Hindawi Publishing Corporation
Journal of Function Spaces
Volume 2015, Article ID 579853, 10 pages
http://dx.doi.org/10.1155/2015/579853

Research Article

On the Null Space Property of l_q-Minimization for $0 < q \leq 1$ in Compressed Sensing

Yi Gao,[1,2] Jigen Peng,[1] Shigang Yue,[3] and Yuan Zhao[1,3]

[1]*School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*
[2]*School of Mathematics and Information Science, Beifang University of Nationalities, Yinchuan, Ningxia 750021, China*
[3]*School of Computer Science, University of Lincoln, Lincoln LN6 7TS, UK*

Correspondence should be addressed to Yi Gao; gaoyimh@163.com

Received 18 December 2014; Revised 1 March 2015; Accepted 2 March 2015

Academic Editor: Yuri Latushkin

The paper discusses the relationship between the null space property (NSP) and the l_q-minimization in compressed sensing. Several versions of the null space property, that is, the l_q stable NSP, the l_q robust NSP, and the $l_{q,p}$ robust NSP for $0 < p \leq q < 1$ based on the standard l_q NSP, are proposed, and their equivalent forms are derived. Consequently, reconstruction results for the l_q-minimization can be derived easily under the NSP condition and its equivalent form. Finally, the l_q NSP is extended to the l_q-synthesis modeling and the mixed l_2/l_q-minimization, which deals with the dictionary-based sparse signals and the block sparse signals, respectively.

1. Introduction

Compressed sensing has been drawing extensive and hot attention as soon as it was proposed since 2006 [1–4]. It is known as a new revolution in signal processing because it can realize sampling and compression of signal at the same time. Its fundamental idea is to recover a high-dimensional signal from remarkably small number of measurements using the sparsity of signals. N-dimensional real-world signal is called sparse if the number of its nonzero coefficients under some representation is absolutely smaller than the signal length N. Suppose that $\|x\|_0$ denotes the number of nonzero representation coefficients of the signal x; if $\|x\|_0 \leq s$, then we say the signal x is s-sparse. Here, we call tentatively $\|x\|_0$ l_0 norm (in fact, it is not a norm, because it does not satisfy the positive homogeneity of norm obviously). Using Σ_s denotes the set of all s-sparse signals: $\Sigma_s = \{x \in R^N : \|x\|_0 \leq s\}$. Suppose the observed data $y \in R^n$; we will recover the signal $x \in R^N$ via a linear system

$$y = Ax, \tag{1}$$

where A is an $n \times N$ ($n < N$) real matrix, known as measurement matrix. Generally speaking, when $n < N$, the underdetermined function $y = Ax$ has infinitely many solutions. In other words, without additional information, it is impossible to recover the signal under this condition. However, compressed sensing focuses on the sparsest solution, and its mathematical modeling is as follows:

$$\begin{aligned} &\min_{x \in R^N} \quad \|x\|_0 \\ &\text{s.t.} \quad y = Ax. \end{aligned} \tag{2}$$

We call the minimization (2) l_0 modeling. Unfortunately, the l_0-minimization (2) is a nonconvex and NP-hard optimization problem [5]. To conquer the hardness, one seeks another tractable mathematical modeling to replace the l_0-minimization. This is called l_q ($q > 0$) modeling. Consider

$$\begin{aligned} &\min_{x \in R^N} \quad \|x\|_q \\ &\text{s.t.} \quad y = Ax, \end{aligned} \tag{3}$$

where

$$\|x\|_q = \begin{cases} \left(\sum_{i=1}^{N} |x_i|^q \right)^{1/q}, & 0 < q < \infty; \\ \max_{i=1,\ldots,N} |x_i|, & q = \infty. \end{cases} \tag{4}$$

For $1 \le q \le \infty$, $\|x\|_q$ is a norm, while for $0 < q < 1$, it is a quasinorm because it does not satisfy the triangle inequality, and it only satisfies the q-triangle inequality:

$$\|x + y\|_q^q \le \|x\|_q^q + \|y\|_q^q, \quad x, y \in R^N. \tag{5}$$

It is natural that the l_2 modeling is to be considered:

$$\begin{aligned} &\min_{x \in R^N} \quad \|x\|_2 \\ &\text{s.t.} \quad y = Ax. \end{aligned} \tag{6}$$

To one's disappointment, the l_2 modeling is unable to exactly recover the sparse signals, even 1-sparse vectors. For example, supposing measurement matrix such as

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \tag{7}$$

we will recover the signal x from the measurement vector $y = Ax = [1, 0, 0]^T$. Obviously, the solutions of this function are $\eta = [1 + t, t, t, t]^T$, $t \in R$. Let $f(t) = (1 + t)^2 + 3t^2$; then $\min_{t \in R} f(t) = f(-1/4) = 3/4 < 1$. But $\eta = [3/4, -1/4, -1/4, -1/4]^T$ is not sparse. In fact, $[1, 0, 0, 0]^T$ is the sparsest solution of the above function. Moreover, for any $q > 1$, although the l_q modeling is also a convex optimization problem, it cannot guarantee the sparsest solution; that is, the l_q modeling for $q > 1$ is unable to exactly recover the sparse signals. Here, sparsity plays a key role in such recovery of signal. The problem is which q we can select to get the sparsest solution exactly.

Candès et al. gave the following l_1 modeling [2, 3, 6–8]:

$$\begin{aligned} &\min_{x \in R^N} \quad \|x\|_1 \\ &\text{s.t.} \quad y = Ax. \end{aligned} \tag{8}$$

The l_1-minimization (8) is a convex optimization problem and can be transformed into a linear programming problem to be solved tractably. A natural question is whether the solutions of (8) are equivalent to those of (2), or what kind of measurement matrices can be constructed to guarantee the equivalence of the solutions of the two problems. Candès and Tao [2] proposed the Restricted Isometry Property (RIP), which is a milestone work in sparse information processing. We say that a matrix A satisfies the Restricted Isometry Property of order s, if there exists a constant $\delta_s \in [0, 1)$, for any $x \in \Sigma_s$, such that

$$(1 - \delta_s) \|x\|_2^2 \le \|Ax\|_2^2 \le (1 + \delta_s) \|x\|_2^2. \tag{9}$$

We call the smallest constant δ_s satisfying the above inequality the Restricted Isometry Constant (RIC). Candès and Tao showed, for instance, that any s-sparse vector is exactly recovered as soon as the measurement matrix A satisfies the RIP with $\delta_{2s} \le 0.4142$ [9], and the solution of (8) is equivalent to that of (2). Later, the RIC has been improved; for example, $\delta_{2s} \le 0.4531$ by Foucart and Lai [10], $\delta_{2s} \le 0.4721$ by Cai et al. [11], $\delta_{2s} \le 0.4931$ by Mo and Li [12], and especially $\delta_{2s} \le 0.5746$ by Zhou et al. [13].

We see that the l_1 modeling is seemingly perfect selection because it not only is convex but also can exactly recover sparse signals. Then, what about the l_q-minimization under $0 < q < 1$? It is amazing that $\|x\|_q$ can also obtain the sparse solutions via hyperplane $y = Ax$, and more and more sparsity is being required as q decreases below 1 [4]. This is natural to induce one's strong interest, and some important results have been obtained. Compared with the l_1 modeling, a sufficiently sparse signal can be recovered perfectly with the l_q $(0 < q < 1)$ modeling under less restrictive RIP requirements than those needed to guarantee perfect recovery with $q = 1$ [14]. Compared with the l_0 modeling, empirical evidence strongly indicates that solving the l_q $(0 < q < 1)$ modeling takes much less time than that with $q = 0$ [15]. These surprising phenomena undoubtedly inspire more people to research the l_q $(0 < q < 1)$ modeling although the l_q-minimization problem for $0 < q < 1$ is also NP-hard in general [16]. Reconstruction sparse signals via l_q $(0 < q < 1)$ modeling have been considered in the literature in a series of papers [10, 17–20]. However, it is not clear if there exists any hidden relationship between the l_q $(0 < q < 1)$ modeling and the l_0 modeling. Recently, [21] showed that the solution of the l_q $(0 < q < 1)$ problem (3) is also equivalent to that of the l_0 problem (2) when q is smaller than a definitive constant $C_{A,y}$, where the definitive constant $C_{A,y}$ depends only on A and y, which revealed the relationship behind these interesting phenomena.

The null space property (NSP) was introduced when one checked the equivalence of the solutions between the l_0 modeling and the l_1 modeling. When we inspect the existence of the solutions of the function $y = Ax$, $x \in R^N$, we often make judgement via the null space of matrix A with linear algebra knowledge. Therefore, it is absolutely important to propose the null space property. Now, we introduce the definition of the null space property [6, 8, 22]. Given a subset S of $[N] := \{1, 2, \ldots, N\}$ and a vector $v \in R^N$, we denote by v_S the vector that coincides with v on S and that vanishes on the complementary set $S^C =: [N] \setminus S$.

Definition 1. For any set $S \subset [N]$ with $\text{card}(S) \le s$, a matrix $A \in R^{n \times N}$ is said to satisfy the null space property of order s if

$$\|v_S\|_1 < \|v_{S^C}\|_1, \tag{10}$$

for all $v \in \text{Ker } A \setminus \{0\}$.

The thoughts of the null space have appeared since one researched the best approximation of L_1 [23], but the name was first used by Cohen et al. in [22], and Donoho and Huo [6] and Gribonval and Nielsen [8] developed the thoughts and gave the following theorem.

Theorem 2. *Given a matrix $A \in R^{n \times N}$, every s-sparse vector $x \in R^N$ is the unique solution of the l_1-minimization with $y = Ax$ if and only if A satisfies the null space property of order s.*

This theorem not only gives an existence condition of the solution of the l_1 modeling but also demonstrates that the solution is right to that of the l_0 modeling. That is, it implies that for every $y = Ax$ with s-sparse x, the solution of the l_1-minimization problem (8) actually solves the l_0-minimization problem (2) when the measurement matrix A satisfies the null space property of order s. Indeed, assume that every s-sparse vector x is recovered via l_1-minimization from $y = Ax$. Let z be the minimizer of the l_0-minimization with $y = Ax$; then $\|z\|_0 \leq \|x\|_0$, so that also z is s-sparse. But since every s-sparse vector is the unique l_1-minimizer, it follows that $z = x$ [16].

As for RIP, the null space property is a necessary and sufficient condition to exactly reconstruct the signal using the l_1-minimization, which is very important in theory. We see that many works [7, 8, 16, 24–26], especially [16], focused on the null space property on the l_1-minimization and gave the stable null space property, the robust null space property, and characterization of the solutions of the l_1-minimization problem (8) under these properties. But, as for the null space property on the l_q-minimization for $0 < q < 1$, there are few papers [18, 27] touched upon it.

This paper focuses on the different types of null space property on the l_q-minimization for $0 < q < 1$. The remainder of the paper is organized as follows. In Section 2, based on the standard l_q null space property, we give the definition of the l_q stable null space property and derive its equivalent form. Then we discuss the approximation of the solutions of the l_q-minimization with $Ax = Az$. In Section 3, we further consider the l_q robust null space property and the $l_{q,p}$ robust null space property for $0 < p \leq q < 1$, respectively. Using these properties, we effectively characterize reconstruction results for the l_q-minimization when the observations are corrupted by noise. In Section 4, according to many practical scenarios, we extend the l_q null space property to the l_q-synthesis modeling and the mixed l_2/l_q-minimization, which deals with the dictionary-based sparse signals and the block sparse signals, respectively. Finally, we relegate the proofs of the main results, that is, Theorems 7, 13, 18, and 26 to Appendix.

2. l_q Stable Null Space Property

In this section, we will discuss the l_q stable null space property; for this purpose, we first introduce the l_q null space property [8, 16].

Definition 3. For any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$, a matrix $A \in R^{n\times N}$ is said to satisfy the l_q null space property of order s if

$$\|v_S\|_q < \|v_{S^C}\|_q, \tag{11}$$

for all $v \in \mathrm{Ker}\, A \setminus \{0\}$; here $0 < q < 1$, if one does not make the special statement, then one always supposes that $0 < q < 1$ in what follows.

Definition 3 generalizes Definition 1. This definition is so important that it can characterize the existence and uniqueness of the solution of the l_q-minimization (3) for $0 < q < 1$.

Theorem 4. *Given a matrix $A \in R^{n\times N}$, every s-sparse vector $x \in R^N$ is the unique solution of the l_q-minimization with $y = Ax$ if and only if A satisfies the l_q null space property of order s.*

The proof of Theorem 4 is implied in [8] and easily found in [18].

However, in more realistic scenarios, we can only claim that the vectors are close to sparse vectors but no absolutely sparse ones. In such cases, we would like to recover a vector $x \in R^N$ with an error controlled by its distance to s-sparse vectors. This property is usually referred to as the stability of the reconstruction scheme with respect to sparsity defect [16]. To better discuss this property, we give the definition of the l_q stable null space property.

Definition 5. For any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$, a matrix $A \in R^{n\times N}$ is said to satisfy the l_q stable null space property of order s with constant $0 < \tau < 1$ if

$$\|v_S\|_q \leq \tau^{1/q} \|v_{S^C}\|_q, \tag{12}$$

for all $v \in \mathrm{Ker}\, A$.

Remark 6. Formula (12) is often replaced by $\|v_S\|_q^q \leq \tau \|v_{S^C}\|_q^q$ with constant $0 < \tau < 1$.

Obviously, Definition 5 is stronger than Definition 3 if $v \in \mathrm{Ker}\, A \setminus \{0\}$. But the following theorem demonstrates that it is also important to introduce Definition 5, because the l_q stable null space property characterizes the distance between two vectors $x \in R^N$ and $z \in R^N$ satisfying $Az = Ax$.

Theorem 7. *For any set $S \subset [N]$ with $\mathrm{card}(S) \leq s$, the matrix $A \in R^{n\times N}$ satisfies the l_q stable null space property of order s with constant $0 < \tau < 1$ if and only if*

$$\|z - x\|_q^q \leq \frac{1+\tau}{1-\tau} \left(\|z\|_q^q - \|x\|_q^q + 2\|x_{S^C}\|_q^q \right), \tag{13}$$

for all vectors $x, z \in R^N$ with $Az = Ax$.

We defer the proof of this theorem to Appendix. Under this theorem, we give the main stability result as follows.

Corollary 8. *Suppose that a matrix $A \in R^{n\times N}$ satisfies the l_q stable null space property of order s with constant $0 < \tau < 1$; then for any $x \in R^N$, a solution x^* of the l_q-minimization with $y = Ax$ approximates the vector x with l_q-error:*

$$\|x - x^*\|_q \leq \left[\frac{2(1+\tau)}{1-\tau} \right]^{1/q} \sigma_s(x)_q, \tag{14}$$

where

$$\sigma_s(x)_q := \inf_{z\in\Sigma_s} \|x - z\|_q, \tag{15}$$

which denotes the error of best s-term approximation to x with respect to the l_q-quasinorm. Obviously, if $x \in \Sigma_s$, then $\sigma_s(x)_q = 0$.

Proof. Take S to be a set of s largest absolute coefficients of x, so that $\sigma_s(x)_q = \|x_{S^C}\|_q$. If x^* is a minimizer of the l_q-minimization, then $\|x^*\|_q \le \|x\|_q$ with $Ax^* = Ax$. So we take $z = x^*$ in inequality (13); then

$$\begin{aligned}\left\|x - x^*\right\|_q^q &\le \frac{1+\tau}{1-\tau}\left(\left\|x^*\right\|_q^q - \|x\|_q^q + 2\sigma_s(x)_q\right)\\ &\le \frac{2(1+\tau)}{1-\tau}\sigma_s(x)_q^q,\end{aligned} \tag{16}$$

which is the desired inequality. □

By Corollary 8, using $\sigma_s(x)_q = 0$ for $x \in \Sigma_s$, we have the following corollary.

Corollary 9. *Suppose that a matrix $A \in R^{n\times N}$ satisfies the l_q stable null space property of order s with constant $0 < \tau < 1$; then every s-sparse vector $x \in R^N$ can be exactly recovered from the l_q-minimization with $y = Ax$.*

Compared with Theorem 4, the l_q stable null space property is only a sufficient condition to exactly recover sparse vector via l_q-minimization.

Corollary 10. *Suppose that a matrix $A \in R^{n\times N}$ satisfies the l_q stable null space property of order s with constant $0 < \tau < 1$; then for any $0 < p < q < 1$ and any $x \in B_p^N := \{x \in R^N : \|x\|_p \le 1\}$, a solution x^* of the l_q-minimization with $y = Ax$ approximates the vector x with l_q-error:*

$$\left\|x - x^*\right\|_q \le \left[\frac{2(1+\tau)}{1-\tau}\right]^{1/q} \frac{C_{p,q}}{s^{1/p-1/q}}\|x\|_p, \tag{17}$$

where $C_{p,q} = [\alpha^\alpha(1-\alpha)^{1-\alpha}]^{1/p} \le 1$, $\alpha = p/q$.

Proof. This result can be derived from the following inequality [16]:

$$\sigma_s(x)_q \le \frac{C_{p,q}}{s^{1/p-1/q}}\|x\|_p, \tag{18}$$

for any $0 < p < q$ and $x \in R^N$. □

Remark 11. Under the condition of Corollary 10, we choose $p = 1/3$, $q = 1/2$, and $\tau = 1/2$; let $\|x\|_p \le 9/48$; then

$$\left\|x - x^*\right\|_{1/2} \le \frac{1}{s}. \tag{19}$$

Let us choose $p = 1/4$, $q = 1/2$, and $\tau = 1/2$; let $\|x\|_p \le 4/9$; then

$$\left\|x - x^*\right\|_{1/2} \le \frac{1}{s^2}. \tag{20}$$

3. l_q Robust Null Space Property

In realistic situations, it is also inconceivable to measure a signal $x \in R^N$ with infinite precision. This means that the measurement vector $y \in R^n$ is only an approximation of the vector $Ax \in R^n$, with

$$\left\|y - Ax\right\|_2 \le \epsilon, \tag{21}$$

for some $\epsilon \ge 0$. In this case, the reconstruction scheme should be required to output a vector $x^* \in R^N$ whose distance to the original vector $x \in R^N$ is controlled by the measurement error $\epsilon \ge 0$. This property is usually referred to as the robustness of the reconstruction scheme with respect to measurement error [16]. We are going to investigate the following modeling:

$$\begin{aligned}&\min_{x\in R^N} \quad \|x\|_q\\ &\text{s.t.} \quad \left\|y - Ax\right\|_2 \le \epsilon.\end{aligned} \tag{22}$$

Definition 12. For any set $S \subset [N]$ with $\text{card}(S) \le s$, a matrix $A \in R^{n\times N}$ is said to satisfy the l_q robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$, if

$$\left\|v_S\right\|_q^q < \tau\left\|v_{S^C}\right\|_q^q + \gamma\|Av\|_2, \tag{23}$$

for all $v \in R^N$.

We see that Definition 12 is broader than Definition 5, because it does not require that $v \in \text{Ker}\, A$. Obviously, if $v \in \text{Ker}\, A$, the l_q robust null space property implies the l_q stable null space property. Similar to Theorem 7, we also give an equivalent form of the l_q robust null space property.

Theorem 13. *For any set $S \subset [N]$ with $card(S) \le s$, the matrix $A \in R^{n\times N}$ satisfies the l_q robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$ if and only if*

$$\begin{aligned}\|z - x\|_q^q \le \frac{1+\tau}{1-\tau}\left(\|z\|_q^q - \|x\|_q^q + 2\left\|x_{S^C}\right\|_q^q\right)\\ + \frac{2\gamma}{1-\tau}\|A(z-x)\|_2,\end{aligned} \tag{24}$$

for all vectors $x, z \in R^N$.

The proof of Theorem 13 is deferred to Appendix, and spirit of this theorem is the following corollary.

Corollary 14. *Suppose that a matrix $A \in R^{n\times N}$ satisfies the l_q robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$; then for any $x \in R^N$, a solution x^* of the l_q-minimization with $y = Ax + e$ and $\|e\|_2 \le \epsilon$ approximates the vector x with l_q-error:*

$$\left\|x - x^*\right\|_q^q \le \frac{2(1+\tau)}{1-\tau}\sigma_s(x)_q^q + \frac{4\gamma}{1-\tau}\epsilon. \tag{25}$$

Proof. Noting that $\|A(z-x)\|_2 \le 2\epsilon$, the rest of proof is similar to that of Corollary 8. □

Remark 15. When $\epsilon = 0$, inequality (25) is the conclusion of Corollary 8.

Remark 16. Using inequality $(a^q + b^q)^{1/q} \le 2^{1/q-1}(a+b)$ for $a, b \ge 0$ and $0 < q \le 1$, inequality (25) can be modified such that

$$\begin{aligned}\left\|x - x^*\right\|_q &\le 2^{1/q-1}\left[\left(\frac{2(1+\tau)}{1-\tau}\right)^{1/q}\sigma_s(x)_q + \left(\frac{4\gamma}{1-\tau}\epsilon\right)^{1/q}\right] \\ &= 2^{2/q-1}\left(\frac{1+\tau}{1-\tau}\right)^{1/q}\sigma_s(x)_q + 2^{3/q-1}\left(\frac{\gamma}{1-\tau}\epsilon\right)^{1/q},\end{aligned} \tag{26}$$

and if we set $\|e\|_2 \le \epsilon^q$, $0 < q \le 1$, then we can get the following inequality:

$$\left\|x - x^*\right\|_q \le C_{q,\tau}\sigma_s(x)_q + C_{q,\tau,\gamma}\epsilon, \tag{27}$$

where $C_{q,\tau} = 2^{2/q-1}((1+\tau)/(1-\tau))^{1/q}$ and $C_{q,\tau,\gamma} = 2^{3/q-1}(\gamma/(1-\tau))^{1/q}$.

Inequality (27) or (14) is the same as that of Theorem 3.1 in [10] except for the constant when we only consider $y = Ax$, but we used the different approach; the latter is based on the measurement matrix A satisfying the following inequality:

$$\gamma_{2t} - 1 < 4\left(\sqrt{2} - 1\right)\left(\frac{t}{s}\right)^{1/q-1/2}, \tag{28}$$

for some integer $t \ge s$, where $\gamma_{2s} := \beta_{2s}^2/\alpha_{2s}^2$, while α_s and β_s are the best constants such that the measurement matrix A satisfies the following inequality:

$$\alpha_s \|x\|_2 \le \|Ax\|_2 \le \beta_s \|x\|_2, \quad x \in \Sigma_s. \tag{29}$$

If we add the robust term, inequality (27) is better than that of Theorem 3.1 in [10].

The rest of this section will focus on another robust null space property, that is, the $l_{q,p}$ robust null space property for $0 < p \le q \le 1$. In the following definition, we will use two quasinorms l_q and l_p rather than the same quasinorm.

Definition 17. Given $0 < p \le q \le 1$, for any set $S \subset [N]$ with $\mathrm{card}(S) \le s$, a matrix $A \in R^{n\times N}$ is said to satisfy the $l_{q,p}$ robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$, if

$$\left\|v_S\right\|_q^q < \frac{\tau}{s^{q/p-1}}\left\|v_{S^C}\right\|_p^p + \frac{\gamma}{s^{q/p-1}}\|Av\|_2, \tag{30}$$

for all $v \in R^N$.

Obviously, when $p = q < 1$, Definition 17 reduces to Definition 12, so we will consider the case $p < q < 1$.

Theorem 18. *Given $0 < p < q \le 1$, for any set $S \subset [N]$ with $card(S) \le s$, suppose that the matrix $A \in R^{n\times N}$ satisfies the $l_{q,p}$ robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$; then, for any $x, z \in R^N$ and $\|z - x\|_p \le 1$, one has*

$$\begin{aligned}\|z - x\|_q^q \le{}& \frac{C_1}{s^{q/p-1}}\left(\|z\|_p^p - \|x\|_p^p + 2\left\|x_{S^C}\right\|_p^p\right) \\ &+ \frac{C_2}{s^{q/p-1}}\|A(z-x)\|_2,\end{aligned} \tag{31}$$

where $C_1 = (1+\tau)^2/(1-\tau)$ and $C_2 = (3+\tau)\gamma/(1-\tau)$.

We defer the proof of this theorem to Appendix. In the proof, we will see that the condition $\|z - x\|_p \le 1$, $x, z \in R^N$, is necessary in Theorem 18.

Corollary 19. *Given $0 < p < q \le 1$, suppose that a matrix $A \in R^{n\times N}$ satisfies the $l_{q,p}$ robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$; then for any $x \in R^N$, a solution x^* of the l_q-minimization with $y = Ax+e$ and $\|e\|_2 \le \epsilon$ approximates the vector x with l_q-error:*

$$\left\|x - x^*\right\|_q^q \le \frac{2C_1}{s^{q/p-1}}\sigma_s(x)_p^p + \frac{2C_2}{s^{q/p-1}}\epsilon, \tag{32}$$

where C_1 and C_2 are the same as those of Theorem 18.

Remark 20. In Corollary 19, if we set $\epsilon = 0$, then l_q-error is controlled by $(1/s^{q/p-1})\sigma_s(x)_p^p$; that is,

$$\left\|x - x^*\right\|_q^q \le \frac{C}{s^{q/p-1}}\sigma_s(x)_p^p. \tag{33}$$

Although the quasinorm also satisfies $\|\cdot\|_q \le \|\cdot\|_p$ for $0 < p < q < 1$, the term $(1/s^{q/p-1})\sigma_s(x)_p^p$ is not worse than $\sigma_s(x)_q^q$, because the reconstruction l_q-error decays in rate $s^{q/p-1}$.

4. Extensions

In this section, we will discuss two extensions of the l_q null space property; one is on the null space property of the l_q-synthesis modeling and the other is on the l_q block null space property of the mixed l_2/l_q-minimization.

4.1. Null Space Property of the l_q-Synthesis Modeling. The techniques above hold for signals which are sparse in the standard coordinate basis or sparse with respect to some other orthonormal basis. However, in practical examples, there are numerous signals which are sparse in an overcomplete dictionary D rather than an orthonormal basis; see [27–32]. Here D is $N \times m$ ($N < m$) matrix that is often rather coherent in applications. We also call D a frame in the sense that the columns of D form a finite frame. In this setting, the signal $x \in R^N$ can be represented as $x = Df$, where f is an s-sparse vector in R^m. Such signal is called dictionary-sparse signals or frame-sparse signals [33]. When the dictionary D is specified, the signals are also called D-sparse signals. Many signals naturally possess frame-sparse coefficients, such as radar images (Gabor frames), cartoon-like images (curvelets), and images with directional features (shearlets) [32–35].

The l_1-synthesis modeling [34] is defined by

$$\begin{aligned}&\widehat{f} = \arg\min_{f\in R^m}\|f\|_1 \quad \text{s.t. } y = ADf, \\ &\widehat{x} = D\widehat{f}.\end{aligned} \tag{34}$$

The above method is called l_1-synthesis modeling due to the second synthesizing step. It is relative to the l_1-analysis

modeling, which finds the solution $\widehat{x}$ directly by solving the problem [32]

$$\widehat{x} = \arg\min_{x\in R^N} \left\|D^* x\right\|_1 \quad \text{s.t. } y = Ax, \tag{35}$$

where D^* denotes the transpose of the D.

Empirical studies show that the l_1-synthesis modeling often provides good recovery; however, it is fundamentally distinct from the l_1-analysis modeling. The geometry of two problems was analyzed in [28], and there it was shown that because these geometrical structures exhibit substantially different properties, there was a large gap between the two formulations. This theoretical gap was also demonstrated by numerical simulations in [28], which showed that the two methods perform very differently on large families of signals. Recently, [33] elaborated why the authors introduced the l_1-synthesis and the relationship of the two methods, and it appeared that the l_1-synthesis is a more thorough method than the l_1-analysis, and the l_1-analysis is a subproblem of the l_1-synthesis via sparse duals. Besides, [33] built up a framework for the l_1-synthesis method and proposed a dictionary-based null space property which is the first sufficient and necessary condition for the success of the l_1-synthesis method.

In this subsection, we only propose the l_q-synthesis modeling and give the null space property of the minimization but did not elaborate the relationship between the l_q-synthesis modeling and the l_q-analysis modeling. The l_q-synthesis modeling is as follows:

$$\begin{aligned} &\widehat{f} = \arg\min_{f\in R^m} \left\|f\right\|_q \quad \text{s.t. } y = ADf, \\ &\widehat{x} = D\widehat{f}. \end{aligned} \tag{36}$$

Given a frame $D \in R^{N\times m}$, $D^{-1}(B)$ denotes the preimage of the set B under the frame D. We introduce the following notation [33]:

$$D\Sigma_s = \left\{x \in R^N : \exists f, \text{ such that } x = Df, \ \left\|f\right\|_0 \le s\right\}. \tag{37}$$

Definition 21. Fix a dictionary $D \in R^{N\times m}$, for any set $S \subset [N]$ with $\text{card}(S) \le s$; a matrix $A \in R^{n\times N}$ is said to satisfy the l_q null space property of frame D of order s if for any $v \in D^{-1}(\ker A \setminus \{0\})$ there exists $u \in \ker D$, such that

$$\left\|v_S + u\right\|_q < \left\|v_{S^C}\right\|_q. \tag{38}$$

This null space property is abbreviated to D-NSP. Here, a natural question is proposed which is why does not AD directly satisfy the l_q null space property? The major difference is that D-NSP is essentially the minimum of $\|v_S + u\|_q$ over all $u \in \ker D$, where $u = 0$, and the D-NSP degenerates into AD having the l_q null space property. Therefore this new condition about the l_q-synthesis modeling is weaker than AD having the l_q null space property.

The following theorem asserts that the null space property of a frame D of order s is a sufficient and necessary condition for the l_q-synthesis modeling (36) to successfully recover all the D-sparse signals with sparsity s.

Theorem 22. *Fix a dictionary $D \in R^{N\times m}$; let $A \in R^{n\times N}$; for any $x_0 \in D\Sigma_s$, one has $\widehat{x} = x_0$, where $\widehat{x}$ is the reconstructed signal from $y = Ax_0$ using the l_q-synthesis modeling if only if A satisfies the l_q null space property of frame D of order s.*

Proof. Combining the proof of Theorem 4.2 in [33] with the proof of Theorem 4 (Lemma 2.2 in [18]), the content adds very little to the work. Here, the proof is omitted. □

Remark 23. On the null space property of the l_q-analysis modeling is implied in [27].

4.2. Block Null Space Property of the Mixed l_2/l_q-Minimization. The conventional compressed sensing only considers the sparsity that the signal x has at most s nonzero elements, which can appear anywhere in the vector, and it does not take into account any further structure. However, in many practical scenarios, the unknown signal x not only is sparse but also exhibits additional structure in the form that the nonzero elements are aligned to blocks rather than being arbitrarily spread throughout the vector. These signals are referred to as the block sparse signals and arise in various applications, for example, DNA microarrays [36], equalization of sparse communication channels [37], and color imaging [38]. To define the block sparsity, it is necessary to introduce some further notations. Suppose that $x \in R^N$ is split into m blocks, $x[1], x[2], \ldots, x[m]$, which are of length $d_1, d_2, \ldots, d_m$, respectively; that is,

$$x = \left[\underbrace{x_1, \ldots, x_{d_1}}_{x[1]}, \underbrace{x_{d_1+1}, \ldots, x_{d_1+d_2}}_{x[2]}, \ldots, \underbrace{x_{N-d_m+1}, \ldots, x_N}_{x[m]}\right]^T, \tag{39}$$

and $N = \sum_{i=1}^m d_i$. A vector $x \in R^N$ is called block s-sparse over $\mathcal{I} = \{d_1, d_2, \ldots, d_m\}$ if $x[i]$ is nonzero for at most s indices i. When $d_i = 1$ for each i, the block sparsity reduces to the conventional definition of a sparse vector. Denote

$$\left\|x\right\|_{2,0} = \sum_{i=1}^m I\left(\left\|x[i]\right\|_2 > 0\right), \tag{40}$$

where $I(\|x[i]\|_2 > 0)$ is an indicator function that obtains the value 1 if $\|x[i]\|_2 > 0$ and 0 otherwise. So a block s-sparse vector x can be defined by $\|x\|_{2,0} \le s$, and $\|x\|_0 = \sum_{i=1}^m \|x[i]\|_0$. Here, to avoid confusion with the above notations, we especially emphasize that s denotes s blocks of the signal out of m which are nonzero but not s entries out of N which are nonzero outlined in the above sections.

To recover a block sparse signal, similar to the standard l_0-minimization, one will pursue the sparsest block sparse vector with the following mixed l_2/l_0 modeling [39–41]:

$$\min_{x\in R^N} \left\|x\right\|_{2,0} \quad \text{s.t. } y = Ax. \tag{41}$$

But the mixed l_2/l_0-minimization problem is also NP-hard. It is natural that one uses the mixed l_2/l_1-minimization to replace the mixed l_2/l_0 model [39–42]:

$$\min_{x\in R^N} \left\|x\right\|_{2,1} \quad \text{s.t. } y = Ax, \tag{42}$$

where

$$\|x\|_{2,1} = \sum_{i=1}^{m} \|x[i]\|_2 . \tag{43}$$

To investigate the performance of this method, Eldar and Mishali [40] proposed the definition of the block Restricted Isometry Property (block RIP) of a measurement matrix $A \in R^{n\times N}$. We say that a matrix A satisfies the block RIP over $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$ of order s with positive constant $\delta_{s|\mathcal{J}}$ if for every block s-sparse vector $x \in R^N$ over $\mathcal{J}$, such that

$$\left(1 - \delta_{s|\mathcal{J}}\right)\|x\|_2^2 \le \|Ax\|_2^2 \le \left(1 + \delta_{s|\mathcal{J}}\right)\|x\|_2^2 . \tag{44}$$

Obviously, the block RIP is a natural extension of the standard RIP, but it is a less stringent requirement comparing with the standard RIP. Besides, the required number of measurements of satisfying the block RIP is less than that of satisfying the standard RIP [43]. Eldar and Mishali [40] proved that the mixed l_2/l_1-minimization can exactly recover any block s-sparse signal when the measurement matrix A satisfies the block RIP with $\delta_{2s|\mathcal{J}} < 0.414$. Recently, Lin and Li [43] improved the sufficient condition to $\delta_{2s|\mathcal{J}} < 0.4931$ and established another sufficient condition $\delta_{s|\mathcal{J}} < 0.307$ for exact recovery. There are also a number of works based on non-RIP-analysis to characterize the theoretical performance of the mixed l_2/l_1-minimization, such as block coherence [39], strong group sparsity [44], and null space characterization [42].

Based on the previous discussion of the performance of the l_q-minimization, it is also natural that one would be interested to make an ongoing effort to extend the l_q-minimization to the setting of block sparse signal recovery. Therefore, the l_2/l_q-minimization was proposed in [38, 45, 46]. Consider

$$\min_{x\in R^N} \|x\|_{2,q} \quad \text{s.t. } y = Ax, \tag{45}$$

where

$$\|x\|_{2,q} = \left(\sum_{i=1}^{m} \|x\,[i]\|_2^q\right)^{1/q} . \tag{46}$$

In [38, 45, 46], some numerical experiments demonstrated that fewer measurements are needed for exact recovery when $0 < q < 1$ than when $q = 1$. Moreover, exact recovery conditions based on block restricted q-isometry property have also been studied [46].

In this subsection, we will extend the l_q null space property to the block sparse signals. As for m-block signal $x \in R^N$, whose structure is like (39), we set $S \subset \{1, 2, \ldots, m\}$ and by S^C we mean the complement of the set S with respect to $\{1, 2, \ldots, m\}$; that is, $S^C = \{1, 2, \ldots, m\} \setminus S$.

Definition 24. For any set $S \subset \{1, 2, \ldots, m\}$ with $\mathrm{card}(S) \le s$, a matrix $A \in R^{n\times N}$ is said to satisfy the l_q block null space property over $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$ of order s, if

$$\|v_S\|_{2,q} < \|v_{S^C}\|_{2,q} , \tag{47}$$

for all $v \in \mathrm{Ker}\, A \setminus \{0\}$, where v_S denotes the vector equal to v on a block index set S and zero elsewhere.

Remark 25. Inequality (47) can be replaced by $\|v_S\|_{2,q}^q < \|v_{S^C}\|_{2,q}^q$ or $\|v_S\|_{2,q}^q < (1/2)\|v\|_{2,q}^q$ for every set $S \subset \{1, 2, \ldots, m\}$ with $\mathrm{card}(S) \le s$.

Using Definition 24, we can easily characterize the existence and uniqueness of the solution of the mixed l_2/l_q-minimization (45).

Theorem 26. *Given a matrix $A \in R^{n\times N}$, every block s-sparse vector $x \in R^N$ is the unique solution of the mixed l_2/l_q-minimization with $y = Ax$ if and only if A satisfies the l_q block null space property of order s.*

We defer the proof of Theorem 26 to Appendix.

So far, we only extend the standard l_q NSP to the l_q-synthesis modeling and the mixed l_2/l_q-minimization, respectively. As for extension of the stable l_q NSP and the robust l_q NSP, we will leave it for our next works, especially, for discussion of the block sparse signals via mixed l_2/l_q-minimization.

Appendix

In this section, we provide the proofs of Theorems 7, 13, 18, and 26. We will need the following lemma.

Lemma A.1. *Given a set $S \subset [N]$ and vector $x, z \in R^N$,*

$$\left\|(x - z)_{S^C}\right\|_q^q \le \|z\|_q^q - \|x\|_q^q + \left\|(x - z)_S\right\|_q^q + 2\left\|x_{S^C}\right\|_q^q . \tag{A.1}$$

If x^* is a solution of the l_q-minimization with $y = Ax$, then

$$\left\|(x - x^*)_{S^C}\right\|_q^q \le \left\|(x - x^*)_S\right\|_q^q + 2\left\|x_{S^C}\right\|_q^q . \tag{A.2}$$

Proof. Using the following inequalities, we can easily obtain inequality (A.1):

$$\begin{aligned} \left\|(x - z)_{S^C}\right\|_q^q &\le \left\|x_{S^C}\right\|_q^q + \left\|z_{S^C}\right\|_q^q , \\ \|x\|_q^q = \left\|x_{S^C}\right\|_q^q + \left\|x_S\right\|_q^q &\le \left\|x_{S^C}\right\|_q^q + \left\|(x - z)_S\right\|_q^q + \left\|z_S\right\|_q^q . \end{aligned} \tag{A.3}$$

□

Proof of Theorem 7. For any set $S \subset [N]$ with $\mathrm{card}(S) \le s$, let us now assume that the matrix $A \in R^{n\times N}$ satisfies the l_q stable null space property of order s with constant $0 < \tau < 1$. For $x, z \in R^N$ with $Az = Ax$, since $v = z - x \in \mathrm{Ker}\, A$, the l_q stable null space property yields

$$\left\|v_S\right\|_q^q < \tau\left\|v_{S^C}\right\|_q^q . \tag{A.4}$$

From Lemma A.1, we get

$$\begin{aligned} \left\|v_{S^C}\right\|_q^q &\le \|z\|_q^q - \|x\|_q^q + \left\|v_S\right\|_q^q + 2\left\|x_{S^C}\right\|_q^q \\ &\le \|z\|_q^q - \|x\|_q^q + \tau\left\|v_{S^C}\right\|_q^q + 2\left\|x_{S^C}\right\|_q^q . \end{aligned} \tag{A.5}$$

Since $0 < \tau < 1$, so

$$\left\|v_{S^C}\right\|_q^q \le \frac{1}{1 - \tau}\left(\|z\|_q^q - \|x\|_q^q + 2\left\|x_{S^C}\right\|_q^q\right) . \tag{A.6}$$

Using (12), we derive

$$\begin{aligned}\|v\|_q^q &= \|v_{S^C}\|_q^q + \|v_S\|_q^q \le (1+\tau)\|v_{S^C}\|_q^q \\ &\le \frac{1+\tau}{1-\tau}\left(\|z\|_q^q - \|x\|_q^q + 2\|x_{S^C}\|_q^q\right),\end{aligned} \tag{A.7}$$

which is the desired inequality.

Conversely, let us assume that the matrix A satisfies (13) for all vectors $x, z \in R^N$ with $Az = Ax$. Given a vector $v \in \operatorname{Ker} A$, since $Av_{S^C} = A(-v_S)$, let $x = -v_S$ and $z = v_{S^C}$; then $x_{S^C} = 0$, so we can get from (13)

$$\|v\|_q^q = \|v_{S^C} + v_S\|_q^q \le \frac{1+\tau}{1-\tau}\left(\|v_{S^C}\|_q^q - \|v_S\|_q^q\right), \tag{A.8}$$

and this can be written as

$$(1-\tau)\left(\|v_{S^C}\|_q^q + \|v_S\|_q^q\right) \le (1+\tau)\left(\|v_{S^C}\|_q^q - \|v_S\|_q^q\right); \tag{A.9}$$

therefore, we can get

$$\|v_S\|_q^q \le \tau\|v_{S^C}\|_q^q. \tag{A.10}$$

That is, $\|v_S\|_q \le \tau^{1/q}\|v_{S^C}\|_q$. □

Proof of Theorem 13. For any set $S \subset [N]$ with $\operatorname{card}(S) \le s$, let us assume that the matrix $A \in R^{n\times N}$ satisfies the l_q robust null space property of order s with constant $0 < \tau < 1$ and $\gamma > 0$. For $x, z \in R^N$, setting $v = z - x$, the robust l_q null space property and Lemma A.1 yield

$$\begin{aligned}\|v_S\|_q^q &< \tau\|v_{S^C}\|_q^q + \gamma\|Av\|_2, \\ \|v_{S^C}\|_q^q &\le 2\|x_{S^C}\|_q^q + \|z\|_q^q - \|x\|_q^q + \|v_S\|_q^q,\end{aligned} \tag{A.11}$$

and combining these two inequalities, we can get

$$\|v_{S^C}\|_q^q \le \frac{1}{1-\tau}\left(2\|x_{S^C}\|_q^q + \|z\|_q^q - \|x\|_q^q\right) + \frac{\gamma}{1-\tau}\|Av\|_2. \tag{A.12}$$

Using (23), we derive

$$\begin{aligned}\|v\|_q^q &= \|v_{S^C}\|_q^q + \|v_S\|_q^q \\ &\le (1+\tau)\|v_{S^C}\|_q^q + \gamma\|Av\|_2 \\ &\le \frac{1+\tau}{1-\tau}\left(2\|x_{S^C}\|_q^q + \|z\|_q^q - \|x\|_q^q\right) \\ &\quad + \frac{\gamma(1+\tau)}{1-\tau}\|Av\|_2 + \gamma\|Av\|_2 \\ &= \frac{1+\tau}{1-\tau}\left(2\|x_{S^C}\|_q^q + \|z\|_q^q - \|x\|_q^q\right) + \frac{2\gamma}{1-\tau}\|Av\|_2,\end{aligned} \tag{A.13}$$

which is the desired inequality.

The proof of sufficiency is similar to that of Theorem 7; here it is omitted. □

Proof of Theorem 18. Let us now assume that the matrix A satisfies the $l_{q,p}$ robust null space property with constant $0 < \tau < 1$ and $\tau > 0$. For $x, z \in R^N$, choosing S as an index set of s largest entries of $z - x$, noticing inequality (18) and $\|z - x\|_p \le 1$, we get

$$\begin{aligned}\|z-x\|_q^q &= \|(z-x)_S\|_q^q + \|(z-x)_{S^C}\|_q^q \\ &= \|(z-x)_S\|_q^q + \|(z-x) - (z-x)_S\|_q^q \\ &\le \frac{\tau}{s^{q/p-1}}\|(z-x)_{S^C}\|_p^p + \frac{\gamma}{s^{q/p-1}}\|A(z-x)\|_2 \\ &\quad + \left(\frac{1}{s^{1/p-1/q}}\|z-x\|_p\right)^q \\ &= \frac{\tau}{s^{q/p-1}}\|(z-x)_{S^C}\|_p^p + \frac{1}{s^{q/p-1}}\|z-x\|_p^q \\ &\quad + \frac{\gamma}{s^{q/p-1}}\|A(z-x)\|_2 \le \frac{\tau+1}{s^{q/p-1}}\|z-x\|_p^p \\ &\quad + \frac{\gamma}{s^{q/p-1}}\|A(z-x)\|_2.\end{aligned} \tag{A.14}$$

Applying (24) to $0 < p < 1$, we can get

$$\begin{aligned}\|z-x\|_q^q &\le \frac{(\tau+1)^2}{(1-\tau)s^{q/p-1}}\left(2\|x_{S^C}\|_p^p + \|z\|_p^p - \|x\|_p^p\right) \\ &\quad + \frac{2\gamma(\tau+1)}{(1-\tau)s^{q/p-1}}\|A(z-x)\|_2 \\ &\quad + \frac{\gamma}{s^{q/p-1}}\|A(z-x)\|_2 \\ &= \frac{(\tau+1)^2}{(1-\tau)s^{q/p-1}}\left(2\|x_{S^C}\|_p^p + \|z\|_p^p - \|x\|_p^p\right) \\ &\quad + \frac{\gamma(\tau+3)}{(1-\tau)s^{q/p-1}}\|A(z-x)\|_2.\end{aligned} \tag{A.15}$$

□

Proof of Theorem 26. In order to prove Theorem 26, we will need the following triangle inequality for $0 < q \le 1$:

$$\|x+y\|_{2,q}^q \le \|x\|_{2,q}^q + \|y\|_{2,q}^q, \quad x, y \in R^N. \tag{A.16}$$

Case 1. We consider the simplest case; suppose that both x and y are split into m blocks, $x = \{x[1], x[2], \ldots, x[m]\}^T$ with length $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$ and $y = \{y[1], y[2], \ldots, y[m]\}^T$ with length $\mathcal{J} = \{l_1, l_2, \ldots, l_m\}$, respectively. We also assume that both $x[i]$ and $y[i]$ $(i = 1, 2, \ldots, m)$ have the same number of entries; that is, $d_i = l_i$ $(i = 1, 2, \ldots, m)$. Then inequality (A.16) is easily obtained by only noticing that $\|x+y\|_{2,q}^q = \sum_{i=1}^m \|x[i] + y[i]\|_2^q$ and using the fact that $\|\cdot\|_2^q$ satisfies the triangle inequality for any $0 < q \le 1$; that is,

$$\|x[i] + y[i]\|_2^q \le \|x[i]\|_2^q + \|y[i]\|_2^q, \quad i = 1, 2, \ldots, m. \tag{A.17}$$

Case 2. Suppose that x is split into m blocks with length $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$, but y is split into k blocks with length

$\mathcal{J} = \{l_1, l_2, \ldots, l_k\}$; we might as well set $k < m$; then in the top k blocks, there exist some blocks in y, whose entries are more than those of corresponding blocks in x. Without loss of generality, we might as well suppose that the first block has the above property; that is, $l_1 > d_1$; then we only supplement $l_1 - d_1$ zeros in block $x[1]$, such that $x[1]$ and $y[1]$ have the same number of entries. By Case 1, inequality (A.17) also holds and

$$\sum_{i=1}^{k} \|x[i] + y[i]\|_2^q \leq \sum_{i=1}^{k} \|x[i]\|_2^q + \sum_{i=1}^{k} \|y[i]\|_2^q . \tag{A.18}$$

We can use the same way to deal with the case if there exist some blocks in x whose entries are more than those of corresponding block in y for the top k blocks. Finally, we supplement again $m - k$ zero-blocks in x; then inequality (A.18) can be written by

$$\sum_{i=1}^{m} \|x[i] + y[i]\|_2^q \leq \sum_{i=1}^{k} \|x[i]\|_2^q + \sum_{i=1}^{m} \|y[i]\|_2^q , \tag{A.19}$$

which is the desired inequality.

Now we give the proof of Theorem 26.

Let us first assume that every block s-sparse vector $x \in R^N$ is the unique solution of the mixed l_2/l_q-minimization with $y = Ax$, and suppose that x is split into m blocks with size set $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$. Given a fixed index set $S \subset \{1, 2, \ldots, m\}$ with $\text{card}(S) \leq s$, for any $v \in \text{Ker}\, A \setminus \{0\}$, since $v = v_S + v_{S^C}$ satisfies $Av = 0$, we have $Av_S = A(-v_{S^C})$. Since v_S is block s-sparse and $v_S \neq -v_{S^C}$, therefore

$$\|v_S\|_{2,q} < \|-v_{S^C}\|_{2,q} = \|v_{S^C}\|_{2,q} . \tag{A.20}$$

Conversely, let us assume that the block null space property relative to S holds. Suppose that x is a block s-sparse vector with m blocks, $x = \{x[1], x[2], \ldots, x[m]\}^T$ with length $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$; if we let $S := \{i : x[i] \neq 0,\ i = 1, 2, \ldots, m\}$, then $\text{card}(S) \leq s$. Let further $z \neq x$ be such that $Az = Ax$; without loss of generality, suppose that z is also split into m blocks in terms of x, $z = \{z[1], z[2], \ldots, z[m]\}^T$ with the same length as $\mathcal{J} = \{d_1, d_2, \ldots, d_m\}$. Then $v = x - z \in \text{Ker}\, A \setminus \{0\}$ and

$$\begin{aligned} \|x\|_{2,q}^q &\leq \|x_S - z_S\|_{2,q}^q + \|z_S\|_{2,q}^q \\ &= \|v_S\|_{2,q}^q + \|z_S\|_{2,q}^q \\ &< \|v_{S^C}\|_{2,q}^q + \|z_S\|_{2,q}^q \\ &= \|-z_{S^C}\|_{2,q}^q + \|z_S\|_{2,q}^q \\ &= \sum_{i \in S^C} \|z[i]\|_2^q + \sum_{i \in S} \|z[i]\|_2^q = \|z\|_{2,q}^q , \end{aligned} \tag{A.21}$$

so we have $\|x\|_{2,q} < \|z\|_{2,q}$. This establishes the required minimality of $\|x\|_{2,q}$. Here, we have made use of inequality (A.16). □

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the anonymous reviewers for their insightful comments and valuable suggestions. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant no. 11131006 and by a Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme, via EYE2E (269118) and LIVCODE (295151), and in part by the Science Research Project of Ningxia Higher Education Institutions under Grant no. NGY20140147.

References

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[3] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.

[6] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[7] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2558–2567, 2002.

[8] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3320–3325, 2003.

[9] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[10] S. Foucart and M. J. Lai, "Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$," *Applied and Computational Harmonic Analysis*, vol. 26, pp. 395–407, 2009.

[11] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1300–1308, 2010.

[12] Q. Mo and S. Li, "New bounds on the restricted isometry constant δ_{2k}," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 460–468, 2011.

[13] S. Zhou, L. Kong, and N. Xiu, "New bounds for RIC in compressed sensing," *Journal of the Operations Research Society of China*, vol. 1, no. 2, pp. 227–237, 2013.

[14] R. Saab and O. Yilmaz, "Sparse recovery by non-convex optimization-instance optimality," *Applied and Computational Harmonic Analysis*, vol. 29, no. 1, pp. 30–48, 2010.

[15] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.

[16] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer Science+Business Media, New York, NY, USA, 2013.

[17] M. Davies and R. Gribonval, "Restricted isometry constants where lp sparse recovery can fail for 0 < q ≤ 1," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2203–2214, 2009.

[18] S. Foucart, A. Pajor, H. Rauhut, and T. Ullrich, "The Gelfand widths of l_p-balls for $0 < q \leq 1$," *Journal of Complexity*, vol. 26, pp. 629–640, 2010.

[19] S. Li and J. H. Lin, "Compressed sensing with coherent tight frames via l_q-minimization for $0 < q \leq 1$," Inverse Problems and Imaging, http://arxiv.org/abs/1105.3299.

[20] Q. Sun, "Recovery of sparsest signals via l_q-minimization," *Applied and Computational Harmonic Analysis*, vol. 32, pp. 329–341, 2012.

[21] J. Peng, S. Yue, and H. Li, "*NP*/*CLP* equivalence: a phenomenon hidden among sparsity models for information processing," http://arxiv.org/abs/1501.02018.

[22] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best *k*-term approximation," *Journal of the American Mathematical Society*, vol. 22, no. 1, pp. 211–231, 2009.

[23] A. Pinkus, *On L^1-Approximation*, Cambridge University Press, Cambridge, UK, 1989.

[24] B. S. Kashin and V. N. Temlyakov, "A remark on compressed sensing," *Mathematical Notes*, vol. 82, no. 5-6, pp. 748–755, 2007.

[25] M. J. Lai and Y. Liu, "The null space property for sparse recovery from multiple measurement vectors," *Applied and Computational Harmonic Analysis*, vol. 30, no. 3, pp. 402–406, 2011.

[26] Z. Q. Xu, "Compressed sensing," *Chinese Science: Maths*, vol. 42, no. 9, pp. 865–877, 2012.

[27] A. Aldroubi, X. Chen, and A. M. Powell, "Perturbations of measurement matrices and dictionaries in compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 33, no. 2, pp. 282–291, 2012.

[28] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, 2007.

[29] R. Gribonval and M. Nielsen, "Highly sparse representations from dictionaries are unique and independent of the sparseness measure," *Applied and Computational Harmonic Analysis*, vol. 22, no. 3, pp. 335–355, 2007.

[30] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2210–2219, 2008.

[31] M. Elad, *Sparse and Redundant Representations, From Theory to Applications in Signal and Image Processing*, Springer, New York, NY, USA, 2010.

[32] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[33] X. Chen, H. Wang, and R. Wang, "A null space analysis of the l1-synthesis method in dictionary-based compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 37, pp. 492–515, 2014.

[34] Y. Liu, S. Li, T. Mi, H. Lei, and W. Yu, "Performance analysis of l1-synthesis with coherent frames," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '12)*, pp. 2042–2046, July 2012.

[35] Y. Liu, T. Mi, and S. Li, "Compressed sensing with general frames via optimal-dual-based l_1-analysis," *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4201–4214, 2012.

[36] F. Parvaresh, H. Vikalo, S. Misra, and B. Hassibi, "Recovering sparse signals using sparse measurement matrices in compressed DNA microarrays," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 275–285, 2008.

[37] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 374–377, 2002.

[38] A. Majumdar and R. K. Ward, "Compressed sensing of color images," *Signal Processing*, vol. 90, no. 12, pp. 3122–3127, 2010.

[39] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[40] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, 2009.

[41] E. Elhamifar and R. Vidal, "Block-sparse recovery via convex optimization," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4094–4107, 2012.

[42] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3075–3085, 2009.

[43] J. H. Lin and S. Li, "Block sparse recovery via mixed $l_2 = l_q$ minimization," *Acta Mathematica Sinica*, vol. 29, no. 7, pp. 1401–1412, 2013.

[44] J. Huang and T. Zhang, "The benefit of group sparsity," *The Annals of Statistics*, vol. 38, no. 4, pp. 1978–2004, 2010.

[45] Y. Wang, J. Wang, and Z. B. Xu, "On recovery of block-sparse signals via mixed $l_2 = l_q (0 < q \leq 1)$ norm minimization," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, article 76, 17 pages, 2013.

[46] Y. Wang, J. Wang, and Z. B. Xu, "Restricted *p*-isometry properties of nonconvex block-sparse compressed sensing," *Signal Processing*, vol. 104, pp. 188–196, 2014.