Quantitative invertibility of random matrices: a combinatorial perspective

Vishesh Jain
Massachusetts Institute of Technology

UCLA Probability Seminar

October 24, 2019
The quantitative invertibility problem

Definition (Least singular value)
The least singular value of an $n \times n$ matrix $M_n$ is defined by

$$s_n(M_n) := \inf_{\mathbf{v} \in S^{n-1}} \| M_n \mathbf{v} \|_2.$$
The quantitative invertibility problem

Definition (Least singular value)
The least singular value of an \( n \times n \) matrix \( M_n \) is defined by

\[
s_n(M_n) := \inf_{\mathbf{v} \in S^{n-1}} \| M_n \mathbf{v} \|_2.
\]

Quantitative invertibility problem
What is the probability that \( s_n(M_n) \) is smaller than \( \eta \geq 0 \)?
Regime I: Invertibility of random discrete matrices

Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$. 

Folklore Conjecture: $\Pr(s_n(M_n) = 0) \approx n^{2/3}$. 

Komlós (1967): $\Pr(s_n(M_n) = 0) = o(n)$. 

Kahn, Komlós, and Szemerédi (1995): $\Pr(s_n(M_n) = 0) \approx 0.999n$. 

Tao and Vu (2006, 2007): $\Pr(s_n(M_n) = 0) \approx 0.75n$. 

Bourgain, Vu, and Wood (2010): $\Pr(s_n(M_n) = 0) \approx (1/\sqrt{2})n$. 

Tikhomirov (2018): $\Pr(s_n(M_n) = 0) \approx (0.5 + o(n))n$. 

Vishesh Jain (MIT)
Regime I: Invertibility of random discrete matrices

Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$.

- Folklore Conjecture: $\Pr(s_n(M_n) = 0) \lesssim n^2 2^{-n}$. 
Regime I: Invertibility of random discrete matrices

Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\text{Pr}(s_n(M_n) = 0)$.

- Folklore Conjecture: $\text{Pr}(s_n(M_n) = 0) \lesssim n^2 2^{-n}$.
- Komlós (1967): $\text{Pr}(s_n(M_n) = 0) = o_n(1)$.
Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$.

- **Folklore Conjecture**: $\Pr(s_n(M_n) = 0) \lesssim n^2 2^{-n}$.
- **Komlós (1967)**: $\Pr(s_n(M_n) = 0) = o_n(1)$.
- **Kahn, Komlós, and Szemerédi (1995)**: $\Pr(s_n(M_n) = 0) \lesssim 0.999^n$. 
Regime I: Invertibility of random discrete matrices

Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$.

- **Folklore Conjecture:** $\Pr(s_n(M_n) = 0) \lesssim n^2 2^{-n}$.
- **Komlós (1967):** $\Pr(s_n(M_n) = 0) = o_n(1)$.
- **Kahn, Komlós, and Szemerédi (1995):** $\Pr(s_n(M_n) = 0) \lesssim 0.999^n$.
- **Tao and Vu (2006, 2007):** $\Pr(s_n(M_n) = 0) \lesssim 0.75^n$.
- **Tikhomirov (2018):** $\Pr(s_n(M_n) = 0) \lesssim (\frac{1}{\sqrt{2}})^n$. 

Vishesh Jain (MIT)
Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$.

- **Folklore Conjecture**: $\Pr(s_n(M_n) = 0) \lesssim n^2 2^{-n}$.
- **Komlós (1967)**: $\Pr(s_n(M_n) = 0) = o_n(1)$.
- **Kahn, Komlós, and Szemerédi (1995)**: $\Pr(s_n(M_n) = 0) \lesssim 0.999^n$.
- **Tao and Vu (2006, 2007)**: $\Pr(s_n(M_n) = 0) \lesssim 0.75^n$.
- **Bourgain, Vu, and Wood (2010)**: $\Pr(s_n(M_n) = 0) \lesssim (1/\sqrt{2})^n$. 
Regime I: Invertibility of random discrete matrices

Suppose that each entry of $M_n$ is an independent Rademacher random variable i.e. $+1$ or $-1$ with probability $1/2$ each. Estimate $\Pr(s_n(M_n) = 0)$.

- **Folklore Conjecture**: $\Pr(s_n(M_n) = 0) \lesssim n^2 2^{-n}$.

- **Komlós (1967)**: $\Pr(s_n(M_n) = 0) = o_n(1)$.

- **Kahn, Komlós, and Szemerédi (1995)**: $\Pr(s_n(M_n) = 0) \lesssim 0.999^n$.

- **Tao and Vu (2006, 2007)**: $\Pr(s_n(M_n) = 0) \lesssim 0.75^n$.

- **Bourgain, Vu, and Wood (2010)**: $\Pr(s_n(M_n) = 0) \lesssim (1/\sqrt{2})^n$.

- **Tikhomirov (2018)**: $\Pr(s_n(M_n) = 0) \lesssim (0.5 + o_n(1))^n$. 
Suppose that each entry of $M_n$ is an independent copy of the standard Gaussian. Estimate $s_n(M_n)$ for a ‘typical’ such matrix.
Suppose that each entry of $M_n$ is an independent copy of the standard Gaussian. Estimate $s_n(M_n)$ for a ‘typical’ such matrix.

- **Edelman (1988), Szarek (1991):** $\Pr(s_n(M_n) \leq \epsilon n^{-1/2}) \leq \epsilon$.
- Hence, for ‘most’ such matrices, $s_n(M_n) = \Omega(n^{-1/2})$. 
Suppose that each entry of $M_n$ is an independent copy of the standard Gaussian. Estimate $s_n(M_n)$ for a ‘typical’ such matrix.

- **Edelman (1988), Szarek (1991):** $\Pr(s_n(M_n) \leq \epsilon n^{-1/2}) \leq \epsilon$.

- Hence, for ‘most’ such matrices, $s_n(M_n) = \Omega(n^{-1/2})$.

- **Sankar, Spielman, and Teng (2006):** $\Pr(s_n(A_n + M_n) \leq \epsilon n^{-1/2}) \lesssim \epsilon$.
  Here, $A_n$ is an arbitrary square matrix.
The Spielman-Teng conjecture

Conjecture (Spielman and Teng, ICM 2002)

Suppose that the entries of $M_n$ are independent Rademacher random variables. There exists some constant $c \in (0, 1)$ such that for all $\eta \geq 0$,

$$\Pr \left( s_n(M_n) \leq \eta \right) \leq \sqrt{n\eta} + c^n.$$  

This combines ‘Gaussian behavior’ with the added possibility of singularity.
Theorem (Rudelson and Vershynin, 2007)

Suppose that the entries of $M_n$ are i.i.d. subgaussian random variables with mean 0 and variance 1. Then, there exists $c \in (0, 1)$ such that for all $\eta \geq 0$,

$$\Pr(s_n(M_n) \leq \eta) \lesssim \sqrt{n\eta} + c^n.$$
Resolution of the Spielman-Teng conjecture

**Theorem (Rudelson and Vershynin, 2007)**

Suppose that the entries of $M_n$ are i.i.d. subgaussian random variables with mean 0 and variance 1. Then, there exists $c \in (0, 1)$ such that for all $\eta \geq 0$,

$$\Pr(s_n(M_n) \leq \eta) \lesssim \sqrt{n\eta} + c^n.$$ 

Theorem (Rudelson and Vershynin, 2007)

Suppose that the entries of $M_n$ are i.i.d. subgaussian random variables with mean 0 and variance 1. Then, there exists $c \in (0, 1)$ such that for all $\eta \geq 0$,

$$\Pr(s_n(M_n) \leq \eta) \lesssim \sqrt{n\eta} + c^n.$$ 

- **Livshyts, Tikhomirov, and Vershynin (2019)**: Removed identically distributed assumption.
Thus far, the **high-dimensional geometric methods** used in the proofs of the previous results have failed to address the following important model of random matrices:
Thus far, the **high-dimensional geometric methods** used in the proofs of the previous results have failed to address the following important model of random matrices:

$$M_n := A_n + N_n,$$

where $A_n$ is a ‘large’ fixed complex matrix, and $N_n$ is a random matrix, each of whose entries is an independent copy of a **complex random variable of mean $0$ and variance $1$**.
Thus far, the **high-dimensional geometric methods** used in the proofs of the previous results have failed to address the following important model of random matrices:

\[ M_n := A_n + N_n, \]

where \( A_n \) is a ‘large’ fixed complex matrix, and \( N_n \) is a random matrix, each of whose entries is an independent copy of a complex random variable of mean 0 and variance 1.

Why is this important?
For the strong circular law, known reductions (Girko, 1984; Bai, 1997; Tao and Vu, 2008) show that we need to study $s_n(M_n)$ for

$$M_n = z \cdot \text{Id}_n + \frac{N_n}{\sqrt{n}},$$

with $z \in \mathbb{C}$ fixed.
For the strong circular law, known reductions (Girko, 1984; Bai, 1997; Tao and Vu, 2008) show that we need to study \( s_n(M_n) \) for

\[
M_n = z \cdot \text{Id}_n + \frac{N_n}{\sqrt{n}},
\]

with \( z \in \mathbb{C} \) fixed.

For numerical linear algebra, the smoothed analysis program of Spielman and Teng (2001) considers

\[
M_n = A_n + N_n,
\]

where \( N_n \) represents the random ‘noise’ in the system.
Least singular value of shifted i.i.d. matrices

- For the strong circular law, known reductions (Girko, 1984; Bai, 1997; Tao and Vu, 2008) show that we need to study $s_n(M_n)$ for

$$M_n = z \cdot \text{Id}_n + \frac{N_n}{\sqrt{n}},$$

with $z \in \mathbb{C}$ fixed.

- For numerical linear algebra, the smoothed analysis program of Spielman and Teng (2001) considers

$$M_n = A_n + N_n,$$

where $N_n$ represents the random ‘noise’ in the system.

- Moreover, for these applications, sharp results are not necessary.
Least singular value of shifted i.i.d. matrices

Prior to our work, the best known result in this setting is due to Tao and Vu, based on deep ideas from additive combinatorics.
Least singular value of shifted i.i.d. matrices

Prior to our work, the best known result in this setting is due to Tao and Vu, based on deep ideas from additive combinatorics.


Let $F_n$ be a fixed $n \times n$ complex matrix, each of whose entries is $O(n^B)$. Suppose that the entries of $N_n$ are i.i.d. complex random variables with mean 0 and variance 1. Then, for any $A > 0$, there exists $C > 0$ such that:

$$\Pr \left( s_n(F_n + N_n) \leq n^{-C} \right) \lesssim n^{-A}.$$
Random matrices with dependent entries

- There are significant additional challenges in dealing with random matrices with dependent entries.
There are significant additional challenges in dealing with random matrices with dependent entries.

As a result, there are big gaps even in our knowledge of the singularity probability for most discrete random matrix models.
There are significant additional challenges in dealing with random matrices with dependent entries.

As a result, there are big gaps even in our knowledge of the singularity probability for most discrete random matrix models.

For instance, in the next simplest model of symmetric random Rademacher matrices, the best known upper bound is only of the form \((1/2)^{\sqrt{n}}\) (Campos, Mattos, Morris, and Morrison, 2019), as compared to the conjectured bound of \((1/2 + o_n(1))^n\).
For combinatorial models, such as adjacency matrices of regular (di)graphs, the situation is much more dire – bounds of the form $n^{-1}$ are not known except in special cases.
For combinatorial models, such as adjacency matrices of regular (di)graphs, the situation is much more dire – bounds of the form $n^{-1}$ are not known except in special cases.

In fact, even $o_n(1)$ bounds were only very recently obtained e.g. Huang (2018), Landon, Sosoe, and Yau (2016), Litvak, Lytova, Tikhomirov, Tomczak-Jaegermann, and Youssef (2015), Cook (2014)...
A motivating result

**Theorem (Vershynin, 2011)**

Suppose that $M_n$ is a symmetric matrix, each of whose above diagonal entries is an independent copy of a subgaussian random variable with mean 0 and variance 1. Then,

$$\Pr(s_n(M_n) \leq \eta) \lesssim (\sqrt{m\eta})^{1/9} + \exp(-n^c).$$
A motivating result

**Theorem (Vershynin, 2011)**
Suppose that $M_n$ is a symmetric matrix, each of whose above diagonal entries is an independent copy of a subgaussian random variable with mean 0 and variance 1. Then,

$$\Pr(s_n(M_n) \leq \eta) \lesssim (\sqrt{n\eta})^{1/9} + \exp(-n^c).$$

- As discussed, for many applications, it suffices to have a result with the (optimal) $\sqrt{n}$ replaced by a larger power of $n$. 
We prove bounds of the form

$$\Pr (s_n(M_n) \leq \eta) \lesssim n^C \eta^\delta + \exp(-n^c)$$

in a **simple and unified** way for quite general random matrix models.
Our goals and results

Summary

We prove bounds of the form

$$\Pr (s_n(M_n) \leq \eta) \lesssim n^C \eta^\delta + \exp(-n^c)$$

in a simple and unified way for quite general random matrix models

To this end, we:

- Introduce new tools, in particular for the so-called ‘counting problem in inverse Littlewood–Offord theory’.
Our goals and results

Summary
We prove bounds of the form

\[ \Pr (s_n(M_n) \leq \eta) \lesssim n^C \eta^\delta + \exp(-n^c) \]

in a **simple and unified** way for quite general random matrix models.

To this end, we:

- Introduce new tools, in particular for the so-called ‘counting problem in inverse Littlewood–Offord theory’.

- Introduce new reductions, some of which can even be used in combination with previously known tools.
Our results - Non centered complex matrices


Let $F_n$ be a fixed $n \times n$ complex matrix with operator norm $O(n^B)$. If the entries of $N_n$ are i.i.d. complex random variables with mean 0 and variance 1, then for all $\eta \geq 0$,

$$\Pr\left( s_n(F_n + N_n) \leq \eta \right) \lesssim n^C \eta^{\delta_\eta} + n^{-\omega(1)}.$$
Our results - Non centered complex matrices


Let $F_n$ be a fixed $n \times n$ complex matrix with operator norm $O(n^B)$. If the entries of $N_n$ are i.i.d. complex random variables with mean 0 and variance 1, then for all $\eta \geq 0$,

$$\Pr \left( s_n(F_n + N_n) \leq \eta \right) \lesssim n^C \eta^{\delta_\eta} + n^{-\omega(1)}.$$

Theorem (J., 2019+)

Under the same assumptions,

$$\Pr \left( s_n(F_n + N_n) \leq \eta \right) \lesssim n^C \eta^{\delta_\eta} + \exp(-n^c).$$
Our results - dependent entries

Theorem (J., 2019+)

Let $M_n$ be an $n \times n$ random matrix with independent rows in $\{0, 1\}^n$, each of which sums to $n/2$. Then, for any $\eta \geq 0$,

$$\Pr (s_n(M_n) \leq \eta) \lesssim n^2 \eta + \exp(-n^c).$$

Theorem (Ferber and J., 2018)

Let $M_n$ be an $n \times n$ symmetric matrix whose above diagonal entries are independent Rademacher random variables. Then,

$$\Pr (s_n(M_n) = 0) \lesssim \exp(-n^{1/4}).$$
Our results - dependent entries

**Theorem (J., 2019+)**

Let $M_n$ be an $n \times n$ random matrix with independent rows in $\{0, 1\}^n$, each of which sums to $n/2$. Then, for any $\eta \geq 0$,

$$\Pr \left( s_n(M_n) \leq \eta \right) \lesssim n^2 \eta + \exp(-n^c).$$

**Theorem (Ferber and J., 2018)**

Let $M_n$ be an $n \times n$ symmetric matrix whose above diagonal entries are independent Rademacher random variables. Then,

$$\Pr \left( s_n(M_n) = 0 \right) \lesssim \exp(-n^{1/4}).$$
In the remainder of this talk, I will sketch some of the ideas and techniques that go into the proofs of our results.

In order to motivate them, I will first present a high-level ‘proof template’.
A proof template

- In the remainder of this talk, I will sketch some of the ideas and techniques that go into the proofs of our results.

- In order to motivate them, I will first present a high-level ‘proof template’.

- To keep technicalities to a minimum, our goal in the next few slides will be to discuss how to obtain upper bounds on the probability that an i.i.d matrix is singular in the ‘light-tailed’ setting.
A proof template: the anti-concentration phenomenon

**Definition (Small ball probability)**

The *$r$-ball probability* of a vector $v := (v_1, \ldots, v_n) \in \mathbb{R}^n$ with respect to a random variable $\xi$ is defined by

$$
\rho_{r,\xi}(v) := \sup_{x \in \mathbb{R}} \Pr (|v_1 \xi_1 + \cdots + v_n \xi_n - x| \leq r),
$$

where $\xi_1, \ldots, \xi_n$ are independent copies of $\xi$. 

---

Vishesh Jain (MIT)  
Quantitative invertibility of random matrices  
October 24, 2019  
17 / 36
Definition (Small ball probability)

The *r-ball probability* of a vector $\mathbf{v} := (v_1, \ldots, v_n) \in \mathbb{R}^n$ with respect to a random variable $\xi$ is defined by

$$\rho_{r, \xi}(\mathbf{v}) := \sup_{x \in \mathbb{R}} \Pr(|v_1 \xi_1 + \cdots + v_n \xi_n - x| \leq r),$$

where $\xi_1, \ldots, \xi_n$ are independent copies of $\xi$.

Examples (when $\xi$ is Rademacher):

- If $\mathbf{v} = (10, 100, 1000, \ldots, 10^n)$, then $\rho_{1/4, \xi}(\mathbf{v}) = 2^{-n}$. 
A proof template: the anti-concentration phenomenon

**Definition (Small ball probability)**

The *r-ball probability* of a vector \( \mathbf{v} := (v_1, \ldots, v_n) \in \mathbb{R}^n \) with respect to a random variable \( \xi \) is defined by

\[
\rho_{r,\xi}(\mathbf{v}) := \sup_{x \in \mathbb{R}} \Pr \left( |v_1\xi_1 + \cdots + v_n\xi_n - x| \leq r \right),
\]

where \( \xi_1, \ldots, \xi_n \) are independent copies of \( \xi \).

**Examples (when \( \xi \) is Rademacher):**

- If \( \mathbf{v} = (10, 100, 1000, \ldots, 10^n) \), then \( \rho_{1/4,\xi}(\mathbf{v}) = 2^{-n} \).
- If \( \mathbf{v} = (1, \ldots, 1) \), then \( \rho_{1/4,\xi}(\mathbf{v}) = 2^{-n} \left( \frac{n}{\lfloor n/2 \rfloor} \right) = \Theta \left( \frac{1}{\sqrt{n}} \right) \).
Let $X_1, \ldots, X_n$ denote the rows of $M_n$.

Let $S$ denote the event that $M_n$ is singular.

Let $S_i$ denote the event that $X_i$ lies in the span of the other rows.

Since $1_S \leq 1_{S_1} + \cdots + 1_{S_n}$,

$$\Pr(S) \leq n \Pr(S_n).$$
A proof template: reduction to anti-concentrating normals

Suppose we could prove the following:

Random normals anti-concentrate

Except with probability \(\exp(-n)\), any \(v \in \mathbb{S}^{n-1}\) which is orthogonal to \(X_1, \ldots, X_{n-1}\) satisfies \(\rho_{0, \xi}(v) \leq \rho\).
Suppose we could prove the following:

**Random normals anti-concentrate**

Except with probability $\exp(-n)$, any $\mathbf{v} \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(\mathbf{v}) \leq \rho$.

Then, $\Pr(S_n) \leq \rho + \exp(-n)$, and we would be done.
Except with probability $\exp(-n)$, any $\mathbf{v} \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(\mathbf{v}) \leq \rho$. 
Random normals anti-concentrate

Except with probability \( \exp(-n) \), any \( v \in \mathbb{S}^{n-1} \) which is orthogonal to \( X_1, \ldots, X_{n-1} \) satisfies \( \rho_{0, \xi}(v) \leq \rho \).

**Proof:** Union bound!
Except with probability \( \exp(-n) \), any \( \mathbf{v} \in S^{n-1} \) which is orthogonal to \( X_1, \ldots, X_{n-1} \) satisfies \( \rho_0,\xi(\mathbf{v}) \leq \rho \).

**Proof:** Union bound!

- Let \( \mathbf{v} \in S^{n-1} \) have \( \rho_0,\xi(\mathbf{v}) \in (\lambda/2, \lambda] \).
Proof template: Random normals anti-concentrate

Random normals anti-concentrate

Except with probability \( \exp(-n) \), any \( \mathbf{v} \in S^{n-1} \) which is orthogonal to \( X_1, \ldots, X_{n-1} \) satisfies \( \rho_{0,\xi}(\mathbf{v}) \leq \rho \).

**Proof:** Union bound!

- Let \( \mathbf{v} \in S^{n-1} \) have \( \rho_{0,\xi}(\mathbf{v}) \in (\lambda/2, \lambda] \). By the independence of \( X_1, \ldots, X_{n-1} \),

\[
\Pr (X_1 \cdot \mathbf{v} = 0 \land \cdots \land X_{n-1} \cdot \mathbf{v} = 0) \leq \lambda^{n-1}.
\]
Proof: Union bound!

- Let $v \in S^{n-1}$ have $\rho_{0,\xi}(v) \in (\lambda/2, \lambda]$. By the independence of $X_1, \ldots, X_{n-1}$,
  \[ \Pr(X_1 \cdot v = 0 \wedge \cdots \wedge X_{n-1} \cdot v = 0) \leq \lambda^{n-1}. \]

- Suppose we could show that the ‘number’ of $v \in S^{n-1}$ with $\rho_{0,\xi}(v) \in (\lambda/2, \lambda]$ is at most $(\lambda^{-1}/n^\gamma)^n$. 

Random normals anti-concentrate

Except with probability $\exp(-n)$, any $v \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(v) \leq \rho$.
Proof template: Random normals anti-concentrate

Random normals anti-concentrate

Except with probability \( \exp(-n) \), any \( \mathbf{v} \in \mathbb{S}^{n-1} \) which is orthogonal to \( X_1, \ldots, X_{n-1} \) satisfies \( \rho_{0,\xi}(\mathbf{v}) \leq \rho \).

**Proof:** Union bound!

- Let \( \mathbf{v} \in \mathbb{S}^{n-1} \) have \( \rho_{0,\xi}(\mathbf{v}) \in (\lambda/2, \lambda] \). By the independence of \( X_1, \ldots, X_{n-1} \),

\[
\Pr(X_1 \cdot \mathbf{v} = 0 \land \cdots \land X_{n-1} \cdot \mathbf{v} = 0) \leq \lambda^{n-1}.
\]

- Suppose we could show that the ‘number’ of \( \mathbf{v} \in \mathbb{S}^{n-1} \) with \( \rho_{0,\xi}(\mathbf{v}) \in (\lambda/2, \lambda] \) is at most \( (\lambda^{-1}/n^\gamma)^n \).

- Then, by a union bound, and ranging over \( \lambda = 1, 2^{-1}, 2^{-2}, \ldots, \rho \),
Proof template: Random normals anti-concentrate

Random normals anti-concentrate

Except with probability $\exp(-n)$, any $\mathbf{v} \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(\mathbf{v}) \leq \rho$.

Proof: Union bound!

- Let $\mathbf{v} \in S^{n-1}$ have $\rho_{0,\xi}(\mathbf{v}) \in (\lambda/2, \lambda]$. By the independence of $X_1, \ldots, X_{n-1}$,

  $$\Pr(X_1 \cdot \mathbf{v} = 0 \land \cdots \land X_{n-1} \cdot \mathbf{v} = 0) \leq \lambda^{n-1}.$$ 

- Suppose we could show that the ‘number’ of $\mathbf{v} \in S^{n-1}$ with $\rho_{0,\xi}(\mathbf{v}) \in (\lambda/2, \lambda]$ is at most $(\lambda^{-1}/n^{\gamma})^n$.

- Then, by a union bound, and ranging over $\lambda = 1, 2^{-1}, 2^{-2}, \ldots, \rho$,

  $$\Pr(BAD_{\rho}) \leq \sum_{\lambda} \left(\frac{\lambda^{-1}}{n^{\gamma}}\right)^n \cdot (\lambda)^{n-1} \lesssim \log(1/\rho) \rho n^{-\gamma n}.$$
Can we count?

Not directly, since $\mathbb{S}^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.
Can we count?

Not directly, since $S^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.

- **High-dimensional geometric framework**
  - Discretize the unit sphere using a net based on a very exact relation between Diophantine approximation and anti-concentration.
Can we count?

Not directly, since $S^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.

- **High-dimensional geometric framework**
  - Discretize the unit sphere using a net based on a very exact relation between Diophantine approximation and anti-concentration.
  - Requires strong control on the operator norm, which is often not available when dealing with non-centered entries.
Can we count?

Not directly, since $\mathbb{S}^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.

- **High-dimensional geometric framework**
  - Discretize the unit sphere using a net based on a very exact relation between Diophantine approximation and anti-concentration.
  - Requires strong control on the operator norm, which is often not available when dealing with non-centered entries.

- **Additive combinatorial framework**
  - Uses deep structure vs. randomness ideas from additive combinatorics to build a net on the sphere.
Can we count?

Not directly, since $\mathbb{S}^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.

- **High-dimensional geometric framework**
  - Discretize the unit sphere using a net based on a very exact relation between Diophantine approximation and anti-concentration.
  - Requires strong control on the operator norm, which is often not available when dealing with non-centered entries.

- **Additive combinatorial framework**
  - Uses deep structure vs. randomness ideas from additive combinatorics to build a net on the sphere.
  - **Only effective for** $\rho \geq n^{-C}$. 

Vishesh Jain (MIT)  
Quantitative invertibility of random matrices  
October 24, 2019  
21 / 36
Can we count?

Not directly, since $\mathbb{S}^{n-1}$ has uncountably many points! Overcoming this obstacle is the heart of the matter.

- **High-dimensional geometric framework**
  - Discretize the unit sphere using a net based on a very exact relation between Diophantine approximation and anti-concentration.
  - Requires strong control on the operator norm, which is often not available when dealing with non-centered entries.

- **Additive combinatorial framework**
  - Uses deep structure vs. randomness ideas from additive combinatorics to build a net on the sphere.
  - **Only effective for** $\rho \geq n^{-C}$.
  - Extensions to dependent models require much more work e.g. quadratic inverse Littlewood–Offord theory of [Nguyen (2011)](https://doi.org/10.1112/S0010437X11005324).
We can count something!

Theorem (Ferber, J., Luh, and Samotij, 2018+; J., 2019+)

For all $\rho \geq \exp(-n^{c_1})$, the number of vectors $v \in \mathbb{Z}^n$ with

$$\|v\|_\infty \leq \exp(n^{c_2}); \quad \rho_{1,\xi}(v) \geq \rho$$

is at most

$$(\rho^{-1}/n^{0.5-\epsilon})^n.$$

Proved by a (perhaps surprisingly!) short and elementary double counting argument!
What can we do with the counting theorem?

Recall that we wanted to prove:

**Random normals anti-concentrate**

Except with probability $\exp(-n)$, any $\mathbf{v} \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ i.e. for which $\sum_{i=1}^{n-1} |X_i \cdot \mathbf{v}|^2 = 0$ satisfies $\rho_{0,\xi}(\mathbf{v}) \leq \rho$. 
What can we do with the counting theorem?

Recall that we wanted to prove:

**Random normals anti-concentrate**

Except with probability $\exp(-n)$, any $\mathbf{v} \in \mathbb{S}^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ i.e. for which $\sum_{i=1}^{n-1} |X_i \cdot \mathbf{v}|^2 = 0$ satisfies $\rho_{0,\xi}(\mathbf{v}) \leq \rho$.

Let us show how to use the counting theorem to prove:

**Random integer ‘approximate normals’ anti-concentrate**

Except with probability $\exp(-n)$, any non-zero $\mathbf{z} \in \mathbb{Z}^n$, $\|\mathbf{z}\|_\infty \leq \exp(n^c)$ for which

$$\sqrt{\sum_{i=1}^{n-1} |X_i \cdot \mathbf{z}|^2} \leq n^{1-2\epsilon}$$

satisfies $\rho_{1,\xi}(\mathbf{z}) \leq \rho$. 
What can we do with the counting theorem?

Random integer ‘approximate normals’ anti-concentrate

Except with probability \( \exp(-n) \), any non-zero \( z \in \mathbb{Z}^n, \|z\|_\infty \leq \exp(n^c) \) for which \( \sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} \leq n^{1-2\epsilon} \) satisfies \( \rho_{1,\xi}(z) \leq \rho \).

Proof: By a similar union bound to what we have seen.
What can we do with the counting theorem?

Random integer ‘approximate normals’ anti-concentrate

Except with probability \( \exp(-n) \), any non-zero \( z \in \mathbb{Z}^n \), \( \|z\|_\infty \leq \exp(n^c) \) for which

\[
\sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} \leq n^{1-2\epsilon}
\]

satisfies \( \rho_{1,\xi}(z) \leq \rho \).

\[\text{Proof: By a similar union bound to what we have seen.}\]

- Let \( z \) be an integer vector with \( \|z\|_\infty \leq \exp(n^c) \) and \( \rho_{1,\xi}(z) \in (\lambda/2, \lambda] \).

- By independence, the probability that the vector \( (X_1 \cdot z, \ldots, X_{n-1} \cdot z) \)
  lies in any fixed hypercube with side length 1 is at most \( \lambda^{n-1} \).
What can we do with the counting theorem?

- **Key point:** Since the volume of the \( n^{1-2\epsilon} \)-ball in \( \mathbb{R}^{n-1} \) is at most \( (n^{1-2\epsilon}/\sqrt{n})^{(n-1)} \),
What can we do with the counting theorem?

- **Key point:** Since the volume of the $n^{1-2\epsilon}$-ball in $\mathbb{R}^{n-1}$ is at most $(n^{1-2\epsilon}/\sqrt{n})^{(n-1)}$, the probability that $\sqrt{\sum_{i=1}^{n-1}|X_i \cdot z|^2} \leq n^{1-2\epsilon}$ is at most

  $$\lambda^{n-1} \cdot n^{(0.5-2\epsilon)n}.$$
What can we do with the counting theorem?

- **Key point:** Since the volume of the $n^{1-2\epsilon}$-ball in $\mathbb{R}^{n-1}$ is at most $(n^{1-2\epsilon}/\sqrt{n})^{(n-1)}$, the probability that $\sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} \leq n^{1-2\epsilon}$ is at most

$$\lambda^{n-1} \cdot n^{(0.5-2\epsilon)n}.$$

- On the other hand, by the counting theorem, the number of such $z$ is at most

$$\left(\lambda^{-1}\right)^n \cdot n^{-0.5+\epsilon n}.$$
What can we do with the counting theorem?

- **Key point:** Since the volume of the $n^{1-2\epsilon}$-ball in $\mathbb{R}^{n-1}$ is at most $(n^{1-2\epsilon} / \sqrt{n})^{n-1}$, the probability that $\sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} \leq n^{1-2\epsilon}$ is at most
  \[ \lambda^{n-1} \cdot n^{(0.5-2\epsilon)n}. \]

- On the other hand, by the counting theorem, the number of such $z$ is at most
  \[ (\lambda^{-1})^n \cdot n^{(-0.5+\epsilon)n}. \]

- Therefore, the contribution of such $z$ to the union bound is at most
  \[ (\lambda^{-1})^n n^{(-0.5+\epsilon)n} \cdot \lambda^{n-1} n^{(0.5-2\epsilon)n} = \lambda^{-1} \cdot n^{-\epsilon n}. \]
Getting around by rounding?

Recall that we wanted to prove:

**Random normals anti-concentrate**

Except with probability $\exp(-n)$, any $v \in S^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(v) \leq \rho$. 
Getting around by rounding?

Recall that we wanted to prove:

**Random normals anti-concentrate**

Except with probability $\exp(-n)$, any $v \in \mathbb{S}^{n-1}$ which is orthogonal to $X_1, \ldots, X_{n-1}$ satisfies $\rho_{0,\xi}(v) \leq \rho$.

Can we reduce it to what we can prove?

**Random integer ‘approximate normals’ anti-concentrate**

Except with probability $\exp(-n)$, any non-zero $z \in \mathbb{Z}^n$, $\|z\|_{\infty} \leq \exp(n^c)$ for which

$$\sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} \leq n^{1-2\epsilon}$$

satisfies $\rho_{1,\xi}(z) \leq \rho$. 
Failed attempt: naïve rounding

By rounding \( \mathbf{v} \in \mathbb{S}^{n-1} \) to the nearest integer multiple of \( 1/\sqrt{n} \), we obtain some \( \mathbf{z} \in \mathbb{Z}^n \) such that \( \| \mathbf{v} - (\mathbf{z}/\sqrt{n}) \|_2 \leq 1/2 \) i.e.

\[
\| \sqrt{n} \mathbf{v} - \mathbf{z} \|_2 \leq \sqrt{n}/2.
\]
**Failed attempt: naïve rounding**

- By rounding \( \mathbf{v} \in \mathbb{S}^{n-1} \) to the nearest integer multiple of \( 1/\sqrt{n} \), we obtain some \( \mathbf{z} \in \mathbb{Z}^n \) such that \( \| \mathbf{v} - (\mathbf{z}/\sqrt{n}) \|_2 \leq 1/2 \) i.e.

  \[
  \| \sqrt{n} \mathbf{v} - \mathbf{z} \|_2 \leq \sqrt{n}/2.
  \]

- Let \( \tilde{M}_{n-1} \) denote the matrix consisting of the first \( n - 1 \) rows of \( M_n \).

- So, if \( \tilde{M}_{n-1} \mathbf{v} = 0 \) and \( \| \tilde{M}_{n-1} \| \leq \sqrt{n} \), we get...
Failed attempt: naïve rounding

- By rounding $\mathbf{v} \in S^{n-1}$ to the nearest integer multiple of $1/\sqrt{n}$, we obtain some $\mathbf{z} \in \mathbb{Z}^n$ such that $\|\mathbf{v} - (\mathbf{z}/\sqrt{n})\|_2 \leq 1/2$ i.e.

  $$\|\sqrt{n}\mathbf{v} - \mathbf{z}\|_2 \leq \sqrt{n}/2.$$ 

- Let $\tilde{M}_{n-1}$ denote the matrix consisting of the first $n - 1$ rows of $M_n$.

- So, if $\tilde{M}_{n-1}\mathbf{v} = 0$ and $\|\tilde{M}_{n-1}\| \leq \sqrt{n}$, we get

  $$\sqrt{\sum_{i=1}^{n-1} |X_i \cdot \mathbf{z}|^2} = \|\tilde{M}_{n-1}\mathbf{z}\|_2 = \|\tilde{M}_{n-1}(\mathbf{z} - \sqrt{n}\mathbf{v})\|_2$$

  $$\leq \|\tilde{M}_{n-1}\| \cdot \|\sqrt{n}\mathbf{v} - \mathbf{z}\|_2$$

  $$\leq \sqrt{n} \cdot \frac{1}{2} \sqrt{n} = \frac{n}{2}.$$
Failed attempt: naïve rounding

- By rounding \( \mathbf{v} \in S^{n-1} \) to the nearest integer multiple of \( 1/\sqrt{n} \), we obtain some \( \mathbf{z} \in \mathbb{Z}^n \) such that \( \| \mathbf{v} - (\mathbf{z}/\sqrt{n}) \|_2 \leq 1/2 \) i.e.

\[
\| \sqrt{n} \mathbf{v} - \mathbf{z} \|_2 \leq \sqrt{n}/2.
\]

- Let \( \tilde{M}_{n-1} \) denote the matrix consisting of the first \( n-1 \) rows of \( M_n \).

- So, if \( \tilde{M}_{n-1} \mathbf{v} = 0 \) and \( \| \tilde{M}_{n-1} \| \leq \sqrt{n} \), we get

\[
\sqrt{\sum_{i=1}^{n-1} |X_i \cdot \mathbf{z}|^2} = \| \tilde{M}_{n-1} \mathbf{z} \|_2 = \| \tilde{M}_{n-1} (\mathbf{z} - \sqrt{n} \mathbf{v}) \|_2 \\
\leq \| \tilde{M}_{n-1} \| \cdot \| \sqrt{n} \mathbf{v} - \mathbf{z} \|_2 \\
\leq \sqrt{n} \cdot \frac{1}{2} \sqrt{n} = \frac{n}{2}.
\]

- But we wanted something of the form \( n^{1-2\epsilon} \) on the right hand side...
Successful attempt: non-trivial rounding available!

- We saw that naïve rounding ‘just’ fails. However, we are not trying to round any $\mathbf{v} \in \mathbb{S}^{n-1}$ but only those with $\rho_{0,\xi}(\mathbf{v}) \geq \rho$. 

Proposition (Diophantine approximation vs. small-ball probability)

For $\mathbf{v} \in \mathbb{S}^{n-1}$ with $\rho_{0,\xi}(\mathbf{v}) \geq \rho \geq \exp(-nc)$, there exists some $\gamma \in [1, \exp(nc)]$ and non-zero $\mathbf{z} \in \mathbb{Z}^n$ such that $\|\gamma \mathbf{v} - \mathbf{z}\|_2 \leq n\delta$. 
Successful attempt: non-trivial rounding available!

- We saw that naïve rounding ‘just’ fails. However, we are not trying to round any \( \mathbf{v} \in \mathbb{S}^{n-1} \) but only those with \( \rho_{0,\xi}(\mathbf{v}) \geq \rho \).

- But such vectors are already special and have a better-than-trivial integer approximation!

**Proposition (Diophantine approximation vs. small-ball probability)**

For \( \mathbf{v} \in \mathbb{S}^{n-1} \) with \( \rho_{0,\xi}(\mathbf{v}) \geq \rho \geq \exp(-n^c) \), there exists some \( \gamma \in [1, \exp(n^c)] \) and non-zero \( \mathbf{z} \in \mathbb{Z}^n \) such that

\[
\| \gamma \mathbf{v} - \mathbf{z} \|_2 \leq n^\delta.
\]
Proposition (Diophantine approximation vs. small-ball probability)

For $\mathbf{v} \in S^{n-1}$ with $\rho_{0,\xi}(\mathbf{v}) \geq \rho \geq \exp(-n^c)$, there exists some $\gamma \in [1, \exp(n^c)]$ and non-zero $\mathbf{z} \in \mathbb{Z}^n$ such that

$$\|\gamma \mathbf{v} - \mathbf{z}\|_2 \leq n^\delta.$$ 

Key point: Don’t require any specific dependence of $\gamma$ on $\rho$. 
Proposition (Diophantine approximation vs. small-ball probability)

For \( \mathbf{v} \in S^{n-1} \) with \( \rho_{0,\xi}(\mathbf{v}) \geq \rho \geq \exp(-n^c) \), there exists some \( \gamma \in [1, \exp(n^c)] \) and non-zero \( \mathbf{z} \in \mathbb{Z}^n \) such that

\[
\| \gamma \mathbf{v} - \mathbf{z} \|_2 \leq n^\delta.
\]

- **Key point:** Don’t require any specific dependence of \( \gamma \) on \( \rho \).
- Direct consequence of the following Fourier analytic bound:

\[
\rho_{1,\xi}(\mathbf{v}) \lesssim \int_{\mathbb{R}} \exp \left( -\text{dist} (\lambda \mathbf{v}, \mathbb{Z}^n)^2 - \lambda^2 \right) d\lambda,
\]

which essentially appears in a classical work of Halász (1977).
Proposition (Diophantine approximation vs. small-ball probability)

For \( v \in S^{n-1} \) with \( \rho_0,\xi(v) \geq \rho \geq \exp(-nc) \), there exists some \( \gamma \in [1, \exp(nc)] \) and non-zero \( z \in \mathbb{Z}^n \) such that \( \|\gamma v - z\|_2 \leq n^\delta \).

- Suppose \( \tilde{M}_{n-1} v = 0 \) for \( v \in S^{n-1} \) satisfying \( \rho_0,\xi(v) \geq \rho \).
Putting everything together

**Proposition (Diophantine approximation vs. small-ball probability)**

For \( \mathbf{v} \in S^{n-1} \) with \( \rho_{0, \xi} (\mathbf{v}) \geq \rho \geq \exp(-n^c) \), there exists some \( \gamma \in [1, \exp(n^c)] \) and non-zero \( \mathbf{z} \in \mathbb{Z}^n \) such that \( \| \gamma \mathbf{v} - \mathbf{z} \|_2 \leq n^\delta \).

- Suppose \( \tilde{M}_{n-1} \mathbf{v} = 0 \) for \( \mathbf{v} \in S^{n-1} \) satisfying \( \rho_{0, \xi} (\mathbf{v}) \geq \rho \).
- Then, by the theorem, we can find a non-zero \( \mathbf{z} \in \mathbb{Z}^n \) with \( \| \gamma \mathbf{v} - \mathbf{z} \|_2 \leq n^\delta \) and \( \| \mathbf{z} \|_{\infty} \leq \exp(n^c) \).
Proposition (Diophantine approximation vs. small-ball probability)

For $v \in \mathbb{S}^{n-1}$ with $\rho_{0,\xi}(v) \geq \rho \geq \exp(-nc)$, there exists some $\gamma \in [1, \exp(nc)]$ and non-zero $z \in \mathbb{Z}^n$ such that $\|\gamma v - z\|_2 \leq n^\delta$.

- Suppose $\tilde{M}_{n-1} v = 0$ for $v \in \mathbb{S}^{n-1}$ satisfying $\rho_{0,\xi}(v) \geq \rho$.
- Then, by the theorem, we can find a non-zero $z \in \mathbb{Z}^n$ with $\|\gamma v - z\|_2 \leq n^\delta$ and $\|z\|_\infty \leq \exp(nc)$.
- Therefore, if $\|\tilde{M}_{n-1}\| \leq \sqrt{n}$, we get

$$\sqrt{\sum_{i=1}^{n-1} |X_i \cdot z|^2} = \|\tilde{M}_{n-1} z\|_2 \leq \|\tilde{M}_{n-1}\| \cdot \|\gamma v - z\|_2$$

$$\leq \sqrt{n} \cdot n^\delta \ll n^{1-2\epsilon},$$

so that $z$ is an integer ‘approximate normal’.

Vishesh Jain (MIT)
Quantitative invertibility of random matrices
October 24, 2019 30 / 36
Putting everything together

**Proposition (Diophantine approximation vs. small-ball probability)**

For \( \mathbf{v} \in S^{n-1} \) with \( \rho_{0, \xi}(\mathbf{v}) \geq \rho \geq \exp(-n^c) \), there exists some \( \gamma \in [1, \exp(n^c)] \) and non-zero \( \mathbf{z} \in \mathbb{Z}^n \) such that \( \|\gamma \mathbf{v} - \mathbf{z}\|_2 \leq n^\delta \).

- Suppose \( \tilde{M}_{n-1} \mathbf{v} = 0 \) for \( \mathbf{v} \in S^{n-1} \) satisfying \( \rho_{0, \xi}(\mathbf{v}) \geq \rho \).
- Then, by the theorem, we can find a non-zero \( \mathbf{z} \in \mathbb{Z}^n \) with \( \|\gamma \mathbf{v} - \mathbf{z}\|_2 \leq n^\delta \) and \( \|\mathbf{z}\|_\infty \leq \exp(n^c) \).
- Therefore, if \( \|\tilde{M}_{n-1}\| \leq \sqrt{n} \), we get

\[
\sqrt{\sum_{i=1}^{n-1} |X_i \cdot \mathbf{z}|^2} = \|\tilde{M}_{n-1} \mathbf{z}\|_2 \leq \|\tilde{M}_{n-1}\| \cdot \|\gamma \mathbf{v} - \mathbf{z}\|_2 \\
\leq \sqrt{n} \cdot n^\delta \ll n^{1-2\epsilon},
\]

so that \( \mathbf{z} \) is an integer ‘approximate normal’.
- This is exactly what we have ruled out with high probability!
Rounding in the presence of heavy tails

How can we round if the entries are only assumed to have finite second moment?
Rounding in the presence of heavy tails

How can we round if the entries are only assumed to have finite second moment?

Control on $\infty$-to-2 norm

Let $A_{n \times m}$ be an $n \times m$ ‘tall’ random matrix, each of whose entries is an independent copy of a random variable with mean 0 and variance 1. Then, except with probability $\exp(-n^c)$,

$$\|A_{n' \times m}\|_{\infty \to 2} \lesssim n^c \cdot \sqrt{n} \cdot \sqrt{m}.$$

Proved using Chernoff bound + standard concentration inequalities on the symmetric group.
Rounding in the presence of heavy tails

Let \( \text{err} := \gamma v - z \). Recall that \( \| \text{err} \|_2 \leq n^\delta \) and we want to show that \( \| \tilde{M}_{n-1} \text{err} \|_2 \leq n^{1-2\epsilon} \).
Rounding in the presence of heavy tails

- Let $\texttt{err} := \gamma v - z$. Recall that $\|\texttt{err}\|_2 \leq n^\delta$ and we want to show that $\|\tilde{M}_{n-1} \texttt{err}\|_2 \leq n^{1-2\epsilon}$.

- Decompose $\texttt{err} = \texttt{err}_{sp} + \texttt{err}_{sm}$, where $\texttt{err}_{sp}$ consists of the largest $n^{0.8}$ coordinates in absolute value.
Let $\text{err} := \gamma v - z$. Recall that $\|\text{err}\|_2 \leq n^\delta$ and we want to show that $\|\tilde{M}_{n-1}\text{err}\|_2 \leq n^{1-2\epsilon}$.

Decompose $\text{err} = \text{err}_{sp} + \text{err}_{sm}$, where $\text{err}_{sp}$ consists of the largest $n^{0.8}$ coordinates in absolute value.

Since $n^{0.8}\|\text{err}_{sm}\|_\infty^2 \leq \|\text{err}_{sp}\|_2^2 \leq \|\text{err}\|_2^2$,

$$\|\text{err}_{sm}\|_\infty \leq n^\delta / n^{0.4}.$$
Rounding in the presence of heavy tails

- Let $\text{err} := \gamma v - z$. Recall that $\|\text{err}\|_2 \leq n^\delta$ and we want to show that $\|\tilde{M}_{n-1}\text{err}\|_2 \leq n^{1-2\epsilon}$.

- Decompose $\text{err} = \text{err}_{sp} + \text{err}_{sm}$, where $\text{err}_{sp}$ consists of the largest $n^{0.8}$ coordinates in absolute value.

- Since $n^{0.8}\|\text{err}_{sm}\|_\infty^2 \leq \|\text{err}_{sp}\|_2^2 \leq \|\text{err}\|_2^2$, we have $\|\text{err}_{sm}\|_\infty \leq n^\delta/n^{0.4}$.

- Therefore,

$$\|\tilde{M}_{n-1}\text{err}_{sm}\|_2 \leq \|\tilde{M}_{n-1}\|_{\infty \rightarrow 2}\|\text{err}_{sm}\|_\infty \leq \frac{n^{1+\epsilon+\delta}}{n^{0.4}} \ll n^{1-2\epsilon}.$$
Rounding in the presence of heavy tails

- It remains to show that $\|\tilde{M}_{n-1} \text{err}_{sp}\|_2 \leq n^{1-2\epsilon}$. 

Note that $\tilde{M}_{n-1} \text{err}_{sp} = (\tilde{M}_{n-1} \text{Proj} J \text{err}_{sp})$, where $|J| = n^0.8$. 

Except with probability at most $\exp(-n^{c})$, $\|\tilde{M}_{n-1} \text{Proj} J \text{err}_{sp}\|_{\infty}^{\rightarrow 2} \leq n^0.9 + \epsilon \approx n^{1-2\epsilon}$. Hence, $\|\tilde{M}_{n-1} \text{err}_{sp}\|_2 \leq \|\tilde{M}_{n-1} \text{Proj} J \text{err}_{sp}\|_{\infty} \leq n^0.9 + \epsilon < n^{1-2\epsilon}$. 

Vishesh Jain (MIT) 
Quantitative invertibility of random matrices 
October 24, 2019 33 / 36
It remains to show that $\|\tilde{M}_{n-1} \text{err}_{sp}\|_2 \leq n^{1-2\epsilon}$.

Note that $\tilde{M}_{n-1} \text{err}_{sp} = (\tilde{M}_{n-1} \text{Proj}_J) \text{err}_{sp}$, where $|J| = n^{0.8}$. 
Rounding in the presence of heavy tails

- It remains to show that \( \| \tilde{M}_{n-1} \text{err}_{sp} \|_2 \leq n^{1-2\epsilon} \).

- Note that \( \tilde{M}_{n-1} \text{err}_{sp} = (\tilde{M}_{n-1} \text{Proj}_J) \text{err}_{sp} \), where \( |J| = n^{0.8} \).

- Except with probability at most \( \exp(-n^c) \),

\[
\| \tilde{M}_{n-1} \text{Proj}_J \|_{\infty \rightarrow 2} \leq n^\epsilon \cdot \sqrt{n} \cdot \sqrt{|J|} = n^{0.9+\epsilon}.
\]
It remains to show that \( \| \tilde{M}_{n-1} \text{err}_{sp} \|_2 \leq n^{1-2\epsilon} \).

Note that \( \tilde{M}_{n-1} \text{err}_{sp} = (\tilde{M}_{n-1} \text{Proj}_J) \text{err}_{sp} \), where \( |J| = n^{0.8} \).

Except with probability at most \( \exp(-n^c) \),
\[
\| \tilde{M}_{n-1} \text{Proj}_J \|_{\infty \rightarrow 2} \leq n^{\epsilon} \cdot \sqrt{n} \cdot \sqrt{|J|} = n^{0.9+\epsilon}.
\]

Hence,
\[
\| \tilde{M}_{n-1} \text{err}_{sp} \|_2 \leq \| \tilde{M}_{n-1} \text{Proj}_J \|_{\infty \rightarrow 2} \| \text{err}_{sp} \|_{\infty} \leq n^{0.9+\epsilon} \ll n^{1-2\epsilon}.
\]


For an introduction to the geometric framework of Rudelson and Vershynin, see their survey *Non-asymptotic theory of random matrices: extreme singular values* from the 2010 ICM Proceedings.

For an introduction to inverse Littlewood-Offord theory and its applications to random matrix theory, see *Small Ball Probability, Inverse Theorems, and Applications*, Nguyen and Vu. Erdős Centennial pp 409-463.
THANK YOU!

References available at math.mit.edu/∼visheshj

For any questions or comments: visheshj@mit.edu