# Sparse approximation property and stable recovery of sparse signals from noisy measurements

Qiyu Sun



***Abstract*—In this paper, we introduce a sparse approximation property of order $s$ for a measurement matrix $\mathbf{A}$:**

$$\|\mathbf{x}_s\|_2 \le D\|\mathbf{A}\mathbf{x}\|_2 + \beta\frac{\sigma_s(\mathbf{x})}{\sqrt{s}} \quad \text{for all } \mathbf{x},$$

**where $\mathbf{x}_s$ is the best $s$-sparse approximation of the vector $\mathbf{x}$ in $\ell^2$, $\sigma_s(\mathbf{x})$ is the $s$-sparse approximation error of the vector $\mathbf{x}$ in $\ell^1$, and $D$ and $\beta$ are positive constants. The sparse approximation property for a measurement matrix can be thought of as a weaker version of its restricted isometry property and a stronger version of its null space property. In this paper, we show that the sparse approximation property is an appropriate condition on a measurement matrix to consider stable recovery of any compressible signal from its noisy measurements. In particular, we show that any compressible signal can be stably recovered from its noisy measurements via solving an $\ell^1$-minimization problem if the measurement matrix has the sparse approximation property with $\beta \in (0,1)$, and conversely the measurement matrix has the sparse approximation property with $\beta \in (0,\infty)$ if any compressible signal can be stably recovered from its noisy measurements via solving an $\ell^1$-minimization problem.**

## I. INTRODUCTION

Given positive integers $m$ and $n$ with $m \le n$ and a measurement matrix $\mathbf{A}$ of size $m\times n$, we consider the problem of compressive sampling in recovering a compressible signal $\mathbf{x}\in\mathbb{R}^n$ from its noisy measurements $\mathbf{z}=\mathbf{A}\mathbf{x}+\mathbf{n}$ via solving the following $\ell^q$-minimization problem:

$$\min \|\mathbf{y}\|_q^q \quad \text{subject to } \|\mathbf{A}\mathbf{y}-\mathbf{z}\|_p \le \epsilon, \tag{I.1}$$

where $0<q\le 1, q\le p\le\infty, \epsilon\ge 0$, and the measurement noise $\mathbf{n}$ satisfies $\|\mathbf{n}\|_p\le\epsilon$ ([1] – [8]). Here $\|\cdot\|_q, 0<q\le\infty$, stand for the "$\ell^q$-norm" on the Euclidean space.

Given a subset $S\subset\{1,\dots,n\}$ and a vector $\mathbf{x}\in\mathbb{R}^n$, denoted by $\mathbf{x}_S$ the vector whose components on $S$ are the same as those of the vector $\mathbf{x}$ and vanish on the complement $S^c$. A vector $\mathbf{x}\in\mathbb{R}^n$ is said to be *$s$-sparse* if $\mathbf{x}=\mathbf{x}_S$ for some subset $S\subset\{1,\dots,n\}$ with its cardinality $\#S$ less than or equal to $s$, where $s\ge 1$. Denote by $\Sigma_s$ the set of all $s$-sparse vectors. Given a vector $\mathbf{x}$, its best $s$-sparse approximation vector $\mathbf{x}_s$ in $\ell^q$ is an $s$-sparse vector which has minimal distance to $\mathbf{x}$ in $\ell^q$; i.e., $\|\mathbf{x}-\mathbf{x}_s\|_q=\sigma_{s,q}(\mathbf{x}):=\inf_{\mathbf{y}\in\Sigma_s}\|\mathbf{x}-\mathbf{y}\|_q$. For $q=1$, we use $\sigma_s(\mathbf{x})$ instead of $\sigma_{s,1}(\mathbf{x})$ for brevity.



Q. Sun is with the Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA. His telephone number is 407-823-4839, fax number is 407-823-6253, and email address is qiyu.sun@ucf.edu

In this paper, we introduce a new property of a measurement matrix $\mathbf{A}$: there exist positive constants $D$ and $\beta$ such that

$$\|\mathbf{x}_s\|_r^q \le D\|\mathbf{A}\mathbf{x}\|_p^q + \beta s^{q/r-1}(\sigma_{s,q}(\mathbf{x}))^q \quad \text{for all } \mathbf{x}\in\mathbb{R}^n, \tag{I.2}$$

where $0<p,q,r\le\infty$, $s$ is a positive integer, and $\mathbf{x}_s$ is the best $s$-sparse approximation of the vector $\mathbf{x}$ in $\ell^q$. The property of a measurement matrix mentioned in the abstract is a special case of the above property where $p=r=2$ and $q=1$. We call the property (I.2) the *sparse approximation property of order* $s$, as it is closely related to the best $s$-sparse approximation. We call the minimal constant $\beta$ such that (I.2) holds the *sparse approximation constant*, and denote it by $\beta_s(\mathbf{A})$.

In this paper, we show that for the stable recovery of a compressible signal $\mathbf{x}$ from its noisy measurements $\mathbf{z}=\mathbf{A}\mathbf{x}+\mathbf{n}$ via solving the $\ell^q$-minimization problem (I.1), the sparse approximation property (I.2) with sparse approximation constant $\beta_s(\mathbf{A})<1$ is **sufficient** while the sparse approximation property (I.2) with finite sparse approximation constant $\beta_s(\mathbf{A})$ is **necessary**. We refer the reader to [2], [3], [7], [9] – [17] and the references therein for other various conditions on a measurement matrix that guarantee the stable recovery of any compressible signal from its noisy measurements via solving the $\ell^q$-minimization problem (I.1).

*Theorem 1.1:* Let $0<q\le 1, q\le r\le\infty, 1\le p\le\infty$, $\epsilon\ge 0$, positive integers $m,n,s$ satisfy $2s\le m\le n$, $\mathbf{A}$ be a matrix of size $m\times n$ having the sparse approximation property (I.2) with $D\in(0,\infty)$ and $\beta\in(0,1)$, $\mathbf{z}=\mathbf{A}\mathbf{x}+\mathbf{n}$ with $\|\mathbf{n}\|_p\le\epsilon$ and $\mathbf{x}\in\mathbb{R}^n$, and let $\mathbf{x}^*$ be the solution of the $\ell^q$-minimization problem (I.1). Then

$$\|\mathbf{x}^*-\mathbf{x}\|_r^q \le \frac{(3+\beta)D}{1-\beta}(2\epsilon)^q + \frac{2(1+\beta)^2}{1-\beta}s^{q/r-1}(\sigma_{s,q}(\mathbf{x}))^q \tag{I.3}$$

and

$$\|\mathbf{x}^*-\mathbf{x}\|_q^q \le \frac{(3+\beta)D}{1-\beta}s^{1-q/p}(2\epsilon)^q + \frac{2(1+\beta)^2}{1-\beta}(\sigma_{s,q}(\mathbf{x}))^q \tag{I.4}$$

if $q<r$, and

$$\|\mathbf{x}^*-\mathbf{x}\|_q^q \le \frac{2D}{1-\beta}(2\epsilon)^q + \frac{2(1+\beta)}{1-\beta}(\sigma_{s,q}(\mathbf{x}))^q \tag{I.5}$$

if $q=r$.

*Theorem 1.2:* Let $0<q,p\le\infty$, positive integers $m,n,s$ satisfy $2s\le m\le n$, and let $\mathbf{A}$ be a matrix of size $m\times n$. If for any $\epsilon\ge 0$ and $\mathbf{x}\in\mathbb{R}^n$, the error between the given vector $\mathbf{x}$ and the solution $\mathbf{x}^*$ of the $\ell^q$-minimization problem (I.1) satisfies

$$\|\mathbf{x}^*-\mathbf{x}\|_p^q \le B_1\epsilon^q + B_2 s^{q/p-1}(\sigma_{s,q}(\mathbf{x}))^q, \tag{I.6}$$

where $B_1$ and $B_2$ are positive constants independent of $\epsilon$ and $\mathbf{x}$, then

$$\|\mathbf{x}\|_p^q \le B_1\|\mathbf{A}\mathbf{x}\|_p^q + B_2 s^{q/p-1}(\sigma_{s,q}(\mathbf{x}))^q \quad \text{for all } \mathbf{x}\in\mathbb{R}^n, \tag{I.7}$$

and hence $\mathbf{A}$ has the sparse approximation property (I.2) with $r=p$, $D=B_1$ and $\beta=B_2$.

The $m\times n$ adjacency matrix $\mathbf{\Phi}$ of an unbalanced $(2s,\alpha)$-expander with left degree $d$ and $\alpha\in(0,1/4)$ satisfies

$$\|\mathbf{x}_s\|_1 \le \frac{1}{d(1-4\alpha)}\|\mathbf{\Phi}\mathbf{x}\|_1 + \frac{2\alpha}{1-4\alpha}\sigma_s(\mathbf{x}) \quad \text{for all } \mathbf{x}\in\mathbb{R}^n, \tag{I.8}$$

(and hence it has the sparse approximation property (I.2) with $p=q=r=1$). The above property for the adjacency matrix $\mathbf{\Phi}$ is established in [12, Lemma 16] implicitly. Then by (I.8) and Theorem 1.1, we have the following result similar to [12, Theorem 17].

*Corollary 1.3:* Let $\epsilon\ge 0$, positive integers $m,n,s$ satisfy $2s\le m\le n$, $\alpha\in(0,1/6)$, $\mathbf{\Phi}$ be the $m\times n$ adjacency matrix of an unbalanced $(2s,\alpha)$-expander with left degree $d$, $\mathbf{z}=\mathbf{\Phi}\mathbf{x}+\mathbf{n}$ with $\|\mathbf{n}\|_1\le\epsilon$ for some $\mathbf{x}\in\mathbb{R}^n$, and let $\mathbf{x}^*$ be the solution of the minimization problem (I.1) with $p=q=1$. Then

$$\|\mathbf{x}^*-\mathbf{x}\|_1 \le \frac{4}{d(1-6\alpha)}\epsilon + \frac{2-4\alpha}{1-6\alpha}\sigma_s(\mathbf{x}). \tag{I.9}$$

The paper is organized as follows. One of two basic properties of a measurement matrix $\mathbf{A}$ in compressive sampling ([18] – [24]) is the *null space property of order* $s$ in $\ell^q, 0<q\le 1$; i.e., there exists a positive constant $\gamma$ such that

$$\|\mathbf{x}_S\|_q^q \le \gamma\|\mathbf{x}_{S^c}\|_q^q \tag{I.10}$$

hold for all vectors $\mathbf{x}$ in the null space $N(\mathbf{A})$ of the matrix $\mathbf{A}$ and all sets $S$ with cardinality $\#S$ less than or equal to $s$. In Section II, we show in Theorem 2.1 that any measurement matrix satisfying (I.2) will have the null space property (I.10). So the sparse approximation property (I.2) of a measurement matrix can be considered as a **stronger** version of the null space property (I.10). The other basic property of a measurement matrix $\mathbf{A}$ in compressive sampling ([1], [2], [7], [18] – [24]) is the *restricted isometry property of order* $s$; i.e., there exists a positive constant $\delta\in(0,1)$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{A}\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \quad \text{for all } \mathbf{x}\in\Sigma_s. \tag{I.11}$$

In Section III, we prove that if a measurement matrix has the restricted isometry property (I.11) of order $2s$ then it has the sparse approximation property (I.2) with $p=r=2$, and furthermore the constant $\beta$ in (I.2) is small when the restricted isometry constant is small, see Theorems 3.1 and 3.2 for details. Thus the sparse approximation property (I.2) of a measurement matrix can also thought of as a **weaker** version of the restricted isometry property (I.11), see also Remarks 3.3 and 3.4. The proofs of all theorems are included in the appendix.

## II. Null space property and sparse approximation property

Let $\mathcal{R}(\mathbf{A})$ be the set of matrices $\mathbf{R}$ satisfying $\mathbf{A}=\mathbf{A}\mathbf{R}\mathbf{A}$, and denote by $\|\mathbf{R}\|_{p\to q}$ the operator norm of a matrix $\mathbf{R}$ from $\ell^p$ to $\ell^q$, i.e., $\|\mathbf{R}\mathbf{x}\|_q\le\|\mathbf{R}\|_{p\to q}\|\mathbf{x}\|_p$ for all vectors $\mathbf{x}$. In this section, we show that any measurement matrix satisfying (I.2) will have the null space property (I.10) with its null space constant less than or equal to the constant $\beta$ in (I.2). Here *null space constant* $\gamma_s(\mathbf{A})$ of a measurement matrix $\mathbf{A}$ is the minimal constant $\gamma$ such that (I.10) holds.

*Theorem 2.1:* Let $0<q\le r\le\infty$, $0<p<\infty$, integers $m,n,s$ satisfy $2\le 2s\le m\le n$, and $\mathbf{A}$ be a matrix of size $m\times n$. Then the following statements hold.

(i) If the matrix $\mathbf{A}$ has the sparse approximation property (I.2), then it has the null space property of order $s$ in $\ell^q$ with its null space constant $\gamma_s(\mathbf{A})\le\beta_s(\mathbf{A})$.

(ii) If the matrix $\mathbf{A}$ has the null space property of order $s$ in $\ell^q$ with the null space constant $\gamma_s(\mathbf{A})$, then it has the sparse approximation property (I.2) with $p=q=r$, $D=\max(1,\gamma_s(\mathbf{A}))\inf_{\mathbf{R}\in\mathcal{R}(\mathbf{A})}\|\mathbf{R}\|_{q\to q}^q$ and $\beta=\gamma_s(\mathbf{A})$; i.e.,

$$\begin{aligned}\|\mathbf{x}_s\|_q^q \ \le\ & \big(\max(1,\gamma_s(\mathbf{A}))\inf_{\mathbf{R}\in\mathcal{R}(\mathbf{A})}\|\mathbf{R}\|_{q\to q}^q\big)\|\mathbf{A}\mathbf{x}\|_q^q\\ &+\gamma_s(\mathbf{A})(\sigma_{s,q}(\mathbf{x}))^q \quad \text{for all } \mathbf{x}\in\mathbb{R}^n.\end{aligned}$$

Applying Theorems 1.1 and 2.1 with $p=q=r=1$, we have the following result on recovering compressible signals from noisy measurements when the measurement matrix has the null space property of order $s$ in $\ell^1$, which is obtained in [8] for the noiseless case.

*Corollary 2.2:* Let $\epsilon\ge 0$, $m,n,s$ be positive integers with $2s\le m\le n$, $\mathbf{A}$ be a matrix of size $m\times n$ satisfying the null space property (I.10) with $q=1$, $\mathbf{z}=\mathbf{A}\mathbf{x}+\mathbf{n}$ with $\|\mathbf{n}\|_1\le\epsilon$ and $\mathbf{x}\in\mathbb{R}^n$, and let $\mathbf{x}^*$ be the solution of the minimization problem (I.1) with $p=q=1$. If the null space constant $\gamma_s(\mathbf{A})\in(0,1)$, then

$$\|\mathbf{x}^*-\mathbf{x}\|_1 \le \frac{4\inf_{\mathbf{R}\in\mathcal{R}(\mathbf{A})}\|\mathbf{R}\|_{1\to 1}}{1-\gamma_s(\mathbf{A})}\epsilon + \frac{2+2\gamma_s(\mathbf{A})}{1-\gamma_s(\mathbf{A})}\sigma_s(\mathbf{x}).$$

*Remark 2.3:* The null space property of a measurement matrix is invariant under preconditioning, i.e., if a measurement matrix $\mathbf{A}$ has the null space property (I.10) then the preconditioned matrix $\mathbf{PA}$ has the null space property (I.10) with the same null space constants, where a preconditioner is a nonsingular matrix $\mathbf{P}$. The sparse approximation property (I.2) is weakly preconditioning-invariant in the sense that if a measurement matrix $\mathbf{A}$ satisfies (I.2) then the preconditioned matrix $\mathbf{PA}$ also satisfies (I.2) with $D$ replaced by $\|\mathbf{P}^{-1}\|_{p\to p}D$. This suggests appropriate preconditioning the measurement matrix (and hence the noisy measurements) before signal recovery from its noisy measurements via solving an $\ell^q$-minimization problem.

*Remark 2.4:* Let the matrix $\mathbf{A}$ of size $m\times n$ have full rank $m$ (which is the case in most of compressive sampling problems) and $\mathbf{A}=\mathbf{U}\Sigma\mathbf{V}^t$ be its singular value decomposition. Here and hereafter $\mathbf{x}^t$ stands for the transpose of a vector or a matrix $\mathbf{x}$. Then $\Sigma=(\Sigma'\ \mathbf{0})$ for some nonsingular diagonal matrix $\Sigma'$. Now we define the conventional preconditioned measurement matrix $\tilde{\mathbf{A}}$ by $\tilde{\mathbf{A}}=\mathbf{PA}$, where $\mathbf{P}=(\Sigma')^{-1}\mathbf{U}^t$. In this case, $\mathbf{R}\in\mathcal{R}(\mathbf{A})$ if and only if $\mathbf{R}=\mathbf{V}\begin{pmatrix}\mathbf{I}\\ \mathbf{B}\end{pmatrix}$, where $\mathbf{I}$ is the unit matrix of size $m\times m$ and $\mathbf{B}$ is an arbitrary matrix of size $(m-n)\times n$. Let $\mathbf{v}_i, 1\le i\le n$, be the column vectors of the matrix $\mathbf{V}$. Then the null space $N(\mathbf{A})$ of the matrix $\mathbf{A}$ is spanned by $\mathbf{v}_i, m+1 \le i \le n$, and the vectors $\mathbf{v}_i, 1 \le i \le m$, form an orthonormal basis for $N(\mathbf{A})^\perp$, the orthogonal complement of the null space $N(\mathbf{A})$ of the matrix $\mathbf{A}$. As the set $\{\mathbf{Rx} : \ \|\mathbf{x}\|_1 \le 1\}$ is a polygon, the maximal $\ell^1$-norm of $\mathbf{Rx}, \|\mathbf{x}\|_1 \le 1$, is then attained at some vertices. Thus

$$\begin{aligned} &\inf_{\mathbf{R}\in\mathcal{R}(\tilde{\mathbf{A}})} \|\mathbf{R}\|_{1\to 1} = \inf_{\mathbf{R}\in\mathcal{R}(\tilde{\mathbf{A}})} \max_{1\le i\le m} \|\mathbf{Re}_i\|_1 \\ &= \inf_{\mathbf{B}} \max_{1\le i\le m} \|\mathbf{v}_i - (\mathbf{v}_{m+1} \ \cdots \ \mathbf{v}_n)\mathbf{Be}_i\|_1 \\ &= \inf_{\mathbf{u}\in N(\mathbf{A})} \max_{1\le i\le m} \|\mathbf{v}_i - \mathbf{u}\|_1, \end{aligned} \tag{II.1}$$

where $\mathbf{e}_i, 1 \le i \le m$, form the standard orthonormal basis of $\mathbb{R}^m$. In other words, the quantity $\inf_{\mathbf{R}\in\mathcal{R}(\tilde{\mathbf{A}})} \|\mathbf{R}\|_{1\to 1}$ is the same as the distance of $\mathbf{v}_i, 1 \le i \le m$, from the null space $N(\mathbf{A})$ in $\ell^1$. From (II.1) it follows that $\inf_{\mathbf{R}\in\mathcal{R}(\tilde{\mathbf{A}})} \|\mathbf{R}\|_{1\to 1} \le \max_{1\le i\le m} \|\mathbf{v}_i\|_1 \le n^{1/2}$. It would be an interesting topic on preconditioning a measurement matrix $\mathbf{A}$ with the null space property (I.10) such that the quantity $\inf_{\mathbf{R}\in\mathcal{R}(\tilde{\mathbf{A}})} \|\mathbf{R}\|_{q\to q}, 0 < q \le 1$, for the preconditioned matrix $\tilde{\mathbf{A}}$ is not a large number.

## III. Restricted isometry property and sparse approximation property

In this section, we prove that if a measurement matrix has the restricted isometry property (I.11) of order $2s$, then it has the sparse approximation property (I.2) with $p = r = 2$, and the sparse approximation constant is small when the restricted isometry constant is small. Here the *restricted isometry constant* $\delta_s(\mathbf{A})$ of a measurement matrix $\mathbf{A}$ is the smallest positive constant $\delta$ that satisfies (I.11).

*Theorem 3.1:* Let $0 < q \le 1$, positive integers $m, n, s$ satisfy $2s \le m \le n$, and the matrix $\mathbf{A}$ of size $m \times n$ have the restricted isometry property (I.11) of order $2s$ with restricted isometry constant $\delta_{2s}(\mathbf{A}) \in (0, 1)$. Then for all $\mathbf{x} \in \mathbb{R}^n$,

$$\begin{aligned} \|\mathbf{x}\|_2^2 \ \le \ & \frac{\sqrt{1+\delta_{2s}(\mathbf{A})} + \sqrt{2\delta_{2s}(\mathbf{A})}}{(1-\delta_{2s}(\mathbf{A}))\sqrt{1+\delta_{2s}(\mathbf{A})}} \|\mathbf{Ax}\|_2^2 \\ & + \Big(\frac{\sqrt{1+\delta_{2s}(\mathbf{A})} + \sqrt{2\delta_{2s}(\mathbf{A})}}{1-\delta_{2s}(\mathbf{A})}\Big)^2 \\ & \times \delta_{2s}(\mathbf{A}) s^{1-2/q} \big(\sigma_{s,q}(\mathbf{x})\big)^2 \end{aligned} \tag{III.1}$$

and

$$\begin{aligned} \|\mathbf{Ax}\|_2^2 \ \le \ & \big(1+\delta_{2s}(\mathbf{A}) + \sqrt{2\delta_{2s}(\mathbf{A})}\big)\|\mathbf{x}\|_2^2 \\ & + \big(1+\sqrt{2\delta_{2s}(\mathbf{A})}\big)\delta_{2s}(\mathbf{A}) \\ & \times s^{1-2/q}(\sigma_{s,q}(\mathbf{x}))^2. \end{aligned} \tag{III.2}$$

*Theorem 3.2:* Let $0 < q \le r \le \infty, 0 < p \le \infty$, positive integers $m, n, s$ satisfy $2s \le m \le n$, and $\mathbf{A}$ be a matrix of size $m \times n$ that has the sparse approximation property (I.2). Then

$$\frac{1}{D}\|\mathbf{x}\|_r^q \le \|\mathbf{Ax}\|_p^q \quad \text{for all } \mathbf{x} \in \Sigma_s, \tag{III.3}$$

and

$$\frac{1-\beta}{2D}\|\mathbf{x}\|_r^q \le \|\mathbf{Ax}\|_p^q \quad \text{for all } \mathbf{x} \in \Sigma_{2s}. \tag{III.4}$$

*Remark 3.3:* From Theorem 3.1, we see that a measurement matrix with small restricted isometry constant will have the sparse approximation property (I.2) with $p = r = 2, D$ close to one and $\beta$ close to zero. Conversely for $p = r = 2$ we obtain from Theorem 3.2 that if a measurement matrix $\mathbf{A}$ has the sparse approximation property (I.2) with $D$ close to one and $\beta$ close to zero, then the first inequality in the restricted isometry property (I.11) holds for some constant $\delta$ close to $1/2$ **only**. For $p = q = r = 1$, the $m \times n$ adjacency matrices $\mathbf{\Phi}$ of unbalanced $(2s, \alpha)$-expander with fixed left degree $d$ has the sparse approximation property (I.2) with small sparse approximation constant (see (I.8)) and the restricted isometry property with respect to $\ell^1$-norm:

$$(1-C\alpha)\|\mathbf{x}\|_1 \le \|\mathbf{\Phi x}\|_1/d \le (1+C\alpha)\|\mathbf{x}\|_1 \quad \text{for all } \mathbf{x} \in \Sigma_{2s}$$

where $C$ is a positive constant (see [12, Theorem 1]), but it does not have the restricted isometry property (I.11) when $m/s^2$ is sufficiently small [27].

*Remark 3.4:* If a measurement matrix $\mathbf{A}$ has the restricted isometry property (I.11) with small restricted isometry constant (see [1], [4], [24], [25], [26] for examples of such measurement matrices), then the preconditioned measurement matrix $\mathbf{PA}$ has the sparse approximation property (I.2) with $p = r = 2$, $D$ close to $\|\mathbf{P}^{-1}\|_{2\to 2}$ and $\beta$ close to zero but it does not have the restricted isometry property (I.11) in general. This observation may suggest that preconditioning procedure could generate new measurement matrices for the stable recovery of compressible signals from their noisy measurements.

## IV. Conclusions and final remarks

In this paper, we introduce the sparse approximation property (I.2) of a measurement matrix and show that it is a sufficient and almost necessary condition that any compressible signal can be stably recovered from its noisy measurements via solving the $\ell^q$-minimization problem (I.1).

The sparse approximation property (I.2) of a measurement matrix with $q \le r$ is a stronger version of the null space property (I.10) with the preconditioning-invariance almost preserved. The sparse approximation property (I.2) with $p = r = 2$ and $0 < q \le 1$ is a weaker version of the restricted isometry property (I.11). The adjacency matrices of some unbalanced expanders have the sparse approximation property (I.2) with $p = q = r = 1$ and small sparse approximation constant, but they do not have the restricted isometry property (I.11). A challenging problem is the construction of measurement matrices, other than random matrices [1], [4], [24], [25], [26] and adjacency matrices of a graph [12], [27], [28], [29], [30], that have sparse approximation property (I.2) with small sparse approximation constant.

## Appendix

### A. Proof of Theorem 1.1

Set $\mathbf{h} = \mathbf{x}^* - \mathbf{x}$. Let $S_0$ be so chosen that $\|\mathbf{x}_{S_0^c}\|_q = \|\mathbf{x} - \mathbf{x}_s\|_q$, $S_1$ be the set of indices of the $s$ largest components, in absolute value, of $\mathbf{h}$ in $S_0^c$, $S_2$ be the set of indices of the next

$s$ largest components, in absolute value, of $\mathbf{h}$ in $(S_0 \cup S_1)^c$, and so on. Then

$$\|\mathbf{A}\mathbf{h}\|_p \le 2\epsilon \quad \text{and} \quad \|\mathbf{h}_{S_0^c}\|_q^q \le \|\mathbf{h}_{S_0}\|_q^q + 2\|\mathbf{x}-\mathbf{x}_s\|_q^q \quad \text{(A.1)}$$

by (I.1), and

$$\|\mathbf{h}_{S_j}\|_{\tilde{r}} \le s^{1/\tilde{r}-1/q}\|\mathbf{h}_{S_{j-1}}\|_q, \quad j \ge 2 \quad \text{(A.2)}$$

by the construction of the sets $S_j, j \ge 1$, where $q \le \tilde{r} \le r$. Combining (I.2) and (A.1) gives

$$\|\mathbf{h}_T\|_r^q \le D(2\epsilon)^q + \beta s^{q/r-1}\|\mathbf{h}_{T^c}\|_q^q \quad \text{(A.3)}$$

for any subset $T$ of $\{1,\ldots,n\}$ with $\#T \le s$. Applying (A.3) with $T$ replaced by $S_0$ and then using the estimate (A.1) for $\|\mathbf{h}_{S_0^c}\|_q^q$,

$$\|\mathbf{h}_{S_0}\|_r^q \le D(2\epsilon)^q + 2\beta s^{q/r-1}\|\mathbf{x}-\mathbf{x}_s\|_q^q + \beta s^{q/r-1}\|\mathbf{h}_{S_0}\|_q^q. \quad \text{(A.4)}$$

By Hölder inequality and the property that $\#S_0 \le s$,

$$\|\mathbf{h}_{S_0}\|_q \le s^{1/q-1/r}\|\mathbf{h}_{S_0}\|_r. \quad \text{(A.5)}$$

Substituting the above inequality into the right-hand side of the inequality (A.4) leads to the first crucial inequality:

$$\|\mathbf{h}_{S_0}\|_r^q \le \frac{D}{1-\beta}(2\epsilon)^q + \frac{2\beta}{1-\beta}s^{q/r-1}\|\mathbf{x}-\mathbf{x}_s\|_q^q. \quad \text{(A.6)}$$

Combining (A.1), (A.5) and (A.6) yields the second crucial inequality:

$$\|\mathbf{h}_{S_0^c}\|_q^q \le \frac{D}{1-\beta}s^{1-q/r}(2\epsilon)^q + \frac{2}{1-\beta}\|\mathbf{x}-\mathbf{x}_s\|_q^q. \quad \text{(A.7)}$$

For $r = q$, the conclusion (I.5) follows from (A.6) and (A.7).

Applying (A.3) with $T$ replaced by $S_1$ yields

$$\|\mathbf{h}_{S_1}\|_r^q \le D(2\epsilon)^q + \beta s^{q/r-1}\|\mathbf{h}_{S_1^c}\|_q^q.$$

This together with (A.1), (A.5), (A.6) and (A.7) leads to the third crucial inequality:

$$\begin{aligned}\|\mathbf{h}_{S_1}\|_r^q &\le D(2\epsilon)^q + \beta s^{q/r-1}\big(\|\mathbf{h}_{S_0}\|_q^q + \|\mathbf{h}_{(S_0\cup S_1)^c}\|_q^q\big)\\ &\le \frac{D(1+\beta)}{1-\beta}(2\epsilon)^q + \frac{2\beta(1+\beta)}{1-\beta}s^{q/r-1}\|\mathbf{x}-\mathbf{x}_s\|_q^q. \end{aligned} \quad \text{(A.8)}$$

Therefore for $q \le \tilde{r} \le r$,

$$\begin{aligned}\|\mathbf{h}\|_{\tilde{r}}^q &\le \|\mathbf{h}_{S_0}\|_{\tilde{r}}^q + \|\mathbf{h}_{S_1}\|_{\tilde{r}}^q + \sum_{j\ge 2}\|\mathbf{h}_{S_j}\|_{\tilde{r}}^q\\ &\le s^{q/\tilde{r}-q/r}\|\mathbf{h}_{S_0}\|_r^q + s^{q/\tilde{r}-q/r}\|\mathbf{h}_{S_1}\|_r^q\\ &\quad + s^{q/\tilde{r}-1}\|\mathbf{h}_{S_0^c}\|_q^q\\ &\le \frac{D(3+\beta)}{1-\beta}s^{q/\tilde{r}-q/r}(2\epsilon)^q\\ &\quad + \frac{2(1+\beta)^2}{1-\beta}s^{q/\tilde{r}-1}\|\mathbf{x}-\mathbf{x}_s\|_q^q, \end{aligned} \quad \text{(A.9)}$$

where the first inequality holds by the triangle inequality for $\|\cdot\|_{q/\tilde{r}}^{q/\tilde{r}}$ as $q \le \tilde{r}$, the second inequality is true by Hölder inequality and (A.2), and the third inequality follows from (A.6), (A.7) and (A.8). Then the conclusions (I.3) and (I.4) follow by letting $\tilde{r} = r$ and $\tilde{r} = q$ in (A.9) respectively.

### B. Proof of Theorem 1.2

The conclusion (I.7) follows from the estimate (I.6) and the observation that the zero vector is the solution of the $\ell^q$-minimization problem (I.1) with $\epsilon = \|\mathbf{A}\mathbf{x}\|_p$ and $\mathbf{z} = \mathbf{A}\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$.

### C. Proof of Theorem 2.1

(i) Take a vector $\mathbf{x} \in \mathbb{R}^n$ with $\mathbf{A}\mathbf{x} = \mathbf{0}$ and let $\mathbf{x}_s$ be its best $s$-sparse approximation in $\ell^q$. Then it is a best $s$-sparse approximation in $\ell^r$. This together with the sparse approximation property (I.2) leads to $\|\mathbf{x}_s\|_q^q \le s^{1-q/r}\|\mathbf{x}_s\|_r^q \le \beta_s(\mathbf{A})(\sigma_{s,q}(\mathbf{x}))^q$. Thus the measurement matrix $\mathbf{A}$ has the null space property of order $s$ with $\gamma_s(\mathbf{A}) \le \beta_s(\mathbf{A})$.

(ii) Take a vector $\mathbf{x} \in \mathbb{R}^n$. Then it suffices to prove that

$$\begin{aligned}\|\mathbf{x}_T\|_q^q &\le \Big(\max(1,\gamma_s(\mathbf{A}))\inf_{\mathbf{R}\in\mathcal{R}(\mathbf{A})}\|\mathbf{R}\|_{p\to q}^q\Big)\|\mathbf{A}\mathbf{x}\|_p^q\\ &\quad + \gamma_s(\mathbf{A})\|\mathbf{x}_{T^c}\|_q^q, \end{aligned} \quad \text{(A.10)}$$

for all subsets $T \subset \{1,\ldots,n\}$ with $\#T \le s$, where $0 < p \le \infty$. Note that $\mathbf{A}(\mathbf{x} - \mathbf{R}\mathbf{A}\mathbf{x}) = (\mathbf{A} - \mathbf{A}\mathbf{R}\mathbf{A})\mathbf{x} = \mathbf{0}$ for all $\mathbf{R} \in \mathcal{R}(\mathbf{A})$. This together with the null space property (I.10) of the measurement matrix $\mathbf{A}$ leads to $\|(\mathbf{x}-\mathbf{R}\mathbf{A}\mathbf{x})_T\|_q^q \le \gamma_s(\mathbf{A})\|(\mathbf{x}-\mathbf{R}\mathbf{A}\mathbf{x})_{T^c}\|_q^q$ for all subsets $T \subset \{1,\ldots,n\}$ with $\#T \le s$ and $\mathbf{R} \in \mathcal{R}(\mathbf{A})$. Hence

$$\|\mathbf{x}_T\|_q^q \le \max(1,\gamma_s(\mathbf{A}))\|\mathbf{R}\|_{p\to q}^q\|\mathbf{A}\mathbf{x}\|_p^q + \gamma_s(\mathbf{A})\|\mathbf{x}_{T^c}\|_q^q.$$

Taking minimum over all matrices $\mathbf{R} \in \mathcal{R}(\mathbf{A})$ in the right-hand side of the above estimate leads to (A.10), and hence proves the second conclusion.

### D. Proof of Theorem 3.1

Take a vector $\mathbf{x} \in \mathbb{R}^n$ and let $\mathbf{x}_s$ be its $s$-sparse approximation in $\ell^2$. We write $\mathbf{x} = \sum_{j\ge 0}\mathbf{x}_{S_j}$, where $S_0$ is the set of indices of the $s$ largest components, in absolute value, of $\mathbf{x}$, $S_1$ is the set of indices of the $s$ largest components, in absolute value, of $\mathbf{x}$ in $S_0^c$, and so on. From the construction of the sets $S_j, j \ge 0$, we obtain that $\mathbf{x}_{S_0} = \mathbf{x}_s$, $\|\mathbf{x}_{S_0^c}\|_q = \sigma_{s,q}(\mathbf{x})$, $\sum_{j\ge 2}\|\mathbf{x}_{S_j}\|_2 \le s^{1/2-1/q}\sigma_{s,q}(\mathbf{x})$,

$$\begin{aligned}\|\mathbf{x}_{S_j}\|_2 &\le s^{1/2-1/q}\|\mathbf{x}_{S_{j-1}}\|_q^{1-q/2}\|\mathbf{x}_{S_j}\|_q^{q/2}\\ &\le s^{1/2-1/q}\|\mathbf{x}_{S_{j-1}}\|_q \end{aligned} \quad \text{(A.11)}$$

for all $j \ge 1$, and

$$\begin{aligned}\|\mathbf{A}\mathbf{x}\|_2^2 &= \|\mathbf{A}(\mathbf{x}_{S_0}+\mathbf{x}_{S_1})\|_2^2 + \sum_{j\ge 2}\|\mathbf{A}\mathbf{x}_{S_j}\|_2^2\\ &\quad + 2\sum_{j\ge 2}\langle \mathbf{A}\mathbf{x}_{S_0}, \mathbf{A}\mathbf{x}_{S_j}\rangle + 2\sum_{j\ge 2}\langle \mathbf{A}\mathbf{x}_{S_1}, \mathbf{A}\mathbf{x}_{S_j}\rangle\\ &\quad + 2\sum_{2\le j<j'}\langle \mathbf{A}\mathbf{x}_{S_j}, \mathbf{A}\mathbf{x}_{S_{j'}}\rangle. \end{aligned} \quad \text{(A.12)}$$

Recalling that $|\langle \mathbf{A}\mathbf{x}_{S_j}, \mathbf{A}\mathbf{x}_{S_{j'}}\rangle| \le \delta_{2s}(\mathbf{A})\|\mathbf{x}_{S_j}\|_2\|\mathbf{x}_{S_{j'}}\|_2$ for all $j' \ne j$ ([1]), and applying the restricted isometry property (I.11), we obtain from (A.11) and (A.12) that

$$\begin{aligned}(1-\delta_{2s}(\mathbf{A}))\|\mathbf{x}\|_2^2 &\le \|\mathbf{A}\mathbf{x}\|_2^2 + \delta_{2s}(\mathbf{A})s^{1-2/q}(\sigma_{s,q}(\mathbf{x}))^2\\ &\quad + 2\sqrt{2}\delta_{2s}(\mathbf{A})s^{1/2-1/q}\|\mathbf{x}\|_2\sigma_{s,q}(\mathbf{x})\\ &\le \|\mathbf{A}\mathbf{x}\|_2^2 + \delta_{2s}(\mathbf{A})\epsilon\|\mathbf{x}\|_2^2\\ &\quad + \delta_{2s}(\mathbf{A})(1+2\epsilon^{-1})s^{1-2/q}(\sigma_{s,q}(\mathbf{x}))^2, \end{aligned}$$

where $\epsilon > 0$. Then (III.1) follows by taking $\epsilon = -2 + \sqrt{2(\delta_{2s}(\mathbf{A}))^{-1} + 2}$.

Similarly we get

$$\|\mathbf{A}\mathbf{x}\|_2^2 \leq \big(1 + \delta_{2s}(\mathbf{A}) + \sqrt{2\delta_{2s}(\mathbf{A})}\big)\|\mathbf{x}\|_2^2 + \big(1 + \sqrt{2\delta_{2s}(\mathbf{A})}\big)\delta_{2s}(\mathbf{A})s^{1-2/q}(\sigma_{s,q}(\mathbf{x}))^2.$$

This proves (III.2).

### E. Proof of Theorem 3.2

Take an $s$-sparse vector $\mathbf{x} \in \mathbb{R}^n$. Then $\mathbf{x} = \mathbf{x}_s$ and $\sigma_{s,q}(\mathbf{x}) = 0$. This together with the sparse approximation property (I.2) gives $\|\mathbf{x}\|_r^q = \|\mathbf{x}_s\|_r^q \leq D\|\mathbf{A}\mathbf{x}\|_p^q + \beta s^{q/p-1}\sigma_{s,q}(\mathbf{x})^q = D\|\mathbf{A}\mathbf{x}\|_p^q$, and hence proves (III.3).

Take a $2s$-sparse vector $\mathbf{x} \in \mathbb{R}^n$ and write $\mathbf{x} = \mathbf{x}_{S_0} + \mathbf{x}_{S_1}$ for some subsets $S_0$ and $S_1$ of $\{1, \ldots, n\}$ with empty intersection and cardinality less than or equal to $s$. Applying (I.2) to the given $2s$-sparse vector $\mathbf{x}$ and replacing $S$ by $S_0$ and $S_1$ respectively, we obtain

$$\|\mathbf{x}_{S_0}\|_r^q \leq D\|\mathbf{A}\mathbf{x}\|_p^q + \beta s^{q/p-1}\|\mathbf{x}_{S_1}\|_q^q \leq D\|\mathbf{A}\mathbf{x}\|_p^q + \beta\|\mathbf{x}_{S_1}\|_r^q \quad \text{(A.13)}$$

and

$$\|\mathbf{x}_{S_1}\|_r^q \leq D\|\mathbf{A}\mathbf{x}\|_p^q + \beta s^{q/p-1}\|\mathbf{x}_{S_0}\|_q^q \leq D\|\mathbf{A}\mathbf{x}\|_p^q + \beta\|\mathbf{x}_{S_0}\|_r^q. \quad \text{(A.14)}$$

Summing up the above estimates (A.13) and (A.14) yields the following inequality:

$$\begin{aligned}(1-\beta)\|\mathbf{x}\|_r^q &= (1-\beta)(\|\mathbf{x}_{S_0}\|_r^r + \|\mathbf{x}_{S_1}\|_r^r)^{q/r} \\ &\leq (1-\beta)(\|\mathbf{x}_{S_0}\|_r^q + \|\mathbf{x}_{S_1}\|_r^q) \leq 2D\|\mathbf{A}\mathbf{x}\|_p^q.\end{aligned}$$

Hence (III.4) follows.

**Acknowledgement.** Part of this work was done when the author visited Ecole Polytechnique Federale de Lausanne. The author would like to thank Professors Michael Unser and Martin Vetterli for the hospitality and fruitful discussion, and also Professor Namrata Vaswani and anonymous reviewers for providing useful remarks and additional references, and for pointing out that the adjacency matrices of unbalanced expanders have the sparse approximation property.

## REFERENCES

[1] E. J. Candes and T. Tao, Decoding by linear programming, *IEEE Trans. Inform. Theory*, **51**(2005), 4203–4215.

[2] E. J. Candes, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math.*, **59**(2006), 1207–1223.

[3] E. J. Candes, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory*, **52**(2006), 489–509.

[4] D. Donoho, Compressive sampling, *IEEE Trans. Inform. Theory*, **52**(2006), 1289–1306.

[5] R. Chartrand, Exact reconstruction of sparse signals via nonconvex minimization, *IEEE Signal Proc. Letter*, **14**(2007), 707–710.

[6] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano and L. Coulot, Sparse sampling of signal innovations, *IEEE Signal Processing Magazine*, **25**(2008), 31–40.

[7] A. Cohen, W. Dahmen and R. DeVore, Compressive sensing and best $k$-term approximation, *J. Amer. Math. Soc.*, **22**(2009), 211–231.

[8] I. Dauchebies, R. DeVore, M. Fornasier, and C. S. Gunturk, Iteratively re-weighted least squares minimization for sparse recovery, *Commun. Pure Appl. Math.*, **63**(2010), 1–38.

[9] Y. Zhang A simple proof for recoverability of $\ell_1$-minimization: go over or under, Technical Report TR05-09, Rice University, http://www.caam.rice.edu/∼yzhang/reports/tr0509.pdf

[10] E. J. Candes, The restricted isometry property and its implications for compressed sensing, *C. R. Acad. Sci. Paris, Ser. I*, **346**(2008), 589–592.

[11] W. Xu and B. Hassibi, Compressed sensing over the Grassmann manifold: A unified analytical framework, In *46th Annual Allerton Conference on Communication, Control, and Computing*, 2008, pp. 562–567.

[12] R. Berinde, A. C. Gilbert, P. Indyk, H. Karloff, and M. J. Strauss, Combining geometry and combinatorics: A unified approach to sparse signal recovery, In *46th Annual Allerton Conference on Communication, Control, and Computing*, 2009, pp. 798–805.

[13] S. Foucart and M.-J. Lai, Sparsest solutions of underdetermined linear system via $\ell_q$-minimization for $0 < q \leq 1$, *Appl. Comput. Harmonic Anal.*, **26**(2009), 395–407.

[14] Y. Zhang, Theory of compressive sensing via $\ell_1$-minimization: A non-RIP analysis and extensions, Technical Report TR08-11 revised, Rice University, 2009.

[15] T. T. Cai, L. Wang and G. Xu, Shifting inequality and recovery of sparse signals, *IEEE Trans. Signal Process.*, **58**(2010), 1300–1308.

[16] S. Foucart, A note on guaranteed sparse recovery via $\ell_1$-minimization, *Appl. Comput. Harmonic Anal.*, **29**(2010), 97–103.

[17] Q. Sun, Recovery of sparse signals via $\ell^q$-minimization, *Appl. Comput. Harmonic Anal.*, revised. arXiv:1005.0267.

[18] D. Donoho and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via $\ell^1$ norm minimization, *Proc. Nat. Acad. Sci. USA*, **100**(2003), 2197–2002.

[19] R. Gribonval and M. Nielsen, Sparse representations in unions of bases, *IEEE Trans. Inform. Theory*, **49**(2003), 3320–3325.

[20] R. Chartrand and V. Staneva, Restricted isometry properties and non-convex compressive sensing, *Inverse Problems*, **24**(2008), 035020 (14 pp).

[21] J. A. Troppa, On the conditioning of random subdictionaries, *Appl. Comput. Harmonic Anal.*, **25**(2008), 1–24.

[22] E. J. Candes and Y. Plan, Near-ideal model selection by $l^1$ minimization, *Ann. Statist.*, **37**(2009), 2145–2177.

[23] S. Gurevich and R. Hadani, The statistical restricted isometry property and the Wigner semicircle distribution of incoherent dictionaries, Preprint 2009.

[24] R. Calderbank, S. Howard, S. and S. Jafarpour, Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property, *IEEE Journal of Selected Topics in Signal Processing*, **4**(2010), 358–374.

[25] H. Rauhut, Stability results for random sampling of sparse trigonometrical polynomials, *IEEE Trans. Inform. Theory*, **54**(2008), 5661–5670.

[26] M. Rudelson and R. Vershynin, On sparse reconstruction from Fourier and Gaussian measurements, *Comm. Pure Appl. Math.*, **61**(2008), 1025–1045.

[27] V. Chander, A negative result concerning explicit matrices with the restricted isometry property, Preprint 2008.

[28] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson, Randomness conductors and constant-degree lossless expanders, *Proceedings of 17th IEEE Annual Conference on Computational Complexity*, 2002, pp. 8.

[29] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank, Efficient and robust compressed sensing using optimized expander graphs, *IEEE Trans. Inform. Theory*, **55**(2009), 4299–4308.

[30] K. D. Ba, P. Indyk, E. Price and D. P. Woodruff, Lower bounds for sparse recovery, In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, Session 9A, Jan. 17-19, 2010, Hyatt Regency Austin, Austin, TX*, pp. 1190–1197.

PLACE PHOTO HERE

**Qiyu Sun** received the BSc and PhD degree in mathematics from Hangzhou University, China in 1985 and 1990 respectively. He is currently with Department of Mathematics, University of Central Florida. His prior position was with Zhejiang University (China), National University of Singapore (Singapore), Vanderbilt University, and University of Houston. His research interest includes sampling theory, Wiener's lemma, multi-band wavelets, frame theory, linear and nonlinear inverse problems, and Fourier analysis. He has published more than 100 papers on mathematics and signal processing, and written a book "*An Introduction to Multiband Wavelets*" (Zhejiang University Press, 2001) with Ning Bi and Daren Huang. He is on the editorial board of the journals "*Advance in Computational Mathematics*" and "*Numerical Functional Analysis and Optimization*".