A stronger bound for the growth factor of a discrete random matrix

Anay Aggarwal

August 2026

Abstract

Let ξ be a real random variable with finite support and let $M_n(\xi)$ be the $n \times n$ random matrix with entries drawn independently from ξ. Mateo, Urschel, and West proved that, with probability bounded away from zero, $M_n(\xi)$ has an LU factorization without pivoting whose growth factor is at most $n^{5/2+\delta}$, but experimentally observed that the correct exponent should be $3/2$. We improve the bound to $Cn^{3/2} \log n$, matching the experimentally observed exponent (with a logarithmic factor tacked on).

1 Introduction

Gaussian elimination is the process by which a matrix A is decomposed into the product of a lower triangular matrix L and upper triangular matrix U, an LU-decomposition. This process is widely used to solve linear systems of equations.

This decomposition is not always possible, but when it is, it is unique. It turns out that one can explicitly describe when the decomposition is impossible. This can be made rigorous by defining the notion of a *strongly-non-singular* matrix (the set of which will be denoted SNS). We say that $A \in \text{SNS}$ if all the leading principal minors of A (excluding $\det(A)$ itself) are nonzero. Then an LU-decomposition is possible if and only if the matrix is SNS.

For the Gaussian elimination algorithm run on an SNS matrix, in floating point arithmetic the accumulated round-off error is governed by the *growth factor*,

$$\text{growth}(A) = \max\left\{ \|L\|_{\max}, \frac{\|U\|_{\max}}{\|A\|_{\max}} \right\},$$

where $\|\bullet\|_{\max}$ denotes the maximum entry of a matrix. The role of the growth factor in backwards error analysis of Gaussian elimination goes back to [7].

In this paper, we work in the setting of Mateo, Urschel, and West [2] and consider the growth factor of discrete random matrices. In particular, we let ξ be a random variable with finite support and consider $M_n(\xi)$, the $n \times n$ random matrix where each entry is independent and identically distributed; drawn from ξ. We let A be a matrix sampled from $M_n(\xi)$, and consider the quantity $\text{growth}(A)$. Two results of [2] are relevant: the probability that A is SNS is bounded below by a constant p_ξ [2, Corollary 1.2], and the following theorem concerning the size of the growth factor.

Theorem 1.1 (Paraphrased [2, Thm 1.1])**.** *For all $\delta > 0$ there is a constant $C_\xi < 1$ such that*

$$\mathbb{P}(\text{growth}(M_n(\xi)) > n^{5/2+\delta}) \leq C_\xi + o_n(1).$$

The 5/2 exponent matches the best known bound even in the smooth Gaussian setting [5], but experiments in [2] consistently point to a conjectured exponent of 3/2.

The main result of this paper is to confirm that the theorem does indeed hold true for an exponent of 3/2 (bringing on a logarithmic factor).

Theorem 1.2. *For all $\varepsilon > 0$, there exists a constant $C_{\xi,\varepsilon} > 0$ such that for sufficiently large n and $A \sim M_n(\xi)$,*

$$\mathbb{P}(A \in \text{SNS},\ \text{growth}(A) \leq C_{\xi,\varepsilon} n^{3/2} \log n) \geq \mathbb{P}(A \in \text{SNS}) - \varepsilon.$$

Using the fact that $\mathbb{P}(A \in \text{SNS}) \geq p_\xi$ from [2], we directly obtain the following corollary, stated in the same language as Theorem 1.1.

Corollary 1.3. *For all $\varepsilon > 0$, there are constants $c_\xi > 0$ and $C_\xi < 1$ such that*

$$\mathbb{P}(\text{growth}(M_n(\xi)) > c_\xi n^{3/2} \log n) \leq C_\xi + o_n(1).$$

This explicitly improves the $n^{5/2+\delta}$ bound of [2] to $n^{3/2} \log n$, done through a more careful analysis but a similar overall idea. We remark that this bound is likely tight, but the task of finding a corresponding lower bound remains open.

Outline of the paper. We establish our notation in Section 2. We then give a helpful outline of the proof in Section 3, which consists of two steps. The first step, restricting to a helpful subset of matrices, is done in Section 4, and the inequalities from this section are put together in Section 5 to complete the proof.

2 Notation

Matrix notation. For a matrix A, we let A_m denote the leading principal $m \times m$ submatrix. We let $[n] = \{1, 2, \ldots, n\}$, and for sets $I, J \subseteq [n]$, we let $A_{I,J}$ denote the submatrix of A given by the A-indexed rows and the B indexed columns. The smallest singular value of A is denoted $\sigma_{\min}(A)$. We let $\mathbf{1}_n \in \mathbb{R}^n$ be the all-ones vector (often dropping the subscript), $J_n \in M_n(\mathbb{R})$ be the all-ones matrix, and e_j be the vector that is one in position j and zero elsewhere. The functions $\|\bullet\|_{\text{op}}, \|\bullet\|_F, \|\bullet\|_{\max}$ represent operator norm, Frobenius norm, and maximum entry. The function $\|\bullet\|$ represents the standard Euclidean norm of a vector. Throughout the paper, we assume the SNS condition for A, and so A_m^{-1} is well-defined for all m. We abuse $c_\bullet$ and $C_\bullet$ quantities denoting constants.

Probability notation. Throughout, ξ is a random variable with finite support. We denote $K_\xi = \max\{|a| : a \in \text{supp}(\xi)\}$ and $\mu = \mathbb{E}\xi$. We denote $M_n(\xi)$ by the $n \times n$ random matrix where each entry is independent and identically distributed; drawn from ξ. For a set S that can be interpreted as an event, we interchange $\mathbb{P}(S)$ and $\mathbb{P}(A \in S)$. For example, $\mathbb{P}(\text{SNS})$ denotes $\mathbb{P}(A \in \text{SNS})$.

3 Proof Outline

The fundamental idea of [2] is that we can recursively compute determinants of principal minors. In block form,

$$A_{m+1} = \begin{pmatrix} A_m & y \\ x^T & \alpha \end{pmatrix},$$

for $x, y \in \mathbb{R}^m$. Hence $\det(A_{m+1}) = \det(A_m)(\alpha - x^T A_m^{-1} y)$ by standard properties of block matrices. Therefore, the singularity of A_{m+1} (and blowup of its norm) is determined by the quantity $\alpha - x^T A_m^{-1} y$. This is the type of quantity we will study. More precisely, by using the Doolittle formula in Schur complement form, there is an explicit LU decomposition with

$$U_{ij} = A_{ij} - A_{i,[m]} A_m^{-1} A_{[m],j}, \; L_{ij} = A_{i,[j]} A_j^{-1} e_j,$$

where $m = i - 1$ (and $U_{1j} = A_{1j}$).

We first restrict our analysis to "nice" matrices. In particular, matrices with A_m^{-1} bounded in various ways. We show that a random matrix is "nice" with high probability, and restrict our study of nice matrices. In particular, we define four sets of matrices:

$$\begin{aligned}
H_1 &: \; \left\|A_m^{-1}\right\|_{\text{op}} \leq Cn\sqrt{m} \text{ for all } m, \\
H_2 &: \; \left\|A_m^{-1}\right\|_F \leq Cn\sqrt{m \log n} \text{ for all } m, \\
H_3 &: \; \left\|A_m^{-1}\mathbf{1}\right\| + \left\|(A_m^{-1})^T\mathbf{1}\right\| \leq Cn\sqrt{m \log n} \text{ for all } m, \\
H_4 &: \; \left|\mathbf{1}^T A_m^{-1}\mathbf{1}\right| \leq Cn\sqrt{m}\log n \text{ for all } m,
\end{aligned}$$

where C is some implicit constant (really, each H_i is a function of C). Then our set of nice matrices will be $H = H_1 \cap H_2 \cap H_3 \cap H_4$. It turns out that this is essentially all we need; the bound in Theorem 1.2 falls out after decoupling x and y in $x^T A_m^{-1} y$. The main savings as compared to [2] comes from decoupling into operator and Frobenius norms, and being precise about singular values rather than using Cauchy-Schwarz. We will use the following three bounds in our proof.

Theorem 3.1 (Decoupled Hanson-Wright, see [4] and [6, Thm 6.2.1])**.** *Let $\zeta, \zeta' \in \mathbb{R}^m$ be independent random vectors, each with independent mean-zero entries bounded by K, and let W be a fixed matrix over $M_m(\mathbb{R})$. Then for any $u \geq 0$,*

$$\mathbb{P}\left(\left|\zeta^T W \zeta'\right| \geq C_1(K)(\|W\|_F \sqrt{u} + \|W\|_{\text{op}} u)\right) \leq 2e^{-u}.$$

Theorem 3.2 (Jain-Sah-Sawhney in square form, see [2, Cor. 2.6] of [1, Thm. 1.4])**.** *There are constants $C_0 > 0$, $\gamma \in (0,1)$, and $k_1 \in \mathbb{N}$, depending on ξ, such that for all $m \geq k_1$ and $t \geq 0$,*

$$\mathbb{P}\left(\sigma_{\min}(A_m) \leq t/\sqrt{m}\right) \leq C_0 t + \gamma^m.$$

Theorem 3.3 (Rudelson-Vershynin in rectangular form, see [3])**.** *Let Z be a $k \times m$ matrix, $m \leq k$, with independent entries drawn from a fixed mean-zero, nondegenerate, bounded distribution. There are constants $\varepsilon_0, c_2 > 0$, depending on the distribution, such that*

$$\mathbb{P}(\sigma_{\min}(Z) \leq \varepsilon_0(\sqrt{k} - \sqrt{m-1})) \leq 2^{-(k-m+1)} + e^{-c_2 k}.$$

Theorem 3.1 will be used directly to bound $x^T A_m^{-1} y$ (after decoupling), Theorem 3.2 will be used to treat H_1 directly, and Theorem 3.3 will be used to bound the singular values of tall matrices (which will show up in treatment of H_2). We will also frequently use Hoeffding's inequality to approximate random variables by their mean, and the union bound to unify bounds over rows and columns.

4 Most Matrices are in H

We treat each H_i separately, with the goal of eventually combining them to earn the following.

Proposition 4.1. *For every $\eta > 0$, there is a constant $C = C_{\xi,\eta}$ and an event $H = H_{n,\eta}$ such that*

$$\mathbb{P}(\mathrm{SNS} \cap H) \geq \mathbb{P}(SNS) - \eta - r_n,$$

where $r_n \to 0$ and, on $SNS \cap H$, the hypotheses H_1 through H_4 are satisfied (with C as their constant).

The nature of H_1 through H_4 will become clear in Section 5. Let us begin by treating H_1 with Theorem 3.2.

Lemma 4.2. *We have for any $\eta > 0$ that*

$$\mathbb{P}(\mathrm{SNS} \cap H_1) \geq \mathbb{P}(\mathrm{SNS}) - \eta - r_{1n},$$

where $r_{1n} \to 0$.

Proof. Let C_0, γ, k_1 be as in Theorem 3.2. Pick $k_2 \geq k_1$ with $\gamma^{k_2}/(1-\gamma) \leq \eta/2$. Let $s_m = \sigma_{\min}(A_m)$. Assume that A is SNS. Applying Theorem 3.2 with $t = \eta/(2C_0 n)$, we have

$$\sum_{m=k_2+1}^{n} \mathbb{P}(s_m \leq t/\sqrt{m}) \leq nC_0 t + \sum_{m>k_2} \gamma^m \leq \eta.$$

In the range $m \leq k_2$, because k_2 is a constant we have that when n is large enough, $s_m \geq c_\eta/(n\sqrt{m})$ automatically. Since $\left\|A_m^{-1}\right\|_{\mathrm{op}} \leq Cn\sqrt{m}$ is equivalent to $s_m \geq c/(n\sqrt{m})$, we obtain the desired result by union bound. □

Now we will treat H_2 assuming that H_1 is true. This is the most nontrivial condition to deal with.

Lemma 4.3. *We have for any $\eta > 0$ that*

$$\mathbb{P}(\mathrm{SNS} \cap H_1 \cap H_2) \geq \mathbb{P}(\mathrm{SNS} \cap H_1) - r_{2n},$$

where $r_{2n} \to 0$.

Proof. Assume that $A \in \mathrm{SNS} \cap H_1$. To bound the Frobenius norm, we bound each singular value separately. Consider $\sigma_{m-j}(A_m)$, the $(m-j)$th singular value of A_m. By Courant-Fisher,

$$\sigma_{m-j}(A_m) \geq \min_{v \in V, \|v\|=1} \|A_m v\|,$$

for any subspace $V \in \mathbb{R}^m$ of dimension $m-j$. The trick is, we will use an explicit case of V. In particular, we let V be the space of vectors in $\mathbb{R}^m$ with last $j-1$ entries zero, and with sum of entries equal to zero. This clearly has dimension $m-j$. The reason this particular subspace is useful is because it reduces A_m to a zero-mean matrix. Decompose $A_m = \mu J_m + Z_m$, where Z_m has independent mean-zero entries. Then for $v \in V$ with norm 1, because $J_m v = 0$,

$$\|A_m v\| = \|Z_m v\| = \left\|Z_{[m],[m-j+1]} v_{[m-j+1]}\right\| \geq \sigma_{\min}(Z_{[m],[m-j+1]}).$$

This allows us to invoke Theorem 3.3, to bound the singular values of the tall matrix $Z_{[m],[m-j+1]}$:

$$\mathbb{P}(\sigma_{\min}(Z_{[m],[m-j+1]}) \le \varepsilon_0 j/(2\sqrt{m})) \le \mathbb{P}(\sigma_{\min}(Z_{[m],[m-j+1]}) \le \varepsilon_0(\sqrt{m}-\sqrt{m-j})) \le 2^{-j} + e^{-c_2 m},$$

because $\sqrt{m}-\sqrt{m-j} \ge j/(2\sqrt{m})$. We will split j at $T = C_3 \log n$. Summing this probability for $j \ge T$ and $m \le n$, we have a $1 - O(n^{-2})$ probability that none of these singular values $\sigma_{m-j}(A_m)$ are at most $\varepsilon_0 j/(2\sqrt{m})$. In this case,

$$\left\|A_m^{-1}\right\|_F^2 = \sum_{j=0}^{m-1} \sigma_{m-j}(A_m)^{-2} \le T s_m^{-2} + \sum_{j>T} \frac{4m}{\varepsilon_0^2 j^2} \le C(s_m^{-2}\log n + m).$$

Because $A \in H_1$, we have that $s_m \ge c/(n\sqrt{m})$, which yields $\left\|A_m^{-1}\right\|_F \le Cn\sqrt{m\log n}$. This implies that

$$\mathbb{P}(\mathrm{SNS} \cap H_1 \cap H_2) \ge \mathbb{P}(\mathrm{SNS} \cap H_1)(1 - O(n^{-2})),$$

which yields the desired. □

We may now kill two birds with one stone by handling H_3 and H_4 together. The main idea is that $\mathbf{1}$ is almost an eigenvector of A_m. Note that we assume that $\mu \neq 0$; in the case that $\mu = 0$ the terms that H_3 and H_4 control vanish in Lemma 5.1 and hence can be ignored.

Lemma 4.4. *We have for any $\eta > 0$ that*

$$\mathbb{P}(\mathrm{SNS} \cap H_3 \cap H_4 \cap H_1) \ge \mathbb{P}(\mathrm{SNS} \cap H_1) - r_{3n},$$

where $r_{3n} \to 0$.

Proof. Assume $A \in \mathrm{SNS} \cap H_1$. Decompose $A_m\mathbf{1} = \mu m\mathbf{1} + w_m$, for $w_m \in \mathbb{R}^m$. By Hoeffding, a given entry of w_m is bounded in magnitude by $C\sqrt{m\log n}$ with probability $1 - O(n^{-D})$. Put $D = 4$, say. Then union bounding over every m and every element of w_m, we have with probability $1 - O(n^{-2})$ that $\|w\|_\infty \le C\sqrt{m\log n}$, and hence $\|w\| \le Cm\sqrt{\log n}$ for all m. We can solve for $A_m^{-1}\mathbf{1}$ as

$$A_m^{-1}\mathbf{1} = \frac{1}{\mu m}(\mathbf{1} - A_m^{-1}w).$$

Hence with probability $1 - O(n^{-2})$ we have

$$\left\|A_m^{-1}\mathbf{1}\right\| \le \frac{1}{|\mu|m}(\sqrt{m} + s_m^{-1}Cm\sqrt{\log n}) \le Cn\sqrt{m\log n}.$$

The same inequality holds for the transpose. For H_4, we note that

$$|\mathbf{1}^T A_m^{-1}\mathbf{1}| = \frac{1}{|\mu|m}\left|m - \mathbf{1}^T A_m^{-1}w\right| \le \frac{1}{|\mu|} + \frac{1}{|\mu|m}\left\|(A_m^{-1})^T\mathbf{1}\right\|\|w\| \le C(\sqrt{\log n} + s_m^{-1}\log n).$$

Assuming H_1, we have $s_m^{-1} \le cn\sqrt{m}$, yielding that $A \in \mathrm{SNS} \cap H_3 \cap H_4 \cap H_1$ with probability $1 - O(n^{-2})$. □

We may now compile these three lemmas to prove Proposition 4.1.

Proof. Let $X = \mathrm{SNS} \cap H_1$. Observe that, by union bound and Lemmas 4.3, 4.4,

$$\begin{aligned}\mathbb{P}(X \backslash (H_2 \cap H_3 \cap H_4)) &\le \mathbb{P}(X \backslash H_2) + \mathbb{P}(X \backslash (H_3 \cap H_4)) \\ &\le o_n(1) + o_n(1) \\ &= o_n(1).\end{aligned}$$

Therefore, $\mathbb{P}(X \cap H_2 \cap H_3 \cap H_4) \ge \mathbb{P}(X) - o_n(1) \ge \mathbb{P}(\mathrm{SNS}) - \eta - o_n(1)$, as wanted. □

5 Bounding the Growth

We begin with an illuminating lemma that bounds $|x^T W y|$ in a way that makes clear the reason H was defined as so.

Lemma 5.1. *Let $W \in M_n(\mathbb{R})$ be a deterministic matrix, and let $x, y \in \mathbb{R}^n$ be independent random vectors with entries drawn independently from ξ. Then for any $D > 0$, there is a constant $C > 0$ such that with probability $1 - O(n^{-D})$ we have the inequality*

$$\left|x^T W y\right| \leq \mu^2 \left|\mathbf{1}^T W \mathbf{1}\right| + C|\mu|\sqrt{\log n}(\|W\mathbf{1}\| + \left\|W^T \mathbf{1}\right\|) + C(\|W\|_F \sqrt{\log n} + \|W\|_{\text{op}} \log n).$$

Proof. In order to bound $\left|x^T W y\right|$ we decouple $x = \mu\mathbf{1} + \tilde{x}$, $y = \mu\mathbf{1} + \tilde{y}$. This gives us

$$\left|x^T W y\right| = \mu^2 \mathbf{1}^T W \mathbf{1} + \mu \tilde{x}^T (W\mathbf{1}) + \mu(\mathbf{1}TW)\tilde{y} + \tilde{x}^T W \tilde{y}.$$

We can use Hoeffding with parameter $\sqrt{2D \log n}$ to bound the middle two terms, of the form $\tilde{x}^T v$ and $v'\tilde{y}$. For the final term we use Theorem 3.1:

$$\mathbb{P}(\left|\tilde{x}^T W \tilde{y}\right| \geq C(\|W\|_F \sqrt{u} + \|W\|_{\text{op}} u)) \leq 2e^{-u}.$$

With $u = D \log n$ this gives us, with $1 - O(n^{-D})$ probability (for any $D > 0$), the inequality

$$\left|x^T W y\right| \leq \mu^2 \left|\mathbf{1}^T W \mathbf{1}\right| + C|\mu|\sqrt{\log n}(\|W\mathbf{1}\| + \left\|W^T \mathbf{1}\right\|) + C(\|W\|_F \sqrt{\log n} + \|W\|_{\text{op}} \log n),$$

where C is some constant depending on D. □

Now all the hypotheses become clear: if $A \in H$, we have that for each $m = i - 1$,

$$\left|A_{i,[m]} A_m^{-1} A_{[m],j}\right| \leq C_\xi n\sqrt{m} \log n \leq C_\xi n^{3/2} \log n,$$

with probability $1 - O(n^{-D})$. We may now prove Theorem 1.2.

Proof. We set $\eta = \varepsilon/2$ and separately bound $\|U\|_{\max}$, $\|A\|_{\max}$, and $\|L\|_{\max}$.

Bounding U. Because each $|A_{ij}| \leq K_\xi$, we have the event E_m of the form

$$E_m : \ |U_{ij}| \leq C_{\xi,\varepsilon} n^{3/2} \log n \text{ for all } j \in [n], \ i = m + 1$$

has probability $p_m = 1 - O(n^{-D_m})$ of occurring, for a $D_m > 0$ that we may choose (which alters $C_{\xi,\varepsilon}$). Thus by the union bound, $\mathbb{P}\left(\bigcap_m E_m\right)$ is of the form $1 - O(n^{-D})$. In other words, we have $\|U\|_{\max} \leq C_{\xi,\varepsilon} n^{3/2} \log n$ with $1 - o_n(1)$ probability for a suitable $C_{\xi,\varepsilon}$.

Bounding A. Note that $\mathbb{P}(\|A\|_{\max} < K_\xi) = (1-p)^{n^2} = o_n(1)$ for a $p > 0$ (namely, $p = \mathbb{P}(\xi = K_\xi)$). Therefore, $\|A\|_{\max} = K_\xi$ with $1 - o_n(1)$ probability.

Bounding L. We use $L_{ij} = A_{i,[j]} A_j^{-1} e_j$ for $i > j$, and condition on A_j, so that A_j^{-1} is deterministic. Denoting $x = A_{i,[j]}^T$, write $x = \mu\mathbf{1} + \tilde{x}$. Then by Hoeffding with parameter proportional to $\sqrt{\log n}$, we have that, with probability $1 - O(n^{-D})$,

$$|L_{ij}| = \left|x^T (A_j^{-1} e_j)\right| \leq |\mu||\mathbf{1}^T A_j^{-1} e_j| + C\sqrt{\log n} \left\|A_j^{-1} e_j\right\|,$$

for a $C > 0$ depending on D. This can be bounded via the two inequalities,

$$\left|\mathbf{1}^T A_j^{-1} e_j\right| = \left|\langle (A_j^{-1})^T \mathbf{1}, e_j\rangle\right| \le \left\|(A_j^{-1})^T \mathbf{1}\right\|,$$

$$\left\|A_j^{-1} e_j\right\| \le \left\|A_j^{-1}\right\|_{\mathrm{op}}.$$

Therefore, for an $A \in H$, we have that

$$|L_{ij}| \le C_{\xi,\varepsilon}(n\sqrt{j\log n} + n\sqrt{j}\sqrt{\log n}) \le C_{\xi,\varepsilon} n^{3/2}\log n,$$

with $1 - O(n^{-D})$ probability for each j. This gives

$$\|L\|_{\max} \le C_{\xi,\varepsilon} n^{3/2}\log n,$$

with probability $1 - o_n(1)$ by the same union bound argument used to bound $\|U\|_{\max}$.

Combining these three bounds, we obtain that, given that $A \in \mathrm{SNS} \cap H$,

$$\mathbb{P}(\mathrm{growth}(A) \le C'_{\xi,\varepsilon} n^{3/2}\log n) = 1 - o_n(1),$$

for a constant $C'_{\xi,\varepsilon}$. By Proposition 4.1, we thus have

$$\mathbb{P}(A \in \mathrm{SNS},\ \mathrm{growth}(A) \le C'_{\xi,\varepsilon} n^{3/2}\log n) \ge \mathbb{P}(A \in \mathrm{SNS}) - \varepsilon/2 - o_n(1),$$

which yields the desired result for large enough n. □

Note on Generative AI Use: This problem was solved in conjunction with Claude Fable 5 (a single web agent). The agent spun its wheels when initially given the problem, and required some additional pushes from the author, e.g. "try to use more past principal minors via Doolittle's formula" to eventually solve the problem. The substance of the argument was then written up in this paper. We would like to stress that every word in this paper was written by the author himself, and the entirety of the proof was human-verified. On the other hand, the result may not have been achieved without the helpful use of AI, which for example found and expertly utilized Theorems 3.1, 3.2, and 3.3. The open lower bound has resisted attacks from various LLMs. This is a beta test for the Large Math Initiative, an initiative to streamline the use of AI in math research.

References

[1] V. Jain, A. Sah, and M. Sawhney, Singularity of discrete random matrices, *Geom. Funct. Anal.* **31** (2021), 1160–1218.

[2] S. Orellana Mateo, J. Urschel, and N. West, LU factorization of discrete random matrices, arXiv:2608.08998 (2026).

[3] M. Rudelson and R. Vershynin, Smallest singular value of a random rectangular matrix, *Comm. Pure Appl. Math.* **62** (2009), 1707–1739.

[4] M. Rudelson and R. Vershynin, Hanson–Wright inequality and sub-gaussian concentration, *Electron. Commun. Probab.* **18** (2013), 1–9.

[5] A. Sankar, D. A. Spielman, and S.-H. Teng, Smoothed analysis of the condition numbers and growth factors of matrices, *SIAM J. Matrix Anal. Appl.* **28** (2006), 446–476.

[6] R. Vershynin, *High-Dimensional Probability*, Cambridge University Press, 2018.

[7] J. H. Wilkinson, Error analysis of direct methods of matrix inversion, *J. ACM* **8** (1961), 281–330.