New sufficient conditions for recovery in compressed sensing

井上, 寛

https://doi.org/10.15017/1441043

出版情報：九州大学, 2013, 博士（機能数理学）, 課程博士
バージョン：

New sufficient conditions for recovery in compressed sensing

井上　寛

New sufficient conditions for recovery in compressed sensing

Hiroshi Inoue

Graduate School of Mathematics

Kyushu University

2014

Acknowledgements

I have been able to complete this doctoral thesis with a lot of support from many people. First of all, I would like to express my sincere gratitude of my supervisor, Professor Yoshiyuki Ninomiya, for his valuable comments, suggestions and encouragement through the whole of my research.

I would like to acknowledge valuable advice and various helps given by Professor Ryuei Nishii, Yoshihiko Maesono, Hiroki Masuda and the members of group of statistics in Kyushu University. I wish to express my great appreciation to Professor Sadanori Konishi of Chuo University for his enormous advice.

I also would like to than Professor Hidefumi Kawasaki of Kyushu University and Dr. Matsuoka Hidetoshi and Dr. Hiroshi Ikeda of FUJITSU LABORATORIES LTD. for my internship.

This thesis is supported by Kyushu University Global COE Program and Education and Research Hub for Mathematics-for-Industry.

Finally, I would like to give special thanks to my parents for their enormous encouragement and financial assistance, and my firends for enormous support and encouragement.

Hiroshi Inoue

February, 2014, Fukuoka

Contents

Chapter 1

Introduction

Compressed sensing (CS) is a rapidly growing field that has attracted considerable attention in applied mathematics, signal processing, statistics and computer science as well as the broader scientific field. Though its initial development was only a few years ago, many results have been obtained both of theoretical and practical nature.

CS asserts that one can recover certain signals and images $\boldsymbol{x}$ from only a few samples or measurements $\boldsymbol{y}$. Here, we consider

$$\boldsymbol{y} = A\boldsymbol{x}, \quad \boldsymbol{x} \in \boldsymbol{R}^n, \tag{1.1}$$

where A is an $m \times n$ matrix. We are interested in the ill-posed problem when $m < n$. To be more specific, in CS, we do not measure $\boldsymbol{x}$ directly, but rather acquire $m < n$ linear measurements of the form (1.1) using an $m \times n$ CS matrix A. Ideally, the matrix is designed to reduce the number of measurements as much as possible while allowing for recovery of a wide class of signals from their measurement vectors $\boldsymbol{y}$. Thus, we would like to choose $m << n$. However, the matrix A is rank-deficient, meaning that it has a nonempty nullspace. This implies that for any particular signal $\boldsymbol{x}_0$, an infinite number of signals $\boldsymbol{x}$ will yield the same measurements $\boldsymbol{y} = A\boldsymbol{x} = A\boldsymbol{x}_0$ for chosen CS matrix A. In order to reconstruct $\boldsymbol{x}$ with good accuracy, we must assume a special class of signals $\boldsymbol{x}$.

The most important signal structure used in CS is that of sparsity. In detail, the sparsity implies that $\boldsymbol{x}$ has only a small number of nonzero values. More generally, CS ideas can be applied when a suitable representation of $\boldsymbol{x}$ is sparse or nearly sparse. The surprising results in CS theory is that if $\boldsymbol{x}$ is s-sparse, that is, it has at most s nonzero elements, then it can be recovered exactly from $\boldsymbol{y} = A\boldsymbol{x}$ under some conditions to A. Furthermore, these can easily be applied to noisy recovery:

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{z}, \tag{1.2}$$

where A is a $m \times n$ matrix $(m < n)$ and $\boldsymbol{z}$ is an unknown noise term.

Our goal is to reconstruct an unknown signal $\boldsymbol{x}$ based on A and $\boldsymbol{y}$ are given. Then we consider reconstructing $\boldsymbol{x}$ as the solution $\boldsymbol{x}^\star$ to the optimization problem:

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|\boldsymbol{y} - A\boldsymbol{x}\|_2 \leq \varepsilon, \tag{1.3}$$

where ε is an upper bound on the the size of the noisy contribution. In CS theory, a crucial issue is to research good conditions in order to achieve our goal:

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq C_0\|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + C_1\varepsilon, \tag{1.4}$$

where C_0 and C_1 are small constants and $\boldsymbol{x}_s$ is the vector consisting of the largest entries of $\boldsymbol{x}$ in magnitude.

Many papers in CS were devoted to the analysis of theoretical guarantees on the CS matrix A in order to enable stable recovery. The field of CS grew out of the work of E.J. Candès, J. Romberg, T. Tao and of D. Donoho, who showed that a finite dimensional signal having a sparse or compressible representation can be recovered from a small set of linear, non-adaptive measurements. In particular, one of the most generally known condition for CS theory is the restricted isometry property(RIP) introduced by E.J. Candès and T. Tao (Candès and Tao, 2005). The RIP needs that the subsets of columns of A for all locations in $\{1, 2, \cdots, n\}$ behave nearly orthonormal system. In detail, a matrix A satisfies the RIP of order s if there exists a constant δ with $0 < \delta < 1$ such that

$$(1 - \delta)\|\boldsymbol{a}\|_2^2 \leq \|A\boldsymbol{a}\|_2^2 \leq (1 + \delta)\|\boldsymbol{a}\|_2^2 \tag{1.5}$$

for all s-sparse vectors $\boldsymbol{a}$. A vector is said to be s-sparse vector if it has at most s nonzero entries. The minimum δ satisfying the above restrictions is said to be the restricted isometry constant and is denoted by δ_s. The CS theory have been studied by many researchers under the assumption of RIP. However, the RIP requires a bounded condition number for all submatrices built by selecting s arbitrary columns and the spectral norm of a matrix is generally difficult to calculate. Therefore, it seems useful to weaken the condition of RIP and to apply to the CS theory.

Theme 1「Sufficient conditions for CS-recovery」

The first method is of rescaling. We assume that any s column vectors of A is linearly independent, that is, $A_T^\star A_T$ is invertible for any location T in $\{1, 2, \cdots, n\}$ with $|T| = s$ and define the s-restrictly norm constant $r_s(A) \equiv \max\{\|A_T\|;\ \ T \subset \{1, 2, \cdots, n\}, |T| = s\}$, where A_T is the $m \times |T|$ matrix composed of these columns for T. Then we consider the following (1.6) instead of (1.2):

$$\tilde{\boldsymbol{y}} = \tilde{A}\boldsymbol{x} + \tilde{\boldsymbol{z}}, \tag{1.6}$$

where $\tilde{\boldsymbol{y}} \equiv \frac{\boldsymbol{y}}{r_s}$, $\tilde{A} \equiv \frac{A}{r_s}$ and $\tilde{\boldsymbol{z}} \equiv \frac{\boldsymbol{z}}{r_s}$. And we have the following result:

Theorem A. If A is s-restrictly ($2s$-restrictly) invertible, then $\tilde{A}$ obeys the RIP of order s ($2s$). In addition, if $\tilde{\delta}_s < 0.472$ ($\tilde{\delta}_{2s} < 0.661$), (1.4) holds, where $\tilde{\delta}_s$ ($\tilde{\delta}_{2s}$) is the restricted isometry constant for $\tilde{A}$ of order s ($2s$).

Furthermore, we show that if $\tilde{\delta}_s < 0.5$, or $\tilde{\delta}_{2s} < 0.828$, then (1.4) holds by applying the results of (Cai and Zhang, 2013a, 2013b). The above sufficient conditions ($\tilde{\delta}_s < 0.472$ and $\tilde{\delta}_{2s} < 0.661$) are less than the sufficient conditions by applying the results of (Cai and Zhang, 2013a, 2013b), however, their error bounds are more accurate in case that the restricted isometry constants are small.

Theme 2「A generalization of the restricted isometry property and applications to compressed sensing」

The second method is of using the notion of weak RIP. E.J. Candès and Y. Plan have introduced the notion of weak RIP which is a generalization of RIP as follows (Candès and Plan, 2010): Let $T_0 \subset \{1, 2, \cdots, n\}$ with $|T_0| = s$ and $1 < r < s$. A obeys the weak RIP with respect to T_0 of order r if there exists $0 < \delta < 1$ such that for any subset $R \subset T_0^c$

with $|R| \leq r$,

$$(1-\delta) \parallel \boldsymbol{x} \parallel_2^2 \leq \parallel A\boldsymbol{x} \parallel_2^2 \leq (1+\delta) \parallel \boldsymbol{x} \parallel_2^2 \tag{1.7}$$

for all $\boldsymbol{x} \in \boldsymbol{R}^n$ with supp $\boldsymbol{x} \subset T_0 \cup R$. The minimum of such constants δ is denoted by $\delta_{T_0,r}$. And they have proved that under the assumptions of isotropy property: $E\left(A_{\{k\}}A^{\star}_{\{k\}}\right) = I$ for every row vector $A_{\{k\}}$ of A and incoherence property a random matrix obeys the weak RIP with high probability $1 - 5e^{-\beta}$ if $m \geq C \log n$. Furthermore, by using the weak RIP, they have evaluated stochastically the solution of LASSO which is similar to the solution of CS in a certain sense (Tibshirani, 1996). Roughly speaking the notion of the weak RIP, we choose a suitable location T_0 with $|T_0| = s$ in the columns of the matrix A. Then, the matrix $A_{T_0 \cup R}$ obeys the inequality (1.5) for any subset R of T_0^c with $|R| = r$. We focus the notion of the weak RIP and obtain the following result:

Theorem B. Suppose A obeys the weak RIP with respect to T_0 of order r. If $\delta_{T_0,r} < \frac{1}{1+\sqrt{\frac{2s}{[\frac{r}{2}]}}}$ and $T_0 = \text{supp } \boldsymbol{x}_s$, where $[\cdot]$ is the floor function, then (1.4) holds.

Theme 3「Weak RIP and its application to compressed sensing」

The third propose of this paper is to give a sufficient condition under which A obeys the weak RIP with respect to T_0 of order r and to evaluate the solution of CS by using a correlation relation $\theta_{T_0,r}$ of the location T_0 and T_0^c. We assume the followings: (i) The submatrix A_{T_0} is nearly isometric, that is, there exists a constant δ $(0 < \delta < 1)$ such that $(1-\delta)\|\boldsymbol{x}\|_2^2 \leq \|A_{T_0}\boldsymbol{x}\|_2^2 \leq (1+\delta)\|\boldsymbol{x}\|_2^2$ for each $\boldsymbol{x} \in \boldsymbol{R}^n$ with supp $\boldsymbol{x} \subset T_0$. The minimum of such constants δ is denoted by $\delta(T_0)$; (ii) $A_{T_0^c}$ obeys the RIP of order r. Let $\delta_r(T_0^c)$ denote the restricted isometry constant of $A_{T_0^c}$. In addition, we consider the correlative relationship of the submatrices A_{T_0} and $A_{T_0^c}$. We put $\mu_{T_0,r} \equiv \max_T\{\sup\{| < A\boldsymbol{x},\ A\boldsymbol{y} > |;\ \text{supp } \boldsymbol{x} \subset T_0, \text{supp } \boldsymbol{y} \subset T, \|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1\}$, where T moves all locations of T_0^c with $|T| = r$, and define the coefficient of correlation $\theta_{T_0,r}$ by $\theta_{T_0,r} = \max(\delta(T_0),\ \delta_r(T_0^c),\ \mu_{T_0,r})$. Then we have the following result:

Theorem C. If $\theta_{T_0,r} < \frac{1}{2}$, then A obeys the weak RIP with respect to T_0 of order r. In addition, if $2\theta_{T_0,\frac{r}{5}} + \sqrt{\frac{5s}{2r}}\theta_{T_0,r} < 1$ and $T_0 = \text{supp } \boldsymbol{x}_s$, then (1.4) holds.

It seems more easier to construct matrices obeying the weak RIP by considering $\theta_{T_0,r}$.

Furthermore, we show that if A satisfies the isotropy property, then $\tilde{A} \equiv \frac{A}{\sqrt{m}}$ obeys the weak RIP with high probability and it is applocable to the CS theory.

Theme 4「RIPless theory for compressed sensing」

The fourth is to weaken the notion of weak RIP and to apply it to the CS theory. And we obtain the following result:

Theorem D. Assume the following (i) and (ii):

(i) $\{\boldsymbol{a}_k;\ k \in T\}$ is linearly independent for each $T \subset \{1, 2, \cdots, n\}$ with $|T| = s$,

(ii) $\mu \equiv \max_T \mu_T = \max_{1 \leq i \neq j \leq n} | < \boldsymbol{a}_i, \boldsymbol{a}_j > | < \frac{1}{s \max\left(\| \left(A_T^{\star} A_T \right)^{-1} \|; |T|=s \right)}$.

Then (1.4) holds.

This thesis is organized as follows:

In Chapter 2, we introduce some results which play an important rule for our studies in Chapter 3-6. The first is of Theorem 2.1 obtained by (Candès, 2008) and of Theorem 2.2 and 2.3 which are generalization of Theorem 2.1. These are one of motivation for our studies. The second is of Theorem 2.4 obtained by (Cai, Wang and Xu, 2010b). The third is of Theorem 2.5 obtained by (Mo and Li, 2011). The fourth are of Theorem 2.6 and Theorem 2.7 obtained by (Cai and Zhang, 2013a, 2013b). Furthermore, we introduce Theorem 2.8 obtained by E.J. Candès and Y. Plan using the notion of weak RIP (Candès and Plan, 2010). This is a motivation of our studies for weak RIP and the RIPless stated in Chapter 4, 5 and 6.

In Chapter 3, we establish the sufficient condition for the restricted isometry constant $\tilde{\delta}_k$ ($k \geq s$) of $\tilde{A} \equiv \frac{A}{r_k(A)}$ under the assumption that A is k-restrictly invertible. This is due to our paper "Sufficient conditions for CS-recovery, 2013".

In Chapter 4, we introduce the notion of weak RIP which is a proposed theory to be more useful for data analysis than RIP and evaluate the solution of CS under the assumption of only the weak RIP. This is due to our paper "A generalization of the restricted isometry property and applications to compressed sensing, 2013".

In Chapter 5, the first purpose is to give a sufficient condition under which A obeys the weak RIP with respect to T_0 of order r and to evaluate the solution of CS by using a correlative relationship $\theta_{T_0,r}$ of the locations T_0 and T_0^c. The second is to show that when an $m \times n$ random matrix A satisfies the isotropy property: $E\left(A_{\{k\}}A_{\{k\}}^{\star}\right) = I$ for every row vector $A_{\{k\}}$ of A, $\tilde{A} \equiv \frac{A}{\sqrt{m}}$ always obeys the weak RIP with high probability and it is applicable to the CS theory. This is due to our paper "Weak RIP and its application to Compressed Sensing, 2013".

In Chapter 6, we discuss the theory for RIPless in compressed sensing under the assumptions of linear independence of any s column vectors in A and of correative relationship μ. This is due to our paper "RIPless Theory for Compressed Sensing, 2013".

Chapter 2

Background

The CS theory has been studied using random matrices by many researchers, for example (Baraniuk, Davenport, DeVore and Wakin, 2008, Rudelson, 1999, Rudekson and Vershynin, 2008, Candès and Plan, 2009, Candès and Tao, 2006). On the other hand, it has been shown under the assumption of RIP that l_1 optimization can recover an unknown signal in noiseless case and noisy case under various sufficient conditions on δ_s or δ_{2s}. For example, E.J. Candès and T. Tao have proved that if $\delta_{2s} < \sqrt{2} - 1$, then an unknown signal can be recovered (Candès and Tao, 2005). Later, S. Foucart and M. Lai have improved the bound to $\delta_{2s} < 0.4531$ (Forcart and Lai, 2009). Others, $\delta_{2s} < 0.4652$ is used by S. Foucart (Foucart, 2010), $\delta_{2s} < 0.4721$ for cases such that s is a multiple of 4 or s is very large by T. Cai. $el.al.$ (Cai, Wang and Xu, 2010b), $\delta_{2s} < 0.4734$ for the case such that s is very large by S. Foucart (Foucart, 2010) and $\delta_s < 0.307$ by T. Cai $el.al.$ (Cai, Wang and Xu, 2010b). Q. Mo and S. Li have improved the sufficient condition to $\delta_{2s} < 0.4931$ for general case and $\delta_{2s} < 0.6569$ in some special case such that $n \leq 4s$ (Mo and Li, 2011). In recent paper, T. Cai and A. Zhang have improved the sufficient condition to $\delta_s < 0.333$ for general case (Cai and Zhang, 2013a). T. Cai and A. Zhang have improved the sufficient condition to δ_k in case of $k \geq \frac{4}{3}s$, in particular $\delta_{2s} < 0.707$ (Cai and Zhang, 2013b). In this chapter, we introduce some important results that play basic rules for our studies in Chapter 3-6.

2.1 The Candès result and its generalization

Our studies stated in Chapter 3-6 for compressed sensing were motivated on the Candès paper (Candès, 2008) in which has been obtained the following results:

Theorem 2.1 (Candès, 2008) Assume that A obeys the RIP of order $2s$ and $\delta_{2s} < \sqrt{2}-1$, and put

$$C_0 = \frac{2(\sqrt{2}-1)\delta_{2s}+2}{1-(1+\sqrt{2})\delta_{2s}}, \quad C_1 = \frac{4\sqrt{1+\delta_{2s}}}{1-(1+\sqrt{2})\delta_{2s}}.$$

Then the following hold:

(1) Noiseless recovery. The solution $\boldsymbol{x}^\star$ to (1.3) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_1 \leq C_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1, \tag{2.1}$$

and

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq C_0 \frac{1}{\sqrt{s}} \| \boldsymbol{x} - \boldsymbol{x}_s \|_1, \tag{2.2}$$

In particular, if $\boldsymbol{x}$ is s-sparse, the recovery is exact.

(2) Noisy recovery. Assume that $\| \boldsymbol{z} \|_2 \leq \varepsilon$. Then the solution $\boldsymbol{x}^\star$ to (1.4) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq C_0 \frac{1}{\sqrt{s}} \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 + C_1 \varepsilon. \tag{2.3}$$

We shall roughly state the idea behind E.J. Candès' results. E.J. Candès obtained the above results by taking first the location T_0 of the s largest coefficients of $\boldsymbol{x}$, next the location $T_1 \subset T_0^c$ of the s largest coefficients of $\boldsymbol{h} \equiv \boldsymbol{x} - \boldsymbol{x}^\star$, and repeating this method, and then investigating $\boldsymbol{h}_{T_0 \cup T_1}$ and $\boldsymbol{h}_{(T_0 \cup T_1)^c}$. In this section, we introduce two results obtained by different uses of the RIP which are generalizations of Theorem 2.1. This is a starting point of our studies. Let $s, s' \in \boldsymbol{N}$ with $s < n$ and $s' < n - s$. Under these assumptions, the following main theorems hold:

Theorem 2.2. Assume that A satisfies the RIP of order $\max(s+s', 2s')$ and

$$\begin{aligned} \left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'} < 1, \quad s' \leq s, \\ \delta_{s+s'} + \sqrt{\frac{2s}{s'}}\delta_{2s'} < 1, \quad s' \geq s. \end{aligned} \tag{2.4}$$

Put

$$D_0 = \begin{cases} \sqrt{\frac{s}{s'}}\,\frac{2(\sqrt{2}-1)\delta_{s+s'}+2}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}} & , s' \leq s \\ \sqrt{\frac{s}{s'}}\,\frac{2\left(\sqrt{2}\delta_{2s'}-\delta_{s+s'}+1\right)}{1-\delta_{s+s'}-\sqrt{\frac{2s}{s'}}\delta_{2s'}} & , s' \geq s, \end{cases} \tag{2.5}$$

and

$$D_1 = \begin{cases} \left(1+\sqrt{\frac{s}{s'}}\right)\,\frac{2\sqrt{1+\delta_{s+s'}}}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}} & , s' \leq s \\ \left(1+\sqrt{\frac{s}{s'}}\right)\,\frac{2\sqrt{1+\delta_{s+s'}}}{1-\delta_{s+s'}-\sqrt{\frac{2s}{s'}}\delta_{2s'}} & , s' \geq s. \end{cases} \tag{2.6}$$

Then the following hold:

(1) Noiseless recovery. The solution $\boldsymbol{x}^\star$ to (1.3) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq \frac{1}{\sqrt{s}} D_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 . \tag{2.7}$$

In particular, if $\boldsymbol{x}$ is s-sparse, the recovery is exact.

(2) Noisy recovery. The solution $\boldsymbol{x}^\star$ to (1.4) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq \frac{1}{\sqrt{s}} D_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 + D_1 \varepsilon. \tag{2.8}$$

Theorem 2.3. Assume that A satisfies the RIP of order $(s+2s')$ and

$$\delta_{s+s'} + \sqrt{\frac{s}{s'}}\delta_{s+2s'} < 1. \tag{2.9}$$

Put

$$E_0 = 2\sqrt{\frac{s}{s'}}\,\frac{1-\delta_{s+s'}+\delta_{s+2s'}}{1-\delta_{s+s'}-\sqrt{\frac{s}{s'}}\delta_{s+2s'}}, \quad E_1 = \frac{2\left(1+\sqrt{\frac{s}{s'}}\right)\sqrt{1+\delta_{s+s'}}}{1-\delta_{s+s'}-\sqrt{\frac{s}{s'}}\delta_{s+2s'}}. \tag{2.10}$$

Then the following hold:

(1) Noiseless recovery. The solution $\boldsymbol{x}^\star$ to (1.3) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq \frac{2}{\sqrt{s}} E_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 . \tag{2.11}$$

(2) Noisy recovery. The solution $\boldsymbol{x}^\star$ to (1.4) obeys

$$\| \boldsymbol{x}^\star - \boldsymbol{x} \|_2 \leq \frac{2}{\sqrt{s}} E_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 + \frac{1}{\sqrt{s'}} E_1 \varepsilon. \tag{2.12}$$

The RIP requires a bounded condition number for all submatrices built by selecting s arbitrary columns and the spectral norm of such a matrix is not generally easy to compute. Hence, it is meaningful to discuss whether we can take the as small as possible order of the isometry constant δ, the as large as possible upper bounds of δ, and the as small as possible values for the constants of D_0, D_1, E_0, and E_1. Furthermore, when we apply the method of compressed sensing in statistics and others fields, it is meaningful to assess the solution $\boldsymbol{x}^\star$ for order $s + s'$. Furthermore, we shall show that our results are more flexible than those of E.J. Candès by taking some appropriate numbers for s'.

As Theorem 2.2 and Theorem 2.3 are start of our studies, we give the proofs.

Proof of Theorem 2.2. It suffices to show the case of noisy recovery. By Lemma 2.1 in (Candès, 2008), we have

$$|\langle A\boldsymbol{a},\ A\boldsymbol{b}\rangle| \leq \delta_{s+s'} \| \boldsymbol{a} \|_2 \| \boldsymbol{b} \|_2 \tag{2.13}$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R}^n$ supported on disjoint subsets $T, T' \subset \{1, 2, \cdots, n\}$ with $|T| \leq s$ and $|T'| \leq s'$. We define

$$\boldsymbol{h} = \boldsymbol{x}^\star - \boldsymbol{x}. \tag{2.14}$$

By the linearity of A and the triangle equality, we have

$$\| A\boldsymbol{h} \|_2 \leq 2\varepsilon. \tag{2.15}$$

For simplicity, we use the following symbol: For any $\boldsymbol{a} \in \boldsymbol{R}^n$ and $T \subset \{1, 2, \cdots, n\}$ we define

$$\boldsymbol{a}_T = \begin{pmatrix} a_1^T \\ \vdots \\ a_n^T \end{pmatrix},\ a_i^T = \begin{cases} a_i & ,\ i \in T \\ 0 & ,\ i \in T^c. \end{cases}$$

Let T_0 be the location of the s largest coefficients of $\boldsymbol{x}$. Then, $\boldsymbol{x}_{T_0} = \boldsymbol{x}_s$ and for any $k \in T_0$ and $j \geq s+1$, we have

$$\left|x_k^{T_0}\right| \geq |x_j| \,. \tag{2.16}$$

Let T_1 be the location of the s' largest coefficients of $\boldsymbol{h}_{T_0^c}$. Then, for any $k \in T_1$ and $i \in (T_0 \cup T_1)^c$,

$$\left|h_k^{T_1}\right| \geq |h_i| \,. \tag{2.17}$$

Repeating this method, we get vectors $\boldsymbol{h}_{T_j}(j \geq 2)$ satisfying the conditions

$$\left|h_k^{T_j}\right| \geq |h_i|, \tag{2.18}$$

for all $k \in T_j$ and $i \in (T_0 \cup T_1 \cup \cdots \cup T_j)^c$. Let $\{1, 2, \cdots, n\} = T_0 \cup T_1 \cup \cdots \cup T_{r-1} \cup T_r$, $|T_r| \leq s'$, and let $2 \leq j \leq r-1$. Clearly,

$$\begin{aligned} \| \boldsymbol{h}_{T_j} \|_2 &= \left(\sum_{k=1}^{s'} \left|h_k^{T_j}\right|^2 \right)^{\frac{1}{2}} \\ &\leq \sqrt{s'} \, \| \boldsymbol{h}_{T_j} \|_\infty \\ &\equiv \sqrt{s'} \max_{k \in T_j} \left|h_k^{T_j}\right|, \end{aligned} \tag{2.19}$$

and it follows from (2.17) and (2.19) that

$$\| \boldsymbol{h}_{T_{j-1}} \|_1 = \sum_{k=1}^{s'} \left|h_k^{T_{j-1}}\right| \geq s' \, \| \boldsymbol{h}_{T_j} \|_\infty, \tag{2.20}$$

which implies that

$$\| \boldsymbol{h}_{T_j} \|_2 \leq \frac{1}{\sqrt{s'}} \, \| \boldsymbol{h}_{T_{j-1}} \|_1, \tag{2.21}$$

for all $2 \leq j \leq r-1$. Furthermore, it follows that

$$\| \boldsymbol{h}_{T_{r-1}} \|_1 = \sum_{k=1}^{s'} |h_k^{T_{r-1}}| \geq s' \, \| \boldsymbol{h}_{T_r} \|_\infty$$

and

$$\begin{aligned} \| \boldsymbol{h}_{T_r} \|_2 &\leq \sqrt{|T_r|} \, \| \boldsymbol{h}_{T_r} \|_\infty \\ &\leq \sqrt{s'} \, \| \boldsymbol{h}_{T_r} \|_\infty \\ &\leq \frac{1}{\sqrt{s'}} \, \| \boldsymbol{h}_{T_{r-1}} \|_1 \,. \end{aligned} \tag{2.22}$$

By (2.21) and (2.22), we have for any $2 \leq j \leq r$

$$\| \boldsymbol{h}_{T_j} \|_2 \leq \frac{1}{\sqrt{s'}} \| \boldsymbol{h}_{T_{j-1}} \|_1 . \tag{2.23}$$

We next show

$$\| \boldsymbol{h}_{(T_0 \cup T_1)^c} \|_2 \leq \frac{1}{\sqrt{s'}} \| \boldsymbol{h}_{T_0^c} \|_1 . \tag{2.24}$$

Indeed, this follows from (2.23), as follows:

$$\begin{aligned}
\| \boldsymbol{h}_{(T_0 \cup T_1)^c} \|_2 &= \| \sum_{j \geq 2} h_{T_j} \|_2 \\
&\leq \sum_{j \geq 2} \| \boldsymbol{h}_{T_j} \|_2 \\
&\leq \frac{1}{\sqrt{s'}} \sum_{j \geq 1} \| \boldsymbol{h}_{T_j} \|_1 \\
&= \frac{1}{\sqrt{s'}} \| \sum_{j \geq 1} \boldsymbol{h}_{T_j} \|_1 \\
&= \frac{1}{\sqrt{s'}} \| \boldsymbol{h}_{T_0^c} \|_1 .
\end{aligned} \tag{2.25}$$

Since

$$\begin{aligned}
\| \boldsymbol{x} \|_1 &\geq \| \boldsymbol{x}^\star \|_1 \\
&= \| \boldsymbol{x} + \boldsymbol{h} \|_1 \\
&= \| \boldsymbol{x}_{T_0} + \boldsymbol{h}_{T_0} + \boldsymbol{x}_{T_0^c} + \boldsymbol{h}_{T_0^c} \|_1 \\
&= \| \boldsymbol{x}_{T_0} + \boldsymbol{h}_{T_0} \|_1 + \| \boldsymbol{x}_{T_0^c} + \boldsymbol{h}_{T_0^c} \|_1 \\
&\geq \| \boldsymbol{x}_{T_0} \|_1 - \| \boldsymbol{h}_{T_0} \|_1 + \| \boldsymbol{h}_{T_0^c} \|_1 - \| \boldsymbol{x}_{T_0^c} \|_1,
\end{aligned}$$

it follows that

$$\begin{aligned}
\| \boldsymbol{h}_{T_0^c} \|_1 &\leq \| \boldsymbol{x} \|_1 - \| \boldsymbol{x}_{T_0} \|_1 + \| \boldsymbol{x}_{T_0^c} \|_1 + \| \boldsymbol{h}_{T_0} \|_1 \\
&= \| \boldsymbol{x}_{T_0^c} \|_1 + \| \boldsymbol{x}_{T_0^c} \|_1 + \| \boldsymbol{h}_{T_0} \|_1 \\
&= 2 \| \boldsymbol{x}_{T_0^c} \|_1 + \| \boldsymbol{h}_{T_0} \|_1 \\
&= 2 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 + \| \boldsymbol{h}_{T_0} \|_1,
\end{aligned} \tag{2.26}$$

which implies by (2.19), (2.25), and the Schwartz inequality that

$$\begin{aligned}
\|\boldsymbol{h}_{(T_0\cup T_1)^c}\|_2 &\le \frac{1}{\sqrt{s'}} \parallel \boldsymbol{h}_{T_0^c} \parallel_1 \\
&\le \frac{1}{\sqrt{s'}} (\parallel \boldsymbol{h}_{T_0} \parallel_1 + 2 \parallel \boldsymbol{x} - \boldsymbol{x}_s \parallel_1) \\
&\le \sqrt{\frac{s}{s'}} \parallel \boldsymbol{h}_{T_0} \parallel_2 + \frac{2}{\sqrt{s'}} \parallel \boldsymbol{x} - \boldsymbol{x}_s \parallel_1 \\
&\le \sqrt{\frac{s}{s'}} \|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \frac{2}{\sqrt{s'}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1. \qquad (2.27)
\end{aligned}$$

Suppose that $s' \le s$. Then we have $\delta_{2s'} \le \delta_{s+s'}$ by the definition of restricted isometry constants. Since $\boldsymbol{h}_{T_0}$ and $\boldsymbol{h}_{T_1}$ are orthogonal, it follows from (2.13) that for any $j \ge 2$

$$\begin{aligned}
\left|\langle A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}_{T_j}\rangle\right| &\le \left|\langle A\boldsymbol{h}_{T_0}, A\boldsymbol{h}_{T_j}\rangle\right| + \left|\langle A\boldsymbol{h}_{T_1}, A\boldsymbol{h}_{T_j}\rangle\right| \\
&\le \delta_{s+s'} \parallel \boldsymbol{h}_{T_0} \parallel_2 \parallel \boldsymbol{h}_{T_j} \parallel_2 + \delta_{2s'} \parallel \boldsymbol{h}_{T_1} \parallel_2 \parallel \boldsymbol{h}_{T_j} \parallel_2 \\
&\le \delta_{s+s'} \parallel \boldsymbol{h}_{T_j} \parallel_2 (\parallel \boldsymbol{h}_{T_0} \parallel_2 + \parallel \boldsymbol{h}_{T_1} \parallel_2) \\
&\le \delta_{s+s'} \parallel \boldsymbol{h}_{T_j} \parallel_2 \sqrt{2} \left(\parallel \boldsymbol{h}_{T_0} \parallel_2^2 + \parallel \boldsymbol{h}_{T_1} \parallel_2^2\right)^{\frac{1}{2}} \\
&= \sqrt{2}\delta_{s+s'} \parallel \boldsymbol{h}_{T_j} \parallel_2 \parallel \boldsymbol{h}_{T_0} + \boldsymbol{h}_{T_1} \parallel_2 \\
&= \sqrt{2}\delta_{s+s'} \parallel \boldsymbol{h}_{T_j} \parallel_2 \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2,
\end{aligned}$$

which implies by (2.15) and (2.16) that

$$\begin{aligned}
\parallel A\boldsymbol{h}_{T_0\cup T_1} \parallel_2^2 &= \langle A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h} - \sum_{j\ge 2} A\boldsymbol{h}_{T_j}\rangle \\
&= \langle A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}\rangle - \langle A\boldsymbol{h}_{T_0\cup T_1}, \sum_{j\ge 2} A\boldsymbol{h}_{T_j}\rangle \\
&\le \|A\boldsymbol{h}_{T_0\cup T_1}\|_2 \|A\boldsymbol{h}\|_2 + \sum_{j\ge 2} \left|\langle A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}_{T_j}\rangle\right| \qquad (2.28) \\
&\le \sqrt{1+\delta_{s+s'}} \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 2\varepsilon + \sqrt{2}\delta_{s+s'} \left(\sum_{j\ge 2} \parallel \boldsymbol{h}_{T_j} \parallel_2\right) \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \\
&= \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \left(2\varepsilon\sqrt{1+\delta_{s+s'}} + \sqrt{2}\delta_{s+s'} \sum_{j\ge 2} \parallel \boldsymbol{h}_{T_j} \parallel_2\right) \\
&\le \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \left(2\varepsilon\sqrt{1+\delta_{s+s'}} + \sqrt{2}\delta_{s+s'} \frac{1}{\sqrt{s'}} \sum_{j\ge 1} \parallel \boldsymbol{h}_{T_j} \parallel_1\right) \\
&= \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \left(2\varepsilon\sqrt{1+\delta_{s+s'}} + \sqrt{2}\delta_{s+s'} \frac{1}{\sqrt{s'}} \parallel \boldsymbol{h}_{T_0^c} \parallel_1\right).
\end{aligned}$$

Hence, we have

$$\begin{aligned}(1-\delta_{s+s'}) \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2^2 &\le \parallel A\boldsymbol{h}_{T_0\cup T_1} \parallel_2^2\\ &\le \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \left(2\varepsilon\sqrt{1+\delta_{s+s'}}+\sqrt{2}\delta_{s+s'}\frac{1}{\sqrt{s'}} \parallel \boldsymbol{h}_{T_0^c} \parallel_1\right),\end{aligned}$$

which implies that

$$\parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \le \frac{2\sqrt{1+\delta_{s+s'}}}{1-\delta_{s+s'}}\varepsilon+\frac{\sqrt{2}\delta_{s+s'}}{1-\delta_{s+s'}}\frac{1}{\sqrt{s'}} \parallel \boldsymbol{h}_{T_0^c} \parallel_1 . \tag{2.29}$$

Furthermore, it follows from (2.26) that

$$\begin{aligned}\parallel \boldsymbol{h}_{T_0^c} \parallel_1 &\le \parallel \boldsymbol{h}_{T_0} \parallel_1 +2 \parallel \boldsymbol{x}-\boldsymbol{x}_s \parallel_1\\ &\le \sqrt{s} \parallel \boldsymbol{h}_{T_0} \parallel_2 +2 \parallel \boldsymbol{x}-\boldsymbol{x}_s \parallel_1\\ &\le \sqrt{s} \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 +2 \parallel \boldsymbol{x}-\boldsymbol{x}_s \parallel_1,\end{aligned} \tag{2.30}$$

so that from (2.29),

$$\left(1-\frac{\sqrt{2}\delta_{s+s'}}{1-\delta_{s+s'}}\sqrt{\frac{s}{s'}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \le \frac{2\sqrt{1+\delta_{s+s'}}}{1-\delta_{s+s'}}\varepsilon+\frac{1}{\sqrt{s'}}\ \frac{2\sqrt{2}\delta_{s+s'}}{1-\delta_{s+s'}}\|\boldsymbol{x}-\boldsymbol{x}_s\|_1.$$

Hence, by the assumption $\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}<1$, we have

$$\parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 \le \frac{2\sqrt{1+\delta_{s+s'}}}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}}\varepsilon+\frac{1}{\sqrt{s'}}\ \frac{2\sqrt{2}\delta_{s+s'}}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}}\|\boldsymbol{x}-\boldsymbol{x}_s\|_1. \tag{2.31}$$

Thus, we have by (2.27) and (2.31) that

$$\begin{aligned}\parallel \boldsymbol{x}-\boldsymbol{x}^\star \parallel_2 &= \parallel \boldsymbol{h} \parallel_2\\ &\le \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 + \parallel \boldsymbol{h}_{(T_0\cup T_1)^c} \parallel_2\\ &\le \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 +\sqrt{\frac{s}{s'}} \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 +\frac{2}{\sqrt{s'}} \parallel \boldsymbol{x}-\boldsymbol{x}_s \parallel_1\\ &= \left(1+\sqrt{\frac{s}{s'}}\right) \parallel \boldsymbol{h}_{T_0\cup T_1} \parallel_2 +\frac{2}{\sqrt{s'}} \parallel \boldsymbol{x}-\boldsymbol{x}_s \parallel_1\\ &\le \left(1+\sqrt{\frac{s}{s'}}\right)\frac{2\sqrt{1+\delta_{s+s'}}}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}}\varepsilon\\ &\quad +\frac{2}{\sqrt{s'}}\ \frac{\left(\sqrt{2}-1\right)\delta_{s+s'}+1}{1-\left(1+\sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}}\|\boldsymbol{x}-\boldsymbol{x}_s\|_1.\end{aligned} \tag{2.32}$$

Hence, we have

$$\| \boldsymbol{x} - \boldsymbol{x}^\star \|_2 \leq \frac{1}{\sqrt{s}} D_0 \| \boldsymbol{x} - \boldsymbol{x}_s \|_1 + D_1 \varepsilon.$$

This completes the proof of the $s' \leq s$ case.

Suppose that $s' \geq s$ and $\delta_{s+s'} + \sqrt{\frac{2s}{s'}} < 1$. We can prove in the same way as the $s' \leq s$ case just proved that

$$\|\boldsymbol{h}_{T_0 \cup T_1}\|_2 \leq \frac{2\sqrt{1+\delta_{s+s'}}}{1 - \delta_{s+s'} - \sqrt{\frac{2s}{s'}}\delta_{2s'}} \varepsilon + \frac{1}{\sqrt{s'}} \frac{2\sqrt{2}\delta_{2s'}}{1 - \delta_{s+s'} - \sqrt{\frac{2s}{s'}}\delta_{2s'}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1,$$

which implies by (2.27) that

$$\begin{aligned}
\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 &\leq \left(1 + \sqrt{\frac{s}{s'}}\right) \|\boldsymbol{h}_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{s'}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 \\
&\leq \frac{1}{\sqrt{s'}} \frac{2\left(\sqrt{2}\delta_{2s'} - \delta_{s+s'}\right) + 2}{1 - \delta_{s+s'} - \sqrt{\frac{2s}{s'}}\delta_{2s'}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 \\
&\quad + \left(1 + \sqrt{\frac{s}{s'}}\right) \frac{2\sqrt{1+\delta_{s+s'}}}{1 - \left(1 + \sqrt{\frac{2s}{s'}}\right)\delta_{s+s'}} \varepsilon \\
&= \frac{1}{\sqrt{s}} D_0 \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + D_1 \varepsilon.
\end{aligned}$$

This completes the proof of Theorem 2.2.

Proof of Theorem 2.3. Using (2.19) on (2.32) directly, we can obtain the inequality

$$\begin{aligned}
\| A\boldsymbol{h}_{T_0 \cup T_1} \|_2^2 &\leq \| A\boldsymbol{h}_{T_0 \cup T_1} \|_2 \| A\boldsymbol{h} \|_2 + \sum_{j \geq 2} \left| \langle A\boldsymbol{h}_{T_0 \cup T_1}, A\boldsymbol{h}_{T_j} \rangle \right| \\
&\leq \sqrt{1+\delta_{s+s'}} \| \boldsymbol{h}_{T_0 \cup T_1} \|_2 \, 2\varepsilon + \sum_{j \geq 2} \delta_{s+2s'} \| \boldsymbol{h}_{T_0 \cup T_1} \|_2 \| \boldsymbol{h}_{T_j} \|_2 \\
&= \| \boldsymbol{h}_{T_0 \cup T_1} \|_2 \left(2\varepsilon\sqrt{1+\delta_{s+s'}} + \delta_{s+2s'} \sum_{j \geq 2} \| \boldsymbol{h}_{T_j} \|_2 \right),
\end{aligned}$$

and by (2.21), (2.30), and the assumption $\delta_{s+s'}\sqrt{\frac{2s}{s'}} < 1$, we have

$$\|\boldsymbol{h}_{T_0 \cup T_1}\|_2 \leq \frac{1}{\sqrt{s'}} \left(\frac{2\delta_{s+2s'}}{1 - \delta_{s+s'} - \sqrt{\frac{s}{s'}}\delta_{s+2s'}} \right) \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + \frac{2\sqrt{1+\delta_{s+s'}}}{1 - \delta_{s+s'} - \sqrt{\frac{s}{s'}}} \varepsilon. \quad (2.33)$$

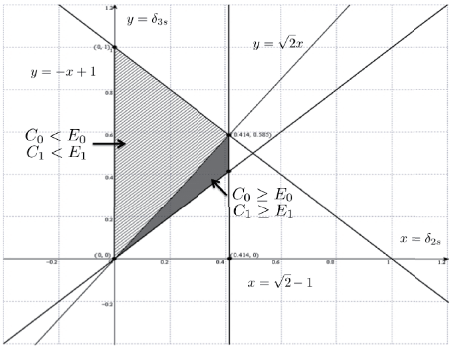

Figure 2.1: (1) C_0, C_1 vs E_0, E_1

Hence, we can show similarly to (2.32) that

$$\begin{aligned}\|\boldsymbol{x} - \boldsymbol{x}_s\|_2 &\leq \frac{1}{\sqrt{s'}} \frac{2\left(\delta_{s+2s'} - \delta_{s+s'} + 1\right)}{1 - \delta_{s+s'} - \sqrt{\frac{s}{s'}}\delta_{s+2s'}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + \left(1 + \sqrt{\frac{s}{s'}}\right) \frac{2\sqrt{1 + \delta_{s+s'}}}{1 - \delta_{s+s'} - \sqrt{\frac{s}{s'}}}\varepsilon \\ &= \frac{1}{\sqrt{s}} E_0 \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + E_1 \varepsilon. \end{aligned} \tag{2.34}$$

This completes the proof. Furthermore, we shall compare the constants D_0, E_0 and D_1, E_1 to the constants C_0 and C_1, respectively.

(1) In Theorem 2.2 and Theorem 2.3, taking $s' = s$, the assumption (2.4) coincides with the assumption

$$\delta_{2s} < \sqrt{2} - 1, \tag{2.35}$$

and the assumption (2.9) coincides with the assumption

$$\delta_{2s} + \delta_{3s} < 1 \tag{2.36}$$

of E.J. Candès (Candès, 2006). Investigating each constant, clearly, $C_0 = D_0,\ C_1 = D_1$, and

$$E_0 = 2\left(\frac{1 - \delta_{2s} + \delta_{3s}}{1 - \delta_{2s} - \delta_{3s}}\right), \quad E_1 = 2\left(\frac{2\sqrt{1 + \delta_{2s}}}{1 - \delta_{2s} - \delta_{3s}}\right).$$

Thus, $C_0 \geq E_0$ and $C_1 \geq E_1$ hold for $(\delta_{2s}, \delta_{3s})$ in set filled and the opposite inequalities $C_0 > E_0$ and $C_1 > E_1$ hold for $(\delta_{2s}, \delta_{3s})$ in set shaded in Fig. 2.1.

(2) Let $s = 2k,\ k \in \boldsymbol{N}$. Taking $s' = k$, the assumption (2.4) becomes

$$\delta_{\frac{3}{2}s} < \frac{1}{3} \tag{2.37}$$

and the assumption (2.11) becomes

$$\delta_{\frac{3}{2}s} + \sqrt{2}\delta_{2s} < 1. \tag{2.38}$$

Since $\delta_{\frac{3}{2}s} + \sqrt{2}\delta_{2s} < \left(1+\sqrt{2}\right)\delta_{2s}$, the condition (2.38) is stricter than the condition $\delta_{2s} < \sqrt{2}-1$ in Theorem 2.1. Investigating each constant, clearly,

$$D_0 = \frac{2}{\sqrt{k}}\frac{\left(\sqrt{2}-1\right)\delta_{3k}+1}{1-3\delta_{3k}}, \quad D_1 = \left(1+\sqrt{2}\right)\frac{2\sqrt{1+\delta_{3k}}}{1-3\delta_{3k}},$$
$$E_0 = 2\sqrt{2}\frac{1-\delta_{3k}+\delta_{4k}}{1-\delta_{3k}-\sqrt{2}\delta_{4k}}, \quad E_1 = \frac{2\left(\sqrt{2}+1\right)\sqrt{1+\delta_{3k}}}{1-\delta_{3k}-\sqrt{2}\delta_{4k}}.$$

Thus, $C_0 < D_0$ and $C_1 < D_1$ hold for $(\delta_{3k}, \delta_{4k})$ in set shaded in Fig. 2.2, and $E_0 < C_0 + 3.14$ and $E_1 < C_1$ hold for $(\delta_{3k}, \delta_{4k})$ in the intersection of the two sets shaded in Figs. 2.3 and 2.4. This means that the result of Theorem 2.3 improves that of Theorem 2.1 with respect to not only the assessment of the upper bounds of the isometry constants but also the values of constants C_0, C_1, E_0, and E_1.

(3) Let $s' = 2s$. Then, Theorem 2.2 holds under the assumption

$$\delta_{4s} < \frac{1}{2} \tag{2.39}$$

and Theorem 2.3 holds if

$$\delta_{3s} + \frac{1}{\sqrt{2}}\delta_{5s} < 1. \tag{2.40}$$

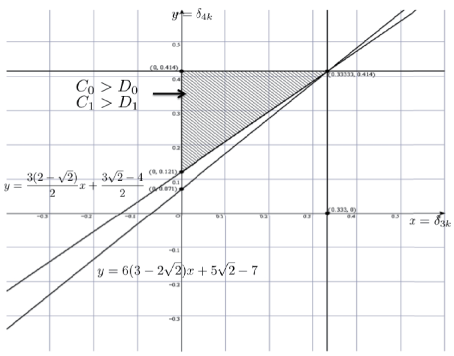

Figure 2.2: (2) C_0, C_1 vs D_0, D_1

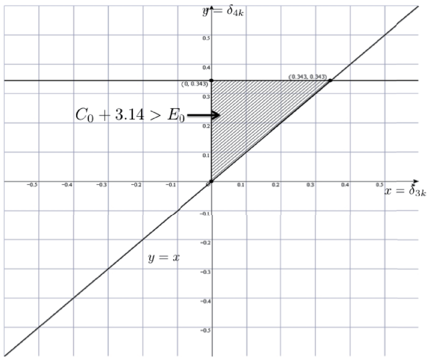

Figure 2.3: (2) C_0 vs E_0

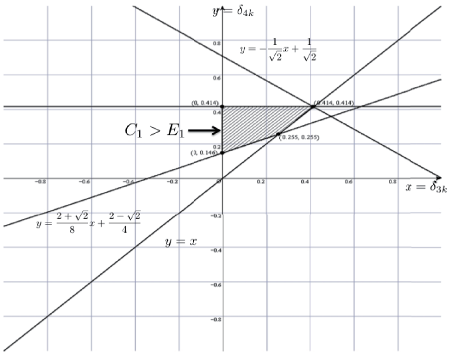

Figure 2.4: (2) C_1 vs E_1

Clearly,

$$D_0 = \frac{\sqrt{2}\left(\sqrt{2}\delta_{4s} - \delta_{3s} + 1\right)}{1 - \delta_{3s} - \delta_{4s}}, \quad D_1 = \left(1 + \frac{1}{\sqrt{2}}\right)\frac{2\sqrt{1+\delta_{3s}}}{1 - \delta_{3s} - \delta_{4s}},$$

$$E_0 = \sqrt{2}\frac{1 - \delta_{3s} + \delta_{5s}}{1 - \delta_{3s} - \frac{1}{\sqrt{2}}\delta_{5s}}, \quad E_1 = \frac{2\left(1 + \frac{1}{\sqrt{2}}\right)\sqrt{1+\delta_{3s}}}{1 - \delta_{3s} - \frac{1}{\sqrt{2}}\delta_{5s}}.$$

Thus, $C_0 > D_0 + 0.05$ and $C_1 - 2.19 > D_1$ hold for $(\delta_{2s}, \delta_{4s})$ in the intersection of the two sets shaded in Fig. 2.5 and 2.6. Here, we omit assessing the relationship between (C_0, C_1) and (E_0, E_1), respectively.

(4) Let $s = 2k+1,\ k \in \boldsymbol{N}$. Taking $s' = k+1$, Theorem 2.2 holds under the assumption

$$\delta_{\frac{3s+1}{2}} < \frac{1}{3} \tag{2.41}$$

and Theorem 2.3 holds under the assumption

$$\delta_{\frac{3s+1}{2}} + \sqrt{2}\delta_{2s+1} < 1. \tag{2.42}$$

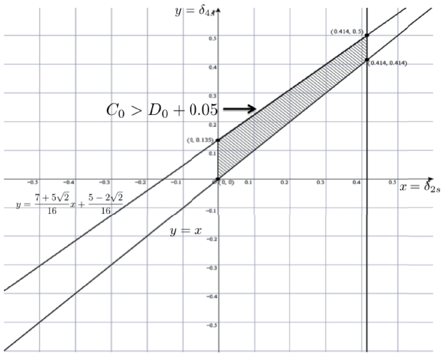

Figure 2.5: (3) C_0 vs D_0

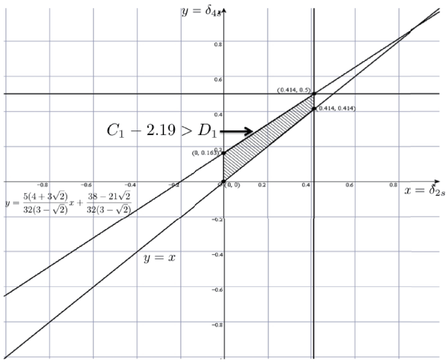

Figure 2.6: (3) C_1 vs D_1

The constants D_0, D_1, E_0, and E_1 are as follows:

$$D_0 = \sqrt{\frac{2k+1}{k+1}}\frac{2(\sqrt{2}-1)\delta_{3k+2}+2}{1-\left(1+\sqrt{\frac{2k+1}{k+1}}\right)\delta_{3k+2}}, \quad D_1 = \left(1+\sqrt{\frac{2k+1}{k+1}}\right)\frac{2\sqrt{1+\delta_{3k+2}}}{1-\left(1+\sqrt{\frac{2k+1}{k+1}}\right)\delta_{3k+2}},$$

$$E_0 = 2\sqrt{\frac{2k+1}{k+1}}\frac{1-\delta_{3k+2}+\delta_{4k+3}}{1-\delta_{3k+2}-\sqrt{\frac{2k+1}{k+1}}\delta_{4k+3}}, \quad E_1 = \frac{2\left(1+\sqrt{\frac{2k+1}{k+1}}\right)\sqrt{1+\delta_{3k+2}}}{1-\delta_{3k+2}-\sqrt{\frac{2k+1}{k+1}}\delta_{4k+3}}.$$

Here we do not discuss to assess the relationship between (C_0, D_0, E_0) and (C_1, D_1, E_1), respectively.

(5) Let $s = 2k,\ k \in \boldsymbol{N}$. Taking $s' = 3k$, Theorem 2.2 holds if

$$\delta_{3s} < \frac{1}{1+\sqrt{\frac{2}{3}}} \tag{2.43}$$

and Theorem 2.3 holds if

$$\delta_{\frac{5}{2}s} + \sqrt{\frac{2}{3}}\delta_{4s} < 1. \tag{2.44}$$

The constants D_0, D_1, E_0, and E_1 are as follows:

$$D_0 = \frac{2\sqrt{2}}{3}\frac{\sqrt{2}\delta_{6k} - \delta_{5k} + 1}{1 - \delta_{5k} - \frac{2}{\sqrt{3}}\delta_{6k}}, \quad D_1 = \left(1 + \sqrt{\frac{2}{3}}\right)\frac{2\sqrt{1+\delta_{5k}}}{1 - \delta_{5k} - \frac{2}{\sqrt{3}}\delta_{6k}},$$

$$E_0 = \frac{2\sqrt{2}}{\sqrt{3}}\frac{1 - \delta_{5k} + \delta_{8k}}{1 - \delta_{5k} - \sqrt{\frac{2}{3}}\delta_{8k}}, \quad E_1 = \frac{2\left(1 + \sqrt{\frac{2}{3}}\right)\sqrt{1+\delta_{5k}}}{1 - \delta_{5k} - \sqrt{\frac{2}{3}}\delta_{8k}}.$$

Here we do not discuss to assess the relationship between (C_0, D_0, E_0) and (C_1, D_1, E_1), respectively.

2.2 The others results

In this section, we introduce the important results (Theorem 2.4, 2.5, 2.6, 2.7 and 2.8) used in Chapter 3-6.

Theorem 2.4 (Cai, Wang and Xu, 2010b) Suppose that A obeys the RIP of order s and $\delta_s < 0.307$. Then

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \le \frac{1}{0.307 - \delta_s}\frac{1}{\sqrt{s}}\|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + \frac{1}{0.307 - \delta_s}\varepsilon. \tag{2.45}$$

Theorem 2.5. (Mo and Li, 2011) Assume that A obeys the RIP of order $2s$ and $\delta_{2s} < 0.4931$. Then

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \le C_0\frac{1}{\sqrt{s}}\|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + C_1\varepsilon,$$

where

$$C_0 = \frac{4}{1-\theta_s}\sqrt{\frac{2(2-\delta_{2s})}{(1-\delta_{2s})(32-25\delta_{2s})}},$$
$$C_1 = \frac{2}{\sqrt{1-\delta_{2s}}}\left(1 + \frac{4}{1-\theta_s}\sqrt{\frac{2-\delta_{2s}}{(1-\delta_{2s})(32-25\delta_{2s})}}\right). \tag{2.46}$$

Here,

$$\theta_s = \sqrt{\frac{(1+\delta_{2s})^2}{8(1-2\delta_{2s})}}. \tag{2.47}$$

Theorem 2.6. (Cai and Zhang, 2013a) Suppse that A obeys the RIP of order s and $\delta_s < \frac{1}{3}$ with $s \geq 2$. Then

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq C_0 \frac{1}{\sqrt{s}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + C_1 \varepsilon,$$

where

$$\begin{aligned} C_0 &= \frac{2\sqrt{2}(2\delta_s + \sqrt{(1-3\delta_s)\delta_s} + 2(1-3\delta_s))}{1-3\delta_s}, \\ C_1 &= \frac{\sqrt{2(1+\delta_s)}}{1-3\delta_s}. \end{aligned} \tag{2.48}$$

Theorem 2.7. (Cai and Zhang, 2013b) Let $t \geq \frac{4}{3}$. Assume that A obeys the RIP of order ts and $\delta_{ts} < \sqrt{\frac{t-1}{t}}$. Then

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq C_0 \frac{1}{\sqrt{s}} \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + C_1 \varepsilon,$$

where

$$\begin{aligned} C_0 &= 2\left(\frac{\sqrt{2}\delta_{ts} + \sqrt{t\left(\sqrt{\frac{t-1}{t}} - \delta_{ts}\right)\delta_{ts}}}{t\left(\sqrt{\frac{t-1}{t}} - \delta_{ts}\right)} + 1 \right), \\ C_1 &= \frac{\sqrt{2(1+\delta_{ts})}}{1 - \sqrt{\frac{t-1}{t}}\delta_{ts}}. \end{aligned} \tag{2.49}$$

Theorem 2.8. (Candès and Plan, 2010) Let T_0 be a subset of $\{1, 2, \cdots, n\}$ with $|T_0| = s$, $\delta > 0$ and $r \in \boldsymbol{N}$. Suppose that a random matrix A satisfies the isotropy property and $m \geq C_\delta \beta \mu \cdot \max\left(s \log(s\mu), r \log n \log^2 r \log(r\mu \log n)\right)$, where μ is the coherence parameter. Then A obeys the weak RIP with respect to T_0 of order r with probability at least $(1 - 5e^{-\beta})$.

Chapter 3

Sufficient conditions for CS-recovery

3.1 Introduction

This chapter introduces the theory of compressed sensing(CS) in rescaling method. Formally, one considers the following model:

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{z}, \tag{3.1}$$

where A is a $m \times n$ matrix($m < n$) and $\boldsymbol{z}$ is an unknown noise term.

Our goal is to reconstruct an unknown signal $\boldsymbol{x}$ based on A and $\boldsymbol{y}$ are given. Then we consider reconstructing $\boldsymbol{x}$ as the solution $\boldsymbol{x}^\star$ to the optimization problem

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1, \quad \text{subject to } \ \|\boldsymbol{y} - A\boldsymbol{x}\|_2 \leq \varepsilon, \tag{3.2}$$

where ε is an upper bound on the the size of the noisy contribution.
In fact, a crucial issue is to research good conditions under which the inequality

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1 \varepsilon, \tag{3.3}$$

for some suitable constants C_0 and C_1, where T_0 is any location of $\{1, 2, \cdots, n\}$ with number $|T_0|$ of elements of T_0 and $\boldsymbol{x}_{T_0}$ is the restriction of $\boldsymbol{x}$ to indices in T_0. One of the most generally known condition for CS theory is the restricted isometry property(RIP) introduced by E.J. Candès and T. Tao (Candès and Tao, 2005). When we discuss our proposed results, it is an important notion. The RIP needs that the subsets of columns of

A for all locations in $\{1, 2, \cdots, n\}$ behave nearly orthonormal system. In detail, a matrix A satisfies the RIP of order s if there exists a constant δ with $0 < \delta < 1$ such that

$$(1-\delta)\|\boldsymbol{a}\|_2^2 \leq \|A\boldsymbol{a}\|_2^2 \leq (1+\delta)\|\boldsymbol{a}\|_2^2 \tag{3.4}$$

for all s-sparse vectors $\boldsymbol{a}$. A vector is said to be an s-sparse vector if it has at most s nonzero entries. The minimum δ satisfying the above restrictions is said to be the restricted isometry constant and is denoted by δ_s.
As stated in Introduction and Background, many researchers has been shown that l_1 optimization can recover an unknown signal in noiseless case and noisy case under various sufficient conditions on δ_s or δ_{2s} when A obeys the RIP

In this chapter, we define the k-restricted norm constant $r_k(A)$ (simply, r_k) by

$$r_k(A) \equiv \max\{\|A_T\|;\ T \subset \{1, 2, \cdots, n\},\ |T| = k\},$$

where A_T is the $m \times |T|$ matrix composed of these columns for T and $\|\cdot\|$ is operator norm and research good conditions under which the inequality (3.3) holds by considering the following equality (3.5) instead of (3.1):

$$\tilde{\boldsymbol{y}} = \tilde{A}\boldsymbol{x} + \tilde{\boldsymbol{z}} \tag{3.5}$$

where $\tilde{\boldsymbol{y}} \equiv \frac{\boldsymbol{y}}{r_k}$, $\tilde{A} \equiv \frac{A}{r_k}$ and $\tilde{\boldsymbol{z}} \equiv \frac{\boldsymbol{z}}{r_k}$. Since $\{\boldsymbol{a} \in \boldsymbol{R}^n;\ \|\boldsymbol{y} - A\boldsymbol{a}\|_2 \leq \varepsilon\} = \{\boldsymbol{a} \in \boldsymbol{R}^n; \|\tilde{\boldsymbol{y}} - \tilde{A}\boldsymbol{a}\|_2 \leq \frac{\varepsilon}{r_k}\}$, it follows that the solution $\boldsymbol{x}^\star$ to the optimization problem (3.2) is the same as that to (3.5). When A is a matrix with uniformly bounded entries, that is, $|a_{ij}| \leq 1$ for each $1 \leq i \leq m$ and $1 \leq j \leq n$, $r_k(A) \leq \sqrt{ms}$. In particular, if A is the DFT matrix with entries:

$$a_{jt} = e^{-i2\pi jt/n} \quad , 0 \leq j, t \leq n-1,$$

then $r_k(A) \leq \min\left(\sqrt{mk},\ \sqrt{n}\right)$. We also need the notion of restricted invertibility of A. The matrix A is called k-restrictly invertible if $(A_T^\star A_T)^{\frac{1}{2}}$ is invertible for any subset T of $\{1, 2, \cdots, n\}$ with $|T| = k$.(if and only if A_T is an injection for any subset T of $\{1, 2, \cdots n\}$ with $|T| = k$ if and only if the column vectors $\{\boldsymbol{a}_i;\ i \in T\}$ of A is independent for any T of $\{1, 2, \cdots, n\}$ with $|T| = k$.) It is shown in Lemma 3.1 that A is k-restrictly invertible if and only if $\tilde{A} \equiv \frac{A}{r_k}$ obey the RIP of order k and $r_k(\tilde{A}) \leq 1$. Hence we can make use of results for RIP and in particular, the E.J. Candès idea, the T. Cai et al. idea and the Q. Mo and S. Li idea. The first main propose of this chapter is to show that if A is s-restrictly invertible and $\tilde{\delta}_s < \frac{2}{2+\sqrt{5}} \approx 0.472$, where $\tilde{\delta}_s$ is the restricted isometry constant for $\tilde{A} \equiv \frac{A}{r_s}$, and if A is $2s$-restrictly invertible and the restricted isometry constant $\tilde{\delta}_{2s} < 0.661$, then inequality (1.3) holds. The second is to obtain the better sufficient conditions of $\tilde{\delta}_s < 0.5$ and $\tilde{\delta}_{2s} < 0.828$ by using the results of (Cai and Zhang, 2013a) and (Cai and Zhang, 2013b).

Our analysis is very simple and elementary. We introduce the proposed results using the E.J. Candès idea, the T. Cai et al. idea, the Q. Mo and S. Li idea and the T. Cai and A. Zhang idea. We regard Theorem 3.1, Theorem 3.2, Theorem 3.3 and Theorem 3.5 as the main results in this chapter. Otherwise, in Section 2, we prepare some notions and lemmas to prove main theorems. In Section 3, we introduce new bounds of δ_s and $\delta_k (k > s)$.

3.2 Preliminaries and some lemmas

In this section, we prepare some lemmas needed for the proofs of Theorem 3.1 and Theorem 3.3.

Lemma 3.1. *Let k be a natural number with $k < n$. Then A is k-restrictly invertible if and only if $\tilde{A} \equiv \frac{A}{r_k}$ obeys the RIP of order k. If this is true, then*

$$\tilde{\delta}_k = 1 - \frac{1}{(r_k w_k)^2}$$

and

$$\left(1 - \tilde{\delta}_k\right) \|\boldsymbol{a}\|_2^2 \leq \|\tilde{A}\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{a}\|_2^2,$$

equivalently

$$r_k^2 \left(1 - \tilde{\delta}_k\right) \|\boldsymbol{a}\|_2^2 \leq \|A\boldsymbol{a}\|_2^2 \leq r_k^2 \|\boldsymbol{a}\|_2^2$$

for all k-sparse vector $\boldsymbol{a}$ in $\boldsymbol{R}^n$, where $\tilde{\delta}_k$ is the restricted isometry constant for $\tilde{A}$ and

$$w_k \equiv \min\left\{\| \left(A_T^{\star} A_T\right)^{-\frac{1}{2}} \|;\ T \subset \{1, 2, \cdots, n\} \text{ with } |T| = k\right\}.$$

Proof. Suppose that A is k-restrictly invertible. Then we have

$$\frac{1}{w_k}\|\boldsymbol{a}\|_2 \leq \|A\boldsymbol{a}\|_2 \leq r_k\|\boldsymbol{a}\|_2$$

for all k-sparse vector $\boldsymbol{a}$ in $\boldsymbol{R}^n$. Hence, $r_k w_k \geq 1$ and

$$\left(1 - \left(1 - \frac{1}{(r_k w_k)^2}\right)\right)\|\boldsymbol{a}\|_2^2 \leq \|\tilde{A}\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{a}\|_2^2$$

for all k-sparse vector $\boldsymbol{a}$, which implies that $\tilde{A}$ obeys the RIP of order k and

$$\tilde{\delta}_k \leq 1 - \frac{1}{(r_k w_k)^2}. \tag{3.6}$$

Conversely, suppose that $\tilde{A}$ obeys the RIP of order k. Then, since

$$\left(1 - \tilde{\delta}_k\right)\|\boldsymbol{a}\|_2^2 \leq \|\tilde{A}\boldsymbol{a}\|_2^2 \leq \|\boldsymbol{a}\|_2^2$$

for all k-sparse vector $\boldsymbol{a}$, it follows that A is k-restrictly invertible and $\tilde{\delta} \geq 1 - \frac{1}{(r_k w_k)^2}$, which implies by (3.6) that $\tilde{\delta}_k = 1 - \frac{1}{(r_k w_k)^2}$. This completes the proof.

Lemma 3.2. Take any $t \geq 1$ and positive integers s', s'' such that ts'' is an integer. Suppose that A is $(s' + s'')$-restrictly invertible. Then,

$$| < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > | \leq \frac{1}{2\sqrt{t}}\tilde{\delta}_{s'+s''}\|\boldsymbol{a}'\|_2\|\boldsymbol{a}''\|_2 \tag{3.7}$$

for any vectors $\boldsymbol{a}', \boldsymbol{a}'' \in \boldsymbol{R}^n$ with disjoint supports and sparsity ts'' and s', respectively.

Proof. We make use of the square root lifting inequality (Cai, Wang and Xu, 2010a):

$$\theta_{s',ts''} \leq \frac{1}{\sqrt{t}}\theta_{s',s''}, \quad t \geq 1, \tag{3.8}$$

where $\theta_{k,k'}$ is the k, k'-restricted orthogonality constant. The k, k'-restricted orthogonality constant is the smallest number that satisfies

$$| < A\boldsymbol{c}, A\boldsymbol{c}' > | \leq \theta_{k,k'}\|\boldsymbol{c}\|_2\|\boldsymbol{c}'\|_2$$

for all k-sparse vector $\boldsymbol{c}$ and k'-sparse vector $\boldsymbol{c}'$ with disjoint supports. Take arbitrary s'-sparse vector $\boldsymbol{a}'$ and s''-sparse vector $\boldsymbol{a}''$ with disjoint supports. Then we show

$$| < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > | \leq \frac{\tilde{\delta}_{s'+s''}}{2} \|\boldsymbol{a}'\|_2 \|\boldsymbol{a}''\|_2. \tag{3.9}$$

Indeed, it is sufficient to show this inequality without loss of generality in case $\|\boldsymbol{a}'\|_2 = \|\boldsymbol{a}''\|_2 = 1$. Since $\tilde{A}$ obeys the RIP of order $(s' + s'')$ and $r_{s'+s''}(\tilde{A}) \leq 1$, it follows that

$$\begin{aligned}
4 < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > &= \|\tilde{A}(\boldsymbol{a}' + \boldsymbol{a}'')\|_2^2 - \|\tilde{A}(\boldsymbol{a}' - \boldsymbol{a}'')\|_2^2 \\
&\leq \|\boldsymbol{a}' + \boldsymbol{a}''\|_2^2 - (1 - \tilde{\delta}_s)\|\boldsymbol{a}' - \boldsymbol{a}''\|_2^2 \\
&= 2\tilde{\delta}_{s'+s''}
\end{aligned}$$

and

$$\begin{aligned}
4 < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > &\geq (1 - \tilde{\delta}_s)\|\boldsymbol{a}' + \boldsymbol{a}''\|_2^2 - \|\boldsymbol{a}' - \boldsymbol{a}''\|_2^2 \\
&= -2\tilde{\delta}_{s'+s''},
\end{aligned}$$

which implies that

$$| < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > | \leq \frac{1}{2}\tilde{\delta}_{s'+s''}.$$

Hence we have

$$\theta_{s',s''} \leq \frac{1}{2}\tilde{\delta}_{s'+s''}. \tag{3.10}$$

By (3.7) and (3.10) we have

$$\begin{aligned}
| < \tilde{A}\boldsymbol{a}', \tilde{A}\boldsymbol{a}'' > | &\leq \theta_{ts'',s'} \|\boldsymbol{a}'\|_2 \|\boldsymbol{a}''\|_2 \\
&\leq \frac{1}{\sqrt{t}}\theta_{s'',s'} \|\boldsymbol{a}'\|_2 \|\boldsymbol{a}''\|_2 \\
&\leq \frac{1}{2\sqrt{t}}\tilde{\delta}_{s'+s''} \|\boldsymbol{a}'\|_2 \|\boldsymbol{a}''\|_2
\end{aligned}$$

for all ts''-sparse vector $\boldsymbol{a}'$ and s'-sparse vector $\boldsymbol{a}''$ with disjoint supports.

Lemma 3.3. For any $\boldsymbol{a} \in \boldsymbol{R}^k$, we have

$$\|\boldsymbol{a}\|_2 \leq \frac{1}{\sqrt{k}}\|\boldsymbol{a}\|_1 + \frac{\sqrt{k}}{4}\left(\max_{1\leq i\leq k}|a_i| - \min_{1\leq i\leq k}|a_i|\right). \tag{3.11}$$

Proof. The proof of this lemma can be obtained by (Cai, Wang and Xu, 2010b; Proposition 2.1).

Suppose $\boldsymbol{x}$ is an original signal we need to recover and $\boldsymbol{x}^\star$ is the solution of CS optimization problem (3.2). Let $\boldsymbol{h} \equiv \boldsymbol{x}^\star - \boldsymbol{x}$ and $\boldsymbol{h} = (h_1, \cdots, h_n)$. For simplicity, we assume that the index of $\boldsymbol{h}$ is sorted by $|h_1| \geq |h_2| \geq \cdots \geq |h_n|$. Throughout this chapter, let T_0 be an arbitrary location of $\{1, 2, \cdots, n\}$ with $|T_0| = s$ and let $\{T_1, T_2, \cdots, T_l\}$ be a decomposition of $\{1, 2, \cdots, n\}$ with $|T_1| = s$, $|T_k| = s'$ $(2 \leq k \leq l-1)$ and $1 \leq |T_l| \equiv r \leq s'$, where $|T|$ is number of elements of T. We consider the decomposition of $\boldsymbol{h}$ as follows:

$$
\begin{aligned}
\boldsymbol{h}_{T_1} &= (h_1^{(T_1)}, h_2^{(T_1)}, \cdots, h_s^{(T_1)}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_2} &= (0, \cdots, 0, h_1^{(T_2)}, \cdots, h_{s'}^{(T_2)}, 0, \cdots, 0) \\
&\vdots \\
\boldsymbol{h}_{T_{l-1}} &= (0, \cdots, 0, h_1^{(T_{l-1})}, \cdots, h_{s'}^{(T_{l-1})}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_l} &= (0, \cdots 0, h_1^{(T_l)}, \cdots, h_r^{(T_l)}).
\end{aligned}
$$

This is due to the T. Cai et al. idea (Cai, Wang and Xu, 2010b) in case of $s = s'$. We have the following Lemma 3.4–Lemma 3.10 for the decomposition $(\boldsymbol{h}_{T_1}, \boldsymbol{h}_{T_2}, \cdots, \boldsymbol{h}_{T_l})$ of $\boldsymbol{h}$. By definition of CS optimization (3.2), we have the following

Lemma 3.4. We have

$$
\|\boldsymbol{h}_{T_0^c}\|_1 \leq 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0}\|_1. \tag{3.12}
$$

Refer to (Candès, 2008) for the proof of Lemma 2.3. T. Cai et al. have obtained a similar result for the location T_1.

Lemma 3.5. For $|T_0| = |T_1| = s$, we have

$$\|\boldsymbol{h}_{T_1^c}\|_1 \le 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_1}\|_1. \tag{3.13}$$

Proof. Since $|T_0^c \cap T_1| = |T_0 \cap T_1^c|$, we have $\|\boldsymbol{h}_{T_0 \cap T_1^c}\|_1 \le \|\boldsymbol{h}_{T_0^c \cap T_1}\|_1$, which implies by (3.12) that

$$\begin{aligned}
\|\boldsymbol{h}_{T_1^c}\|_1 &= \|\boldsymbol{h}_{T_0 \cap T_1^c}\|_1 + \|\boldsymbol{h}_{T_0^c}\|_1 - \|\boldsymbol{h}_{T_1 \cap T_0^c}\|_1 \\
&\le 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_1}\|_1 \\
&\quad +2\left(\|\boldsymbol{h}_{T_0 \cap T_1^c}\|_1 - \|\boldsymbol{h}_{T_1 \cap T_0^c}\|_1\right) \\
&\le 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_1}\|_1.
\end{aligned}$$

Lemma 3.6. We have

$$\begin{aligned}
\sum_{i \ge 2} \|\boldsymbol{h}_{T_i}\|_2 \le \frac{2}{\sqrt{s'}} &\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \\
&+ \left(\frac{\sqrt{s}}{\sqrt{s'}} + \frac{\sqrt{s'}}{4\sqrt{s}}\right) \|\boldsymbol{h}_{T_1}\|_2.
\end{aligned} \tag{3.14}$$

Proof. By using Lemma 3.3, we have

$$\begin{aligned}
\|\boldsymbol{h}_{T_i}\|_2 \le \frac{1}{\sqrt{s'}} \|\boldsymbol{h}_{T_i}\|_1 + \frac{\sqrt{s'}}{4} &\left(|h_1^{(T_i)}| - |h_1^{(T_{i+1})}|\right) \\
&, 3 \le i \le l-1,
\end{aligned}$$

which implies by Lemma 3.6 that

$$\begin{aligned}
\sum_{i \ge 2} \|\boldsymbol{h}_{T_i}\|_2 &\le \frac{1}{\sqrt{s'}} \sum_{i \ge 2} \|\boldsymbol{h}_{T_i}\|_1 + \frac{\sqrt{s'}}{4} |h_1^{(T_2)}| \\
&\le \frac{2}{\sqrt{s'}} \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \left(\frac{\sqrt{s}}{\sqrt{s'}} + \frac{\sqrt{s'}}{4\sqrt{s}}\right) \|\boldsymbol{h}_{T_1}\|_2.
\end{aligned} \tag{3.15}$$

Similarly we have the following

Lemma 3.7. Let $s' < s$. We consider the decomposition $T_1 = \{T_1', T_1''\}$ of T_1 with $|T_1'| = s'$ and $|T_1''| = s''$. Then, $s' = (1-t)s,\ s'' = ts$ for some $t \in (0,1)$ and

$$\sum_{i\geq 2} \|\boldsymbol{h}_{T_i}\|_2 \leq \frac{2}{\sqrt{s(1-t)}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \left(\frac{1}{\sqrt{1-t}} + \frac{\sqrt{1-t}}{4}\right)\|\boldsymbol{h}_{T_1}\|_2. \tag{3.16}$$

We put $\|\boldsymbol{h}_{T_2}\|_1 \equiv p\sum_{i\geq 2}\|\boldsymbol{h}_{T_i}\|_1 = p\|\boldsymbol{h}_{T_1^c}\|_1$. Then $0 \leq p \leq 1$ and $\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_1 = (1-p)\|\boldsymbol{h}_{T_1^c}\|_1$. Then the following Lemma 3.8 is easily shown and Lemma 3.9 is also easily shown by using the inequality (3.15).

Lemma 3.8. We have

$$\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2^2 < \frac{p(1-p)}{s'}\|\boldsymbol{h}_{T_1^c}\|_1^2. \tag{3.17}$$

Proof. This follows from

$$\begin{aligned}\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2^2 &\leq |h_1^{(T_3)}|\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_1 \\ &\leq \frac{1}{s'}\|\boldsymbol{h}_{T_2}\|_1\left(\sum_{i\geq 2}\|\boldsymbol{h}_{T_i}\|_1 - \|\boldsymbol{h}_{T_2}\|_1\right) \\ &= \frac{p}{s'}(1-p)\|\boldsymbol{h}_{T_1^c}\|_1^2.\end{aligned}$$

Lemma 3.9. We have

$$\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2 < \frac{1-3p/4}{\sqrt{s'}}\|\boldsymbol{h}_{T_1^c}\|_1. \tag{3.18}$$

Proof. By (3.15), we have

$$\begin{aligned}\sqrt{s'}\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2 &\leq \sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_1 + \frac{s'}{4}|h_1^{(T_3)}| \\ &\leq \sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_1 + \frac{1}{4}\|\boldsymbol{h}_{T_2}\|_1 \\ &= \left(1 - \frac{3}{4}p\right)\|\boldsymbol{h}_{T_1^c}\|_1.\end{aligned}$$

Lemma 3.10. (i) Let $s' \leq s$. Suppose that A is $(s+s')$-restrictly invertible. Then,

$$\|\sum_{i\geq 3} \tilde{A}\boldsymbol{h}_{T_i}\|_2^2 \leq \frac{1}{2s'}\left((2-\delta_{s+s'})\,p(1-p)\right.$$
$$\left.+\delta_{s+s'}\left(1-\frac{3}{4}p\right)^2\right)\|\boldsymbol{h}_{T_1^c}\|_1^2. \tag{3.19}$$

(ii) Let $s' > s$. Suppose that A is $2s'$-restrictly invertible. Then,

$$\|\sum_{i\geq 3} \tilde{A}\boldsymbol{h}_{T_i}\|_2^2 \leq \frac{1}{2s'}\left((2-\delta_{2s'})\,p(1-p)\right.$$
$$\left.+\delta_{2s'}\left(1-\frac{3}{4}p\right)^2\right)\|\boldsymbol{h}_{T_1^c}\|_1^2. \tag{3.20}$$

Proof. (i) Since $\tilde{A}$ obeys the RIP of order $(s+s')$ and $r_{2s'}(A) \leq 1$, it follows from Lemma 3.2, Lemma 3.8 and Lemma 3.9 that

$$\begin{aligned}
&\|\sum_{i\geq 3} \tilde{A}\boldsymbol{h}_{T_i}\|_2^2 \\
&= \sum_{i\geq 3}\|\tilde{A}\boldsymbol{h}_{T_i}\|_2^2 + 2\sum_{3\leq i<j\leq l} < \tilde{A}\boldsymbol{h}_{T_i}, \tilde{A}\boldsymbol{h}_{T_j} > \\
&\leq \sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2^2 + 2\sum_{3\leq i<j\leq l} \frac{\tilde{\delta}_{2s'}}{2}\|\boldsymbol{h}_{T_i}\|_2\|\boldsymbol{h}_{T_j}\|_2 \\
&\leq \sum_{i\geq 3}\left(1-\frac{\tilde{\delta}_{s+s'}}{2}\right)\|\boldsymbol{h}_{T_i}\|_2^2 + \frac{\tilde{\delta}_{s+s'}}{2}\left(\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2^2\right. \\
&\quad \left.+2\sum_{3\leq i<j\leq l}\frac{\tilde{\delta}_{2s'}}{2}\|\boldsymbol{h}_{T_i}\|_2\|\boldsymbol{h}_{T_j}\|_2\right) \\
&= \left(1-\frac{\tilde{\delta}_{s+s'}}{2}\right)\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2^2 + \frac{\tilde{\delta}_{s+s'}}{2}\left(\sum_{i\geq 3}\|\boldsymbol{h}_{T_i}\|_2\right)^2 \\
&\leq \left(1-\frac{\tilde{\delta}_{s+s'}}{2}\right)\frac{p(1-p)}{s'}\|\boldsymbol{h}_{T_i^c}\|_1^2 \\
&\quad +\frac{\tilde{\delta}_{s+s'}}{2s'}\left(1-\frac{3}{4}p\right)^2\|\boldsymbol{h}_{T_i^c}\|_1^2 \\
&= \frac{1}{2s'}\left(\left(2-\tilde{\delta}_{s+s'}\right)p(1-p) + \tilde{\delta}_{s+s'}\left(1-\frac{3}{4}p\right)^2\right)\|\boldsymbol{h}_{T_i^c}\|_1^2.
\end{aligned} \tag{3.21}$$

(ii) This is shown similarly to (i).

3.3 Main results

In this section, we introduce the main results of the sufficient condition of δ_k $(k \geq s)$ under the assumption that A is restrictly invertible.

3.3.1 Bound for $\tilde{\delta}_s$

We have established the sufficient condition $\tilde{\delta}_s$ for CS optimization problem in Theorem 3.1 and Theorem 3.2.

Theorem 3.1. Assume that A is s-restrictly invertible and $\tilde{\delta}_s < \frac{2}{2+\sqrt{5}} \approx 0.472$. Then, the solution $\boldsymbol{x}^\star$ to (3.2) obeys

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1 \varepsilon, \tag{3.22}$$

where

$$C_0 = \frac{9\tilde{\delta}_s}{2\sqrt{5s}\left(1 - \frac{\sqrt{5}+2}{2}\tilde{\delta}_s\right)}, \quad C_1 = \frac{2}{r_s\left(1 - \frac{\sqrt{5}+2}{2}\tilde{\delta}_s\right)}.$$

Proof. Let $\{T_1, T_2, \cdots, T_l\}$ be a decomposition of $\{1, 2, \cdots, n\}$ with $|T_1| = s$, $|T_k| = s' < s (2 \leq k \leq l-1)$ and $1 \leq |T_l| \equiv r \leq s'$. We consider the decomposition $T_1 = \{T_1', T_1''\}$ of T_1 with $|T_1'| = s'$ and $|T_1''| = s''$. Then, $s' = (1-t)s$ and $s'' = ts$ for some $t \in (0,1)$. Since $\tilde{A} \equiv \frac{A}{r_s}$ obeys the RIP of order $s = s' + s'' = \frac{1}{t}s''$, it follows from Lemma 3.2 that

$$| < \tilde{A}\boldsymbol{h}_{T_1}, \tilde{A}\boldsymbol{h}_{T_j} > | \leq \frac{1}{2\sqrt{t}}\tilde{\delta}_s \|\boldsymbol{h}_{T_1}\|_2 \|\boldsymbol{h}_{T_j}\|_2, \quad j \geq 2,$$

which implies by $r_s(\tilde{A}) \leq 1$ that

$$
\begin{aligned}
(1-\tilde{\delta}_s)\|\boldsymbol{h}_{T_1}\|_2^2 &\leq < \tilde{A}\boldsymbol{h}_{T_1}, \tilde{A}\boldsymbol{h} - \sum_{j\geq 2} \tilde{A}\boldsymbol{h}_{T_j} > \\
&\leq 2\|\tilde{A}\boldsymbol{h}_{T_1}\|_2 \varepsilon + \sum_{j\geq 2} | < \tilde{A}\boldsymbol{h}_{T_1}, \tilde{A}\boldsymbol{h}_{T_j} > | \\
&\leq \frac{2}{r_s}\varepsilon \|\boldsymbol{h}_{T_1}\|_2 \\
&\quad + \frac{1}{2\sqrt{t}}\tilde{\delta}_s \|\boldsymbol{h}_{T_1}\|_2 \left(\sum_{j\geq 2} \|\boldsymbol{h}_{T_j}\|_2 \right).
\end{aligned}
$$

Thus, by Lemma 3.7 and the above inequality, we have

$$
\begin{aligned}
(1-\tilde{\delta}_s)\|\boldsymbol{h}_{T_1}\|_2 &\leq \frac{2}{r_s}\varepsilon + \frac{\tilde{\delta}_s}{\sqrt{(1-t)ts}} \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \\
&\quad + \frac{1}{2\sqrt{t}} \left(\frac{1}{\sqrt{1-t}} + \frac{\sqrt{1-t}}{4} \right) \tilde{\delta}_s \|\boldsymbol{h}_{T_1}\|_2.
\end{aligned} \tag{3.23}
$$

Here, put $f(t) = \frac{1}{\sqrt{t}} \left(\frac{1}{\sqrt{1-t}} + \frac{\sqrt{1-t}}{4} \right)$. Then, f is increasing when $\frac{5}{9} < t < 1$ and decreasing when $0 < t < \frac{5}{9}$. Thus, when $t = \frac{5}{9}$, we have

$$
\begin{aligned}
(1-\tilde{\delta}_s)\|\boldsymbol{h}_{T_1}\|_2 &\leq \frac{2}{r_s}\varepsilon + \frac{9}{2\sqrt{5s}}\tilde{\delta}_s \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \\
&\quad + \frac{\sqrt{5}}{2}\tilde{\delta}_s \|\boldsymbol{h}_{T_1}\|_2,
\end{aligned} \tag{3.24}
$$

so that by assumption $\tilde{\delta}_s < \frac{2}{2+\sqrt{5}} \approx 0.472$,

$$
\|\boldsymbol{h}_{T_1}\|_2 \leq \frac{1}{1-(\sqrt{5}+1)\tilde{\delta}_s} \left(\frac{2}{r_s}\varepsilon + \frac{9}{2\sqrt{5s}}\tilde{\delta}_s \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \right). \tag{3.25}
$$

Furthermore, it follows from Lemma 3.7 that

$$
\|\boldsymbol{h}_{T_1^c}\|_2 \leq \sum_{j\geq 2} \|\boldsymbol{h}_{T_j}\|_2 \leq \frac{3}{\sqrt{s}} \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \frac{5}{3}\|\boldsymbol{h}_{T_1}\|_2, \tag{3.26}
$$

which implies by (3.25) that

$$
\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq \|\boldsymbol{h}_{T_1}\|_2 + \|\boldsymbol{h}_{T_1^c}\|_2 \leq C_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1\varepsilon.
$$

This completes the proof.

By using Theorem 3.3 in (Cai and Zhang, 2013a), we obtain a better result than that of Theorem 3.1.

Theorem 3.2. Assume that A is s-restrictly invertible and $\tilde{\delta}_s < \frac{1}{2} = 0.5$. Then, the solution $\boldsymbol{x}^\star$ to (3.2) obeys

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq D_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + D_1\varepsilon,$$

where

$$D_0 = \frac{2\left(\sqrt{\tilde{\delta}_s(1-2\tilde{\delta}_s)} - (2-\sqrt{2})\tilde{\delta}_s + 1\right)}{(1-2\tilde{\delta}_s)\sqrt{s}},$$
$$D_1 = \frac{2\sqrt{2}}{(1-2\tilde{\delta}_s)r_s}.$$

Proof. Since A is s-restrictly invertible, we have

$$\frac{1}{w_s}\|\boldsymbol{a}\|_2 \leq \|A\boldsymbol{a}\|_2 \leq r_s\|\boldsymbol{a}\|_2$$

for all s-sparse vectors $\boldsymbol{a}$ in $\boldsymbol{R}^n$. We here put $\hat{A} \equiv A/\frac{\sqrt{(r_s w_s)^2+1}}{\sqrt{2}w_s}$ and $\hat{\delta}_s = \frac{(r_s w_s)^2-1}{(r_s w_s)^2+1}$. Then we have

$$(1-\hat{\delta}_s)\|\boldsymbol{a}\|_2^2 \leq \|\hat{A}\boldsymbol{a}\|_2^2 \leq (1+\hat{\delta}_s)\|\boldsymbol{a}\|_2^2$$

for all s-sparse vectors $\boldsymbol{a}$ in $\boldsymbol{R}^n$, that is, $\hat{A}$ obeys the RIP of order s with the restricted isometry constant $\hat{\delta}_s$. Furthermore, by Lemma 3.1 we have

$$\tilde{\delta}_s = 1 - \frac{1}{(r_s w_s)^2} < \frac{1}{2},$$

which implies

$$\hat{\delta}_s = 1 - \frac{2}{1+(r_s w_s)^2} < \frac{1}{3}.$$

Considering the following equality (3.27) instead of (3.1):

$$\hat{\boldsymbol{y}} = \hat{A}\boldsymbol{x} + \hat{\boldsymbol{z}}, \tag{3.27}$$

where $\hat{\boldsymbol{y}} = \frac{\boldsymbol{y}}{\sqrt{q}}$, $\hat{A} = \frac{A}{\sqrt{q}}$ and $\hat{\boldsymbol{z}} = \frac{\boldsymbol{z}}{\sqrt{q}}$ (here, q denotes $\frac{\sqrt{(r_s w_s)^2+1}}{\sqrt{2}w_s}$), it follows from Theorem 3.3 in (Cai and Zhang, 2013a) that

$$\begin{aligned}
\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \leq & \frac{2\sqrt{2}\left(2\hat{\delta}_s + \sqrt{(1-3\hat{\delta}_s)\hat{\delta}_s}\right) + 2(1-3\hat{\delta}_s)}{1-3\hat{\delta}_s} \frac{\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1}{\sqrt{s}} \\
& + \frac{2\sqrt{2(1+\hat{\delta}_s)}}{1-3\hat{\delta}_s} \frac{\varepsilon}{q} \\
= & D_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + D_1\varepsilon.
\end{aligned}$$

This completes the proof.

3.3.2 Bound for $\tilde{\delta}_k$ $(s < k)$

Using the E.J. Candès decomposition $\{T_1, T_2, \cdots, T_q\}$ of T_0^c with $|T_k| = s$ $(k = 1, \cdots, q)$ and $|h_1^{(T_1)}| \geq |h_2^{(T_1)}| \geq \cdots \geq |h_s^{(T_1)}| \geq |h_1^{(T_2)}| \geq |h_2^{(T_2)}| \geq \cdots$, Q. Mo and S. Li have obtained a new bound of the isometry constant δ_{2s} (Mo and Li, 2011). In Theorem 3.3, using the decomposition $\{T_1, T_2, \cdots, T_l\}$ of $\{1, 2, \cdots, n\}$ stated in Section 3.2 and taking an arbitrary natural number s', we have obtained a bound of the isometry constant $\tilde{\delta}_k$ $(s < k)$ under the assumption that A is k-restrictly invertible..

Theorem 3.3. (i) Let $\frac{1}{8}s \leq s' \leq s$. We assume A is $(s+s')$-restrictly invertible and $\frac{\sqrt{s}}{\sqrt{s'}}\alpha_{s,s'} < 1$, where

$$\alpha_{s,s'} = \sqrt{\frac{4(2 + 3\tilde{\delta}_{s+s'} - 4\tilde{\delta}_{s+s'}^2)}{(1-\tilde{\delta}_{s+s'})(64 - 57\tilde{\delta}_{s+s'})}}.$$

Then,

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq E_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + E_1\varepsilon, \tag{3.28}$$

where

$$\begin{aligned}
E_0 &= \frac{2\left(1 + \left(1 + \frac{\sqrt{s'}}{4\sqrt{s}}\right)\alpha_{s,s'}\right)}{\sqrt{s'}\left(1 - \sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}, \\
E_1 &= \frac{2\sqrt{2}\left(1 + \sqrt{\frac{s}{s'}} + \frac{\sqrt{s'}}{4\sqrt{s}}\right)}{r_{s+s'}\sqrt{1-\tilde{\delta}_{s+s'}}\left(1 - \sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}.
\end{aligned}$$

(ii) Let $s' \geq s$. We assume that A is $2s'$-restrictly invertible and $\frac{\sqrt{s}}{\sqrt{s'}}\alpha_{s'} < 1$, where

$$\alpha_{s'} = \sqrt{\frac{4(2+3\tilde{\delta}_{2s'} - 4\tilde{\delta}_{2s'}^2)}{(1-\tilde{\delta}_{2s'})(64-57\tilde{\delta}_{2s'})}}.$$

Then,

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq E_0'\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + E_1'\varepsilon, \tag{3.29}$$

where

$$E_0' = \frac{2\left(1 + \left(1 + \frac{\sqrt{s'}}{4\sqrt{s}}\right)\alpha_{s'}\right)}{\sqrt{s'}\left(1 - \sqrt{\frac{s}{s'}}\alpha_{s'}\right)},$$

$$E_1' = \frac{2\sqrt{2}\left(1 + \sqrt{\frac{s}{s'}} + \frac{\sqrt{s'}}{4\sqrt{s}}\right)}{r_{2s'}\sqrt{1-\tilde{\delta}_{2s'}}\left(1 - \sqrt{\frac{s}{s'}}\alpha_{s'}\right)}.$$

Proof. (i) Let $\frac{1}{8}s \leq s' \leq s$. By the definition of RIP and Lemma 3.10, we have

$$\begin{aligned}
&\left(1-\tilde{\delta}_{s+s'}\right)\|\boldsymbol{h}_{T_1}\|_2^2 \\
&= (1-\tilde{\delta}_{s+s'})\|\boldsymbol{h}_{T_1\cup T_2}\|_2^2 - (1-\tilde{\delta}_{s+s'})\|\boldsymbol{h}_{T_2}\|_2^2 \\
&\leq \|\tilde{A}\boldsymbol{h}_{T_1\cup T_2}\|_2^2 - (1-\tilde{\delta}_{s+s'})\|\boldsymbol{h}_{T_2}\|_2^2 \\
&\leq \|\tilde{A}\boldsymbol{h} - \sum_{j\geq 3}\tilde{A}\boldsymbol{h}_{T_j}\|_2^2 - \frac{(1-\tilde{\delta}_{s+s'})}{s'}p^2\|\boldsymbol{h}_{T_1^c}\|_1^2 \\
&\leq \left(\frac{2}{r_{s+s'}}\varepsilon + \|\sum_{j\geq 3}\tilde{A}\boldsymbol{h}_{T_j}\|_2\right)^2 - \frac{(1-\tilde{\delta}_{s+s'})}{s'}p^2\|\boldsymbol{h}_{T_1^c}\|_1^2 \\
&\leq \frac{4}{r_{s+s'}^2}\varepsilon^2 + \frac{4}{r_{s+s'}}\varepsilon\frac{1}{\sqrt{2s'}} \\
&\times \sqrt{(2-\tilde{\delta}_{s+s'})p(1-p) + \tilde{\delta}_{s+s'}\left(1-\frac{3}{4}p\right)^2}\|\boldsymbol{h}_{T_1^c}\|_1 \\
&+ \frac{1}{2s'}\left((2-\tilde{\delta}_{s+s'})p(1-p) + \tilde{\delta}_{s+s'}\left(1-\frac{3}{4}p\right)^2\right. \\
&\quad\left. -2(1-\tilde{\delta}_{s+s'})p^2\right)\|\boldsymbol{h}_{T_1^c}\|_1^2.
\end{aligned}$$

Since

$$\begin{aligned}
&\sqrt{(2-\tilde{\delta}_{s+s'})p(1-p) + \tilde{\delta}_{s+s'}\left(1-\frac{3}{4}p\right)^2} \\
&\leq \sqrt{\frac{8(2-\tilde{\delta}_{s+s'})}{32-25\tilde{\delta}_{s+s'}}},
\end{aligned}$$

$$(2-\tilde{\delta}_{s+s'})p(1-p)+\tilde{\delta}_{s+s'}\left(1-\frac{3}{4}p\right)^2$$
$$-2(1-\tilde{\delta}_{s+s'})p^2 \leq \frac{8(2+3\tilde{\delta}_{s+s'}-4\tilde{\delta}_{s+s'}^2)}{64-57\tilde{\delta}_{s+s'}}$$

and

$$\frac{8(2-\tilde{\delta}_{s+s'})}{32-25\tilde{\delta}_{s+s'}} \leq 2\frac{8(2+3\tilde{\delta}_{s+s'}-4\tilde{\delta}_{s+s'}^2)}{64-57\tilde{\delta}_{s+s'}},$$

we have

$$\begin{aligned}&(1-\tilde{\delta}_{s+s'})\|\boldsymbol{h}_{T_1}\|_2^2\\ &\leq \left(\frac{2\sqrt{2}}{r_{s+s'}}\varepsilon+\sqrt{\frac{1}{2s'}}\sqrt{\frac{8(2+3\tilde{\delta}_{s+s'}-4\tilde{\delta}_{s+s'}^2)}{64-57\tilde{\delta}_{s+s'}}}\|\boldsymbol{h}_{T_1^c}\|_1\right)^2,\end{aligned}$$

which implies by Lemma 3.5 that

$$\|\boldsymbol{h}_{T_1}\|_2 \leq \frac{2\sqrt{2}}{r_{s+s'}\sqrt{1-\tilde{\delta}_{s+s'}}}\varepsilon+\frac{\alpha_{s,s'}}{\sqrt{s'}}\left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1+\|\boldsymbol{h}_{T_1}\|_1\right).$$

By the assumption $\frac{\sqrt{s}}{\sqrt{s'}}\alpha_{s,s'}<1$, we have

$$\begin{aligned}\|\boldsymbol{h}_{T_1}\|_2 \leq& \frac{2\sqrt{2}}{r_{s+s'}\sqrt{1-\tilde{\delta}_{s+s'}}\left(1-\sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}\varepsilon\\ &+\frac{2\alpha_{s,s'}}{\sqrt{s'}\left(1-\sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1,\end{aligned} \tag{3.30}$$

which implies by Lemma 3.6 that

$$\begin{aligned}\|\boldsymbol{x}-\boldsymbol{x}^\star\|_2 \leq& \|\boldsymbol{h}_{T_1}\|_2+\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2\\ \leq& \left(1+\sqrt{\frac{s}{s'}}+\frac{\sqrt{s'}}{4\sqrt{s}}\right)\|\boldsymbol{h}_{T_1}\|_2\\ &+\frac{2}{\sqrt{s'}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1\\ \leq& \frac{2\left(1+\left(1+\frac{\sqrt{s'}}{4\sqrt{s}}\right)\alpha_{s,s'}\right)}{\sqrt{s'}\left(1-\sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1\\ &+\frac{2\sqrt{2}\left(1+\sqrt{\frac{s}{s'}}+\frac{\sqrt{s'}}{4\sqrt{s}}\right)}{r_{s+s'}\sqrt{1-\tilde{\delta}_{s+s'}}\left(1-\sqrt{\frac{s}{s'}}\alpha_{s,s'}\right)}\varepsilon.\end{aligned} \tag{3.31}$$

(ii) This is shown similarly to (i).

We here find the bound of the restricted isometry $\tilde{\delta}.$ by the condition $\sqrt{\frac{s}{s'}}\alpha. < 1$ in Theorem 3.2. Let $p \equiv \frac{s'}{s} > \frac{1}{8}$. Then the equality

$$(57p+16)t^2 - (121p+12)t + 64p - 8 = 0, \quad 0 < t < 1 \tag{3.32}$$

has the unique solution t_p and the following hold:
When $\frac{1}{8} < p \leq 1$, the condition $\sqrt{\frac{s}{s'}}\alpha_{s,s'} < 1$ in Theorem 3.2, (i) holds if and only if

$$\delta_{(1+p)s} < t_p. \tag{3.33}$$

When $p \geq 1$, the condition $\sqrt{\frac{s}{s'}}\alpha_{s'} < 1$ in Theorem 3.2, (ii) holds if and only if

$$\delta_{2ps} < t_p. \tag{3.34}$$

Putting $p = 1$ in (3.34), we obtain the following result for bound for $\tilde{\delta}_{2s}$.

Corollary 3.4. We assume that A is $2s$-restrictly invertible and $\tilde{\delta}_{2s} < \frac{133-\sqrt{1337}}{146} \approx 0.661$. Then,

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq E_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + E_1\varepsilon,$$

where

$$\begin{aligned} E_0 &= \frac{2\left(1 + \frac{5}{4}\alpha_s\right)}{\sqrt{s}(1-\alpha_s)}, \\ E_1 &= \frac{9\sqrt{2}}{2r_{2s}\sqrt{1-\delta_{2s}}(1-\alpha_s)}, \\ \alpha_s &= 2\sqrt{\frac{2+3\delta_{2s}-4\delta_{2s}^2}{(1-\delta_{2s})(64-57\delta_{2s})}}. \end{aligned}$$

By taking various numbers p, we can obtain some others $\tilde{\delta}_k (k > s)$. Here we give several bounds of the isometry constants $\tilde{\delta}_k$.

Example

(i) Let $p = \frac{2}{3}$. Then $t_{\frac{2}{3}} = 0.551$ and so $\tilde{\delta}_{\frac{5}{3}s} < 0.551$.

(ii) Let $p = \frac{3}{4}$. Then $t_{\frac{3}{4}} = 0.585$ and so $\tilde{\delta}_{\frac{7}{4}s} < 0.585$.

(iii) Let $p = \frac{3}{2}$. Then $t_{\frac{3}{2}} = 0.749$ and so $\tilde{\delta}_{3s} < 0.749$.

(iv) Let $p = 2$. Then $t_2 = 0.8$ and so $\tilde{\delta}_{4s} < 0.8$.

Using Theorem 2.1 in (Cai and Zhang, 2013b), we have the following

Theorem 3.5. Assume that A is ts-restrictly invertible for any $t \geq \frac{4}{3}$ and

$$\tilde{\delta}_{ts} < 2\sqrt{t-1}\left(\sqrt{t} - \sqrt{t-1}\right). \tag{3.35}$$

Then, the solution $\boldsymbol{x}^\star$ to (3.2) obeys

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \leq F_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + F_1 \varepsilon,$$

where

$$F_0 = \left(\frac{\sqrt{2}\tilde{\delta}_{ts} + \sqrt{t\left(\sqrt{\frac{t-1}{t}}(2-\tilde{\delta}_{ts}) - \tilde{\delta}_{ts}\right)\tilde{\delta}_{ts}}}{t\left(\sqrt{\frac{t-1}{t}}(2-\tilde{\delta}_{ts}) - \tilde{\delta}_{ts}\right)} + 1 \right) \frac{2}{\sqrt{s}}$$

$$F_1 = \frac{8\sqrt{t(t-1)(2-\tilde{\delta}_{ts})(1+\tilde{\delta}_{ts})}}{t\left(\sqrt{\frac{t-1}{t}}(2-\tilde{\delta}_{ts}) - \tilde{\delta}_{ts}\right)} \frac{1}{r_{ts}}.$$

Proof. This is proved similarly to Theorem 3.2. Since A is ts-restrictly invertible, $\hat{A} \equiv A / \frac{\sqrt{(r_{ts}w_{ts})^2+1}}{\sqrt{2}w_{ts}}$ obeys the RIP of order ts with the restricted isometry constant $\hat{\delta}_{ts} = \frac{(r_{ts}w_{ts})^2-1}{(r_{ts}w_{ts})^2+1}$. Then since

$$\tilde{\delta}_{ts} = 1 - \frac{1}{(r_{ts}w_{ts})^2} < 2\sqrt{t-1}(\sqrt{t} - \sqrt{t-1}),$$

it follows that $\hat{\delta}_{ts} < \sqrt{\frac{t-1}{t}}$, which implies by Theorem 2.1 in (Cai and Zhang, 2013b) that

$$
\begin{aligned}
\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le & \left(\frac{\sqrt{2}\hat{\delta}_{ts} + \sqrt{t\left(\sqrt{\frac{t-1}{t}} - \hat{\delta}_{ts}\right)\hat{\delta}_{ts}}}{t\left(\sqrt{\frac{t-1}{t}} - \hat{\delta}_{ts}\right)} + 1 \right) \frac{2}{\sqrt{s}} \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \\
& + \frac{4\sqrt{2t(t-1)(1+\hat{\delta}_{ts})}}{t\left(\sqrt{\frac{t-1}{t}} - \hat{\delta}_{ts}\right)} \frac{\sqrt{2} w_{ts}}{\sqrt{(r_{ts} w_{ts})^2 + 1}} \varepsilon \\
= & F_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + F_1 \varepsilon.
\end{aligned}
$$

This completes the proof.

Putting $t = 2$ in (3.35), we obtain the following result for bound for $\tilde{\delta}_{2s}$.

Corollary 3.6. We assume that A is $2s$-restrictly invertible and $\tilde{\delta}_{2s} < 2(\sqrt{2} - 1) \approx 0.828$. Then

$$
\|\boldsymbol{x} - \boldsymbol{x}\|_2 \le F_0 \|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + F_1 \varepsilon,
$$

where

$$
\begin{aligned}
F_0 &= \left(\frac{2(\sqrt{2} - \tilde{\delta}_{2s}) + \sqrt{\left(2\sqrt{2} - (2+\sqrt{2})\tilde{\delta}_{2s}\right)\tilde{\delta}_{2s}}}{2\sqrt{2} - (2+\sqrt{2})\tilde{\delta}_{2s}} \right) \frac{2}{\sqrt{s}}, \\
F_1 &= \frac{8\sqrt{2(2-\tilde{\delta}_{2s})(1+\tilde{\delta}_{2s})}}{2\sqrt{2} - (2+\sqrt{2})\tilde{\delta}_{2s}} \frac{1}{r_{2s}}.
\end{aligned}
$$

Chapter 4

A generalization of the restricted isometry property and applications to compressed sensing

4.1 Introduction

This chapter shows that it is possible to apply CS theory to various fields. For example, when we apply CS to a statistical model, we define A as a basis function matrix and $\boldsymbol{x}$ as a coefficient vector. We have to estimate the coefficient vector and assess this model. In this case, if A is a random matrix, we can not interpret the estimated model. Thus, in order to interpret models, it is important to discuss the method of using a matrix according to the structure of the data and the assessment of estimators. However, the RIP requires a bounded condition number for all submatrices built by selecting s arbitrary columns and the spectral norm of a matrix is generally difficult to calculate. Therefore, it seems useful to weaken the condition of RIP. E.J. Candès and Y. Plan have introduced the notion of weak RIP which is a generalization of RIP as follows (Candès and Plan, 2010):

Definition 4.1. (Weak RIP) Let $T_0 \subset \{1, 2, \cdots, n\}$ with $|T_0| = s$ and $1 < r < s$. A obeys the weak RIP with respect to T_0 of order r if there exists $0 < \delta < 1$ such that for any subset $R \subset T_0^c$ with $|R| \leq r$,

$$(1-\delta) \parallel \boldsymbol{x}_{T_0 \cup R} \parallel_2^2 \leq \parallel A\boldsymbol{x}_{T_0 \cup R} \parallel_2^2 \leq (1+\delta) \parallel \boldsymbol{x}_{T_0 \cup R} \parallel_2^2 \tag{4.1}$$

for all $\boldsymbol{x} \in \boldsymbol{R}^n$. The minimum of such constants δ is denoted by $\delta_{T_0,r}$.
Roughly speaking the notion of the weak RIP, we choose a suitable location T_0 with $|T_0| = s$ in the columns of the matrix A. We remark that A obeys the RIP of order r, but it does not necessarily obey the RIP of order $(s+r)$. Furthermore, the matrix $A_{T_0 \cup R}$ obeys the inequality (4.1) for any subset R of T_0^c with $|R| = r$. In (Candès and Plan, 2010), they have proved that under the assumptions of isotropy property and incoherence property a random matrix obeys the weak RIP with high probability $1 - 5e^{-\beta}$ if $m \geq C \log n$ (where C is a constant which only depends on β, δ, s, r and the coherent parameter μ), and have evaluated stochastically the solution of LASSO (Tibshirani, 1996) using the weak RIP and the other properties (the existence of inexact dual vector, the noise correlation bound and etc.) In this chapter, we focus on this notion and evaluate the solution of CS under the assumption of only the weak RIP without the probability, and obtain almost the same results (the following Theorem 4.1 and Theorem 4.2) as for the case of the RIP. In case that we have some information about the data, that is, we have a good location T_0, it seems better to analyze data using the weak RIP because it is much easier to construct matrices obeying the weak RIP than matrices obeying the RIP.
Throughout this section, let A be an $m \times n$ matrix. For matrix A and a subset $T \subset \{1, 2, \cdots, n\}$, A_T denotes the $m \times |T|$ matrix with column indices in T. Also, $A_{\{i\}}$ is the i-th column of A. Likewise, for a vector $\boldsymbol{a} \in \boldsymbol{R}^n$, $\boldsymbol{a}_T$ is the restriction of $\boldsymbol{a}$ to indices in T. Thus, if $\boldsymbol{a}$ is supported on T, $A\boldsymbol{a} = A_T \boldsymbol{a}_T$. Furthermore, the identity matrix, in any dimension, is denoted I, and the operator norm of A is denoted $\|A\|$.

Theorem 4.1. Let T_0 be a fixed subset of $\{1, 2, \cdots, n\}$. Assume that A obeys the weak RIP with respect to T_0 of order r, and $\delta_{T_0.r} < \frac{1}{1+\sqrt{\frac{s}{[\frac{r}{2}]}}}$, where $[\cdot]$ is the floor function. Then, the solution $\boldsymbol{x}^\star$ to (1.4) obeys

$$\parallel \boldsymbol{x} - \boldsymbol{x}^\star \parallel_2 \leq D_0 \parallel \boldsymbol{x} - \boldsymbol{x}_{T_0} \parallel_1 + D_1 \varepsilon, \tag{4.2}$$

where

$$D_0 = \frac{2}{\sqrt{[\frac{r}{2}]}\left(1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta_{T_0,r}\right)},$$

$$D_1 = \left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\frac{2\sqrt{1+\delta_{T_0,r}}}{\left(1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta_{T_0,r}\right)}.$$

In particular, if $\boldsymbol{x}$ is a T_0-sparse vector, then $\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le D_1\varepsilon$.

The following theorem is only a slight generalization of Theorem 4.1; however, it is useful in order to construct a matrix A obeying the weak RIP (see Section 4.3 for examples).

Theorem 4.2. Let T_0 be a subset of $\{1, 2, \cdots, n\}$ with $|T_0| = s$, and r be a natural number such that $1 < r < s$. Suppose that

(i) $\{A_{\{i\}};\ i \in T_0 \cup R\}$ is linearly independent for any subset R of T_0^c with $|R| = r$;

(ii) $\delta_0 \equiv 1-\left(\frac{\min_R c_R}{\max_R \|A_{T_0\cup R}\|}\right)^2 < \frac{1}{1+\frac{1}{2}\sqrt{\frac{s}{[\frac{r}{2}]}}}$, where $c_R = \max\{c > 0;\ c\|\boldsymbol{x}_{T_0\cup R}\|_2^2 \le \|A\boldsymbol{x}_{T_0\cup R}\|_2^2,\ \boldsymbol{x} \in \boldsymbol{R}^n\}$.

Then,

$$\|\boldsymbol{x} - \boldsymbol{x}^\star\|_2 \le E_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + E_1\varepsilon,$$

where

$$E_0 = \frac{1}{\sqrt{[\frac{r}{2}]}}\left(\frac{2-\delta_0}{1-\left(1+\frac{1}{2}\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta_0}\right),$$

$$E_1 = \frac{2\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)}{\max_R \|A_{T_0\cup R}\|\left(1-\left(1+\frac{1}{2}\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\right)\delta_0}.$$

4.2 Proofs

In this section, we prove Theorem 4.1 and Theorem 4.2. Here, we simply set $\delta = \delta_{T_0\cup R}$. By definition 4.1, we have the following:

Lemma 4.1. Take arbitrary subsets R_1 and R_2 of T_0^c such that $R_1 \cap R_2 = \emptyset$ and $|R_1| + |R_2| \leq r$. Then,

$$|< A\boldsymbol{a}, A\boldsymbol{b} >| \leq \delta \|\boldsymbol{a}\|_2 \|\boldsymbol{b}\|_2 \tag{4.3}$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R}^n$ with supp $\boldsymbol{a} \subset T_0 \cup R_1$ and supp $\boldsymbol{b} \subset R_2$, where $< \cdot, \cdot >$ denotes the inner product.

Proof. Take arbitrary $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{R}^n$ with supp $\boldsymbol{a} \subset T_0 \cup R_1$ and supp $\boldsymbol{b} \subset R_2$. We may suppose $\|\boldsymbol{a}\|_2 = \|\boldsymbol{b}\|_2 = 1$ without loss of generality to show the inequality (4.3). We put $R = R_1 \cup R_2$. Since A obeys the weak RIP with respect to T_0 of order r and $\boldsymbol{a}$ is orthogonal to $\boldsymbol{b}$, it follows that

$$\begin{aligned}
4 \langle A\boldsymbol{a}, A\boldsymbol{b} \rangle &= \|A_{T_0 \cup R}(\boldsymbol{a} + \boldsymbol{b})\|_2^2 - \|A_{T_0 \cup R}(\boldsymbol{a} - \boldsymbol{b})\|_2^2 \\
&\leq (1 + \delta)\|\boldsymbol{a} + \boldsymbol{b}\|_2^2 - (1 - \delta)\|\boldsymbol{a} - \boldsymbol{b}\|_2^2 \\
&= 4\delta
\end{aligned}$$

and

$$\begin{aligned}
4 \langle A\boldsymbol{a}, A\boldsymbol{b} \rangle &\geq (1 - \delta)\|\boldsymbol{a} + \boldsymbol{b}\|_2^2 - (1 + \delta)\|\boldsymbol{a} - \boldsymbol{b}\|_2^2 \\
&= -4\delta,
\end{aligned}$$

which implies that

$$|\langle A\boldsymbol{a}, A\boldsymbol{b} \rangle| \leq \delta.$$

[Proof of Theorem 4.1]

We set $\boldsymbol{h} = \boldsymbol{x}^\star - \boldsymbol{x}$. By the linearity of A and the triangle equality, we have

$$\|A\boldsymbol{h}\|_2 \leq 2\varepsilon. \tag{4.4}$$

Let T_1 be the locations of the $[\frac{r}{2}]$ largest coefficients of $\boldsymbol{h}_{T_0^c}$. Repeating this method, $\{1, 2, , \cdots, n\}$ is decomposed as $\{1, 2, \cdots, n\} = T_0 \cup T_1 \cup \cdots \cup T_{l-1} \cup T_l$, $|T_l| \leq [\frac{r}{2}]$. We may assume without loosing generality that

$$
\begin{aligned}
\boldsymbol{h}_{T_0} &= (h_1^{(T_0)}, h_2^{(T_0)}, \cdots, h_s^{(T_0)}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_1} &= (0, \cdots, 0, h_1^{(T_1)}, \cdots, h_{[\frac{r}{2}]}^{(T_1)}, 0, \cdots, 0) \\
&\vdots \\
\boldsymbol{h}_{T_{l-1}} &= (0, \cdots, 0, h_1^{(T_{l-1})}, \cdots, h_{[\frac{r}{2}]}^{(T_{l-1})}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_l} &= (0, \cdots 0, h_1^{(T_l)}, \cdots, h_{|T_l|}^{(T_l)}).
\end{aligned}
$$

Then, since

$$
\left| h_k^{(T_{j-1})} \right| \geq \max_{k \in T_j} \left| h_k^{(T_j)} \right|, \quad 2 \leq j \leq l, \ 1 \leq k \leq \left[\frac{r}{2} \right],
$$

it follows that for any j such that $2 \leq j \leq l$,

$$
\left\| \boldsymbol{h}_{T_j} \right\|_2 \leq \frac{1}{\sqrt{\left[\frac{r}{2}\right]}} \left\| \boldsymbol{h}_{T_{j-1}} \right\|_1, \tag{4.5}
$$

which implies that

$$
\begin{aligned}
\left\| \boldsymbol{h}_{(T_0 \cup T_1)^c} \right\|_2 &= \| \sum_{j \geq 2} \boldsymbol{h}_{T_j} \|_2 \\
&\leq \frac{1}{\sqrt{\left[\frac{r}{2}\right]}} \| \boldsymbol{h}_{T_0^c} \|_1.
\end{aligned} \tag{4.6}
$$

Since

$$
\begin{aligned}
\|\boldsymbol{x}\|_1 &\geq \|\boldsymbol{x}^\star\|_1 \\
&= \|\boldsymbol{x}_{T_0} + \boldsymbol{h}_{T_0} + \boldsymbol{x}_{T_0^c} + \boldsymbol{h}_{T_0^c}\|_1 \\
&\geq \|\boldsymbol{x}_{T_0}\|_1 - \|\boldsymbol{h}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0^c}\|_1 - \|\boldsymbol{x}_{T_0^c}\|_1,
\end{aligned}
$$

it follows that

$$
\|\boldsymbol{h}_{T_0^c}\|_1 \leq 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0}\|_1, \tag{4.7}
$$

which implies by (4.6) and the Schwartz inequality

$$
\begin{aligned}
\|\boldsymbol{h}_{(T_0\cup T_1)^c}\|_2 &\le \frac{1}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{h}_{T_0}\|_1 + \frac{2}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 \\
&\le \sqrt{\frac{s}{\left[\frac{r}{2}\right]}}\|\boldsymbol{h}_{T_0}\|_2 + \frac{2}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 \\
&\le \sqrt{\frac{s}{\left[\frac{r}{2}\right]}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \frac{2}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.
\end{aligned}
\tag{4.8}
$$

Furthermore, it follows from Lemma 4.1 that

$$
\left|\left\langle A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}_{T_j}\right\rangle\right| \le \delta\|\boldsymbol{h}_{T_j}\|_2\|\boldsymbol{h}_{T_0\cup T_1}\|_2,
$$

which implies by (4.4) and (4.5) that for any $j\ge 2$,

$$
\begin{aligned}
\|A\boldsymbol{h}_{T_0\cup T_1}\|_2^2 &=< A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}-\sum_{j\ge2}A\boldsymbol{h}_{T_j} > \\
&\le \|A\boldsymbol{h}_{T_0\cup T_1}\|_2\|A\boldsymbol{h}\|_2 + \sum_{j\ge 2}\left|< A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}_{T_j} >\right| \\
&\le \sqrt{1+\delta}\|\boldsymbol{h}_{T_0\cup T_1}\|_2\; 2\varepsilon + \delta\left(\sum_{j\ge2}\|\boldsymbol{h}_{T_j}\|_2\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \\
&\le \|\boldsymbol{h}_{T_0\cup T_1}\|_2\left(2\varepsilon\sqrt{1+\delta}+\delta\frac{1}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{h}_{T_0^c}\|_1\right).
\end{aligned}
\tag{4.9}
$$

Hence, it follows from (4.1) and (4.7) that

$$
\begin{aligned}
(1-\delta)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 &\le 2\varepsilon\sqrt{1+\delta} \\
&+ \frac{\delta}{\sqrt{\left[\frac{r}{2}\right]}}\left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_0\cup T_1}\|_2\right),
\end{aligned}
$$

so that

$$
\begin{aligned}
&\left(1-\left(1+\sqrt{\frac{s}{\left[\frac{r}{2}\right]}}\right)\delta\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \\
&\le 2\varepsilon\sqrt{1+\delta}+\frac{2\delta}{\sqrt{\left[\frac{r}{2}\right]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.
\end{aligned}
\tag{4.10}
$$

By assumption: $\delta < \frac{1}{1+\sqrt{\frac{s}{[\frac{r}{2}]}}}$ (if and only if, $1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta > 0$), we have

$$\begin{aligned}
&\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \\
&\leq \frac{2\sqrt{1+\delta}}{1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta}\varepsilon \\
&+ \frac{2\delta}{\left(1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta\right)\sqrt{[\frac{r}{2}]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.
\end{aligned} \tag{4.11}$$

Thus, we have by (4.8) and (4.11)

$$\begin{aligned}
\|\boldsymbol{x}-\boldsymbol{x}^\star\|_2 &= \|\boldsymbol{h}\|_2 \\
&\leq \|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \|\boldsymbol{h}_{(T_0\cup T_1)^c}\|_2 \\
&\leq \|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \sqrt{\frac{s}{[\frac{r}{2}]}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \\
&+ \frac{2}{\sqrt{[\frac{r}{2}]}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 \\
&\leq \left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\frac{2\sqrt{1+\delta}}{\left(1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta\right)}\varepsilon \\
&+ \frac{2}{\sqrt{[\frac{r}{2}]}\left(1-\left(1+\sqrt{\frac{s}{[\frac{r}{2}]}}\right)\delta\right)}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.
\end{aligned} \tag{4.12}$$

This completes the proof.

[Proof of Theorem 4.2]

Take an arbitrary $R \subset T_0^c$ with $|R| = r$. By (i), the operator $A_{T_0\cup R}$ is injective, and so there exists a positive constant $c > 0$ such that

$$c\|\boldsymbol{x}_{T_0\cup R}\|_2 \leq \|A_{T_0\cup R}\boldsymbol{x}_{T_0\cup R}\|_2$$

for each $\boldsymbol{x} \in \boldsymbol{R}^n$. We set $c_R = \max\{c > 0;\ c\|\boldsymbol{x}_{T_0\cup R}\|_2^2 \leq \|A\boldsymbol{x}_{T_0\cup R}\|_2^2,\ \boldsymbol{x} \in \boldsymbol{R}^n\}$. Then,

$$\begin{aligned}
\left(\min_R c_R\right)\|\boldsymbol{x}_{T_0\cup R}\|_2 &\leq \|A\boldsymbol{x}_{T_0\cup R}\|_2 \\
&\leq \left(\max_R \|A_{T_0\cup R}\|\right)\|\boldsymbol{x}_{T_0\cup R}\|_2
\end{aligned} \tag{4.13}$$

for all $\boldsymbol{x} \in \boldsymbol{R}^n$. Here, we put $\delta_0 = 1 - \left(\frac{\min_R c_R}{\max_R \|A_{T_0 \cup R}\|}\right)^2$. By (4.13), we have

$$\begin{aligned} &\left(\max_R \|A_{T_0 \cup R}\|^2\right)(1-\delta_0)\,\|\boldsymbol{x}_{T_0 \cup R}\|_2^2 \leq \|A\boldsymbol{x}_{T_0 \cup R}\|_2^2 \\ \leq &\left(\max_R \|A_{T_0 \cup R}\|^2\right)\|\boldsymbol{x}_{T_0 \cup R}\|_2^2 \end{aligned} \tag{4.14}$$

for all $\boldsymbol{x} \in R^n$, which implies that

$$\begin{aligned} &\left|\left\langle A\boldsymbol{h}_{T_0 \cup T_1}, A\boldsymbol{h}_{T_j}\right\rangle\right| \\ \leq &\frac{\left(\max_R \|A_{T_0 \cup R}\|^2\right)\delta_0}{2}\|\boldsymbol{h}_{T_j}\|_2 \|\boldsymbol{h}_{T_0 \cup T_1}\|_2 \end{aligned} \tag{4.15}$$

for all $j \geq 2$. Using (4.14) and (4.15), we can prove Theorem 4.2 in the same way as Theorem 4.1.

4.3 Examples

In this section, we give simple examples of $m \times n$ matrices obeying the weak RIP.

Example 4.1. Let T_0 be a subset of $\{1, 2, \cdots, n\}$ with $|T_0| = s$ and r be a natural number with $0 < r < s$. Suppose that A_{T_0} and $A_{T_0^c}$ satisfy independently the following (i) and (ii), respectively:

(i) A_{T_0} obeys the RIP of order s. We denote by δ_{T_0} the isometry constant of A_{T_0}.

(ii) $A_{T_0^c}$ obeys the RIP of order r. We denote by δ_r the isometry constant of $A_{T_0^c}$.

Furthermore, suppose that A_{T_0} and $A_{T_0^c}$ have the following relation (iii):

(iii) A obeys the mutual incoherence property (MIP), which requires that the maximum pairwise correlation of columns of $(A_{T_0}, A^c_{T_0})$ is small, i.e., $\left| < A_{\{i\}}, A_{\{j\}} > \right| < \varepsilon$ for any $i \in T_0$ and $j \in T_0^c$, where $0 < \varepsilon < \frac{1-\max(\delta_{T_0}, \delta_r)}{sr}$.

Many researchers have studied about the MIP. For example, D.L. Donoho and X. Huo have introduced the property of the MIP (Donoho and Huo, 2001, Candès and Romberg, 2007) and T. Cai *et.al.* have introduced the connections between the RIP and the MIP (Cai, Xu and Zhang, 2009).

Then the inequality (4.1) holds for all matrices $A_{T_0 \cup R}$ constructed by mixed locations

$T_0 \cup R$ of T_0 and T_0^c, which means that A obeys the weak RIP with respect to T_0 of order r. Indeed, take arbitrary $R = \{n_1', n_2', \cdots, n_r'\} \subset T_0^c$ and $\boldsymbol{x} \in \boldsymbol{R}^n$. Since

$$
\begin{aligned}
&\left| < A_{T_0}^\star A_R \boldsymbol{x}_R, \boldsymbol{x}_{T_0} > \right| \\
&\leq \sum_{j=1}^{r} \sum_{i=1}^{s} \left| < A_{\{n_i\}}, A_{\{n_j'\}} > \right| \left| x_{n_i} x_{n_j'} \right| \\
&\leq sr\varepsilon \|\boldsymbol{x}_{T_0}\|_2 \|\boldsymbol{x}_R\|_2,
\end{aligned}
$$

it follows that

$$
\begin{aligned}
&\left| < \left(A_{T_0 \cup T_R}^\star A_{T_0 \cup R} - I \right) \boldsymbol{x}, \boldsymbol{x} > \right| \\
&= \left| < \left(A_{T_0}^\star A_{T_0} - I \right) \boldsymbol{x}_{T_0}, \boldsymbol{x}_{T_0} > \right| \\
&+ 2 \left| < A_{T_0}^\star A_R \boldsymbol{x}_R, \boldsymbol{x}_{T_0} > \right| + \left| < \left(A_R^\star A_R - I \right) \boldsymbol{x}_R, \boldsymbol{x}_R > \right| \\
&\leq \delta_{T_0} \|\boldsymbol{x}_{T_0}\|_2^2 + \delta_r \|\boldsymbol{x}_R\|_2^2 + 2sr\varepsilon \|\boldsymbol{x}_{T_0}\|_2 \|\boldsymbol{x}_R\|_2 \\
&\leq (\max(\delta_{T_0}, \delta_r) + sr\varepsilon) \left(\|\boldsymbol{x}_{T_0}\|_2^2 + \|\boldsymbol{x}_R\|_2^2 \right) \\
&= \delta \|\boldsymbol{x}\|_2^2,
\end{aligned}
$$

where $\delta \equiv \max(\delta_{T_0}, \delta_r) + sr\varepsilon$. By (ii), we have $0 < \delta < 1$. Thus, A obeys the weak RIP with respect to T_0 of order r.

Example 4.2. Let T_0 and r be as in Example 3.1. Suppose the following hold:

(i) $\max\{\|A_{T_0 \cup R}\|;\ R \subset T_0^c \text{ and } |R| = r\} \leq 1$.

(ii) $\{A_{\{i\}};\ i \in T_0\}$ is linearly independent.

(iii) For any $R \subset T_0^c$ with $|R| = r$, $\{A_{\{j\}};\ j \in R\}$ is linearly independent.

It follows from (i) and (ii) that A_{T_0} obeys the RIP of order s and from (iii) that $A_{T_0^c}$ obeys the RIP of order r. We denote by δ_{T_0} and δ_r the isometry constants of A_{T_0} and $A_{T_0^c}$, respectively. It is easily shown that $\delta_{T_0} = 1 - c_{T_0}^2$ and $\delta_r = 1 - \min_R c_R^2$. Furthermore, suppose the following (iv) and (v) hold:

(iv) For $k \in \{1, 2, \cdots, n\}$, adding a vector $A_{\{k\}}'$ in $\boldsymbol{R}^{n-m}$ to the vector $A_{\{k\}}$ in $\boldsymbol{R}^m$, we construct a vector in $\boldsymbol{R}^n$ as follows:

$$B_{\{k\}} = \begin{pmatrix} A_{\{k\}} \\ A'_{\{k\}} \end{pmatrix},$$
$$\|B_{\{k\}}\|_2 = 1.$$

For any $R \subset T_0^c$ with $|R| = r$,

$$\{B_{\{i\}};\ i \in T_0\} \perp \{B_{\{j\}}; j \in R\}.$$

(v) For any $i \in T_0$,

$$1 - \|A_{\{i\}}\|_2^2 \leq \left(\frac{1 - \max(\delta_{T_0}, \delta_r)}{sr}\right)^2.$$

Then A obeys the weak RIP with respect to T of order r. Indeed, it follows from (iv) and (v) that for any $i \in T_0$ and $j \in T_0^c$

$$\begin{aligned} \|A'_{\{i\}}\|_2^2 &= \|B_{\{i\}}\|_2^2 - \|A_{\{i\}}\|_2^2 \\ &= 1 - \|A_{\{i\}}\|_2^2 \\ &\leq \left(\frac{1 - \max(\delta_{T_0}, \delta_r)}{sr}\right)^2, \end{aligned}$$

which implies by (iv) that

$$\begin{aligned} \langle A_{\{i\}}, A_{\{j\}} \rangle &= \left|\langle B_{\{i\}}, B_{\{j\}} \rangle - \langle A'_{\{i\}}, A'_{\{j\}} \rangle\right| \\ &= \left|\langle A_{\{i\}}, A_{\{j\}} \rangle\right| \\ &\leq \|A'_{\{i\}}\|_2 \|A'_{\{j\}}\|_2 \\ &\leq \|A'_{\{i\}}\|_2 \\ &\leq \frac{1 - \max(\delta_{T_0}, \delta_r)}{sr}. \end{aligned}$$

Hence, (i), (ii) and (iii) in Example 4.1 hold, and so A obeys the weak RIP with respect to T of order r.

Example 4.3. Let T_0 and r be as in Example 3.1. Suppose that the following hold:

(i) $\max\{\|A_{T_o \cup R}\|;\ R \subset T_0^c \text{ and } |R| = r\} \leq 1$.

(ii) $\{A_{\{i\}};\ i \in T_0\}$ is an orthogonal system in $\boldsymbol{R}^m$.

(iii) For any $R \subset T_0^c$ with $|R| = r$, $\{A_{\{j\}};\ j \in R\}$ is linearly independent in $\boldsymbol{R}^m$ and orthogonal to $\{A_{\{i\}};\ i \in T_0\}$.

Then, A obeys the weak RIP with respect to T_0 of order r. Indeed, this follows from

$$\begin{aligned}
\|A\boldsymbol{x}_{T_0\cup R}\|_2^2 &= \|A\boldsymbol{x}_{T_0}\|_2^2 + \|A\boldsymbol{x}_R\|_2^2 + 2\langle A\boldsymbol{x}_{T_0}, A\boldsymbol{x}_R\rangle \\
&\geq (1-\delta_{T_0})\|\boldsymbol{x}_{T_0}\|_2^2 + (1-\delta_r)\|\boldsymbol{x}_R\|_2^2 \\
&\geq (1-\max(\delta_{T_0},\delta_r))\|\boldsymbol{x}\|_2^2.
\end{aligned}$$

We remark that $\delta_{T_0} = 1 - \min\{\|A_{\{i\}}\|_2^2;\ i \in T_0\}$ and $\delta_r = 1 - \min_R c_R^2$.

The following example is a special case of Example 4.3 and it is useful itself.

Example 4.4. Let T_0 be a subset of $\{1, 2, \cdots n\}$ with $|T_0| = s$ and R_0 be a subset of T_0^c with $|R_0| = r$. Suppose that an $m \times n$ matrix A satisfies the following conditions:

(i) $\|A_{\{k\}}\|_2 \leq \frac{1}{\sqrt{s+r}}, \quad k = 1, 2, \cdots, n.$

(ii) $\{A_{\{i\}};\ i \in T_0 \cup R_0\}$ is an orthogonal system in $\boldsymbol{R}^m$.

(iii) $\{A_{\{j\}};\ j \in (T_0 \cup R_0)^c\}$ is contained in the linear span of $\{A_{\{i\}};\ i \in R_0\}$.

(iv) $\{A_{\{j\}};\ j \in R\}$ is linearly independent for each subset R of $(T_0 \cup R_0)^c$ with $|R| = r$.

Then A obeys the weak RIP with respect to T_0 of order r. Indeed, (i) in Example 4.3 follows from (i) and (iii) in Example 4.3 follows from (ii)-(iv). Hence, by Example 4.3 A obeys the weak RIP with respect to T_0 of order r.

4.4 Discussions

When analyzing data by using compressed sensing, it is common to use random matrices with no data structure. When matrices with data structure are used, it is very difficult to investigate whether these matrices obey RIP. Therefore, it seems useful to weaken the conditions of RIP. E.J. Candès and Y. Plan have defined the notion of the weak RIP as a generalization of the RIP (Candès and Plan, 2010). In this chapter, we have obtained almost the same results as for the case of the RIP. This is significant because it is much easier to construct matrices obeying the weak RIP than matrices obeying the RIP. In the case that we have some information about the data, that is, we know good locations T_0, it seems better to analyze data using the weak RIP. We believe the proposed definition has more potential applications to statistics and other fields than the original RIP.

Chapter 5

Weak RIP and its application to compressed sensing

5.1 Introduction

As shown in Section 4, it seems that the notion of weak RIP is useful in case that we have some information about the data, that is, we have a good location T_0, and it seems better to analyze data using the weak RIP because it is much easier to construct matrices obeying the weak RIP than matrices obeying the RIP. In this chapter, we give a sufficient condition under which A obeys the weak RIP with respect to T_0 of order r and evaluate the solution of CS by using a correlative relationship $\theta_{T_0,r}$ of the locations T_0 and T_0^c defined in (5.6). Furthermore, we apply this result to the case of a random matrix satisfying the isotropy property.

5.2 The weak RIP and CS

Throughout this chapter, let T_0 be a subset of $\{1, 2, \cdots, n\}$ with $|T_0| = s$ and r be a natural number with $0 < r < s$. In this section, we define the coefficient of correlation $\theta_{T_0,r}$ of A_{T_0} and $A_{T_0^c}$ and give a sufficient condition of $\theta_{T_0,r}$ under which A obeys the weak

RIP with respect to T_0 of order r and evaluate the solution of CS. We assume the following (i) and (ii):

(i) The submatrix A_{T_0} is nearly isometric, that is, there exists a constant δ $(0 \leq \delta < 1)$ such that

$$(1-\delta)\|\boldsymbol{x}\|_2^2 \leq \|A_{T_0}\boldsymbol{x}\|_2^2 \leq (1+\delta)\|\boldsymbol{x}\|_2^2 \tag{5.1}$$

for each $\boldsymbol{x} \in \boldsymbol{R}^n$ with supp $\boldsymbol{x} \subset T_0$. The minimum of such constants δ is denoted by $\delta(T_0)$.

The matrix A_{T_0} is nearly isometric if and only if it obeys the RIP of order s. It is easily shown that

$$0 < \|A_{T_0}^\star A_{T_0}\| < 2, \quad \| \left(A_{T_0}^\star A_{T_0}\right)^{-1} \| > 1 \tag{5.2}$$

and

$$\delta(T_0) = \begin{cases} 1 - \frac{1}{\|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|} \quad \text{if} \quad \|A_{T_0}^\star A_{T_0}\| \leq 1 \\ \max\left(\|A_{T_0}^\star A_{T_0}\| - 1,\ 1 - \frac{1}{\|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|}\right) \quad = \quad \max(\lambda_1 - 1,\ 1 - \lambda_s) \\ \qquad\qquad\qquad\qquad \text{if} \quad \|A_{T_0}^\star A_{T_0}\| > 1, \end{cases} \tag{5.3}$$

where λ_1 and λ_s are the maximum eigenvalue and the minimum eigenvalue of the positive matrix $A_{T_0}^\star A_{T_0}$, respectively.

(ii) $A_{T_0^c}$ obeys the RIP of order r. Let $\delta_r(T_0^c)$ denote the restricted isometry constant of $A_{T_0^c}$.

We consider the correlative relationship of the submatrices A_{T_0} and $A_{T_0^c}$. Let T be any location of T_0^c with $|T| = r$. Then we define the coefficient of correlation $\theta_{T_0,r}$ of A_{T_0} and A_T by

$$\begin{aligned} &\mu\left(T_0,\ T\right) = \sup\left\{\left| < A\boldsymbol{x},\ A\boldsymbol{y} > \right|;\ \ \text{supp}\ \boldsymbol{x} \subset T_0, \text{supp}\ \boldsymbol{y} \subset T, \|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1\right\}, \\ &\qquad \mu_{T_0,r} = \max\left\{\mu(T_0, T);\ \ T \subset T_0^c \ \text{ with }\ |T| = r\right\}. \end{aligned} \tag{5.4}$$

Then we have

$$\mu_{T_0,r} = \max\left\{\|A_T^\star A_{T_0}\|;\ \ T \subset T_0^c \text{ with } |T| = r\right\}. \tag{5.5}$$

Here we define the coefficient of correlation $\theta_{T_0,r}$ as follows:

$$\theta_{T_0,r} = \max\left(\delta(T_0),\ \delta_r(T_0^c),\ \mu_{T_0,r}\right). \tag{5.6}$$

It is easily shown that whenever $r' \leq r$

$$\delta_{r'}(T_0^c) \leq \delta_r(T_0^c) \ \text{ and } \ \mu_{T_0,r'} \leq \mu_{T_0,r}, \tag{5.7}$$

so that

$$\theta_{T_0,r'} \leq \theta_{T_0,r} \tag{5.8}$$

Suppose that A obeys the weak RIP with respect to T_0 of order r. Then it is clear that A_{T_0} is nearly isometric with $\delta(T_0) \leq \delta_{T_0,r}$ and $A_{T_0^c}$ obeys the RIP of order r with $\delta_r(T_0^c) \leq \delta_{T_0,r}$. Furthermore, since

$$|< A\boldsymbol{x}, A\boldsymbol{y} >| \leq \delta_{T_0,r}\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 \tag{5.9}$$

for each $\boldsymbol{x},\ \boldsymbol{y} \in \boldsymbol{R}^n$ with $\operatorname{supp} \boldsymbol{x} \subset T_0$ and $\operatorname{supp} \boldsymbol{y} \subset T_0^c$ with $|\operatorname{supp} \boldsymbol{y}| \leq r$, it follows that $\mu_{T_0,r} \leq \delta_{T_0,r}$. Hence we have

$$\theta_{T_0,r} \leq \delta_{T_0,r}. \tag{5.10}$$

Conversely we have the following

Theorem 5.1. Suppose that A_{T_0} is nearly isometric, $A_{T_0^c}$ obeys the RIP of order r and $\theta_{T_0,r} < \frac{1}{2}$. Then A obeys the weak RIP with respect to T_0 of order r and $\theta_{T_0,r} \leq \delta_{T_0,r} \leq 2\theta_{T_0,r}$.

Proof. Take arbitrary $\boldsymbol{x},\ \boldsymbol{y} \in \boldsymbol{R}^n$ such that $\operatorname{supp} \boldsymbol{x} \subset T_0$, $\operatorname{supp} \boldsymbol{y} \subset T_0^c$ and $|\operatorname{supp} \boldsymbol{y}| = r$.

Then, since

$$
\begin{aligned}
\|A(\boldsymbol{x}+\boldsymbol{y})\|_2^2 &= \|A\boldsymbol{x}\|_2^2 + 2\left\langle A\boldsymbol{x}, A\boldsymbol{y}\right\rangle + \|A\boldsymbol{y}\|_2^2 \\
&\le (1+\delta(T_0))\|\boldsymbol{x}\|_2^2 + 2\mu_{T_0,r}\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 + (1+\delta_r(T_0^c))\|\boldsymbol{y}\|_2^2 \\
&\le (1+2\theta_{T_0,r})\left(\|\boldsymbol{x}\|_2^2+\|\boldsymbol{y}\|_2^2\right) \\
&= (1+2\theta_{T_0,r})\|\boldsymbol{x}+\boldsymbol{y}\|_2^2 \qquad (5.11)
\end{aligned}
$$

and

$$
\begin{aligned}
\|A(\boldsymbol{x}+\boldsymbol{y})\|_2^2 &\ge \|A\boldsymbol{x}\|_2^2 - 2| < A\boldsymbol{x}, A\boldsymbol{y} > | + \|A\boldsymbol{y}\|_2^2 \\
&\ge (1-\delta(T_0))\|\boldsymbol{x}\|_2^2 - 2\mu_{T_0,r}\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2 + (1-\delta_r(T_0))\|\boldsymbol{y}\|_2^2 \\
&\ge (1-2\theta_{T_0,r})\|\boldsymbol{x}+\boldsymbol{y}\|_2^2, \qquad (5.12)
\end{aligned}
$$

it follows that

$$
(1-2\theta_{T_0,r})\|\boldsymbol{x}+\boldsymbol{y}\|_2^2 \le \|A(\boldsymbol{x}+\boldsymbol{y})\|_2^2 \le (1+2\theta_{T_0,r})\|\boldsymbol{x}+\boldsymbol{y}\|_2^2, \qquad (5.13)
$$

which implies Theorem 5.1.

We have the following result for an evaluation of the solution of CS.

Theorem 5.2. Suppose A_{T_0} is nearly isometric, $A_{T_0^c}$ obeys the RIP of order r and

$$
2\theta_{T_0,\frac{r}{5}} + \sqrt{\frac{5s}{2r}}\theta_{T_0,r} < 1. \qquad (5.14)
$$

Then A obeys the weak RIP with respect to T_0 of order r and

$$
\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le C_0\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + C_1\varepsilon,
$$

where

$$
\begin{aligned}
C_0 &= \sqrt{\frac{5}{r}}\left(\frac{1-2\theta_{T_0,\frac{r}{5}}+\sqrt{2}\theta_{T_0,r}}{1-2\theta_{T_0,\frac{r}{5}}-\sqrt{\frac{5s}{2r}}\theta_{T_0,r}}\right), \\
C_1 &= \frac{2\left(1+\sqrt{\frac{5s}{2r}}\right)\sqrt{1+\theta_{T_0,\frac{r}{5}}}}{1-2\theta_{T_0,\frac{r}{5}}-\sqrt{\frac{5s}{2r}}\theta_{T_0,r}}. \qquad (5.15)
\end{aligned}
$$

In particular, if $\boldsymbol{x}$ is a T_0-sparse, that is, supp $\boldsymbol{x} \subset T_0$, then

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \leq C_1\varepsilon.$$

Proof. We put $\boldsymbol{h} \equiv \boldsymbol{x}^\star - \boldsymbol{x}$. Then we have

$$\|\boldsymbol{h}\|_2 \leq 2\varepsilon \tag{5.16}$$

and by definition of CS optimization

$$\|\boldsymbol{h}_{T_0^c}\|_1 \leq 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0}\|_1. \tag{5.17}$$

We consider the following decomposition of $\{1, 2, \cdots, n\}$: Let T_1 be the location of the r' largest coefficients of $\boldsymbol{h}_{T_0^c}$, T_2 the location of the $r'' \equiv r - r'$ largest coefficients of $\boldsymbol{h}_{(T_0\cup T_1)^c}$ and T_3 the location of the r'' largest coefficients of $\boldsymbol{h}_{(T_0\cup T_1 \cup T_2)^c}$. Repeating this method, $\{1, 2, \cdots, n\} = T_0 \cup T_1 \cup \cdots \cup T_{l-1}$, $|T_l| \leq r''$. Then, since

$$|h_k^{T_{j-1}}| \geq \max_{k\in T_j} |h_k^{T_j}|, \ \ 2 \leq j \leq l, \ \ 1 \leq k \leq r'',$$

it follows from Proposition 2.1. in (Cai, Wang and Xu, 2010b) that

$$\|\boldsymbol{h}_{T_j}\|_2 \leq \frac{1}{\sqrt{r''}}\|\boldsymbol{h}_{T_j}\|_1 + \frac{\sqrt{r''}}{4}\left(|h_1^{T_j}| - |h_{r''}^{T_j}|\right), \quad j \geq 2$$

and

$$\begin{aligned}
\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2 &\leq \frac{1}{\sqrt{r''}}\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_1 + \frac{\sqrt{r''}}{4}|h_1^{T_2}| \\
&\leq \frac{1}{\sqrt{r''}}\|\boldsymbol{h}_{(T_0\cup T_1)^c}\|_1 + \frac{\sqrt{r''}}{4r'}\|\boldsymbol{h}_{T_1}\|_1 \\
&= \frac{1}{\sqrt{r''}}\|\boldsymbol{h}_{T_0^c}\|_1 - \left(\frac{1}{\sqrt{r''}} - \frac{\sqrt{r''}}{4r'}\right)\|\boldsymbol{h}_{T_1}\|_1.
\end{aligned} \tag{5.18}$$

By taking $r' = \frac{1}{5}r$ and $r'' = \frac{4}{5}r$, we can obtain the decompsition $\{T_1, T_2, \cdots, T_l\}$ of T_0^c, which is better than those of (Cai, Wang and Xu, 2010b) and (Candès, 2008). Then it follows from (5.17) and (5.18) that

$$\begin{aligned}
\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2 &\leq \frac{1}{\sqrt{\frac{4}{5}r}}\|\boldsymbol{h}_{T_0^c}\|_1 \\
&\leq \sqrt{\frac{5}{r}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \sqrt{\frac{5}{4r}}\|\boldsymbol{h}_{T_0}\|_1 \\
&\leq \sqrt{\frac{5}{r}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \sqrt{\frac{5s}{4r}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2.
\end{aligned} \tag{5.19}$$

Since $A_{T_0^c}$ obeys the RIP of order r, it follows that

$$|< A\boldsymbol{h}_{T_1}, A\boldsymbol{h}_{T_j} >| \le \delta_r(T_0^c)\|\boldsymbol{h}_{T_1}\|_2\|\boldsymbol{h}_{T_j}\|_2, \quad j \ge 2. \tag{5.20}$$

By the assumption (5.14), we have $\theta_{T_0,r} < \sqrt{\frac{2r}{5s}} < \frac{1}{2}$. Hence it follows from Theorem 5.1 that A obeys the weak RIP with respect to T_0 of order r. Furthermore, we can show similarly to (5.11) and (5.12) that

$$\left(1 - 2\theta_{T_0,\frac{r}{5}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2^2 \le \|A\boldsymbol{h}_{T_0\cup T_1}\|_2^2 \le \left(1 + 2\theta_{T_0,\frac{r}{5}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2^2, \tag{5.21}$$

which implies by (5.1), (5.7) and (5.8) that

$$\begin{aligned}
&\left(1 - 2\theta_{T_0,\frac{r}{5}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2^2 \\
&\le \|A\boldsymbol{h}_{T_0\cup T_1}\|_2^2 \\
&\le < A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h} > + \sum_{j\ge 2} |< A\boldsymbol{h}_{T_0\cup T_1}, A\boldsymbol{h}_{T_j} >| \\
&\le 2\varepsilon\sqrt{1 + 2\theta_{T_0,\frac{r}{5}}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \sum_{j\ge 2} |< A\boldsymbol{h}_{T_0}, A\boldsymbol{h}_{T_j} >| + \sum_{j\ge 2} |< A\boldsymbol{h}_{T_1}, A\boldsymbol{h}_{T_j} >| \\
&\le 2\varepsilon\sqrt{1 + 2\theta_{T_0,\frac{r}{5}}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \mu_{T_0,\frac{4}{5}r}\|\boldsymbol{h}_{T_0}\|_2\left(\sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2\right) + \delta_r(T_0^c)\|\boldsymbol{h}_{T_1}\|_2\left(\sum_{j\ge 2}\boldsymbol{h}_{T_j}\|_2\right) \\
&\le 2\varepsilon\sqrt{1 + 2\theta_{T_0,\frac{r}{5}}}\|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \sqrt{2}\theta_{T_0,r}\|\boldsymbol{h}_{T_0\cup T_1}\|_2\left(\sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2\right),
\end{aligned} \tag{5.22}$$

and by (5.19)

$$\left(1 - 2\theta_{T_0,\frac{r}{5}} - \sqrt{\frac{5s}{2r}}\theta_{T_0,r}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 \le 2\sqrt{1 + \theta_{T_0,\frac{r}{5}}}\varepsilon + \sqrt{\frac{10}{r}}\theta_{T_0,r}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1. \tag{5.23}$$

Hence we have by (5.19) and (5.23)

$$\begin{aligned}
\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 &\le \|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \|\boldsymbol{h}_{(T_0\cup T_1)^c}\|_2 \\
&\le \|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2 \\
&\le \left(1 + \sqrt{\frac{5s}{4r}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2 + \sqrt{\frac{5}{r}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 \\
&\le C_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1\varepsilon.
\end{aligned}$$

This completes the proof of Theorem 5.2.

We next consider a special case that A satisfies a restricted norm condition:

$$\|A_T\| \le 1, \tag{5.24}$$

for each $T \subset T_0^c$ with $|T| \le r$. Then we have the following

Theorem 5.3. Suppose that $A_{T_0}^\star A_{T_0}$ is invertible, $A_T^\star A_T$ is invertible for every $T \subset T_0^c$ with $|T| \le r$ and

$$2\theta_{T_0,\frac{r}{5}} + \sqrt{\frac{5s}{2r}} \max\left(\theta_{T_0,\frac{4}{5}r},\ \frac{1}{2}\theta_{T_0,r}\right) < 1. \tag{5.25}$$

Then we have

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le D_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + D_1\varepsilon,$$

where

$$D_0 = \sqrt{\frac{5}{r}}\left(\frac{1 - 2\theta_{T_0,\frac{r}{5}} + \sqrt{2}\max\left(\theta_{T_0,\frac{4}{5}r}, \frac{1}{2}\theta_{T_0,r}\right)}{1 - 2\theta_{T_0,\frac{r}{5}} - \sqrt{\frac{5s}{2r}}\max\left(\theta_{T_0,\frac{4}{5}r}, \frac{1}{2}\theta_{T_0,r}\right)}\right),$$

$$D_1 = \frac{2\sqrt{2}\left(1 + \sqrt{\frac{5s}{4r}}\right)}{1 - 2\theta_{T_0,\frac{r}{5}} - \sqrt{\frac{5s}{2r}}\max\left(\theta_{T_0,\frac{4}{5}r}, \frac{1}{2}\theta_{T_0,r}\right)}. \tag{5.26}$$

Proof. Since $\|A_{T_0}\| \le 1$ and $A_{T_0}^\star A_{T_0}$ is invertible, it follows that A_{T_0} is nearly isometric and $\delta(T_0) = 1 - \frac{1}{\|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|}$. Since $\|A_T\| \le 1$ and $A_T^\star A_T$ is invertible for every $T \subset T_0^c$ with $|T| \le r$, it follows that $A_{T_0^c}$ obeys the RIP of order r and

$$\delta_r(T_0^c) = 1 - \frac{1}{\min\{\|\left(A_T^\star A_T\right)^{-1}\|;\ T \subset T_0^c \text{ and } |T| \le r\}}. \tag{5.27}$$

Since $\|A_{T_0}\| \le 1$ and $\|A_T\| \le 1$ for each $T \subset T_0^c$ with $|T| \le r$, it follows that

$$| < A\boldsymbol{h}_{T_1}, A\boldsymbol{h}_{T_j} > | \le \frac{1}{2}\delta_r(T_0^c)\|\boldsymbol{h}_{T_1}\|_2\|\boldsymbol{h}_{T_j}\|_2 \quad , j \ge 2 \tag{5.28}$$

and similarly to (4.22)

$$
\begin{aligned}
&\left(1-2\theta_{T_0,\frac{r}{5}}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2^2\\
&\le \|A\boldsymbol{h}_{T_0\cup T_1}\|_2^2\\
&\le 2\sqrt{2}\varepsilon\|\boldsymbol{h}_{T_0\cup T_1}\|_2+\mu_{T_0,\frac{4}{5}r}\|\boldsymbol{h}_{T_0}\|_2\left(\sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2\right)+\frac{1}{2}\delta_r(T_0^c)\|\boldsymbol{h}_{T_1}\|_2\left(\sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2\right)\\
&\le 2\sqrt{2}\varepsilon\|\boldsymbol{h}_{T_0\cup T_1}\|_2+\sqrt{2}\max\left(\theta_{T_0,\frac{4}{5}r},\frac{1}{2}\theta_{T_0,r}\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2\left(\sum_{j\ge 2}\|\boldsymbol{h}_{T_j}\|_2\right),
\end{aligned}
$$

which implies by (4.19)

$$
\begin{aligned}
&\left(1-2\theta_{T_0,\frac{r}{5}}-\sqrt{\frac{5s}{2r}}\max\left(\theta_{T_0,\frac{4}{5}r},\frac{1}{2}\theta_{T_0,r}\right)\right)\|\boldsymbol{h}_{T_0\cup T_1}\|_2\\
&\le 2\sqrt{2}\varepsilon+\sqrt{\frac{10}{r}}\max\left(\theta_{T_0,\frac{4}{5}r},\frac{1}{2}\theta_{T_0,r}\right)\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.
\end{aligned}
$$

Hence we have

$$\|\boldsymbol{x}^\star-\boldsymbol{x}\|_2\le D_0\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1+D_1\varepsilon.$$

Remark. In case that A does not satisfy the restricted norm condition (5.24), we may obtain a similar result to Theorem 5.3 using a rescaled matrix $\tilde{A}\equiv\frac{A}{\sigma_r(T_0^c)}$, where

$$\sigma_r(T_0^c)=\sup\left\{\|A_T\|;\ \ T\subset T_0^c \text{ and } |T|=r\right\}.$$

Such rescaling technique has been used by many authors. In particular, in Section 3 it is has shown that it is useful to apply the above rescaled matrix $\tilde{A}$ to the theory of CS (Inoue, 2013b).

5.3 The weak RIP and CS with probability

In this section, we next evaluate the solution of CS in case that a random $m\times n$ matrix A satisfies the isotropy property:

$$E\left(A_{\{k\}}A_{\{k\}}^\star\right)=I \tag{5.29}$$

for every row vector $A_{\{k\}}$ of A. We put $|a_{ij}|\le\rho(A)$ (simply, ρ), $1\le i\le m$, $1\le j\le n$. Then we have the following

Theorem 5.4. For any $0 < \delta < \frac{1}{2}$, $\tilde{A} \equiv \frac{A}{\sqrt{m}}$ obeys the weak RIP with respect to T_0 order r and $\tilde{\theta}_{T_0,r} \leq \delta$ with probability at least $1 - e^{-\beta}$ provided with

$$m \geq \max \left\{ \begin{array}{c} \frac{2\left(1+\frac{\delta}{3}\right)}{\delta^2}(\rho s - 1)(\beta + \log\ 2s), \\ \frac{2\left(1+\frac{\delta}{3}\right)}{\delta^2}(\rho r - 1)\left(\beta + r \log\ \frac{n-s}{r} + r + \log\ 2r\right), \\ \frac{8\rho s r}{\delta^2}\left(\beta + \log\ (n-s) + \frac{1}{4}\right) \end{array} \right\}. \tag{5.30}$$

We consider the following equality (5.31) instead of (3.1):

$$\tilde{\boldsymbol{y}} = \tilde{A}\boldsymbol{x} + \tilde{\boldsymbol{z}}, \tag{5.31}$$

where $\tilde{\boldsymbol{y}} \equiv \frac{\boldsymbol{y}}{\sqrt{m}}$, $\tilde{A} \equiv \frac{A}{\sqrt{m}}$ and $\tilde{\boldsymbol{z}} \equiv \frac{\boldsymbol{z}}{\sqrt{m}}$. Since $\{\boldsymbol{a} \in \boldsymbol{R}^n;\ \|\boldsymbol{y} - A\boldsymbol{a}\|_2 \leq \varepsilon\} = \{\boldsymbol{a} \in \boldsymbol{R}^n;\ \|\tilde{\boldsymbol{y}} - \tilde{A}\boldsymbol{a}\|_2 \leq \frac{\varepsilon}{\sqrt{m}}\}$, it follows that the solution $\boldsymbol{x}^\star$ to the optimization problem is the same as that to (3.2).

Theorem 5.5. The solution $\boldsymbol{x}^\star$ to (3.2) obeys

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \leq C_0\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1\frac{\varepsilon}{\sqrt{m}},$$

with probability with $1 - e^{-\beta}$ provided with

$$m \geq \max \left\{ \begin{array}{c} \frac{3\left(2+\sqrt{\frac{5s}{2r}}\right)^2}{2\left(5+\sqrt{\frac{5s}{2r}}\right)}(\rho s - 1)(\beta + \log\ 2s), \\ \frac{3\left(2+\sqrt{\frac{5s}{2r}}\right)^2}{2\left(5+\sqrt{\frac{5s}{2r}}\right)}(\rho r - 1)\left(\beta + r \log\ \frac{n-s}{r} + r + \log\ 2r\right), \\ 8\left(2 + \sqrt{\frac{5s}{2r}}\right)^2 \rho s r\left(\beta + \log\ n + \frac{1}{4}\right) \end{array} \right\}. \tag{5.32}$$

We prepare some lemmas to prove Theorem 5.4 and Theorem 5.5. We assume that a random $m \times n$ matrix A satisfies the isotropy property and put $\tilde{A} = \frac{A}{\sqrt{m}}$. Let δ be any real number with $0 < \delta < 1$.

Lemma 5.1. Let T be any subset of $\{1, 2, \cdots, n\}$ with $|T| = k$. Then $\tilde{A}_T$ is nearly isometric for δ with probability $1 - 2k \exp\left\{-\frac{m}{(\rho k - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)}\right\}$.

Proof. This is due to Lemma 2.1 in (Candès and Plan, 2010). We simply give the proof. It is shown that

$$\tilde{A}_T^\star \tilde{A}_T - I = \frac{1}{m}\sum_{i=1}^{m}\left(\left(A_{\{i\}T}\right)\left(A_{\{i\}T}\right)^\star - I\right).$$

Putting $X_i = A_{\{i\}T}A_{\{i\}T}^\star - I$, $i = 1, 2, \cdots, m$, we can show that $E(X_i) = 0$, $\|X_i\| \le \rho k - 1$ and $0 \le E(X_i^2) \le (\rho k - 1)I$. Hence it follows from the matrix Bernstein inequality (Gross, 2011, Tropp, 2012) that

$$P\left(\|\tilde{A}_T^\star \tilde{A}_T - I\| \ge \delta\right) \le 2k \exp\left\{-\frac{m}{(\rho k - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)}\right\}. \tag{5.33}$$

Lemma 5.2. The following statements (i) and (ii) hold:

(i) $\tilde{A}_{T_0}$ is nearly isometric for δ with probability at least $1 - e^{-\beta}$ provided with

$$m \ge \frac{2\left(1+\frac{\delta}{3}\right)}{\delta^2}(\rho s - 1)(\beta + \log\ 2s). \tag{5.34}$$

(ii) $\tilde{A}_{T_0^c}$ obeys the RIP of order r and $\delta_r(T_0^c) \le \delta$ with probability at least $1 - e^{-\beta}$ provided with

$$\begin{aligned} m &\ge \frac{2\left(1+\frac{\delta}{3}\right)}{\delta^2}(\rho r - 1)(\beta + \log\ {}_{n-s}C_r + \log 2r) \\ &\ge \frac{2\left(1+\frac{\delta}{3}\right)}{\delta^2}(\rho r - 1)\left(\beta + r\log\ \frac{n-s}{r} + r + \log 2r\right). \end{aligned} \tag{5.35}$$

Proof. (i) This follows from Lemma 5.1.

(ii) By Lemma 5.1 we have

$$\begin{aligned} & P\left(\bigcup_T \left\{\|\tilde{A}_T^\star \tilde{A}_T - I\| \ge \delta\right\}\right) \\ \le & \sum_T P\left(\|\tilde{A}_T^\star \tilde{A}_T - I\| \ge \delta\right) \\ \le & {}_{n-s}C_r 2r \exp\left(-\frac{m}{(\rho r - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)}\right) \\ = & \exp\left(-\frac{m}{(\rho r - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)} + \log\ \left({}_{n-s}C_r 2r\right)\right) \\ \le & \exp\left(-\frac{m}{(\rho r - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)} + \log\ \left(\left(\frac{e(n-s)}{r}\right)^r 2r\right)\right) \\ = & \exp\left(-\frac{m}{(\rho r - 1)}\frac{\delta^2}{2\left(1+\frac{\delta}{3}\right)} + r\log\ \frac{n-s}{r} + r + \log 2r\right), \end{aligned} \tag{5.36}$$

where T moves all subsets of T_0^c with $|T| = r$, which implies (ii).

Lemma 5.3. We have

$$\tilde{\theta}_{T_0,r} \leq \delta \tag{5.37}$$

with probability at least $1 - e^{-\beta}$ provided with

$$m \geq \frac{8\rho s r}{\delta^2}\left(\beta + \log(n-s) + \frac{1}{4}\right). \tag{5.38}$$

Proof. By Lemma 2.5 in (Candès and Plan, 2010), we have

$$P\left(\max_{i \in T_0^c} \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \geq t\right) \leq (n-s)\exp\left(-\frac{mt^2}{8\rho s} + \frac{1}{4}\right). \tag{5.39}$$

Hence we have

$$\begin{aligned}
P\left(\max_{i \in T_0^c} \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \geq t\right) &\geq P\left(\max_{i \in T} \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \geq t\right) \\
&= P\left(\left(\bigcap_{i \in T}\left\{\|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 < t\right\}\right)^c\right) \\
&= 1 - P\left(\bigcap_{i \in T}\left\{\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 < t\right\}\right),
\end{aligned} \tag{5.40}$$

where T moves all subset of T_0^c with $|T| = r$, which implies by (5.39) that

$$\begin{aligned}
P\left(\bigcap_{i \in T}\left\{\|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 < t\right\}\right) &\geq 1 - P\left(\max_{i \in T_0^c} \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \geq t\right) \\
&\geq 1 - (n-s)\exp\left(-\frac{mt^2}{8\rho s} + \frac{1}{4}\right).
\end{aligned} \tag{5.41}$$

Since

$$\begin{aligned}
\|\tilde{A}_{T_0}^{\star} \tilde{A}_T \boldsymbol{w}\|_2 &= \|\sum_{i \in T} w_i \tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \\
&\leq \sum_{i \in T} |w_i| \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2 \\
&\leq \|\boldsymbol{w}\|_2 \left(\sum_{i \in T} \|\tilde{A}_{T_0}^{\star} \boldsymbol{a}_i\|_2^2\right)^{\frac{1}{2}}
\end{aligned} \tag{5.42}$$

for each $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)^T \in \boldsymbol{R}^n$, it follows from (5.41) that

$$\tilde{\theta}_{T_0,r} = \max\left\{\|\tilde{A}_{T_0}^{\star} \tilde{A}_T\|;\ \ T \subset T_0^c \ \text{ with } \ |T| = r\right\} \leq t\sqrt{r} \tag{5.43}$$

with at least probability $1-(n-s)\exp\left(-\frac{mt^2}{8\rho s}+\frac{1}{4}\right)$. Putting $t=\frac{\delta}{\sqrt{r}}$, Lemma 4.3 holds.

Proof of Theorem 5.4. This follows from Lemma 5.2, Lemma 5.3 and Theorem 5.1.

Proof of Theorem 5.5. We put $\delta=\frac{1}{2+\sqrt{\frac{5s}{2r}}}$. Then Theorem 5.5 follows from Theorem 5.2 and Theorem 5.4.

Remark. In (Candès and Plan, 2010), E.J. Candès and Y. Plan have shown that if A satisfies the isotropy property and

$$m \geq C_\delta \beta\rho \max\ \left(s\log\ s\rho, r\log\ n\left(\log\ r\right)^2 \log\left(r\rho\log\ n\right)\right), \tag{5.44}$$

then A obeys the weak RIP with respect to T_0 of order r with probability $\left(1-5e^{-\beta}\right)$. In Theorem 5.5, we have obtained by a simple proof that A obeys the weak RIP with respect to T_0 of order r with probability $\left(1-e^{-\beta}\right)$ if m satisfies the inequality (5.30) and the condition (5.30) of m is better than (5.44) if n is sufficient large for s (for example, $n\geq e^{(\log r)^s}$).

Chapter 6

RIPless theory for compressed sensing

6.1 Introduction

This chapter introduces the RIPless theory of compressed sensing(CS). Let $\boldsymbol{x}$ be not a sparse vector. Compressed sensing aims to recover high-dimensional signal (for example: images signal, voice signal, code signal...etc.) from only a few samples or linear measurements. Formally, one considers the following model in noiseless case:

$$\boldsymbol{y} = A\boldsymbol{x}, \tag{6.1}$$

where A is a $m \times n$ matrix($m < n$).

Our goal is to reconstruct an unknown signal $\boldsymbol{x}$ based on A and $\boldsymbol{y}$ are given. Then we consider reconstructing $\boldsymbol{x}$ as the solution $\boldsymbol{x}^\star$ to the optimization problem

$$\min_{\tilde{\boldsymbol{x}}} \|\tilde{\boldsymbol{x}}\|_1, \quad \text{subject to } \ \boldsymbol{y} = A\tilde{\boldsymbol{x}}. \tag{6.2}$$

Furthermore, one considers the following model in noisy case:

$$\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{z}, \tag{6.3}$$

where $\boldsymbol{z}$ is an unknown noise term.

In this context, we consider reconstructing $\boldsymbol{x}$ as the solution $\boldsymbol{x}^\star$ to the optimization problem

$$\min_{\tilde{\boldsymbol{x}}} \|\tilde{\boldsymbol{x}}\|_1, \quad \text{subject to } \ \|\boldsymbol{y} - A\tilde{\boldsymbol{x}}\|_2 \leq \varepsilon, \tag{6.4}$$

where ε is an upper bounded on the size of the noisy contribution.
The RIP requires a bounded condition number for all submatrices built by selecting s arbitrary columns and the spectral norm of a matrix is generally difficult to calculate. Therefore, it seems useful to weaken the condition of RIP.

In this chapter, we propose the RIPless theory and the method of an unknown signal recovery in CS. There are main benefits for considering the RIPless theory. First, we do not suppose that a matrix satisfies the condition of RIP. Moreover, we do not suppose the condition of sparsity. Practically, it is very difficult to know the condition of RIP and the sufficient condition of isometry constants. Likewise, we can not know the sparsity of $\boldsymbol{x}$. Second, the assessments of various cases lead to developments for signal analysis or other analysis. We introduce the proposed results using most simple approach. We expect that more efficient approaches are suggested as developments for many analysis.

In this chapter, suppose $\boldsymbol{x}$ is an original signal we need to recover and $\boldsymbol{x}^\star = (x_1^\star, \cdots, x_n^\star)$ is the solution of CS optimization problem (6.2) or (6.4). Let $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$, where $\boldsymbol{a}_1 = (a_{11}, a_{21}, \cdots, a_{m1})'$, $\boldsymbol{a}_2 = (a_{12}, a_{22}, \cdots, a_{m2})'$,$\cdots$,$\boldsymbol{a}_n = (a_{1n}, a_{2n}, \cdots, a_{mn})'$. For a subset $T \subset \{1, 2, \cdots, n\}$ we denote by $|T|$ number of elements of T. For a matrix A and a subset $T \subset \{1, 2, \cdots, n\}$, A_T denotes the $m \times |T|$ matrix with column indices in T.

6.2 RIPless theory

In this section, we first introduce a simple case in the RIPless theory. Suppose that the rank of $A = r$ $(s < r < m)$ and for simplicity denote by $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$ a r column vectors of A which is linearly independent. Furthermore, suppose that the others column vectors $\{\boldsymbol{a}_{r+1}, \boldsymbol{a}_{r+2}, \cdots, \boldsymbol{a}_n\}$ of A is contained in the orthogonal complement $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}^\perp$ in $\boldsymbol{R}^m$. Then we have the following:

Theorem 6.1. Suppose that $\boldsymbol{x}$ is any s-sparse vector such that support of $\boldsymbol{x}$ in $\{1, 2, \cdots, r\}$.

Then $\boldsymbol{x}$ is exactly recovered by the solution $\boldsymbol{x}^\star$ to (6.2), that is, $\boldsymbol{x} = \boldsymbol{x}^\star$.

Proof. Let $\boldsymbol{x}$ be a s-sparse vector such that support of $\boldsymbol{x}$ in $\{1, 2, \cdots, r\}$. We may put $\boldsymbol{x} = (x_1, x_2, \cdots, x_s, 0, \cdots, 0)$ without loss of generality. Since

$$\begin{aligned} 0 = \|A\boldsymbol{x}^\star - A\boldsymbol{x}\|_2^2 &= \|(x_1^\star - x_1)\boldsymbol{a}_1 + \cdots + (x_r^\star - x_r)\boldsymbol{a}_r\|_2^2 \\ &\quad + \|x_{r+1}^\star \boldsymbol{a}_{r+1} + \cdots + x_n^\star \boldsymbol{a}_n\|_2^2, \end{aligned} \tag{6.5}$$

we have

$$(x_1^\star - x_1)\boldsymbol{a}_1 + \cdots + (x_r^\star - x_r)\boldsymbol{a}_r = 0.$$

Since $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$ is linear independent, we have $x_1^\star = x_1, \cdots, x_r^\star = x_r$ and $x_k^\star = x_k = 0$, $k = s+1, \cdots, r$. We now put $\boldsymbol{x}_s = (x_1^\star, x_2^\star, \cdots, x_s^\star, 0, \cdots, 0)$. Then

$$A\boldsymbol{x}_s^\star = A\boldsymbol{x} = \boldsymbol{y}.$$

By the definition of CS optimization, we have

$$\|\boldsymbol{x}^\star\|_1 \leq \|\boldsymbol{x}_s^\star\|_1.$$

Furthermore, by the definition of $\boldsymbol{x}_s^\star$, we have

$$\|\boldsymbol{x}_s^\star\|_1 \leq \|\boldsymbol{x}^\star\|_1.$$

Thus, we have $\boldsymbol{x}^\star = \boldsymbol{x}_s^\star$ and $x_{s+1}^\star = \cdots = x_n^\star = 0$. By these discussions, we have

$$\boldsymbol{x}^\star = \boldsymbol{x}_s^\star = \boldsymbol{x}.$$

This completes the proof.

We give an example of A satisfying the conditions in Theorem 6.1:
Let $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$ $(s < r < m)$ be r column vectors of A which is mutually orthogonal and the others column vectors are contained in its orthogonal complement. For example, $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$ is the Fourier orthogonal system, that is,

$$\boldsymbol{a}_k = (e^{-i2\pi jk/m}), \quad k = 1, 2, \cdots r \text{ and } j = 1, \cdots, m$$

and

$$\boldsymbol{a}_t \in \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}^{\perp}, \quad t = r+1, \cdots, n.$$

In Section 4 and Section 5, we have focused on the notion of weak RIP and evaluate the solution of CS under the assumption of only the weak RIP and obtained almost the same results as for the case of the RIP. Thus it seems that the notion of weak RIP is useful in case that we have some information about the data, that is, we have a good location.

In Theorem 6.2 and Theorem 6.3, we discuss a RIPless theory without the assumption of weak RIP and of course RIP. We first consider in case that we have the knowledges of data, that is, we know a good location T_0. Let $K_\varepsilon \equiv \{\boldsymbol{x} \in \boldsymbol{R}^n;\ \|\boldsymbol{y} - A\boldsymbol{x}\|_2 \le \varepsilon\}$, $\varepsilon > 0$ and $K_0 \equiv \{\boldsymbol{x} \in \boldsymbol{R}^n; \boldsymbol{y} = A\boldsymbol{x}\}$. Assume that $K_\varepsilon \ne \emptyset$.

Theorem 6.2. Let T_0 be a location in $\{1, 2, \cdots, n\}$ with $|T_0| = s$. Suppose
(i) $\{\boldsymbol{a}_k; k \in T_0\}$ is linearly independent;
(ii)

$$\begin{aligned}\mu_{T_0} &\equiv \max\{| < \boldsymbol{a}_k,\ \boldsymbol{a}_j > |;\quad k \in T_0,\ j \in T_0^c\} \\ &< \frac{1}{s\|\left(A_{T_0^\star}A_{T_0}\right)^{-1}\|} = \frac{\lambda_s}{s},\end{aligned}$$

where $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_s > 0$ are eigenvalues of $A_{T_0}^\star A_{T_0}$.
Then, for every $\boldsymbol{x} \in K_\varepsilon$ we have

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le C_0^{(T_0)}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + C_1^{(T_0)}\varepsilon, \tag{6.6}$$

where

$$\begin{aligned}C_0^{(T_0)} &= \frac{2\left(1 + \mu_{T_0}\sqrt{s}\|(A_{T_0}^\star A_{T_0})^{-1}\|\right)}{1 - \mu_{T_0}s\|(A_{T_0}^\star A_{T_0})^{-1}\|}, \\ C_1^{(T_0)} &= \frac{2(1+\sqrt{s})\sqrt{\|(A_{T_0}^\star A_{T_0})^{-1}\|}}{1 - \mu_{T_0}s\|(A_{T_0}^\star A_{T_0})^{-1}\|}.\end{aligned} \tag{6.7}$$

Proof. Let T_0 be a location in $\{1, 2, \cdots, n\}$ with $|T_0| = s$. We may take $T_0 = \{1, 2, \cdots, s\}$ without loss of generality. Take an arbitrary $\boldsymbol{x} \in K_\varepsilon$ and put $\boldsymbol{h} \equiv \boldsymbol{x}^\star - \boldsymbol{x}$. Then we have

$$\begin{aligned}\|A\boldsymbol{h}_{T_0}\|_2^2 &= \left\langle A\boldsymbol{h}_{T_0},\ A\boldsymbol{h} - A\boldsymbol{h}_{T_0^c}\right\rangle \\ &\le 2\varepsilon\|A\boldsymbol{h}_{T_0}\|_2 + \left\langle A\boldsymbol{h}_{T_0},\ A\boldsymbol{h}_{T_0^c}\right\rangle\end{aligned} \tag{6.8}$$

and

$$\begin{aligned}
\left|\left\langle A\boldsymbol{h}_{T_0},\ A\boldsymbol{h}_{T_0^c}\right\rangle\right| &= \left|\sum_{k=1}^{s}\sum_{j=s+1}^{n} h_k h_j < \boldsymbol{a}_k,\ \boldsymbol{a}_j >\right| \\
&\le \mu_{T_0}\sum_{k=1}^{s}\sum_{j=s+1}^{n} |h_k h_j| \\
&\le \mu_{T_0}\|\boldsymbol{h}_{T_0}\|_1\|\boldsymbol{h}_{T_0^c}\|_1 \\
&\le \mu_{T_0}\sqrt{s}\|\boldsymbol{h}_{T_0}\|_2\|\boldsymbol{h}_{T_0^c}\|_1. \qquad (6.9)
\end{aligned}$$

By definition of CS optimization (6.4) we have

$$\begin{aligned}
\|\boldsymbol{h}_{T_0^c}\|_1 &\le 2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0}\|_1 \\
&\le 2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_0}\|_2, \qquad (6.10)
\end{aligned}$$

and so by (6.7),

$$\left|\left\langle A\boldsymbol{h}_{T_0},\ A\boldsymbol{h}_{T_0^c}\right\rangle\right| \le \mu_{T_0}\sqrt{s}\|\boldsymbol{h}_{T_0}\|_2\left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1+\sqrt{s}\|\boldsymbol{h}_{T_0}\|_2\right). \qquad (6.11)$$

We also have

$$\|\boldsymbol{h}_{T_0}\|_2 \le \sqrt{\|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|}\|A\boldsymbol{h}_{T_0}\|_2. \qquad (6.12)$$

Indeed, this follows from

$$\begin{aligned}
\|\boldsymbol{h}_{T_0}\|_2^2 &= \left\langle \left(A_{T_0}^\star A_{T_0}\right)^{-1}\left(A_{T_0}^\star A_{T_0}\right)\boldsymbol{h}_{T_0},\ \boldsymbol{h}_{T_0}\right\rangle \\
&= \left\langle \left(A_{T_0}^\star A_{T_0}\right)^{-1}\left(A_{T_0}^\star A_{T_0}\right)^{\frac{1}{2}}\boldsymbol{h}_{T_0},\ \left(A_{T_0}^\star A_{T_0}\right)^{\frac{1}{2}}\boldsymbol{h}_{T_0}\right\rangle \\
&\le \|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|\ \|\left(A_{T_0}^\star A_{T_0}\right)^{\frac{1}{2}}\boldsymbol{h}_{T_0}\|_2^2 \\
&= \|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|\ \|A\boldsymbol{h}_{T_0}\|_2^2.
\end{aligned}$$

Using (6.8), (6.11) and (6.12), we get

$$\|A\boldsymbol{h}_{T_0}\|_2 \le 2\varepsilon + \mu_{T_0}\sqrt{r_{T_0}s}\left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1+\sqrt{s}\|\boldsymbol{h}_{T_0}\|_2\right),$$

where $r_{T_0} \equiv \|\left(A_{T_0}^\star A_{T_0}\right)^{-1}\|$. By using (6.12) again, we have

$$\|\boldsymbol{h}_{T_0}\|_2 \le 2\sqrt{r_{T_0}}\varepsilon + 2\mu_{T_0}r_{T_0}\sqrt{s}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \mu_{T_0}r_{T_0}s\|\boldsymbol{h}_{T_0}\|_2.$$

Therefore, we obtain

$$(1-\mu_{T_0} r_{T_0} s)\,\|\boldsymbol{h}_{T_0}\|_2 \leq 2\sqrt{r_{T_0}}\varepsilon + 2\mu_{T_0} r_{T_0}\sqrt{s}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1. \tag{6.13}$$

By the assumption (ii), we have

$$\|\boldsymbol{h}_{T_0}\|_2 \leq \frac{2\sqrt{r_{T_0}}}{1-\mu_{T_0} r_{T_0} s}\varepsilon + \frac{2\mu_{T_0} r_{T_0}\sqrt{s}}{1-\mu_{T_0} r_{T_0} s}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1. \tag{6.14}$$

Since

$$\begin{aligned}\|\boldsymbol{h}_{T_0^c}\|_2 \leq \|\boldsymbol{h}_{T_0^c}\|_1 &\leq 2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_0}\|_1 \\ &\leq 2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_0}\|_2,\end{aligned}$$

it follows from (6.14) that

$$\begin{aligned}\|\boldsymbol{x}^\star-\boldsymbol{x}\|_2 &\leq \|\boldsymbol{h}_{T_0}\|_2 + \|\boldsymbol{h}_{T_0^c}\|_2 \\ &\leq 2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + (1+\sqrt{s})\|\boldsymbol{h}_{T_0}\|_2 \\ &\leq \frac{2(1+\sqrt{s})\sqrt{r_{T_0}}}{1-\mu_{T_0} r_{T_0} s}\varepsilon + \frac{2(1+\mu_{T_0} r_{T_0}\sqrt{s})}{1-\mu_{T_0} r_{T_0} s}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.\end{aligned}$$

This completes the proof.

By Theorem 6.2 we have the following:

Corollary 6.1. Suppose that A satisfies the conditions (i) and (ii) in Theorem 6.2. Then, the following results are obtained.

Noiseless case: For any T_0-sparse vector in K_0 we have

$$\boldsymbol{x} = \boldsymbol{x}^\star.$$

Noisy case: For any T_0-sparse vector in K_ε we have

$$\|\boldsymbol{x}^\star-\boldsymbol{x}\|_2 \leq C_1^{(T_0)}\varepsilon.$$

For a general case that we do not know any good location we have the following:

Theorem 6.3. Suppose that $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$ satisfies the following

(i) $\{\boldsymbol{a}_k;\ k \in T\}$ is linearly independent for each $T \subset \{1, 2, \cdots, n\}$ with $|T| = s$;

(ii) $\mu \equiv \max_T \mu_T = \max_{1 \le i \ne j \le n} | < \boldsymbol{a}_i, \boldsymbol{a}_j > | < \frac{1}{s \max\left(\| \left(A_T^\star A_T \right)^{-1} \|; |T| = s \right)} = \frac{\lambda_s^T}{s}$, where $\lambda_1^T \ge \lambda_2^T \ge \cdots \ge \lambda_s^T > 0$ are eigenvalues of $A_T^\star A_T$.

Then, for any vector $\boldsymbol{x} \in K_\varepsilon$, we have

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le C_0 \|\boldsymbol{x} - \boldsymbol{x}_s\|_1 + C_1 \varepsilon,$$

where $\boldsymbol{x}_s$ is the vector consisting of the s largest entries of $\boldsymbol{x}$ in magnitude and $r = \max\{r_T;\ T \subset \{1, 2, \cdots, n\}$ with $|T| = s\}$ and

$$C_0 = \frac{4 + 5\mu r s}{2\sqrt{s}(1 - \mu r s)},$$
$$C_1 = \frac{9\sqrt{r}}{2(1 - \mu r s)}.$$

Proof. The proof of this theorem can be obtained based on Theorem 6.2. We show the modifications (6.16)-(6.19). For simplicity, we assume that the index of $\boldsymbol{h}$ is sorted by $|h_1| \ge |h_2| \ge \cdots \ge |h_n|$. Take an arbitrary location T_0 of $\{1, 2, \cdots, n\}$ with $|T_0| = s$ and let $\{T_1, T_2, \cdots, T_l\}$ be a decomposition of $\{1, 2, \cdots, n\}$ with $|T_k| = s$ $(1 \le k \le l - 1)$ and $1 \le |T_l| \le s$, where $|T|$ is number of elements of T. We consider a decomposition of $\boldsymbol{h}$ as follows:

$$\begin{aligned}
\boldsymbol{h}_{T_1} &= (h_1^{(T_1)}, h_2^{(T_1)}, \cdots, h_s^{(T_1)}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_2} &= (0, \cdots, 0, h_1^{(T_2)}, \cdots, h_s^{(T_2)}, 0, \cdots, 0) \\
&\vdots \\
\boldsymbol{h}_{T_{l-1}} &= (0, \cdots, 0, h_1^{(T_{l-1})}, \cdots, h_s^{(T_{l-1})}, 0, \cdots, 0) \\
\boldsymbol{h}_{T_l} &= (0, \cdots 0, h_1^{(T_l)}, \cdots, h_{|T_l|}^{(T_l)}).
\end{aligned}$$

This decomposition is due to the T. Cai *et.al.* idea (Cai, Wang and Xu, 2010b) and they have obtained the following inequality:

$$\|\boldsymbol{h}_{T_1^c}\|_1 \le 2\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \|\boldsymbol{h}_{T_1}\|_1 \tag{6.15}$$

Then we can show similaly to (6.11)-(6.13) in Theorem 6.2 that

$$\begin{aligned}\|\boldsymbol{h}_{T_1}\|_2 &\leq \sqrt{\|\left(A_{T_1}^\star A_{T_1}\right)^{-1}\|}\|A\boldsymbol{h}_{T_1}\|_2 \\ &\leq \sqrt{r}\|A\boldsymbol{h}_{T_1}\|_2 \,. \end{aligned} \tag{6.16}$$

$$\begin{aligned}< A\boldsymbol{h}_{T_1}, \sum_{j\geq 2} A\boldsymbol{h}_{T_j} > &\leq \mu\sqrt{s}\|\boldsymbol{h}_{T_1}\|_2 \left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_1}\|_2\right) \\ &\leq \mu\sqrt{rs}\|A\boldsymbol{h}_{T_1}\|_2 \left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_1}\|_2\right), \end{aligned} \tag{6.17}$$

$$\frac{\|\boldsymbol{h}_{T_1}\|_2}{\sqrt{r}} \leq \|A\boldsymbol{h}_{T_1}\|_2 \leq 2\varepsilon + \mu\sqrt{rs}\left(2\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \sqrt{s}\|\boldsymbol{h}_{T_1}\|_2\right). \tag{6.18}$$

Therefore, we have

$$(1-\mu rs)\|\boldsymbol{h}_{T_1}\|_2 \leq 2\sqrt{r}\varepsilon + 2\mu r\sqrt{s}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1.$$

By the assumption (ii), we have

$$\|\boldsymbol{h}_{T_1}\|_2 \leq \frac{2\sqrt{r}}{1-\mu rs}\varepsilon + \frac{2\mu r\sqrt{s}}{1-\mu rs}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1. \tag{6.19}$$

Furthermore, by using the T. Cai $et.al.$ result of Proposition 2.1 in (Cai, Wang and Xu, 2010b):

$$\|\boldsymbol{a}\|_2 \leq \frac{1}{\sqrt{k}}\|\boldsymbol{a}\|_1 + \frac{\sqrt{k}}{4}\left(\max_{1\leq i\leq k}|a_i| - \min_{1\leq i\leq k}|a_i|\right), \quad \boldsymbol{a}\in\boldsymbol{R}^k,$$

we have

$$\begin{aligned}\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2 &\leq \frac{1}{\sqrt{s}}\sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_1 + \frac{\sqrt{s}}{4}|h_1^{(T_2)}| \\ &= \frac{1}{\sqrt{s}}\|\boldsymbol{h}_{T_1^c}\|_1 + \frac{\sqrt{s}}{4}|h_1^{(T_2)}|,\end{aligned}$$

which implies by (6.15) that

$$\begin{aligned}\|\boldsymbol{h}_{T_1^c}\|_2 &\leq \sum_{j\geq 2}\|\boldsymbol{h}_{T_j}\|_2 \\ &\leq \frac{2}{\sqrt{s}}\|\boldsymbol{x}-\boldsymbol{x}_{T_0}\|_1 + \frac{5}{4}\|\boldsymbol{h}_{T_1}\|_2. \end{aligned} \tag{6.20}$$

By (6.19) and (6.20), we have

$$\begin{aligned}
\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 &\le \|\boldsymbol{h}_{T_1}\|_2 + \|\boldsymbol{h}_{T_1^c}\|_2 \\
&\le \frac{2}{\sqrt{s}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \frac{9}{4}\|\boldsymbol{h}_{T_1}\|_2 \\
&\le \frac{2}{\sqrt{s}}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \frac{9}{4}\left(\frac{2\sqrt{r}}{1-\mu rs}\varepsilon + \frac{2\mu r\sqrt{s}}{1-\mu rs}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1\right) \\
&= \frac{4+5\mu rs}{2\sqrt{s}(1-\mu rs)}\|\boldsymbol{x} - \boldsymbol{x}_{T_0}\|_1 + \frac{9\sqrt{r}}{2(1-\mu rs)}\varepsilon.
\end{aligned}$$

Thus, Theorem 6.3 is obtained by putting $T_0 = \operatorname{supp} \boldsymbol{x}_s$. This completes the proof.

Corollary 6.2. Suppose that A satisfies the conditions (i) and (ii) in Theorem 6.3. Then, the following results are obtained.

Noiseless case: For any s-sparse vector $\boldsymbol{x}$ in K_0, we have

$$\boldsymbol{x} = \boldsymbol{x}^\star.$$

Noisy case: For any s-sparse vector $\boldsymbol{x}$ in K_ε, we have

$$\|\boldsymbol{x}^\star - \boldsymbol{x}\|_2 \le C_1\varepsilon.$$

Bibliography

R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, Constructive Approximation. 28(3) (2008) 253-263.

T. Cai, G. Xu, J. Zhang, On recovery of sparse signals via l_1 minimization, IEEE Transactions on Information Theory. 55 (2009) 3388-3397.

T. Cai, L. Wang, G. Xu, Shifting inequality and recovery of sparse signals, IEEE Transactions on Signal Processing. 58 (2010a) 1300-1308.

T. Cai, L. Wang, G. Xu, New bounds for restricted isometry constants, IEEE Transactions on Information Theory. 56 (2010b) 4388-4394.

T. Cai, A. Zhang, Sharp RIP Bound for Sparse Signal and Low-Rank Matrix Recovery, Applied and Computational Harmonic Analysis. 35 (2013a) 74-93.

T. Cai, A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, Technical Report. (2013b).

E.J. Candès, The restricted isometry property and its implications for compressed sensing, C. R. Acad. Sci. Ser. I 346 (2008) 589-592.

E.J. Candès, Y. Plan, Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements, IEEE Transactions on Information Theory. 57(4) (2009) 2342-2359.

E.J. Candès, Y. Plan, A probabilistic and RIPless theory of compressed sensing, IEEE Transactions on Information Theory. 57(11) (2010) 7235-7254.

E.J. Candès, J. Romberg, Sparsity and incoherence in compressive sampling, Inverse Problems. 23(3) (2007) 969-985.

E.J. Candès, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Communications on Pure and Applied Mathematics. 59(8) (2006a) 1207-1223.

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

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

E.J. Candès, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, IEEE Transactions on Information Theory. 52(12) (2006) 5406-5425.

D.L. Donoho, Compressed sensing, IEEE Transactions on Information Theory. 52,(2006) 1289-1306.

D.L. Donoho, X. Huo, Uncertainty principles and ideal atomic decomposition, IEEE Transactions on Information Theory. 47 (2001) 2845-2862.

S. Foucart, A note on guaranteed sparse recovery via l_1-minimization, Applied and Computational Harmonic Analysis. 29 (2010) 97-103.

D. Gross, Recovering low-rank matrices from few coefficients in any basis. IEEE Transactions on Information Theory. 57(3) (2011) 1548-1566.

H. Inoue, RIPless Theory for Compressed Sensing, MI preprint Series. 2013-5 (2013).

H. Inoue, A generalization of the restricted isometry property and applications to compressed sensing, Journal of Math-for-Industry. 5 (2013B) 129-133.

H. Inoue, Sufficient conditions for CS-recovery, British Journal of Mathematics & Computer Science. 4(2) (2014) 184-198 (In press).

H. Inoue, Weak RIP and its application to Compressed Sensing, British Journal of Mathematics & Computer Science. (In press).

Q. Mo, S. Li, New bounds on the restricted isometry constant δ_{2k}, Applied and Computational Harmonic Analysis. 31 (2011) 460-468.

M. Rudelson, Random vectors in the isotropic position, Journal of Functional Analysis. 164(1) (1999) 60-72.

M. Rudelson, R. Vershynin, On sparse reconstruction from Fourier and Gaussian measurements, Communications on Pure and Applied Mathematics. 61(8) (2008) 1025-1045.

J.A. Tropp, User-friendly tail bounds for sums of random matrices, Foundations of Computational Mathematics. 12(4) (2012) 389-434.

R. Tibshirani, Regression shrinkage and selection via the Lasso, Journal of the Royal Statistical Society Series B. 58 (1996) 267-288.