Stability Analysis for a Class of Sparse Optimization Problems

Jialiang Xu[a] and Yun-Bin Zhao[b]

[a,b] School of Mathematics, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

ARTICLE HISTORY
Compiled April 29, 2019

ABSTRACT
The sparse optimization problems arise in many areas of science and engineering, such as compressed sensing, image processing, statistical and machine learning. The ℓ_0-minimization problem is one of such optimization problems, which is typically used to deal with signal recovery. The ℓ_1-minimization method is one of the plausible approaches for solving the ℓ_0-minimization problems, and thus the stability of such a numerical method is vital for signal recovery. In this paper, we establish a stability result for the ℓ_1-minimization problems associated with a general class of ℓ_0-minimization problems. To this goal, we introduce the concept of restricted weak range space property (RSP) of a transposed sensing matrix, which is a generalized version of the weak RSP of the transposed sensing matrix introduced in [Zhao et al., Math. Oper. Res., 44(2019), 175-193]. The stability result established in this paper includes several existing ones as special cases.

KEYWORDS
Sparsity optimization; ℓ_1-minimization; stability; optimality condition; Hoffman theorem; restricted weak range space property.

1. Introduction

The sparsity is a useful assumption under which the sparse optimization models arise frequently in many areas in science and engineering. Let $A \in R^{m\times n}(m \ll n)$, $B \in R^{l\times n}(l < n)$ and $U \in R^{m\times h}(m \ll h)$ be three given full-row-rank matrices. Let $y \in R^m$ and $b \in R^l$ be given vectors and ε be a positive number. Consider the following sparse optimization model:

$$\begin{array}{ll} \min\limits_{x\in R^n} & \|x\|_0 \\ \text{s.t.} & a_1\|y-Ax\|_2 + a_2\left\|U^T(Ax-y)\right\|_\infty + a_3\left\|U^T(Ax-y)\right\|_1 \le \varepsilon \\ & Bx \le b, \end{array} \tag{1}$$

where $\|x\|_0$ is called the 'ℓ_0-norm' which counts the number of nonzero components of x, and a_1, a_2 and a_3 are given nonnegative parameters satisfying $\sum_{i=1}^3 a_i = 1$. Many problems in signal and image processing (see, e.g., [6, 13, 17]) and statistical regressions [23] can be formulated as the form (1) or its special cases. In problem (1), the constraint $Bx \le b$ is motivated by some practical applications. For instance, many

Yun-Bin Zhao. Email: y.zhao.2@bham.ac.uk

signal recovery models might need to include certain constraints reflecting special structures of the target signal. For simplicity, we define

$$\phi(x) = U^T(Ax - y),$$

and write the problem (1) as

$$\min_{x \in R^n} \left\{ \|x\|_0 : \ a_1 \|y - Ax\|_2 + a_2 \|\phi(x)\|_\infty + a_3 \|\phi(x)\|_1 \leq \varepsilon, Bx \leq b \right\}.$$

The following ℓ_0-minimization models are clearly the special cases of (1):

(C1) $\min_x \{\|x\|_0 : \ y = Ax\}$; (C2) $\min_x \{\|x\|_0 : \ \|y - Ax\|_2 \leq \varepsilon\}$;
(C3) $\min_x \{\|x\|_0 : \ \left\|U^T(Ax - y)\right\|_1 \leq \varepsilon\}$; (C4) $\min_x \{\|x\|_0 : \ \left\|U^T(Ax - y)\right\|_\infty \leq \varepsilon\}$.

The problem (C1) is often called the standard ℓ_0-minimization problem [8, 17, 28]. Two structured sparsity models, called the nonnegative sparsity model [7, 8, 17, 28] and the monotonic sparsity model (isotonic regression) [23, 24], are also the special cases of the model (1).

It is well known that ℓ_1-minimization is a useful method to solve the ℓ_0-minimization problem. By replacing the ℓ_0-norm with the ℓ_1-norm in problem (1), we immediately obtain the ℓ_1-minimization problem

$$\min_x \{\|x\|_1 : \ a_1 \|y - Ax\|_2 + a_2 \|\phi(x)\|_\infty + a_3 \|\phi(x)\|_1 \leq \varepsilon, Bx \leq b\}. \tag{2}$$

Similar to its ℓ_0 counterpart, the problem (2) includes the following special cases:

(D1) $\min_x \{\|x\|_1 : \ y = Ax\}$; (D2) $\min_x \{\|x\|_1 : \ \|y - Ax\|_2 \leq \varepsilon\}$;
(D3) $\min_x \{\|x\|_1 : \ \left\|U^T(Ax - y)\right\|_1 \leq \varepsilon\}$; (D4) $\min_x \{\|x\|_1 : \ \left\|U^T(Ax - y)\right\|_\infty \leq \varepsilon\}$.

The problem (D2) is often called quadratically constrained basis pursuit [10, 17, 28], and it reduces to (D1) if $\varepsilon = 0$, which is called standard ℓ_1-minimization or the basis pursuit [8, 12, 17, 19, 26]. The problem (D4) is the type of Dantzig Selectors [9, 17].

From both numerical and theoretical viewpoints, it is important to know how close the solutions of ℓ_0- and ℓ_1-minimization problems are. To address this question, one needs to study the stability of ℓ_1-minimization methods. The stability of a sparse optimization method can be described as follows: For any $x \in R^n$ in the feasible set of a sparse optimization problem, the solution $x^\#$ generated by the method satisfies the following bound:

$$\left\|x - x^\#\right\|_2 \leq C_1 \sigma_k(x)_1 + C_2 \varepsilon \tag{3}$$

where C_1 and C_2 are constants, and $\sigma_k(x)_1$ is called the error of the best k-term approximation of the vector x (see, e.g., [12, 17]):

$$\sigma_k(x)_1 = \min_z \{\|x - z\|_1 : \ \|z\|_0 \leq k\}.$$

In this paper, we establish a stability result for the ℓ_1-minimization method (2). The stability of (D1) and (D2) has been investigated by Donoho, Candès, Tao, Romberg and others [3, 6–8, 12–14, 16, 25] under various assumptions such as the so-called restricted isometry property (RIP) of order k, mutual coherence, stable null space property (NSP) of order k or robust NSP of order k. The RIP of order k was introduced by Candès and Tao [8] to study the stability of ℓ_1-minimization. The singular-value-property-based stability analysis for (D1), (D2) and the Dantzig Selector have also been performed by Tang and Nehorai in [22].

A new and unified stability analysis for ℓ_1-minimization methods has been developed by Zhao, Jiang and Luo [29] under the assumption of weak RSP of order k, which has been proven as a necessary and sufficient condition for the standard ℓ_1-minimization to be stable. The main differnece between the weak-RSP-based-analysis and existing ones lies in the constants C_1 and C_2 in (3). Specifically, the constants C_1 and C_2 in (3) are determined by the RIP or NSP constant in existing analysis [3, 8, 17]. However, in [28, 29], these constants are determined by the so-called Robinson's constant. Motivated by the new analysis tool introduced in [29], we develop the stability result for the model (2) in this paper under the assumption of restricted weak range space property (RSP) of order k (which will be introduced in next section). Our result extends the stability theorem for ℓ_1-minimization established by Zhao et al. [28–30].

This paper is organized as follows. In Section 2, we introduce the concept of restricted weak RSP of order k. An approximation of the solution set of (2) will be discussed in Section 3. Then, in Section 4, we show the main stability result of this paper. Finally, some special cases are discussed in Section 5.

Notation

The field of real numbers is denoted by R and the n-dimensional Euclidean space is denoted by R^n. Let R^n_+ and R^n_- be the sets of nonnegative and nonpositive vectors, respectively. Unless otherwise stated, the identity matrix of suitable size is denoted by I. Given a vector $u \in R^n$, $|u|$, $(u)^+$ and $(u)^-$ denote the vectors with components $|u|_j = |u_j|$, $[(u)^+]_j = \max\{u_j, 0\}$ and $[(u)^-]_j = \min\{u_j, 0\}$, $j = 1, ..., n$, respectively. The cardinality of the set S is denoted by $|S|$ and the complementary set of $S \subseteq \{1, ..., n\}$ is denoted by $\bar{S}$, i.e., $\bar{S} = \{1, ..., n\} \setminus S$. For a given vector $x \in R^n$, x_S denotes the vector supported on S. $a_{i,j}$ denotes the entry of the matrix A in row i and column j. For the set $S \subseteq \{1, ..., n\}$, A_S denotes the submatrix of $A \in R^{m \times n}$ obtained by deleting the columns indexed by $\bar{S}$. For a matrix $A = (a_{i,j})$, $|A|$ represents the absolute version of A, i.e., $|A| = (|a_{i,j}|)$. $\mathcal{R}\left(A^T\right) = \{A^T y : y \in R^m\}$ is the range space of A^T. $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$, where $p \geq 1$, is a norm, called the ℓ_p-norm of x. $\|x\|_\infty = \max_{i=1}^n |x_i|$ is called the ℓ_∞-norm of x. For $1 \leq p, q \leq \infty$, $\|A\|_{p \to q} = \sup_{\|x\|_p \leq 1} \|Ax\|_q$ is the matrix norm induced by ℓ_p- and ℓ_q-norms.

2. Restricted weak range space property

The RSP of order k of a transposed matrix was first introduced in [26, 27] to develop a necessary and sufficient condition for the uniform recovery of sparse signals via ℓ_1-minimization. Zhao et al. [29] generalised the RSP of order k to the following weak RSP of order k to develop a stability theory for convex optimization algorithms:

Definition 2.1 (weak RSP of order k)**.** Given a matrix $A \in R^{m \times n}$, A^T is said to satisfy the weak RSP order k if for any two disjoint sets $J_1, J_2 \subseteq \{1, ..., n\}$ satisfying

$|J_1| + |J_2| \leq k$, there exists a vector $\eta \in \mathcal{R}\left(A^T\right)$ such that

$$\begin{cases} \eta_i = 1 & \text{if } i \in J_1, \\ \eta_i = -1 & \text{if } i \in J_2, \\ |\eta_i| \leq 1 & \text{if } i \notin J_1 \cup J_2. \end{cases}$$

In [28, 29], it was shown that the weak RSP of order k is a sufficient condition for the stability of many convex optimization methods, and it is also a necessary stability condition for many optimization methods.

Different from the problems (D1)-(D4), the problem (2) is more general than these models. To investigate the stability of the problem (2), we need to extend the notion of weak RSP of order k to the so-called restricted weak RSP of order k, which is defined as follows:

Definition 2.2 (Restricted weak RSP of order k)**.** Given matrices $A \in R^{m\times n}$ and $B \in R^{l\times n}$, the pair $\left(A^T, B^T\right)$ is said to satisfy the restricted weak RSP of order k if for any two disjoint sets $J_1, J_2 \subseteq \{1, ..., n\}$ satisfying $|J_1| + |J_2| \leq k$, there exists a vector $\eta \in \mathcal{R}\left(A^T, B^T\right)$ such that $\eta = \left(A^T, B^T\right)\begin{pmatrix} \nu \\ h \end{pmatrix}$ where $\nu \in R^m$, $h \in R^l_-$ and

$$\begin{cases} \eta_i = 1 & \text{if } i \in J_1, \\ \eta_i = -1 & \text{if } i \in J_2, \\ |\eta_i| \leq 1 & \text{if } i \notin J_1 \cup J_2. \end{cases}$$

It is worth mentioning that a generalized version of the RSP of order k is also used in [31] to study the exact sign recovery in 1-bit compressive sensing.

3. Approximation of (2) and its solution set

By introducing the slack variables r, s, ξ and v, the problem (2) can be rewritten as

$$\begin{array}{ll} \min\limits_{(x,r,s,\xi,v)} & \|x\|_1 \\ \text{s.t.} & a_1 s + a_2\xi + a_3 \left(e^h\right)^T v \leq \varepsilon, \\ & r \in s\mathcal{B},\ r = y - Ax,\ (s, \xi, v) \geq 0, \\ & \|\phi(x)\|_\infty \leq \xi,\ |\phi(x)| \leq v,\ Bx \leq b, \end{array} \tag{4}$$

where e^h is the vector of ones in R^h and $\mathcal{B}$ is the unit ℓ_2-ball defined as $\mathcal{B} = \{z \in R^m : \|z\|_2 \leq 1\}$. The unit ball $\mathcal{B}$ can be also described as

$$\mathcal{B} = \bigcap_{\|\mathbf{a}\|_2=1} \{z \in R^m : \mathbf{a}^T z \leq 1\}. \tag{5}$$

Denote the set E by

$$\begin{aligned} \mathrm{E} = \{(x, s, \xi, v) : a_1 s + a_2\xi + a_3(e^h)^T v \leq \varepsilon,\ Bx \leq b,\ \|\phi(x)\|_\infty \leq \xi, \\ |\phi(x)| \leq v,\ (s, \xi, v) \geq 0\}, \end{aligned}$$

and hence the solution set of (4) can be represented as

$$\Omega^* = \{(x, r, s, \xi, v) : \|x\|_1 \leq \theta^*,\ r \in s\mathcal{B},\ r = y - Ax,\ (x, s, \xi, v) \in \mathrm{E}\}, \tag{6}$$

where θ^* is the optimal value of (4). By replacing $\mathcal{B}$ in (6) with a polytope $P \supseteq \mathcal{B}$, we can get the relaxation of Ω^*, denoted by Ω_P, i.e.,

$$\Omega_P = \{(x, r, s, \xi, v) : \|x\|_1 \leq \theta^*,\ r \in sP,\ r = y - Ax,\ (x, s, \xi, v) \in \mathrm{E}\}. \tag{7}$$

The polytope Ω_P can approximate Ω^* to any level of accuracy provided that P is chosen suitably. Recall the Hausdorff metric of two sets $M_1, M_2 \subseteq R^m$:

$$\delta^{\mathcal{H}}(M_1, M_2) = \max\left\{ \sup_{x \in M_1} \inf_{z \in M_2} \|x - z\|_2 , \sup_{z \in M_2} \inf_{x \in M_1} \|x - z\|_2 \right\}.$$

Following the analysis in [28, 29] (see Lemmas 5.1, 5.2 and 5.3 in [29]), we can obtain the following lemma:

Lemma 3.1. *Let ε be the given number in problem (2). Then for any $\varepsilon' \leq \varepsilon$, there exists a polytope approximation P of $\mathcal{B}$ satisfying $P \supseteq \mathcal{B}$ and*

$$\delta^{\mathcal{H}}(\Omega^*, \Omega_P) \leq \varepsilon'. \tag{8}$$

In the remainder of this paper, we fix $\varepsilon' \in (0, \varepsilon]$ and choose the polytope P such that Ω_P and Ω^* satisfy (8). The polytope P can be represented as the intersection of a finite number of half spaces:

$$P = \{z \in R^m : (\mathbf{a}^i)^T z \leq 1, 1 \leq i \leq L\},$$

where $\mathbf{a}^i$, $1 \leq i \leq L$ are some unit vectors (i.e., $\|\mathbf{a}^i\|_2 = 1$), and L is an integer number. By adding the $2m$ half spaces

$$(\beta^j)^T z \leq 1,\ \ -(\beta^j)^T z \leq 1,\ j = 1, ..., m$$

to P, where β^j is the jth column of the $m \times m$ identity matrix, we obtain the following polytope:

$$\begin{aligned} P_0 &= P \cap \left\{ z \in R^m : (\beta^j)^T z \leq 1, -(\beta^j)^T z \leq 1, j = 1, ..., m \right\} \\ &= \left\{ z \in R^m : (\mathbf{a}^i)^T z \leq 1, 1 \leq i \leq L; (\beta^j)^T z \leq 1, -(\beta^j)^T z \leq 1, j = 1, ..., m \right\}. \end{aligned} \tag{9}$$

We define T as the collection of the vectors $\mathbf{a}^i$ and $\pm\beta^j$ in P_0, that is,

$$T := \{\mathbf{a}^i : 1 \leq i \leq L\} \cup \{\pm\beta^j : 1 \leq j \leq m\}.$$

Clearly, P_0 still satisfies (8) in Lemma 3.1, i.e.,

$$\delta^{\mathcal{H}}(\Omega^*, \Omega_{P_0}) \leq \varepsilon'.$$

In the remainder of the chapter, we use the above defined polytope P_0. Let $N = |T|$, and let M_{P_0} be the matrix with column vectors in T. Thus P_0 can be written as

$$P_0 = \{z \in R^m : (M_{P_0})^T z \leq e^N\},$$

where e^N is the vector of ones in R^N.

By replacing $\mathcal{B}$ by P_0, we obtain the following approximation of the optimal value θ^* of (2):

$$\begin{aligned} \theta^*_{P_0} := & \min_{(x,r,s,\xi,v)} \{\|x\|_1 : \ r \in sP_0, \ r = y - Ax, \ (x,s,\xi,v) \in \mathrm{E}\} \\ = & \min_{(x,s,\xi,v)} \{\|x\|_1 : \ (M_{P_0})^T (y - Ax) \leq se^N, \ (x,s,\xi,v) \in \mathrm{E}\}. \end{aligned}$$

The associated approximation problem of (2) can be written as

$$\min_{(x,s,\xi,v)} \{\|x\|_1 : (M_{P_0})^T (y - Ax) \leq se^N, \ (x,s,\xi,v) \in \mathrm{E}\}. \tag{10}$$

The solution set of (10) is

$$\Omega^*_{P_0} = \{x \in R^n : \ \|x\|_1 \leq \theta^*_{P_0}, \ r \in sP_0, \ r = y - Ax, \ (x,s,\xi,v) \in \mathrm{E}\}. \tag{11}$$

Note that $\mathcal{B} \subseteq P_0$ implies that $\theta^* \geq \theta^*_{P_0}$. So we can see that $\Omega^*_{P_0} \subseteq \Omega_{P_0}$. By the definition of P_0, we also have $\Omega^* \subseteq \Omega_{P_0}$. In the next section, we prove the main result for the problem (2).

4. Main result

Introducing a variable t yields the following equivalent form of (10):

$$\begin{aligned} \min_{(x,t,s,\xi,v)} \quad & e^T t \\ \text{s.t.} \quad & a_1 s + a_2 \xi + a_3 \left(e^h\right)^T v \leq \varepsilon, \ Bx \leq b, \ |x| \leq t, \\ & (M_{P_0})^T (y - Ax) \leq se^N, \ (t,s,\xi,v) \geq 0, \\ & \|\phi(x)\|_\infty \leq \xi, |\phi(x)| \leq v. \end{aligned} \tag{12}$$

The solution set of (12) is given as (11). Note that the above optimization problem is equivalent to a linear programming problem. In fact, the constraint $\|\phi(x)\|_\infty \leq \xi$ can be rewritten as $|\phi(x)| \leq \xi e^h$, where e^h is the vector of ones in R^h. Thus the model (12) can be rewritten explicitly as the linear programming problem

$$\begin{aligned} \min_{(x,t,s,\xi,v)} \quad & e^T t \\ \text{s.t.} \quad & x + t \geq 0, \ \ -x + t \geq 0, \\ & -a_1 s - a_2 \xi - a_3 \left(e^h\right)^T v \geq -\varepsilon, M^T_{P_0} Ax + e^N s \geq M^T_{P_0} y, \\ & U^T Ax + \xi e^h \geq U^T y, -U^T Ax + \xi e^h \geq -U^T y, \\ & U^T Ax + v \geq U^T y, -U^T Ax + v \geq -U^T y, \\ & -Bx \geq -b, \ (t,s,\xi,v) \geq 0. \end{aligned} \tag{13}$$

The dual problem of (13) is given as follows:

$$
\begin{array}{ll}
\max\limits_{w} & -\varepsilon w_3 + y^T M_{P_0} w_4 + y^T U(w_5 - w_6 + w_7 - w_8) - b^T w_9 \\
\text{s.t.} & w_1 - w_2 + A^T M_{P_0} w_4 + A^T U(w_5 - w_6 + w_7 - w_8) - B^T w_9 = 0, \\
& w_1 + w_2 \le e, \\
& -a_1 w_3 + (e^N)^T w_4 \le 0, \\
& -a_2 w_3 + (e^h)^T (w_5 + w_6) \le 0, \\
& -a_3 w_3 e^h + w_7 + w_8 \le 0, \\
& w_1, w_2 \in R^n_+, \ w_3 \in R_+, \ w_4 \in R^N_+, \ w_{5-8} \in R^h_+, \ w_9 \in R^l_+.
\end{array}
\tag{14}
$$

The optimality condition yields the following lemma:

Lemma 4.1. *Denote by $u = (x, t, s, \xi, v, w)$. Then x^* is an optimal solution of* (10) *if and only if there exists a vector $u^* = (x^*, t^*, s^*, \xi^*, v^*, w^*) \in \Theta$, where Θ is the set given as*

$$
\Theta = \left\{ u : \begin{array}{l}
-x - t \le 0, \ x - t \le 0, \ a_1 s + a_2 \xi + a_3 \left(e^h\right)^T v \le \varepsilon, \\
-M_{P_0}^T Ax - e^N s \le -M_{P_0}^T y, \ Bx \le b, \\
-U^T Ax - \xi e^h \le -U^T y, \ U^T Ax - \xi e^h \le U^T y, \\
-U^T Ax - v \le -U^T y, \ U^T Ax - v \le U^T y, \\
w_1 - w_2 + A^T M_{P_0} w_4 + A^T U(w_5 - w_6 + w_7 - w_8) - B^T w_9 = 0, \\
w_1 + w_2 \le e, \quad -a_1 w_3 + (e^N)^T w_4 \le 0, \ (t, s, \xi, v, w) \ge 0, \\
-a_2 w_3 + (e^h)^T (w_5 + w_6) \le 0, \quad -a_3 e^h w_3 + w_7 + w_8 \le 0, \\
e^T t = -\varepsilon w_3 + y^T M_{P_0} w_4 + y^T U(w_5 - w_6 + w_7 - w_8) - b^T w_9
\end{array} \right\}.
$$

Clearly, $|x^*| = t^*$ holds for every $u^* \in \Theta$. The set Θ can be written as the form

$$
\Theta = \left\{ u : M_1' u \le p', \ M_2' u = q' \right\}, \tag{15}
$$

where the vectors $q' = 0$ and

$$
p' = \left[\begin{array}{cccccccccccccc} 0 & 0 & \varepsilon & -M_{P_0}^T y & b & -U^T y & U^T y & -U^T y & U^T y & e & 0 & 0 & 0 & 0 \end{array} \right. \\
\left. \begin{array}{cccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]^T.
$$

The matrices M_1' and M_2' in (15) are given as follows:

$$
M_1' = \begin{bmatrix} D^1 & 0 \\ 0 & D^2 \\ D^3 & 0 \\ 0 & -\widetilde{I} \end{bmatrix}, \ M_2' = \left[\begin{array}{cc} M_* & M_{**} \end{array} \right], \tag{16}
$$

where the matrices M_*, M_{**}, D^1, D^2 and D^3 and $\widetilde{I}$ are given as follows:

$$
M_* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & I & -I & 0 \\ 0 & e^T & 0 & 0 & 0 & 0 & 0 & \epsilon \end{bmatrix},
$$

$$M_{**} = \begin{bmatrix} A^T M_{P_0} & A^T U & -A^T U & A^T U & -A^T U & -B^T \\ -y^T M_{P_0} & -y^T U & y^T U & -y^T U & y^T U & b^T \end{bmatrix},$$

$$D^1 = \begin{bmatrix} -I & -I & 0 & 0 & 0 \\ I & -I & 0 & 0 & 0 \\ 0 & 0 & a_1 & a_2 & a_3 e^T \\ -M_{P_0}^T A & 0 & -e^N & 0 & 0 \\ B & 0 & 0 & 0 & 0 \\ -U^T A & 0 & 0 & -e^h & 0 \\ U^T A & 0 & 0 & -e^h & 0 \\ -U^T A & 0 & 0 & 0 & -I^h \\ U^T A & 0 & 0 & 0 & -I^h \end{bmatrix}, \; D^3 = \begin{bmatrix} 0 & -I & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -I^h \end{bmatrix},$$

$$D^2 = \begin{bmatrix} I & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_1 & (e^N)^T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_2 & 0 & (e^h)^T & (e^h)^T & 0 & 0 & 0 \\ 0 & 0 & -a_3 e^h & 0 & 0 & 0 & I^h & I^h & 0 \end{bmatrix}, \; \widetilde{I} = I^{2n+1+N+4h+l}.$$

In the above matrices, 0's are zero matrices with suitable sizes and I, I^h and $\widetilde{I}$ are the $n \times n$, $h \times h$ and $(2n+1+N+4h+l) \times (2n+1+N+4h+l)$ identity matrices, respectively.

To prove the main stability result, we also need the next two Lemmas.

Lemma 4.2 (Hoffman [18, 21]). *Let $M_1 \in R^{m \times n}$ and $M_2 \in R^{l \times n}$ be two given matrices and the set $\mathcal{Q}$ be given as*

$$\mathcal{Q} = \{x \in R^n : M_1 x \le p, \;\; M_2 x = q\}.$$

For any vector $x \in R^n$, there exists a vector $x^* \in \mathcal{Q}$ satisfying

$$\|x - x^*\|_2 \le \sigma(M_1, M_2) \left\| \begin{bmatrix} (M_1 x - p)^+ \\ M_2 x - q \end{bmatrix} \right\|_1,$$

where $\sigma(M_1, M_2)$ is a constant determined by M_1 and M_2.

The constant $\sigma(M_1, M_2)$ is also called the Robinson constant. We also use the following lemma in the proof of the main result in this section.

Lemma 4.3 ([28, 30]). *Let $\pi_S(x)$ be the projection of x into the convex set S, i.e., $\pi_S(x) = \arg\min_{z \in S} \|x - z\|_2$. Let the three convex compact sets T_1, T_2 and T_3 satisfy that $T_1 \subseteq T_2$ and $T_3 \subseteq T_2$. Then for any $x \in R^n$ and any $z \in T_3$ the following holds:*

$$\|x - \pi_{T_1}(x)\|_2 \le \delta^{\mathcal{H}}(T_1, T_2) + 2\|x - z\|_2 .$$

We also define two types of constants. Let

$$C = \left[A^T, B^T\right]^T = \begin{bmatrix} A \\ B \end{bmatrix} \tag{18}$$

be a matrix with full row rank. Given three positive numbers $c, d, \widehat{d} \in [1, \infty]$, we define the constants $\Upsilon(d, \widehat{d})$ and $\vartheta(c)$ as follows:

$$\Upsilon(d, \widehat{d}) = \max_{\mho \subseteq \{1,\ldots,h\}, |\mho|=m} \|U_{\mho}^{-1}\|_{\widehat{d}\to d} \left\| \left(CC^T\right)^{-1} C \right\|_{\infty\to\widehat{d}}, \tag{19a}$$

$$\vartheta(c) = \left\| \left(CC^T\right)^{-1} C \right\|_{\infty\to c}. \tag{19b}$$

We will use the above constants together with the specific constants $\Upsilon(1,1), \Upsilon(\infty,\infty)$ and $\vartheta(1)$ in the stability analysis of (2). The main result is given as follows.

Theorem 4.4. *Let the problem data $(U, A, B, \varepsilon, a_1, a_2, a_3, b, y)$ of (2) be given, and the matrix $C \in R^{(m+l)\times n}$ be given in (18) with full row rank. Let P_0 be the polytope given in (9) satisfying (8). If $\left(A^T, B^T\right)$ satisfies the restricted weak* RSP *of order k, then for any $x \in R^n$, there is an optimal solution x^* of (2) satisfying the bound*

$$\begin{aligned} \|x - x^*\|_2 \ \leq \ & \varepsilon' + 2\sigma' \Big\{ 2\sigma_k(x)_1 + \varepsilon\hat{\Upsilon} + \|(Bx-b)^+\|_1 + + \|Bx-b\|_{c'} \vartheta(c) + \\ & \|\phi(x)\|_{d'} \Upsilon(d,\widehat{d}) + (a_1 \|y - Ax\|_2 + a_2 \|\phi(x)\|_\infty + a_3 \|\phi(x)\|_1 - \varepsilon)^+ \Big\}. \end{aligned} \tag{20}$$

where σ' is the Robinson constant determined by (M_1', M_2') in (16), $\Upsilon(d, \widehat{d})$ and $\vartheta(c)$ are the constants given in (19a) and (19b), and $\hat{\Upsilon} = \max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}$. $\widehat{d}, d, c, d', c' \in [1, +\infty]$ are five given positive numbers (allowing to be ∞) satisfying

$$\frac{1}{c} + \frac{1}{c'} = 1 \text{ and } \frac{1}{d} + \frac{1}{d'} = 1. \tag{21}$$

In particular, if x is a feasible solution of (2), then there is an optimal solution x^* of (2) such that

$$\|x - x^*\|_2 \leq \varepsilon' + 2\sigma' \Big\{ \varepsilon\hat{\Upsilon} + 2\sigma_k(x)_1 + \|\phi(x)\|_{d'} \Upsilon(d,\widehat{d}) + \|Bx-b\|_{c'} \vartheta(c) \Big\}. \tag{22}$$

Proof. Let x be any given vector in R^n and P_0 be the fixed polytope given in (9) satisfying (8) in Lemma 3.1. We let (t, s, ξ, v) satisfy that

$$t = |x|, \ s = \left\|(M_{P_0})^T (y - Ax)\right\|_\infty, \ \xi = \left\|U^T(y - Ax)\right\|_\infty, \ v = \left|U^T(y - Ax)\right|. \tag{23}$$

With such a choice of (t, s, ξ, ν), we have

$$\begin{array}{l} (-x-t)^+ = 0, \ (x-t)^+ = 0, \ \left(M_{P_0}^T(y-Ax) - e^N s\right)^+ = 0, \\ \left(U^T(y-Ax) - \xi e^h\right)^+ = 0, \ \left(-U^T(y-Ax) - \xi e^h\right)^+ = 0, \\ \left(U^T(y-Ax) - v\right)^+ = 0, \ \left(-U^T(y-Ax) - v\right)^+ = 0. \end{array} \tag{24}$$

Let J be the support set of k largest absolute entries of x, and J_1 and J_2 be the sets such that

$$J_1 = \{i : \; x_i > 0, \; i \in J\}, \; J_2 = \{i : \; x_i < 0, \; i \in J\}.$$

Clearly, $|J_1 \cup J_2| = |J| = |J_1| + |J_2| \le k$. Let J_3 be the complementary set of J. Clearly, J_1, J_2 and J_3 are disjoint. Under the assumption of restricted weak RSP of order k, there exists a vector $\eta \in R\left(A^T, B^T\right)$ such that $\eta = A^T \nu^* + B^T h^*$ for some $\nu^* \in R^m$ and $h^* \in R^l_-$ satisfying

$$\eta_i = 1 \text{ for } i \in J_1; \; \eta_i = -1 \text{ for } i \in J_2; \; |\eta_i| \le 1 \text{ for } i \in J_3. \tag{25}$$

Now we construct a feasible solution $w = (w_1, ..., w_9)$ to the dual problem (14).

Constructing (w_1, w_2). Set w_1 and w_2 as follows:

$$\begin{cases} (w_1)_i = 0, \; (w_2)_i = 1, & i \in J_1; \\ (w_1)_i = 1, \; (w_2)_i = 0, & i \in J_2; \\ (w_1)_i = \frac{1-\eta_i}{2}, \; (w_2)_i = \frac{1+\eta_i}{2}, & i \in J_3. \end{cases}$$

Such w_1 and w_2 satisfy that

$$w_1 + w_2 \le e, \; w_2 - w_1 = \eta, \; w_1, w_2 \ge 0. \tag{26}$$

Constructing $(w_5–w_8)$. Note that U is a matrix with full row rank. There must exist an invertible $m \times m$ matrix of U, denoted by $U_\mho$, where $\mho \subseteq \{1, ..., h\}$ with $|\mho| = m$. Denote the complementary set of $\mho$ by $\bar{\mho} = \{1, ..., h\} \setminus \mho$. Then we construct a vector $g \in R^h$ satisfying $g_\mho = U_\mho^{-1} \nu^*$ and $g_{\bar{\mho}} = 0$, which imply that

$$Ug = \nu^*. \tag{27}$$

Let g^+ (g^-) be the vector obtained by keeping the positive (negative) components of g and setting the remaining components to 0. By using the vector g, w_5–w_8 can be constructed as follows:

$$w_5 = a_2 g^+, \; w_6 = -a_2 g^-, \; w_7 = a_3 g^+, \; w_8 = -a_3 g^-, \tag{28}$$

which implies that

$$w_5 - w_6 + w_7 - w_8 = (a_2 + a_3) g, \; w_5, w_6, w_7, w_8 \ge 0. \tag{29}$$

Constructing w_4. Without loss of generality, we suppose that the first m columns in M_{P_0} are $\beta_j, \; j = 1, ..., m$, and $-\beta_j, \; j = 1, ..., m$ are the second m columns of M_{P_0}. The components of w_4 can be assigned as follows:

$$\begin{cases} (w_4)_j = a_1 \nu_j^*, & \text{if } \; \nu_j^* > 0, \; j = 1, ..., m; \\ (w_4)_{j+m} = -a_1 \nu_j^*, & \text{if } \; \nu_j^* < 0, \; j = 1, ..., m; \\ 0, & \text{otherwise.} \end{cases}$$

From this choice of w_4, we can see that

$$M_{P_0} w_4 = a_1 \nu^*, \quad \|w_4\|_1 = a_1 \|\nu^*\|_1 \text{ and } w_4 \geq 0. \tag{30}$$

Constructing w_3. Let $w_3 = \max\{\|\nu^*\|_1, \|g\|_1, \|g\|_\infty\}$. Such a choice of w_3 together with the choice of w_4–w_8 implies that

$$\begin{cases} \left(-a_1 w_3 + (e^N)^T w_4\right)^+ & \leq \left(-a_1 \|\nu^*\|_1 + (e^N)^T w_4\right)^+ = 0, \\ \left(-a_2 w_3 + e^T(w_5 + w_6)\right)^+ & \leq (-a_2 \|g\|_1 + a_2 \|g\|_1)^+ = 0, \\ (-a_3 e w_3 + w_7 + w_8)^+ & \leq (-a_3 e \|g\|_\infty + a_3 |g|)^+ = 0. \end{cases} \tag{31}$$

Constructing w_9. Let $w_9 = -h^*$. Clearly, $w_9 \geq 0$ due to $h^* \leq 0$.

With the above choice of w, we deduce from (26), (29), (30) and (31) that

$$\begin{cases} w_1 - w_2 + A^T M_{P_0} w_4 + A^T U(w_5 - w_6 + w_7 - w_8) - B^T w_9 = 0, \\ (w_1 + w_2 - e)^+ = 0, \ \left(-a_1 w_3 + (e^N)^T w_4\right)^+ = 0, \\ \left(-a_2 w_3 + e^T(w_5 + w_6)\right)^+ = 0, \ (-a_3 e w_3 + w_7 + w_8)^+ = 0, \\ t^- = 0, \ s^- = 0, \ \xi^- = 0, \ v^- = 0, \ w^- = 0. \end{cases} \tag{32}$$

Let $\mathcal{X}$ and $\mathcal{Y}$ be defined as follows:

$$\begin{cases} \mathcal{X} = e^T t + \varepsilon w_3 - y^T M_{P_0} w_4 - y^T U(w_5 - w_6 + w_7 - w_8) + b^T w_9, \\ \mathcal{Y} = \left(a_1 s + a_2 \xi + a_3 e^T v - \varepsilon\right)^+. \end{cases}$$

For the vector $u = (x, t, s, \xi, \nu, w)$ where (t, s, ξ, ν, w) is constructed above, by Lemma 4.2, there exists a vector $\hat{u} \in \Theta$, where Θ is given in Lemma 4.1 and written as (15), such that

$$\|u - \hat{u}\|_2 \leq \sigma' \left\| \begin{bmatrix} \mathcal{X} \\ \mathcal{Y} \\ (Bx - b)^+ \\ (-x - t)^+ \\ (x - t)^+ \\ \left(M_{P_0}^T (y - Ax) - e^N s\right)^+ \\ \left(U^T(y - Ax) - \xi e^h\right)^+ \\ \left(-U^T(y - Ax) - \xi e^h\right)^+ \\ \left(U^T(y - Ax) - v\right)^+ \\ \left(-U^T(y - Ax) - v\right)^+ \\ (w_1 + w_2 - e)^+ \\ \left(-a_1 w_3 + (e^N)^T w_4\right)^+ \\ \left(-a_2 w_3 + e^T(w_5 + w_6)\right)^+ \\ (-a_3 e w_3 + w_7 + w_8)^+ \\ \{w_1 - w_2 + A^T M_{P_0} w_4 + \\ A^T U(w_5 - w_6 + w_7 - w_8) - B^T w_9\} \\ (t^-, s^-, \xi^-, v^-, w^-) \end{bmatrix} \right\|_1 \tag{33}$$

where σ' is the Robinson constant determined by (M_1', M_2') given by (16). Since the vector (x, t, s, ξ, v, w) satisfies (24) and (32), the inequality (33) can be simplified to

$$\|u-\hat{u}\|_2 \leq \sigma(M_1', M_2')\{|\mathcal{Y}| + \left\|(Bx-b)^+\right\|_1 + |\mathcal{X}|\}. \tag{34}$$

In the reminder of the proof, we estimate the terms on the right-hand side of (34). Note that the vectors in T are unit vectors. It is easy to see that

$$\max_{1\leq i\leq N} \left|(M_{P_0})^T(Ax-y)\right|_i \leq \|y-Ax\|_2 .$$

The value of s in (23) implies that $s \leq \|y-Ax\|_2$. Therefore we have

$$\mathcal{Y} \leq \left(a_1 \|y-Ax\|_2 + a_2 \left\|U^T(y-Ax)\right\|_\infty + a_3 \left\|U^T(y-Ax)\right\|_1 - \varepsilon\right)^+ . \tag{35}$$

Due to (27), (29) and (30), we have

$$\begin{aligned} |\mathcal{X}| &= \left|e^T t + \varepsilon w_3 - y^T \nu^* - b^T h^*\right| \\ &= \left|e^T t + \varepsilon w_3 - x^T A^T \nu^* + (\phi(x))^T g + (Bx-b)^T h^* - x^T B^T h^*\right| . \end{aligned}$$

The fact $A^T\nu^* + B^T h^* = \eta$ (due to the restricted weak RSP of order k) and the triangle inequality imply that

$$|\mathcal{X}| \leq \left|e^T t - x^T \eta\right| + \varepsilon\,|w_3| + \left|(\phi(x))^T g\right| + \left|(Bx-b)^T h^*\right| . \tag{36}$$

Now we deal with the right-hand side of the above inequality. First, by using the index sets J and J_3, we have

$$\left|e^T t - x^T \eta\right| = \left|e_J^T t_J + e_{J_3}^T t_{J_3} - x_J^T \eta_J - x_{J_3}^T \eta_{J_3}\right| .$$

It follows from $t = |x|$ and (25) that

$$\begin{aligned} \left|e_J^T t_J + e_{J_3}^T t_{J_3} - x_J^T \eta_J - x_{J_3}^T \eta_{J_3}\right| &= \left|e_{J_3}^T t_{J_3} - x_{J_3}^T \eta_{J_3}\right| \leq \left|e_{J_3}^T t_{J_3}\right| + \left|x_{J_3}^T \eta_{J_3}\right| \\ &\leq \|x_{J_3}\|_1 + \left|x_{J_3}^T\right| |\eta_{J_3}| \leq \|x_{J_3}\|_1 + \left|x_{J_3}^T\right| e \\ &= 2\,\|x_{J_3}\|_1 . \end{aligned}$$

Then we obtain

$$\left|e^T t - x^T \eta\right| \leq 2\,\|x_{J_3}\|_1 = 2\sigma_k(x)_1. \tag{37}$$

By using the restricted weak RSP of order k, we have

$$\|\nu^*\|_1 \leq \left\|\begin{bmatrix} \nu^* \\ h^* \end{bmatrix}\right\|_1 \leq \left\|\left(CC^T\right)^{-1} C\eta\right\|_1 \leq \left\|\left(CC^T\right)^{-1} C\right\|_{\infty\to 1} \|\eta\|_\infty \leq \vartheta(1),$$

where $C = \left[A^T, B^T\right]^T \in R^{(m+l)\times n}$ and $\vartheta(1)$ is defined in (19b). Moreover, we have

$$\|g\|_1 = \|g_\mho\|_1 = \left\|U_\mho^{-1}\nu^*\right\|_1 \leq \left\|U_\mho^{-1}\right\|_{1\to 1} \|\nu^*\|_1 \leq \left\|U_\mho^{-1}\right\|_{1\to 1} \vartheta(1)$$

Recall that $\Upsilon(1,1)$ is determined in (19a). Then $\|g\|_1 \leq \Upsilon(1,1)$. Similarly, $\|g\|_\infty \leq \Upsilon(\infty,\infty)$ can be obtained. Due to $w_3 = \max\{\|\nu^*\|_1, \|g\|_1, \|g\|_\infty\}$, we have

$$\varepsilon |w_3| \leq \varepsilon \max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}. \tag{38}$$

Let $c, d, \widehat{d} \in [1, +\infty]$ be three given positive numbers and d, d' be two given numbers satisfying (21). For the term $\left|(\phi(x))^T g\right|$ in (36), it follows from Hölder inequalities that

$$\begin{aligned} \left|(\phi(x))^T g\right| &\leq \|\phi(x)\|_{d'} \|g\|_d \\ &= \|\phi(x)\|_{d'} \left\|U_\mho^{-1}\nu^*\right\|_d \\ &\leq \|\phi(x)\|_{d'} \left\|U_\mho^{-1}\right\|_{\widehat{d}\to d} \|\nu^*\|_{\widehat{d}} \\ &\leq \|\phi(x)\|_{d'} \left\|U_\mho^{-1}\right\|_{\widehat{d}\to d} \left\|\left(CC^T\right)^{-1} C\right\|_{\infty\to\widehat{d}}. \end{aligned} \tag{39}$$

Let $\Upsilon(d, \widehat{d})$ be given as (19a), i.e.,

$$\Upsilon(d, \widehat{d}) = \max_{\mho\subseteq\{1,\ldots,h\}, |\mho|=m} \left\|U_\mho^{-1}\right\|_{\widehat{d}\to d} \left\|\left(CC^T\right)^{-1} C\right\|_{\infty\to\widehat{d}}.$$

Thus we have

$$\left|(\phi(x))^T g\right| \leq \Upsilon(d, \widehat{d}) \|\phi(x)\|_{d'}. \tag{40}$$

Similarly, the following inequalities holds

$$\begin{aligned} \left|(Bx-b)^T h^*\right| &\leq \|Bx-b\|_{c'} \|h^*\|_c \\ &\leq \|Bx-b\|_{c'} \left\|(CC^T)^{-1}C\right\|_{\infty\to c} \|\eta\|_\infty \\ &\leq \|Bx-b\|_{c'} \left\|(CC^T)^{-1}C\right\|_{\infty\to c} \\ &= \vartheta(c) \|Bx-b\|_{c'}. \end{aligned} \tag{41}$$

Due to (37), (38), (40) and (41), the inequality (36) is reduced to

$$|\mathcal{X}| \leq \varepsilon\hat{\Upsilon} + 2\sigma_k(x)_1 + \|\phi(x)\|_{d'} \Upsilon(d, \widehat{d}) + \|Bx-b\|_{c'} \vartheta(c), \tag{42}$$

where $\hat{\Upsilon} = \max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}$.

Note that $\|x-\hat{x}\|_2 \leq \|u-\hat{u}\|_2$. It follows from (34), (35) and (42) that

$$\begin{aligned} \|x-\hat{x}\|_2 \leq \sigma' \Big\{ & 2\sigma_k(x)_1 + \|(Bx-b)^+\|_1 + \varepsilon\hat{\Upsilon} + \|\phi(x)\|_{d'} \Upsilon(d,\widehat{d}) + \\ & (a_1 \|y-Ax\|_2 + a_2 \|\phi(x)\|_\infty + a_3 \|\phi(x)\|_1 - \varepsilon)^+ + \|Bx-b\|_{c'} \vartheta(c) \Big\}. \end{aligned} \tag{43}$$

We recall the three sets Ω^*, Ω_{P_0} and $\Omega^*_{P_0}$, where Ω^* and $\Omega^*_{P_0}$ are the solution sets of (2) and (10), given as (6) and (11), respectively, and Ω_{P_0} is given as (7) with $P = P_0$. Clearly, $\hat{x} \in \Omega^*_{P_0}$. Let x^* denote the projection of x onto Ω^*, that is,

$$x^* = \pi_{\Omega^*}(x).$$

Note that the three sets are compact convex sets satisfying $\Omega^* \subseteq \Omega_{P_0}$ and $\Omega^*_{P_0} \subseteq \Omega_{P_0}$.

Then by applying Lemma 4.3 with $T_1 = \Omega^*$, $T_2 = \Omega_{P_0}$ and $T_3 = \Omega^*_{P_0}$, we have

$$\|x - \pi_{\Omega^*}(x)\|_2 = \|x - x^*\|_2 \leq \delta^{\mathcal{H}}(\Omega^*, \Omega_{P_0}) + 2\|x - \hat{x}\|_2 .$$

Since P_0 satisfies (8), it implies that

$$\|x - x^*\|_2 \leq \varepsilon' + 2\|x - \hat{x}\|_2 .$$

Let $\hat{\Upsilon} = \max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}$. Combination of the above inequality and (43) yields the desired results (20). If x is the feasible solution of (2), then $\|(Bx - b)^+\|_1 = 0$ and

$$(a_1\|y - Ax\|_2 + a_2\|\phi(x)\|_\infty + a_3\|\phi(x)\|_1 - \varepsilon)^+ = 0,$$

and thus the desired error bound (22) is also obtained. □

Based on Theorem 4.4, the error bound for the solutions of (1) and (2) can be stated as follows.

Corollary 4.5. *For any optimal solution x of* (1)*, there is an optimal solution x^* of* (2) *estimating x with the error:*

$$\|x - x^*\|_2 \leq \varepsilon' + 2\sigma'\left\{\varepsilon\hat{\Upsilon} + \sigma_k(x)_1 + \|\phi(x)\|_{d'}\Upsilon(d, \widehat{d}) + \|Bx - b\|_{c'}\vartheta(c)\right\},$$

where the constants ε', $\hat{\Upsilon}$, σ', $\Upsilon(d, \widehat{d})$ and $\vartheta(c)$ are given as in Theorem 4.4.

5. Special cases

Firstly, by setting different values of a_1, a_2 and a_3, the problem (2) can reduce to several special cases, and the corresponding stability results for these special cases can be obtained from (20) and (22) immediately. Note that if any of a_1, a_2 and a_3 is zero, the constant $\hat{\Upsilon} = \max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}$ in (20) and (22) will be simplified as well. For example, if $a_1 = 0$, the constant $\hat{\Upsilon}$ is reduced to $\max\{\Upsilon(1,1), \Upsilon(\infty,\infty)\}$. The following table shows the form of the constant $\hat{\Upsilon}$ for different choices of a_1, a_2 and a_3. Note that for any case with $a_1 = 0$, we have $\Omega^* = \Omega_{P_0} = \Omega^*_{P_0}$ so that $\hat{x} = x^*$ where

Table 1. The constant $\hat{\Upsilon}$.

a_i	$\hat{\Upsilon}$
$a_1 + a_2 = 0$	$\Upsilon(\infty,\infty)$
$a_1 + a_3 = 0$	$\Upsilon(1,1)$
$a_2 + a_3 = 0$	$\vartheta(1)$
$a_1 = 0$	$\max\{\Upsilon(1,1), \Upsilon(\infty,\infty)\}$
$a_2 = 0$	$\max\{\Upsilon(\infty,\infty), \vartheta(1)\}$
$a_3 = 0$	$\max\{\Upsilon(1,1), \vartheta(1)\}$
$a_1, a_2, a_3 \neq 0$	$\max\{\Upsilon(1,1), \Upsilon(\infty,\infty), \vartheta(1)\}$

$\hat{x} \in \Omega^*_{P_0}$ and $x^* \in \Omega^*$. Thus instead of using Lemma 4.3, the stability results can be immediately obtained from (43).

Secondly, without matrix B, the problem (2) is reduced to

$$\begin{array}{ll} \min\limits_{x\in R^n} & \|x\|_1 \\ \text{s.t.} & a_1\|y-Ax\|_2 + a_2\|\phi(x)\|_\infty + a_3\|\phi(x)\|_1 \le \varepsilon. \end{array}$$

In this case, the restricted weak RSP of order k is reduced to the standard weak RSP of order k, which means $A^T\nu^* = \eta$. In fact, the upper bound of $\left|(\phi(x))^T g\right|$ in (39) can be improved to

$$\begin{array}{rcl} \left|(\phi(x))^T g\right| & \le & \|\phi(x)\|_{d'}\|g\|_d \\ & = & \|\phi(x)\|_{d'}\left\|U_\mho^{-1}\nu^*\right\|_d \\ & \le & \|\phi(x)\|_{d'}\left\|U_\mho^{-1}(AA^T)^{-1}A\eta\right\|_d \\ & \le & \|\phi(x)\|_{d'}\left\|U_\mho^{-1}(AA^T)^{-1}A\right\|_{\infty\to d}. \end{array}$$

Then in order to obtain a tighter bound, $\Upsilon(d,\widehat{d})$ can be replaced by

$$\Upsilon'(d) = \max_{\mho\subseteq\{1,\ldots,h\},|\mho|=m}\left\|U_\mho^{-1}\left(AA^T\right)^{-1}A\right\|_{\infty\to d}.$$

Thus we have $\left|(\phi(x))^T g\right| \le \|\phi(x)\|_{d'}\Upsilon'(d)$. Similarly, the constants $\Upsilon(1,1)$ and $\Upsilon(\infty,\infty)$ are replaced by $\Upsilon'(1)$ and $\Upsilon'(\infty)$, respectively. Clearly, in this case, $\vartheta(c) = \left\|(AA^T)^{-1}A\right\|_{\infty\to c}$. Let $\hat{\Upsilon}' = \max\{\Upsilon'(1),\Upsilon'(\infty),\vartheta(1)\}$. Then the bound (22) is reduced to

$$\|x-x^*\|_2 \le \varepsilon' + 2\sigma'\left\{\varepsilon\hat{\Upsilon}' + 2\sigma_k(x)_1 + \|\phi(x)\|_{d'}\Upsilon'(d)\right\}.$$

Similarly, we list the constants $\hat{\Upsilon}'$ for different choices of a_i, $i = 1,2,3$ in the following table. Note that when $a_1 = 0$, we have $\hat{\Upsilon}' = \Upsilon'(1)$ due to the fact

Table 2. The constant $\hat{\Upsilon}'$.

a_i	$\hat{\Upsilon}'$
$a_1+a_2=0$	$\Upsilon'(\infty)$
$a_1+a_3=0$	$\Upsilon'(1)$
$a_2+a_3=0$	$\vartheta(1)$
$a_1=0$	$\max\{\Upsilon'(1),\Upsilon'(\infty)\}$
$a_2=0$	$\max\{\Upsilon'(\infty),\vartheta(1)\}$
$a_3=0$	$\max\{\Upsilon'(1),\vartheta(1)\}$
$a_1,a_2,a_3\ne 0$	$\max\{\Upsilon'(1),\Upsilon'(\infty),\vartheta(1)\}$

$\left\|U_\mho^{-1}\left(AA^T\right)^{-1}A\right\|_{\infty\to 1} \ge \left\|U_\mho^{-1}\left(AA^T\right)^{-1}A\right\|_{\infty\to\infty}$. Moreover, in this case, setting $d=1$ yields

$$\|x-x^*\|_2 \le \sigma'\left\{\varepsilon\Upsilon'(1) + 2\sigma_k(x)_1 + \|\phi(x)\|_\infty\Upsilon'(1)\right\},$$

which is the bound for the following ℓ_1-minimization established by Zhao and Li [30] (see also in Zhao [28]):

$$\min\{\|x\|_1 :\ a_2\|\phi(x)\|_\infty + a_3\|\phi(x)\|_1 \le \varepsilon\}.$$

Last but not least, our analysis can also apply to 1-bit basis pursuit [31], which can be viewed as a special case of our model (2). The stability result for the 1-bit basis pursuit in [31] can be obtained immediately from Theorem 4.4 by setting $a_2 = a_3 = 0$.

6. Conclusion

In this paper, we have studied the stability issue of the ℓ_1-minimization method (2). To establish our results, we introduced the restricted weak RSP of order k which is a mild assumption governing the stability of sparsity-seeking algorithms. Under this assumption, we use the classic Hoffman theorem and Lemma 4.3 to show that the ℓ_1-minimization method (2) is stable and thus the error between the solutions of the problems (1) and (2) can be measured in terms of the best k term approximation and the problem data (see Theorem 4.4). The result developed in this paper can apply to a range of problems with constraints defined by ℓ_1-, ℓ_2-, and ℓ_∞-norms.

References

[1] J. Andersson and J. O. Strömberg, *On the theorem of uinform recoery of structured random matrices*, IEEE Trans. Inform. Theory 60 (2014), pp. 1700–1710.

[2] P. T. Boufounos, *Greedy sparse signal reconstruction from sign measurements*, Porc. 43rd Asilomar Conf. Signals, Systems and Computers, 2009, pp. 1305–1309.

[3] J. Cahill, X. Chen, and R. Wang, *The gap between the null space property and restricted isometry property*, Linear Algebra Appl. 501 (2016), pp. 363–375.

[4] T. Cai, L. Wang, and G. Xu, *New bounds for restricted isometry constants*, IEEE Transactions on Information Theory. 56 (2010), pp. 4388–4394.

[5] T. Cai and A. Zhang, *Sharp RIP bound for sparse signal and low-rank matrix recovery*, Applied and Computational Harmonic Analysis. 35 (2013), pp. 74–93.

[6] E. Candès, *Compressive sampling*, Proceedings of the International Congress of Mathematicians. 3 (2006), pp. 1433–1452.

[7] E. Candès, J. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Communications on Pure and Applied Mathematics 59 (2006), pp. 1207–1223.

[8] E. Candès and T. Tao, *Decoding by linear programming*, IEEE Transactions on Information Theory 51 (2005), pp. 4203–4215.

[9] E. Candès and T. Tao, *The Dantzig selector: Statistical estimation when p is much larger than n*, The Annals of Statistics 35 (2007), pp. 2313–2351.

[10] D. Chen and R. Plemmons, *Nonnegativity constraints in numerical analysis*, Symposium on the Birth of Numerical Analysis, 2007.

[11] S. Chen, D. Donoho, and M. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput 20 (1998), pp. 33–61.

[12] A. Cohen, W. Dahmen and R. DeVore, *Compressed sensing and best k-term approximation*, Journal of the American Mathematical Society 22 (2009), pp. 211–231.

[13] D. Donoho, *Compressed sensing*, IEEE Transactions on Information Theory. 52 (2006), pp. 1289–1306.

[14] D. Donoho and M. Elad, *Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization*, Proceedings of the National Academy of Sciences. 100 (2003), pp. 2197–2202.

[15] R. Dudley, *Metric entropy of some classes of sets with differentiable boundaries*, Journal of Approximation Theory 10 (1974), pp. 227–236.

[16] Y. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, Cambridge University Press, 2012.

[17] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer, NY, 2013.

[18] A. J. Hoffman, *On approximate solutions of systems of linear inequalities*, J. Res. Nat. Bur. Standards 49 (1952), pp. 263–265.

[19] X. Huo and D. Donoho, *Uncertainty principles and ideal atomic decomposition*, IEEE Transactions on Information Theory 47 (2001), pp. 2845–2862.

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

[21] S. M. Robinson, *Bounds for error in the solution set of a perturbed linear program*, Linear Algebra and Its Applications 6 (1973), pp. 69–81.

[22] G. Tang and A. Nehorai, *Performance analysis of sparse recovery based on constrained minimal singular values*, IEEE Transactions on Signal Processing 59 (2011), pp. 5734–5745.

[23] R. Tibshirani, M. Wainwright, and T. Hastie, *Statistical Learning with Sparsity: The Lasso and Generalizations*, Chapman and Hall/CRC, 2015.

[24] R. Tibshirani$_2$, H. Hoefling, and R. Tibshirani, *Nearly-isotonic regression*, Technometrics 53 (2011), pp. 54–61.

[25] Y. Zhang, *A simple proof for recoverability of ℓ_1-minimization (ii): the nonnegative case*, Technical Report, Rice University, 2005.

[26] Y. B. Zhao, *RSP-based analysis for sparsest and least ℓ_1-norm solutions to underdetermined linear systems*, IEEE Transactions on Signal Processing 61 (2013), pp. 5777–5788.

[27] Y. B. Zhao, *Equivalence and strong equivalence between the sparsest and least ℓ_1-norm nonnegative solutions of linear systems and their applications*, Journal of the Operations Research Society of China 2 (2014), pp. 171–193.

[28] Y. B. Zhao, *Sparse Optimization Theory and Methods*, CRC Press, Taylor & Francis Group, 2018.

[29] Y. B. Zhao, H. Jiang, and Z. Q. Luo, *Weak stability of ℓ_1-minimization methods in sparse data reconstruction*, Mathematics of Operations Reseach 44 (2019), pp. 173–195.

[30] Y. B. Zhao and D. Li, *A theoretical analysis of sparse recovery stability of Dantzig selector and Lasso*. Available at arXiv: 1711.03783, 2017.

[31] Y. B. Zhao and C. Xu, *1-Bit compressive sensing: Reformulation and RRSP-based sign recovery theory*, Science China Mathematics 59 (2016), pp. 2049–2074.