Geometry determined by random matrices associated to high-dimensional convex bodies

Nicole Tomczak-Jaegermann

Banach space theory workshop

BIRS, March 5 – March 9, 2012

Motivation

Series of joint papers by (subsets of):
Radosław Adamczak,
Olivier Guédon,
Rafał Latała,
Alexander Litvak,
Alain Pajor,
NT-J.

Motivation

Series of joint papers by (subsets of):
Radosław Adamczak,
Olivier Guédon,
Rafał Latała,
Alexander Litvak,
Alain Pajor,
NT-J.

Motivation coming from

- Convexity and Computational Geometry;
- Compressed sensing, in particular approximate reconstruction problems and the Restricted Isometry Property;
- Point of view of Random Matrix Theory.

Notation, Isotropicity

$\mathbb{R}^n$ with the canonical inner product $\langle \cdot, \cdot \rangle$. $|\cdot|$ is the natural Euclidean norm, also the normalized volume on $\mathbb{R}^n$, or the cardinality of a set.

By a random vector $X \in \mathbb{R}^n$, we mean a measurable function defined on a probability space and taking values in $\mathbb{R}^n$. $\mathbb{E}$ is the expectation.

Notation, Isotropicity

$\mathbb{R}^n$ with the canonical inner product $\langle \cdot, \cdot \rangle$. $|\cdot|$ is the natural Euclidean norm, also the normalized volume on $\mathbb{R}^n$, or the cardinality of a set.

By a random vector $X \in \mathbb{R}^n$, we mean a measurable function defined on a probability space and taking values in $\mathbb{R}^n$. $\mathbb{E}$ is the expectation.

A random vector $X \in \mathbb{R}^n$ is called isotropic if

$$\mathbb{E}\langle X, y\rangle = 0, \quad \mathbb{E}\,|\langle X, y\rangle|^2 = |y|^2 \quad \text{for all } y \in \mathbb{R}^n.$$

In other words, if X is centered and its covariance matrix is the identity:

$$\mathbb{E}\, X \otimes X = \mathrm{Id}.$$

($X \otimes X$ is the linear operator on $\mathbb{R}^n$ given by the $n \times n$ matrix $[x_i\, x_j]_{i,j}$ where x_i is the ith coordinnate of X.)

For every random vector X not supported on any $n-1$ dimensional hyperplane, there exists an affine map $T\colon \mathbb{R}^n \to \mathbb{R}^n$ such that TX is isotropic.

Notation, log-concavity

A measure μ on $\mathbb{R}^n$ is log-concave if for any measurable subsets A, B of $\mathbb{R}^n$ and any $\theta \in [0, 1]$,

$$\mu(\theta A + (1-\theta)B) \geqslant \mu(A)^{\theta}\mu(B)^{(1-\theta)}$$

whenever the following *Minkowski sum* is measurable:

$$\theta A + (1-\theta)B = \{\theta x_1 + (1-\theta)x_2 \, : \, x_1 \in A, \; x_2 \in B\}$$

[Borell] Log-concave measures not supported by any $(n-1)$ dimensional hyperplanes are exactly those which are absolutely continuous w.r. to the Lebesgue measure, and have log-concave densities, that is, densities of the form $\exp(-V(x))$, where $V \colon \mathbb{R}^n \to (-\infty, \infty]$ is convex.

Examples

1. Let $K \subset \mathbb{R}^n$ be a convex body ($=$ compact convex, with non-empty interior) (symmetric means $-K = K$).
X a random vector uniformly distributed in K. Then the corresponding probability measure on $\mathbb{R}^n$

$$\mu_K(A) = \frac{|K \cap A|}{|K|}$$

is log-concave (by **Brunn-Minkowski**).
Moreover, for every convex body K there exists an affine map T such that μ_{TK} is isotropic.

Examples

1. Let $K \subset \mathbb{R}^n$ be a convex body ($=$ compact convex, with non-empty interior) (symmetric means $-K = K$).
X a random vector uniformly distributed in K. Then the corresponding probability measure on $\mathbb{R}^n$

$$\mu_K(A) = \frac{|K \cap A|}{|K|}$$

is log-concave (by **Brunn-Minkowski**).
Moreover, for every convex body K there exists an affine map T such that μ_{TK} is isotropic.

2. The Gaussian vector $G = (g_1, ..., g_n)$, where g_i's have $\mathcal{N}(0, 1)$ distribution, is isotropic and log-concave.

Examples

1. Let $K \subset \mathbb{R}^n$ be a convex body ($=$ compact convex, with non-empty interior) (symmetric means $-K = K$).
X a random vector uniformly distributed in K. Then the corresponding probability measure on $\mathbb{R}^n$

$$\mu_K(A) = \frac{|K \cap A|}{|K|}$$

is log-concave (by **Brunn-Minkowski**).
Moreover, for every convex body K there exists an affine map T such that μ_{TK} is isotropic.

2. The Gaussian vector $G = (g_1, ..., g_n)$, where g_i's have $\mathcal{N}(0, 1)$ distribution, is isotropic and log-concave.

3. Similarly the vector $X = (\xi_1, ..., \xi_n)$, where ξ_i's have exponential distribution (i.e., with density $f(t) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|t|)$, for $t \in \mathbb{R}$)
is isotropic and log-concave.

KLS question, I

$n, N \geqslant 1$, $X \in \mathbb{R}^n$ isotropic log-concave, $(X_i)_{i \leqslant N}$ independent copies of X.
By law of large numbers, the empirical covariance matrix converges to Id.

$$\frac{1}{N}\sum_{i=1}^{N} X_i \otimes X_i \longrightarrow \mathrm{Id} \qquad \text{as } N \to \infty \quad \text{a.s.}$$

KLS question, II

Kannan-Lovász-Simonovits asked (around 1995), motivated by a problem of complexity in computing volume in high dimension:
$X \in \mathbb{R}^n$ *isotropic log-concave. Given* $\varepsilon \in (0,1)$ *estimate* N *for which*

$$\left\| \frac{1}{N} \sum_{i=1}^{N} X_i \otimes X_i - \mathrm{Id} \right\| \leqslant \varepsilon$$

holds with high probability.

KLS question, II

Kannan-Lovász-Simonovits asked (around 1995), motivated by a problem of complexity in computing volume in high dimension:

$X \in \mathbb{R}^n$ *isotropic log-concave. Given* $\varepsilon \in (0,1)$ *estimate* N *for which*

$$\sup_{y \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} |\langle X_i, y \rangle|^2 - 1 \right| = \left\| \frac{1}{N} \sum_{i=1}^{N} X_i \otimes X_i - \mathrm{Id} \right\| \leqslant \varepsilon$$

holds with high probability.

KLS question, II

Kannan-Lovász-Simonovits asked (around 1995), motivated by a problem of complexity in computing volume in high dimension:
$X \in \mathbb{R}^n$ *isotropic log-concave. Given* $\varepsilon \in (0,1)$ *estimate* N *for which*

$$\sup_{y \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} |\langle X_i, y \rangle|^2 - 1 \right| = \left\| \frac{1}{N} \sum_{i=1}^{N} X_i \otimes X_i - \mathrm{Id} \right\| \leqslant \varepsilon$$

holds with high probability.

KLS showed that for any $\varepsilon, \delta \in (0,1)$ (under a finite third moment assumption), $N \geqslant (C/\varepsilon\delta) n^2$ gives the required approximation, with probability $1 - \delta$.

Bourgain (1996): for any $\varepsilon, \delta \in (0,1)$, there exists $C(\varepsilon, \delta) > 0$ such that $N = C(\varepsilon, \delta) n \log^3 n$ gives the approximation with probability $1 - \delta$.

The question generated a lot of activity; Improvement of powers of logarithms by: Rudelson, Giannopoulos, Paouris...

Random matrices with i.i.d.log-concave columns

$n, N \geqslant 1$, $X \in \mathbb{R}^n$ isotropic log-concave, $(X_i)_{i \leqslant N}$ independent copies of X.
Γ is a $n \times N$ matrix with X_i as columns;

$$\sup_{y \in S^{n-1}} \langle (\sum_{i=1}^N X_i \otimes X_i) y, y \rangle = \sup_{y \in S^{n-1}} \langle \Gamma \Gamma^* y, y \rangle = \|\Gamma : \ell_2^N \to \ell_2^n\|^2,$$

with the operator norm $\|\Gamma\|$.

Question An upper bound for $\|\Gamma\|$, with some (large?) probability?

Random matrices with i.i.d.log-concave columns

$n, N \geqslant 1$, $X \in \mathbb{R}^n$ isotropic log-concave, $(X_i)_{i \leqslant N}$ independent copies of X.
Γ is a $n \times N$ matrix with X_i as columns;
$\sup_{y \in S^{n-1}} \langle (\sum_{i=1}^N X_i \otimes X_i) y, y \rangle = \sup_{y \in S^{n-1}} \langle \Gamma \Gamma^* y, y \rangle = \|\Gamma : \ell_2^N \to \ell_2^n\|^2$,
with the operator norm $\|\Gamma\|$.

Question An upper bound for $\|\Gamma\|$, with some (large?) probability?

Note $\mathbb{E}\|\Gamma\|^2 \geqslant \mathbb{E}|X_i|^2 = n$ (comparing with norms of columns);
denoting the rows of Γ by Y_j we also have

$$\mathbb{E}\|\Gamma\|^2 \geqslant \max_{j \leqslant N} \mathbb{E}|Y_j|^2 \geqslant \frac{1}{n} \sum_{j=1}^{n} \mathbb{E}|Y_j|^2 \geqslant N.$$

Consequently, with positive probability, $\|\Gamma\| \geqslant \max(\sqrt{n}, \sqrt{N})$.

We will show: with overwhelming probability, this is asymptotically the right order:

$$\|\Gamma\| \leqslant C \max(\sqrt{n}, \sqrt{N}).$$

The same behaviour as e.g., random Gaussian matrix (with $N(0,1)$ independent entries)

Norm of matrices with i.i.d. log-concave columns

Given $1 \leqslant k \leqslant N$, we let

$$\Gamma(k) = \sup_{\substack{z \in S^{N-1} \\ |\operatorname{supp} z| \leqslant k}} |\Gamma z|,$$

the norm of Γ on k-sparse vectors.

Theorem 1 [ALPT] *Let $n \geqslant 1$ and $1 \leqslant N \leqslant e^{\sqrt{n}}$. Let $X \in \mathbb{R}^n$ be isotropic log-concave random vector and $X_1, \dots, X_N$ be i.i.d. copies of X. Let Γ be the $n \times N$ random matrix having the X_is as the columns. Then for any $t \geqslant 1$, the following holds with probability $\geqslant 1 - e^{-ct\sqrt{n}}$:*

$$\forall k \leqslant N \,:\, \Gamma(k) \leqslant Ct \left(\sqrt{n} + \sqrt{k} \log \frac{2N}{k} \right),$$

where $C, c > 0$ are absolute constants. In particular, with the same probability,

$$\|\Gamma\| \leqslant Ct(\sqrt{n} + \sqrt{N}).$$

The norm estimate is optimal, up to universal constants, as seen for the exponential distribution.

Paouris' large deviation estimate

Let $X \in \mathbb{R}^n$ be isotropic log-concave random vector.
Consider the matrix that consists of just one column then its norm is equal to $|X|$.

Note that $(\mathbb{E}|X|^2)^{1/2} = \sqrt{n}$. What is the large deviation
$\mathbb{P}\left\{ |X| \geqslant t\sqrt{n} \right\} \leqslant ?$ for $t \geqslant C_0$ where $C_0 > 1$ an absolute constant.

Paouris' large deviation estimate

Let $X \in \mathbb{R}^n$ be isotropic log-concave random vector.
Consider the matrix that consists of just one column then its norm is equal to $|X|$.

Note that $(\mathbb{E}|X|^2)^{1/2} = \sqrt{n}$. What is the large deviation
$\mathbb{P}\left\{ |X| \geqslant t\sqrt{n} \right\} \leqslant ?$ for $t \geqslant C_0$ where $C_0 > 1$ an absolute constant.

Paouris' large deviation theorem (2005)**:** *There exist constants* $C_0 > 1$, $c > 0$ *such that the following holds: Let* X *be an isotropic, log-concave random vector on* $\mathbb{R}^n$*. Then for all* $t \geqslant C_0$,

$$\mathbb{P}\left\{|X| \geqslant t\sqrt{n}\right\} \leqslant \exp(-ct\sqrt{n}).$$

Note that $\mathbb{E}|X| \sim \sqrt{n} = (\mathbb{E}|X|^2)^{1/2}$.

Norm of random matrices, for specialists

Theorem 1 remains valid for a larger class of distributions. Let $1 \leqslant p < \infty$. For a (real valued) random variable Z,

$$\|Y\|_{\psi_1} = \inf\{C > 0\,;\ \mathbb{E}\exp\left(|Z|^p/C\right) \leqslant 2\}.$$

For $p = 2$ so-called *subgaussian*; for $p = 1$ a large class – for example, all log-concave distributions are ψ_1.

Theorem 2 *Let n, N be integers and $X_1, \ldots, X_N$ be independent random vectors in $\mathbb{R}^n$ such that*

$$\sup_{i \leqslant N} \sup_{y \in S^{n-1}} \|\langle X_i, y\rangle\|_{\psi_1} \leqslant \psi.$$

Then for every $k \leqslant N$ and $t \geqslant 1$ one has

$$\begin{aligned} \mathbb{P}\left(\Gamma(k) \geqslant \max_{i \leqslant N} |X_i| + C\psi t\sqrt{k}\log\frac{2N}{k}\right) \\ \leqslant \quad (1 + 2\log m)\exp\left(-t\sqrt{k}\log\frac{2N}{k}\right). \end{aligned}$$

Empirical moments, Answer to KLS question

Theorem 3 [ALPT] *Let $(X_i)_{i \leqslant N}$ be independent isotropic log-concave random vectors on $\mathbb{R}^n$. Then with probability at least $1 - 2\exp(-c\sqrt{n})$ one has*

$$\sup_{y \in S^{n-1}} \left| \frac{1}{N} \sum_{i=1}^{N} \left(|\langle X_i, y\rangle|^2 - \mathbb{E}|\langle X_i, y\rangle|^2 \right) \right| \leqslant C\sqrt{n/N},$$

where $C, c > 0$ are universal constants.

It uses the full strength of Theorem 1, which provided deviation inequalities for *norms on sparse vectors*. So in fact we reduced a concentration inequality above to deviation inequalities.
The proof by a natural approach for empirical processes.
This of course implies that for every $\varepsilon \in (0, 1)$ the appropriate difference is $\leqslant \varepsilon$, whenever $N \geqslant C'n/\varepsilon^2$.

Point of view of Random Matrix Theory, I

Random Matrix Theory studies matrices of finite size whose entries are random variables, traditionally they are *i.i.d.*; we look for **limiting** results as the size $\to \infty$.

In AGA: we consider finite matrices with a fixed size; typically we expect results in form of inequalities or estimates with constants independent on the size; the size might be required to be “sufficiently large” depending on parameters of the problem, and in this sense we study **asymptotic** behaviour.

This approach is recently actively developed in various frameworks. Notable contributions to this general direction by Mark Rudelson and Roman Vershynin

It follows from our results that properties of log-concave random vectors and spectral properties of matrices with independent log-concave rows (or columns); in high dimensions behave similarly as if the coordinates were independent; or even independent Gaussian.

Point of view of Random Matrix Theory, II

I will just give one example.
Fix $\beta \in (0,1)$ and let $\lim_n \frac{n}{N} = \beta$.
Bai-Yin showed that if random $n \times N$ matrices $A^{(n)}$ have i.i.d entries (which satisfy some mild moment assumptions) then

$$\lim \lambda_n(A^{(n)*}A^{(n)})/N = (1-\sqrt{\beta})^2 \qquad \lim \lambda_1(A^{(n)*}A^{(n)})/N = (1+\sqrt{\beta})^2$$

Point of view of Random Matrix Theory, II

I will just give one example.
Fix $\beta \in (0, 1)$ and let $\lim_n \frac{n}{N} = \beta$.
Bai-Yin showed that if random $n \times N$ matrices $A^{(n)}$ have i.i.d entries (which satisfy some mild moment assumptions) then

$$\lim \lambda_n(A^{(n)*}A^{(n)})/N = (1 - \sqrt{\beta})^2 \qquad \lim \lambda_1(A^{(n)*}A^{(n)})/N = (1 + \sqrt{\beta})^2$$

In contrast, Theorem 3 implies quantitative estimates:
There is $C \geqslant 1$ such that setting $\beta = \frac{n}{N} \in (0, 1)$, we get with overwhelming probability and for every $1 \leqslant j \leqslant n$,

$$\left| \frac{\sqrt{\lambda_j}}{\sqrt{N}} - 1 \right| \leqslant C\sqrt{\beta}.$$

Singular Values

Let $1 \leqslant n \leqslant N$. Let $X \in \mathbb{R}^n$ be an isotropic log-concave random vector
let $X_1, \ldots, X_N$ be i.i.d. copies of X.
Let Γ be the $n \times N$ random matrix having the X_is as the columns.
$\Gamma : \ell_2^N \to \ell_2^n$ and $\Gamma^* : \ell_2^n \to \ell_2^N$

Singular values of Γ = eigenvalues of $\sqrt{\Gamma\Gamma^*}$.

$$\|\Gamma\| = s_1(\Gamma) \geqslant \ldots \geqslant s_n(\Gamma) = \frac{1}{\|(\Gamma^*)^{-1}\|}.$$

The smallest $\neq 0$ singular value of Γ = the smallest $\neq 0$ singular value of $\Gamma^* = \inf_{y \in S^{n-1}} |\Gamma^* y|$

Smallest singular numbers, square matrices

Theorem 4 [AGLPT] *Let Γ be an $n \times n$ matrix whose columns are i.i.d. distributed acording to an isotropic log-concave random vector in $\mathbb{R}^n$. For every $\varepsilon \in (0, 1)$,*

$$\mathbb{P}\left(\inf_{y \in S^{n-1}} |\Gamma^* y| \leqslant c\varepsilon n^{-1/2}\right) \leqslant C \min\left\{n\varepsilon, \varepsilon + e^{-c\sqrt{n}}\right\} \leqslant C\varepsilon^{1/2},$$

if n sufficiently large, and c, C are absolute positive constants.

Smallest singular numbers, square matrices

Theorem 4 [AGLPT] *Let Γ be an $n \times n$ matrix whose columns are i.i.d. distributed acording to an isotropic log-concave random vector in $\mathbb{R}^n$. For every $\varepsilon \in (0,1)$,*

$$\mathbb{P}\left(\inf_{y \in S^{n-1}} |\Gamma^* y| \leqslant c\varepsilon n^{-1/2}\right) \leqslant C \min\left\{n\varepsilon, \varepsilon + e^{-c\sqrt{n}}\right\} \leqslant C\varepsilon^{1/2},$$

if n sufficiently large, and c, C are absolute positive constants.

One of the points is that we get probability small, with fixed n, by choice of ε.

This gives a lower bound for $s_n(\Gamma)$ with large probability.
Note that for a Gaussian matrix the same type of estimate is valid, with the only change in probability, which is $\leqslant C\varepsilon$ (independently by Edelstein and Szarek).

Random polytopes

Let $1 \leqslant m \leqslant n \leqslant N$and let $X_1, \ldots, X_N \in \mathbb{R}^n$. Denote by Γ the $n \times N$ matrix with $X_1, \ldots, X_N$ as columns and by $K(\Gamma) = K(X_1, \ldots, X_N)$ the convex hull of $\pm X_1, \ldots, \pm X_N$.
Recall that a centrally symmetric convex polytope is m-centrally-neighborly if any set of less than m vertices containing no-opposite pairs, is the vertex set of a face.
D. Donoho proved that the following are equivalent:

i) $K(\Gamma)$ has $2N$ vertices and is m-neighborly

ii) given $y \in \mathbb{R}^n$ of a form $y = \Gamma z$ for some $z \in \mathbb{R}^N$ having at most m non-zero coordinates (in other words z is m-sparse), then z is the unique solution of the problem

$$(P) \qquad \min \|t\|_{\ell_1}, \quad \Gamma t = y.$$

Here the ℓ_1-norm is defined by $\|t\|_{\ell_1} = \sum_{i=1}^N |t_i|$ for any $t = (t_i)_{i=1}^N \in \mathbb{R}^N$.

Restricted Isometry Property

introduced by E. Candes and T. Tao (2005):

Let M be a $n \times N$ matrix. For any $1 \leqslant m \leqslant \min(n, N)$, the isometry constant of M is defined as the smallest number $\delta_m = \delta_m(M)$ so that

$$(1 - \delta_m)|z|^2 \leqslant |Mz|^2 \leqslant (1 + \delta_m)|z|^2$$

holds for all m-sparse vectors $z \in \mathbb{R}^N$. The matrix M is said to satisfy the Restricted Isometry Property of order m with parameter δ, if $0 \leqslant \delta_m(M) < \delta$.

It provides a quantitative sufficient condition for the basis pursuit condition (ii). Huge literature and many statements of the type: if $\delta_{2m}(M) < \sqrt{2} - 1$ then (ii) is satisfied (hence also (i)) (Candes, 2008)

Log-concave random polytopes and RIP

Let $1 \leqslant m \leqslant n \leqslant N$*. Let* $X_1, \dots, X_N \in \mathbb{R}^n$ *be i.i.d. log-concave random vectors and let* Γ *be the matrix having the* X_i*'s as columns. Then, for any* $N \leqslant \exp(\sqrt{n})$*, with probability at least* $1 - C\exp(-c\sqrt{n})$*, the polytope* $K(\Gamma)$ *is* m*-centrally-neighborly, whenever*

$$m \leqslant cn / \log^2(CN/n),$$

where $C, c > 0$ *are universal constants.*

From Theorem 3 it follows that Γ satisfies the RIP of order m.

Note that the definition of the RIP agrees with the structure of Γ given by independent column vectors.

Matrices with log-concave rows

We assume as before that $n \leqslant N$, but now we wish to define the $n \times N$ matrix by rows rather than columns.
Let $Y_1, \ldots, Y_n \in \mathbb{R}^N$ be independent log-concave random vectors and let A be the $n \times N$ random matrix with rows Y_i.
This may be the case of an "exact reconstruction" problem when we consider a small number – namely n – of random measurements in $\mathbb{R}^N$, and we might be interested in the solution of an ℓ_1-minimiization algorithm.
But the RIP condition is expressed in terms of any subset of m columns of A, which destroys the row structure.

Matrices with log-concave rows, II

A is an $n \times N$ matrix. Let $1 \leqslant k \leqslant N$ and $1 \leqslant m \leqslant n$. Set

$$A_{k,m}^2 = \sup_{\substack{y \in S^{N-1} \\ |\operatorname{supp} y| \leqslant m}} \sup_{\substack{I \subset \{1,\dots,n\} \\ |I| = k}} \sum_{i \in I} |\langle Y_i, y \rangle|^2.$$

Theorem 5 [ALLPT] *Let $1 \leqslant n \leqslant N$, and let A be an $n \times N$ random matrix with independent isotropic log-concave rows. For any integers $k \leqslant n$, $m \leqslant N$ and any $t \geqslant 1$, we have*

$$\mathbb{P}\left(A_{k,m} \geqslant Ct\lambda\right) \leqslant \exp(-t\lambda/\sqrt{\log(3m)}),$$

where $\lambda = \sqrt{\log\log(3m)}\sqrt{m}\log(eN/m) + \sqrt{k}\log(en/k)$.

The estimate is optimal, up to the factor of $\sqrt{\log\log(3m)}$. Assuming additionally unconditionality of the distributions of the rows, we can remove this factor and get a sharp estimate.

RIP

Theorem 6 [ALLPT] *Let* $0 < \theta < 1$, $1 \leqslant n \leqslant N$. *Let* A *be an* $n \times N$ *random matrix with independent isotropic log-concave rows. There exists* $c(\theta) > 0$ *such that* $\delta_m(A/\sqrt{n}) \leqslant \theta$ *with overwhelming probability whenever*

$$m \log^2(2N/m) \log\log 3m \leqslant c(\theta) n.$$

Uniform deviation theorem

We extend Paouris's theorem to the following bound on deviations of norm of projections of an isotropic log-concave vector, uniform over all coordinate projections P_I of a fixed rank.

Theorem 7 [ALLPT] *Let* $m \leqslant N$ *and* X *be an isotropic log-concave vector in* $\mathbb{R}^N$*. Then for every* $t \geqslant 1$ *one has*

$$\mathbb{P}\left(\sup_{\substack{I\subset\{1,\dots,N\}\\ |I|=m}} |P_I X| \geqslant Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \leqslant \exp\left(-t\frac{\sqrt{m}}{\sqrt{\log(em)}}\log\left(\frac{eN}{m}\right)\right).$$

Actually our applications require a stronger result in which the bound for probability is improved by involving the parameter σ_X and its inverse σ_X^{-1}.

$$\sigma_X(p) = \sup_{t\in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}.$$