho} (\parallel Az + f - b \parallel_2^2 - \parallel Ax + f - b \parallel_2^2).
\]
(19)
By the above inequalities, we have
\[
0 \geq \parallel \hat{x} \parallel_1 + \lambda \parallel f \parallel_1 + \frac{1}{2\rho} \parallel Az + h \parallel_2^2 - \frac{1}{2} \parallel z_{max(s)} \parallel_1 + \parallel z_{max(s)} \parallel_1
\]
\[
+ \frac{1}{2\rho} (\parallel Az + h \parallel_2^2 - \parallel Ax + f - b \parallel_2^2).
\]
(20)
where (1) is from (14), (15) (18) with (19). Therefore
\[
\frac{1}{\rho} \parallel Az + h \parallel_2^2 + \parallel z_{max(s)} \parallel_1 + \lambda \parallel h_{max(t)} \parallel_1 \leq (3\parallel z_{max(s)} \parallel_1 + 4 \parallel \hat{x} - x \parallel_1)
\]
\[
+ \lambda (3\parallel h_{max(t)} \parallel_1 + 4 \parallel f_{-max(t)} \parallel_1),
\]
which completes the proof.

III. STABLE RECOVERY VIA EXTENDED DANTZIG SELECTOR MODEL

In this section, we consider to recover signals from extended Dantzig selector model (11).

Theorem 1. Let \(\lambda \in \left[\frac{1}{c_1^2} \sqrt{\frac{1}{\epsilon^2} + c_2^2}, \sqrt{\frac{1}{\epsilon^2} + c_2^2} \right] \) with \(c_1 \geq 1, c_2 \geq 1. \) If \(\Phi = [A, I] \) satisfies \((2s, 2t)\)-RIP with
\[
\delta = \delta_{2s,2t} + 2c_1c_2\delta_{2s,2t} < 1,
\]
then the solution \((\hat{x}, \hat{f})\) of optimization problem (11) satisfies
\[
\parallel \hat{x} - \hat{x} \parallel_2^2 + \parallel \hat{f} - f \parallel_2^2 \leq \frac{2}{\sqrt{\epsilon^2 + 1}} \left(1 + \frac{1}{\sqrt{\epsilon^2 + 1}} \right)
\]
\[
x - \hat{x} \parallel_\infty + \frac{\parallel \hat{f} - f \parallel_\infty}{\sqrt{\epsilon^2 + 1}}.
\]

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
where $\tilde{x} \in \mathbb{R}^n$ is the signal we want to recover, and $b = A\tilde{x} + f + e$ satisfies $||A^t e||_{\infty} \leq \eta$.

Proof: Please see Appendix A.

Corollary 1. Let $\lambda \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$ with $c_1 \geq 1, c_2 \geq 1$. If $\tilde{x} \in \mathbb{R}^n$ is $s$-sparse, $f \in \mathbb{R}^m$ is $t$-sparse, and $\Phi = [A, I]$ satisfies $(2s, 2t)$-RIP with

$$\tilde{\delta} = \delta_{2s, 2t} + 2c_1c_2\delta_{2s, 2t} < 1,$$

then the solution $(\tilde{x}, \tilde{f})$ of optimization problem (11) satisfies

$$\sqrt{||\tilde{x} - \hat{x}||^2 + ||\tilde{f} - \hat{f}||^2} \leq \frac{2\sqrt{2}s + t}{1 - \delta} \eta.$$ 

By Lemma 1 and Corollary 1, we have the following conclusion.

Corollary 2. Let $\lambda \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$ with $c_1 \geq 1, c_2 \geq 1$. If $\tilde{x} \in \mathbb{R}^n$ is $s$-sparse, $f \in \mathbb{R}^m$ is $t$-sparse, and

$$m \geq K_1(s + t) \log \left(\frac{n + m}{s + t}\right),$$

then $(\tilde{x}, \tilde{f})$ can be stably recovered via (11) with probability exceeding $1 - 3e^{-K_2m}$, where $K_1$ and $K_2$ are constants that depend only on the desired $c_1, c_2$ and matrix $A$.

Remark 1. If we take $c_1 = c_2 = 2$, then our condition $\delta_{2s, 2t} < 1/(2c_1c_2 + 1) = 1/9$ coincides with that of [17, Lemma 2.3]. But Theorem 1 provides the conclusion for proximately sparse signal $x$, and Li [17] only considered the sparse signal recovery.

IV. STABLE RECOVERY VIA EXTENDED LASSO MODEL

In this section, we consider to recover signals via extended Lasso model (10).

Theorem 2. Let $\lambda \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$ with $c_1 \geq 1, c_2 \geq 1$, $||A^t e||_{\infty} \leq \frac{\eta}{2}$ and $||e||_{\infty} \leq \frac{\hat{\delta}A}{2}$. If $\Phi = [A, I]$ satisfies $(2s, 2t)$-RIP with

$$\hat{\delta} = 12c_1c_2\delta_{2s, 2t} < 1,$$

then the solution $(\hat{x}, \hat{f})$ of (10) satisfies

$$\sqrt{||x - \hat{x}||^2 + ||f - \hat{f}||^2} \leq \frac{3\sqrt{2}c_1c_2 + 1}{2\varepsilon(1 - \delta)} \rho,$$

where

$$\varepsilon = \frac{2}{3\sqrt{2} + \delta_{2s, 2t}} \frac{6c_1c_2 - 1}{2} \delta_{2s, 2t}.$$ 

Proof: Please see Appendix B.

Corollary 3. Let $\lambda \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$ with $c_1 \geq 1, c_2 \geq 1$, $||A^t e||_{\infty} \leq \frac{\eta}{2}$ and $||e||_{\infty} \leq \frac{\hat{\delta}A}{2}$. If $\tilde{x} \in \mathbb{R}^n$ is $s$-sparse, $f \in \mathbb{R}^m$ is $t$-sparse, $b = A\tilde{x} + f + e$, and $\Phi = [A, I]$ satisfies $(2s, 2t)$-RIP with

$$\delta = 12c_1c_2\delta_{2s, 2t} < 1,$$

then the solution $(\hat{x}, \hat{f})$ of (10) satisfies

$$\sqrt{||x - \hat{x}||^2 + ||f - \hat{f}||^2} \leq \frac{3\sqrt{2}c_1c_2 + 1}{2\varepsilon(1 - \delta)} \rho,$$

where

$$\varepsilon = \frac{2}{3\sqrt{2} + \delta_{2s, 2t}} \frac{6c_1c_2 - 1}{2} \delta_{2s, 2t}.$$ 

By Lemma 1 and Corollary 3, we have the following conclusion.

Corollary 4. Let $\lambda \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$ with $c_1 \geq 1, c_2 \geq 1$. If $x \in \mathbb{R}^n$ is $s$-sparse, $f \in \mathbb{R}^m$ is $t$-sparse, and

$$m \geq K_1(s + t) \log \left(\frac{n + m}{s + t}\right),$$

then the solution $(\hat{x}, \hat{f})$ of (10) with probability exceeding $1 - 3e^{-K_2m}$, where $K_1$ and $K_2$ are constants that depend only on the desired $c_1, c_2$ and matrix $A$.

Remark 2. We notice that Shen, Han and Braverman [21] considered a special extended Lasso (in the case of ignoring the tight frames),

$$\min_{x \in \mathbb{R}^n, f \in \mathbb{R}^m} ||x||_1 + ||f||_1 + \frac{1}{2\rho} \|Ax + f - b\|_2^2,$$

(21)

i.e., (10) with $\lambda = 1$. In Theorem 2, we are always able to take $c_1 \geq 1, c_2 \geq 1$ such that $1 \in [\frac{1}{c_1}, c_2] \frac{\sqrt{T}}{2}$. Although the condition $\delta_{2s, 2t} < 1/(12c_1c_2)$ is stronger than their sufficient condition $\delta_{2(s+t)} < 0.2$, our result is more general in some sense.

V. CONCLUSIONS AND DISCUSSION

In this paper, we consider the stable recovery of sparse or proximately sparse signals $x \in \mathbb{R}^n$ from highly corrupted measurements. We propose an extended Dantzig selector model (11) and give a sufficient condition under generalized restricted isometry property to guarantee signal stable recovery (Theorem 1). We also consider to recover signals from extended Lasso model (10). And we establish a sufficient condition guaranteeing the stable recovery of sparse or approximated sparse signals (Theorem 2).

We notice that Li, Sun and Chi [18] considered that the set of $m$ measurements may be corrupted by either arbitrary outliers $f \in \mathbb{R}^m$ or bounded noise $e \in \mathbb{R}^m$, which can be represented as

$$b = A(X) + f + e,$$
where $X \in \mathbb{R}^{n \times n}$ is a rank-$r$ positive semidefinite matrix. The linear mapping $A : \mathbb{R}^{n \times n} \to \mathbb{R}^n$ is defined as $[A(X)]_j = \langle a_j, a_j^T, X \rangle$, where $a_j \in \mathbb{R}^n$ is the $j$th sensing vector composed of i.i.d. standard Gaussian entries. The vector $f$ denotes the outlier vector, which is assumed to be sparse whose entries can be arbitrarily large. And the vector $e$ denotes the additive noise, which is assumed bounded. This problem is called low-rank positive semidefinite matrix recovery from corrupted rank-one measurements. Therefore, one of the future works is to recover low-rank matrix from corrupted measurements.

**Acknowledgement:** Wengu Chen is supported by Natural Science Foundation of China (No.11871109), NSF (Grant No.U1830107) and Science Challenge Project (TZ2018001).

**APPENDIX A
PROOF OF THEOREM 1**

Before proving main theorems, we first state a technical estimate. For any vector $z$, let $V_0$ be an index set with $|V_0| = s$ satisfying $\text{supp}(z_{\max(s)}) \subseteq V_0$ and $V_0' = \bigcup_{j=1}^s V_j$, where $V_1$ is the index set of the $s$ largest entries of $z_{\max(s)}$, $V_2$ is the index set of the next $s$ largest entries of $z_{\max(s)}$, and so on. We divide $V_0'$ into $l_2$ sets, $|T_j| = t$, $1 \leq j < l_2$ and $|T_{l_2}| = t$. We set

$$w := \begin{bmatrix} z \\ h \end{bmatrix}.$$ 

Let

$$w_{01} := w_0 + w_1 = \begin{bmatrix} z_{V_0} \\ h_{T_0} \end{bmatrix} + \begin{bmatrix} z_{V_1} \\ h_{T_1} \end{bmatrix} =: \begin{bmatrix} z_{V_{01}} \\ h_{T_{01}} \end{bmatrix},$$

and

$$w_{02} = w - w_{01}.$$ 

Owing to

$$||w||_2 \leq ||w_{01}||_2 + ||w_{02}||_2,$$ (25)

we first estimate $||w_{01}||_2$. We consider the following identity

$$||\Phi w_{01}, \Phi w|| = ||\Phi w_0, \Phi w_{01}|| + ||\Phi w_{01}, \Phi w_{02}||.$$ (26)

Firstly, we estimate the lower bound of $||\Phi w_0, \Phi w||$. By (26), we know

$$||\Phi w_{01}, \Phi w|| \geq ||\Phi w_{01}||_2 - ||\Phi w_{01}, \Phi w_{02}||,$$ (27)

where (1) follows from the fact that $||z_{V_0}||_2 \leq 2s$, $||h_{T_0}||_2 \leq 2t$ and $\Phi$ satisfies $(2s, 2t)$-RIP.

Let $l = \max\{l_1, l_2\}$ and $z = \sum_{j=0}^l z_{V_j}$ with

$$z_{V_j} = \begin{cases} z_{V_j}, & \text{if } j = 0, 1, \ldots, l_1 \\ 0, & \text{if } j = l_1 + 1, \ldots, l \end{cases}.$$ 

And we denote $\bar{h} = \sum_{j=0}^l h_{T_j}$ in the same way. Then we can denote $\bar{w} = \sum_{j=0}^l w_j$ with

$$\bar{w}_j = \begin{bmatrix} z_{V_j} \\ h_{T_j} \end{bmatrix}.$$ 

Then by the definition of $\bar{w}$, we get

$$||\Phi w_{01}, \Phi w_{02}|| \leq \sum_{k=2}^l \left( ||\Phi w_{01}, \Phi h_{k}|| \right),$$

where $\Phi w_{k} := \Phi w_{0} + \Phi w_{1}$. Then we have

$$\leq \sum_{k=2}^l \left( ||\Phi w_{0}, \Phi h_{k}|| + ||\Phi w_{1}, \Phi h_{k}|| \right).$$ (28)

And from Lemma 3, one has the following cone constraint inequality

$$\left( ||z_{\max(s)}||_1 + \lambda ||h_{\max(s)}||_1 \right) \leq \left( ||z_{\max(s)}||_1 + 2||\bar{h}_{\max(s)}||_1 \right) + \lambda \left( ||h_{\max(s)}||_1 + 2||\bar{f}_{\max(s)}||_1 \right).$$ (24)

Assume $\text{supp}(z_{\max(s)}) \subseteq V_0$ and $\text{supp}(h_{\max(s)}) \subseteq T_0$. We divide $V_0' = \bigcup_{j=1}^{l_1} V_j$ and $T_0' = \bigcup_{j=1}^{l_2} T_j$ just as done before.

**Theorem 1.** Let $V_0'$ be an index set with $|V_0'| = s$ satisfying $\text{supp}(z_{\max(s)}) \subseteq V_0$ and $V_0' = \bigcup_{j=1}^s V_j$, where $V_1$ is the index set of the $s$ largest entries of $z_{\max(s)}$, $V_2$ is the index set of the next $s$ largest entries of $z_{\max(s)}$, and so on. We divide $V_0'$ into $l_2$ sets, $|T_j| = t$, $1 \leq j < l_2$ and $|T_{l_2}| = t$. We set
where (1) is from \( \|z_{V_0} + z_{V_2}\| \leq 2s \) and \( \|h_{T_0} + h_{T_2}\| \leq 2t \) \((j = 0, 1)\) and the fact that \( \Phi \) satisfies \((2s, 2t)\)-RIP and Lemma 2, and (2) is due to the definitions of \( \hat{z} \) and \( \hat{h} \).

In fact, we also have

\[
\|w_{-01}\| \leq \sum_{i=2}^{l_1} \|z_{V_i}\| + \sum_{j=2}^{l_2} \|h_{T_j}\| \\
\leq \frac{1}{\sqrt{s}} \|z_{V_0}\| + \frac{1}{\sqrt{t}} \|h_{T_0}\| \\
= \frac{1}{\sqrt{s}} (\|z_{\max(s)}\| + \sqrt{\frac{s}{1 - \lambda}} \|h_{\max(t)}\|) \\
= \frac{2c_1}{\sqrt{s}} (\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\|) ,
\]

where (1) is from (22), and (2) is due to \( \lambda \geq \frac{1}{\sqrt{t}} \). Then (24) leads to us that

\[
\|w_{-01}\| \leq \sum_{i=2}^{l_1} \|z_{V_i}\| + \sum_{j=2}^{l_2} \|h_{T_j}\| \\
\leq \frac{c_1}{\sqrt{s}} (\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\|) \\
+ 2\|\hat{x}_{\max(s)}\| + 2\lambda \|\hat{f}_{\max(t)}\| ,
\]

\[
\leq \frac{c_1}{\sqrt{s}} (\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\|) \\
+ 2\|\hat{x}_{\max(s)}\| + \lambda \|\hat{f}_{\max(t)}\| ,
\]

\[
\leq c_1 (\|z_{\max(s)}\| + \lambda \sqrt{\frac{t}{s}} \|h_{\max(t)}\|) \\
+ \frac{c_1}{\sqrt{s}} (\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\|) \\
\leq c_1 c_2 (\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\|) \\
+ 2c_1 \left( \sqrt{\frac{t}{s}} \|\hat{x}_{\max(s)}\| + \lambda \left( \frac{\sqrt{s}}{t} \|\hat{f}_{\max(t)}\| \right) \right) ,
\]

\[
\leq \sqrt{2c_1 c_2} \|w_{01}\| \\
+ 2c_1 c_2 \left( \sqrt{\frac{t}{s}} \|\hat{x}_{\max(s)}\| + \lambda \left( \frac{\sqrt{s}}{t} \|\hat{f}_{\max(t)}\| \right) \right) \\
= \sqrt{2c_1 c_2} \|w_{01}\| + 2c_1 c_2 \alpha ,
\]

where

\[
\alpha = \sqrt{\frac{t}{s}} \|\hat{x}_{\max(s)}\| + \lambda \sqrt{\frac{s}{t}} \|\hat{f}_{\max(t)}\| ,
\]

and (3) comes from \( \|z_{\max(s)}\| \leq \sqrt{s} \|z_{\max(s)}\| \leq 2 \|z_{\max(s)}\| \) and \( \|h_{\max(t)}\| \leq \sqrt{t} \|h_{\max(t)}\| \) and (4) is due to \( \lambda \leq c_2 \sqrt{s/t} \),

and (5) follows from the fact that \( \|z_{\max(s)}\| + \|h_{\max(t)}\| \leq \sqrt{2} \|w_0\| \leq \sqrt{2} \|w_{01}\| \) and \( \lambda \leq c_2 \sqrt{s/t} \). Substituting (29) into (28), we get

\[
\langle \Phi w_{01}, \Phi w_{-01} \rangle \\
\leq 2c_1 c_2 \delta_{2s, 2t} \|w_{01}\| + 2\sqrt{2c_1 c_2 d_{2s, 2t} \alpha} \|w_{01}\| .
\]

Substituting (30) into (27).

\[
\|\Phi w_{01} \| \leq \|\Phi w_{01} \| + \|\Phi w_{-01} \| \\
\leq 2c_1 c_2 \delta_{2s, 2t} \|w_{01}\| + \sqrt{\sqrt{2c_1 c_2 d_{2s, 2t} \alpha}} \|w_{01}\| .
\]

Next, we estimate the upper bound of \( \|\Phi w_{01}, \Phi w\| \) as follows

\[
\|\Phi w_{01}, \Phi w\| = \|\Phi^T \Phi w, w_{01}\| \\
\leq \|A, I^T \Phi w, w_{01}\| \\
= \|A, I^T (Az + h)\| \|w_{01}\| \\
\leq 2\eta \|w_{01}\| \\
\leq 2\sqrt{2} \|w_{01}\| ,
\]

where (1) is from (23) and (2) is from \( \|w_{01}\| \leq 2(s + t) \) and \( \|w_{01}\| \leq 2(1 - \delta) \|w_{01}\| .

Combining the upper bound of (32) with the lower bound (31), one has

\[
(1 - \delta_{2s, 2t} - 2c_1 c_2 \delta_{2s, 2t}) \|w_{01}\| \leq 2\sqrt{2c_1 c_2 d_{2s, 2t} \alpha} \|w_{01}\| \\
\leq 2\sqrt{2} \|w_{01}\| .
\]

Therefore,

\[
\|w_{01}\| \leq \frac{1}{1 - \delta} \|w_{01}\| \\
\leq \frac{1}{1 - \delta} \|w_{01}\| .
\]

Last, we get the estimate of \( \|w\| \) as follows

\[
\|w\| \leq \frac{1}{1 + \sqrt{2c_1 c_2}} \|w_{01}\| + 2c_1 c_2 \alpha \\
\leq (1 + \sqrt{2c_1 c_2}) \left( \frac{2\sqrt{2} \sqrt{s + t}}{1 - \delta} \eta + \frac{2\sqrt{2c_1 c_2 d_{2s, 2t} \alpha}}{1 - \delta} \right) \\
+ 2c_1 c_2 \alpha \\
\leq \frac{2\sqrt{2} \sqrt{s + t}}{1 - \delta} \eta \\
+ \frac{2\sqrt{2c_1 c_2} \sqrt{1 + \sqrt{2c_1 c_2} \delta_{2s, 2t} + 2c_1 c_2}}{1 - \delta} \left( \|\hat{x}_{\max(s)}\| + \|\hat{f}_{\max(t)}\| \right) ;
\]

where (1) comes from (25) and (29), and (2) comes from (33). Thus we finish the proof.

\textbf{APPENDIX B}

\textbf{PROOF OF THEOREM 2}

\textit{Proof}: Let \( z = \hat{x} - \hat{x} \) and \( h = \hat{f} - \hat{f} \). Then by Lemma 4, we get a modified cone constraint inequality as follows

\[
\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\| \\
\leq (3) \|z_{\max(s)}\| + 4 \|\hat{x}_{\max(s)}\| \\
+ \lambda (3) \|h_{\max(t)}\| + 4 \|\hat{f}_{\max(t)}\| ,
\]

\textbf{APPENDIX C}

\textbf{PROOF OF THEOREM 3}

\textit{Proof}: Let \( z = \hat{x} - \hat{x} \) and \( h = \hat{f} - \hat{f} \). Then by Lemma 4, we get a modified cone constraint inequality as follows

\[
\|z_{\max(s)}\| + \lambda \|h_{\max(t)}\| \\
\leq (3) \|z_{\max(s)}\| + 4 \|\hat{x}_{\max(s)}\| \\
+ \lambda (3) \|h_{\max(t)}\| + 4 \|\hat{f}_{\max(t)}\| ,
\]
and
\[ \frac{1}{\rho} \| A z + h \|_2^2 \leq (3\| z_{max}(s) \|_1 + 4\| \bar{x}_{-max}(s) \|_1) + \lambda (3\| h_{max}(t) \|_1 + 4\| \bar{f}_{-max}(t) \|_1), \] (35)
instead of (24) and (23) in the proof of Theorem 1.

As the same as the proof of Theorem 1, in order to estimate \( \| w_{01} \|_2 \), we also consider identity (26). But the lower bound and upper bound for \( \| \Phi w_{01}, \Phi w \| \) are different from that in the proof of Theorem 1.

First, we estimate the lower bound of \( \| \Phi w_{01}, \Phi w \| \). Note that (27) and (28) still hold. We estimate \( \| w_{-01} \|_2 \) as follows
\[
\| w_{-01} \|_2 \leq \sum_{i=2}^{t_1} \| z_{V_i} \|_2 + \sum_{j=2}^{t_2} \| h_{T_j} \|_2 \\
\leq \frac{c_1}{\sqrt{s}} (\| z_{-max}(s) \|_1 + \lambda \| h_{-max}(t) \|_1)
\leq \frac{c_1}{\sqrt{s}} (3\| z_{max}(s) \|_1 + 3\lambda \| h_{max}(t) \|_1)
+ 4\| \bar{x}_{-max}(s) \|_1 + 4\| \bar{f}_{-max}(t) \|_1)
\leq 3c_1 \left( \| z_{max}(s) \|_1 + \lambda \right)
+ 4c_1 \left( \| \bar{x}_{-max}(s) \|_1 + \lambda \| \bar{f}_{-max}(t) \|_1 \right)
\leq (2)
\leq 3c_1 c_2 (\| z_{max}(s) \|_2 + \| h_{max}(t) \|_2)
+ 4c_1 \left( \| \bar{x}_{-max}(s) \|_1 + \lambda \| \bar{f}_{-max}(t) \|_1 \right)
\leq 3\sqrt{2} c_1 c_2 \| w_{01} \|_2
+ 4c_1 c_2 \left( \| \bar{x}_{-max}(s) \|_1 + \| \bar{f}_{-max}(t) \|_1 \right)
\leq 3\sqrt{2} c_1 c_2 \| w_{01} \|_2 + 4c_1 c_2 \alpha,
\] (36)
where
\[ \alpha = \frac{\| \bar{x}_{-max}(s) \|_1 + \| \bar{f}_{-max}(t) \|_1}{\sqrt{t}}, \]
and (1) is due to (34), and (2) comes from \( \| z_{max}(s) \|_1 \leq \sqrt{s}\| z_{max}(s) \|_2 \) and \( \| h_{max}(t) \|_1 \leq \sqrt{t}\| h_{max}(t) \|_2 \), and (3) is due to \( \lambda \leq c_2 \sqrt{s}/t \), and (4) follows the fact that \( \| z_{max}(s) \|_2 + \| h_{max}(t) \| \leq \sqrt{2} \| w_{01} \|_2 \), and \( \lambda \leq c_2 \sqrt{s}/t \).

Substituting (36) into (28), we get
\[
\| \Phi w_{01}, \Phi w \|_{-01} \leq (2)
\leq 6c_1 c_2 \delta_{s,2t} \| w_{01} \|_2 + 4\sqrt{2} c_1 c_2 \delta_{s,2t} \alpha \| w_{01} \|_2, \]
(37)
instead of (30). Therefore,
\[
\| \Phi w_{01}, \Phi w \| \geq (1 - \delta_{s,2t} - 6c_1 c_2 \delta_{s,2t}) \| w_{01} \|_2^2.
\]
which gives a lower bound of \( \| \Phi w_{01}, \Phi w \| \).

Next we estimate the upper bound of \( \| \Phi w_{01}, \Phi w \| \), which is totally different from that in the proof of Theorem 1.
\[
\| \Phi w_{01} \|_2 \leq \| \Phi w \|_2 \| w_{01} \|_2 \leq (1)
\leq A z + h \|_2 \sqrt{1 + \delta_{s,2t}} \| w_{01} \|_2, \] (39)
where (1) is from \( \| \Phi w \|_2 = \| A z + h \|_2 \), and \( \| z_{01} \|_0 \leq 2s \), \( \| h_{T_0} \|_0 \leq 2t \) and the fact \( \Phi \) satisfies \( (2s, 2t) \)-RIP.

In order to estimate the upper bound of \( \| \Phi w_{01}, \Phi w \| \), we need to estimate \( \| A z + h \|_2 \). Note that
\[
\| A z + h \|_2^2 = (1)
\leq \| A \| \| z \|_2 \| h \|_2
\leq (2)
\leq \frac{1}{\rho} \| z \|_2 \| h \|_2
\leq (3)
\| z \|_2 \| h \|_2
\leq (4)
\| z \|_2 \| h \|_2
\leq (5)
\| z \|_2 \| h \|_2
\leq (6)
\| z \|_2 \| h \|_2
\leq (7)
\| z \|_2 \| h \|_2
\leq (8)
\| z \|_2 \| h \|_2
\leq (9)
\| z \|_2 \| h \|_2
\leq (10)
\| z \|_2 \| h \|_2
\leq (11)
\| z \|_2 \| h \|_2
\leq (12)
\| z \|_2 \| h \|_2
\leq (13)
\| z \|_2 \| h \|_2
\leq (14)
\| z \|_2 \| h \|_2
\leq (15)
\| z \|_2 \| h \|_2
\leq (16)
\| z \|_2 \| h \|_2
\leq (17)
\| z \|_2 \| h \|_2
\leq (18)
\| z \|_2 \| h \|_2
\leq (19)
\| z \|_2 \| h \|_2
\leq (20)
\| z \|_2 \| h \|_2
\leq (21)
\| z \|_2 \| h \|_2
which gives an upper bound of $\langle \Phi u_0, \Phi w \rangle$.

By the fact that the lower bound (38) is less than the upper bound (41), we have

\[
(1 - \delta_{2s,2t} - 6c_1c_2\delta_{2s,2t})\|w_0\|^2_2 - 4\sqrt{2}c_1c_2\delta_{2s,2t}\|w_0\|_2
\leq \frac{3\sqrt{2}\sqrt{1 + \delta_{2s,2t}}}{2}\|w_0\|^2_2
+ \left( \frac{\rho}{2\varepsilon} + 2\sqrt{s_c}\varepsilon\alpha \right) \sqrt{1 + \delta_{2s,2t}}\|w_0\|_2,
\]

i.e.,

\[
(1 - \delta_{2s,2t} - 6c_1c_2\delta_{2s,2t})\|w_0\|^2_2 - \frac{3\sqrt{2}\sqrt{1 + \delta_{2s,2t}}}{2}\|w_0\|^2_2
\leq \left( \frac{\rho}{2\varepsilon} + 2\sqrt{s_c}\varepsilon\alpha \right) \sqrt{1 + \delta_{2s,2t}} + 4\sqrt{2}c_1c_2\delta_{2s,2t}\|w_0\|_2.
\]

(42)

Taking

\[
\varepsilon = \frac{2}{3\sqrt{2}\sqrt{1 + \delta_{2s,2t}}\sqrt{s_c}}(6c_1c_2 - 1)\delta_{2s,2t},
\]

then (42) can be formulated as

\[
(1 - \hat{\delta})\|w_0\|^2_2
\leq \left( \frac{\rho}{2\varepsilon} + 2\sqrt{s_c}\varepsilon\alpha \right) \sqrt{1 + \delta_{2s,2t}} + 4\sqrt{2}c_1c_2\delta_{2s,2t}\|w_0\|_2,
\]

(43)

where

\[
\hat{\delta} = 12c_1c_2\delta_{2s,2t}.
\]

Owing to $\delta < 1$, we can get

\[
\|w_0\|^2_2 \leq \frac{\sqrt{1 + \delta_{2s,2t}}}{2\varepsilon(1 - \hat{\delta})}\rho
+ \left( 2\sqrt{s_c}\varepsilon\sqrt{1 + \delta_{2s,2t}} + 4\sqrt{2}c_1c_2\delta_{2s,2t}\right)\|w_0\|_2,
\]

which gives an estimate of $\|w_0\|^2_2$.

Last, we begin to estimate $\|w\|^2_2$.

\[
\|w\|^2_2 \leq \|w_0\|^2_2 + \|w - w_0\|^2_2
\leq (3\sqrt{2}c_1c_2 + 1)\|w_0\|^2_2 + 4c_1c_2\alpha
\leq (3\sqrt{2}c_1c_2 + 1)\sqrt{1 + \delta_{2s,2t}}\|w_0\|^2_2
\leq \frac{\rho}{2\varepsilon(1 - \hat{\delta})}\rho
+ \left( (3\sqrt{2}c_1c_2 + 1)\sqrt{1 + \delta_{2s,2t}} + 4\sqrt{2}c_1c_2\delta_{2s,2t}\right)\|w_0\|_2,
\]

(44)

where (1) is in view of (36), and (2) is by virtue of (44). Thus we finish the proof.

---

References