Appl. Comput. Harmon. Anal. 33 (2012) 282–291

Contents lists available at SciVerse ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Perturbations of measurement matrices and dictionaries in compressed sensing

Akram Aldroubi [*,1], Xuemei Chen, Alexander M. Powell [2]

Department of Mathematics, Vanderbilt University, Nashville, TN 37240, United States

ARTICLE INFO

Article history:
Received 31 January 2011
Revised 19 November 2011
Accepted 12 December 2011
Available online 20 December 2011
Communicated by Albert Cohen

Keywords:
Compressed sensing
Sparse representation
Stability
Redundant dictionary

ABSTRACT

The compressed sensing problem for redundant dictionaries aims to use a small number of linear measurements to represent signals that are sparse with respect to a general dictionary. Under an appropriate restricted isometry property for a dictionary, reconstruction methods based on ℓ^q minimization are known to provide an effective signal recovery tool in this setting. This note explores conditions under which ℓ^q minimization is robust to measurement noise, and stable with respect to perturbations of the sensing matrix A and the dictionary D. We propose a new condition, the D null space property, which guarantees that ℓ^q minimization produces solutions that are robust and stable against perturbations of A and D. We also show that ℓ^q minimization is jointly stable with respect to imprecise knowledge of the measurement matrix A and the dictionary D when A satisfies the restricted isometry property.

1. Introduction

Compressed sensing addresses the problem of recovering an unknown signal $z_0 \in \mathbb{R}^d$ from a small number of linear measurements based on an underlying structure of sparsity or compressibility. A vector $z_0 \in \mathbb{R}^d$ is said to be *s-sparse* if it has at most s nonzero entries, and the class of all s-sparse vectors in $\mathbb{R}^d$ is denoted by Σ_s^d. A primary problem of interest is to represent signals $z_0 \in \Sigma_s^d$ using nonadaptive linear measurements $y = Az_0 \in \mathbb{R}^m$ where the number of measurements m is typically less than the ambient dimension d, i.e., $m \ll d$. The $m \times d$ matrix A is referred to as a *measurement matrix* or *sensing matrix*.

Suppose $z_0 \in \mathbb{R}^d$ is unknown and that one is given the noisy measurements $y = Az_0 + e \in \mathbb{R}^m$ with noise level $\|e\|_2 \leqslant \epsilon$. If z_0 is approximately sparse then ℓ^q minimization is known to provide a method for recovering z_0 from y in a robust and stable way as described below. For fixed $0 < q \leqslant 1$, the following minimization problem reconstructs an approximation $\widetilde{z}$ to z_0 from the measurements y:

$$\widetilde{z} = \arg\min \|z\|_q^q \quad \text{subject to } \|Az - y\|_2 \leqslant \epsilon. \tag{1.1}$$

If A is appropriately chosen and z_0 is approximately sparse then minimizers $\widetilde{z}$ are known to be good approximations to z. Moreover under certain conditions on A, the program (1.1) is known to be stable with respect to noise in the measurement vector [1–7].

* Corresponding author.
E-mail addresses: akram.aldroubi@vanderbilt.edu (A. Aldroubi), xuemei.chen@vanderbilt.edu (X. Chen), alexander.m.powell@vanderbilt.edu (A.M. Powell).

1 A. Aldroubi was supported in part by NSF DMS Grant 0807464 and DMS 1108631.
2 A.M. Powell was supported in part by NSF DMS Grant 0811086.

1063-5203/$ – see front matter
doi:10.1016/j.acha.2011.12.002

The matrix A satisfies the *restricted isometry property* (*RIP*) of order s if there exists a positive constant $0 < \delta < 1$ such that

$$\forall x \in \Sigma_s^d, \quad (1-\delta)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1+\delta)\|x\|_2^2. \tag{1.2}$$

The smallest constant $\delta > 0$ that satisfies (1.2) is denoted by $\delta_s = \delta_s(A)$ and is known as the *restricted isometry constant* (RIC). It has been shown in [1,2,4,8–10] that if A satisfies RIP then (1.1) provides robust and stable recovery in the following sense: for every $\epsilon \geqslant 0$ and every $z_0 \in \mathbb{R}^d$, the reconstructed signal $\widetilde{z}$ satisfies

$$\|\widetilde{z} - z_0\|_2 \leqslant C_1\epsilon + C_2\sigma_s(z_0)_q, \tag{1.3}$$

where $\sigma_s(z_0)_q = \min_{z\in\Sigma_s^d} \|z - z_0\|_q$ is the best s-term ℓ^q approximation error for z_0. Here *stability* refers to the first term on the right-hand side of the inequality while *robustness* refers to the second term. The constants C_1, C_2 may depend on the sparsity s and the measurement matrix A. In particular, (1.3) shows that if $x \in \Sigma_s^d$ is s-sparse and there is no noise $\epsilon = 0$ then (1.1) is able to exactly recover z_0 from y.

The bound (1.3) addresses stability of the ℓ^q problem (1.1) when additive measurement noise is present and when z_0 is only approximately sparse. Other works, such as [11,12], provide an extension of this to deal with imprecise knowledge of the measurement matrix A by considering a perturbed sensing matrix of the form of $A = B + E$. It was shown in [11] that if A satisfies a certain restricted isometry property, then taking $q = 1$ and letting ϵ be a combined error to account for both measurements noise e and matrix perturbation E allows (1.1) to stably recover approximately sparse signals:

$$\|\widetilde{z} - z_0\|_2 \leqslant C_1\epsilon + C_2\sigma_s(z_0)_1. \tag{1.4}$$

A recent direction of interest in compressed sensing concerns problems where signals are sparse in an overcomplete dictionary D instead of a basis (e.g., the canonical basis associated to Σ_s^d), see [13–15]. This is motivated by the widespread use of overcomplete dictionaries in signal processing and data analysis. In this setting the signal $z_0 \in \mathbb{R}^d$ can be represented as $z_0 = Dx_0$, where $x_0 \in \Sigma_s^n$ and D is a $d \times n$ matrix with $n \geqslant d$. The columns of D may be thought of as an overcomplete frame or dictionary for $\mathbb{R}^d$. The work in [13] shows that a modified ℓ^q minimization approach can be used for sparse signal recovery in the setting of redundant dictionaries. If $z_0 \in \mathbb{R}^d$ is unknown and one is given the noisy measurements $y = Az_0 + e \in \mathbb{R}^m$ with noise level $\|e\|_2 \leqslant \epsilon$, one reconstructs $\widetilde{z} \in \mathbb{R}^d$ by

$$\widetilde{z} = \arg\min \left\| D^* z\right\|_1 \quad \text{subject to } \|Az - y\|_2 \leqslant \epsilon, \tag{1.5}$$

where D^* is the transpose of D. Error bounds for (1.5) are proven in [13] when D is a tight frame and it is shown that the reconstruction is robust and stable with respect to the measurement noise and non-sparsity of D^*z_0.

The null space property [2,3,16,17] is another well-known condition on measurement matrices. Specifically, a matrix A satisfies the ℓ^q *null space property* (NSP_q) of order s if

$$\forall v \in \ker A\backslash\{0\},\ \forall |T| \leqslant s, \quad \|v_T\|_q^q < \|v_{T^c}\|_q^q, \tag{1.6}$$

where $v_T \in \mathbb{R}^d$ denotes the vector whose entries are the same as $v \in \mathbb{R}^d$ on the index set T and equal to zero on the complementary index set T^c.

The importance of the null space property is that it is the necessary and sufficient condition under which ℓ^q recovery is exact for s-sparse signals, e.g., see [2]. Specifically, for a fixed $0 < q \leqslant 1$, suppose that A is an $m \times d$ matrix, $z_0 \in \Sigma_s^d$ and $y = Az_0$. Then z_0 is exactly recovered from y by

$$\widetilde{z} = \arg\min \|z\|_q^q \quad \text{subject to } Az = y \tag{1.7}$$

if and only if A satisfies NSP_q.

It is also shown in [2] that the null space property can stably recover approximately sparse signals when there is no measurement noise. A stronger condition called the *sparse approximation property* recently introduced by Sun [3] also provides stable and robust recovery. This sparse approximation property is weaker than RIP and stronger than NSP_q.

We shall restrict this work to the setting of real valued signals $z \in \mathbb{R}^d$. For perspective, it is known that compressed sensing results such as (1.3) are also valid for complex valued signals $z \in \mathbb{C}^d$, e.g., [18].

1.1. Overview and main results

This note will focus on ℓ^q recovery of signals that are sparse with respect to a general dictionary D. Our main goal is to investigate conditions under which the ℓ^q version, $0 < q \leqslant 1$, of problem (1.5) is robust and stable when there is measurement noise and imprecise knowledge of both the measurement matrix A and the dictionary D.

Properties of the measurement matrix A play an important role in our stability analysis. Our first result, Theorem 3.1, provides analysis under the null space property. As far as we know, there are no results on the robustness when measurement noise is present under the null space property assumption. Our second result, Theorem 3.4, requires that A satisfies a restricted isometry property as in [13]. Both of these ℓ^q stability results are valid for the full range of parameters $0 < q \leqslant 1$.

The effect of a perturbed measurement matrix A satisfying the restricted RIP has previously been considered in the classical case of sparsity with respect to a basis, see [11]. We provide an extension of this to the case of sparsity with respect to a redundant dictionary, see Theorem 3.4 when A satisfies D-RIP, and Theorem 3.1 when A satisfies a null space property only. We also investigate a second, not previously considered, type of stability to address imprecision in the dictionary D. Our performance analysis for the ℓ^q recovery method (1.5) will typically require that D is chosen to satisfy a design condition such as D-NSP$_q$ in (2.2). However, in practice it may only be possible to use a perturbed version of D for which there are no a priori guarantees that the desired design condition holds. For example, D may be viewed as a real reconstruction device which in practice will differ from its exact specifications. We prove that ℓ^q minimization is stable with respect to imprecisions in the dictionary D, see Theorems 3.1 and 3.4.

The main contribution of this note may thus be summarized as follows:

It is shown that ℓ^q recovery is robust to measurement noise and jointly stable with respect to compressible signals and imprecisions in the dictionary D and the measurement matrix A when A satisfies either the null space property or an appropriate restricted isometry property.

The precise statements of our main results are given in Theorems 3.1 and 3.4. The remainder of the paper is organized as follows. In Section 2 we present background on the restricted isometry property and the null space property. In Section 3 we state our main stability results. Proofs of the main theorems are given in Section 4.

2. Background

2.1. The restricted isometry property

The classical restricted isometry property was modified for the setting of sparsity in a dictionary in [13]. Let D be a given $d \times n$ matrix. The $m \times d$ matrix A satisfies the *restricted isometry property with respect to D* (D-RIP) of order s if there exists a constant $\delta > 0$ such that

$$\forall x \in \Sigma_s^n, \quad (1-\delta)\|Dx\|_2^2 \leqslant \|ADx\|_2^2 \leqslant (1+\delta)\|Dx\|_2^2. \tag{2.1}$$

The smallest value of $\delta > 0$ for which (2.1) holds is denoted by δ_s.

Similar to the standard restricted isometry property, random matrices provide examples that satisfy D-RIP, see [13].

2.2. The null space property

If M is an $n \times d$ matrix then M_T is the $n \times d$ matrix that satisfies $M_T v = (Mv)_T$ for all $v \in \mathbb{R}^d$, i.e., M_T is obtained by replacing the rows of M corresponding to T^c by zero rows.

We introduce the following modified null space property to address sparsity with respect to redundant dictionaries. Let D be a given $d \times n$ dictionary matrix. The matrix A satisfies the *ℓ^q null space property of order s relative to D* (D-NSP$_q$) if

$$\forall z \in \ker A \backslash \{0\}, \ \forall |T| \leqslant s, \quad \left\| D_T^* z \right\|_q^q < \left\| D_{T^c}^* z \right\|_q^q. \tag{2.2}$$

Here $D_T^* = (D^*)_T$. A simple compactness argument, e.g., see [2], shows that D-NSP$_q$ is equivalent to the existence of a constant c, $0 < c < 1$, such that

$$\forall z \in \ker A, \ \forall |T| \leqslant s, \quad \left\| D_T^* z \right\|_q^q \leqslant c \left\| D_{T^c}^* z \right\|_q^q. \tag{2.3}$$

The smallest value of the constant c in (2.3) is referred to as the *null space constant* (NSC).

3. Main theorems

In this section we describe our main stability theorems for ℓ^q recovery of signals that are sparse in a dictionary. We initially assume the following set-up:

- D is a $d \times n$ dictionary matrix for $\mathbb{R}^d$ (thus $n \geqslant d$),
- B is an $m \times d$ measurement matrix for $\mathbb{R}^d$,
- $D^* z_0$ is approximately s-sparse.

The assumption that $D^* z_0$ is approximately sparse is justified in applications since practical signal classes often have sparse frame coefficients, for example, with respect to wavelets, curvelets, edgelets, shearlets, [19–21].

At this point, one is given the noisy measurements $y = Bz_0 + e \in \mathbb{R}^m$ with noise level $\|e\|_2 \leqslant \epsilon$, and one wishes to recover z_0 from y. We assume that one only has approximate knowledge of B, for example, due to a nonideal measurement device or because of computational limitations. We also assume perturbations of the dictionary D. For example, the intended

D in (1.5) might have been carefully designed to satisfy a hypothesis such as D-NSP, but computational necessities, or quantization errors, could result in the use of a perturbed $\widetilde{D}$ as in the ℓ^q minimization in (3.1) below. So, we further assume that:

- $\widetilde{D}$ is a $d \times n$ dictionary (perturbation of the intended dictionary D),
- A is an $m \times d$ full rank measurement matrix (our knowledge of the true matrix B).

The full rank condition is justified when redundant measurements are excluded. For fixed $0 < q \leqslant 1$, the following ℓ^q minimization problem reconstructs the approximation $\widetilde{z}$ to z_0 based on the noisy measurements y and the perturbations $\widetilde{D}$ and A of D and B, respectively

$$\widetilde{z} = \arg\min \|\widetilde{D}^* z\|_q^q \quad \text{subject to } \|Az - y\|_2 \leqslant \epsilon. \tag{3.1}$$

The matrix A will satisfy hypotheses such as D-NSP$_q$ or D-RIP, but the perturbed matrix $\widetilde{D}$ used in (3.1) introduces uncertainty and distortion into these hypotheses.

Our first theorem, Theorem 3.1, provides stability analysis under the assumption of the null space property. For this result, we assume that the dictionary D satisfies the frame inequality

$$\forall z \in \mathbb{R}^d, \quad \alpha\|z\|_2 \leqslant \|D^* z\|_2 \leqslant \beta\|z\|_2,$$

with frame constants $\beta \geqslant \alpha > 0$. If M is an $n \times d$ matrix then $\|M\|_{op}$ denotes the operator norm of M as a mapping from $(\mathbb{R}^d, \|\cdot\|_2)$ to $(\mathbb{R}^n, \|\cdot\|_2)$.

Theorem 3.1. *Suppose that D is a $d \times n$ dictionary with frame constants $\beta \geqslant \alpha > 0$ and suppose that the $m \times d$ matrix A satisfies* D-NSP$_q$ *with null space constant c. Moreover, suppose that the $d \times n$ matrix $\widetilde{D}$ satisfies $\|D^* - \widetilde{D}^*\|_{op} \leqslant \frac{5\alpha}{2^{1/q} n^{1/q-1/2}} (\frac{1-c}{10})^{1/q}$ and that B is an $m \times d$ measurement matrix.*

If $z_0 \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$ satisfy $\|y - Bz_0\|_2 \leqslant \epsilon$ then any solution $\widetilde{z}$ to (3.1) *satisfies*

$$\begin{aligned}\|\widetilde{z} - z_0\|_2 \leqslant{}& \frac{2\beta}{5\nu_A} C_1 n^{1/q-1/2}\epsilon + 2^{1/q} C_1 \sigma_s(D^* z_0)_q + 2^{1/q} C_1 n^{1/q-1/2} \|D^* - \tilde{D}^*\|_{op} \|z_0\|_2 \\ &+ \frac{C_1}{\nu_A} n^{1/q-1/2} \big(\beta(1 + 2^{1/q}) + 2^{1/q}\|D^* - \tilde{D}^*\|_{op}\big)\|A - B\|_{op}\|z_0\|_2. \end{aligned} \tag{3.2}$$

Here ν_A is the smallest positive singular value of A. The constant C_1 is quantified in (4.8) *and* (4.13).

When D is the canonical basis, D-NSP$_q$ becomes the standard NSP$_q$ and as mentioned in the introduction, NSP$_q$ characterizes the exact recovery of any sparse signal from its noiseless observation y and sensing matrix A via ℓ^q minimization [2]. Thus we get the following corollary.

Corollary 3.2. *Let D be the canonical basis for $\mathbb{R}^d$ given by the $d \times d$ identity matrix $D = I$. A has* NSP$_q$ *of order s is a necessary and sufficient condition to robustly and stably recover any approximately sparse signal with respect to perturbations on the measurement vector and the sensing matrix using program* (1.1)*, i.e. given any vector z_0 in $\mathbb{R}^d$ and the measurement vector y such that $\|Az_0 - y\|_2 \leqslant \epsilon$, we have*

$$\|\widetilde{z} - z_0\|_q \leqslant C_2\epsilon + C_3\sigma_s(z_0)_q + C_4\|A - B\|_{op},$$

where $\widetilde{z}$ is any minimizer of (1.1).

Remark 3.3. It is known that NSP$_q$ is a necessary condition on A for exact recovery for s-sparse signals using ℓ^q minimization. Surprisingly, as stated in the above corollary, this necessary condition is also sufficient for robustness and stability via ℓ^q minimization. Specifically, the ℓ^q minimization (1.1) recovers the s-sparse signal in the ideal case (i.e., $\epsilon = 0$) if and only if ℓ^q minimization is also stable and robust (i.e., $\|\widetilde{z} - z_0\|_2 \leqslant C_1\epsilon + C_2\sigma_s(z_0)_q$). Notice that the last equivalence is interesting since there is no mention of the NSP$_q$ property.

For direct comparison with [13], the next result assumes that the dictionary D satisfies the *Parseval frame* condition $DD^* = I$, but as noted in [13] there are extensions to general frames.

Theorem 3.4. *Suppose that D is a $d \times n$ Parseval frame matrix and that the $d \times n$ matrix $\widetilde{D}$ satisfies $\|D^* - \widetilde{D}^*\|_{op} \leqslant \frac{1}{\sqrt{2}K_2}(\frac{n}{s})^{1/2-1/q}$ for some constant K_2. Suppose that A and B are $m \times d$ matrices and that A satisfies* D-RIP *with $\delta_{7s} < \frac{6-3(2/3)^{2/q-2}}{6-(2/3)^{2/q-2}}$.*

If $z_0 \in \mathbb{R}^d$ and $y \in \mathbb{R}^m$ satisfy $\|y - Bz_0\|_2 \leqslant \epsilon$ then any solution $\widetilde{z}$ to (3.1) *satisfies*:

$$\|\widetilde{z} - z_0\|_2 \leqslant C_5\epsilon + C_6 s^{1/2-1/q}\sigma_s\big(D^*z_0\big)_q + C_7\left(\frac{n}{s}\right)^{1/q-1/2}\big\|D^* - \widetilde{D}^*\big\|_{op}\|z_0\|_2$$
$$+\left(\frac{n}{s}\right)^{1/q-1/2}\frac{1}{\nu_A}\big(C_8 + C_9\big\|D^* - \tilde{D}^*\big\|_{op}\big)\|A - B\|_{op}\|z_0\|_2. \tag{3.3}$$

Here ν_A is the smallest positive singular value of A. Quantitative bounds on the constants C_5, C_6, C_7, C_8, C_9 and K_2 are contained in the proof, see (4.30), (4.31), (4.32).

It is possible to formulate Theorems 3.1 and 3.4 using different choices of norms. Except for the term $\sigma_s(D^*z_0)_q$, the bounds in (3.2) and (3.3) are stated using the ℓ^2 norm and the associated operator norm and hence incur the discouraging constants $n^{1/q-1/2}$. Note that if we use $\sigma_s(D^*z_0)_2$ instead of the standard $\sigma_s(D^*z_0)_q$, we would also incur the constant $n^{1/q-1/2}$ in front of this term as well. Furthermore, $n^{1/q-1/2}$ is multiplied by the factor $1/\nu_A$ in the 4th term on the right-hand side of (3.2) and (3.3) which is essentially $(\frac{m}{d})^{1/2}$. Indeed in the case where A is an $m \times d$ Gaussian random matrix with i.i.d. $\mathcal{N}(0, 1/m)$ entries, it is known that this choice of A satisfies D-RIP with high probability, see [13], when $m \gtrsim s\log(d/s)$. Moreover, the smallest singular value ν_A satisfies $\nu_A \gtrsim (\frac{d}{m})^{1/2}$ with high probability greater than $1 - 2e^{-d/8}$, e.g., see Corollary 35 in [22].

Remark 3.5. We conclude this section with the following remarks:

(i) In the noise free case $\epsilon = 0$, if A and D are exactly known (unperturbed), and D^*z_0 is exactly s-sparse, then $\widetilde{z}$ exactly reconstructs z_0, i.e., $\widetilde{z} = z_0$.
(ii) With no perturbations on the sensing matrix or the dictionary, and $q = 1$, we obtain the case studied in [13] and gain the same result. Furthermore, if D is the canonical basis, we obtain the now classical result (1.3).
(iii) When $D = I$ is the canonical basis and there are no perturbations of $D = I$, we obtain a result related to the one in [11].
(iv) If $D = I$, our proofs can be used to show that the set of sensing matrices satisfying the NSP_q is open in the operator norm topology [23]. Thus, if A satisfies NSP_q, then B satisfies NSP_q as well if $\|A - B\|_{op}$ is small. However, if A satisfies D-NSP_q, we do not know if A satisfies $\tilde{\mathrm{D}}$-NSP_q even if $\|\tilde{D} - D\|_{op}$ is small.

4. Proofs of the main theorems

Both proofs involve some properties of the ℓ^q quasinorm which must be recalled. Namely, for any vectors $u \in \mathbb{R}^N$

$$\|u\|_p \leqslant \|u\|_q \leqslant N^{1/q-1/p}\|u\|_p, \quad 0 < q \leqslant p \leqslant \infty. \tag{4.1}$$

The following lemma plays an important role in our proofs. The lemma follows from standard properties of the singular value decomposition.

Lemma 4.1. *Suppose A is an $m \times d$ matrix where $m \leqslant d$, then any vector $h \in \mathbb{R}^d$ can be decomposed as $h = a + \eta$ with $a \in \ker A$, $\eta \perp \ker A$, and $\|\eta\| \leqslant \frac{1}{\nu_A}\|Ah\|$, where ν_A is the smallest positive singular value of A.*

Proof of Theorem 3.1. Set $h = \tilde{z} - z_0$. There are two main inequalities. One obtained from the null space property. The other from the ℓ^q minimization which is essentially the reverse of the null space property. Combining these two, we obtain an upper bound on $\|D^*h\|_2$ in terms of the perturbations, and thus an upper bound for $\|h\|_2$ since D is a frame.

Step 1: Approximate D-NSP$_q$ for h. Note that h is expected to be almost in the null space of A. Thus we will decompose h as $h = a + \eta$ where $a \in \ker A$ and η small since, by Lemma 4.1, $\|\eta\|_2 \leqslant \frac{1}{\nu_A}\|Ah\|_2$.

Since $a \in \ker A$ and A has D-NSP_q, let T be any index set such that $|T| \leqslant s$,

$$\big\|D_T^*h\big\|_q^q \leqslant \big\|D_T^*a\big\|_q^q + \big\|D_T^*\eta\big\|_q^q \leqslant c\big\|D_{T^c}^*a\big\|_q^q + \big\|D_T^*\eta\big\|_q^q \leqslant c\big\|D_{T^c}^*h\big\|_q^q + \big\|D^*\eta\big\|_q^q.$$

Thus, we get the approximate D-NSP_q for h

$$\big\|D_T^*h\big\|_q^q \leqslant c\big\|D_{T^c}^*h\big\|_q^q + \big\|D^*\eta\big\|_q^q. \tag{4.2}$$

Step 2: An approximate reversed inequality for h from ℓ^q minimization. Since A is a perturbation of B, $\|y - Az_0\|_2$ is not necessarily less than ϵ, i.e., z_0 is not necessarily feasible for program (3.1). However, we can find a vector $z_0 + w$ that is very close to z_0 and is feasible. Specifically, since A is full rank by assumption, there exists w such that $Aw = (B - A)z_0$. Thus $\|A(z_0 + w) - y\|_2 = \|Bz_0 - y\|_2 \leqslant \epsilon$, and $z_0 + w$ is feasible in Program (3.1). Moreover, w is small since, by Lemma 4.1, we can pick w such that

$$\|w\|_2 \leqslant \frac{1}{\nu_A}\|Aw\|_2 = \frac{1}{\nu_A}\big\|(B-A)z_0\big\|_2. \tag{4.3}$$

Since $\tilde{z}$ minimizes (3.1) we have

$$\big\|\tilde{D}^*\tilde{z}\big\|_q^q \leqslant \big\|\tilde{D}^*(z_0+w)\big\|_q^q = \big\|\tilde{D}_T^* z_0 + \tilde{D}_T^* w\big\|_q^q + \big\|\tilde{D}_{T^c}^* z_0 + \tilde{D}_{T^c}^* w\big\|_q^q.$$

Moreover

$$\begin{aligned}\big\|\tilde{D}^*\tilde{z}\big\|_q^q &= \big\|\tilde{D}^*(h+z_0)\big\|_q^q = \big\|\tilde{D}_T^* h + \tilde{D}_T^* z_0\big\|_q^q + \big\|\tilde{D}_{T^c}^* h + \tilde{D}_{T^c}^* z_0\big\|_q^q \\ &\geqslant \big\|\tilde{D}_T^* z_0 + \tilde{D}_T^* w\big\|_q^q - \big\|\tilde{D}_T^* h - \tilde{D}_T^* w\big\|_q^q + \big\|\tilde{D}_{T^c}^* h\big\|_q^q - \big\|\tilde{D}_{T^c}^* z_0\big\|_q^q.\end{aligned}$$

Combining the above two inequalities we get

$$\big\|\tilde{D}_{T^c}^* h\big\|_q^q \leqslant \big\|\tilde{D}_T^* h\big\|_q^q + 2\big\|\tilde{D}_{T^c}^* z_0\big\|_q^q + \big\|\tilde{D}^* w\big\|_q^q. \tag{4.4}$$

Using the triangle inequality and (4.4) we obtain the desired inequality:

$$\big\|D_{T^c}^* h\big\|_q^q \leqslant \big\|D_T^* h\big\|_q^q + \big\|D^*h - \tilde{D}^*h\big\|_q^q + 2\big\|\tilde{D}_{T^c}^* z_0\big\|_q^q + \big\|\tilde{D}^* w\big\|_q^q. \tag{4.5}$$

Step 3: Estimation of $\|D^*h\|_q$. Our ultimate goal is to estimate $\|h\|_2$. However, this can be done by first estimating $\|D^*h\|_q$ and thereby $\|D^*h\|_2$ and hence $\|h\|_2$ since D is a frame, by assumption. We do this, by combining the two inequalities in Steps 1 and 2, we get

$$\big\|D_T^* h\big\|_q^q \leqslant \frac{c}{1-c}\big\|D^*h - \tilde{D}^*h\big\|_q^q + \frac{2c}{1-c}\big\|\tilde{D}_{T^c}^* z_0\big\|_q^q + \frac{c}{1-c}\big\|\tilde{D}^* w\big\|_q^q + \frac{1}{1-c}\big\|D^*\eta\big\|_q^q. \tag{4.6}$$

By (4.5) and (4.6) we obtain

$$\begin{aligned}\big\|D^*h\big\|_q^q &= \big\|D_T^* h\big\|_q^q + \big\|D_{T^c}^* h\big\|_q^q \leqslant 2\big\|D_T^* h\big\|_q^q + \big\|D^*h - \tilde{D}^*h\big\|_q^q + 2\big\|\tilde{D}_{T^c}^* z_0\big\|_q^q + \big\|\tilde{D}^* w\big\|_q^q \\ &\leqslant \frac{1+c}{1-c}\big\|D^*h - \tilde{D}^*h\big\|_q^q + \frac{2+2c}{1-c}\big\|\tilde{D}_{T^c}^* z_0\big\|_q^q + \frac{2}{1-c}\big\|D^*\eta\big\|_q^q + \frac{1+c}{1-c}\big\|\tilde{D}^* w\big\|_q^q.\end{aligned} \tag{4.7}$$

Step 4: Estimation of $\|h\|_2$. Rewriting the term $\tilde{D}_{T^c}^* z_0$ in (4.7) and using the fact that D is a frame and the inequality (4.1) we get

$$\begin{aligned}\|h\|_2 &\leqslant \frac{1}{\alpha}\big\|D^*h\big\|_2 \leqslant \frac{1}{\alpha}\big\|D^*h\big\|_q \leqslant Cn^{1/q-1/2}\big\|D^* - \tilde{D}^*\big\|_{op}\|h\|_2 + C\big\|D^*\eta\big\|_q \\ &\quad + 2^{1/q}C\big[\big\|\tilde{D}^*z_0 - D^*z_0\big\|_q + \big\|D_{T^c}^* z_0\big\|_q\big] + C\big\|\tilde{D}^* w\big\|_q,\end{aligned}$$

where

$$C = \frac{1}{5\alpha}\left(\frac{10}{1-c}\right)^{1/q}. \tag{4.8}$$

This leads to the estimation of $\|h\|_2$ in terms of the perturbations

$$(1-\rho)\|h\|_2 \leqslant C\big\|D^*\eta\big\|_q + 2^{1/q}C\big[\big\|\tilde{D}^*z_0 - D^*z_0\big\|_q + \big\|D_{T^c}^* z_0\big\|_q\big] + C\big\|\tilde{D}^* w\big\|_q, \tag{4.9}$$

where $\rho := 2^{1/q}Cn^{1/q-1/2}\|D^* - \tilde{D}^*\|_{op}$.

Step 5: Estimation of the perturbations. (1) *Estimation of* $\|D^*\eta\|_q$. Using the fact that $\|\eta\|_2 \leqslant \frac{1}{\nu_A}\|Ah\|_2$, and

$$\|Ah\|_2 = \|A\tilde{z} - Az_0\|_2 \leqslant \|A\tilde{z} - y\|_2 + \|y - Bz_0\|_2 + \|Bz_0 - Az_0\|_2 \leqslant 2\epsilon + \big\|(A-B)z_0\big\|_2,$$

we get

$$\big\|D^*\eta\big\|_q \leqslant n^{1/q-1/2}\big\|D^*\eta\big\|_2 \leqslant n^{1/q-1/2}\beta\|\eta\|_2 \leqslant n^{1/q-1/2}\frac{\beta}{\nu_A}\big(2\epsilon + \big\|(A-B)z_0\big\|_2\big). \tag{4.10}$$

(2) *Estimation of* $\|\tilde{D}^*w\|_q$. Using the upper frame bound β of D we get

$$\begin{aligned}\big\|\tilde{D}^*w\big\|_q^q &\leqslant \big\|\tilde{D}^*w - D^*w\big\|_q^q + \big\|D^*w\big\|_q^q \leqslant \big(n^{1/q-1/2}\big\|\tilde{D}^*w - D^*w\big\|_2\big)^q + \big(n^{1/q-1/2}\big\|D^*w\big\|_2\big)^q \\ &\leqslant n^{1-q/2}\|w\|_2^q\big(\big\|\tilde{D}^* - D^*\big\|_{op}^q + \beta^q\big),\end{aligned}$$

from which we get (using (4.3))

$$\|\tilde{D}^* w\|_q \leqslant \frac{(2n)^{1/q-1/2}}{\nu_A}\left(\|\tilde{D}^* - D^*\|_{op} + \beta\right)\|(B-A)z_0\|_2. \tag{4.11}$$

Step 6: Final estimate of $\|h\|_2$. Substitute (4.10) and (4.11) into (4.9) and letting T be the index set corresponding to the s largest magnitude entries of $D^* z_0$, we get

$$\begin{aligned}\|h\|_2 \leqslant{}& \frac{2\beta}{5\nu_A} C_1 n^{1/q-1/2}\epsilon + 2^{1/q} C_1 \sigma_s\left(D^* z_0\right)_q + 2^{1/q} C_1 n^{1/q-1/2}\|D^* - \tilde{D}^*\|_{op}\|z_0\|_2 \\ &+ \frac{C_1}{\nu_A} n^{1/q-1/2}\left(\beta\left(1+2^{1/q}\right) + 2^{1/q}\|\tilde{D}^* - D^*\|_{op}\right)\|A-B\|_{op}\|z_0\|_2,\end{aligned} \tag{4.12}$$

where

$$C_1 = \frac{C}{1 - 2^{1/q} C n^{1/q-1/2}\|D^* - \tilde{D}^*\|_{op}} \tag{4.13}$$

is positive if $\|D^* - \tilde{D}^*\|_{op} < 2^{-1/q} C^{-1} n^{1/2-1/q}$. □

Proof of Theorem 3.4. This proof is inspired by and follows closely the proof of Theorem 1.4 in [13]. Set $h = \tilde{z} - z_0$.

Step 1: Consequence of the ℓ^q minimization. As in Step 2 of the proof of Theorem 3.1, let T be any index set such that $|T| \leqslant s$, we get

$$\|\tilde{D}^*_{T^c} h\|_q^q \leqslant \|\tilde{D}^*_T h\|_q^q + 2\|\tilde{D}^*_{T^c} z_0\|_q^q + \|\tilde{D}^* w\|_q^q, \tag{4.14}$$

where as before $Aw = (B-A)z_0$, $z_0 + w$ is feasible and

$$\|w\|_2 \leqslant \frac{1}{\nu_A}\|Aw\|_2 = \frac{1}{\nu_A}\|(B-A)z_0\|_2. \tag{4.15}$$

As typically done in compressed sensing proofs using RIP, we divide the coordinates T^c into sets of size M (to be chosen later) in order of decreasing magnitude of $\tilde{D}^*_{T^c} h$. Call these sets $T_1, T_2, \ldots, T_r$ and for simplicity set $T_{01} = T \cup T_1$. By construction:

$$\|\tilde{D}^*_{T_{j+1}} h\|_\infty \leqslant \|\tilde{D}^*_{T_j} h\|_1 / M \leqslant M^{1-1/q}\|\tilde{D}^*_{T_j} h\|_q / M, \quad j \geqslant 1$$

which yields

$$\|\tilde{D}^*_{T_{j+1}} h\|_2^2 \leqslant M^{1-2/q}\|\tilde{D}^*_{T_j} h\|_q^2. \tag{4.16}$$

Using the triangle inequality, (4.1), (4.14) and (4.16), we have

$$\begin{aligned}\sum_{j\geqslant 2}\|D^*_{T_j} h\|_2^q &\leqslant \sum_{j\geqslant 2}\left(\|D^*_{T_j} h - \tilde{D}^*_{T_j} h\|_2 + M^{1/2-1/q}\|\tilde{D}^*_{T_{j-1}} h\|_q\right)^q \\ &\leqslant \sum_{j\geqslant 2}\|D^*_{T_j} h - \tilde{D}^*_{T_j} h\|_2^q + \sum_{j\geqslant 1} M^{q/2-1}\|\tilde{D}^*_{T_j} h\|_q^q \\ &\leqslant r^{1-q/2}\left(\sum_{j\geqslant 2}\|D^*_{T_j} h - \tilde{D}^*_{T_j} h\|_2^2\right)^{q/2} + \sum_{j\geqslant 1} M^{q/2-1}\|\tilde{D}^*_{T_j} h\|_q^q \\ &= r^{1-q/2}\|D^*_{T^c_{01}} h - \tilde{D}^*_{T^c_{01}} h\|_2^q + M^{q/2-1}\|\tilde{D}^*_{T^c} h\|_q^q \\ &\leqslant r^{1-q/2}\|D^*_{T^c_{01}} h - \tilde{D}^*_{T^c_{01}} h\|_2^q + M^{q/2-1}\left(\|\tilde{D}^*_T h\|_q^q + 2\|\tilde{D}^*_{T^c} z_0\|_q^q + \|\tilde{D}^* w\|_q^q\right).\end{aligned}$$

Taking the qth root of the previous inequality, writing $\tilde{D}^*_T h = \tilde{D}^*_T h - D^*_T h + D^*_T h$, and using the triangle inequality we get

$$\sum_{j\geqslant 2}\|D^*_{T_j} h\|_2 \leqslant \left(\sum_{j\geqslant 2}\|D^*_{T_j} h\|_2^q\right)^{1/q} = \rho\left(\|D^*_T h\|_2 + \eta\right), \tag{4.17}$$

where

$$\rho = 4^{1/q-1}(s/M)^{1/q-1/2} \tag{4.18}$$

and

$$\eta = \left(\frac{n}{s}\right)^{1/q-1/2} \left\| D^*_{T^c_{01}} h - \tilde{D}^*_{T^c_{01}} h \right\|_2 + \left\| \tilde{D}^*_T h - D^*_T h \right\|_2 + s^{1/2-1/q} \left(2^{1/q} \left\| \tilde{D}^*_{T^c} z_0 \right\|_q + \left\| \tilde{D}^* w \right\|_q\right). \tag{4.19}$$

The term η can be made small by controlling $\|D^* - \tilde{D}^*\|_{op}$, and w (through $\|A - B\|_{op}$) since the remaining term $\|\tilde{D}^*_{T^c} z_0\|_q$ is small by assumption.

Step 2: The use of D-RIP. The inequality (4.17) is exactly the same as the one in Lemma 2.2 of [13] except that the expressions for ρ and η are different since these expressions now contain terms that are due to perturbations of D and B. Thus, using Lemmas 2.4, 2.5 and 2.6 of [13], and the use of D-RIP combined with (4.17) will give the following two inequalities

$$\sqrt{1-\delta_{s+M}} \left\| D D^*_{T_{01}} h \right\|_2 \leqslant \rho \sqrt{1+\delta_M} \left(\|h\|_2 + \eta\right) + 2\epsilon + \left\| (A-B) z_0 \right\|_2, \tag{4.20}$$

$$\sqrt{1 - \frac{c_1}{2} - \rho^2 - \rho^2 c_2} \|h\|_2 \leqslant \frac{1}{\sqrt{2c_1}} \left\| D D^*_{T_{01}} h \right\|_2 + \rho\eta \sqrt{1 + \frac{1}{c_2}}, \tag{4.21}$$

where we have used $\|Ah\|_2 \leqslant 2\epsilon + \|(A-B)z_0\|_2$, instead of $\|Ah\|_2 \leqslant 2\epsilon$ in Lemma 2.3 of [13].

Combining (4.20) and (4.21) to eliminate $\|D D^*_{T_{01}} h\|$ yields

$$\|h\|_2 \leqslant K_1 \left(2\epsilon + \left\|(A-B) z_0\right\|_2\right) + K_2 \eta, \tag{4.22}$$

where

$$K_1 = \frac{\sqrt{1-\delta_{s+M}}}{\sqrt{2c_1(1-\delta_{s+M})(1-\frac{c_1}{2}-\rho^2-\rho^2 c_2)} - \rho\sqrt{1+\delta_M}}, \tag{4.23}$$

$$K_2 = \frac{\rho\sqrt{1+\delta_M} + \rho\sqrt{2c_1(1-\delta_{s+M})(1+1/c_2)}}{\sqrt{2c_1(1-\delta_{s+M})(1-\frac{c_1}{2}-\rho^2-\rho^2 c_2)} - \rho\sqrt{1+\delta_M}}, \tag{4.24}$$

and the particular choice of the free parameters c_1, c_2, M making the expressions for K_1 and K_2 valid and positive will be chosen at the end of the proof.

Step 3: $\|h\|_2$ is small if $\|D^* - \tilde{D}^*\|_{op}$ is small. Inequality (4.22) is not the desired estimate of $\|h\|_2$ yet since h is still included in the term η. Therefore we need to estimate η. Obviously $(\frac{n}{s})^{1/q-1/2} \geqslant 1$, so

$$\begin{aligned} \eta &\leqslant \sqrt{2}\left(\frac{n}{s}\right)^{1/q-1/2} \left\| D^* h - \tilde{D}^* h \right\|_2 + s^{1/2-1/q}\left(2^{1/q}\left\|\tilde{D}^*_{T^c} z_0\right\|_q + \left\|\tilde{D}^* w\right\|_q\right) \\ &\leqslant \sqrt{2}\left(\frac{n}{s}\right)^{1/q-1/2} \left\| D^* - \tilde{D}^* \right\|_{op} \|h\|_2 + s^{1/2-1/q}\left(2^{1/q}\left\|\tilde{D}^*_{T^c} z_0\right\|_q + \left\|\tilde{D}^* w\right\|_q\right). \end{aligned} \tag{4.25}$$

Substituting (4.25) into (4.22) and combining $\|h\|_2$ terms gives

$$(1-l)\|h\|_2 \leqslant K_1\left(2\epsilon + \left\|(A-B)z_0\right\|_2\right) + K_2 s^{1/2-1/q}\left(2^{1/q}\left\|\tilde{D}^*_{T^c} z_0\right\|_q + \left\|\tilde{D}^* w\right\|_q\right), \tag{4.26}$$

where

$$l = \sqrt{2}\left(\frac{n}{s}\right)^{1/q-1/2} K_2 \left\| D^* - \tilde{D}^* \right\|_{op}. \tag{4.27}$$

Therefore (4.26) gives an upper bound of $\|h\|_2$ if $\|D^* - \tilde{D}^*\|_{op}$ is small enough such that $l < 1$.

Step 4: Estimation of perturbations. The estimation of $\|\tilde{D}^* w\|_q$ is the same as (4.11) in Step 5 of the proof of Theorem 3.1, except here $\beta = 1$:

$$\left\|\tilde{D}^* w\right\|_q \leqslant \frac{(2n)^{1/q-1/2}}{\nu_A}\left(\left\|\tilde{D}^* - D^*\right\|_{op} + 1\right)\left\|(B-A)z_0\right\|_2. \tag{4.28}$$

For $\|\tilde{D}^*_{T^c} z_0\|_q$ we have

$$\left\|\tilde{D}^*_{T^c} z_0\right\|_q^q \leqslant \left\|\tilde{D}^*_{T^c} z_0 - D^*_{T^c} z_0\right\|_q^q + \left\|D^*_{T^c} z_0\right\|_q^q \leqslant n^{1-q/2}\left\|\tilde{D}^* - D^*\right\|_{op}^q \|z_0\|_2^q + \left\|D^*_{T^c} z_0\right\|_q^q.$$

Taking the qth root we get

$$\left\|\tilde{D}_T^* z_0\right\|_q \leqslant (2n)^{1/q-1/2}\left\|\tilde{D}^* - D^*\right\|_{op}\|z_0\|_2 + 2^{1/q-1}\left\|D_{T^c}^* z_0\right\|_q. \tag{4.29}$$

Step 5: Final estimate of $\|h\|_2$. Substituting (4.28) and (4.29) into (4.26) and letting T be the index set corresponding to the s largest magnitude entries of $D^* z_0$ yields

$$\begin{aligned}\|\tilde{z} - z_0\|_2 \leqslant{} & \frac{2K_1}{1-l}\epsilon + \frac{K_2}{1-l}(s/4)^{1/2-1/q}\left\|D_{T^c}^* z_0\right\|_q \\ & + \left(\frac{K_1}{1-l} + \frac{K_2}{\nu_A(1-l)}\left(\frac{2n}{s}\right)^{1/q-1/2}\left(1 + \left\|D^* - \tilde{D}^*\right\|_{op}\right)\right)\|B - A\|_{op}\|z_0\|_2 \\ & + \frac{\sqrt{2}K_2}{1-l}\left(\frac{4n}{s}\right)^{1/q-1/2}\left\|D^* - \tilde{D}^*\right\|_{op}\|z_0\|_2.\end{aligned} \tag{4.30}$$

Step 6: The choice of the parameters for K_1 and K_2 in Step 2. It only remains to choose the parameters c_1, c_2 and M so that K_1 and K_2 are positive. The same as in [13], we choose $c_1 = 1$, $M = 6s$ and take c_2 arbitrarily small so that the denominator of K_1 and K_2 is positive if

$$\delta_{7s} < a(q) := \frac{6 - 3(2/3)^{2/q-2}}{6 - (2/3)^{2/q-2}}.$$

In this case,

$$K_1 = \frac{\sqrt{1-\delta_{7s}}}{\sqrt{2(1-\delta_{7s})(\frac{1}{2} - \frac{3}{8}(\frac{2}{3})^{2/q}(1+c_2))} - \frac{\sqrt{6}}{4}(\frac{2}{3})^{1/q}\sqrt{1+\delta_{7s}}}, \tag{4.31}$$

$$K_2 = \frac{\frac{\sqrt{6}}{4}(\frac{2}{3})^{1/q}[\sqrt{1+\delta_{7s}} + \sqrt{2(1-\delta_{7s})(1+1/c_2)}\,]}{\sqrt{2(1-\delta_{7s})(\frac{1}{2} - \frac{3}{8}(\frac{2}{3})^{2/q}(1+c_2))} - \frac{\sqrt{6}}{4}(\frac{2}{3})^{1/q}\sqrt{1+\delta_{7s}}} \tag{4.32}$$

(choose c_2 so that K_1, K_2 are positive).

$a(1) = 0.6$ which coincides the result in [13]. Notice $a(q)$ tends to be 1 as $q \to 0$. For example, $a(q) = 0.84$ when $q = 1/2$. □

Acknowledgments

The authors thank Professors Simon Foucart and Qiyu Sun for valuable discussions related to the material. A. Powell also thanks the Academia Sinica (Taipei, Taiwan) for its hospitality and support.

References

[1] S. Foucart, M. Lai, Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leqslant 1$, Appl. Comput. Harmon. Anal. 26 (3) (2009) 395–407.
[2] S. Foucart, Notes on compressed sensing, 2009.
[3] Q. Sun, Sparse approximation property and stable recovery of sparse signals from noisy measurements, IEEE Trans. Signal Process. 19 (2011) 5086–5090.
[4] R. Chartrand, Exact reconstruction of sparse signals via nonconvex minimization, IEEE Signal Process. Lett. 14 (2007) 707–710.
[5] E. Candes, T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory 51 (12) (2005) 4203–4215.
[6] E. Candes, J. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math. 59 (2006) 1207–1223.
[7] D. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (4) (2006) 1289–1306.
[8] R. Chartrand, V. Staneva, Restricted isometry properties and non-convex compressive sensing, Inverse Problems 24 (035020) (2008) 1–14.
[9] S. Foucart, A note on guaranteed sparse recovery via l_1-minimization, Appl. Comput. Harmon. Anal. 29 (1) (2010) 97–103.
[10] R. Saab, O. Yilmaz, Sparse recovery by non-convex optimization-instance optimality, Appl. Comput. Harmon. Anal. 29 (1) (2010) 30–48.
[11] M. Herman, T. Strohmer, General deviants: An analysis of perturbations in compressed sensing, IEEE J. Sel. Topics Signal Process. 4 (2) (2010) 342–349.
[12] M. Rosenbaum, A.B. Tsybakov, Sparse recovery under matrix uncertainty, Ann. Statist. 38 (5) (2010) 2620–2651.
[13] E. Candes, Y.C. Eldar, D. Needell, P. Randall, Compressed sensing with coherent and redundant dictionaries, Appl. Comput. Harmon. Anal. 31 (1) (2010) 59–73.
[14] H. Rauhut, K. Schnass, P. Vandergheynst, Compressed sensing and redundant dictionaries, IEEE Trans. Inform. Theory 54 (5) (2008) 2210–2219.
[15] R. Gribonval, M. Nielsen, Highly sparse representations from dictionaries are unique and independent of the sparseness measure, Appl. Comput. Harmon. Anal. 22 (3) (2007) 335–355.
[16] R. Gribonval, M. Nielsen, Sparse decompositions in unions of bases, IEEE Trans. Inform. Theory 49 (12) (2003) 3320–3325.
[17] A. Cohen, W. Dahmen, R. DeVore, Compressed sensing and best k-term approximation, J. Amer. Math. Soc. 22 (2009) 211–231.
[18] S. Foucart, Real versus complex null space properties for sparse vector recovery, C. R. Math. Acad. Sci. Paris 348 (15–16) (2010) 863–865.
[19] G. Kutyniok, D. Labate, Shearlets. The first five years, Oberwolfach Rep. 44 (2010).

[20] K. Guo, D. Labate, Optimally sparse representations of 3D data with C^2 surface singularities using Parseval frames of shearlets, 2010, submitted for publication.
[21] B. Boufounos, G. Kutyniok, H. Rauhut, Sparse recovery from combined fusion frame measurements, IEEE Trans. Inform. Theory 57 (6) (2011) 3864–3876.
[22] R. Vershynin, Introduction to the non-asymptotic analysis of random matrices, in: Y. Eldar, G. Kutyniok (Eds.), Compressed Sensing, Theory and Applications, August 2010, pp. 210–268, Chapter 5.
[23] A. Aldroubi, X. Chen, A. Powell, Stability and robustness of ℓ^q minimization using null space property, in: Proceedings of SampTA, May 2011.