

STABLE RECOVERY OF SPARSELY CORRUPTED SIGNALS THROUGH JUSTICE PURSUIT DE-NOISING*

Ningning Li and Wengu Chen¹⁾

Institute of Applied Physics and Computational Mathematics, Beijing 100088, China

Email: imliningning@163.com, chenwg@iapcm.ac.cn

Huanmin Ge

Beijing Sport University, Beijing 100084, China

Email: gehuanmin@163.com

Abstract

This paper considers a corrupted compressed sensing problem and is devoted to recover signals that are approximately sparse in some general dictionary but corrupted by a combination of interference having a sparse representation in a second general dictionary and measurement noise. We provide new restricted isometry property (RIP) analysis to achieve stable recovery of sparsely corrupted signals through Justice Pursuit De-Noising (JPDN) with an additional parameter. Our main tool is to adapt a crucial sparse decomposition technique to the analysis of the Justice Pursuit method. The proposed RIP condition improves the existing representative results. Numerical simulations are provided to verify the reliability of the JPDN model.

Mathematics subject classification: 65F22, 68W25, 90C25.

Key words: Justice Pursuit De-Noising, Restricted isometry property, Corrupted compressed sensing, Signal recovery.

1. Introduction

The theory of compressed sensing (CS) has been a very active research field in recent years and has attracted much attention in signal processing, electrical engineering and statistics. CS is concerned with recovering high-dimensional sparse signals from a small number of linear measurements. Specifically, sparse recovery from fewer noiseless observations $\mathbf{y} = \mathbf{A}\mathbf{x}$ in standard CS is in general an ill-posed problem and can be solved by some optimization algorithms. A well-known effective algorithm is Basis Pursuit (BP)

$$\min \|\hat{\mathbf{x}}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\hat{\mathbf{x}}, \quad (1.1)$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a sensing matrix with $m \ll n$, $\mathbf{y} \in \mathbb{R}^m$ is the measurement vector. In the case of bounded measurement noise, one observes $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{e} \in \mathbb{R}^m$ denotes a bounded noise vector with $\|\mathbf{e}\|_2 \leq \varepsilon$. A widely used approach is the following Basis Pursuit De-Noising (BPDN) method [9]:

$$\min \|\hat{\mathbf{x}}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2 \leq \varepsilon. \quad (1.2)$$

A central goal of CS is to recover the unknown sparse or nearly sparse signal $\mathbf{x} \in \mathbb{R}^n$ exactly or stably from the constrained method (1.1) or (1.2) based on \mathbf{A} and \mathbf{y} . Here, a vector $\mathbf{x} \in \mathbb{R}^n$

* Received November 22, 2021 / Revised version received February 7, 2022 / Accepted April 2, 2022 /
Published online February 28, 2023 /

¹⁾ Corresponding author

is called s -sparse if the number of nonzero elements in \mathbf{x} is at most s and a vector $\mathbf{v} \in \mathbb{R}^n$ is said nearly s -sparse if the error of its best s -term approximation decays quickly in s [16]. Many recovery guarantees about the methods (1.1) and (1.2) have been well developed and readers can refer to [5, 8–12, 17, 34, 35, 42].

Different from classical CS, corrupted compressed sensing can deal with unbounded noises that appear in many settings, such as impulse noise, narrowband interference, malfunctioning hardware and transmission errors in the case where signals was sent over a noisy channel. In these cases, the measurement error may be sparse or approximately sparse with unbounded value. Its mathematical model can be represented as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e} = [\mathbf{A} \ \mathbf{I}] \begin{bmatrix} \mathbf{x} \\ \mathbf{f} \end{bmatrix} + \mathbf{e}, \tag{1.3}$$

where the sparse corruption vector $\mathbf{f} \in \mathbb{R}^m$ may have extremely large elements and $\mathbf{I} \in \mathbb{R}^{m \times m}$ denotes the identity matrix. Several papers [7, 22, 26, 27, 30, 36, 38] considered the following constraint problem for the model (1.3):

$$\min_{\hat{\mathbf{x}} \in \mathbb{R}^n, \hat{\mathbf{f}} \in \mathbb{R}^m} \|\hat{\mathbf{x}}\|_1 + \lambda \|\hat{\mathbf{f}}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - (\mathbf{A}\hat{\mathbf{x}} + \hat{\mathbf{f}})\|_2 \leq \varepsilon, \tag{1.4}$$

where $\lambda > 0$ is a balance parameter. For the model (1.3), Lin *et al.* [27] proposed new algorithms for reconstructing signals that are nearly sparse in terms of a tight frame in the presence of bounded noise combined with sparse noise, and presented corresponding recovery guarantees. Li *et al.* [24] established sufficient conditions based on the restricted isometry property, which guarantee stable signal recovery from extended Dantzig selector and extended Lasso models. Foygel and Mackey in [15] used a convex programming method to recover \mathbf{x} and \mathbf{f} with or without prior information and provided new bounds for the Gaussian complexity of sparse signals, leading to a sharper recovery guarantee. Adcock *et al.* in [1] showed that the signal \mathbf{x} and its corruption \mathbf{f} can be recovered stably if the matrix \mathbf{A} satisfies the generalized RIP of order $(2s_1, 2s_2)$ with

$$\delta_{2s_1, 2s_2} < \frac{1}{\sqrt{1 + \left(\frac{1}{2\sqrt{2}} + \sqrt{\eta}\right)^2}}, \quad \eta = \frac{s_1 + \lambda^2 s_2}{\min\{s_1, \lambda^2 s_2\}},$$

(s_1 and s_2 are the sparsity of \mathbf{x} and \mathbf{f} , respectively) and the generalized RIP will be defined below (see Definition 2.2).

Note that the measurement corruption may be sparse in some bases. Mathematically, we have

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{f} = \mathbf{\Phi} \begin{bmatrix} \mathbf{x} \\ \mathbf{f} \end{bmatrix}, \tag{1.5}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times l}$, $\mathbf{\Phi} = [\mathbf{A} \ \mathbf{B}]$, $\mathbf{x} \in \mathbb{R}^n$ is a sparse unknown signal, $\mathbf{f} \in \mathbb{R}^l$ is a sparse corruption. For instance, when the measurements are corrupted by 60Hz hum, the corruption noise is sparse in the discrete Fourier basis. To reconstruct \mathbf{x} (and \mathbf{f}) from the model (1.5), a popular Justice Pursuit (JP) method has been introduced by Laska *et al.* [23]

$$\min_{\hat{\mathbf{x}} \in \mathbb{R}^n, \hat{\mathbf{f}} \in \mathbb{R}^l} \|\hat{\mathbf{x}}\|_1 + \|\hat{\mathbf{f}}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{f}}. \tag{1.6}$$

Laska *et al.* [23] demonstrated that the algorithm JP (1.6) can achieve exact recovery from measurements corrupted with sparse noise. The references [21, 31–33] considered deterministic and coherence-based results for the model (1.5). Studer and Baraniuk in [32] obtained effective recovery and separation of signals through the optimization method (1.6) based on the support of \mathbf{x} and \mathbf{f} and the coherent parameters of matrices \mathbf{A} and \mathbf{B} .

Besides, in the presence of additive white Gaussian noise or quantization error, namely,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{f} + \mathbf{e}, \quad (1.7)$$

Laska *et al.* [23] also investigated Justice Pursuit De-Noising (JPDN) method as follows:

$$\min_{\hat{\mathbf{x}} \in \mathbb{R}^n, \hat{\mathbf{f}} \in \mathbb{R}^l} \|\hat{\mathbf{x}}\|_1 + \|\hat{\mathbf{f}}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{f}})\|_2 \leq \varepsilon. \quad (1.8)$$

When the corruption is sparse on certain sparsifying domain, Zhang *et al.* [41] proved the uniform recovery guarantee of the problem (1.7) for two classes of structured sensing matrices.

In this paper, we study the minimization program JPDN with a parameter λ

$$\min_{\hat{\mathbf{x}} \in \mathbb{R}^n, \hat{\mathbf{f}} \in \mathbb{R}^l} \|\hat{\mathbf{x}}\|_1 + \lambda \|\hat{\mathbf{f}}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - (\mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\hat{\mathbf{f}})\|_2 \leq \varepsilon, \quad (1.9)$$

where the parameter $\lambda \geq 0$ trades off $\|\hat{\mathbf{x}}\|_1$ and $\|\hat{\mathbf{f}}\|_1$. The advantage of using extra λ in (1.9) will be discussed in the main theorem and numerical simulations. The optimization problem (1.9) is suitable for many problems such as saturated or clipped signals [2, 3, 22], corrupted signals [26, 31, 37] and sparsity-based super-resolution [13, 29]. Separation of signals into two distinct components also fits into our framework. Signal separation problems arise in applications such as the separation of texture from cartoon parts in images [4, 14] and the separation of neuronal calcium transients from smooth signals caused by astrocytes in calcium imaging [20]. When $\mathbf{A} = \mathbf{B}$, Lin *et al.* [28] showed that the distinct subcomponents, which are (approximately) sparse in morphologically different (redundant) dictionaries, can be reconstructed by solving the split-analysis algorithm, provided that the dictionaries satisfy a mutual coherence (between the different dictionaries) condition and the measurement matrix satisfies a restricted isometry property adapted to a composed dictionary. And Li *et al.* [25] proposed an iterative hard thresholding algorithm adapted to dictionaries. Then they showed that under the usual assumptions that the measurement system satisfies a restricted isometry property (adapted to a composed dictionary) condition and the dictionaries satisfy a mutual coherence condition, the algorithm can approximately reconstruct the distinct subcomponents after a fixed number of iterations.

The main aim of this paper is to establish new RIP analysis for the optimization problem (1.9). In particular, we provide recovery guarantees based on a key tool established independently in [6, 40]. In classical CS, the technique that signals can be decomposed as a convex combination of sparse signals is a key tool to get recovery guarantees. Especially, Cai and Zhang [6] used this technical tool to establish sharp sufficient RIP conditions for signal recovery via l_1 -minimization. Zhang and Li [43] provided optimal RIP bounds which can guarantee sparse signal recovery via l_p -minimization for $p \in (0, 1]$. Now, we extend this tool to the problem of corrupted CS. Furthermore, we show that the proposed recovery condition in this paper is weaker than that in [26] and also weaker than that in [1] within a certain range of λ .

The rest of the paper is arranged as follows. In Section 2, we introduce the RIP and its generalization and recall some supporting lemmas needed in this paper. We state the main theorem and provide its proof in Section 3. In Section 4, we illustrate the performance of the proposed method by numerical simulations. Conclusions are given in Section 5.

2. Preliminaries

In order to properly state our results, we review the RIP definitions and several key lemmas needed throughout this paper.

2.1. Notations

We use the following notations. Boldface lowercase and uppercase letters stand for vectors and matrices, respectively. For any vector $\mathbf{v} \in \mathbb{R}^n$, we denote $\mathbf{v}_{\max(s)}$ as the vector \mathbf{v} with all but the largest s entries in absolute value set to zero. And $\mathbf{v}_{-\max(s)} = \mathbf{v} - \mathbf{v}_{\max(s)}$.

2.2. Definitions and supporting lemmas

First, we review the restrict isometry property (RIP) introduced by Candés and Tao in [9].

Definition 2.1. For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and an integer $1 \leq s \leq n$, \mathbf{A} is said to satisfy the RIP of order s if there exists a constant δ_s such that

$$(1 - \delta_s)\|\mathbf{v}\|_2^2 \leq \|\mathbf{A}\mathbf{v}\|_2^2 \leq (1 + \delta_s)\|\mathbf{v}\|_2^2 \tag{2.1}$$

holds for all s -sparse signals $\mathbf{v} \in \mathbb{R}^n$. The smallest constant δ_s is called the restricted isometry constant (RIC) of order s for \mathbf{A} .

The generalized RIP was introduced by Li in [26], which plays an important role in corrupted compressed sensing.

Definition 2.2. For a matrix $\Phi \in \mathbb{R}^{m \times (n+l)}$ and integers $1 \leq s_1 \leq n$, $1 \leq s_2 \leq l$, define the (s_1, s_2) -order RIP constant δ_{s_1, s_2} as the smallest number δ such that

$$(1 - \delta)\left\| \begin{bmatrix} \mathbf{x} \\ \mathbf{f} \end{bmatrix} \right\|_2^2 \leq \left\| \Phi \begin{bmatrix} \mathbf{x} \\ \mathbf{f} \end{bmatrix} \right\|_2^2 \leq (1 + \delta)\left\| \begin{bmatrix} \mathbf{x} \\ \mathbf{f} \end{bmatrix} \right\|_2^2 \tag{2.2}$$

holds for all $\mathbf{x} \in \mathbb{R}^n$ with $|\text{supp}(\mathbf{x})| \leq s_1$ and all $\mathbf{f} \in \mathbb{R}^l$ with $|\text{supp}(\mathbf{f})| \leq s_2$.

Remark 2.1. For a matrix $\Phi = [\mathbf{A} \ \mathbf{B}] \in \mathbb{R}^{m \times (n+l)}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times l}$, if Φ satisfies $(s_1 + s_2)$ -order RIP, then it also satisfies (s_1, s_2) -order RIP by the Definitions 2.1 and 2.2 and we can get $\delta_{s_1, s_2} \leq \delta_{s_1 + s_2}$.

It is clear from [23] that if matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ have elements \mathbf{A}_{ij} drawn according to $\mathcal{N}(0, 1/m)$, \mathbf{B} is an $m \times l$ matrix with orthonormal columns and the number of measurements satisfy

$$m \geq K_1(s_1 + s_2) \log \left(\frac{n + l}{s_1 + s_2} \right),$$

then $[\mathbf{A} \ \mathbf{B}]$ satisfies the RIP of order (s_1, s_2) with high probability.

As in the compressed sensing, the importance of weakening the RIP or modified RIP conditions is that we could use fewer observations to ensure the signal recovery. For the corrupted sensing, this can be seen from the following results.

Theorem 2.1 ([1, 26]). Suppose $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e}$ with $\|\mathbf{e}\|_2 \leq \varepsilon$. Let $0 < \delta, \varepsilon < 1$, $1 \leq s_1 \leq n$, $1 \leq s_2 \leq m$ and suppose that

$$m \gtrsim \delta^{-2}(s_1 \cdot \log(2n/s_1) + \log(2\varepsilon^{-1})), \tag{2.3}$$

$$m \gtrsim \delta^{-2} \cdot s_2 \cdot \log(\delta^{-1}). \tag{2.4}$$

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be a matrix whose entries are independent Gaussian random variables with mean zero and variance 1. Then with probability at least $1 - \varepsilon$, the matrix $\frac{1}{\sqrt{m}}\mathbf{A}$ has the RIP for the sparse corruptions problem of order (s_1, s_2) with constant $\delta_{s_1, s_2} \leq \delta$.

Theorem 2.1 implies that we can reduce the required number of measurements m by weakening the modified RIP condition.

Zhang *et al.* [41] also presented a bound on the required number of measurements m such that the corresponding matrix Φ has the (s_1, s_2) -RIP constant satisfying $\delta_{s_1, s_2} \leq \delta$ for any $\delta \in (0, 1)$.

Theorem 2.2 ([41]). Suppose $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e}$ with $\Phi = [\mathbf{A} \ \mathbf{I}] \in \mathbb{R}^{m \times (n+m)}$, $\mathbf{A} = \mathbf{U}\mathbf{D}\tilde{\mathbf{B}}$ ($\tilde{\mathbf{B}} \in \mathbb{R}^{\tilde{n} \times n}$, $\tilde{n} \geq n$, represents a column-wise orthonormal matrix, i.e. $\tilde{\mathbf{B}}^T \tilde{\mathbf{B}} = \mathbf{I}$) and $\mu(\mathbf{U}) \sim 1/\sqrt{m}$. If, for $\delta \in (0, 1)$,

$$\begin{aligned} m &\geq c_3 \delta^{-2} s_1 \tilde{n} \mu^2(\tilde{\mathbf{B}}) \log^2 s_1 \log^2 \tilde{n}, \\ m &\geq c_4 \delta^{-2} s_2 \log^2 s_2 \log^2 \tilde{n}, \end{aligned}$$

where c_3 and c_4 are some absolute constants, then with probability at least $1 - \tilde{n}^{-\log^2 s_1 \log \tilde{n}}$, the (s_1, s_2) -RIP constant of Φ satisfies $\delta_{s_1, s_2} \leq \delta$.

The following lemma states the cone constrained inequality (see [8, 26]).

Lemma 2.1. For any $\mathbf{x}, \hat{\mathbf{x}} \in \mathbb{R}^n$, $\mathbf{f}, \hat{\mathbf{f}} \in \mathbb{R}^l$, $\mathbf{z} = \hat{\mathbf{x}} - \mathbf{x}$, $\mathbf{h} = \hat{\mathbf{f}} - \mathbf{f}$, if $\|\hat{\mathbf{x}}\|_1 + \lambda \|\hat{\mathbf{f}}\|_1 \leq \|\mathbf{x}\|_1 + \lambda \|\mathbf{f}\|_1$, then for any positive integers $s_1 \leq n$, $s_2 \leq l$,

$$\begin{aligned} &\|\mathbf{z}_{-\max(s_1)}\|_1 + \lambda \|\mathbf{h}_{-\max(s_2)}\|_1 \\ &\leq \|\mathbf{z}_{\max(s_1)}\|_1 + 2\|\mathbf{x}_{-\max(s_1)}\|_1 + \lambda(\|\mathbf{h}_{\max(s_2)}\|_1 + 2\|\mathbf{f}_{-\max(s_2)}\|_1) \\ &= \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \lambda \mathbf{h}_{\max(s_2)} \end{bmatrix} \right\|_1 + 2 \left\| \begin{bmatrix} \mathbf{x}_{-\max(s_1)} \\ \lambda \mathbf{f}_{-\max(s_2)} \end{bmatrix} \right\|_1. \end{aligned}$$

One inspiration of this paper comes from the sparse representation of a polytope firstly established in [6, 40]. We review the sparse representation of a polytope as in [6], which enables non-sparse vectors to be represented by sparse vectors, and then the RIP condition can be applied to non-sparse vectors.

Lemma 2.2. For a positive number α and a positive integer s , define the polytope $T(\alpha, s) \subset \mathbb{R}^p$ by

$$T(\alpha, s) = \{\mathbf{v} \in \mathbb{R}^p : \|\mathbf{v}\|_\infty \leq \alpha, \|\mathbf{v}\|_1 \leq s\alpha\}.$$

For any $\mathbf{v} \in \mathbb{R}^p$, define the set of sparse vectors $U(\alpha, s, \mathbf{v}) \subset \mathbb{R}^p$ by

$$U(\alpha, s, \mathbf{v}) = \{\mathbf{u} \in \mathbb{R}^p : \text{supp}(\mathbf{u}) \subseteq \text{supp}(\mathbf{v}), \|\mathbf{u}\|_0 \leq s, \|\mathbf{u}\|_1 = \|\mathbf{v}\|_1, \|\mathbf{u}\|_\infty \leq \alpha\}.$$

Then $\mathbf{v} \in T(\alpha, s)$ if and only if \mathbf{v} is in the convex hull of $U(\alpha, s, \mathbf{v})$. In particular, any $\mathbf{v} \in T(\alpha, s)$ can be expressed as

$$\mathbf{v} = \sum_{i=1}^N \lambda_i \mathbf{u}_i \quad \text{and} \quad 0 \leq \lambda_i \leq 1, \quad \sum_{i=1}^N \lambda_i = 1 \quad \text{and} \quad \mathbf{u}_i \in U(\alpha, s, \mathbf{v}).$$

The following lemma will play an important role in the proof of the main result.

Lemma 2.3 ([5]). *Suppose $m \geq r$, $a_1 \geq a_2 \geq \dots \geq a_m \geq 0$, $\sum_{i=1}^r a_i \geq \sum_{i=r+1}^m a_i$, then for all $\alpha \geq 1$,*

$$\sum_{j=r+1}^m a_j^\alpha \leq \sum_{i=1}^r a_i^\alpha.$$

More generally, suppose $a_1 \geq a_2 \geq \dots \geq a_m \geq 0$, $\lambda \geq 0$ and $\sum_{i=1}^r a_i + \lambda \geq \sum_{i=r+1}^m a_i$, then for all $\alpha \geq 1$,

$$\sum_{j=r+1}^m a_j^\alpha \leq r \left(\sqrt[\alpha]{\frac{\sum_{i=1}^r a_i^\alpha}{r}} + \frac{\lambda}{r} \right)^\alpha.$$

3. Main Results

In this section, we first state two existing representative results for corrupted compressed sensing and then we will present new RIP bound to achieve the stable recovery of sparsely corrupted signals via the minimization program (1.9).

3.1. Two representative results for corrupted CS

Theorem 3.1 ([26]). *Suppose that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e}$ and $\Phi = [\mathbf{A} \ \mathbf{I}] \in \mathbb{R}^{m \times (n+m)}$. If Φ has the $(2s_1, 2s_2)$ -RIP constants satisfying*

$$\delta_{2s_1, 2s_2} < \frac{1}{18} \tag{3.1}$$

and

$$\lambda \in \left[\frac{1}{2} \sqrt{\frac{s_1}{s_2}}, 2 \sqrt{\frac{s_1}{s_2}} \right], \tag{3.2}$$

then for any $\mathbf{x} \in \mathbb{R}^n$ with $|\text{supp}(\mathbf{x})| \leq s_1$, any $\mathbf{f} \in \mathbb{R}^m$ with $|\text{supp}(\mathbf{f})| \leq s_2$, and $\mathbf{e} \in \mathbb{R}^m$ with $\|\mathbf{e}\|_2 \leq \varepsilon$, the solution $(\hat{\mathbf{x}}, \hat{\mathbf{f}})$ to the penalized optimization problem (1.4) satisfies

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \|\hat{\mathbf{f}} - \mathbf{f}\|_2 \leq \frac{4\sqrt{13 + 13\delta_{2s_1, 2s_2}}}{1 - 9\delta_{2s_1, 2s_2}} \varepsilon.$$

Theorem 3.2 ([1]). *Suppose $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e}$ and $\Phi = [\mathbf{A} \ \mathbf{I}] \in \mathbb{R}^{m \times (n+m)}$ has the $(2s_1, 2s_2)$ -RIP constant $\delta_{2s_1, 2s_2}$ satisfying*

$$\delta_{2s_1, 2s_2} < \frac{1}{\sqrt{1 + \left(1/(2\sqrt{2}) + \sqrt{\eta}\right)^2}} \tag{3.3}$$

with $\eta = \frac{s_1 + \lambda^2 s_2}{\min\{s_1, \lambda^2 s_2\}}$. Then for $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{f} \in \mathbb{R}^m$, and $\mathbf{e} \in \mathbb{R}^m$ with $\|\mathbf{e}\|_2 \leq \varepsilon$, the solution $(\hat{\mathbf{x}}, \hat{\mathbf{f}})$ to the penalized optimization problem (1.4) satisfies

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \|\hat{\mathbf{f}} - \mathbf{f}\|_2 \leq c_1 \left(1 + \eta^{\frac{1}{4}}\right) \left(\frac{\|\mathbf{x}_{-\max(s_1)}\|_1}{\sqrt{s_1}} + \frac{\|\mathbf{f}_{-\max(s_2)}\|_1}{\sqrt{s_2}} \right) + c_2 \left(1 + \eta^{\frac{1}{4}}\right) \varepsilon.$$

3.2. New RIP bound for corrupted CS

Theorem 3.3. Consider $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{f} + \mathbf{e}$ with $\|\mathbf{e}\|_2 \leq \varepsilon$. If the sensing matrix $\Phi = [\mathbf{A} \ \mathbf{B}]$ has the (t_1s_1, t_2s_2) -RIP constant $\delta_{t_1s_1, t_2s_2}$ satisfying

$$\delta_{t_1s_1, t_2s_2} < \frac{1}{\sqrt{1+2c_1}} \tag{3.4}$$

with $t_1 > 1, t_2 > 1, c_1 = \frac{\max\{s_1, \lambda^2s_2\}}{\min\{s_1(t_1-1), \lambda^2s_2(t_2-1)\}}$, then the solution $(\hat{\mathbf{x}}, \hat{\mathbf{f}})$ of (1.9) satisfies

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \|\hat{\mathbf{f}} - \mathbf{f}\|_2 \leq C_1\varepsilon + C_2 \left(\frac{\|\mathbf{x}_{-\max(s_1)}\|_1}{\sqrt{s_1}} + \frac{\|\mathbf{f}_{-\max(s_2)}\|_1}{\sqrt{s_2}} \right),$$

where

$$C_1 = \left(\sqrt{2} + \max \left\{ \lambda \sqrt{\frac{s_2}{s_1}}, \frac{1}{\lambda} \sqrt{\frac{s_1}{s_2}} \right\} \right) \frac{2\sqrt{1+\delta}}{1-\delta\sqrt{1+2c_1}},$$

$$C_2 = 2 \left[\left(\sqrt{2} + \max \left\{ \lambda \sqrt{\frac{s_2}{s_1}}, \frac{1}{\lambda} \sqrt{\frac{s_1}{s_2}} \right\} \right) \left(\frac{2\delta c_1 + \sqrt{[\sqrt{1+2c_1} - \delta(1+2c_1)]\delta c_1}}{\sqrt{1+2c_1} - \delta(1+2c_1)} \right) \right. \\ \left. + \left(\max \left\{ 1, \lambda \sqrt{\frac{s_2}{s_1}} \right\} + \max \left\{ \frac{1}{\lambda} \sqrt{\frac{s_1}{s_2}}, 1 \right\} \right) \right],$$

and $\delta_{t_1s_1, t_2s_2}$ denotes as δ for the sake of simplicity.

Theorem 3.3 provides a new bound on the (t_1s_1, t_2s_2) -RIP under which stable signal recovery can be obtained via the minimization program JPND with a parameter λ , i.e. (1.9). To make the modified RIP condition weakest, we need to make the right-hand side of (3.4) largest by adjusting the extra parameter λ to make c_1 as small as possible. If $t_1 = t_2 = 2$, then c_1 becomes $\hat{c}_1 = \frac{\max\{s_1, \lambda^2s_2\}}{\min\{s_1, \lambda^2s_2\}}$. In this special case, the best choice of λ is $\lambda^2 = \frac{s_1}{s_2}$, then $\hat{c}_1 = 1$ and $\delta_{(2s_1, 2s_2)} < \frac{\sqrt{3}}{3}$.

Remark 3.1. If $\lambda = 1$, the optimization problem (1.9) exactly is JPND. Let $\lambda = 1$ in Theorem 3.3, we obtain the recovery guarantee for JPND.

If $\lambda = 1$ and $t_1 = t_2 = t > 1$, the sufficient condition (3.4) for the recovery guarantee reduces to

$$\delta_{(ts_1, ts_2)} < \frac{1}{\sqrt{1+2c'_1}} \tag{3.5}$$

with $c'_1 = \frac{\max\{s_1, \lambda^2s_2\}}{(t-1)\min\{s_1, \lambda^2s_2\}}$ and the sufficient condition given by Cai and Zhang [6] can be reformulated as

$$\delta_{t(s_1+s_2)} < \sqrt{\frac{t-1}{t}}. \tag{3.6}$$

Even though $\sqrt{\frac{t-1}{t}} > \frac{1}{\sqrt{1+2c_1}}$, we also have $\delta_{t(s_1+s_2)} > \delta_{(ts_1, ts_2)}$, so these two results are mutually exclusive.

By the result developed in Theorem 3.3, we easily obtain the following sufficient conditions for the exact recovery of all (s_1, s_2) -sparse signals in the noiseless case.

Corollary 3.1. *Suppose that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{f}$, where $\mathbf{x} \in \mathbb{R}^n$ is an s_1 -sparse signal and $\mathbf{f} \in \mathbb{R}^l$ is an s_2 -sparse corruption vector. If the matrix $[\mathbf{A} \ \mathbf{B}]$ satisfies the RIP condition (3.4) for $t_1 > 1$ and $t_2 > 1$, then the minimizer $(\hat{\mathbf{x}}, \hat{\mathbf{f}})$ of (1.9) with $\varepsilon = 0$ recovers (\mathbf{x}, \mathbf{f}) exactly.*

Comparing Theorem 3.3 with two known representative results Theorems 3.1 and 3.2 leads to the following conclusions.

Remark 3.2. For $t_1 = t_2 = 2$, the condition (3.4) reduces to

$$\delta_{2s_1, 2s_2} < \frac{1}{\sqrt{1 + 2\hat{c}_1}}, \tag{3.7}$$

where $\hat{c}_1 = \frac{\max\{s_1, \lambda^2 s_2\}}{\min\{s_1, \lambda^2 s_2\}}$. We show that any s_1 -sparse vector $\mathbf{x} \in \mathbb{R}^n$ and any s_2 -sparse vector $\mathbf{f} \in \mathbb{R}^l$ can be recovered if the RIP condition holds. By computing directly, we can see that

$$\frac{1}{18} < \frac{1}{\sqrt{1 + 2\hat{c}_1}}$$

always holds under the condition (3.2) for any $s_1, s_2 \neq 0$. This implies that our condition (3.7) is always weaker than the RIP condition (3.1). Specifically, if $\lambda = \sqrt{\frac{s_1}{s_2}}$, then $\hat{c}_1 = 1$, one gets

$$\frac{1}{18} < \frac{1}{\sqrt{1 + 2\hat{c}_1}} = \frac{1}{\sqrt{3}}.$$

If $\frac{1}{2}\sqrt{\frac{s_1}{s_2}} \leq \lambda < \sqrt{\frac{s_1}{s_2}}$, i.e., $s_1 > \lambda^2 s_2$, then $\hat{c}_1 = \frac{s_1}{\lambda^2 s_2}$, we have

$$\frac{1}{\sqrt{1 + 2\hat{c}_1}} = \frac{1}{\sqrt{1 + \frac{2s_1}{\lambda^2 s_2}}} \geq \frac{1}{\sqrt{1 + \frac{2s_1}{\frac{s_1}{4s_2} s_2}}} = \frac{1}{3} > \frac{1}{18}.$$

If $\sqrt{\frac{s_1}{s_2}} < \lambda \leq 2\sqrt{\frac{s_1}{s_2}}$, i.e., $s_1 < \lambda^2 s_2$, then $\hat{c}_1 = \frac{\lambda^2 s_2}{s_1}$, we have

$$\frac{1}{\sqrt{1 + 2\hat{c}_1}} = \frac{1}{\sqrt{1 + \frac{2\lambda^2 s_2}{s_1}}} \geq \frac{1}{\sqrt{1 + \frac{8s_1}{\frac{s_2}{s_1} s_2}}} = \frac{1}{3} > \frac{1}{18}.$$

Remark 3.3. For $t_1 = t_2 = 2$, the condition (3.4) reduces to (3.7). And we infer that

$$\frac{1}{\sqrt{1 + (\frac{1}{2\sqrt{2}} + \sqrt{\eta})^2}} < \frac{1}{\sqrt{1 + 2\hat{c}_1}}$$

for any $s_1, s_2 \neq 0$,

$$\frac{6\sqrt{2} - 4}{7} \sqrt{\frac{s_1}{s_2}} < \lambda < \frac{2 + 3\sqrt{2}}{4} \sqrt{\frac{s_1}{s_2}} \quad \text{and} \quad \eta = \frac{s_1 + \lambda^2 s_2}{\min\{s_1, \lambda^2 s_2\}},$$

which means our condition (3.7) for $\mathbf{B} = \mathbf{I}$ is weaker than the RIP condition (3.3).

In particular, if $\lambda = \sqrt{\frac{s_1}{s_2}}$, then $\hat{c}_1 = 1$ and $\eta = 2$, there holds

$$\sqrt{2c_1} = \sqrt{2} < \frac{1}{2\sqrt{2}} + \sqrt{\eta} = \frac{1}{2\sqrt{2}} + \sqrt{2}.$$

If $\lambda < \sqrt{\frac{s_1}{s_2}}$, i.e., $s_1 > \lambda^2 s_2$, then $\hat{c}_1 = \frac{s_1}{\lambda^2 s_2}$ and $\eta = \frac{s_1 + \lambda^2 s_2}{\lambda^2 s_2}$. We observe that

$$\sqrt{2c_1} < \frac{1}{2\sqrt{2}} + \sqrt{\eta} \implies \lambda > \frac{6\sqrt{2}-4}{7} \sqrt{\frac{s_1}{s_2}}.$$

If $\lambda > \sqrt{\frac{s_1}{s_2}}$, i.e., $s_1 < \lambda^2 s_2$, then $\hat{c}_1 = \frac{\lambda^2 s_2}{s_1}$ and $\eta = \frac{s_1 + \lambda^2 s_2}{s_1}$. We have that

$$\sqrt{2c_1} < \frac{1}{2\sqrt{2}} + \sqrt{\eta} \implies \lambda < \frac{2+3\sqrt{2}}{4} \sqrt{\frac{s_1}{s_2}}.$$

From what has been discussed above, we have

$$\frac{1}{\sqrt{1 + \left(\frac{1}{2\sqrt{2}} + \sqrt{\eta}\right)^2}} < \frac{1}{\sqrt{1 + 2\hat{c}_1}} \quad \text{for any} \quad \frac{6\sqrt{2}-4}{7} \sqrt{\frac{s_1}{s_2}} < \lambda < \frac{2+3\sqrt{2}}{4} \sqrt{\frac{s_1}{s_2}} \quad (s_1, s_2 > 0).$$

Now, we prove Theorem 3.3 by using Lemmas 2.1-2.3. The proof of Theorem 3.3 is not a trivial consequence of [6, 40]. The novelty of the proof technique is how to concatenate the two sparse representations obtained by applying Lemma 2.2 to \mathbf{x} and \mathbf{f} respectively. This leads to the sum of the coefficients of the new sparse representation is 2 instead of 1 as in Lemma 2.2. Therefore, we obtain a new Eq. (3.14) which induces different choices of other parameters from the classical cases. For example, $c = 1$ in our proof instead of $1/2$ in classical case.

Proof. Without loss of generality, we first assume that $t_1 s_1$ and $t_2 s_2$ are integers. Suppose that $\mathbf{z} = \hat{\mathbf{x}} - \mathbf{x}$, $\mathbf{h} = \hat{\mathbf{f}} - \mathbf{f}$, we have

$$\left\| \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\|_2 \leq \left\| \mathbf{y} - \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{x} \\ \mathbf{f} \end{bmatrix} \right\|_2 + \left\| \mathbf{y} - \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \hat{\mathbf{x}} \\ \hat{\mathbf{f}} \end{bmatrix} \right\|_2 \leq 2\varepsilon \quad (3.8)$$

and the following cone constrained inequality by Lemma 2.1:

$$\|\mathbf{z}_{-\max(s_1)}\|_1 + \lambda \|\mathbf{h}_{-\max(s_2)}\|_1 \leq \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \lambda \mathbf{h}_{\max(s_2)} \end{bmatrix} \right\|_1 + 2 \left\| \begin{bmatrix} \mathbf{x}_{-\max(s_1)} \\ \lambda \mathbf{f}_{-\max(s_2)} \end{bmatrix} \right\|_1.$$

Define

$$\beta = \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \lambda \mathbf{h}_{\max(s_2)} \end{bmatrix} \right\|_1 + 2 \left\| \begin{bmatrix} \mathbf{x}_{-\max(s_1)} \\ \lambda \mathbf{f}_{-\max(s_2)} \end{bmatrix} \right\|_1$$

and

$$\alpha_1 = \frac{\|\mathbf{z}_{-\max(s_1)}\|_1}{s_1}, \quad \alpha_2 = \frac{\|\lambda \mathbf{h}_{-\max(s_2)}\|_1}{s_2}.$$

Next, we divide $\begin{bmatrix} \mathbf{z}_{-\max(s_1)} \\ \lambda \mathbf{h}_{-\max(s_2)} \end{bmatrix}$ into two parts, i.e.,

$$\begin{bmatrix} \mathbf{z}_{-\max(s_1)} \\ \lambda \mathbf{h}_{-\max(s_2)} \end{bmatrix} = \begin{bmatrix} \mathbf{z}^{(1)} \\ \lambda \mathbf{h}^{(1)} \end{bmatrix} + \begin{bmatrix} \mathbf{z}^{(2)} \\ \lambda \mathbf{h}^{(2)} \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{z}^{(1)} &= \mathbf{z}_{-\max(s_1)} \cdot \mathbf{1}_{\{i \mid |\mathbf{z}_{-\max(s_1)}(i)| > \frac{\alpha_1}{t_1 - 1}\}}, & \lambda \mathbf{h}^{(1)} &= \lambda \mathbf{h}_{-\max(s_2)} \cdot \mathbf{1}_{\{i \mid |\lambda \mathbf{h}_{-\max(s_2)}(i)| > \frac{\alpha_2}{t_2 - 1}\}}, \\ \mathbf{z}^{(2)} &= \mathbf{z}_{-\max(s_1)} \cdot \mathbf{1}_{\{i \mid |\mathbf{z}_{-\max(s_1)}(i)| \leq \frac{\alpha_1}{t_1 - 1}\}}, & \lambda \mathbf{h}^{(2)} &= \lambda \mathbf{h}_{-\max(s_2)} \cdot \mathbf{1}_{\{i \mid |\lambda \mathbf{h}_{-\max(s_2)}(i)| \leq \frac{\alpha_2}{t_2 - 1}\}}. \end{aligned}$$

Let

$$|\text{supp}(\mathbf{z}^{(1)})| = \|\mathbf{z}^{(1)}\|_0 = m_1, \quad |\text{supp}(\lambda\mathbf{h}^{(1)})| = \|\lambda\mathbf{h}^{(1)}\|_0 = m_2.$$

We have

$$\|\mathbf{z}^{(1)}\|_1 \leq \|\mathbf{z}_{-\max(s_1)}\|_1 = s_1\alpha_1, \quad \|\mathbf{z}^{(1)}\|_1 \geq m_1 \frac{\alpha_1}{t_1 - 1}$$

and

$$\|\lambda\mathbf{h}^{(1)}\|_1 \leq \|\lambda\mathbf{h}_{-\max(s_2)}\|_1 = s_2\alpha_2, \quad \|\lambda\mathbf{h}^{(1)}\|_1 \geq m_2 \frac{\alpha_2}{t_2 - 1}.$$

Then $m_1 \leq s_1(t_1 - 1)$ and $m_2 \leq s_2(t_2 - 1)$. Therefore, we derive that

$$\begin{aligned} \|\mathbf{z}^{(2)}\|_1 &= \|\mathbf{z}_{-\max(s_1)}\|_1 - \|\mathbf{z}^{(1)}\|_1 \\ &\leq s_1\alpha_1 - m_1 \frac{\alpha_1}{t_1 - 1} \\ &= [s_1(t_1 - 1) - m_1] \frac{\alpha_1}{t_1 - 1}, \\ \|\mathbf{z}^{(2)}\|_\infty &\leq \frac{\alpha_1}{t_1 - 1}, \\ \|\lambda\mathbf{h}^{(2)}\|_1 &= \|\lambda\mathbf{h}_{-\max(s_2)}\|_1 - \|\lambda\mathbf{h}^{(1)}\|_1 \\ &\leq s_2\alpha_2 - m_2 \frac{\alpha_2}{t_2 - 1} \\ &= [s_2(t_2 - 1) - m_2] \frac{\alpha_2}{t_2 - 1}, \\ \|\lambda\mathbf{h}^{(2)}\|_\infty &\leq \frac{\alpha_2}{t_2 - 1}. \end{aligned}$$

Now, we respectively obtain a combination of sparse vectors of $\mathbf{z}^{(2)}$ and $\lambda\mathbf{h}^{(2)}$ by applying Lemma 2.2. Then

$$\mathbf{z}^{(2)} = \sum_{i=1}^{N_1} \lambda_i^z \mathbf{u}_i^z,$$

where every \mathbf{u}_i^z is $(s_1(t_1 - 1) - m_1)$ -sparse and $\|\mathbf{u}_i^z\|_1 = \|\mathbf{z}^{(2)}\|_1$, $\|\mathbf{u}_i^z\|_\infty \leq \frac{\alpha_1}{t_1 - 1}$, $\text{supp}(\mathbf{u}_i^z) \subseteq \text{supp}(\mathbf{z}^{(2)})$, $\|\mathbf{u}_i^z\|_2 \leq \sqrt{\|\mathbf{u}_i^z\|_0} \|\mathbf{u}_i^z\|_\infty \leq \sqrt{\frac{s_1}{t_1 - 1}} \alpha_1$. And

$$\lambda\mathbf{h}^{(2)} = \sum_{i=1}^{N_2} \lambda_i^h \lambda \mathbf{u}_i^h,$$

where every \mathbf{u}_i^h is $(s_2(t_2 - 1) - m_2)$ -sparse and $\|\lambda \mathbf{u}_i^h\|_1 = \|\lambda\mathbf{h}^{(2)}\|_1$, $\|\lambda \mathbf{u}_i^h\|_\infty \leq \frac{\alpha_2}{t_2 - 1}$, $\text{supp}(\mathbf{u}_i^h) \subseteq \text{supp}(\mathbf{h}^{(2)})$, $\|\lambda \mathbf{u}_i^h\|_2 \leq \sqrt{\frac{s_2}{t_2 - 1}} \alpha_2$. Furthermore, one has

$$\sum_{i=1}^{N_1} \lambda_i^z \|\mathbf{u}_i^z\|_2^2 \leq \frac{s_1}{t_1 - 1} \alpha_1^2, \quad (3.9)$$

$$\sum_{i=1}^{N_2} \lambda_i^h \|\lambda \mathbf{u}_i^h\|_2^2 \leq \frac{s_2}{t_2 - 1} \alpha_2^2. \quad (3.10)$$

Since $\begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \lambda \mathbf{h}_{\max(s_2)} \end{bmatrix} + \begin{bmatrix} \mathbf{z}^{(1)} \\ \lambda \mathbf{h}^{(1)} \end{bmatrix}$ is $(s_1 t_1, s_2 t_2)$ -sparse,

$$\begin{aligned} & \left\langle [\mathbf{A} \ \mathbf{B}] \left(\begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \mathbf{h}_{\max(s_2)} \end{bmatrix} + \begin{bmatrix} \mathbf{z}^{(1)} \\ \mathbf{h}^{(1)} \end{bmatrix} \right), [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\rangle \\ & \leq \left\| [\mathbf{A} \ \mathbf{B}] \left(\begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \mathbf{h}_{\max(s_2)} \end{bmatrix} + \begin{bmatrix} \mathbf{z}^{(1)} \\ \mathbf{h}^{(1)} \end{bmatrix} \right) \right\|_2 \left\| [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\|_2 \\ & \leq 2\varepsilon \sqrt{1 + \delta_{s_1 t_1, s_2 t_2}} \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \mathbf{h}_{\max(s_2)} \end{bmatrix} + \begin{bmatrix} \mathbf{z}^{(1)} \\ \mathbf{h}^{(1)} \end{bmatrix} \right\|_2, \end{aligned} \quad (3.11)$$

where the second inequality is due to (2.2) and (3.8). In order to take advantage of the RIP of the matrix $[\mathbf{A} \ \mathbf{B}]$, let

$$\begin{aligned} \hat{\mathbf{z}} &= \begin{bmatrix} \mathbf{z}^{(2)} \\ \mathbf{0} \end{bmatrix} = \sum_{i=1}^{N_1} \lambda_i^z \begin{bmatrix} \mathbf{u}_i^z \\ \mathbf{0} \end{bmatrix} = \sum_{i=1}^{N_1} \lambda_i^z \hat{\mathbf{u}}_i^z, \\ \hat{\mathbf{h}} &= \begin{bmatrix} \mathbf{0} \\ \mathbf{h}^{(2)} \end{bmatrix} = \sum_{i=1}^{N_2} \lambda_i^h \begin{bmatrix} \mathbf{0} \\ \mathbf{u}_i^h \end{bmatrix} = \sum_{i=1}^{N_2} \lambda_i^h \hat{\mathbf{u}}_i^h. \end{aligned}$$

Then

$$\begin{bmatrix} \mathbf{z}^{(2)} \\ \mathbf{h}^{(2)} \end{bmatrix} = \hat{\mathbf{z}} + \hat{\mathbf{h}} = \sum_{i=1}^{N_1+N_2} \lambda_i \mathbf{u}_i, \quad (3.12)$$

where

$$\mathbf{u}_i = \begin{cases} \hat{\mathbf{u}}_i^z, & i \leq N_1, \\ \hat{\mathbf{u}}_{i-N_1}^h, & N_1 + 1 \leq i \leq N_1 + N_2, \end{cases}$$

$$\lambda_i = \begin{cases} \lambda_i^z, & i \leq N_1, \\ \lambda_{i-N_1}^h, & N_1 + 1 \leq i \leq N_1 + N_2. \end{cases}$$

Note that $\sum_{i=1}^{N_1} \lambda_i^z = 1$ and $\sum_{i=1}^{N_2} \lambda_i^h = 1$, which implies that $\sum_{i=1}^{N_1+N_2} \lambda_i = 2$. Suppose that

$$X = \|\mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)}\|_2 + \|\mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)}\|_2, \quad P = 2 \left\| \begin{bmatrix} \frac{\mathbf{z}_{-\max(s_1)}}{\sqrt{s_1}} \\ \frac{\mathbf{f}_{-\max(s_2)}}{\sqrt{s_2}} \end{bmatrix} \right\|_1.$$

By (3.9), (3.10) and (3.12), we have

$$\begin{aligned} & \sum_{i=1}^{N_1+N_2} \lambda_i \|\mathbf{u}_i\|_2^2 = \sum_{i=1}^{N_1} \lambda_i^z \|\mathbf{u}_i^z\|_2^2 + \sum_{i=1}^{N_2} \lambda_i^h \|\mathbf{u}_i^h\|_2^2 \\ & \leq \frac{s_1}{t_1 - 1} \alpha_1^2 + \frac{s_2}{\lambda^2(t_2 - 1)} \alpha_2^2 \\ & = \frac{1}{s_1(t_1 - 1)} \|\mathbf{z}_{-\max(s_1)}\|_1^2 + \frac{1}{\lambda^2 s_2(t_2 - 1)} \|\lambda \mathbf{h}_{-\max(s_2)}\|_1^2 \\ & \leq \frac{1}{\min\{s_1(t_1 - 1), \lambda^2 s_2(t_2 - 1)\}} \left(\|\mathbf{z}_{-\max(s_1)}\|_1^2 + \|\lambda \mathbf{h}_{-\max(s_2)}\|_1^2 \right) \\ & \leq \frac{1}{\min\{s_1(t_1 - 1), \lambda^2 s_2(t_2 - 1)\}} \beta^2 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\min\{s_1(t_1 - 1), \lambda^2 s_2(t_2 - 1)\}} \left(\left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} \\ \lambda \mathbf{h}_{\max(s_2)} \end{bmatrix} \right\|_1 + 2 \left\| \begin{bmatrix} \mathbf{x}_{-\max(s_1)} \\ \lambda \mathbf{f}_{-\max(s_2)} \end{bmatrix} \right\|_1 \right)^2 \\
 &\leq \frac{1}{\min\{s_1(t_1 - 1), \lambda^2 s_2(t_2 - 1)\}} \left(\sqrt{s_1} \|\mathbf{z}_{\max(s_1)}\|_2 + \lambda \sqrt{s_2} \|\mathbf{h}_{\max(s_2)}\|_2 \right. \\
 &\quad \left. + 2 \|\mathbf{x}_{-\max(s_1)}\|_1 + 2\lambda \|\mathbf{f}_{-\max(s_2)}\|_1 \right)^2 \\
 &\leq \frac{\max\{s_1, \lambda^2 s_2\}}{\min\{s_1(t_1 - 1), \lambda^2 s_2(t_2 - 1)\}} (X + P)^2 \\
 &\triangleq c_1 (X + P)^2.
 \end{aligned}$$

Now suppose $\mu \geq 0$ and $c \geq 0$ which are to be determined. Denote

$$\boldsymbol{\beta}_i = \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} + \mu \mathbf{u}_i,$$

we obtain

$$\sum_{j=1}^{N_1+N_2} \lambda_j \boldsymbol{\beta}_j - c \boldsymbol{\beta}_i = (2 - \mu - c) \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} - c \mu \mathbf{u}_i + \mu \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix}. \tag{3.13}$$

Notice that $\boldsymbol{\beta}_i$ and $\sum_{j=1}^{N_1+N_2} \lambda_j \boldsymbol{\beta}_j - c \boldsymbol{\beta}_i - \mu \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix}$ are both $(t_1 s_1, t_2 s_2)$ -sparse vectors.

The following identity holds in l_2 norm:

$$\begin{aligned}
 &\sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left(\sum_{j=1}^{N_1+N_2} \lambda_j \boldsymbol{\beta}_j - c \boldsymbol{\beta}_i \right) \right\|_2^2 - \sum_{i=1}^{N_1+N_2} \lambda_i (2 - c)^2 \left\| [\mathbf{A} \ \mathbf{B}] \boldsymbol{\beta}_i \right\|_2^2 \\
 &\quad + (2 - 2c) \sum_{1 \leq i < j \leq N_1+N_2} \lambda_i \lambda_j \left\| [\mathbf{A} \ \mathbf{B}] (\boldsymbol{\beta}_i - \boldsymbol{\beta}_j) \right\|_2^2 = 0. \tag{3.14}
 \end{aligned}$$

Set $c = 1$, $\mu = \frac{\sqrt{1+2c_1}-1}{c_1}$ and substitute (3.13) into (3.14), we get

$$\begin{aligned}
 0 &= \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left((1 - \mu) \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} - \mu \mathbf{u}_i + \mu \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right) \right\|_2^2 - \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \boldsymbol{\beta}_i \right\|_2^2 \\
 &= \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left((1 - \mu) \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} - \mu \mathbf{u}_i \right) \right\|_2^2 + \mu^2 \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\|_2^2 \\
 &\quad + 2\mu \sum_{i=1}^{N_1+N_2} \lambda_i \left\langle [\mathbf{A} \ \mathbf{B}] \left((1 - \mu) \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} - \mu \mathbf{u}_i \right), [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\rangle \\
 &\quad - \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left(\begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} + \mu \mathbf{u}_i \right) \right\|_2^2 \\
 &= \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left((1 - \mu) \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} - \mu \mathbf{u}_i \right) \right\|_2^2 \\
 &\quad + 2\mu(2 - \mu) \left\langle [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix}, [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{z} \\ \mathbf{h} \end{bmatrix} \right\rangle
 \end{aligned}$$

$$\begin{aligned}
& - \sum_{i=1}^{N_1+N_2} \lambda_i \left\| [\mathbf{A} \ \mathbf{B}] \left(\begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} + \mu \mathbf{u}_i \right) \right\|_2^2 \\
\leq & (1 + \delta) \sum_{i=1}^{N_1+N_2} \lambda_i \left((1 - \mu)^2 \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} \right\|_2^2 + \mu^2 \|\mathbf{u}_i\|_2^2 \right) \\
& + 4\mu(2 - \mu)\varepsilon\sqrt{1 + \delta} \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} \right\|_2 \\
& - (1 - \delta) \sum_{i=1}^{N_1+N_2} \lambda_i \left(\left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} \right\|_2^2 + \mu^2 \|\mathbf{u}_i\|_2^2 \right) \\
= & \sum_{i=1}^{N_1+N_2} \lambda_i \left\{ [(1 + \delta)(1 - \mu)^2 - (1 - \delta)] \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} \right\|_2^2 \right\} + 2\delta\mu^2 \sum_{i=1}^{N_1+N_2} \lambda_i \|\mathbf{u}_i\|_2^2 \\
& + 4\mu(2 - \mu)\varepsilon\sqrt{1 + \delta} \left\| \begin{bmatrix} \mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)} \\ \mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)} \end{bmatrix} \right\|_2 \\
\leq & \sum_{i=1}^{N_1+N_2} \lambda_i \left\{ [(1 + \delta)(1 - \mu)^2 - (1 - \delta)] \left(\|\mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)}\|_2^2 + \|\mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)}\|_2^2 \right) \right\} \\
& + 4\mu(2 - \mu)\varepsilon\sqrt{1 + \delta} \left(\|\mathbf{z}_{\max(s_1)} + \mathbf{z}^{(1)}\|_2 + \|\mathbf{h}_{\max(s_2)} + \mathbf{h}^{(1)}\|_2 \right) \\
& + 2\delta\mu^2 \sum_{i=1}^{N_1+N_2} \lambda_i \|\mathbf{u}_i\|_2^2. \tag{3.15}
\end{aligned}$$

By the definitions of X and P , the inequality (3.15) can be written as follows:

$$\begin{aligned}
0 & \leq 2 \left\{ [(1 + \delta)(1 - \mu)^2 - (1 - \delta)] X^2 + 2\mu(2 - \mu)\varepsilon\sqrt{1 + \delta} X + \delta\mu^2 c_1 (X + P)^2 \right\} \\
& = 2 \left\{ [\mu^2 - 2\mu + \delta(2 - 2\mu + \mu^2 + c_1\mu^2)] X^2 \right. \\
& \quad \left. + (2\delta c_1 \mu^2 P + 2\varepsilon\mu(2 - \mu)\sqrt{1 + \delta}) X + \delta c_1 \mu^2 P^2 \right\} \\
& = \left[\frac{1 + 2c_1 - (1 + c_1)\sqrt{1 + 2c_1}}{c_1^2} + \delta \left(\frac{(1 + 2c_1)(1 + c_1) - (1 + 2c_1)\sqrt{1 + 2c_1}}{c_1^2} \right) \right] X^2 \\
& \quad + \left(\frac{(1 + c_1) - \sqrt{1 + 2c_1}}{c_1} 2\delta P + 2\varepsilon\sqrt{1 + \delta} \frac{(1 + c_1)\sqrt{1 + 2c_1} - (1 + 2c_1)}{c_1^2} \right) X \\
& \quad + \frac{1 + c_1 - \sqrt{1 + 2c_1}}{c_1} \delta P^2 \\
& = \frac{1 + c_1 - \sqrt{1 + 2c_1}}{c_1^2} \left[(\delta(1 + 2c_1) - \sqrt{1 + 2c_1}) X^2 \right. \\
& \quad \left. + (2\delta c_1 P + 2\sqrt{(1 + \delta)(1 + 2c_1)}\varepsilon) X + \delta c_1 P^2 \right]. \tag{3.16}
\end{aligned}$$

By solving this second-order inequality (3.16) for X , we obtain

$$\begin{aligned}
X & \leq \left\{ (\delta c_1 P + \sqrt{(1 + \delta)(1 + 2c_1)}\varepsilon) + \left[(\delta c_1 P + \sqrt{(1 + \delta)(1 + 2c_1)}\varepsilon)^2 \right. \right. \\
& \quad \left. \left. + (\sqrt{1 + 2c_1} - \delta(1 + 2c_1))\delta c_1 P^2 \right]^{\frac{1}{2}} \right\} \cdot \left(\sqrt{1 + 2c_1} - \delta(1 + 2c_1) \right)^{-1}
\end{aligned}$$

$$\leq \frac{2\sqrt{1+\delta}}{1-\delta\sqrt{1+2c_1}}\varepsilon + \frac{2\delta c_1 + \sqrt{(\sqrt{1+2c_1}-\delta(1+2c_1))\delta c_1}}{\sqrt{1+2c_1}-\delta(1+2c_1)}P. \quad (3.17)$$

Since

$$\begin{aligned} & \|z_{-\max(s_1)}\|_1 + \lambda\|h_{-\max(s_2)}\|_1 \\ & \leq \|z_{\max(s_1)}\|_1 + 2\|x_{-\max(s_1)}\|_1 + \lambda(\|h_{\max(s_2)}\|_1 + 2\|f_{-\max(s_2)}\|_1), \end{aligned}$$

we have

$$\|z_{-\max(s_1)}\|_1 \leq \|z_{\max(s_1)}\|_1 + \lambda_1,$$

where

$$\begin{aligned} \lambda_1 &= 2\|x_{-\max(s_1)}\|_1 + \lambda\|h_{\max(s_2)}\|_1 + 2\lambda\|f_{-\max(s_2)}\|_1 \\ &\leq \lambda\|h_{\max(s_2)}\|_1 + \max\{\sqrt{s_1}, \lambda\sqrt{s_2}\}P. \end{aligned}$$

Then we have the following inequality by Lemma 2.3:

$$\begin{aligned} \|z_{-\max(s_1)}\|_2^2 &\leq s_1 \left(\frac{\|z_{\max(s_1)}\|_2}{\sqrt{s_1}} + \frac{\lambda_1}{s_1} \right)^2 \\ &= \left(\|z_{\max(s_1)}\|_2 + \lambda \frac{\|h_{\max(s_2)}\|_1}{\sqrt{s_1}} + \max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\}P \right)^2. \end{aligned} \quad (3.18)$$

Similarly,

$$\lambda\|h_{-\max(s_2)}\|_1 \leq \lambda\|h_{\max(s_2)}\|_1 + \lambda_2,$$

where

$$\begin{aligned} \lambda_2 &= \|z_{\max(s_1)}\|_1 + 2\|x_{-\max(s_1)}\|_1 + 2\lambda\|f_{-\max(s_2)}\|_1 \\ &\leq \|z_{\max(s_1)}\|_1 + \max\{\sqrt{s_1}, \lambda\sqrt{s_2}\}P. \end{aligned}$$

Then Lemma 2.3 implies that

$$\begin{aligned} \lambda^2\|h_{-\max(s_2)}\|_2^2 &\leq s_2 \left(\frac{\lambda\|h_{\max(s_2)}\|_2}{\sqrt{s_2}} + \frac{\lambda_2}{s_2} \right)^2 \\ &= \left(\lambda\|h_{\max(s_2)}\|_2 + \frac{\|z_{\max(s_1)}\|_1}{\sqrt{s_2}} + \max\left\{\sqrt{\frac{s_1}{s_2}}, \lambda\right\}P \right)^2. \end{aligned} \quad (3.19)$$

Furthermore, from (3.18) and (3.19), we have

$$\begin{aligned} \|z\|_2 &= \sqrt{\|z_{\max(s_1)}\|_2^2 + \|z_{-\max(s_1)}\|_2^2} \\ &\leq \sqrt{\|z_{\max(s_1)}\|_2^2 + \left(\|z_{\max(s_1)}\|_2 + \lambda \frac{\|h_{\max(s_2)}\|_1}{\sqrt{s_1}} + \max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\}P \right)^2} \\ &\leq \sqrt{2\|z_{\max(s_1)}\|_2^2} + \lambda \frac{\|h_{\max(s_2)}\|_1}{\sqrt{s_1}} + \max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\}P, \end{aligned} \quad (3.20)$$

and

$$\begin{aligned} \|h\|_2 &= \sqrt{\|h_{\max(s_2)}\|_2^2 + \|h_{-\max(s_2)}\|_2^2} \\ &\leq \sqrt{\|h_{\max(s_2)}\|_2^2 + \left(\|h_{\max(s_2)}\|_2 + \frac{\|z_{\max(s_1)}\|_1}{\lambda\sqrt{s_2}} + \max\left\{\frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}, 1\right\}P \right)^2} \\ &\leq \sqrt{2\|h_{\max(s_2)}\|_2^2} + \frac{\|z_{\max(s_1)}\|_1}{\lambda\sqrt{s_2}} + \max\left\{\frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}, 1\right\}P. \end{aligned} \quad (3.21)$$

Combining (3.17), (3.20) and (3.21), one has

$$\begin{aligned}
 \|\mathbf{z}\|_2 + \|\mathbf{h}\|_2 &\leq \sqrt{2}(\|\mathbf{z}_{\max(s_1)}\|_2 + \|\mathbf{h}_{\max(s_2)}\|_2) + \lambda \frac{\|\mathbf{h}_{\max(s_2)}\|_1}{\sqrt{s_1}} + \max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\}P \\
 &\quad + \frac{\|\mathbf{z}_{\max(s_1)}\|_1}{\lambda\sqrt{s_2}} + \max\left\{\frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}, 1\right\}P \\
 &\leq \sqrt{2}X + \max\left\{\lambda\sqrt{\frac{s_2}{s_1}}, \frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}\right\}X + \max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\}P + \max\left\{\frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}, 1\right\}P \\
 &\leq \left(\sqrt{2} + \max\left\{\lambda\sqrt{\frac{s_2}{s_1}}, \frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}\right\}\right) \\
 &\quad \times \left(\frac{2\sqrt{1+\delta}}{1-\delta\sqrt{1+2c_1}}\varepsilon + \frac{2\delta c_1 + \sqrt{(\sqrt{1+2c_1}-\delta)(1+2c_1))}\delta c_1}{\sqrt{1+2c_1}-\delta(1+2c_1)}P\right) \\
 &\quad + \left(\max\left\{1, \lambda\sqrt{\frac{s_2}{s_1}}\right\} + \max\left\{\frac{1}{\lambda}\sqrt{\frac{s_1}{s_2}}, 1\right\}\right)P. \tag{3.22}
 \end{aligned}$$

When t_1s_1 and t_2s_2 are not integers, we observe that $t_1 < \frac{\lceil t_1s_1 \rceil}{s_1}$ and $t_2 < \frac{\lceil t_2s_2 \rceil}{s_2}$. Furthermore, let $t'_1 = \frac{\lceil t_1s_1 \rceil}{s_1}$ and $t'_2 = \frac{\lceil t_2s_2 \rceil}{s_2}$, then we have

$$\begin{aligned}
 \delta_{\lceil t_1s_1 \rceil, \lceil t_2s_2 \rceil} &= \delta_{s_1t_1, s_2t_2} < \frac{1}{\sqrt{1 + 2\frac{\max\{s_1, \lambda^2s_2\}}{\min\{s_1(t_1-1), \lambda^2s_2(t_2-1)\}}}}} \\
 &< \frac{1}{\sqrt{1 + 2\frac{\max\{s_1, \lambda^2s_2\}}{\min\{s_1(t'_1-1), \lambda^2s_2(t'_2-1)\}}}}},
 \end{aligned}$$

which finishes the proof. □

4. Numerical Experiments

In this section, we verify the reliability of the model (1.9) with numerical simulations. Let $s_1 = \|\mathbf{x}\|_0$, $s_2 = \|\mathbf{f}\|_0$. In each experiment, we fix $n = 500$, $\mathbf{B} = \mathbf{I}$ and $\|\mathbf{e}\|_2 = 0.001$ for the noisy case. We generate $\mathbf{A} \in \mathbb{R}^{m \times n}$ randomly drawn from i.i.d standard Gaussian distribution. The position of the nonzero entries of \mathbf{x} and \mathbf{f} are randomly generated, while the values of nonzero elements of the original vector \mathbf{x} and \mathbf{f} are generated from a standard Gaussian distribution. In order to compare the average recovery error of each experiment, CVX package [18, 19] for MATLAB was used. For every point on each graph, 100 repetitions were performed.

In Fig. 4.1, we fix $s_1 = 10$, $s_2 = 15$, and vary $\|\mathbf{f}\|_2$ and λ . We use solid point and circle to represent $\|\mathbf{f}\|_2 = 0.05$ and $\|\mathbf{f}\|_2 = 0.3$, respectively. For noiseless case (a), we observe that the average error of BPDN does not decay to zero no matter how large m we set. As m/n increases, each of the solid lines reaches a minimum value greater than zero and does not decay further. In contrast, when λ is 1 or 1.5, JP with the parameter λ reaches exact recovery in all tests. For noisy case (b), the average error of BPDN does not decay to a very small number but JPDN with $\lambda = 1, 1.5$ does.

In Fig. 4.2, we show the effect of the balance parameter λ on the recovery error. Fix $s_1 = 10$, $s_2 = 30$, and let the l_1 -norm of \mathbf{f} be three times the l_1 -norm of \mathbf{x} . From these two figures, we can see that the balance parameter λ has an obvious effect on the recovery error. When the l_1 -norm of \mathbf{f} is several times the l_1 -norm of \mathbf{x} , $\lambda = 0.7$ presents the best recovery performance.

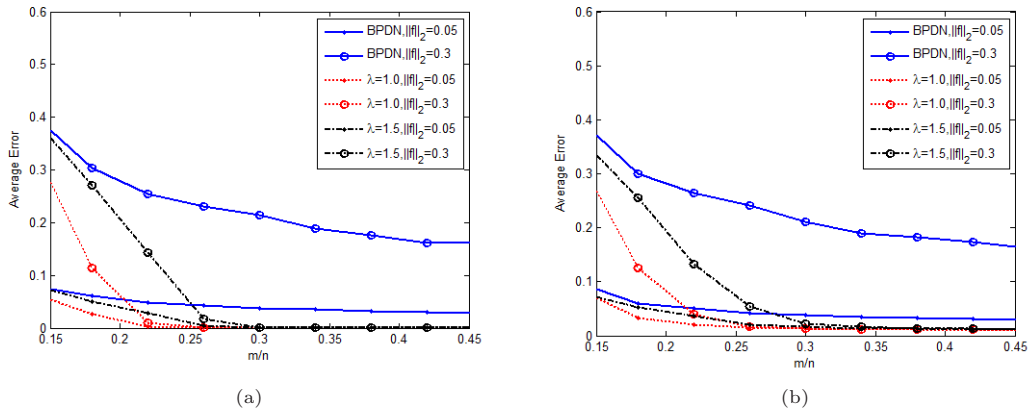


Fig. 4.1. Comparison of the average recovery error $\|\mathbf{x} - \hat{\mathbf{x}}\|_2$ between BPDN and JP (JPDN) with a parameter λ under different values of m/n . In both (a) and (b), we fix $s_1 = \|\mathbf{x}\|_0 = 10$ and $s_2 = \|\mathbf{f}\|_0 = 15$. The solid points and circles indicate that the l_2 norm of \mathbf{f} is 0.05 and 0.3, respectively. (a) In the case of $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f}$, we compare the average errors of BPDN and JP with $\lambda = 1, 1.5$. (b) In the case of $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{f} + \mathbf{e}$, we compare the average errors of BPDN and JPDN with $\lambda = 1, 1.5$.

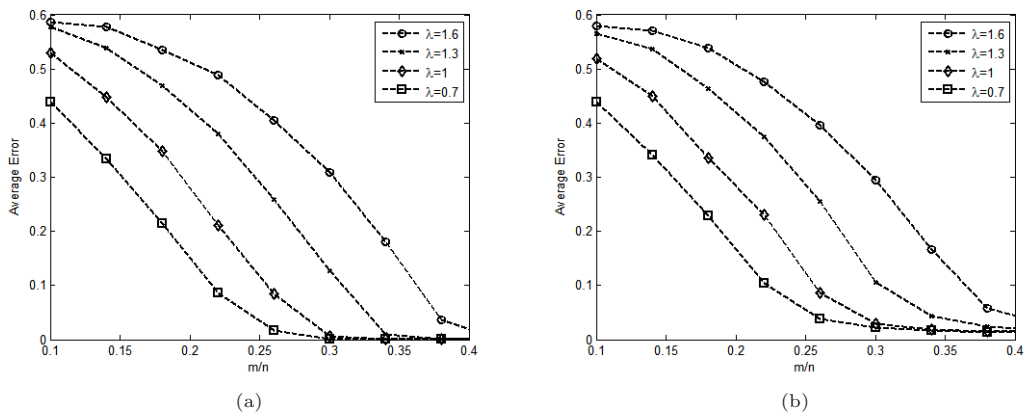


Fig. 4.2. Fix $s_1 = 10$, $s_2 = 30$, and let the l_1 -norm of \mathbf{f} be three times the l_1 -norm of \mathbf{x} .

5. Conclusions

In this paper, we investigated the signal recovery from measurements corrupted by a combination of interference and measurement noise. A new restricted isometry constant bound on $\delta_{t_1 s_1, t_2 s_2}$ ($t_1 > 1, t_2 > 1$) for the exact and stable recovery of sparse signals is proposed and the proposed RIP condition improves the existing representative results. This was accomplished by adapting a crucial sparse decomposition technique to the analysis of the Justice Pursuit method.

Acknowledgements. The authors are very grateful to the referees for their thorough reading and helpful comments that greatly improve the presentation of this paper.

This work was supported by the NSF of China (Grant Nos. 12271050, 11871109, 11901037), by the CAEP Foundation (Grant No. CX20200027) and by the Key Laboratory of Computational Physics Foundation (Grant No. 6142A05210502).

References

- [1] B. Adcock, A. Bao, J.D. Jakeman, and A. Narayan, Compressed sensing with sparse corruptions: Fault-tolerant sparse collocation approximations, *SIAM/ASA J. Uncertain. Quantif.*, **6**:4 (2018), 1424–1453.
- [2] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. Plumbley, A constrained matching pursuit approach to audio declipping, in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Prague, Czech Republic*, (2011), 329–332.
- [3] A. Adler, V. Emiya, M. Jafari, M. Elad, R. Gribonval, and M. Plumbley, Audio inpainting, *IEEE Trans. Audio, Speech, Lang. Process.*, **20**:3 (2012), 922–932.
- [4] J. Cai, S. Osher, and Z. Shen, Split Bregman methods and frame based image restoration, *Multiscale Model. Simul.*, **8**:2 (2009), 337–369.
- [5] T. Cai, and A. Zhang, Sharp RIP bound for sparse signal recovery and low-rank matrix recovery, *Appl. Comput. Harmon. Anal.*, **35** (2013), 74–93.
- [6] T. Cai, and A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, *IEEE Trans. Inf. Theory*, **60** (2014), 122–132.
- [7] E.J. Candés, X. Li, Y. Ma, and J. Wright, Robust principal component analysis?, *J. ACM*, **58**:3 (2011), 1–37.
- [8] E.J. Candés, J. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math.*, **59**:8 (2006), 1207–1223.
- [9] E.J. Candés, and T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory*, **51**:12 (2005), 4203–4215.
- [10] E.J. Candés, and T. Tao, The dantzig selector: Statistical estimation when p is much large than n , *Ann. Statist.*, **35** (2007), 2313–2351.
- [11] D.L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, **52**:4 (2006), 1289–1306.
- [12] D.L. Donoho, and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization, *Proc. Natl. Acad. Sci. USA*, **100**:5 (2003), 2197–2202.
- [13] M. Elad, and Y. Hel-Or, A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur, *IEEE Trans. Image Process.*, **10**:8 (2001), 1187–1193.
- [14] M. Elad, J.L. Starck, P. Querre, and D.L. Donoho, Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA), *Appl. Comput. Harmon. Anal.*, **19** (2005), 340–358.
- [15] R. Foygel, and L. Mackey, Corrupted sensing: Novel guarantees for separating structured signals, *IEEE Trans. Inf. Theory*, **60**:2 (2014), 1223–1247.
- [16] S. Foucart, and H. Rauhut, A Mathematical Introduction to Compressive Sensing, *New York, NY, USA: Springer*, 2013.
- [17] J.J. Fuchs, Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. Inf. Theory*, **51**:10 (2005), 3601–3608.
- [18] M. Grant, and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta, <http://cvxr.com/cvx>, Sept. 2012.
- [19] M. Grant, and S. Boyd, Graph Implementations for Nonsmooth Convex Programs, in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited, 2008.
- [20] W. Göbel, and F. Helmchen, In vivo calcium imaging of neural network function, *Physiology*, **22**:6 (2007), 358–365.
- [21] P. Kuppinger, G. Durisi, and H. Bölcskei, Uncertainty relations and sparse signal recovery for pairs of general signal sets, *IEEE Trans. Inf. Theory*, **58** (2012), 263–277.
- [22] J. Laska, P. Boufounos, M. Davenport, and R. Baraniuk, Democracy in action: Quantization, saturation, and compressive sensing, *Appl. Comput. Harmon. Anal.*, **31**:3 (2011), 429–443.
- [23] J. Laska, M. Davenport, and R. Baraniuk, Exact signal recovery from sparsely corrupted measure-

- ments through the pursuit of justice, *Asilomar Conference on Signals Systems and Computers*, (2009), 1556–1560.
- [24] N. Li, W. Chen, and P. Li, Stable recovery of signals from highly corrupted measurements, *IEEE Access*, **6** (2018), 62865–62873.
- [25] S. Li, J. Lin, D. Liu and W. Sun, Iterative hard thresholding for compressed data separation, *Journal of Complexity*, **59**:2 (2020), 101469.
- [26] X. Li, Compressed sensing and matrix completion with constant proportion of corruptions, *Constr. Approx.*, **37**:1 (2013), 73–99.
- [27] J. Lin, and S. Li, Sparse recovery with coherent tight frames via analysis Dantzig selector and analysis LASSO, *Appl. Comput. Harmon. Anal.*, **37** (2014), 126–139.
- [28] J. Lin, S. Li, and Y. Shen, Compressed data separation with redundant dictionaries, *IEEE Trans. Inf. Theory*, **59**:7 (2013), 4309–4315.
- [29] S. Mallat, and G. Yu, Super-resolution with sparse mixing estimators, *IEEE Trans. Image Process.*, **19**:11 (2010), 2889–2900.
- [30] N. Nguyen, and T. Tran, Exact recoverability from dense corrupted observations via l_1 minimization, *IEEE Trans. Inf. Theory*, **59**:4 (2011), 2017–2035.
- [31] G. Pope, A. Bracher and C. Studer, Probabilistic recovery guarantees for sparsely corrupted signals, *IEEE Trans. Inf. Theory*, **59**:5 (2014), 3104–3116.
- [32] C. Studer, and R.G. Baraniuk, Stable restoration and separation of approximately sparse signals, *Appl. Comput. Harmon. Anal.*, **37** (2014), 12–35.
- [33] C. Studer, P. Kuppinger, G. Pope, and H. Bölckei, Recovery of sparsely corrupted signals, *IEEE Trans. Inf. Theory*, **58**:5 (2012), 3115–3130.
- [34] J.A. Tropp, Greed is good: Algorithmic results for sparse approximation, *IEEE Trans. Inf. Theory*, **50**:10 (2004), 2231–2242.
- [35] J.A. Tropp, Just relax: Convex programming methods for identifying sparse signals in noise, *IEEE Trans. Inf. Theory*, **52**:3 (2006), 1030–1051.
- [36] Y. Tsaig, and D. Donoho, Extensions of compressed sensing, *Signal Process.*, **86**:3 (2006), 549–571.
- [37] S. Vaseghi, and R. Frayling-Cork, Restoration of old gramophone recordings, *J. Audio Eng.*, **40**:10 (1992), 791–801.
- [38] J. Wright, and Y. Ma, Dense error correction via l_1 -minimization, *IEEE Trans. Inf. Theory*, **56** (2010), 3540–3560.
- [39] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, Robust face recognition via sparse representation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **31**:2 (2009), 210–227.
- [40] G. Xu, and Z. Xu, On the l_1 -norm invariant convex k -sparse decomposition of signals, *J. Oper. Res. Soc. China*, **1**:4 (2013), 537–541.
- [41] P. Zhang, L. Gan, C. Ling, and S. Sun, Uniform recovery bounds for structured random matrices in corrupted compressed sensing, *IEEE Trans. Signal Process.*, **66**:8 (2018), 2086–2097.
- [42] R. Zhang, and S. Li, A proof of conjecture on restricted isometry property constants δ_{tk} ($0 < t < \frac{4}{3}$), *IEEE Trans. Inf. Theory*, **64**:3 (2018), 1699–1705.
- [43] R. Zhang, and S. Li, Optimal RIP bounds for sparse signals recovery via minimization, *Appl. Comput. Harmon. Anal.*, **47**:3 (2019), 566–584.