Uniform uncertainty principle for Bernoulli and subgaussian ensembles

Shahar MENDELSON[1] Alain PAJOR[2]
Nicole TOMCZAK-JAEGERMANN[3]

August 21, 2006

1 Introduction

In [CT1] Candes and Tao studied problems of approximate and exact reconstruction of sparse signals from incomplete random measurements and related them to the eigenvalue behavior of submatrices of matrices of random measurements. In particular they introduced the notion they called the *uniform uncertainty principle* (UUP, defined below) and studied it for Gaussian, Bernoulli and Fourier ensembles. This notion was further refined in [CT2, CRT]. In this context they asked ([T]) whether rectangular $k \times n$ Bernoulli matrices (with $k < n$) have the property that by arbitrarily extracting m (with $m < k$) columns one can make so obtained submatrices arbitrarily close to (multiples of) isometries of a Euclidean space (of course m would then depend on the required degree of "closeness" and dimensions k and n).

A different–geometric–approach to approximate and exact reconstruction problems was proposed in [MPT1, MPT2]. Although in these articles the notion of UUP was not considered, an application of one of the main general results there in a simple particular case implied an immediate affirmative

[0] *2000 MSC-classification:* 46B07, 41A45, 94B75, 52B05, 62G99

[1] Partially supported by an Australian Research Council Discovery grant.
[2] Partially supported by an Australian Research Council grant.

[3] This author holds the Canada Research Chair in Geometric Analysis.

answer to the Candes-Romberg-Tao's question (see Corollary 3.5 in [MPT2] and the comments afterwards).

The common roots of the geometric approach of [MPT1, MPT2], as well as the UUP or other related properties, revolve around the fact that various "random projection" operators may act as "almost norm preserving" on various subsets of the sphere; with the UUP associated to the subset of "sparse" vectors on the sphere (denoted later by U_m).

In this note we observe that the results of [MPT1, MPT2] can be applied to a number of other sets with a very simple geometry to get interesting conclusions for the Gaussian, Bernoulli, and more generally, any subgaussian ensemble. Since the proofs of the general results of [MPT1, MPT2] are not easily accessible to non-specialists, we also provide an alternative elementary argument, which works for the specific sets we are interested in.

Let us recall the following notation. By $|\cdot|$ we denote the Euclidean norm on $\mathbb{R}^n$, by $\langle \cdot, \cdot \rangle$ the corresponding inner product, by B_2^n and S^{n-1} the unit Euclidean ball and the unit sphere, respectively. For $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ we let $\operatorname{supp} x = \{i : x_i \neq 0\}$. For a finite set A, the cardinality of A is denoted by $|A|$, and for a set $A \subset \mathbb{R}^n$, $\operatorname{conv} A$ denotes the convex hull of A. Throughout, all absolute constants are fixed, positive numbers, which are denoted by c, C, c', etc. Their value may change from line to line.

We will work with the following (slightly refined) definition of the uniform uncertainty principle ([CT1]).

Definition 1.1 *A $k \times n$ (random) measurement matrix Γ obeys the uniform uncertainty principle with accuracy $0 < \theta < 1$ and oversampling factor $\lambda > 1$, if the following statement is true with probability close to 1: for all subsets $A \subset \{1, \ldots, n\}$ with $|A| \leq k/\lambda$, the matrix Γ_A, obtained by extracting from Γ the columns corresponding to A, satisfies*

$$1 - \theta \leq \lambda_{\min}\left(\frac{\Gamma_A^* \Gamma_A}{k}\right) \leq \lambda_{\max}\left(\frac{\Gamma_A^* \Gamma_A}{k}\right) \leq 1 + \theta, \tag{1.1}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimal and maximal eigenvalues, respectively. Equivalently,

$$(1-\theta)|x|^2 \leq \frac{|\Gamma x|^2}{k} \leq (1+\theta)|x|^2, \tag{1.2}$$

for all vectors $x \in \mathbb{R}^n$ with $|\operatorname{supp} x| \leq k/\lambda$.

We shall use a shorthand notation of $uup(\theta, \lambda)$ for the above property.

In this language and for the Bernoulli ensemble, Candes and Tao showed ([CT1, CT2]) that there exist two absolute constants $0 < \theta_0 < 1$ and $c > 0$ such that for all $k < n$, $k \times n$ Bernoulli random matrices satisfy $uup(\theta_0, \lambda)$ for $\lambda = c \log(cn/k)$, and they asked ([T]) whether an analogous result is true for *every* $0 < \theta < 1$. We formally state their question as follows:

Question 1.2 *Let $1 \leq k < n$ and set Γ to be a $k \times n$ random Bernoulli matrix. Let $0 < \theta < 1$ be arbitrary. Can one find λ depending on θ, k and n only and satisfying $1 \leq \lambda \leq c_\theta \log(c_\theta n/k)$, where $c_\theta > 0$ depends only on θ, such that for n "large enough" and any $k < n$, Γ satisfies $uup(\theta, \lambda)$ with probability close to 1?*

As already mentioned earlier, a positive answer to this question for (more general) subgaussian measurements follows immediately from the results of [MPT1, MPT2], and we explain this connection in the next section. We then show how one can obtain a similar estimate using elementary methods, which can also be used to solve the approximate reconstruction problem in certain simple (but central for the applications) cases (see Section 3 for more details).

Acknowledgement: Part of the work on this article was conducted during the Trimester "Phenomena in High Dimensions" held at the Centre Emile Borel, Institute Henri Poincaré, Paris in the Spring of 2006. We are grateful to the Institute for its hospitality and the excellent working atmosphere it provided.

2 Subgaussian matrices and geometry of the set of sparse vectors

We first recall a few definitions. Let X be a random vector in $\mathbb{R}^n$; X is called isotropic if for every $y \in \mathbb{R}^n$, $\mathbb{E}| \langle X, y \rangle |^2 = |y|^2$, and is ψ_2 with a constant α if for every $y \in \mathbb{R}^n$,

$$\| \langle X, y \rangle \|_{\psi_2} := \inf \left\{ s : \mathbb{E} \exp \left(\langle X, y \rangle^2 / s^2 \right) \leq 2 \right\} \leq \alpha |y|.$$

The most important examples for us are the Gaussian vector $(g_1, \ldots, g_n)$ where the g_i's are independent $N(0,1)$ Gaussian variables and the random sign vector $(\varepsilon_1, \dots, \varepsilon_n)$ where the ε_i's are independent, symmetric ± 1 (Bernoulli) random variables; in both these cases the random vectors are isotropic with a ψ_2 constant $\alpha = c_0'$, for a suitable absolute constant $c_0' \geq 1$.

A subgaussian or ψ_2 operator is a random operator $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ of the form

$$\Gamma = \sum_{i=1}^{k} \langle X_i, \cdot \rangle \, e_i, \tag{2.1}$$

where $X_1, \dots, X_k$ are independent copies of an isotropic ψ_2 vector X on $\mathbb{R}^n$.

Note that if $X_i = (x_{i,j})_{j=1}^n$ then Γ is represented by a matrix whose rows are $(X_i)_{i=1}^k$. However, although the rows of the matrix are independent random vectors, the entries within each row may be dependent.

Finally, for a subset $T \subset \mathbb{R}^n$ we set

$$\ell_*(T) = \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} g_i t_i \right|, \tag{2.2}$$

where $t = (t_i)_{i=1}^n \in \mathbb{R}^n$ and $g_1, ..., g_n$ are independent $N(0,1)$ Gaussian random variables.

The following fact was proved in [MPT2] (Corollary 2.7) as a consequence of one of the main results of [MPT1, MPT2].

Theorem 2.1 *Let $1 \leq k \leq n$ and $0 < \theta < 1$. Let X be an isotropic ψ_2 random vector on $\mathbb{R}^n$ with constant α, set $X_1, \dots, X_k$ to be independent copies of X, put Γ as defined by (2.1) and let $T \subset S^{n-1}$. If k satisfies*

$$k \geq (c' \alpha^4 / \theta^2) \ell_*(T)^2,$$

then with probability at least $1 - \exp(-\bar{c}\, \theta^2 k / \alpha^4)$, for all $x \in T$,

$$1 - \theta \leq \frac{|\Gamma x|^2}{k} \leq 1 + \theta, \tag{2.3}$$

where $c', \bar{c} > 0$ are absolute constants.

Let us explain the meaning of Theorem 2.1, and for the sake of simplicity, assume that α is an absolute constant (in particular independent on the dimension n), as this is the situation for Gaussian or Bernoulli random vectors. The parameter $\ell_*(T)$ is a complexity measure of the set T; in this context, it measures the extent in which probabilistic bounds on the concentration of individual random variables of the form $|\Gamma x|^2$ around their mean can be combined to form a bound that holds uniformly for every $x \in T$. The assertion of Theorem 2.1 is that as long as $k \geq c\ell_*^2(T)/\theta^2$, the random operator $\Gamma/\sqrt{k}$ maps with overwhelming probability all the points in T in an almost norm preserving way.

Let us note that the method used in the proof of Theorem 2.1 is called *generic chaining* (see [Ta] for the most recent survey on this subject). As we show in Section 3, if the set T is "very simple" one can combine the concentration of individual variables around their means and obtain a uniform bound using a far simpler approach.

The prime example for which we would like to apply Theorem 2.1 are the sets U_m consisting of sparse vectors, which are defined for $1 \leq m \leq n$ by

$$U_m := \left\{x \in S^{n-1} : |\mathrm{supp}\, x| \leq m\right\}. \tag{2.4}$$

We shall also consider the analogous subset of the Euclidean ball,

$$\tilde{U}_m := \{x \in B_2^n : |\mathrm{supp}\, x| \leq m\}. \tag{2.5}$$

The reason for our interest in the set U_m is clear: the ability to map it in an almost norm preserving way is equivalent to the UUP. To that end, and in light of Theorem 2.1, one has to bound $\ell_*(U_m)$ in order to control $uup(\theta, \lambda)$.

The sets U_m and $\tilde{U}_m$ have particularly simple structure: they are the unions of the unit spheres, and unit balls, respectively, supported on m-dimensional coordinate subspaces of $\mathbb{R}^n$. Furthermore, for any $0 < r \leq 1$,

$$\tilde{U}_m \cap rB_2^n = r\tilde{U}_m. \tag{2.6}$$

It turns out that a simple geometric property of U_m plays a crucial role in the present context.

Let $T \subset \mathbb{R}^n$. Recall that a set $\Lambda \subset \mathbb{R}^n$ is an ε cover of T with respect to the Euclidean metric if

$$T \subset \bigcup_{x \in \Lambda} (x + \varepsilon B_2^n),$$

where $A + B := \{a + b : a \in A, b \in B\}$ is the Minkowski sum of the sets A and B. (Λ is often called an ε-net for T.) It is well-known and easy to see that if Λ is an ε cover of T with respect to the Euclidean metric then there exists another ε cover of T, say Λ_1, such that $\Lambda_1 \subset T$ and $|\Lambda_1| \leq |\Lambda|$.

The following fact is well-known and standard (see, for example, [P], Lemma 4.10 for a part of the argument). For the convenience of the non-specialist reader we provide a short proof.

Lemma 2.2 *Let $m \geq 1$ and $\varepsilon > 0$. There exists an ε cover $\Lambda \subset B_2^m$ of B_2^m with respect to the Euclidean metric such that $B_2^m \subset (1-\varepsilon)^{-1} \operatorname{conv} \Lambda$ and $|\Lambda| \leq (1+2/\varepsilon)^m$. Similarly, there exists $\Lambda' \subset S^{n-1}$ which is an ε cover of the sphere S^{m-1} and $|\Lambda'| \leq (1+2/\varepsilon)^m$.*

Proof. Let $\Lambda \subset B_2^m$ be a maximal subset such that $|x-y| > \varepsilon$ for all $x \neq y \in \Lambda$. By maximality, Λ is an ε cover for B_2^m. If $x \neq y \in \Lambda$ then the two balls $x + (\varepsilon/2)B_2^m$ and $y + (\varepsilon/2)B_2^m$ have disjoint interiors and $\bigcup_{x \in \Lambda}(x + (\varepsilon/2)B_2^m) \subset (1+\varepsilon/2)B_2^m$. Comparing volumes we get the estimate for $|\Lambda|$. Applying the same argument to the sphere S^{m-1} we get a set the desired set Λ'. Finally, every $z \in B_2^m$ can be written as $z = x_0 + \varepsilon z_1$, where $x_0 \in \Lambda$ and $z_1 \in B_2^m$. Iterating this we get that $z = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \ldots$, with $x_i \in \Lambda$, implying $B_2^m \subset (1-\varepsilon)^{-1} \operatorname{conv} \Lambda$, as required. ∎

The structure of U_m immediately implies similar facts as in the lemma above for U_m and $\tilde{U}_m$.

Lemma 2.3 *There exists an absolute constant c for which the following holds. For every $0 < \varepsilon \leq 1/2$ and every $1 \leq m \leq n$ there is a set $\Lambda \subset B_2^n$ which is an ε cover of $\tilde{U}_m$, such that $\tilde{U}_m \subset 2 \operatorname{conv} \Lambda$ and $|\Lambda|$ is at most*

$$\exp\left(c\, m \log\left(\frac{cn}{m\varepsilon}\right)\right). \tag{2.7}$$

Moreover, there exists an ε cover $\Lambda' \subset S^{n-1}$ of U_m with cardinality at most (2.7).

Furthermore, for any $0 < r \leq 1$ there exists $\bar{\Lambda} \subset rB_2^n$ such that $(U_m - U_m) \cap rB_2^n \subset 2 \operatorname{conv} \bar{\Lambda}$ and $|\bar{\Lambda}|$ is at most (2.7).

Proof. Considering all subsets $A \subset \{1, \ldots, n\}$ with $|A| \leq m$, it is clear that the required sets Λ and Λ' can be obtained as unions of the corresponding sets supported on coordinates from A. By Lemma 2.2 the cardinalities of these sets are at most $(5/\varepsilon)^m \binom{n}{m}$.

To prove the last statement, note that $U_m - U_m \subset 2\tilde{U}_{2m}$, which, by (2.6), implies

$$(U_m - U_m) \cap rB_2^n \subset r\tilde{U}_{2m}.$$

By the first part of the lemma, construct a set $\Lambda \subset \tilde{U}_{2m}$ such that $\tilde{U}_{2m} \subset 2 \operatorname{conv} \Lambda$ and $|\Lambda|$ admits a suitable upper bound. Finally, set $\bar{\Lambda} = r\Lambda$, completing the proof. ∎

Theorem 2.4 *There exist* $c_1, c_2, \bar{c}, C_1 > 0$ *such that the following holds. Let* n*,* θ*,* X *and* Γ *be as in Theorem 2.1. Fix* $1 \le k \le n$*, let* $T \subset S^{n-1}$ *and assume that* $T \subset 2\,\mathrm{conv}\,\Lambda$ *for some* $\Lambda \subset B_2^n$ *with* $|\Lambda| \le \exp(c_1(\theta^2/\alpha^4)k)$*. Then with probability at least* $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$*, for all* $x \in T$*,*

$$1 - \theta \le \frac{|\Gamma x|^2}{k} \le 1 + \theta. \tag{2.8}$$

Furthermore, if

$$m \le \frac{c_2 \theta^2 k}{\alpha^4 \log(C_1 n \alpha^4/\theta^2 k)},$$

then (2.8) holds for $T = U_m$*. In particular, for every* $0 < \theta < 1$*, and with probability at least* $1 - \exp(-\bar{c}\,\theta^2 k/\alpha^4)$*,* Γ *satisfies uup*(θ, λ) *for*

$$\lambda = \frac{\log(n/a'k)}{a},$$

where both $a, a' > 0$ *are of the form* $c\theta^2/\alpha^4$ *for some absolute constant* c*.*

The main point in the proof is that if $T \subset 2\,\mathrm{conv}\,\Lambda$ for $\Lambda \subset B_2^n$ and there is a reasonable control on the cardinality of Λ, then $\ell_*(T)$ may be bounded from above. The rest is just a direct application of Theorem 2.1.

Proof. Let $c', \bar{c} > 0$ be constants from Theorem 2.1. It is well known (see, for example, [LT]) that there exists an absolute constant $c'' > 0$ such that for every $\Lambda \subset B_2^n$,

$$\ell_*(\mathrm{conv}\,\Lambda) = \ell_*(\Lambda) \le c'' \sqrt{\log(|\Lambda|)},$$

and since $T \subset 2\,\mathrm{conv}\,\Lambda$ then

$$\ell_*(T) \le 2\ell_*(\mathrm{conv}\,\Lambda) \le c'' \left(c_1(\theta^2/\alpha^4)k\right)^{1/2}.$$

Choosing $c_1 = 1/(c'c''^2)$ we conclude the proof of (2.8) by applying Theorem 2.1.

As for the "furthermore" part, set Λ to be a $1/2$ cover of $\tilde{U}_m$ provided by Lemma 2.3. Then $T = U_m \subset \tilde{U}_m \subset 2\,\mathrm{conv}\,\Lambda$. Also, by (2.7) and our choice of m (that includes appropriate choices of constants c_2 and C_1), $|\Lambda|$ admits the upper bound required in the first part of the theorem. Finally, the last statement follows from (2.8) for U_m and Definition 1.1 of the UUP. ■

3 Elementary approach

3.1 The uniform uncertainty principle

The aim of this subsection is to obtain a positive answer to Question 1.2 using elementary methods and without resorting to Theorem 2.1. Such a proof is possible mainly because, as already discussed in the preceding section, the geometry of the sets U_m is particularly simple. The price one pays for the simple proof is a slightly worse dependence on the accuracy θ.

The first step in the elementary proof is obtaining an analog of Theorem 2.1, where the complexity measure $\ell_*(T)$ is replaced by estimates on covering numbers.

Consider a set of random $k \times n$ matrices $\tilde{\Gamma}$ satisfying two conditions. First,

$$\mathbb{E}|\tilde{\Gamma}x|^2 = 1 \qquad \text{for all } x \in S^{n-1}, \tag{3.1}$$

that is, on average, $\tilde{\Gamma}$ preserves the norm of each individual x.

The second condition asserts the concentration of the random variable $|\tilde{\Gamma}x|^2$ around its expectation: there exists an absolute constant c_0 such that for every $x \in \mathbb{R}^n$ we have

$$\mathbb{P}\left(\left||\tilde{\Gamma}x|^2 - |x|^2\right| \geq t|x|^2\right) \leq e^{-c_0 t^2 k} \qquad \text{for all } 0 < t \leq 1. \tag{3.2}$$

Let us note that (multiples of) subgaussian matrices considered in Section 2 satisfy (3.1) and (3.2). Indeed, let $(X_i)_{i=1}^k$ be independent copies of an isotropic ψ_2 vector with a constant α and set

$$\tilde{\Gamma} = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i, \cdot \rangle \, e_i,$$

where $(e_i)_{i=1}^n$ are the standard unit vectors in $\mathbb{R}^n$. By the isotropicity assumption, $\mathbb{E}|\tilde{\Gamma}x|^2 = 1$ for every $x \in S^{n-1}$. Moreover, by fixing $x \in S^{n-1}$ and applying Bernstein's inequality (see, e.g. [LT, VW]) to the average of k independent copies of the random variable $\langle X, x \rangle^2$, it is evident that for every $t > 0$,

$$\mathbb{P}\left(\left|\frac{1}{k}\sum_{i=1}^{k} \langle X_i, x \rangle^2 - 1\right| > t\right) \leq 2\exp\left(-ck \min\left\{\frac{t^2}{\alpha^4}, \frac{t}{\alpha^2}\right\}\right),$$

where c is an absolute constant. Since $\alpha \geq 1$, $\tilde{\Gamma}$ satisfies (3.2) for $c_0 = c/\alpha^4$.

Let us formulate the elementary version of Theorem 2.1.

Theorem 3.1 *Consider a set of random $k \times n$ matrices $\tilde{\Gamma}$ satisfying (3.1) and (3.2). Let $T \subset S^{n-1}$ and $0 < \theta < 1$, and assume the following:*

(i) There exists $\Lambda' \subset S^{n-1}$ which is a $\theta/5$-cover of T and satisfies $|\Lambda'| \leq \exp(c_0\theta^2 k/50)$.

(ii) There exists $\Lambda \subset (\theta/5)B_2^n$ such that $(T-T) \cap (\theta/5)B_2^n \subset 2\,\mathrm{conv}\,\Lambda$ and $|\Lambda| \leq \exp(c_0 k/2)$.

Then with probability at least $1 - 2\exp(-c_0\theta^2 k/50)$, for all $x \in T$,

$$1 - \theta \leq |\tilde{\Gamma}x|^2 \leq 1 + \theta. \tag{3.3}$$

Remark 3.2 There is nothing special in the constant 2 in front of $\mathrm{conv}\,\Lambda$ in (ii), and it could be replaced by any constant strictly larger than 1.

The idea behind the proof of Theorem 3.1 is to show that $\tilde{\Gamma}$ acts on Λ' in an almost norm preserving way. This is the case because the degree of concentration of each variable $|\tilde{\Gamma}x|^2$ around its mean defeats the cardinality of Λ'. Then one shows that $\tilde{\Gamma}(\mathrm{conv}\,\Lambda)$ is contained in a small ball - thanks to a similar argument.

Proof. Set $\varepsilon = \theta/5$ and consider the set of $\tilde{\Gamma}$ on which

$$\left| |\tilde{\Gamma}x_0| - 1 \right| \leq \left| |\tilde{\Gamma}x_0|^2 - 1 \right| \leq \varepsilon \qquad \text{for all } x_0 \in \Lambda', \tag{3.4}$$

and

$$|\tilde{\Gamma}z| \leq 2|z| \qquad \text{for all } z \in \Lambda. \tag{3.5}$$

Note that this set has probability larger than or equal to $1 - \exp(-c_0\varepsilon^2 k/2) - \exp(-c_0 k/2) \geq 1 - 2\exp(-c_0\varepsilon^2 k/2)$.

Let $x \in T$ and consider $x_0 \in \Lambda'$ such that $|x - x_0| \leq \varepsilon$. Then

$$|\tilde{\Gamma}x_0| - |\tilde{\Gamma}(x - x_0)| \leq |\tilde{\Gamma}x| \leq |\tilde{\Gamma}x_0| + |\tilde{\Gamma}(x - x_0)|.$$

Since $x - x_0 \in (T - T) \cap \varepsilon B_2^n$, then by the definition of Λ and (3.5) it follows that

$$|\tilde{\Gamma}(x - x_0)| \leq 2 \sup_{z \in \Lambda} |\tilde{\Gamma}z| \leq 4\varepsilon. \tag{3.6}$$

Combining this with (3.4) implies that $1 - 5\varepsilon \leq |\tilde{\Gamma}x| \leq 1 + 5\varepsilon$, completing the proof, by the definition of ε. ∎

We are now ready for an elementary solution to Question 1.2, contained in the following corollary.

Corollary 3.3 *Let* $\tilde{\Gamma}$ *satisfy* (3.1) *and* (3.2). *Then, there are constants* c_1, c_1' *and* c_2 *depending only on* c_0 *from* (3.2) *for which the following holds. For every* $0 < \theta < 1$, *with probability at least* $1 - 2\exp(-c_2\theta^2 k)$, $\tilde{\Gamma}$ *satisfies uup*(θ, λ) *for*

$$\lambda = \frac{c_1 \log(c_1' n / k\theta^3)}{\theta^2}.$$

In particular, there is are absolute constants c_1, c_2 *and* c_3 *for which the following holds. If* X *is an isotropic,* ψ_2 *vector with constant* α *then with probability at least* $1 - \exp(-c_1\theta^2 k/\alpha^4)$, *the operator* $\tilde{\Gamma} = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i, \cdot \rangle e_i$ *satisfies uup*(θ, λ) *for*

$$\lambda = \frac{c_2 \alpha^4}{\theta^2 \log(c_3 n \alpha^4 / k\theta^3)}.$$

Proof. The main part of the proof is to show that there exists $c' > 0$ such that, given $0 < \theta < 1$, if m and k satisfy

$$k \geq \frac{c' m}{\theta^2} \log\left(\frac{c' n}{m\theta}\right), \tag{3.7}$$

then (3.3) holds, that is, $\tilde{\Gamma}$ acts on U_m in an almost norm preserving way. To that end we need to exhibit the sets Λ and Λ'. For the latter set, one can choose c' in such a way that the set Λ' constructed in the moreover part of Lemma 2.3 for $\varepsilon = \theta/5$ satisfies the required condition (i) for $T = U_m$. For the former set, apply the third part of Lemma 2.3 with $r = \theta/5$ to get $\bar{\Lambda}$; adjusting the choice of c' in (3.7), $\bar{\Lambda}$ satisfies (ii).

Now the conclusion follows from (3.7) by a straightforward computation. ■

Remark 3.4 Note that the price for using the elementary approach in the case of U_m - and thus for Question 1.2 is not very high - a slightly worse power of θ in the logarithm. However, there are many cases of sets $T \subset S^{n-1}$ in which this elementary approach would not be enough to show that $\tilde{\Gamma}$ acts in an almost norm preserving way on T.

3.2 The approximate reconstruction problem

Next, we show how the elementary approach can be used to solve the approximate reconstruction problem in several cases that have been considered in [CT1, CT2, D, MPT1, MPT2, BDDW], among others. Let us recall the formulation of this problem.

Question 3.5 *Suppose that $\tilde{T} \subset \mathbb{R}^n$ and fix $t_0 \in \tilde{T}$. Let Γ be a $k \times n$ random matrix, and suppose that one is given the data vector Γt_0, that is, the set of linear measurements $(\langle X_i, t_0 \rangle)_{i=1}^k$. Is it possible to find (with high probability) some $x \in \mathbb{R}^n$, such that $|x - t_0|$ is small?*

In [CT1] this problem has been studied by using the UUP and for particular sets – B_1^n, the unit ball in ℓ_1^n and $B_{p,\infty}^n$ for $0 < p < 1$, the unit balls in weak ℓ_p spaces. In [MPT1, MPT2], a geometric approach was introduced which solved this problem for an arbitrary symmetric quasi-convex subset of $\mathbb{R}^n$. (Recall that a (centrally) symmetric set $\tilde{T}$ is quasi-convex with constant $a \geq 1$, if $\tilde{T} + \tilde{T} \subset 2a\tilde{T}$ and $\tilde{T}$ is star-shaped, i.e., $s\tilde{T} \subset \tilde{T}$ for $0 < s < 1$.)

The geometric idea at the heart of [MPT1, MPT2] is essentially the following: let $\tilde{T} \subset \mathbb{R}^n$ and suppose that one can find ε_k and show that with high probability,

$$\operatorname{diam}\left(\ker(\Gamma) \cap \tilde{T}\right) \leq \varepsilon_k.$$

Since $\tilde{T}$ is quasi-convex, then $\tilde{T} - \tilde{T} \subset 2a\tilde{T}$, for some $a \geq 1$. Hence, $\operatorname{diam}\left(\ker(\Gamma) \cap (\tilde{T} - \tilde{T})\right) \leq 2a\varepsilon_k$. In particular, with high probability, if x is in $\tilde{T}$ and it satisfies $\Gamma x = \Gamma t_0$ then $|x - t_0| \leq 2a\varepsilon_k$, as required.

In other words, the approximate reconstruction problem is reduced to finding an upper estimate on the diameter of the intersection of the kernel of Γ with $\tilde{T}$ that holds with high probability. This parameter has been studied in asymptotic geometry and in approximation theory for certain notions of randomness, and is the random k-th *Gelfand number* of T associated with the random matrix Γ.

Theorem 3.1 provides a method for estimating the diameter of $\ker(\Gamma) \cap \tilde{T}$ in the following way. For $\tilde{T} \subset \mathbb{R}^n$ star-shaped let $T_\rho = \tilde{T} \cap \rho S^{n-1}$. Then if (a multiple of) Γ acts on T_ρ in an almost norm preserving way, then $\ker(\Gamma) \cap \tilde{T} \subset \rho B_2^n$, and thus $\operatorname{diam}\left(\ker(\Gamma) \cap \tilde{T}\right) \leq \rho$. Indeed, if not, there would be a point $t \in \tilde{T}$ of norm greater than ρ which is mapped to 0. Hence, $\rho t/|t| \in T$ will also be mapped to 0, which contradicts the fact that (a multiple of) Γ is almost norm preserving on $\tilde{T}$.

This proves the following Corollary.

Corollary 3.6 *Let $\tilde{\Gamma}$ be as in Theorem 3.1. Let $\tilde{T} \subset \mathbb{R}^n$ be star-shaped. Let $T = \rho^{-1}(\tilde{T} \cap \rho S^{n-1})$ and assume that T satisfies the hypothesis of Theorem 3.1 for some $0 < \theta < 1$ (say, $\theta = 1/2$). Then $\operatorname{diam}\left(\ker(\tilde{\Gamma}) \cap \tilde{T}\right) \le \rho$, with probability at least $1 - 2\exp(-ck)$, where $c > 0$ is an absolute constant.*

To illustrate this corollary, we consider examples of $\tilde{T}$: the unit ball in ℓ_1^n, denoted by B_1^n, and the unit balls in ℓ_p^n and the weak-ℓ_p^n spaces $\ell_{p,\infty}^n$ for $0 < p < 1$, denoted by B_p^n and $B_{p,\infty}^n$, respectively. Recall that $B_{p,\infty}^n$ is the set of all $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ such that the cardinality $|\{i : |x_i| \ge s\}| \le s^{-p}$ for all $s > 0$. Note that $B_p^n \subset B_{p,\infty}^n$ so we can restrict ourselves to considering the balls $B_{p,\infty}^n$ only.

We will require two lemmas. The first lemma comes from [MPT2] and it combines a reformulation of Lemma 3.2 and (3.1) from that article.

Lemma 3.7 *Let $0 < p < 1$, $1 \le m \le n$ and set $r = (1/p - 1)m^{1/p-1/2}$. Then, for every $x \in \mathbb{R}^n$,*

$$\sup_{z \in rB_{p,\infty}^n \cap B_2^n} \langle x, z \rangle \le 2 \left(\sum_{i=1}^m x_i^{*2} \right)^{1/2},$$

where $(x_i^*)_{i=1}^n$ is a non-increasing rearrangement of $(|x_i|)_{i=1}^n$. Equivalently,

$$rB_{p,\infty}^n \cap B_2^n \subset 2 \operatorname{conv} \tilde{U}_m. \tag{3.8}$$

Furthermore,

$$\sqrt{m} B_1^n \cap B_2^n \subset 2 \operatorname{conv} \tilde{U}_m. \tag{3.9}$$

The second lemma shows that $m^{1/p-1/2} B_{p,\infty}^n \cap S^{n-1}$ is well approximated by vectors on the sphere with a relatively short support.

Lemma 3.8 *Let $0 < p < 2$ and $\delta > 0$, set $\varepsilon = 2(2/p - 1)^{-1/2} \delta^{1/p-1/2}$. Then $U_{\lceil m/\delta \rceil}$ is an ε-cover of $m^{1/p-1/2} B_{p,\infty}^n \cap S^{n-1}$ with respect to the Euclidean metric.*

Proof. Let $x \in m^{1/p-1/2} B^n_{p,\infty} \cap S^{n-1}$ and assume without loss of generality that $x_1 \geq x_2 \geq \ldots \geq x_n \geq 0$. Define z' by $z_i = x_i$ for $1 \leq i \leq \lceil m/\delta \rceil$ and $z'_i = 0$, otherwise. Then

$$|x - z'|^2 = \sum_{i > m/\delta} |x_i|^2 \leq m^{2/p-1} \sum_{i > m/\delta} 1/i^{2/p} \leq (2/p-1)^{-1}\, \delta^{2/p-1}.$$

Thus $1 \geq |z'| \geq 1 - (2/p-1)^{-1/2}\, \delta^{1/p-1/2}$. Put $z = z'/|z'|$. Then $z \in U_{\lceil m/\delta \rceil}$ and

$$|z - z'| = 1 - |z'| \leq (2/p-1)^{-1/2}\, \delta^{1/p-1/2}.$$

By the triangle inequality $|x - z| \leq \varepsilon$, completing the proof. ■

Let $0 < p < 1$. Fix $1 \leq m \leq n$, set $\tilde{T} := m^{1/p-1/2} B^n_{p,\infty}$ and $T := \tilde{T} \cap S^{n-1}$. We shall show that for appropriately chosen m, T satisfies the hypothesis of Theorem 3.1 for $\theta = 1/2$. To that end, we need to show that the complexity of the set T as captured by the sets Λ and Λ' is small.

First note, to simplify the calculations a little, that by Lemma 3.8, for $\delta > 0$ the set $U_{\lceil m/\delta \rceil}$ is an ε cover for T, where $\varepsilon = 2\sqrt{\delta}$. (That is, the dependence of ε on δ is universal in the range of p considered here.) Use this fact for $\delta = 1/40^2$ and combine it with the "moreover part" of Lemma 2.3 (for $\varepsilon = 1/20$) which provides us with a set $\Lambda' \subset S^{n-1}$ which is $1/20$ cover of $U_{\lceil m/\delta \rceil}$. Hence, by the triangle inequality, Λ' is $1/20 + 1/20 = 1/10$ cover of T. Moreover, by (2.7), $|\Lambda'| \leq \exp(c_1 m \log(c_1 n/m))$, where $c_1 > 0$ is an absolute constant.

It is easy to check that that $B_{p,\infty}$ is quasi-convex with constant $2^{1/p}$ and therefore

$$(T - T) \cap \frac{1}{10} B^n_2 \subset \left(2^{1+1/p} \tilde{T} \cap 2B^n_2\right) \cap \frac{1}{10} B^n_2 = 2^{1+1/p} \tilde{T} \cap \frac{1}{10} B^n_2 = \frac{1}{10} A,$$

where

$$A := \left(10 \cdot 2^{1+1/p}\right) \tilde{T} \cap B^n_2 = \left(10 \cdot 2^{1+1/p}\, m^{1/p-1/2}\right) B^n_{p,\infty} \cap B^n_2.$$

Set $m_1 = \max\left(c'_p m, m\right)$ where ${c'_p}^{1/p-1/2} = (1/p-1)^{-1} 20 \cdot 2^{1/p}$, so that

$$10 \cdot 2^{1+1/p}\, m^{1/p-1/2} \leq (1/p-1) m_1^{1/p-1/2}.$$

Then, by (3.8), $A \subset 2 \operatorname{conv} \tilde{U}_{m_1}$. By the first part of Lemma 2.3 there is a subset $\Lambda_1 \subset B^n_2$ such that $\tilde{U}_{m_1} \subset 2 \operatorname{conv} \Lambda_1$ and $|\Lambda_1| \leq \exp(c'_1 m_1 \log(c'_1 n/m_1))$, where $c'_1 > 0$ is an absolute constant. Letting $\Lambda = \frac{1}{10}\Lambda_1$ yields $(T - T) \cap \frac{1}{10}B_2^n \subset 4\,\mathrm{conv}\,\Lambda$ and $|\Lambda| \le \exp\left(c''_p m \log(c'_1 n/m)\right)$, where $c''_p \ge 1$ depends on p only. (The precise form of c''_p can be easily calculated from the form of c'_p but we shall not do it here.)

Considering the upper bounds for Λ' and Λ yields the existence of $c_p \ge 1$, depending on p only, and of an absolute constant $c'_1 > 0$ such that whenever k satisfies

$$k \ge c_p m \log\left(\frac{c'_1 n}{m}\right), \tag{3.10}$$

then $T = m^{1/p-1/2}B_{p,\infty}^n \cap S^{n-1}$ satisfies assumptions (i) and (ii) of Theorem 3.1. Therefore, by Corollary 3.6, $\mathrm{diam}\left(\ker(\tilde{\Gamma}) \cap B_{p,\infty}^n\right) \le m^{1/2-1/p}$, with high probability.

For $p = 1$, an analogous result holds for B_1^n: if k and m satisfy (3.10) (with c_p replaced by a certain absolute constant) then $\mathrm{diam}\left(\ker(\tilde{\Gamma}) \cap B_1^n\right) \le m^{-1/2}$, with high probability.

A straightforward calculation then leads to the following estimates for the diameters of the intersection of $\ker(\tilde{\Gamma})$ with the balls B_p^n and $B_{p,\infty}^n$ (for $0 < p < 1$) and of B_1^n.

Corollary 3.9 *Let $\tilde{\Gamma}$ be as in Theorem 3.1. Let $0 < p < 1$. There exist a constant c_p depending only on p, a constant c depending on c_0 and an absolute constant c_1, such that, with probability at least $1 - \exp(-ck)$,*

$$\mathrm{diam}\left(\ker(\tilde{\Gamma}) \cap B_p^n\right) \le \mathrm{diam}\left(\ker(\tilde{\Gamma}) \cap B_{p,\infty}^n\right) \le c_p \left(\frac{\log(c_1 n/k)}{k}\right)^{1/p-1/2}.$$

In particular, of $t_0 \in B_{p,\infty}^n$ and $\tilde{\Gamma}x = \tilde{\Gamma}t_0$ then with high probability,

$$|x - t_0| \le c'_p \left(\frac{\log(c_1 n/k)}{k}\right)^{1/p-1/2}.$$

For $p = 1$, an analogous result holds for the ball B_1^n replacing $B_{p,\infty}^n$ and c_p and c'_p being replaced by an absolute constant.

References

[BDDW] Baraniuk, R., Davenport, M., De Vore, R., & Wakin, M.: *The Johnson-Lindenstrauss Lemma meets compressed sensing*, Preprint.

[CT1] Candes, E. & Tao, T.: *Near optimal recovery from random projections: universal encoding strategies*, IEEE Trans. Inform. Theory, to appear

[CT2] Candes, E. & Tao, T.: *Decoding by linear programming*, IEEE Trans. Inform. Theory, to appear

[CRT] Candes, E., Romberg, J. & Tao, T.: *Stable Signal Recovery from Incomplete and Inaccurate Measurements*, Comm. Pure Appl. Math., to appear.

[D] Donoho D. L. *Compressed sensing*, IEEE Trans. on Inform. Theory, 52 (2006), 1289–1306.

[LT] Ledoux, M. & Talagrand, M. Probability in Banach spaces. Isoperimetry and processes, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 23. Springer-Verlag, Berlin, 1991.

[MPT1] Mendelson, S., Pajor A. & Tomczak-Jaegermann N.: *Reconstruction and subgaussian processes*, C. R. Acad. Sci. Paris, Sér. I Math., 340 (2005), 885-888.

[MPT2] Mendelson, S., Pajor A. & Tomczak-Jaegermann N.: *Reconstruction and subgaussian operators in Asymptotic Geometric Analysis*, Geometric and Functional Analysis, to appear.

[P] Pisier, G. The volume of convex bodies and Banach space geometry, (1989), Cambridge University Press.

[Ta] Talagrand, M. The generic chaining, Springer, 2005.

[T] Tao, T. personal communication

[VW] Van der Vaart, A. W. & Wellner, J. A. Weak convergence and empirical processes, Springer Verlag, 1996.

S. Mendelson Centre for Mathematics and its Applications, The Australian National University, Canberra, ACT 0200, Australia
`shahar.mendelson@anu.edu.au`

A. Pajor Laboratoire d'Analyse et Mathématiques Appliquées, Université de Marne-la-Vallée, 5 boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallee, Cedex 2, France
`alain.pajor@univ-mlv.fr`

N. Tomczak-Jaegermann Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1
`nicole@ellpspace.math.ualberta.ca`