Exact recovery of sparse signals with side information

xiaohu luo (lxh20070701@126.com)
Southwest University https://orcid.org/0000-0002-1897-6832
Nianci Feng
Faculty of Computer and Information Science,Southwest University
Xuhui Guo
School of Information Engineering Ning University
Zili Zhang
School of Information Technology,Deakin University

Research

Keywords: Compressed sensing, side information, ℓ1-ℓ1 method, ℓ1-ℓ2 method, null space property

Posted Date: July 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-660529/v1

Exact recovery of sparse signals with side information

Xiaohu Luo[1]*, Nianci Feng[1]†, Zili Zhang[1,2]‡

[1]Faculty of Computer and Information Science, Southwest University, Chongqing, 400715, China

[2]School of Information Technology, Deakin University, VIC, 3217, Australia

Abstract Compressed sensing has captured considerable attention of researchers in the past decades. In this paper, with the aid of the powerful null space property, some deterministic recovery conditions are established for the previous ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods to guarantee the exact sparse recovery when the side information of the desired signal is available. These obtained results provide a useful and necessary complement to the previous investigation of the ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods that are based on the statistical analysis. Moreover, one of our theoretical findings also shows that, the sharp conditions previously established for the classical ℓ_1 method remain suitable for the ℓ_1-ℓ_1 method to guarantee the exact sparse recovery. Numerical experiments are also carried out to further verify the recovery performance of both ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods.

Key words Compressed sensing, side information, ℓ_1-ℓ_1 method, ℓ_1-ℓ_2 method, null space property.

1 Introduction

Over the past decades, the problem of sparse signal recovery, now termed as compressed sensing (CS) [1–3], has been greatly developed and widely used in many domains such as pattern processing [4], image processing [5–7], medical image [8] and camera design [9], etc. Simply

*E-mail: lxh20070701@126.com

†E-mail: fnc1992sc@163.com

‡Corresponding author, E-mail: zhangzl@swu.edu.cn

speaking, we say a signal $\widehat{\boldsymbol{x}} \in \mathbb{R}^n$ is sparse if and only if it has fewer nonzero components than its length, and if there are at most $k(\ll n)$ nonzero entries in $\widehat{\boldsymbol{x}}$, $\widehat{\boldsymbol{x}}$ is said to be a k-sparse signal. In fact, one of the key goals of CS is the recovery of such a k-sparse signal $\widehat{\boldsymbol{x}}$ from fewer observations $\boldsymbol{b} = A\widehat{\boldsymbol{x}}$, where $A \in \mathbb{R}^{m\times n}(m < n)$ is a pre-designed measurement matrix and $\boldsymbol{b} \in \mathbb{R}^m$ is the resultant observed signal. To realize this goal, it is often suggested to solve the following ℓ_1 method [10, 11]

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \ \|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|, \quad \text{s.t.} \quad A\boldsymbol{x} = \boldsymbol{b}. \tag{1.1}$$

Since model (1.1) is convex, it can be efficiently solved by lots of convex optimization algorithms [12, 13], and many recovery guarantees, including the recovery conditions as well as their resultant recovery error estimates, have also been obtained for this method in the past years, see, e.g., [14–16].

Unfortunately, the ℓ_1 method do not incorporate any side information of $\widehat{\boldsymbol{x}}$ due to the equal treatment of the ℓ_1 norm for the components of the variable $\boldsymbol{x}$. Considering that such side information is often available in many real-world applications, it is naturally expected that the performance of model (1.1) can be further improved if the side information is well integrated. In general, there are two types of common side information in filed of CS. The first one takes the form of a known support estimate. To deal with this type of side information, the authors in [17] first modeled the known support as a set T, and then integrated it into the ℓ_1 norm, leading to the model

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \ \|\boldsymbol{x}_{T^c}\|_1 \triangleq \sum_{i\in T^c} |x_i|, \quad \text{s.t.} \quad A\boldsymbol{x} = \boldsymbol{b},$$

where T^c models the complement set of T in $\{1, 2, 3, \ldots, n\}$. Their work also showed that the resultant recovery conditions are weaker than those without side information. In [18], the authors considered a more general weight rather than a constant weight in the known support estimate. In [19], a variant IHT algorithm was proposed by incorporating the partially known support information, and some theoretical analysis was also established for this algorithm. In [20], OMP, as an iterative greedy algorithm, was extended by using the partially known support. The authors of [21] also considered embedding the known support information into the IRLS algorithm at each iteration, leading a reduction of the number of the measurements as well as the computational cost. Recently, some new recovery conditions were obtained by Ge, et al. in [22]

Another type of the side information takes the form of a similar way to the original signal $\widehat{\boldsymbol{x}}$. The side information of this type usually comes from applications such as MRI [23], video acquisition [24, 25] and estimate problems [26]. In [27], by introducing two ℓ_1 norm and ℓ_2 norm approximation terms to model (1.1), respectively, Mota, et al. proposed to solve an ℓ_1-ℓ_1 method

$$\boldsymbol{x}^\sharp = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\ \|\boldsymbol{x}\|_1 + \beta \parallel \boldsymbol{x} - \boldsymbol{w} \parallel_1 \quad \text{s.t. } A\boldsymbol{x} = \boldsymbol{b} \tag{1.2}$$

and an ℓ_1-ℓ_2 method

$$\boldsymbol{x}^\diamond = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\arg\min}\ \|\boldsymbol{x}\|_1 + \frac{\beta}{2}\|\boldsymbol{x} - \boldsymbol{w}\|_2^2 \quad \text{s.t. } A\boldsymbol{x} = \boldsymbol{b}, \tag{1.3}$$

where β is a positive parameter and $\boldsymbol{w} \in \mathbb{R}^n$ is the referenced signal that models the side information. For simple, $\boldsymbol{w}$ is assumed to obey $\mathrm{Supp}(\boldsymbol{w}) \subset \mathrm{Supp}(\widehat{\boldsymbol{x}})$, where $\mathrm{Supp}(\boldsymbol{w}) = \{i : |\boldsymbol{w}_i| \neq 0, i = 1, 2, \cdots, n\}$. Based on some statistical tools, the authors affirmatively answer how many measurements one required to ensure the exact recovery of any k-sparse signal $\widehat{\boldsymbol{x}}$. Some convincing experiments are also conducted to support their claims.

In this paper, we revisit the above ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods for exact sparse recovery with side information. Different from the pioneering work of [27], this paper aims at investigating both the ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods in a deterministic way. To do so, by means of the powerful null space property (NSP), we established two kind of deterministic sufficient and necessary condition for these two methods. Our obtained theoretical results not only well complement the previous work [27] that was based on the statistical analysis, but also surprisingly find that the sharp exact recovery conditions of model (1.1) are still suitable for the ℓ_1-ℓ_1 model (1.2). Moreover, the resultant numerical experiments show that the recovery performance of the ℓ_1-ℓ_1 method is superior to other methods in terms of the number of the measurements required by incorporating the side information.

The rest of this paper is organized as follows: the main theoretical results are presented in Section 2, and the resultant numerical experiments are provided in Section 3. Finally, we conclude this paper in Section 4.

2 The deterministic analysis of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods

Our main results will be presented in this section, which include the exact recovery guarantees of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods. Before moving on, we first introduce the following two key definitions.

Definition 2.1 (NSP, see, e.g., [28]). *For any subsets $K \subset \{1, 2, ...n\}$ with $|K| \leq k$ and any $\boldsymbol{h} \in Ker(A)\backslash\{0\}$, we say $A \in \mathbb{R}^{m\times n}$ satisfies the k-order NSP if it holds that*

$$\|\boldsymbol{h}_K\|_1 < \|\boldsymbol{h}_{K^c}\|_1. \tag{2.4}$$

Furthermore, if it holds that

$$\|\boldsymbol{h}_K\|_1 \leq \alpha\|\boldsymbol{h}_{K^c}\|_1 \tag{2.5}$$

for certain $0 < \alpha < 1$, then we say A satisfies the k-order stable NSP with constant α.

Definition 2.2 (Restricted isometry property, see, e.g., [3]). *A matrix A is said to satisfy the k-order restricted isometry property (RIP) if there exists $0 < \delta < 1$ such that*

$$(1 - \delta)\|\boldsymbol{h}_K\|_2^2 \leq \|A\boldsymbol{h}_K\|_2 \leq (1 + \delta)\|\boldsymbol{h}_K\|_2^2, \tag{2.6}$$

holds for all k-sparse signals $\boldsymbol{h} \in \mathbb{R}^n$ and subsets $K \subset \{1, 2, ...n\}$ with $|K| \leq k$. Moreover, the smallest δ obeying (2.6) is denoted by δ_k, i.e., the known k-order restricted isometry constant (RIC).

Now we are ready to present our main results. We start with giving the first one, which provides a sufficient and necessary condition for model (1.2) to guarantee the exact recovery of any k-sparse signal, and one can find it in Theorem 2.3.

Theorem 2.3. *The ℓ_1-ℓ_1 model (1.2) has a unique k-sparse solution if and only if A obeys the k-order NSP.*

Remark 2.4. The k-order NSP has been demonstrated to be a necessary and sufficient condition for the classical ℓ_1 method to ensure the exactly k-sparse signal recovery. Surprisingly, according to our Theorem 2.3, this condition also holds true for the ℓ_1-ℓ_1 model (1.2). On the other hand, it has also been shown in [29, 30] that if A obeys the k-order stable NSP with constant α, then α can be expressed by tk-order RIC δ_{tk} with $t > 1$ as follows:

$$\alpha = \frac{\delta_{tk}}{\sqrt{\left(1 - (\delta_{tk})^2\right)(t-1)}}.$$

If one further restricts $\alpha < 1$, then we will get

$$\delta_{tk} < \sqrt{\frac{t-1}{t}}. \tag{2.7}$$

Note that condition (2.7) has been proved to be sharp for the classical ℓ_1 to exactly recover any k-sparse signal. Again, condition (2.7) is also suitable to the ℓ_1-ℓ_1 model (1.2). As far as we know, the RIC-based sufficient conditions have not been established for the ℓ_1-ℓ_1 method before.

Remark 2.5. It should be noted that, before presenting Theorem 2.3, we have assumed that $\boldsymbol{w}$ obeys $\mathrm{Supp}(\boldsymbol{w}) \subset \mathrm{Supp}(\widehat{\boldsymbol{x}})$. In other words, Theorem 2.3 may not hold when the condition $\mathrm{Supp}(\boldsymbol{w}) \subset \mathrm{Supp}(\widehat{\boldsymbol{x}})$ is violated. This also indicates that one should carefully select the referenced signal $\boldsymbol{w}$. Once a bad $\boldsymbol{w}$ is used, the recovery performance of the model (1.2) may be unstable. In such cases, it is suggested to select the parameter β as small as possible, and only in this can the negative influence of the improper $\boldsymbol{w}$ be removed. Given this, it becomes very important and necessary to establish some general selection strategies for the referenced signal $\boldsymbol{w}$, which will be one of our future work.

Proof of Theorem 2.3. First, we prove the sufficiency. Pick any k-sparse vectors $\widehat{\boldsymbol{x}}$. Let $K = Supp(\widehat{\boldsymbol{x}})$. Since $\boldsymbol{h} \in \mathrm{Ker}(A)$, it holds that $A(\widehat{\boldsymbol{x}} + \boldsymbol{h}) = A\widehat{\boldsymbol{x}} = \boldsymbol{b}$. And

$$\begin{aligned} \|\widehat{\boldsymbol{x}} + \boldsymbol{h}\|_1 &+ \beta\|\widehat{\boldsymbol{x}} - \boldsymbol{w} + \boldsymbol{h}\|_1 = \|\widehat{\boldsymbol{x}}_K + \boldsymbol{h}_K\|_1 + \beta\|\widehat{\boldsymbol{x}}_K + \boldsymbol{h}_K - \boldsymbol{w}\|_1 + \|\boldsymbol{h}_{K^c}\|_1 + \beta\|\boldsymbol{h}_{K^c}\|_1 \\ &\geq \|\widehat{\boldsymbol{x}}_K\|_1 - \|\boldsymbol{h}_K\|_1 + \beta\|\widehat{\boldsymbol{x}}_K - \boldsymbol{w}\|_1 - \beta\|\boldsymbol{h}_K\|_1 + \|\boldsymbol{h}_{K^c}\|_1 + \beta\|\boldsymbol{h}_{S^c}\|_1 \\ &= \|\widehat{\boldsymbol{x}}_K\|_1 + \beta\|\widehat{\boldsymbol{x}}_K - \boldsymbol{w}\|_1 + (1+\beta)(\|\boldsymbol{h}_{K^c}\|_1 - \|\boldsymbol{h}_K\|_1), \end{aligned} \tag{2.8}$$

where we have used the triangle inequality in the first inequality. Recall that $\|\boldsymbol{h}_K\|_1 < \|\boldsymbol{h}_{K^c}\|_1$, $\beta > 0$, and we have assumed that A obeys the k-order NSP, we get $\|\widehat{\boldsymbol{x}} + \boldsymbol{h}\|_1 + \beta\|\widehat{\boldsymbol{x}} + \boldsymbol{h} - \boldsymbol{w}\|_1 > \|\widehat{\boldsymbol{x}}\|_1 + \beta\|\widehat{\boldsymbol{x}} - \boldsymbol{w}\|_1$. Hence, the sufficiency is proved.

Now, we prove the necessary. To do so, we first assume that the ith nonzero component of $\widehat{\boldsymbol{x}}$ obeys $\widehat{\boldsymbol{x}}_i = -\mathrm{sign}(\|\boldsymbol{h}\|_\infty)$ for $i \in K$. Then, we can obtain the following properties: $\|\widehat{\boldsymbol{x}}\|_0 \leq k$,

$\|\widehat{\boldsymbol{x}}_K+\tau\boldsymbol{h}_K\|_1=\|\widehat{\boldsymbol{x}}_K\|_1-\|\tau\boldsymbol{h}_K\|_1$ holds for all $0<\tau\leq 1$. Now, by replacing $\boldsymbol{h}$ in (2.8) with $\tau\boldsymbol{h}$, and noting that $\|\widehat{\boldsymbol{x}}+\tau\boldsymbol{h}\|_1+\beta\|\widehat{\boldsymbol{x}}+\tau\boldsymbol{h}-\boldsymbol{w}\|_1>\|\widehat{\boldsymbol{x}}\|_1+\beta\|\widehat{\boldsymbol{x}}-\boldsymbol{w}\|_1$ since $\widehat{\boldsymbol{x}}$ is the exact solution, we can easily deduce that

$$(1+\beta)(\|\tau\boldsymbol{h}_{K^c}\|_1-\|\tau\boldsymbol{h}_K\|_1)>0$$

for any $0<\tau\leq 1$ and $\beta>0$, which requires (2.4) to hold. □

In what follows, we establish the stable NSP condition of order k for the ℓ_1-ℓ_2 model (1.3).

Theorem 2.6. *Assume that $\|\widehat{\boldsymbol{x}}\|_\infty$ is fixed. $\boldsymbol{w}$ is a side information. The ℓ_1-ℓ_2 model* (1.3) *has a unique k-sparse solution $\widehat{\boldsymbol{x}}$ with the fixed $\|\widehat{\boldsymbol{x}}\|_\infty$ and $\|\boldsymbol{w}\|_\infty$ if and only if A obeys k-order stable NSP with α being*

$$\alpha=\frac{1}{1+\beta\left(\|\widehat{\boldsymbol{x}}\|_\infty+\|\boldsymbol{w}\|_\infty\right)} \tag{2.9}$$

Remark 2.7. Compared with the previous NSP condition for the ℓ_1-ℓ_1 model (1.2), the obtained stable NSP for ℓ_1-ℓ_2 model (1.3) performs a bit loose. Besides, it is also affected by the infinite norm of the desired k-sparse signal $\widehat{\boldsymbol{x}}$ and the referenced signal $\boldsymbol{w}$. From this point of view, the ℓ_1-ℓ_2 model is less effective than the ℓ_1-ℓ_1 model in theoretical aspect. If one takes a close look at the ℓ_1-ℓ_2 model (1.3), one will find that it is a strong convex optimization problem, which is much easier to be solved than the convex model (1.2). When the problem scale is large and the recovery precision is not in urgent need, it is suggested to use model (1.3) to construct the fast algorithm to realize the sparse signal recovery.

Proof of Theorem 2.6. Our proof is partially inspired by [31]. We start with proving the sufficiency. Pick any k-sparse vectors $\widehat{\boldsymbol{x}}$. Let $K=Supp(\widehat{\boldsymbol{x}})$. Sicne $\boldsymbol{h}\in\mathrm{Ker}(A)$, we first have $A(\widehat{\boldsymbol{x}}+\boldsymbol{h})=A\widehat{\boldsymbol{x}}=\boldsymbol{b}$, and

$$\begin{aligned}
&\|\widehat{\boldsymbol{x}}+\boldsymbol{h}\|_1+\frac{\beta}{2}\|\widehat{\boldsymbol{x}}+\boldsymbol{h}-\boldsymbol{w}\|_2^2=\|\widehat{\boldsymbol{x}}_K+\boldsymbol{h}_K\|_1+\frac{\beta}{2}\|\widehat{\boldsymbol{x}}_K+\boldsymbol{h}_K-\boldsymbol{w}\|_2^2+\|\boldsymbol{h}_{K^c}\|_1+\frac{\beta}{2}\|\boldsymbol{h}_{K^c}\|_2^2\\
&\geq\|\widehat{\boldsymbol{x}}_K\|_1-\|\boldsymbol{h}_K\|_1+\frac{\beta}{2}\|\widehat{\boldsymbol{x}}_K-\boldsymbol{w}\|_2^2+\frac{\beta}{2}\|\boldsymbol{h}_K\|_2^2+\beta\langle\widehat{\boldsymbol{x}}_K-\boldsymbol{w},\boldsymbol{h}_K\rangle+\|\boldsymbol{h}_{K^c}\|_1+\frac{\beta}{2}\|\boldsymbol{h}_{K^c}\|_2^2\\
&\geq[\|\widehat{\boldsymbol{x}}_K\|_1+\frac{\beta}{2}\|\widehat{\boldsymbol{x}}_K-\boldsymbol{w}\|_2^2]+(\|\boldsymbol{h}_{K^c}\|_1-\|\boldsymbol{h}_K\|_1+\beta\langle\widehat{\boldsymbol{x}}_K-\boldsymbol{w},\boldsymbol{h}_K\rangle)+\frac{\beta}{2}\|\boldsymbol{h}\|_2^2\\
&\geq\|\widehat{\boldsymbol{x}}_K\|_1+\frac{\beta}{2}\|\widehat{\boldsymbol{x}}_K-\boldsymbol{w}\|_2^2+\|\boldsymbol{h}_{K^c}\|_1-\|\boldsymbol{h}_K\|_1-\beta(\|\widehat{\boldsymbol{x}}\|_\infty+\|\boldsymbol{w}\|_\infty)\|\boldsymbol{h}_K\|_1+\frac{\beta}{2}\|\boldsymbol{h}\|_2^2
\end{aligned} \tag{2.10}$$

where we have used $\|\boldsymbol{h}_K\|_2^2 + \|\boldsymbol{h}_{K^c}\|_2^2 = \|\boldsymbol{h}\|_2^2$ and $\langle \widehat{\boldsymbol{x}}_K - \boldsymbol{w}, \boldsymbol{h}_K \rangle \geq -(\|\widehat{\boldsymbol{x}}\|_\infty + \|\boldsymbol{w}\|_\infty)\|\boldsymbol{h}_K\|_1$ in the second inequality. Since $\|\boldsymbol{h}\|_2^2 > 0$ and $[1 + \beta\,(\|\widehat{\boldsymbol{x}}\|_\infty + \|\boldsymbol{w}\|_\infty)]\|\boldsymbol{h}_K\|_1 < \|\boldsymbol{h}_{K^c}\|_1$, we get $\|\widehat{\boldsymbol{x}}+\boldsymbol{h}\|_1 + \frac{\beta}{2}\|\widehat{\boldsymbol{x}}+\boldsymbol{h}-\boldsymbol{w}\|_2^2 > \|\widehat{\boldsymbol{x}}\|_1 + \frac{\beta}{2}\|\widehat{\boldsymbol{x}}-\boldsymbol{w}\|_2^2$. Hence, we prove that $\widehat{\boldsymbol{x}}$ is the unique minimizer of (1.3).

As for the necessary, it is sufficient to show that for any given nonzero $\boldsymbol{h} \in \mathrm{Ker}(A)$ and K with $|K| \leq k$ the stable NSP of order k with α given by (2.9) holds. Similarly with Theorem 2.3, we can obtain $\|\widehat{\boldsymbol{x}}\|_0 \leq k$, $\|\widehat{\boldsymbol{x}}_K + \tau\boldsymbol{h}_K\|_1 = \|\widehat{\boldsymbol{x}}_K\|_1 - \|\tau\boldsymbol{h}_K\|_1$. Furthermore, assume that the scales $\widehat{\boldsymbol{x}}$ and $\boldsymbol{w}$ have fixed values of $\|\widehat{\boldsymbol{x}}\|_\infty$ and $-\|\boldsymbol{w}^0\|_\infty$, respectively, then we get $\langle \widehat{\boldsymbol{x}}_K - \boldsymbol{w}, \tau\boldsymbol{h}_K \rangle = -(\|\widehat{\boldsymbol{x}}\|_\infty + \|\boldsymbol{w}\|_\infty)\|\boldsymbol{h}_K\|_1$ for any $0 < \tau \leq 1$. Now, we use $\tau\boldsymbol{h}$ in the equation array instead of $\boldsymbol{h}$ and observe that both of inequalities of (2.10) now hold with equality. Since $\widehat{\boldsymbol{x}}$ is the exact recovery, it requires $\|\widehat{\boldsymbol{x}} + \tau\boldsymbol{h}\|_1 + \frac{\beta}{2}\|\widehat{\boldsymbol{x}} + \tau\boldsymbol{h} - \boldsymbol{w}\|_2^2 > \|\widehat{\boldsymbol{x}}\|_1 + \frac{\beta}{2}\|\widehat{\boldsymbol{x}} - \boldsymbol{w}\|_2^2$, so we get

$$[\|\tau\boldsymbol{h}_{K^c}\|_1 - \|\tau\boldsymbol{h}_K\|_1 - \beta(\|\widehat{\boldsymbol{x}}\|_\infty + \|\boldsymbol{w}\|_\infty)\|\tau\boldsymbol{h}_K\|_1] + \frac{\beta}{2}\|\tau\boldsymbol{h}\|_2^2 > 0$$

for any $0 < 1 \leq \tau$, which proves the necessary. □

3 Numerical simulations experiments and results

In this section, We conduct some numerical experiments are carried out to demonstrate the performance of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 models with side information.An IRLS1 algorithm is frist proposed to solve the induced ℓ_1 minimization problem (1.1), We then compare ℓ_1-ℓ_1 and ℓ_1-ℓ_2 model analysis method with side information.

3.1 Methods and their theoretical analysis

In order to solve the ℓ_1-ℓ_1 (1.2) and ℓ_1-ℓ_2 (1.3) analysis problem with side information. We derive an efficient analysis-style IRLS1 algorithm.According to the analysis of sparsity and measurement, we compare the performance of the problems (1.2), (1.3), and the ℓ_1 minimization problem (1.1) as well as the unconstrained smoothed ℓ_q minimization (with $q = 0.5$) in [32], which takes the form of $\min \|\boldsymbol{x}\|_{q,\varepsilon}^q + \frac{1}{2\lambda}\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$. The recovery is regarded as successful if $\|\boldsymbol{x} - \boldsymbol{x}^*\|_2 \leq 10^{-3}\|\boldsymbol{x}^*\|_2$.

Algorithm 1 The modified IRLS1 algorithm

1 : Input the vector b, the measurement $A \in \mathbb{R}^{200\times1000}$, the known support set T.

2 : Choose appropriate parameters λ $(0 < \lambda < 1\ )$.

3 : Initialize x^0, satisfying $Ax^0 = b$, set $\varepsilon_0 = 1$.

4 : For t = 0,1,2, . . . , solve the following question for $x^{(t)}$.

$|\boldsymbol{x}^{(t)}\|_{q,\varepsilon}^q + \frac{1}{2\lambda}\|A\boldsymbol{x}^{(t+1)} - \boldsymbol{b}\|_2^2$ =0 .

5 : When x satisfying certain stopping criterion , x will be output, otherwise ,it needs to carry out next step.

6 : Set $t = t + 1$ and return the fourth step.

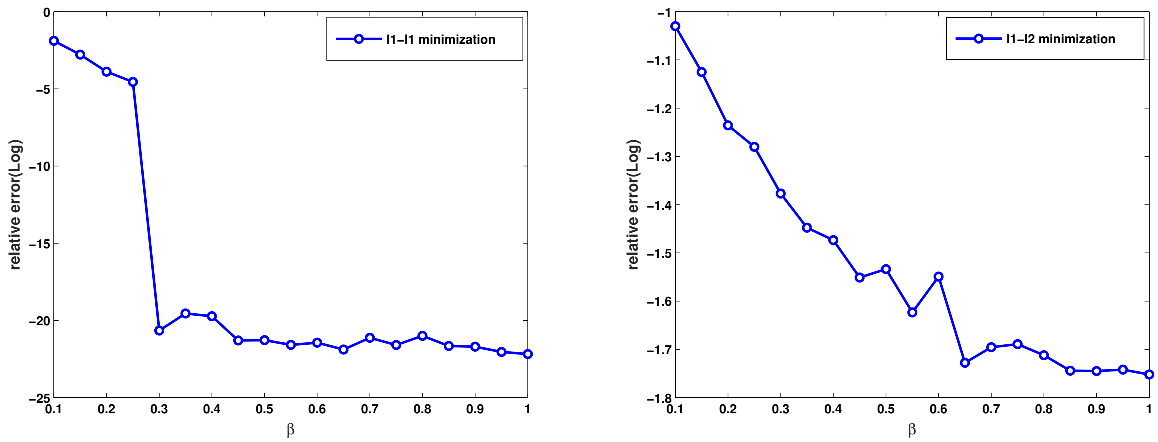
7 : Output vector $x^{(t)} \in\ R^n$.

3.2 Experimental settings

Throughout the experiments, the k-sparse signals $\boldsymbol{x}^*$ $(\boldsymbol{x}^* \in \mathbb{R}^{1000})$ is generated with their nonzero components being chosen from a standard normal distribution and the known side information $\boldsymbol{w}$ is generated by $\boldsymbol{w} = \boldsymbol{x}^* + \boldsymbol{z}$, where $\boldsymbol{z}$ is a 28-sparse signal whose nonzero components are drawn from Gaussian distribution $\mathcal{N}\,(0, 0.8^2)$. Besides, we always assume that the support of $\boldsymbol{x}^*$ and $\boldsymbol{z}$ coincided in 22 positions and differ in 6, and we also limit $\boldsymbol{x}^*$ and $\boldsymbol{w}$ as $\frac{\|\boldsymbol{x}^*-\boldsymbol{w}\|_2}{\|\boldsymbol{x}^*\|_2} \le 0.5$. The measurement matrix A has i.i.d components drawn from a standard Gaussian distribution with normalised columns. We average 100 repetitions of each experiment varying the amplitude of the signals.In the analysis of problems (1.2) and (1.3), should be chosen the appropriate parameter β . Thus, the performance of different value β $(0 < \beta \le 1)$ is tested. First, the measurement matrix $A \in \mathbb{R}^{200\times1000}$ and sparsity $k = 70$ will be set.

3.3 Experimental results

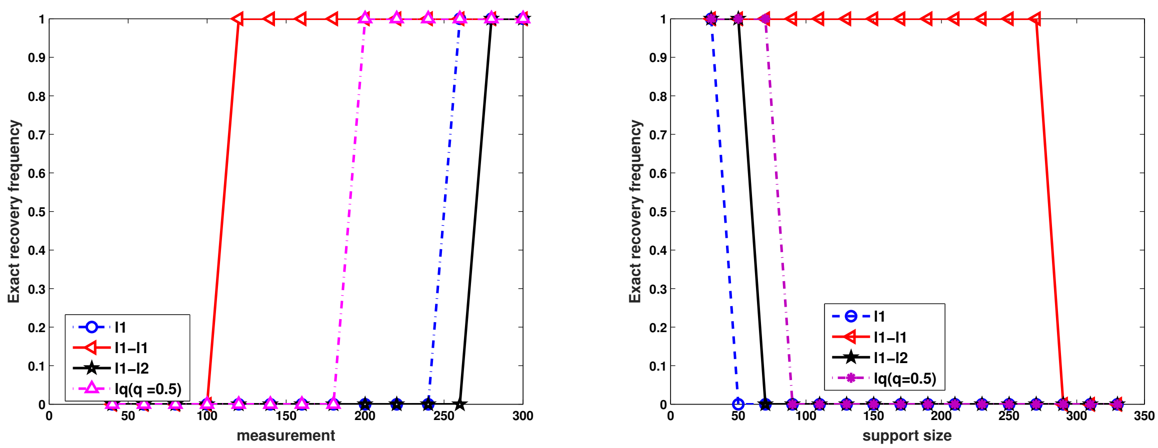
In order to find the value of parameters that minimize the relative error,we conduct three sets of trials.In Figure 1, the relative error of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 models with side information as a function of the parament β are constructed, where the relative error is defined as $\frac{\|\boldsymbol{x}-\boldsymbol{x}^*\|_2}{\|\boldsymbol{x}^*\|_2}$ for a recovered $\boldsymbol{x}^*$. The illustrate that the relative error is the smallest of $\beta = 1$ and the effect around $\beta = 1$ is more stable than others. So we choose $\beta = 1$ uniformly in the following experiments.In Figure 2, we plot the exact recovery performance of four models affected by the number of the



(a) ℓ_1-ℓ_1 minimization (b) ℓ_1-ℓ_2 minimization

Figure 1: Performance of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 minimization with side information.

measurements and the support size the desired signals, respectively. In figure 2(2), the support size of the desired signals are fixed to be $k = 70$. From the Figure 2(a), we can obtain that

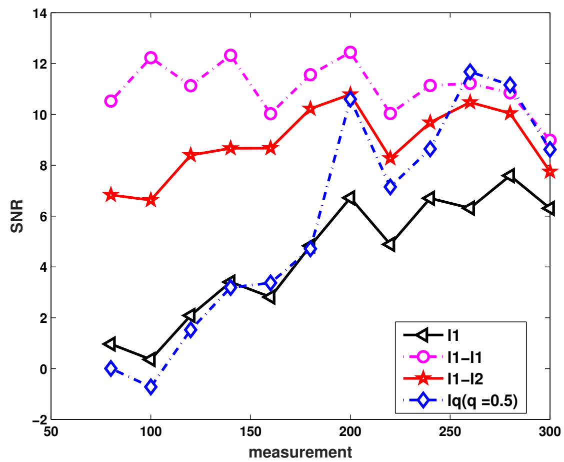


(a) (b)

Figure 2: Comparison results of recoverability: ℓ_1-ℓ_1 and ℓ_1-ℓ_2 models with side information compares with $\ell_q(q = 0.5)$, ℓ_1, varying sparsity and measurement.

ℓ_1-ℓ_1 minimization is superior to other models, and it only needs the fewest measurements to reconstruct the desired sparse signals. Besides, figure 2(b) plots the success frequency versus sparsity K for these four models. In this set of experiments, we always set $A \in \mathbb{R}^{200\times 1000}$. From Figure 2(b), it is easy to conclude that the ℓ_1-ℓ_1 model with side information performs best among all the models.

In the last experiment, the reconstruction signal to noise ratio (SNR) is designed, varying the measurement level where SNR is calculated as $20\log_{10}\left(\frac{\|\boldsymbol{x}\|_2}{\|\boldsymbol{x}-\boldsymbol{x}^*\|_2}\right)$,and one can find the results in figure 3. We compare the effect of (1.2), (1.3) with ℓ_1 and $\ell_q(q=0.5)$ minimization in noise situation. We discover SNR increase with large size of the measurement level. In addition, it is obvious that ℓ_1-ℓ_1 model with side information give better performance than others.



(a)

Figure 3: Performance of ℓ_1-ℓ_1 and ℓ_1-ℓ_2 minimization with side information in terms of SNR, varying the measurement level.

4 Conclusion and discussion

In this paper, we establish two NSP-based sufficient and necessary conditions for two ℓ_1-ℓ_1 and ℓ_1-ℓ_2 methods in the case that the side information of the desired signal becomes available. These deterministic theoretical results provide good complement for the previous work on these two methods that are based on the statistical analysis. Besides, some experiments demonstrate ℓ_1-ℓ_1 method with side information is better than ℓ_1-ℓ_2, $\ell_q(q=0.5)$, ℓ_1 methods in terms of signal recovery and the measurement required.

Note that we only consider the general sparse recovery with side information, which indicates that the obtained theoretical results can not be directly applied to some structured sparse recovery cases. However, by integrating the proposed two methods into some structured sparse models, our established results actually can be easily extened to deal with the structured sparse recovery with side information. On the other hand, it is also very important to estsblish some theoretical results that are based on the coherence tool due to its simplicity. All these considerations will naturally be our future works.

Abbreviations

CS: Compressed sensing;IHT: Iterative hard thresholding;NSP: null space property;RIP: restricted isometry property;RIC: restricted isometry constant;SNR: signal to noise ratio; MRI: magnetic resonance imaging;OMP: orthogonal matching pursuit;IRLS: iteratively reweighted least squares

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors read and approved the final manuscript.

Acknowledgment

This paper is subsidized by the project of the Natural science of china (No. 62063028).

Availability of data and materials

Please contact any of the authors for data and materials.

Author details

[1]Faculty of Computer and Information Science, Southwest University, Chongqing, 400715, China, [2]School of information Technology,Deakin University,VIC,3217, Australia.

References

[1] D.L. Donoho, Compressed sensing. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.

[2] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.

[3] E.J. Candès, T. Tao, Decoding by linear programming. IEEE Transactions on Information Theory, 2005, 51(12): 4203-4215.

[4] J. Wright, Y. Ma and J. Mairal, et al., Sparse representation for computer vision and pattern recognition, Proceedings of the IEEE. 2010, 98(6), 1031-1044.

[5] R. Baraniuk and P. Steeghs, Compressive radar imaging, 2007 IEEE Radar Conference, 2007, 128-133

[6] S. Archana, K.A. Narayanankutty, A. Kumar, Brain mapping using compressed sensing with graphical connectivity maps, International Journal of Computer Applications, 2012, 54(11) :35-39.

[7] T. Wan and Z.C. Qin, An application of compressive sensing for image fusion, International Journal of Computer Mathematics, 2011, 88(18): 3915-3930.

[8] M. Lustig, D. Donoho, and J. Pauly, Sparse MRI: The application of compressed sensing to rapid MR imaging, Magnetic Resonance in Medicine, 2007, 58(6): 1182-1195.

[9] M. Duarte, M. Davenport, D. Takhar, et al., Single-pixel imaging via compressive sampling, IEEE Signal Processing Magazine, 2008, 25(2): 83-91.

[10] R.G. Baraniuk, Compressive sensing, IEEE Signal Processing Magazine, 2007, 24(4): 118-121.

[11] E.J. Candès, M.B. Wakin, S.P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization. Journal of Fourier Analysis and Applications, 2008, 14(5): 877-905.

[12] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.

[13] H. Esmaeili, M. Rostami, M. Kimiaei, Combining line search and trust-region methods for ℓ_1 minimization, International Journal of Computer Mathematics, 2018, 95(10): 1950-1972.

[14] T. Cai, L. Wang, and G.W. Xu, Stable recovery of sparse signals and an oracle inequality, IEEE Transactions on Information Theory, 2010, 56(7): 3516-3522

[15] T. Cai, and A.R. Zhang, Compressed sensing and affine rank minimization under restricted isometry, IEEE Transactions on Signal Processing, 2013, 61(13): 3279-3290

[16] R. Zhang, and S. Li, A proof of conjecture on restricted isometry property constants $\delta_{tk}(0 < t < \frac{4}{3})$, IEEE Transactions on Information Theory, 2018, 64(3): 1699-1705

[17] G.H. Chen, J. Tang, S. Leng, Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. Medical Physics, 2008, 35(2): 660-663.

[18] D. Needell, R. Saab, T. Woolf, Weighted ℓ_1-minimization for sparse recovery under arbitrary prior information, Information and Inference: A Journal of the IMA, 2017, 6(3): 284-309.

[19] L.F. Polania, K.E. Barner, Iterative hard thresholding for compressed sensing with partially known support, IEEE International Conference on Acoustics, Speech and Signal Processing, 2011: 4028-4031

[20] R.E. Carrillo, L.F. Polania and K.E. Barner, Iterative algorithms for compressed sensing wth partially known support, IEEE International Conference on Acoustics, Speech and Signal Processing, 2010: 3654-3657. 23(3): 3654-3657.

[21] C.J. Miosso, R. Borries, M. Argaez, et al., Compressive sensing reconstruction with prior information by iteratively reweighted least-square, IEEE Transactions on Signal Processing, 2009, 57(6): 2424-2431.

[22] H.M. Ge, W. G. Chen, M. K. Ng, New RIP bounds for recovery of sparse signals with partial support information via weighted ℓ_p -minimization, IEEE Transactions on Information Theory, 2020, 66(6): 3914-3928.

[23] L. Weizman, Y. Eldar, D. Bashat, Compressed sensing for longitudinal MRI: An adaptive-weighted approach, Medical Physics, 2015, 42(9): 5195-5208.

[24] V. Stankovic, L. Stankovic, S. Cheng, Compressive image sampling with side information, IEEE International Conference on Image Processing, 2009: 3001-3004.

[25] L.-W. Kang, C.-S. Lu, Distributed compressive video sensing, IEEE International Conference on Acoustics, Speech, and Signal Processing, 2009: 1169-1172.

[26] A. Charles, M. Asif, J. Romberg, C. Rozell, Sparsity penalties in dynamical system estimation, IEEE Conference on Information Sciences and Systems, 2011: 1-6.

[27] J.F.C. Mota, N. Deligiannis, M.R.D. Rodrigues, Compressed sensing with prior information: Strategies, geometry, and bounds, IEEE Transactions on Information Theory, 2017, 63(7): 4472-4496.

[28] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing, Birkhäuser, Basel, Switzerland, 2013

[29] H.M. Ge, J.M. Wen, et al., Stable sparse recovery with three unconstrained analysis based approaches, Available: http://alpha.math.uga.edu/ mjlai/papers/20180126.pdf

[30] W.D. Wang, J.J. Wang, Robust recovery of signals with partially known support information using weighted BPDN, Analysis and Applications, 2020, 18(6): 1025-1055.

[31] M.J. Lai, W.T. Yin, Augmented ℓ_1 and nuclear-norm models with a globally linearly convergent algorithm, SIAM Journal on Imaging Sciences, 2012, 6(2): 1059-1091.

[32] M.J. Lai, Y.Y. Xu, W.T. Yin, Improved iteratively reweighted least squares for unconstrained smoothed ℓ_q minimization, SIAM Journal on Numerical Analysis. 2013, 51(2): 927-957.