Sparse Recovery with Pre-Gaussian Random Matrices

Simon Foucart*
Laboratoire Jacques-Louis Lions
Université Pierre et Marie Curie
Paris, 75013, France

Ming-Jun Lai†
Department of Mathematics
University of Georgia
Athens, GA 30602, USA

August 7, 2009

Abstract

We show that a matrix whose entries are independent copies of a symmetric pre-Gaussian random variable possesses, with overwhelming probability, a Modified Restricted Isometry Property in q-quasinorms for $0 < q \leq 1/3$. We then prove that, if the matrix of an underdetermined linear system of equations satisfies this property, then the sparsest solution of the system can be found using ℓ_q-minimization.

1 Introduction

The field of Compressed Sensing, which has generated a wealth of research activity in recent years, asks for some concrete protocols that make it possible to reconstruct sparse vectors $\mathbf{x} \in \mathbb{R}^N$ from the mere knowledge of measurement vectors $\mathbf{y} = A\mathbf{x} \in \mathbb{R}^m$ with $m \ll N$. In other words, one seeks $m \times N$ measurement matrices A and reconstruction algorithms that enable to find the sparsest solutions of the underdetermined linear system $A\mathbf{x} = \mathbf{y}$. The groundbreaking works of Donoho [7] and of Candès and Tao [5] successfully tackled these questions. The problem of choosing suitable matrices was settled using probabilistic arguments, with the conclusion that most matrices chosen at random allow for an efficient reconstruction of sparse vectors. The reconstruction in question consists in solving the convex optimization problem

$$\underset{\mathbf{z} \in \mathbb{R}^N}{\text{minimize}} \ \|\mathbf{z}\|_1 \quad \text{subject to } A\mathbf{z} = \mathbf{y}, \tag{P$_1$}$$

in place of the unpractical combinatorial problem

$$\underset{\mathbf{z} \in \mathbb{R}^N}{\text{minimize}} \ \|\mathbf{z}\|_0 \quad \text{subject to } A\mathbf{z} = \mathbf{y}. \tag{P$_0$}$$

Here $\|\mathbf{z}\|_1 = \sum_{j=1}^N |z_j|$ stands for the usual ℓ_1-norm of a vector $\mathbf{z} \in \mathbb{R}^N$, while $\|\mathbf{z}\|_0$ represents its sparsity, i.e. the number of its nonzero components. A much favored tool

*foucart@ann.jussieu.fr, This work was initiated when the author was at Vanderbilt University.
†mjlai@math.uga.edu, This research is supported by the National Science Foundation under grant #DMS 0713807.

in the study of the equivalence between (P_0) and (P_1) was introduced by Candès and Tao in [5]. It is said that an $m \times N$ matrix A has the t-th order Restricted Isometry Property if there is a constant $0 \le \delta < 1$ such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|A\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \le t. \tag{1}$$

The smallest such constant is denoted by δ_t and is called the t-th order Restricted Isometry Constant of the matrix A. There are many conditions on the δ_t's that guarantee the recovery of all s-sparse vectors $\mathbf{x} \in \mathbb{R}^N$ as solutions of (P_1) with $\mathbf{y} = A\mathbf{x}$. The arguably most natural ones are only in terms of δ_{2s}. For instance, Candès established the sufficient condition $\delta_{2s} < \sqrt{2} - 1 \approx 0.4142$ in [3], and this was latter improved by the present authors to $\delta_{2s} < 2/(3+\sqrt{2}) \approx 0.4531$ in [8]. Regardless of the sufficient condition called upon, the crucial point is that it is met with overwhelming probability for random matrices whose number m of rows, much smaller than its number N of columns, scales like the sparsity s times the log factor $\ln(N/s)$. The original article [5] dealt with Gaussian random matrices. Since then, many clarifications and extensions have been put forward. As an example, the paper [1] provides a simple proof of the Restricted Isometry Property for random matrices satisfying a concentration inequality (see [10] for some background on measure concentration theory). As another example, the paper [11] establishes the Restricted Isometry Property for sub-Gaussian random matrices.

In this paper, we modify slightly the ℓ_1-minimization strategy to consider instead an ℓ_q-minimization problem with $0 < q \le 1$, namely

$$\underset{\mathbf{z} \in \mathbb{R}^N}{\text{minimize}} \; \|\mathbf{z}\|_q \quad \text{subject to } A\mathbf{z} = \mathbf{y}. \tag{P_q}$$

Although this optimization problem has the drawback of not being convex anymore, the ℓ_q-minimization strategy offer some theoretical advantages. The first one is that the success of the ℓ_q-recovery increases as q decreases. This fact, intuitively anticipated since $\|\mathbf{z}\|_q^q$ approaches the sparsity $\|\mathbf{z}\|_0$ as q decreases to zero, was proved by Gribonval and Nielsen in [9]. Along the same lines, the sufficient conditions for ℓ_q-success that are available in the literature become weaker when q decreases, see e.g. [8]. The ℓ_q-strategy offers a second advantage in that it requires less measurements. Chartrand and Staneva indeed showed in [6] that, for an $m \times N$ Gaussian random matrix, the recovery of s-sparse vectors by ℓ_q-minimization is guaranteed as soon as

$$m \ge c_1\, s + c_2\, q\, s\, \ln(N/s),$$

where c_1 and c_2 are absolute positive constants. In their proof, they used an ℓ_q-variant of the Restricted Isometry Property, namely

$$(1-\delta)\|\mathbf{x}\|_2^q \le \|A\mathbf{x}\|_q^q \le (1+\delta)\|\mathbf{x}\|_2^q \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \le t. \tag{2}$$

A third advantage of the ℓ_q-strategy is demonstrated in this paper, namely that it applies to a wider class of random matrices. We will indeed prove the success of the ℓ_q-strategy for $q \le 1/3$ when the matrix A is populated by independent copies of a symmetric pre-Gaussian random variable. Let us recall, see e.g. [2], that a random variable ξ is pre-Gaussian if and only if $\mathbf{E}(\xi) = 0$ and there is a constant $b > 0$ such that

$$\mathbf{E}\big(|\xi|^k\big) \leq k! b^k \qquad \text{for all integers } k \geq 1.$$

For instance, a special emphasis can be put on Laplace random variables ξ with probability density functions

$$f(t) = \frac{1}{2\lambda} \exp\left(-\frac{|t|}{\lambda}\right), \qquad \lambda > 0,$$

which satisfy $\mathbf{E}\big(|\xi|^k\big) = k!\lambda^k$ for all integers $k \geq 1$, and more generally

$$\mathbf{E}\big(|\xi|^r\big) = \Gamma(r+1)\lambda^r \qquad \text{for all } r > 0.$$

Our main theorem reads as follows.

Theorem 1.1 *Suppose that the entries of an $m \times N$ matrix A are independent copies of a symmetric pre-Gaussian random variable. If $0 < q \leq 1/3$, then the probability that every s-sparse vector $\mathbf{x} \in \mathbb{R}^N$ is recovered as a solution of (P_q) with $\mathbf{y} = A\mathbf{x}$ exceeds $1 - 2\exp(-c_1 m)$, provided that $m \geq c_2 s \ln(N/s)$. The constants c_1 and c_2 depend only on the pre-Gaussian distribution.*

Here is a rapid outline of our arguments, which essentially relies on a variation of the classical Restricted Isometry Property (1). Throughout the paper, the probability density function of a pre-Gaussian random variable is denoted by f, and we require it to be an even function. For $p > 0$ and $\mathbf{x} \in \mathbb{R}^N$, we introduce the quantity

$$/\!\!/\mathbf{x}/\!\!/_{f,p} := \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big|\sum_{j=1}^{N} t_j x_j\Big|^p f(t_1)\cdots f(t_N)\, dt_1 \ldots dt_N\right]^{1/p}. \tag{3}$$

We shall see that it turns out to be a norm for $p \geq 1$ and a quasinorm for $0 < p < 1$. Note that we prefer the notation q if the exponent p is less than one. We will say that an $m \times N$ matrix A satisfies the t-th order p-Modified Restricted Isometry Property if there exists a constant $0 \leq \delta < 1$ such that

$$(1-\delta)m \,/\!\!/\, \mathbf{x}/\!\!/_{f,p}^p \leq \|A\mathbf{x}\|_p^p \leq (1+\delta)m \,/\!\!/\, \mathbf{x}\,/\!\!/_{f,p}^p \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \leq t. \tag{4}$$

When $p = 2$, we will see that this is nothing else than a scaled version of the classical Restricted Isometry Property (1). When the function f is a Gaussian probability density function, we will also see that the q-Modified Restricted Isometry Property for $0 < q \leq 1$ coincides with the variation (2) of the Restricted Isometry Property used by Chartrand and Staneva. An important role will be played by the upper and lower constants c_p and C_p in the inequality

$$c_p \|\mathbf{x}\|_2^p \leq /\!\!/\mathbf{x}/\!\!/_{f,p}^p \leq C_p \|\mathbf{x}\|_2^p,$$

which will be established in Section 2. This section contains other lemmas that are to be used later. We then prove in Section 3 that, under some conditions on the constants c_p and C_p, obviously satisfied in our case of interest, the tail probability

$$\mathbf{P}\Big(\big|\|A\mathbf{x}\|_p^p - m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,p}^p\big| > \epsilon m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,p}^p\Big)$$

decays exponentially fast with m for a fixed $\mathbf{x} \in \mathbb{R}^N$. In Section 4, we show that this tail estimate implies the q-Modified Restricted Isometry Property (4). We complete the argument in Section 5 by proving that a matrix satisfying the q-Modified Restricted Isometry Property allows for the exact reconstruction of sparse vector by ℓ_q-minimization.

2 Preliminaries

We first establish some basic properties of the quantity introduced in (3). These involve the absolute moments of the probability density function f defined as

$$\sigma_r := \int_{-\infty}^{\infty} |t|^r f(t)dt, \qquad r > 0.$$

Lemma 2.1 *Let f be an even probability density function. For $p \geq 1$, the expression* (3) *defines a norm on $\mathbb{R}^N$ provided $\sigma_p < \infty$. For $0 < p < 1$, it defines a quasinorm on $\mathbb{R}^N$ provided $\sigma_1 < \infty$. In this case, one has $/\!\!/\mathbf{x}+\mathbf{y}/\!\!/_{f,p}^p \leq /\!\!/\mathbf{x}/\!\!/_{f,p}^p + /\!\!/\mathbf{y}/\!\!/_{f,p}^p$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$.*

Proof. We first show that $/\!\!/\mathbf{x}/\!\!/_{f,p}$ is well-defined for each vector $\mathbf{x} \in \mathbb{R}^N$. For $p \geq 1$, we use the convexity of the function $t \mapsto t^p$ to write

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p}^p &\leq \|\mathbf{x}\|_1^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\sum_{j=1}^N |t_j| \frac{|x_j|}{\|\mathbf{x}\|_1}\right]^p f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N \\
&\leq \|\mathbf{x}\|_1^p \sum_{j=1}^N \frac{|x_j|}{\|\mathbf{x}\|_1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |t_j|^p f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N \\
&= \|\mathbf{x}\|_1^p \sigma_p^p < \infty.
\end{aligned}$$

For $0 < p < 1$, we use the increase with r of the L_r-quasinorm to write

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p} &\leq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big|\sum_{j=1}^N t_j x_j\Big| f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N \\
&\leq \sum_{j=1}^N |x_j| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |t_j| f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N \\
&= \|\mathbf{x}\|_1 \sigma_1 < \infty.
\end{aligned}$$

Next, given $p > 0$, it is readily seen that $/\!\!/\mathbf{x}/\!\!/_{f,p} = 0$ if and only if $\mathbf{x} = 0$ and that $/\!\!/\lambda\mathbf{x}/\!\!/_{f,p} = |\lambda| \,/\!\!/\, \mathbf{x}/\!\!/_{f,p}$ for all $\lambda \in \mathbb{R}$ and all $\mathbf{x} \in \mathbb{R}^N$. It remains to prove that, if $p \geq 1$, one has $/\!\!/\mathbf{x}+\mathbf{y}/\!\!/_{f,p} \le /\!\!/\mathbf{x}\,/\!\!/_{f,p} + /\!\!/\,\mathbf{y}/\!\!/_{f,p}$ for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^N$, and that, if $0 < p < 1$, one has $/\!\!/\mathbf{x}+\mathbf{y}/\!\!/_{f,p}^p \le /\!\!/\mathbf{x}\,/\!\!/_{f,p}^p + /\!\!/\,\mathbf{y}/\!\!/_{f,p}^p$. If $p \ge 1$, the result follows by applying the triangle inequality for the L_p-norm in the expression of $/\!\!/\mathbf{x}+\mathbf{y}/\!\!/_{f,p}$. If $0 < p < 1$, it follows by applying the triangle inequality for the p-th power of the ℓ_p-quasinorm in the expression of $/\!\!/\mathbf{x}+\mathbf{y}/\!\!/_{f,p}^p$. ■

The following generalization of [13, Lemma 4.10] applies to $\|\cdot\|_q$ and $/\!\!/\cdot/\!\!/_{f,q}$ alike.

Lemma 2.2 *For $0 < q \le 1$, let $|\cdot|_q$ be a quasinorm on $\mathbb{R}^n$ satisfying $|\mathbf{x}+\mathbf{y}|_q^q \le |\mathbf{x}|_q^q + |\mathbf{y}|_q^q$ for all $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$. With $\mathcal{S}_q := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}|_q = 1\}$ denoting the unit sphere of $\mathbb{R}^n$ relative to this quasinorm, for any $\delta > 0$, there exists a finite set $\mathcal{U}_q \subseteq \mathcal{S}_q$ with*

$$\min_{\mathbf{u}\in\mathcal{U}_q} |\mathbf{x}-\mathbf{u}|_q^q \le \delta \quad \text{for all } \mathbf{x} \in \mathcal{S}_q \qquad \text{and} \qquad \operatorname{card}(\mathcal{U}_q) \le \left(1+\frac{2}{\delta}\right)^{n/q}.$$

Proof. Let $(\mathbf{u}_1,\ldots,\mathbf{u}_k)$ be a set of k points on the sphere $\mathcal{S}_q$ such that $|\mathbf{u}_i-\mathbf{u}_j|_q^q > \delta$ for all $i \ne j$. We choose k as large as possible. Thus, it is clear that

$$\min_{1\le i\le k} |\mathbf{x}-\mathbf{u}_i|_q^q \le \delta \quad \text{for all } \mathbf{x} \in \mathcal{S}_q.$$

Let $\mathcal{B}_q := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}|_q \le 1\}$ be the unit ball of $\mathbb{R}^n$ relative to the quasinorm $|\cdot|_q$. It is easy to see that the $(\delta/2)$-balls centered at $\mathbf{u}_i$,

$$\mathbf{u}_i + \left(\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q, \quad 1 \le i \le k,$$

are disjoint. Indeed, if $\mathbf{x}$ would belong to the $(\delta/2)$-ball centered at $\mathbf{u}_i$ as well as the $(\delta/2)$-ball centered at $\mathbf{u}_j$, we would have

$$|\mathbf{u}_i-\mathbf{u}_j|_q^q \le |\mathbf{u}_i-\mathbf{x}|_q^q + |\mathbf{u}_j-\mathbf{x}|_q^q \le \frac{\delta}{2}+\frac{\delta}{2} = \delta,$$

which is a contradiction. Besides, it is easy to see that

$$\mathbf{u}_i + \left(\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q \subseteq \left(1+\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q, \quad 1 \le i \le k.$$

By comparison of volumes, we get

$$k\mathrm{Vol}\bigg(\left(\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q\bigg) = \sum_{i=1}^{k}\mathrm{Vol}\bigg(\mathbf{u}_i + \left(\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q\bigg) \le \mathrm{Vol}\bigg(\left(1+\frac{\delta}{2}\right)^{1/q}\mathcal{B}_q\bigg).$$

Then, by homogeneity of the volumes, we have

$$k\left(\frac{\delta}{2}\right)^{n/q}\mathrm{Vol}\left(\mathcal{B}_q\right) \le \left(1+\frac{\delta}{2}\right)^{n/q}\mathrm{Vol}\left(\mathcal{B}_q\right),$$

which implies that $k \le \left(1+\frac{2}{\delta}\right)^{n/q}$. This completes the proof. ■

Next, we want to compare the (quasi)norm $/\!\!/\mathbf{x}/\!\!/_{f,p}$ with the standard ℓ_2-norm $\|\mathbf{x}\|_2$. This is made possible by Khintchine's inequality with optimal constants, see e.g. [12].

Theorem 2.1 *Given $p > 0$, for any integer N and any vector $\mathbf{x} \in \mathbb{R}^N$, one has*

$$A_p \Big[\sum_{j=1}^N x_j^2\Big]^{p/2} \;\le\; \frac{1}{2^N} \sum_{\epsilon_1,\dots,\epsilon_N=\pm 1} \Big|\sum_{j=1}^N \epsilon_j x_j\Big|^p \;\le\; B_p \Big[\sum_{j=1}^N x_j^2\Big]^{p/2},$$

where

$$A_p := 2^{(p-2)/2} \min\left(1, \frac{\Gamma\big((p+1)/2\big)}{\Gamma\big(3/2\big)}\right), \qquad B_p := 2^{(p-2)/2} \max\left(1, \frac{\Gamma\big((p+1)/2\big)}{\Gamma\big(3/2\big)}\right).$$

Lemma 2.3 *Suppose that f is an even probability density function with bounded absolute moments. Given $p > 0$, there exist positive constants c_p and C_p such that, for all $\mathbf{x} \in \mathbb{R}^N$,*

$$c_p \|\mathbf{x}\|_2^p \le /\!\!/\mathbf{x}/\!\!/_{f,p}^p \le C_p \|\mathbf{x}\|_2^p.$$

Proof. We notice first that

$$\begin{aligned} /\!\!/\mathbf{x}/\!\!/_{f,p}^p &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big|\sum_{j=1}^N t_j x_j\Big|^p f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N \\ &= \sum_{\epsilon_1,\dots,\epsilon_N=\pm 1} \epsilon_1 \cdots \epsilon_N \int_0^{\epsilon_1 \infty} \cdots \int_0^{\epsilon_N \infty} \Big|\sum_{j=1}^N t_j x_j\Big|^p f(t_1)\cdots f(t_N)\, dt_1 \dots dt_N. \end{aligned}$$

Then using the symmetry of the probability density function f, we obtain

$$\begin{aligned} /\!\!/\mathbf{x}/\!\!/_{f,p}^p &= \sum_{\epsilon_1,\dots,\epsilon_N=\pm 1} \int_0^{\infty} \cdots \int_0^{\infty} \Big|\sum_{j=1}^N \epsilon_j u_j x_j\Big|^p f(u_1)\cdots f(u_N)\, du_1 \dots du_N \\ &= 2^N \int_0^{\infty} \cdots \int_0^{\infty} \frac{1}{2^N} \sum_{\epsilon_1,\dots,\epsilon_N=\pm 1} \Big|\sum_{j=1}^N \epsilon_j u_j x_j\Big|^p f(u_1)\cdots f(u_N)\, du_1 \dots du_N. \end{aligned}$$

Calling upon Khintchine's inequality, we have

$$/\!\!/\mathbf{x}\,/\!\!/_{f,p}^p \begin{cases} \ge\; A_p\, 2^N \displaystyle\int_0^{\infty} \cdots \int_0^{\infty} \Big[\sum_{j=1}^N u_j^2 x_j^2\Big]^{p/2} f(u_1)\cdots f(u_N)\, du_1 \dots du_N, \\ \quad = A_p \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big[\sum_{j=1}^N u_j^2 x_j^2\Big]^{p/2} f(u_1)\cdots f(u_N)\, du_1 \dots du_N, \\ \le\; B_p\, 2^N \displaystyle\int_0^{\infty} \cdots \int_0^{\infty} \Big[\sum_{j=1}^N u_j^2 x_j^2\Big]^{p/2} f(u_1)\cdots f(u_N)\, du_1 \dots du_N, \\ \quad = B_p \displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big[\sum_{j=1}^N u_j^2 x_j^2\Big]^{p/2} f(u_1)\cdots f(u_N)\, du_1 \dots du_N. \end{cases}$$

Let us suppose first that $p \geq 2$. For the lower estimate, we use the increase with r of the L_r-norm to derive

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p}^{p} &\geq A_p \Big[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{j=1}^{N} u_j^2 x_j^2 \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \Big]^{p/2} \\
&= A_p \Big[\sum_{j=1}^{N} x_j^2 \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u_j^2 \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \Big]^{p/2} \\
&= A_p \Big[\sum_{j=1}^{N} x_j^2 \, \sigma_2 \Big]^{p/2} = A_p \, \sigma_2^{p/2} \, \|\mathbf{x}\|_2^p.
\end{aligned}$$

For the upper estimate, we use the convexity of the function $t \mapsto t^{p/2}$ to derive

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p}^{p} &\leq B_p \, \|\mathbf{x}\|_2^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big[\sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \, u_j^2 \Big]^{p/2} f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \\
&\leq B_p \, \|\mathbf{x}\|_2^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \, |u_j|^p \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \\
&= B_p \, \|\mathbf{x}\|_2^p \sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |u_j|^p \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \\
&= B_p \, \|\mathbf{x}\|_2^p \sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \, \sigma_p = B_p \, \sigma_p \, \|\mathbf{x}\|_2^p.
\end{aligned}$$

Let us now suppose that $p \leq 2$. For the lower estimate, we use the concavity of the function $t \mapsto t^{p/2}$ to derive

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p}^{p} &\geq A_p \, \|\mathbf{x}\|_2^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big[\sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \, u_j^2 \Big]^{p/2} f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \\
&\geq A_p \, \|\mathbf{x}\|_2^p \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{j=1}^{N} \frac{x_j^2}{\|\mathbf{x}\|_2^2} \, |u_j|^p \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \\
&= A_p \, \sigma_p \, \|\mathbf{x}\|_2^p.
\end{aligned}$$

For the upper estimate, we use the increase with r of the L_r-norm to derive

$$\begin{aligned}
/\!\!/\mathbf{x}/\!\!/_{f,p}^{p} &\leq B_p \Big[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \sum_{j=1}^{N} u_j^2 x_j^2 \, f(u_1) \cdots f(u_N) \, du_1 \ldots du_N \Big]^{p/2} \\
&= B_p \, \sigma_2^{p/2} \, \|\mathbf{x}\|_2^p.
\end{aligned}$$

The proof is now complete. ■

Remark 2.1 *Using* $\Gamma(3/2) = \sqrt{\pi}/2$*, we notice that the constants* c_p *and* C_p *are explicitly given by*

$$c_p = \begin{cases} A_p\,\sigma_2^{p/2} & = & \dfrac{2^{p/2}\sigma_2^{p/2}}{\sqrt{\pi}}\min\big(\Gamma(3/2),\Gamma((p+1)/2)\big) & \text{if } p \geq 2, \\ \\ A_p\,\sigma_p & = & \dfrac{2^{p/2}\sigma_p}{\sqrt{\pi}}\min\big(\Gamma(3/2),\Gamma((p+1)/2)\big) & \text{if } p \leq 2, \end{cases}$$

$$C_p = \begin{cases} B_p\,\sigma_p & = & \dfrac{2^{p/2}\sigma_p}{\sqrt{\pi}}\max\big(\Gamma(3/2),\Gamma((p+1)/2)\big) & \text{if } p \geq 2, \\ \\ B_p\,\sigma_2^{p/2} & = & \dfrac{2^{p/2}\sigma_2^{p/2}}{\sqrt{\pi}}\max\big(\Gamma(3/2),\Gamma((p+1)/2)\big) & \text{if } p \leq 2. \end{cases}$$

For $p = 2$*, we observe that* $c_p = C_p$*. This explains the link between the Restricted Isometry Properties* (1) *and* (4) *in this case.*

Remark 2.2 *If* f *is the probability density function of a standard Gaussian random variable* ξ*, we observe that* $/\!\!/\mathbf{x}/\!\!/_{f,p}^p$ *is the expectation of* $\big|\sum_{j=1}^N x_j\xi_j\big|^p$ *for independent copies* $\xi_1, \ldots, \xi_N$ *of* ξ*. But since* $\sum_{j=1}^N x_j\xi_j$ *is a zero-mean Gaussian random variable with variance* $\|\mathbf{x}\|_2^2$*, we can compute* $/\!\!/\mathbf{x}/\!\!/_{f,p}^p$ *exactly as* $/\!\!/\mathbf{x}/\!\!/_{f,p}^p = 2^{p/2}\Gamma((p+1)/2)\|\mathbf{x}\|_2^p/\sqrt{\pi}$*. This explains the link between the Restricted Isometry Properties* (2) *and* (4) *in this case.*

3 Exponential Decay for the Tail Probability

In order to establish the q-Modified Restricted Isometry Property(4), we consider first individual vectors $\mathbf{x} \in \mathbb{R}^N$ and study the tail probability

$$\mathbf{P}\Big(\big|\|A\mathbf{x}\|_q^q - m /\!\!/ \mathbf{x} /\!\!/_{f,q}^q\big| > \epsilon m /\!\!/ \mathbf{x} /\!\!/_{f,q}^q\Big)$$

The following lemma is at the basis of our argument.

Lemma 3.1 *Let* $\eta_1, \ldots, \eta_m$ *be independent random variables with same expectation* ν*. Suppose that there is a constant* $B \geq 1$ *such that, for all integers* $k \geq 2$*,*

$$\mathbf{E}\big(|\eta_i - \nu|^k\big) \leq (k-1)!!(B\nu)^k, \quad 1 \leq i \leq m. \tag{5}$$

Then, for any $\epsilon \in (0, 1)$*,*

$$\mathbf{P}\Big(\Big|\sum_{i=1}^m \eta_i - m\nu\Big| > \epsilon m\nu\Big) \leq 2\exp\left(-\frac{\epsilon^2 m}{6B^2}\right).$$

Proof. For all $u > 0$, using Markov's inequality and the independence of the η_i, we obtain

$$\mathbf{P}\left(\sum_{i=1}^{m}(\eta_i - \nu) > \epsilon m\nu\right) = \mathbf{P}\left(\exp\left(u\sum_{i=1}^{m}(\eta_i - \nu)\right) > \exp(u\epsilon m\nu)\right)$$
$$\leq \frac{\mathbf{E}\left(\exp\left(u\sum_{i=1}^{m}(\eta_i - \nu)\right)\right)}{\exp(u\epsilon m\nu)} = \prod_{i=1}^{m}\frac{\mathbf{E}\left(\exp\left(u(\eta_i - \nu)\right)\right)}{\exp(u\epsilon\nu)}. \quad (6)$$

In view of $\mathbf{E}(\eta_i - \nu) = 0$ and of (5), we have

$$\begin{aligned}
\mathbf{E}\left(\exp\left(u(\eta_i - \nu)\right)\right) &= 1 + \sum_{k=2}^{\infty}\frac{\mathbf{E}\left(\left(u(\eta_i - \nu)\right)^k\right)}{k!} \leq 1 + \sum_{k=2}^{\infty}\frac{(k-1)!!}{k!}(uB\nu)^k \\
&= 1 + \sum_{k=1}^{\infty}\frac{1}{2^k k!}(uB\nu)^{2k} + \sum_{k=1}^{\infty}\frac{1}{(2k+1)!!}(uB\nu)^{2k+1} \\
&\leq 1 + \sum_{k=1}^{\infty}\frac{1}{k!}\frac{(uB\nu)^{2k}}{2^k} + \frac{(uB\nu)^3}{3}\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\frac{(uB\nu)^{2(k-1)}}{2^{k-1}} \\
&= \exp\left(\frac{(uB\nu)^2}{2}\right) + \frac{(uB\nu)^3}{3}\exp\left(\frac{(uB\nu)^2}{2}\right) \\
&\leq \exp\left(\frac{(uB\nu)^2}{2}\right)\left(1 + \frac{(uB\nu)^3}{3}\right) \leq \exp\left(\frac{(uB\nu)^2}{2}\right)\exp\left(\frac{(uB\nu)^3}{3}\right).
\end{aligned}$$

It then follows that

$$\frac{\mathbf{E}\left(\exp\left(u(\eta_i - \nu)\right)\right)}{\exp(u\epsilon\nu)} \leq \exp\left(\frac{(uB\nu)^2}{2} + \frac{(uB\nu)^3}{3} - u\epsilon\nu\right).$$

We make the choice $u = \epsilon/(B^2\nu)$ to obtain

$$\frac{\mathbf{E}\left(\exp\left(u(\eta_i - \nu)\right)\right)}{\exp(u\epsilon\nu)} \leq \exp\left(\frac{\epsilon^2}{2B^2} + \frac{\epsilon^3}{3B^3} - \frac{\epsilon^2}{B^2}\right) = \exp\left(-\frac{\epsilon^2}{2B^2}\left(1 - \frac{2\epsilon}{3B}\right)\right).$$

Substituting the latter into (6), while taking into account that $1 - 2\epsilon/3B \geq 1 - 2/3 = 1/3$, we derive

$$\mathbf{P}\left(\sum_{i=1}^{m}(\eta_i - \nu) > \epsilon m\nu\right) \leq \exp\left(-\frac{\epsilon^2 m}{6B^2}\right). \quad (7)$$

We would likewise obtain

$$\mathbf{P}\left(\sum_{i=1}^{m}(\eta_i - \nu) < -\epsilon m\nu\right) = \mathbf{P}\left(\sum_{i=1}^{m}(-\eta_i + \nu) > \epsilon m\nu\right) \leq \exp\left(-\frac{\epsilon^2 m}{6B^2}\right). \quad (8)$$

The required result follows from both (7) and (8). ■

Theorem 3.1 *Suppose that the entries of the $m \times N$ matrix A are independent copies of a random variable with an even probability density function f. Given $0 < q \leq 1$, if there is a constant $\alpha > 0$ such that the bounds c_q and C_q satisfy*

$$\frac{C_{qj}}{c_q^j} \leq (j-1)!!\alpha^j \quad \text{for all intergers } j \geq 1, \tag{9}$$

then, for any $\mathbf{x} \in \mathbb{R}^N$ and any $\epsilon \in (0,1)$,

$$\mathbf{P}\Big(\big| \|A\mathbf{x}\|_q^q - m \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^q \big| > \epsilon m \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^q \Big) \leq 2 \exp\left(-\frac{\epsilon^2 m}{6(1+\alpha)^2} \right).$$

Proof. Let $\mathbf{x}$ be a fixed vector in $\mathbb{R}^N$. We introduce the independent random variables $\eta_1, \ldots, \eta_m$ defined by

$$\eta_i := |(A\mathbf{x})_i|^q = \Big| \sum_{j=1}^N a_{i,j} x_j \Big|^q.$$

Let us notice that their expectations are independent of i, since

$$\mathbf{E}(\eta_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \Big| \sum_{j=1}^N t_j x_j \Big|^q f(t_1) \cdots f(t_N) dt_1 \cdots dt_N = \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^q =: \nu.$$

We calculate, for an integer $k \geq 2$,

$$\begin{aligned} \mathbf{E}\big(|\eta_i - \nu|^k\big) &= \mathbf{E}\Big(\Big| \sum_{\ell=0}^k \binom{k}{\ell} (-1)^{k-\ell} \nu^{k-\ell} \eta_i^\ell \Big| \Big) \\ &\leq \sum_{\ell=0}^k \binom{k}{\ell} \nu^{k-\ell} \mathbf{E}(\eta_i^\ell) = \sum_{\ell=0}^k \binom{k}{\ell} \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^{q(k-\ell)} \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q\ell}^{q\ell}. \end{aligned}$$

Calling upon Lemma 2.3 and Condition (9), we obtain

$$\begin{aligned} \mathbf{E}\left(|\eta_i - \nu|^k \right) &\leq \sum_{\ell=0}^k \binom{k}{\ell} \frac{C_{q\ell}}{c_q^\ell} \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^{qk} \leq \sum_{\ell=0}^k \binom{k}{\ell} (\ell-1)!!\alpha^\ell \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^{qk} \\ &= (k-1)!!(1+\alpha)^k \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^{qk}. \end{aligned}$$

Lemma 3.1 can thus be applied to obtain

$$\mathbf{P}\Big(\Big| \sum_{i=1}^m \eta_i - m \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^q \Big| > \epsilon m \mathbin{/\!/} \mathbf{x} \mathbin{/\!/}_{f,q}^q \Big) \leq 2 \exp\left(-\frac{\epsilon^2 m}{6(1+\alpha)^2} \right).$$

The observation that $\|A\mathbf{x}\|_q^q = \sum_{i=1}^m \eta_i$ finishes the proof. ∎

This theorem can now be applied to symmetric pre-Gaussian random variables.

Corollary 3.1 *Suppose that the entries of the $m \times N$ matrix A are independent copies of a pre-Gaussian random variable ξ with an even probability density function f. Given $0 < q \le 1/3$, for any $\mathbf{x} \in \mathbb{R}^N$ and any $\epsilon \in (0,1)$, one has*

$$\mathbf{P}\Big(\big| \|A\mathbf{x}\|_q^q - m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \big| > \epsilon m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \Big) \le 2 \exp\left(-\kappa \epsilon^2 m\right), \tag{10}$$

where κ is a positive constant depending on q and on f.

Proof. Our aim is to prove that the inequality (9) holds when $0 < q \le 1/3$ for the random variable ξ. Since the latter is pre-Gaussian, there exists a positive constant b such that

$$\sigma_k = \mathbf{E}\big(|\xi|^k\big) \le k! b^k \qquad \text{for all integers } k \ge 1.$$

According to Remark 2.1, we need to separate two cases: $qj \ge 2$ and $qj < 2$. We assume first that $qj \ge 2$. Let $n \ge 1$ denote the smallest integer greater than or equal to $qj/2$, so that $2(n-1) < qj \le 2n$. From the increase of $\sigma_r^{1/r}$ with $r > 0$ and the increase of $\Gamma(t)$ with $t \ge 3/2$, we derive

$$\begin{aligned} C_{qj} &= \frac{2^{qj/2}\sigma_{qj}}{\sqrt{\pi}} \max\big(\Gamma(3/2), \Gamma((qj+1)/2)\big) \le \frac{2^{qj/2}}{\sqrt{\pi}} \sigma_{2n}^{qj/2n} \Gamma((2n+1)/2) \\ &\le \frac{2^{qj/2}}{\sqrt{\pi}} ((2n)!)^{qj/2n} b^{qj} \Gamma((2n+1)/2). \end{aligned}$$

Recalling that

$$\Gamma((2k+1)/2) = \frac{\sqrt{\pi}}{2^k}(2k-1)!! \qquad \text{for all integers } k \ge 0,$$

and using $(2n-1)!! \le 2^n n!$, we deduce that

$$C_{qj} \le \frac{2^{qj/2} b^{qj}}{\sqrt{\pi}} (2n)! \sqrt{\pi} n! = 2^{qj/2} b^{qj} (2n)! n!.$$

In view of $2(n-1) < qj \le j/3$, hence $6n - 6 \le j - 1$, we observe that

$$\begin{aligned} &(2n)!n! = (2n)(2n-1)(2n-2)(2n-3)(2n-4)(2n-5)\cdots n(n-1)(n-2)(n-3)\cdots \\ &= \frac{(2n)(2n-1)}{3^{2n-2}}(6n-6)(6n-9)(6n-12)(6n-15)\cdots \frac{n(n-1)}{6^{n-2}}(6n-12)(6n-18)\cdots \\ &\le \frac{(2n)(2n-1)n(n-1)}{2^{n-2}3^{3n-4}}(6n-6)(6n-8)(6n-12)(6n-14)\cdots(6n-10)(6n-16)\cdots \\ &= \frac{n^2(n-1/2)(n-1)}{2^{2n-4}3^{3n-4}}(j-1)(j-3)(j-7)(j-9)\cdots(j-5)(j-11)\cdots \\ &\le \beta^j (j-1)!! \end{aligned}$$

for some positive constant β. We therefore obtain

$$\frac{C_{qj}}{c_q^j} \le \left(\frac{2^{q/2} b^q \beta}{c_q}\right)^j (j-1)!!.$$

We now consider the case $qj < 2$. Here, we simply use $\Gamma(t) \leq \Gamma(1/2) = \sqrt{\pi}$ for all $t \in [1/2, 3/2]$ to derive

$$C_{qj} = \frac{2^{qj/2}\sigma_{qj}}{\sqrt{\pi}} \max\big(\Gamma(3/2), \Gamma((qj+1)/2)\big) \leq 2^{qj/2}\sigma_2^{qj/2} \leq 2^{qj} b^{qj},$$

and then, since $(j-1)!! \geq 1$,

$$\frac{C_{qj}}{c_q^j} \leq \left(\frac{2^q b^q}{c_q}\right)^j (j-1)!!.$$

This concludes the proof, since Condition (9) always holds for some α depending on q and on f. ■

4 Modified Restricted Isometry Property

In this section, we show how to pass from an estimate (10) valid for one vector to an estimate valid for all sparse vectors. The following theorem says that the q-Modified Restricted Isometry Property is satisfied with overwhelming probability.

Theorem 4.1 *Suppose that the entries of the $m \times N$ matrix A are independent copies of a pre-Gaussian random variable with an even probability density function f. Given $0 < q \leq 1/3$ and $\delta \in (0,1)$, there exist constants k_1, k_2, and k_3 depending on f and q such that*

$$\mathbf{P}\Big(\big|\|A\mathbf{x}\|_q^q - m /\!\!/ \mathbf{x} /\!\!/_{f,q}^q\big| < m\delta /\!\!/ \mathbf{x} /\!\!/_{f,q}^q \text{ for all } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \leq t\Big) \geq 1 - 2\exp(-mk_1\delta)$$

provided that

$$m \geq k_2 \frac{t}{\delta^3} + k_3 \frac{t}{\delta^2} \ln(eN/t).$$

Proof. We mainly follow the ideas of [1], which we adapted to the present situation with the use of Lemma 2.2, for instance. The details are included for the reader's convenience. We start by considering a fixed index set $T \subseteq \{1, \ldots, N\}$ of cardinality t and we denote by $\mathbb{R}^T$ the space of vectors in $\mathbb{R}^N$ supported on T. In view of Lemma 2.2, we can find a subset $\mathcal{U}_{T,q}$ of the unit sphere $\mathcal{S}_{T,q}$ of $\mathbb{R}^T$ relative to the quasinorm $/\!\!/ \cdot /\!\!/_{f,q}$ such that

$$\min_{\mathbf{u} \in \mathcal{U}_{T,q}} /\!\!/ \mathbf{x} - \mathbf{u} /\!\!/_{f,q}^q \leq \frac{\delta}{4} \quad \text{for all } \mathbf{x} \in \mathcal{S}_{T,q} \qquad \text{and} \qquad \operatorname{card}(\mathcal{U}_{T,q}) \leq \left(1 + \frac{8}{\delta}\right)^{t/q}.$$

Using Corollary 3.1 and a union bound, we obtain

$$\mathbf{P}\Big(\big|\|A\mathbf{u}\|_q^q - m /\!\!/ \mathbf{u} /\!\!/_{f,q}^q\big| > m\frac{\delta}{2} /\!\!/ \mathbf{u} /\!\!/_{f,q}^q \quad \text{for some } \mathbf{u} \in \mathcal{U}_{T,q}\Big)$$
$$\leq \left(1 + \frac{8}{\delta}\right)^{t/q} 2\exp\left(-\frac{\kappa\delta^2 m}{4}\right) \leq 2\exp\left(-\frac{\kappa\delta^2 m}{4} + \frac{8t}{q\delta}\right).$$

This means that with high probability one has

$$\big| \|A\mathbf{u}\|_q^q - m \,/\!\!/\, \mathbf{u} \,/\!\!/_{f,q}^q \big| < m\frac{\delta}{2} \,/\!\!/\, \mathbf{u} \,/\!\!/_{f,q}^q \quad \text{for all } \mathbf{u} \in \mathcal{U}_{T,q}.$$

Let us assume that the matrix A is drawn so that the latter holds, in other words, so that

$$m\left(1 - \frac{\delta}{2}\right) /\!\!/\, \mathbf{u} /\!\!/_{f,q}^q \leq \|A\mathbf{u}\|_q^q \leq m\left(1 + \frac{\delta}{2}\right) /\!\!/\, \mathbf{u} \,/\!\!/_{f,q}^q \quad \text{for all } \mathbf{u} \in \mathcal{U}_{T,q}. \tag{11}$$

Let $\tilde{\delta}$ be the smallest positive constant such that

$$\|A\mathbf{x}\|_q^q \leq m(1 + \tilde{\delta}) \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \qquad \text{for all } \mathbf{x} \in \mathcal{S}_{T,q}. \tag{12}$$

On the one hand, given $\mathbf{z} \in \mathcal{S}_{T,q}$, picking $\mathbf{u} \in \mathcal{U}_{T,q}$ with $/\!\!/ \mathbf{x} - \mathbf{u} /\!\!/_{f,q}^q \leq \delta/4$, we derive

$$\begin{aligned} \|A\mathbf{x}\|_q^q &\leq \|A\mathbf{u}\|_q^q + \|A(\mathbf{x} - \mathbf{u})\|_q^q \leq m\left(1 + \frac{\delta}{2}\right) + m(1 + \tilde{\delta}) \,/\!\!/\, \mathbf{x} - \mathbf{u} \,/\!\!/_{f,q}^q \\ &\leq m\left(1 + \frac{\delta}{2} + (1 + \tilde{\delta})\frac{\delta}{4}\right). \end{aligned}$$

The minimality of $\tilde{\delta}$ implies that

$$1 + \tilde{\delta} \leq 1 + \frac{\delta}{2} + (1 + \tilde{\delta})\frac{\delta}{4} \leq 1 + \frac{3\delta}{4} + \frac{\tilde{\delta}}{4}, \qquad \text{i.e.} \quad \tilde{\delta} \leq \delta.$$

Substituting into (12), we obtain the upper estimate

$$\|A\mathbf{x}\|_q^q \leq m(1 + \delta) \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \qquad \text{for all } \mathbf{x} \in \mathbb{R}^T. \tag{13}$$

On the other hand, given $\mathbf{x} \in \mathcal{S}_{T,q}$, still picking $\mathbf{u} \in \mathcal{U}_{T,q}$ with $/\!\!/ \mathbf{x} - \mathbf{u} /\!\!/_{f,q}^q \leq \delta/4$, we have

$$\begin{aligned} \|A\mathbf{x}\|_q^q &\geq \|A\mathbf{u}\|_q^q - \|A(\mathbf{x} - \mathbf{u})\|_q^q \geq m\left(1 - \frac{\delta}{2}\right) - m(1 + \delta) \,/\!\!/\, \mathbf{x} - \mathbf{u} \,/\!\!/_{f,q}^q \\ &\geq m\left(1 - \frac{\delta}{2} - (1 + \delta)\frac{\delta}{4}\right) \geq m(1 - \delta). \end{aligned}$$

Thus we obtain the lower estimate

$$\|A\mathbf{x}\|_q^q \geq m(1 - \delta) \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \qquad \text{for all } \mathbf{x} \in \mathbb{R}^T. \tag{14}$$

Since both estimates (13) and (14) hold as soon as (11) is fulfilled, we have

$$\mathbf{P}\Big(\big| \|A\mathbf{x}\|_q^q - m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \big| > m\delta \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \quad \text{for some } \mathbf{x} \in \mathbb{R}^T\Big) \leq 2\exp\left(-\frac{\kappa\delta^2 m}{4} + \frac{8t}{q\delta}\right).$$

We now observe that the set of t-sparse vectors is the union of $\binom{N}{t} \leq (eN/t)^t$ spaces $\mathbb{R}^T$ to deduce

$$\begin{aligned} &\mathbf{P}\Big(\big| \|A\mathbf{x}\|_q^q - m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \big| > m\delta \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \quad \text{for some } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \leq t\Big) \\ &\qquad \leq 2\binom{N}{t}\exp\left(-\frac{\kappa\delta^2 m}{4} + \frac{8t}{q\delta}\right) \leq 2\exp\left(-\frac{\kappa\delta^2 m}{4} + \frac{8t}{q\delta} + t\ln\left(\frac{eN}{t}\right)\right). \end{aligned}$$

If we impose, say, that

$$\frac{8t}{q\delta} + t\ln\left(\frac{eN}{t}\right) \leq \frac{1}{2}\frac{\kappa\delta^2 m}{4}, \qquad \text{i.e.} \qquad m \geq \frac{64}{\kappa q}\frac{t}{\delta^3} + \frac{8}{\kappa}\frac{t}{\delta^2}\ln\left(\frac{eN}{t}\right),$$

we have

$$\mathbf{P}\Big(\big|\|A\mathbf{x}\|_q^q - m \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \big| > m\delta \,/\!\!/\, \mathbf{x} \,/\!\!/_{f,q}^q \quad \text{for some } \mathbf{x} \in \mathbb{R}^N \text{ with } \|\mathbf{x}\|_0 \leq t\Big)$$
$$\leq 2\exp\left(-\frac{\kappa\delta^2 m}{8}\right),$$

which is the required result with $k_1 := \kappa/8$, $k_2 := 64/\kappa q$, and $k_3 := 8/\kappa$. ■

5 Sparse Recovery

In this final section, we prove that ℓ_q-recovery is guaranteed for matrices satisfying the q-Modified Restricted Isometry Property, which enables us to establish our main theorem.

Theorem 5.1 *Given the even probability density function f and given $0 < q \leq 1$, if there is an integer $t \geq s$ such that a matrix A satisfies the q-Modified Restricted Isometry Property* (4) *with*

$$\delta_{s+t} < \frac{K-1}{K+1}, \qquad \text{where} \quad K := \frac{c_q}{C_q}\left(\frac{t}{s}\right)^{1-q/2} > 1, \tag{15}$$

then any s-sparse vector $\mathbf{x}$ is exactly recovered as a solution of (P_q) *with $\mathbf{y} = A\mathbf{x}$.*

Proof. Let $\mathbf{x} \in \mathbb{R}^N$ be an s-sparse vector and let $\mathbf{x}^\star$ be a solution of (P_q) with $\mathbf{y} = A\mathbf{x}$. Let us set $\mathbf{v} := \mathbf{x} - \mathbf{x}^\star$, which is an element of the null space of A. We partition $\{1, \ldots, N\}$ as $S_0 \cup S_1 \cup S_2 \cup \ldots$, where

S_0 is an index set of s largest absolute-value entries of $\mathbf{v}$,
S_1 is an index set of t next largest absolute-value entries of $\mathbf{v}$,
S_2 is an index set of t next largest absolute-value entries of $\mathbf{v}$,

etc. In view of the q-Modified Restricted Isometry Property (4), we have

$$\begin{aligned}
/\!\!/\mathbf{v}_{S_0} + \mathbf{v}_{S_1}/\!\!/_{f,q}^q &\leq \frac{1}{m(1-\delta_{s+t})}\|A(\mathbf{v}_{S_0} + \mathbf{v}_{S_1})\|_q^q \\
&= \frac{1}{m(1-\delta_{s+t})}\|A(-\mathbf{v}_{S_2} - \mathbf{v}_{S_3} - \cdots)\|_q^q \\
&\leq \frac{1}{m(1-\delta_{s+t})}\left[\|A\mathbf{v}_{S_2}\|_q^q + \|A\mathbf{v}_{S_3}\|_q^q + \cdots\right] \\
&\leq \frac{m(1+\delta_{s+t})}{m(1-\delta_{s+t})}\left[\,/\!\!/\,\mathbf{v}_{S_2}\,/\!\!/_{f,q}^q + /\!\!/\,\mathbf{v}_{S_3}\,/\!\!/_{f,q}^q + \cdots\right] \\
&\leq \frac{1+\delta_{s+t}}{1-\delta_{s+t}}C_q\left[\|\mathbf{v}_{S_2}\|_2^q + \|\mathbf{v}_{S_3}\|_2^q + \cdots\right].
\end{aligned}$$

It is classical to observe that, for $k \geq 2$,

$$\|\mathbf{v}_{S_k}\|_2 \leq \frac{1}{t^{1/q-1/2}} \|\mathbf{v}_{S_{k-1}}\|_q.$$

This is derived from the inequalities $|v_i|^q \leq |v_j|^q$, $i \in S_k$, $j \in S_{k-1}$, by averaging over j, raising to the power $2/q$, and summing over i. Thus, with $\overline{S}_0$ denoting the complement of S_0 in $\{1, \ldots, N\}$, we obtain

$$/\!\!/\mathbf{v}_{S_0} + \mathbf{v}_{S_1}/\!\!/_{f,q}^q \leq \frac{1+\delta_{s+t}}{1-\delta_{s+t}} C_q \frac{1}{t^{1-q/2}} \left[\|\mathbf{v}_{S_1}\|_q^q + \|\mathbf{v}_{S_2}\|_q^q + \cdots\right] \leq \frac{1+\delta_{s+t}}{1-\delta_{s+t}} C_q \frac{1}{t^{1-q/2}} \|\mathbf{v}_{\overline{S}_0}\|_q^q \tag{16}$$

Next, we take into account that

$$\|\mathbf{v}_{S_0}\|_q^q \leq s^{1-q/2} \|\mathbf{v}_{S_0}\|_2^q \leq s^{1-q/2} \|\mathbf{v}_{S_0} + \mathbf{v}_{S_1}\|_2^q \leq \frac{s^{1-q/2}}{c_q} /\!\!/ \mathbf{v}_{S_0} + \mathbf{v}_{S_1} /\!\!/_{f,q}^q . \tag{17}$$

Combining (16) and (17), we obtain

$$\|\mathbf{v}_{S_0}\|_q^q \leq \frac{1+\delta_{s+t}}{1-\delta_{s+t}} \frac{C_q}{c_q} \left(\frac{s}{t}\right)^{1-q/2} \|\mathbf{v}_{\overline{S}_0}\|_q^q.$$

Supposing that $\mathbf{v} \neq 0$, the assumption (15) now implies

$$\|\mathbf{v}_{S_0}\|_q^q < \|\mathbf{v}_{\overline{S}_0}\|_q^q.$$

Since this holds for the index set S_0 consisting of s largest absolute-value entries of $\mathbf{v}$, it also holds for any index set of s entries of $\mathbf{v}$. In particular, if S represents the support of $\mathbf{x}$, we have

$$\|\mathbf{v}_S\|_q^q < \|\mathbf{v}_{\overline{S}}\|_q^q. \tag{18}$$

But we also have, because $\mathbf{x}^\star$ is a minimizer of (P_q),

$$\|\mathbf{x}\|_q^q \geq \|\mathbf{x}^\star\|_q^q = \|\mathbf{x}^\star_S\|_q^q + \|\mathbf{x}^\star_{\overline{S}}\|_q^q.$$

By the triangle inequality for the q-th power of the ℓ_q-quasinorm, we obtain

$$\|\mathbf{x}\|_q^q \geq \|\mathbf{x}_S\|_q^q - \|\mathbf{v}_S\|_q^q + \|\mathbf{v}_{\overline{S}}\|_q^q - \|\mathbf{x}_{\overline{S}}\|_q^q = \|\mathbf{x}\|_q^q - \|\mathbf{v}_S\|_q^q + \|\mathbf{v}_{\overline{S}}\|_q^q.$$

Rearranging the latter yields

$$\|\mathbf{v}_S\|_q^q \geq \|\mathbf{v}_{\overline{S}}\|_q^q. \tag{19}$$

The inequalities (18) and (19) are contradictory, unless $\mathbf{v} = 0$. We have therefore proved that $\mathbf{x}^\star = \mathbf{x}$, as expected. ■

It remains to put all the arguments together to validate Theorem 1.1. We start by fixing q to $1/3$, which is legitimate since the ℓ_q-recovery succeeds for $0 < q \leq 1/3$ if it succeeds for $q = 1/3$. We then choose an integer t roughly proportional to s so that $K \geq 2$ in (15), e.g. the smallest integer larger than or equal to ds, where $d := (2C_{1/3}/c_{1/3})^{6/5}$ depends only on f. In this case, Theorem 5.1 ensures the ℓ_q-recovery of every s-sparse vector provided that $\delta_{s+t} < 1/3$, say $\delta_{s+t} \leq 1/4$. But according to Theorem 4.1, when the matrix A is populated by independent copies of the pre-Gaussian random variable with probability density function f, this is satisfied with probability $\geq 1 - 2\exp(-mk_1')$ as soon as $m \geq k_2'(s+t) + k_3'(s+t)\ln\big(eN/(s+t)\big)$, with constants k_1', k_2', and k_3' depending only on f. This condition is of course guaranteed when $m \geq k_4' s \ln\big(N/s\big)$ for some constant k_4' depending only on f.

Remark 5.1 *In the spirit of Compressed Sensing, one also needs to control the error between an almost sparse vector and the vector which is reconstructed using slightly flawed measurements. This was originally done in [4] for the ℓ_1-minimization, and later in [8] for the ℓ_q-minimization. A reader familiar with Compressed Sensing can easily perform the appropriate modifications to obtain such a result. We chose not to include it here for the sake of clarity.*

References

[1] Baraniuk, R., M. Davenport, R. DeVore, and M. B. Wakin, A simple proof of the restricted isometry property for random matrices, Constructive Approximation, 28(2008), 253–263.

[2] Buldygin, V. V. and Yu. V. Kozachenko, Metric Characterization of Random Variables and Random Processes, AMS Publication, Providence, 2000.

[3] Candès, E. J., The restricted isometry property and its implications for compressed sensing, Comptes Rendus de l'Académie des Sciences, Série I, 346(2008), 589–592.

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

[5] Candès, E. and T. Tao, Decoding by linear programing, IEEE Trans. Inf. Theory 51(2005), 4203–4215.

[6] Chartrand, R. and V. Staneva, Restricted isometry properties and nonconvex compressive sensing, Inverse Problems, 2008(24), 1–14.

[7] Donoho, D., For most large underdetermined systems of linear equations, the minimal ℓ_1-norm solution is also the sparsest solution, Comm. Pure Appl. Math., 59(2006), 797–829.

[8] Foucart, S. and M. J. Lai, Sparsest Solutions of Underdetermined Linear Systems via ℓ_q-minimization for $0 < q \leq 1$, Applied and Comput. Harmonic Analysis, 26(2009), 395–407.

[9] Gribonval, R. and M. Nielsen, Highly sparse representations from dictionaries are unique and independent of the sparseness measure, Applied and Comput. Harmonic Analysis, 22(2007), 335–355.

[10] Ledoux, M., *The Concentration of Measure Phenomenon*, American Math. Society, Providence, 2001.

[11] Mendelson, S., A. Pajor and N. Tomczak-Jaegermann, Uniform uncertainty principal for Bernoulli and sub-Gaussian ensembles, Constructive Approximation, 28(2008), 277–289.

[12] Nazarov, F. and A. Podkorytov, Ball, Haagerup, and distribution functions, Complex analysis, operators, and related topics, 247–267, Oper. Theory Adv. Appl., 113, Birkhuser, Basel, 2000.

[13] Pisier, G., *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge University Press, 1999.