ENTRY GROWTH IN GAUSSIAN ELIMINATION

RIKHAV SHAH AND JOHN URSCHEL

Abstract. Gaussian elimination is one of the oldest algorithms in mathematics, and the most popular method for solving an unstructured linear system. Its stability in finite precision is controlled by its growth factor, which measures how large the entries produced during elimination can become. Understanding the worst-case behavior of this quantity has been a central problem in numerical analysis since the 1940s.

Here we make a significant leap in that understanding, settling several open problems. In particular, we determine the asymptotic behavior of the maximum growth factor under complete and rook pivoting, proving that both are quasi-polynomial in dimension. We also show that the exponential growth under partial pivoting persists for sparse matrices and that randomized partial pivoting suffers the same instability. In contrast, we show that every non-singular matrix has a row permutation with polynomial growth, though finding the optimal row permutation is NP-complete.

1. Introduction

Solving a linear system $A\boldsymbol{x} = \boldsymbol{b}$ is one of the oldest problems in mathematics, and the technique now known as Gaussian elimination is one of the oldest and most popular approaches. In Gaussian elimination, an $n \times n$ matrix A is factored into the product of a lower and upper triangular matrix L and U by a sequence of rank-one transformations, resulting in the sequence of matrices $A^{(0)} = A$ and $A^{(k)} = \big(a_{ij}^{(k)}\big)_{i,j=1}^{n-k}$ for k equals 1 to $n-1$, satisfying

$$A^{(k)} = Z_k - Y_k W_k^{-1} X_k, \qquad \text{where } A = \begin{array}{cc} \begin{array}{cc} \;\; k & \;\; n-k \end{array} & \\ \begin{bmatrix} W_k & X_k \\ Y_k & Z_k \end{bmatrix} & \begin{array}{c} k \\ n-k \end{array} \end{array}. \tag{1.1}$$

The sequence $A^{(0)}, \dots, A^{(n)}$ is shift-invariant, i.e., $A^{(k+\ell)}$ is the ℓ^{th} step of Gaussian elimination applied to $A^{(k)}$, and defining $A^{(k+1)}$ as a function of $A^{(k)}$ recovers the Gaussian elimination algorithm itself. The resulting LU factorization of A is encoded by the first row and column of each of the iterates. In particular,

$$L_{ij} = \frac{a_{i-j+1,1}^{(j-1)}}{a_{11}^{(j-1)}} \quad \text{for} \quad i \geq j \qquad \text{and} \qquad U_{ij} = a_{1,j-i+1}^{(i-1)} \quad \text{for} \quad i \leq j.$$

The LU factorization of a matrix, conditional on normalizing L to be unit lower triangular, is unique if it exists. Unfortunately, a matrix does not have an LU factorization if one of its leading submatrices W_k is singular, in which case a permutation of the rows (and/or columns) may be necessary. What's more, performing these computations in finite precision can produce cascading round-off errors. This error is controlled by a quantity known as the growth factor

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

where $\|\cdot\|_{\max}$ is the entry-wise infinity norm. In order for the Gaussian elimination procedure to compute the LU factorization of a matrix in a numerically stable way, the growth factor of that matrix must not be too large. Roughly speaking, the number of bits of precision required for stable computation scales as $\log(\mathrm{growth}(A))$, see [GVL13, Theorem 3.3.1] and [Hig02, Theorem 9.3 & Lemma 9.6] for details.

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA.
E-mail addresses: rdshah@mit.edu, urschel@mit.edu.

2020 *Mathematics Subject Classification.* 65F05, 15A23.
Key words and phrases. Gaussian elimination, growth factor.

For this reason, in practice the matrix A is replaced by the matrix PAQ, where P and Q are permutation matrices that can be chosen quickly, with the idea that growth(PAQ) may be significantly smaller than growth(A). The application of P and Q can then be undone by downstream tasks after Gaussian elimination is completed. The rule governing which permutation matrices P and Q are chosen for a given matrix A is referred to as a pivoting strategy, and the top left entry of each step of Gaussian elimination is referred to as the pivot. In this work, we prove several results related to the growth factor of Gaussian elimination under various pivoting strategies.

1.1. **Pivoting Strategies.** Before presenting our contributions, we first formally define the pivoting strategies under consideration. Complete pivoting, which takes the largest magnitude entry of the matrix as the pivot, is the most traditional pivoting strategy. Its analysis dates back to the seminal 1947 *Numerical Inverting of Matrices of High Order* by Goldstine and von Neumann [vNG47], often referred to as the "first modern paper on numerical analysis" [Gol93], and the stability of Gaussian elimination under complete pivoting has been a major open problem ever since. In 1961, Wilkinson produced a quasi-polynomial upper bound for the growth factor under complete pivoting of $2\sqrt{n}n^{\frac{1}{4}\log(n)}$ and remarked that "no matrix has been encountered for which [the growth factor for complete pivoting] was as large as 8 [Wil61]." Bisain, Edelman, and Urschel recently improved this upper bound to approximately $n^{0.207\log n}$ [BEU25]. In the same decade that Wilkinson commented on the small growth factors encountered in practice, a folklore conjecture that the growth factor under complete pivoting was at most n arose. It "became one of the most famous open problems in numerical analysis" [Hig02, pg. 181]. This conjecture later was shown to be false in dimension 13 in floating point arithmetic by Gould [Gou91] (extended to exact arithmetic by Edelman [Ede92]), and Edelman and Urschel recently showed it to be false for all $n \geq 11$ by a multiplicative constant [EU24]. Afterward, Fedorovskii produced examples of matrices with growth on the order of $n^{\log_{7/2}(4)}$ under complete pivoting [Fed25]. However, the gap between lower and upper bounds for the growth factor under complete pivoting in the literature remained large. We refer the reader to [Hig02, EU24, Urs25] for more details.

In contrast, partial pivoting, which takes the largest magnitude entry in the first column as the pivot, is completely understood in the worst case. In his 1965 text *The Algebraic Eigenvalue Problem* [Wil65], Wilkinson proved the upper bound of 2^{n-1} and produced an example matrix, commonly referred to as the Wilkinson matrix, that achieves it. Later, Higham and Higham characterized the full class of matrices with growth 2^{n-1} [HH89]. An exponentially large growth factor is a prohibitive barrier to accurate computation in floating point arithmetic. For example, Gaussian elimination with partial pivoting is completely inaccurate when solving a 100×100 linear system involving the Wilkinson matrix in double precision [Urs25]. Despite this, due to the low cost and memory access requirements of partial pivoting, Gaussian elimination with partial pivoting is the premier computational technique for solving a linear system and is the algorithm utilized by the built-in linear system solver in nearly every programming language. Much of the analysis and debate regarding partial pivoting regards whether large growth can occur in practice, and to what extent large growth is unstable (i.e., a small random perturbation leads to small growth). We refer the reader to [TS90, TB22, Urs25] for more details.

Due to the horrible worst case behavior of partial pivoting, a number of alternate pivoting strategies have been suggested. One of the most notable is rook pivoting, which takes the largest magnitude entry in its row and column as the pivot. Rook pivoting serves as a compromise between partial and complete pivoting: the strategy appears to be relatively fast in practice [Fos97, PN00], like partial pivoting, but as Foster proved in [Fos97], it is known to have at most a quasi-polynomial growth factor $\frac{3}{2}\,n^{3\log(n)/4}$, like complete pivoting. However, the largest known growth factor for rook pivoting is only of order $n^{1.669}$, proven by Edelman and Urschel [EU24]. As with complete pivoting, the gap between lower and upper bounds in the literature remained quite large. Another alternative which has received recent attention is the use of a randomized version of partial pivoting, which chooses the pivot from the first column probabilistically, with weights depending on the relative magnitudes of the entries.

For the purposes of worst-case analysis, when we consider pivoting strategies such as partial, rook, or complete pivoting, we need not concern ourselves with the choice of permutation matrices during the algorithm, only with the final row and column permuted matrix. For this reason, it suffices to restrict ourselves to the set of matrices which require no pivoting under a given pivoting strategy, and analyze the growth factor without pivoting over those sets. To that end, we define the following three classes:

(1) $\mathrm{CP}_n(\mathbb{R})$: $A \in \mathrm{GL}_n(\mathbb{R})$ is completely-pivoted if for all k,

$$\left|a_{11}^{(k)}\right| \geq \max_{i,j\in[n-k]} \max\left|a_{ij}^{(k)}\right|,$$

(2) $\mathrm{RP}_n(\mathbb{R})$: $A \in \mathrm{GL}_n(\mathbb{R})$ is rook-pivoted if for all k,

$$\left|a_{11}^{(k)}\right| \geq \max_{j\in[n-k]} \max\left\{\left|a_{1j}^{(k)}\right|, \left|a_{j1}^{(k)}\right|\right\},$$

(3) $\mathrm{PP}_n(\mathbb{R})$: $A \in \mathrm{GL}_n(\mathbb{R})$ is partially-pivoted if for all k,

$$\left|a_{11}^{(k)}\right| \geq \max_{j\in[n-k]} \max\left|a_{j1}^{(k)}\right|,$$

where $[n] := \{1, \ldots, n\}$. In addition, in this work we analyze two randomized pivoting strategies, which correspond to distributions on permutation matrices. Each permutation matrix P can be identified with the permutation sequence $(\pi_1, \ldots, \pi_n)$, where $\boldsymbol{e}_j^\top P \boldsymbol{e}_{\pi_j} = 1$. Let $A[S,T]$ denote the submatrix of a matrix A indexed by row indices S and column indices T. The two randomized pivoting strategies we consider are volume pivoting, characterized by certain marginals

$$P \sim \textsc{Vol}(A): \qquad \Pr(\{\pi_1, \ldots, \pi_k\} = S) \propto (\det A[S, [k]])^2, \tag{1.2}$$

and randomized partial pivoting, which is fully defined

$$P \sim \textsc{Rand}\mathrm{PP}_p(A): \qquad \Pr(\pi_{k+1} = i \,|\, \{\pi_1, \ldots, \pi_k\} = S) \propto \left|\frac{\det A[S + \{i\}, [k+1]]}{\det A[S, [k]]}\right|^p. \tag{1.3}$$

It is not obvious from the definition of $\textsc{Vol}$ that a distribution with those prescribed marginals exist; indeed part of our contribution is showing that this is the case (Lemma 3.2).

1.2. **Our Contributions.** Here we make major progress on a number of questions regarding the growth factor under various deterministic and randomized pivoting strategies.

First, we prove that the growth factor under partial pivoting can remain exponentially large even for sparse matrices. The worst-case bound of 2^{n-1} is achievable by matrices with as few as $4n-4$ non-zero entries, and this is tight. Matrices that are locally sparse, i.e., with at most k non-zeros per row and column, can also have exponentially large growth under partial pivoting. For example, there exist matrices with at most three non-zeros per row and column with growth at least φ^{n-1}, where φ is the golden ratio. In contrast, any partially pivoted matrix with at most two non-zeros per row and column has growth factor at most two.

Theorem 1.1 (Theorem 2.2 & Theorem 2.3 / Sparse exponential growth of partial pivoting)**.** *For every $n \geq 2$, there is a matrix $A \in \mathrm{PP}_n(\mathbb{R})$ with $\mathrm{nnz}(A) = 4n-4$ and $\mathrm{growth}(A) = 2^{n-1}$, and this is the minimum value of $\mathrm{nnz}(A)$ possible for any matrix in $\mathrm{PP}_n(\mathbb{R})$ with growth 2^{n-1}. In addition, for every $k > 1$ and $n > k$, there is a matrix $A \in \mathrm{PP}_n(\mathbb{R})$ with at most $k+1$ non-zero entries per row and column and $\mathrm{growth}(A) = r_k^{n-1} \geq (2-2^{1-k})^{n-1}$, where r_k is the unique solution to $x^k = x^{k-1} + \ldots + x + 1$ in $(2-2^{1-k}, 2)$. Any matrix $A \in \mathrm{PP}_n(\mathbb{R})$ with at most two non-zero entries per row and column has $\mathrm{growth}(A) \leq 2$.*

Despite this, we prove the existence of a randomized row pivoting strategy with polynomially bounded growth factor in expectation. In particular, we produce polynomial estimates for the Frobenius norms of L and U, which translate immediately to polynomial estimates for growth. We compare these estimates to Frobenius norm lower bounds produced by Sylvester Hadamard matrices for any row permutation and lower bounds produced by arbitrary Hadamard matrices for any row and column permutation.

Theorem 1.2 (Theorem 3.1 / Existence of a low growth row permutation)**.** *For every n and matrix $A \in \mathrm{GL}_n(\mathbb{R})$, $\textsc{Vol}(A)$ is a well-defined probability distribution on permutations. If $P \sim \textsc{Vol}(A)$ then PA has an LU factorization with probability one, $\mathbb{E}\,\|L\|_F^2 \leq (\frac{1}{6} + o(1))n^3$, $\mathbb{E}\,\|U\|_F^2 \leq (\frac{1}{6} + o(1))n^3\|A\|_{1\to 2}^2$, and $\mathbb{E}\,\mathrm{growth}(PA) \leq (\frac{1}{\sqrt{6}} + o(1))n^2$.*

Theorem 1.3 (Theorem 3.4 & Theorem 3.5 / Lower bounds over all pivoting strategies)**.** *Let $n = 2^k$, $H \in \mathrm{GL}_n(\mathbb{R})$ be the Sylvester Hadamard matrix, and P be any permutation matrix such that PH has an LU factorization. Then $\|L\|_F^2 = n^{\log_2 3}$ and $\|U\|_F^2 = n^{\log_2 6}$. Furthermore, for any Hadamard matrix A of any dimension $n > 1$ and any permutation matrices P and Q such that PAQ has an LU factorization, we have $\|L\|_F^2 \geq \frac{1}{e} n \log n$ and $\|U\|_F^2 \geq \frac{e}{2} n^2$.*

While the existence of a row permutation with polynomial growth is an encouraging result, we also note that, in general, finding the row permutation that minimizes the growth factor is NP-complete. Furthermore, we show that randomized partial pivoting, the best candidate for a fast row pivoting strategy with polynomial growth factor, can have a growth factor that is nearly exponential.

Theorem 1.4 (Theorem 4.2 / Hardness of minimizing growth)**.** *Let $\mathcal{L}$ be the language consisting of tuples $(A, g) \in \mathrm{GL}_n(\mathbb{Z}) \times \frac{1}{2}\mathbb{Z}_+$ for which there exists a permutation P such that* $\mathrm{growth}(PA) \leq g$*. $\mathcal{L}$ is NP-complete.*

Theorem 1.5 (Theorem 5.2 / Large growth of randomized partial pivoting)**.** *For every n and $p \in (0, \infty)$, there is a matrix $A \in \mathrm{GL}_n(\mathbb{R})$ such that if $P \sim \mathrm{RandPP}_p(A)$, then* $\mathrm{growth}(PA) \geq e^{\Omega_p(n/\log n)}$ *with high probability.*

Finally, we turn our attention to pivoting strategies that allow both row and column permutations. It was long thought that rook and complete pivoting were both likely to have polynomial growth factor. Known upper bounds for both are currently quasi-polynomial, while the known lower bounds for both are sub-quadratic. Here we prove the existence of matrices for which rook and complete pivoting have quasi-polynomial growth.

Theorem 1.6 (Theorem 6.2 / Quasi-polynomial growth of rook pivoting)**.** *For every n, there is a matrix $A \in \mathrm{RP}_n(\mathbb{R})$ with* $\mathrm{growth}(A) \geq n^{(\frac{1}{4} - o(1)) \log_2(n)}$*.*

Theorem 1.7 (Theorem 7.11 / Quasi-polynomial growth of complete pivoting)**.** *For every n, there is a matrix $A \in \mathrm{CP}_n(\mathbb{R})$ with* $\mathrm{growth}(A) \geq n^{\Omega(\log_2(n))}$*.*

This characterization of the asymptotic behavior of the growth factor under complete pivoting answers one of the oldest and most fundamental open questions in numerical analysis.

2. Sparse Matrix Growth Under Partial Pivoting

The growth factor under partial pivoting is known to be at most 2^{n-1}, as the entries of successive iterates can only double in size at each of the $n-1$ steps of Gaussian elimination. In his 1965 text *The Algebraic Eigenvalue Problem* [Wil65], Wilkinson showed that this upper bound is achieved by

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & 1 & 0 & 1 \\ -1 & \cdots & -1 & 1 & 1 \\ -1 & \cdots & -1 & -1 & 1 \end{pmatrix},$$

now commonly referred to as the Wilkinson matrix. Later, Higham and Higham characterized the subset of $\mathrm{PP}_n(\mathbb{R})$ that achieves the 2^{n-1} growth factor in the following theorem

Theorem 2.1 ([HH89, Theorem 2.2])**.** *Every $A \in \mathrm{PP}_n(\mathbb{R})$ with* $\mathrm{growth}(A) = 2^{n-1}$ *is of the form*

$$A = DM \begin{bmatrix} \widetilde{T} & \theta \boldsymbol{d} \end{bmatrix}, \qquad \text{where} \quad \widetilde{T} = \begin{bmatrix} T \\ 0 \end{bmatrix},$$

$D \in \mathrm{GL}_n(\mathbb{R})$ is a ± 1 diagonal matrix, $M \in \mathrm{GL}_n(\mathbb{R})$ is unit lower triangular with $M_{ij} = -1$ for $i > j$, $T \in \mathrm{GL}_{n-1}(\mathbb{R})$ is an upper triangular matrix, $\boldsymbol{d} = (1, 2, 4, \ldots, 2^{n-1})^\top \in \mathbb{R}^n$, and θ is a scalar such that $\theta = |A_{1n}| = \|A\|_{\max}$.

The exponential growth factor exhibited by both the Wilkinson matrix and the class of matrices described in the above theorem leads to catastrophic errors in floating point arithmetic [EUZ24, Urs25]. Within the numerical analysis community, there has been an ongoing discussion as to whether such matrices ever arise in practice. Examples of exponential growth coming from application were produced by Wright for two-point boundary value problems [Wri93] and by Foster for quadrature for Volterra integral equations [Fos94]. One notable aspect of Wright's example is that, unlike the Wilkinson matrix, the matrices in question were sparse, with at most three non-zero entries per row and growth factor approaching $2^{n/2}$.

Here, we make progress in two directions. First, we exactly characterize how sparse a matrix can be and still attain growth 2^{n-1}. Second, we prove lower bounds for how large the growth factor of a matrix with at most k non-zeros per row and column can be.

Theorem 2.2. *Let $n \geq 2$ and $A \in \mathrm{PP}_n(\mathbb{R})$ with $\mathrm{growth}(A) = 2^{n-1}$. Then $\mathrm{nnz}(A) \geq 4n-4$, with equality achieved by the matrix*

$$A = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & & & 1 \\ \frac{1}{2} & -1 & \ddots & & \vdots \\ \vdots & & \ddots & \frac{1}{2} & 1 \\ \vdots & & & -1 & 1 \\ \frac{1}{2} & & & & 1 \end{pmatrix}.$$

Proof. First, we note that the theorem statement is immediate when $n = 2$. Now, let $n \geq 3$ and consider the desired lower bound $\mathrm{nnz}(A) \geq 4n-4$ using the formulation of $A \in \mathrm{PP}_n(\mathbb{R})$ with $\mathrm{growth}(A) = 2^{n-1}$ provided by Theorem 2.1. In Theorem 2.1, the matrix D simply re-signs rows, and leaves the sparsity structure unchanged. Therefore, we may instead analyze $B := M \begin{bmatrix} \widetilde{T} & \theta \boldsymbol{d} \end{bmatrix}$.

The first column of B equals $T_{11} M \boldsymbol{e}_1$ for non-zero T_{11}, and the last column equals $\theta \mathbf{1}$. This already accounts for $2n$ non-zero entries in B. It suffices to show that the remaining $n-2$ columns all have at least two non-zero entries in them.

Consider column j for $1 < j < n$. Because $\widetilde{T}$ is upper triangular,

$$\widetilde{T}_{j+1,j} = \boldsymbol{e}_{j+1}^\top M^{-1} B \boldsymbol{e}_j = 0,$$

where $\boldsymbol{e}_{j+1}^\top M^{-1} = (2^{j-1}, 2^{j-2}, \ldots, 2, 1, 1, 0, \ldots, 0)$ is non-zero in its first $j+1$ entries. Therefore, either $\mathrm{nnz}(B\boldsymbol{e}_j) \geq 2$ or the first $j+1$ entries of $B\boldsymbol{e}_j$ are zero. But, by the non-singularity of the leading $j \times j$ submatrix of M, the first j entries of $\widetilde{T}\boldsymbol{e}_j$ are also zero. In particular, $T_{jj} = 0$, a contradiction.

Finally, we verify that the example matrix A with $\mathrm{nnz}(A) = 4n-4$ in the theorem statement is partially pivoted and has growth factor 2^{n-1}. Indeed, the iterates of this matrix are given by

$$A^{(k)} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & & & 2^k \\ \frac{1}{2} & -1 & \ddots & & \vdots \\ \vdots & & \ddots & \frac{1}{2} & 2^k \\ \vdots & & & -1 & 2^k \\ \frac{1}{2} & & & & 2^k \end{pmatrix} \quad \text{for} \quad k \in [n-3], \qquad A^{(n-2)} = \begin{pmatrix} -\frac{1}{2} & 2^{n-2} \\ \frac{1}{2} & 2^{n-2} \end{pmatrix}, \qquad A^{(n-1)} = \begin{pmatrix} 2^{n-1} \end{pmatrix}.$$

□

Theorem 2.3. *Let $k > 1$, $n > k$, r_k be the unique solution to $x^k = x^{k-1} + \ldots + x + 1$ in the interval $(2 - 2^{1-k}, 2)$, $s_j = \sum_{i=1}^{k+1-j} r_k^{-i}$ for $j = 1, \ldots, k$, and $A \in \mathrm{Mat}_n(\mathbb{R})$ be given by*

$$A_{ij} = \begin{cases} 1 & i = j, j < n \\ -1 & 1 \leq i - j \leq k \\ s_i & j = n, i \leq k \\ 0 & \textit{otherwise} \end{cases}.$$

Then $A \in \mathrm{PP}_n(\mathbb{R})$, has at most $k+1$ non-zero entries in each row and column, and

$$\mathrm{growth}(A) = r_k^{n-1} \geq \left(2 - 2^{1-k}\right)^{n-1}.$$

In addition, any $A \in \mathrm{PP}_n(\mathbb{R})$ with at most two non-zero entries in each row and column has $\mathrm{growth}(A) \leq 2$.

Example 2.4. When $n = 6$ and $k = 2$ and $k = 3$, respectively, the matrices associated with Theorem 2.3 are given by

$$\begin{pmatrix} 1 & & & & & 1 \\ -1 & 1 & & & & 1/\varphi \\ -1 & -1 & 1 & & & \\ & -1 & -1 & 1 & & \\ & & -1 & -1 & 1 & \\ & & & -1 & -1 & \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & & & & & 1 \\ -1 & 1 & & & & 1/\psi + 1/\psi^2 \\ -1 & -1 & 1 & & & 1/\psi \\ -1 & -1 & -1 & 1 & & \\ & -1 & -1 & -1 & 1 & \\ & & -1 & -1 & -1 & \end{pmatrix},$$

where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio and $\psi = \frac{1}{3}\big(1 + \sqrt[3]{19 + 3\sqrt{33}} + \sqrt[3]{19 - 3\sqrt{33}}\big)$.

Proof of Theorem 2.3. First, we verify that the given matrix A, parameterized by k, has the desired properties. By inspection, A has at most $k+1$ non-zero entries in each row and column, and, because A is lower triangular save for the last column, with entries below the diagonal bounded by the diagonal entries, A is partially pivoted. To observe the growth factor, it suffices to give the LU factorization explicitly. Let L be unit lower triangular with -1 on the first k subdiagonals and zeros elsewhere and U be upper triangular with $U_{ii} = 1$ for $i < n$, $U_{in} = r_k^{i-1}$ for all i and zeros elsewhere. Then, because U is the identity matrix in its first $n-1$ columns, and the first $n-1$ columns of L agree with the first $n-1$ columns of A, it suffices to analyze the last column of A. We have

$$(LU)_{in} = r_k^{i-1} - \sum_{j=1}^{\min\{k,i-1\}} r_k^{i-1-j}.$$

For $i \le k$, using the identity $1 = r_k^{-1} + \ldots + r_k^{-k}$,

$$(LU)_{in} = r_k^{i-1} - \sum_{j=1}^{i-1} r_k^{i-1-j} = r_k^{i-1}\left(1 - \sum_{j=1}^{i-1} r_k^{-j}\right) = r_k^{i-1}\sum_{j=i}^{k} r_k^{-j} = \sum_{\ell=1}^{k+1-i} r_k^{-\ell} = s_i,$$

and, for $i > k$,

$$(LU)_{in} = r_k^{i-1} - \sum_{j=1}^{k} r_k^{i-1-j} = r_k^{i-1-k}\left(r_k^k - r_k^{k-1} - r_k^{k-2} - \cdots - 1\right) = 0.$$

In addition, by definition, $1 = s_1 > s_2 > \ldots > s_k = r_k^{-1} > 0$, so $\text{growth}(A) = r_k^{n-1}$. That the root r_k is unique and indeed in the interval $(2 - 2^{1-k}, 2)$ follows by Rouche's theorem on $(x^k - x^{k-1} - \cdots - 1)\cdot\frac{x-1}{x^k}$ (see, e.g., [Sha22, Lemma 3.4]).

Now consider a matrix $A \in \mathrm{PP}_n(\mathbb{R})$ with at most two non-zero entries per row and column. We induct on the following claim: $A^{(k)}$ has at most two non-zero entries per row and column, every row and column of $A^{(k)}$ with two non-zero entries has infinity norm at most $\|A\|_{\max}$, and every row and column of $A^{(k)}$ with one non-zero entry has infinity norm at most $2\|A\|_{\max}$. The matrix A certainly satisfies this claim. Now suppose $A^{(k)}$ also satisfies this claim for some $k < n-1$. If the pivot is the only non-zero entry in either the first row or column, $A^{(k+1)}$ is simply $A^{(k)}$ with the first row and column removed, and therefore also satisfies the claim. If $a_{i1}^{(k)} \neq 0$ and $a_{1j}^{(k)} \neq 0$ for some i and j, then, in this case, $A^{(k+1)}$ is simply $A^{(k)}$ with the first row and column removed and only the entry $a_{ij}^{(k)}$ updated. The update $-a_{i1}^{(k)}a_{1j}^{(k)}/a_{11}^{(k)}$ is at most $\|A\|_{\max}$ in magnitude, as, by the claim, $\big|a_{1j}^{(k)}\big| \le \|A\|_{\max}$, and, by $A \in \mathrm{PP}_n(\mathbb{R})$, $\big|a_{i1}^{(k)}/a_{11}^{(k)}\big| \le 1$. If $a_{ij}^{(k)} = 0$, then $\big|a_{i-1,j-1}^{(k+1)}\big| \le \|A\|_{\max}$, and still there are at most two non-zero entries in its row and column, as the column containing $a_{i1}^{(k)}$ and the row containing $a_{1j}^{(k)}$ have been removed. If $a_{ij}^{(k)} \neq 0$, then $\big|a_{i-1,j-1}^{(k+1)}\big| \le \big|a_{ij}^{(k)}\big| + \big|a_{i1}^{(k)}a_{1j}^{(k)}/a_{11}^{(k)}\big| \le 2\|A\|_{\max}$. However, because $a_{ij}^{(k)} \neq 0$, there are no other non-zero entries in the row or column of $a_{i-1,j-1}^{(k+1)}$. Therefore, $A^{(k+1)}$ also satisfies the desired claim, thus proving the claim for all k.

From here, it follows immediately that $\text{growth}(A) \le 2$, as $A \in \mathrm{PP}_n(\mathbb{R})$ implies that $\|L\|_{\max} = 1$ and our proven claim implies that $\|U\|_{\max} \le 2\|A\|_{\max}$.

∎

3. Existence of Low Growth Row Exchanges

Partial pivoting exchanges rows so that the resulting lower triangular matrix L has max norm one. Unfortunately, this can result in the entries of U becoming as large as 2^{n-1}. This tradeoff is lessened by using complete pivoting, for which $\|L\|_{\max} = 1$ and the entries of U grow at worst quasi-polynomially. However, this requires row and column exchanges, and still doesn't produce a polynomial growth factor. Here, we consider a fundamental existence question: does there always a exist a row permutation that produces polynomial growth? By analyzing the randomized pivoting algorithm $P \sim \textsc{Vol}(A)$, we answer this question affirmatively.

This question shares many characteristics with the concept of rank revealing factorizations and maximum volume pivoting in numerical linear algebra (see [GE96, Pan00, SG20, DGTY24]). A $k \times k$ matrix locally maximizes volume if its squared determinant cannot be increased by swapping out a row or column. This is naturally related to Gaussian elimination, as the entries of L and U are ratios of determinants:

$$L_{ij} = \frac{\det A[[j-1]+\{i\},[j]]}{\det A[[j],[j]]} \quad \text{for} \quad i \geq j \qquad \text{and} \qquad U_{ij} = \frac{\det A[[i],[i-1]\cup\{j\}]}{\det A[[i-1],[i-1]]} \quad \text{for} \quad i \leq j. \tag{3.1}$$

Unfortunately, there is a natural obstruction to applying these results directly. For any fixed k, one can find a large volume set of k rows in the first k columns of A, and these choices do not need to be nested as k varies. However, a row pivoting strategy for Gaussian elimination requires a single chain of choices $S_1 \subset S_2 \subset \ldots \subset S_n$, where $|S_k| = k$. Using $\textsc{Vol}$ pivoting, we build a random chain with a determinantal distribution based on $\det(A[S_k,[k]])^2$, commonly referred to as volume sampling [DRVW06, DR10] or a fixed size determinantal point process (DPP) [KT12].

The only structural result about DPPs that we require is the stochastic monotonicity of projection processes. If $\mathcal{H}_1 \subseteq \mathcal{H}_2$ are subspaces of $\ell^2(E)$ for some countable set E and $\mathbb{P}^{\mathcal{H}_i}$ denotes the projection DPP associated with the orthogonal projection onto $\mathcal{H}_i$, then $\mathbb{P}^{\mathcal{H}_1}$ is stochastically dominated by $\mathbb{P}^{\mathcal{H}_2}$, see [Lyo03, Theorem 6.2]. By Strassen's coupling characterization of stochastic domination [Str65], the two processes therefore admit a coupling (S_1, S_2) such that $S_1 \subseteq S_2$ almost surely. Applying this result successively to the nested column spaces

$$\mathrm{Col}(A[:,[1]]) \subseteq \mathrm{Col}(A[:,[2]]) \subseteq \cdots \subseteq \mathrm{Col}(A)$$

produces the required random nested chain, while preserving the volume-sampling marginal at every level.

Theorem 3.1. *For every $A \in \mathrm{GL}_n(\mathbb{R})$, $\textsc{Vol}(A)$ is a well-defined measure on permutations and if $P \sim \textsc{Vol}(A)$ then $PA = LU$ satisfies*

$$\mathbb{E}\,\|L\|_F^2 \leq \frac{n^3+5n}{6} \qquad \text{and} \qquad \mathbb{E}\,\|U\|_F^2 \leq \frac{n^3+3n^2+2n}{6}\,\|A\|_{1\to 2}^2\,.$$

Since $\|\cdot\|_{\max} \leq \|\cdot\|_F$ and $\|A\|_{1\to 2} \leq \sqrt{n}\,\|A\|_{\max}$, this additionally implies that $\mathbb{E}\,\mathrm{growth}(PA)^2 \leq \frac{n^4}{6} + O(n^3)$.

Proof. The existence of the distribution $\textsc{Vol}$ with the desired marginals (1.2) is given by Lemma 3.2. The expected squared Euclidean norms for each column of L and row of U are given by Lemma 3.3, so we simply sum these expressions,

$$\begin{aligned}
\mathbb{E}\,\|L\|_F^2 &= \sum_{k=1}^n \mathbb{E}\,\|L[:,k]\|^2 \leq n + \sum_{k=1}^n k(n-k) = \frac{n^3+5n}{6}, \\
\mathbb{E}\,\|U\|_F^2 &= \mathbb{E}\,\|U[1,:]\| + \sum_{k=1}^n \mathbb{E}\,\|U[k+1,:]\|^2 \leq \|A\|_{1\to 2} + \|A\|_{1\to 2}\sum_{k=1}^n (k+1)(n-k) = \frac{n^3+3n^2+2n}{6}
\end{aligned} \tag{3.2}$$

. □

Lemma 3.2. *Given $A \in \mathrm{GL}_n(\mathbb{R})$, let Π_k be the orthogonal projection onto the span of the leading k columns of A. Then there is a distribution, which we denote $\textsc{Vol}$, on permutations $(\pi_1, \ldots, \pi_n)$ of $[n]$ such that*

$$\Pr(\{\pi_1, \ldots, \pi_k\} = S) = \det \Pi_k[S,S]. \tag{3.3}$$

for all $k \in [n]$. Furthermore, these distributions are equivalently formulated as

$$\Pr(\{\pi_1, \ldots, \pi_k\} = S) \propto (\det A[S,[k]])^2.$$

Proof. The equivalence of the forms comes from the orthogonal projection formula,

$$\Pi_k = A[:, [k]](A[:, [k]]^\top A[:, [k]])^{-1} A[:, [k]]^\top.$$

Looking at the $[S, S]$ submatrix of both sides and taking determinants gives

$$\det(\Pi_k[S, S]) = \frac{(\det A[S, [k]])^2}{\det(A[:, [k]]^\top A[:, [k]])}.$$

which indeed defines a probability distribution by Cauchy-Binet. For each k, let S_k be the random element of $\binom{[n]}{k}$ such that

$$\Pr(S_k = S) = \det \Pi_k[S, S].$$

By [Lyo03, Theorem 6.2], S_{k+1} stochastically dominates S_k, and, by Strassen's coupling theorem [Str65], this is equivalent to the existence of a coupling such that $S_k \subset S_{k+1}$. By applying this theorem pairwise, we have an entire nested sequence

$$\emptyset = S_0 \subset S_1 \subset S_2 \subset \cdots \subset S_n = [n], \qquad |S_k| = k.$$

Set π_i to be the unique item in $S_i \backslash S_{i-1}$ to obtain the desired distribution $\textsc{Vol}(A)$ on permutations. $\square$

Lemma 3.3. *Let P be a permutation sampled according to $\textsc{Vol}(A)$ of Lemma 3.2. Then PA admits an LU factorization, $PA = LU$, where L is unit lower triangular and U is upper triangular, and*

$$\mathbb{E}\, \|L[:, k]\|^2 \le 1 + k(n-k) \qquad \text{and} \qquad \mathbb{E}\, \|U[k, :]\|^2 \le k \sum_{i=k}^{n} \|Ae_i\|^2$$

for all $k \in [n]$.

Proof. Set $C = [k]$. In order for an LU factorization to exist, one needs each leading pivot block to be invertible. That is, if S is the set of the first k pivots, then we need $\det A[S, C] \neq 0$. From the definition of $\textsc{Vol}(A)$, we see that if $\det A[S, C] = 0$, S will be the set of first k pivots with probability 0; i.e. $\det A[S, C] \neq 0$ for the S sampled by $\textsc{Vol}(A)$ occurs with probability 1.

Let $(\pi_1, \ldots, \pi_n)$ be the permutation from Lemma 3.2 and set $S = \{\pi_1, \ldots, \pi_k\}$, $S' = \{\pi_1, \ldots, \pi_{k+1}\}$. By Cramer's rule,

$$L_{ik} = \frac{\det A[S - \{\pi_k\} + \{\pi_i\}, C]}{\det A[S, C]}, \quad U_{k+1,j} = \frac{\det A[S', C + \{c_j\}]}{\det A[S, C]}.$$

We first handle L.

$$\begin{aligned}
\|L[:, k]\|^2 = 1 + \sum_{i=k+1}^{n} |L_{ik}|^2 &= 1 + \sum_{i=k+1}^{n} \left| \frac{\det A[S - \{\pi_k\} + \{\pi_i\}, C]}{\det A[S, C]} \right|^2 \\
&\le 1 + \sum_{\pi \in S} \sum_{\pi' \notin S} \left| \frac{\det A[S - \{\pi\} + \{\pi'\}, C]}{\det A[S, C]} \right|^2.
\end{aligned} \tag{3.4}$$

Notice taking expectations with $\Pr(S) \propto (\det A[S, C])^2$ the denominator cancels giving us

$$\mathbb{E}\, \|L[:, k]\|^2 \le 1 + \frac{\sum_S \sum_{\pi \in S} \sum_{\pi' \notin S} (\det A[S - \{\pi\} + \{\pi'\}, C])^2}{\sum_S (\det A[S, C])^2}$$

where the outer sum is over $S \in \binom{[n]}{k}$. Notice $S - \{\pi\} + \{\pi'\}$ is just another element of $\binom{[n]}{k}$, and it is counted exactly $k(n-k)$ times which gives $\|L[:, k]\|^2 \le 1 + k(n-k)$. We now handle U in a similar fashion.

$$\begin{aligned}
\|U[k+1, :]\|^2 = \sum_{j=k+1}^{n} |U_{k+1,j}|^2 &= \sum_{j=k+1}^{n} \left| \frac{\det A[S', C + \{c_j\}]}{\det A[S, C]} \right|^2 \\
&\le \sum_{c' \notin C} \sum_{\pi' \notin S} \left| \frac{\det A[S + \{\pi'\}, C + \{c'\}]}{\det A[S, C]} \right|^2
\end{aligned} \tag{3.5}$$

Once again taking expectations with $\Pr(S) \propto (\det A[S,C])^2$ causes the denominator to cancel giving us

$$\mathbb{E}\,\|U[k+1,:]\|^2 \leq \sum_{c' \notin C} \frac{\sum_S \sum_{\pi' \notin S} |\det A[S+\{\pi'\}, C+\{c'\}]|^2}{\sum_S |\det A[S,C]|^2} \tag{3.6}$$

where the outer sum is over $S \in \binom{[n]}{k}$. Notice $S+\{\pi'\}$ is an element of $\binom{[n]}{k+1}$, and it is counted exactly $k+1$ times in the sum. Applying Cauchy-Binet to the numerator and denominator gives for $B = A^\top A$,

$$\mathbb{E}\,\|U[k+1,:]\|^2 \leq (k+1) \sum_{c' \notin C} \frac{\det B[C+\{c'\}, C+\{c'\}]}{\det B[C,C]} \leq (k+1) \sum_{c' \notin C} B_{c',c'}$$

where the final step used $\det B[C+\{c'\}, C+\{c'\}] \leq \det(B[C,C]) \cdot B_{c',c'}$, a consequence of Hadamard's maximum determinant inequality. Replacing k with $k-1$ gives the stated result. ☐

Here, we provide a lower bound as well, giving some indication as to how close Theorem 3.1 is to being tight. There is a well known lower bound of n for the growth of an $n \times n$ matrix over any pivoting strategy. Indeed, for any Hadamard matrix H (note any row or column permutation of a Hadamard matrix is also Hadamard) with an LU decomposition $H = LU$, we have $|U_{n,n}| = n$ [HH89]. In order to produce lower bounds for $\|L\|_F$ and $\|U\|_F$ under row permutations we utilize Sylvester Hadamard matrices.

Theorem 3.4. *Let $k \in \mathbb{N}$, $H_k = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes k}$ be the $n = 2^k$ dimensional Sylvester Hadamard matrix, and P be any permutation matrix for which PH_k has an LU factorization $PH_k = LU$. Then*

$$\|L\|_F^2 = 3^k = n^{\log_2 3} \qquad \text{and} \qquad \|U\|_F^2 = 6^k = n^{\log_2 6}.$$

Proof. We have

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}$$

and so, by the Kronecker mixed product property,

$$H_k = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{\otimes k} = \left[\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \right]^{\otimes k} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\otimes k} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}^{\otimes k} = LU.$$

This completes the proof for H_k.

What remains is to show that PH_k, if it has an LU factorization, has the exact same Frobenius norms for L and U. We proceed by induction on k for the larger class of signed and row permuted Sylvester matrices $A = DPH_k$, where D is a ± 1 diagonal matrix. We induct on the identities $\|L\|_F^2 = \|L^{-1}\|_F^2 = 3^k$. This immediately implies our desired result, as $AA^T = 2^k I$, so $UU^T = 2^k L^{-1} L^{-T}$ and $\|U\|_F^2 = 2^k \|L^{-1}\|_F^2$.

The base case of $k = 0$ is immediate. Now, let $n = 2^{k-1}$. The rows of H_k consists of the n pairs $(\boldsymbol{e}_i^T H_{k-1}, \boldsymbol{e}_i^T H_{k-1})$ and $(\boldsymbol{e}_i^T H_{k-1}, -\boldsymbol{e}_i^T H_{k-1})$ for $i \in [n]$. The first n rows of A consists of precisely one row from each of these pairs and possibly re-signs them, and the last n rows consists of the other row from each pair, otherwise the leading $n \times n$ principal minor of A would be zero. Therefore A takes the form

$$A = \begin{bmatrix} \widetilde{H} & \widetilde{D}\widetilde{H} \\ \widehat{H} & \widehat{D}\widehat{H} \end{bmatrix},$$

where $\widetilde{H}$ are $\widehat{H}$ are signed row permutations of H_{k-1}, and $\widetilde{D}$ and $\widehat{D}$ are ± 1 diagonal matrices. Let Q be the signed permutation matrix that reorders and re-signs $\widetilde{H}$ to $\widehat{H}$, i.e., $\widehat{H} = Q\widetilde{H}$. Then $Q\widetilde{D}Q^T = -\widehat{D}$, and so $Q\widetilde{D}\widetilde{H} = -\widehat{D}\widehat{H}$. We have

$$A = \begin{bmatrix} \widetilde{H} & \widetilde{D}\widetilde{H} \\ \widehat{H} & \widehat{D}\widehat{H} \end{bmatrix} = \begin{bmatrix} I & 0 \\ Q & I \end{bmatrix} \begin{bmatrix} \widetilde{H} & \widetilde{D}\widetilde{H} \\ 0 & 2\widehat{D}\widehat{H} \end{bmatrix}.$$

Because A has an LU factorization and the leading principal minors of A equal the leading principal minors of the second matrix in the above decomposition, both $\widetilde{H}$ and $\widehat{D}\widehat{H}$ have LU factorizations, which we induct on. Let $\widetilde{H} = \widetilde{L}\widetilde{U}$ and $\widehat{D}\widehat{H} = \widehat{L}\widehat{U}$. Then A has the LU factorization

$$\begin{bmatrix} \widetilde{H} & \widetilde{D}\widetilde{H} \\ \widehat{H} & \widehat{D}\widehat{H} \end{bmatrix} = \begin{bmatrix} \widetilde{L} & 0 \\ Q\widetilde{L} & \widehat{L} \end{bmatrix} \begin{bmatrix} \widetilde{U} & \widetilde{L}^{-1}\widetilde{D}\widetilde{H} \\ 0 & \widehat{U} - \widehat{L}^{-1}Q\widetilde{D}\widetilde{H} \end{bmatrix}$$

and

$$L^{-1} = \begin{bmatrix} \widetilde{L}^{-1} & 0 \\ -\widehat{L}^{-1}Q & \widehat{L}^{-1} \end{bmatrix}.$$

The matrix Q is orthogonal, and so, by induction,

$$\|L\|_F^2 = \|\widetilde{L}\|_F^2 + \|Q\widetilde{L}\|_F^2 + \|\widehat{L}\|_F^2 = 2\|\widetilde{L}\|_F^2 + \|\widehat{L}\|_F^2 = 3 \times 3^{k-1} = 3^k$$

and

$$\|L^{-1}\|_F^2 = \|\widetilde{L}^{-1}\|_F^2 + \|\widehat{L}^{-1}Q\|_F^2 + \|\widehat{L}^{-1}\|_F^2 = 2\|\widetilde{L}^{-1}\|_F^2 + \|\widehat{L}^{-1}\|_F^2 = 3 \times 3^{k-1} = 3^k.$$

□

Given the opportunity to reorder both rows and columns, the uniformity of the Frobenius norms of L and U no longer persists for Sylvester Hadamard matrices. However, we can produce a lower bound for any Hadamard matrix, thus providing a lower bound for any Hadamard matrix under row and column pivoting (as this class is closed under both operations).

Theorem 3.5. *Let $n \geq 2$ and H be an $n \times n$ Hadamard matrix with an LU factorization $H = LU$. Then*

$$\|L\|_F^2 \geq n \sum_{k=1}^{n} \frac{(k-1)^{k-1}}{k^k} \geq \frac{n \log n}{e} \qquad \text{and} \qquad \|U\|_F^2 \geq (n-1) + \sum_{k=1}^{n} \frac{k^k}{(k-1)^{k-1}} \geq \frac{en^2}{2}.$$

Proof. Let $\Delta_0^2 := 1$ and $\Delta_k^2 = \prod_{i=1}^k U_{ii}^2$, $k = 1, \ldots, n$, be the squared determinant of the leading $k \times k$ principal submatrix of U. By summing the squares of the diagonal entries and the $n-1$ off-diagonal entries in the first row of U, we have

$$\|U\|_F^2 \geq (n-1) + \sum_{k=1}^{n} \frac{\Delta_k^2}{\Delta_{k-1}^2}$$

and, summing the squares of the diagonal entries of U^{-1},

$$\|U^{-1}\|_F^2 \geq \sum_{k=1}^{n} \frac{\Delta_{k-1}^2}{\Delta_k^2}.$$

By Hadamard's inequality, and the fact that H is Hadamard matrix, $\Delta_k^2 \leq k^k$ for $k \in [n-1]$, and $\Delta_n^2 = n^n$. Now, define $\boldsymbol{x}_k = \log \Delta_k^2$. We have $\boldsymbol{x}_0 = 0$, $\boldsymbol{x}_n = n \log n$, and $0 \leq \boldsymbol{x}_k \leq k \log k$ for $k \in [n-1]$. Consider the functions

$$f(\boldsymbol{x}) = \sum_{k=1}^{n} e^{x_k - x_{k-1}} \qquad \text{and} \qquad g(\boldsymbol{x}) = \sum_{k=1}^{n} e^{-(x_k - x_{k-1})}.$$

Both are the sum of exponentials of linear functions, and so are convex functions over the convex set $0 \leq \boldsymbol{x}_k \leq k \log k$ for $k \in [n-1]$. Let $\boldsymbol{x}^*$ be the corner point $\boldsymbol{x}_k = k \log k$ for $k \in [n-1]$. We have

$$\frac{\partial f(\boldsymbol{x}^*)}{\partial \boldsymbol{x}_k} = e^{\boldsymbol{x}_k - \boldsymbol{x}_{k-1}} - e^{\boldsymbol{x}_{k+1} - \boldsymbol{x}_k} = \frac{k^k}{(k-1)^{k-1}} - \frac{(k+1)^{k+1}}{k^k} < 0$$

and

$$\frac{\partial g(\boldsymbol{x}^*)}{\partial \boldsymbol{x}_k} = -e^{-(\boldsymbol{x}_k - \boldsymbol{x}_{k-1})} + e^{-(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k)} = -\frac{(k-1)^{k-1}}{k^k} + \frac{k^k}{(k+1)^{k+1}} < 0.$$

Therefore, $\nabla f(\boldsymbol{x}^*)(\boldsymbol{y} - \boldsymbol{x}^*) \geq 0$ and $\nabla g(\boldsymbol{x}^*)(\boldsymbol{y} - \boldsymbol{x}^*) \geq 0$ for all feasible $\boldsymbol{y}$, and so $\boldsymbol{x}^*$ is the minimizer of both f and g over the convex set. This gives the lower bounds

$$\|U\|_F^2 \geq (n-1) + \sum_{k=1}^{n} \frac{k^k}{(k-1)^{k-1}} \qquad \text{and} \qquad \|U^{-1}\|_F^2 \geq \sum_{k=1}^{n} \frac{(k-1)^{k-1}}{k^k}.$$

Finally, because H is a Hadamard matrix, $\|L\|_F^2 = \|HU^{-1}\|_F^2 = n\|U^{-1}\|_F^2$. □

4. Hardness of Optimal Row Exchanges

The majority of research regarding pivoting in Gaussian elimination involves fixing a pivot strategy and analyzing the resulting element growth. One notable exception is the theorem [HH89, Theorem 2.1] of Higham and Higham, which provides a lower bound for growth over all row and column permutations. Using this theorem, they provided numerous examples of matrices that have $\mathrm{growth}(PAQ) \geq n/2$ for all permutation matrices P and Q [HH89, Equations 2.1 to 2.6]. Here, we show that recognizing matrices that admit a bounded growth row ordering is NP-complete, even for the fixed threshold of growth equal to one.

Several related elimination-ordering problems are known to be computationally hard. For instance, computing a minimum-fill ordering for sparse symmetric elimination is NP-hard [Yan81], determining the pivots selected by partial or complete pivoting is P-complete [Vav89], recognizing matrices admitting a perfect partial elimination scheme is NP-hard [BKS13], and global maximum volume column selection, which underlies maximum-volume and rank-revealing pivoting, is NP-hard [ÇMI09].

Our reduction uses incidence matrices between variables and clauses. Let $i \in [n]$ index variables, $j \in [m]$ index clauses, and, for an assignment $\boldsymbol{b} \in \{0,1\}^n$, let $M_{\boldsymbol{b}} \in \{0,1\}^{n\times m}$ be such that $(M_{\boldsymbol{b}})_{ij} = 1$ when setting $x_i = \boldsymbol{b}_i$ implies that the jth clause is true. The j^{th} coordinate of $\mathbf{1}^\top M_{\boldsymbol{b}}$ is the number of literals of the j^{th} clause satisfied by $\boldsymbol{b}$. The use of such a matrix in a hardness reduction is not novel; our new ingredient here is to place $M_{\boldsymbol{b}}$ inside a block Gaussian elimination gadget so that elimination produces the row

$$\frac{3}{2}\mathbf{1}^\top - \frac{1}{2}\mathbf{1}^\top M_{\boldsymbol{b}}.$$

The key property of this vector is that its j^{th} entry is at most 1 exactly when the j^{th} clause is satisfied.

Definition 4.1. Let f be a 3-CNF formula with variables x_i indexed by $[n]$ and clauses c_j indexed by $[m]$. For an assignment $\boldsymbol{b} \in \{0,1\}^n$, let $M_{\boldsymbol{b}}$ be the matrix $(M_{\boldsymbol{b}})_{ij} = [[(x_i = b_i) \implies c_j]]$ and define

$$A_f = \begin{array}{c} \begin{array}{ccccc} \scriptstyle 1 & \scriptstyle n & \scriptstyle 1 & \scriptstyle n & \scriptstyle m \end{array} \\ \begin{bmatrix} 1 & & & & -\mathbf{1}^\top \\ & I & & \frac{1}{2}I & M_{\mathbf{0}} \\ & I & & -\frac{1}{2}I & M_{\mathbf{1}} \\ \frac{1}{2} & \frac{1}{2}\mathbf{1}^\top & 1 & & \mathbf{1}^\top \\ & & & & I \end{bmatrix} \end{array} \begin{array}{c} \scriptstyle 1 \\ \scriptstyle n \\ \scriptstyle n \\ \scriptstyle 1 \\ \scriptstyle m \end{array} \in \mathrm{Mat}_{2n+m+2}(\tfrac{1}{2}\mathbb{Z}).$$

Notice that A_f is of polynomial order in n and m and can be constructed in polynomial time given f.

Theorem 4.2. *If f is satisfiable, then there is a permutation P such that* $\mathrm{growth}(PA_f) = 1$*. If f is not satisfiable, then for every permutation P, one has* $\mathrm{growth}(PA_f) \geq \frac{3}{2}$*.*

Proof. Suppose that f is satisfiable. We construct a permutation P such that $\mathrm{growth}(PA_f) = 1$. First note that $\|A_f\|_{\max} = 1$, so it suffices to bound $\|L\|_{\max}$ and $\|U\|_{\max}$ by one. For an assignment $\boldsymbol{b} \in \{0,1\}^n$, let $P_{\boldsymbol{b}}$ be the associated permutation that transposes rows $1+i$ and $1+i+n$ if and only if $b_i = 1$. Then

$$P_{\boldsymbol{b}}A_f = \begin{array}{c} \begin{array}{ccccc} \scriptstyle 1 & \scriptstyle n & \scriptstyle 1 & \scriptstyle n & \scriptstyle m \end{array} \\ \begin{bmatrix} 1 & & & & -\mathbf{1}^\top \\ & I & & \frac{1}{2}D & M_{\boldsymbol{b}} \\ & I & & -\frac{1}{2}D & M_{\mathbf{1}-\boldsymbol{b}} \\ \frac{1}{2} & \frac{1}{2}\mathbf{1}^\top & 1 & & \mathbf{1}^\top \\ & & & & I \end{bmatrix} \end{array} \begin{array}{c} \scriptstyle 1 \\ \scriptstyle n \\ \scriptstyle n \\ \scriptstyle 1 \\ \scriptstyle m \end{array},$$

where $D = I - 2\,\mathrm{diag}(\boldsymbol{b})$ is a diagonal ± 1 matrix. We apply one further permutation, which swaps the third and fourth block rows, to form $\widetilde{P}_{\boldsymbol{b}}$:

$$\widetilde{P}_{\boldsymbol{b}} A_f = \begin{array}{cc} \begin{array}{ccccc} 1 & n & 1 & n & m \end{array} & \\ \begin{bmatrix} 1 & & & & -\mathbf{1}^\top \\ & I & & \frac{1}{2}D & M_{\boldsymbol{b}} \\ \frac{1}{2} & \frac{1}{2}\mathbf{1}^\top & 1 & & \mathbf{1}^\top \\ & I & & -\frac{1}{2}D & M_{\mathbf{1}-\boldsymbol{b}} \\ & & & & I \end{bmatrix} & \begin{array}{c} 1 \\ n \\ 1 \\ n \\ m \end{array} \end{array}.$$

By inspection, $\widetilde{P}_{\boldsymbol{b}} A_f = LU$, where

$$L = \begin{array}{cc} \begin{array}{ccccc} 1 & n & 1 & n & m \end{array} & \\ \begin{bmatrix} 1 & & & & \\ & I & & & \\ \frac{1}{2} & \frac{1}{2}\mathbf{1}^\top & 1 & & \\ & I & & I & \\ & & & & I \end{bmatrix} & \begin{array}{c} 1 \\ n \\ 1 \\ n \\ m \end{array} \end{array} \quad \text{and} \quad U = \begin{array}{cc} \begin{array}{ccccc} 1 & n & 1 & n & m \end{array} & \\ \begin{bmatrix} 1 & & & & -\mathbf{1}^\top \\ & I & & \frac{1}{2}D & M_{\boldsymbol{b}} \\ & & 1 & -\frac{1}{4}\mathbf{1}^\top D & \frac{3}{2}\mathbf{1}^\top - \frac{1}{2}\mathbf{1}^\top M_{\boldsymbol{b}} \\ & & & -D & M_{\mathbf{1}-\boldsymbol{b}} - M_{\boldsymbol{b}} \\ & & & & I \end{bmatrix} & \begin{array}{c} 1 \\ n \\ 1 \\ n \\ m \end{array} \end{array}.$$

We have $\|L\|_{\max} = 1$ and $\|U\|_{\max} = \max\{1, \left\|\frac{3}{2}\mathbf{1} - \frac{1}{2}M_b^\top \mathbf{1}\right\|_\infty\}$. If $\boldsymbol{b}$ is indeed a satisfying assignment, then $M_{\boldsymbol{b}}$ has $1, 2$, or 3 non-zero entries in each column, all equal to 1, and so $M_{\boldsymbol{b}}^\top \mathbf{1} \in \{1,2,3\}^m$. This means $\|U\|_{\max} = 1$ and $\mathrm{growth}(\widetilde{P}_{\boldsymbol{b}} A_f) = 1$, as desired.

Conversely, suppose that there exists a permutation P with $\mathrm{growth}(PA_f) < \frac{3}{2}$. The first pivot must come from the first block, as using the fourth block leads to $\|L\|_{\max} \geq 2$. Now, consider the next n pivots chosen. Again, during these n steps, using the fourth block leads to $\|L\|_{\max} \geq 2$, so the $(1+i)^{\text{th}}$ pivot must have index $1+i$ or $1+i+n$, corresponding to the choice b_i equals 0 or 1. In particular, the first $n+1$ rows of the permutation P must agree with a permutation $P_{\boldsymbol{b}}$ for some $\boldsymbol{b}$. There is only one choice for the $(n+2)^{nd}$ pivot, the entry in the third block row, and so the third block row and fifth block column of the resulting U must contain $\frac{3}{2}\mathbf{1}^\top - \frac{1}{2}\mathbf{1}^\top M_{\boldsymbol{b}}$. In order for $\mathrm{growth}(PA_f) < \frac{3}{2}$, we must have $\left\|\frac{3}{2}\mathbf{1} - \frac{1}{2}M_b^\top \mathbf{1}\right\|_\infty < \frac{3}{2}$. In particular, every entry of $M_b^\top \mathbf{1}$ must be positive, which means that $\boldsymbol{b}$ is a satisfying assignment of f. $\square$

Theorem 4.2 immediately implies our desired Theorem 1.4. Indeed, the factorization provided in the above proof is valid for any choice of $\boldsymbol{b}$, independent of whether it satisfies f or not, and provides a certification that $A_f \in \mathrm{GL}_{2n+m+2}(\mathbb{R}) \cap \mathrm{Mat}_{2n+m+2}(\frac{1}{2}\mathbb{Z})$, as the diagonal blocks 1, I, 1, $-D$, and I of U are all non-singular. To see that this language is a member of NP, we note that the computation of a minor of a matrix can be done in polynomial time, and its bit length is bounded. The existence of an LU factorization of a matrix and the entries of the LU factorization, if it exists, are therefore computable in polynomial time.

5. Large Growth of Randomized Partial Pivoting

Recall that randomized partial pivoting, $\textsc{RandPP}_p$, is a procedure for performing row exchanges where the next pivot index π_{k+1} is selected according to the marginal distribution

$$\Pr(\pi_{k+1} = i \,|\, \{\pi_1, \ldots, \pi_k\} = S) \propto \left| \frac{\det A[S + \{i\}, [k+1]]}{\det A[S, [k]]} \right|^p, \tag{5.1}$$

where $\pi_1, \ldots, \pi_k$ are the pivot rows chosen at steps 1 through k of the algorithm. By (1.1),

$$A^{(k)} = A[S^c, [k]^c] - A[S^c, [k]] A[S, [k]]^{-1} A[S, [k]^c],$$

and so, indexing the rows and columns of $A^{(k)}$ using the indices of S^c and $[k]^c$, respectively,

$$a_{ij}^{(k)} = a_{ij} - A[\{i\}, [k]] A[S, [k]]^{-1} A[S, \{j\}].$$

Therefore, by Schur's formula,

$$\Pr(\pi_{k+1} = i \mid A^{(k)}) \propto \left|\frac{\det A[S+\{i\},[k+1]]}{\det A[S,[k]]}\right|^p = \left|a_{ik}^{(k)}\right|^p.$$

When $p \to \infty$, one recovers ordinary deterministic partial pivoting (up to random selection among ties); when $p \to 0$, one obtains uniform pivoting. Neither extreme is desirable: partial pivoting has exponential growth in U in the worst case, and uniform pivoting is liable to select tiny pivots, leading to arbitrarily large growth in L. For this reason, the numerical analysis community has focused on intermediate values of p, most naturally $p = 1$ and $p = 2$, which favors large pivots, and, hence, usually moderate multipliers, while also occasionally deviating from the greedy trajectory.

A folklore expectation that followed naturally was that this process should introduce enough randomness to ensure polynomial growth with high probability. We are not aware of a formal statement of this conjecture in the published literature, but similar ideas have appeared. The complete pivoting analog to $\textsc{RandPP}_p$ is to select both a row and column index, in proportion to a power of the entry, $\propto |a_{ij}^{(k)}|^p$; this was introduced for $p = 2$ by [GW26]. In addition, there is a more specialized analog in the positive-definite setting, where one samples the row and column indices to be the same, in proportion to a power of the diagonal entry, $\propto |a_{ii}^{(k)}|^p$ [CETW25].

Separately, randomization is known to yield results elsewhere in the elimination process. Average-case and smoothed analyses randomize the input matrix and establish polynomial growth or precision bounds for Gaussian inputs and Gaussian perturbations [TS90, CP07, SST06, HT24]. Randomized complete pivoting instead uses a random sketch of the Schur complement to locate an approximately maximal pivot, obtaining complete pivoting-type growth bounds with high probability [MG15]. In both cases, the source and role of the randomness differ from that of $\textsc{RandPP}_p$, though they highlight the apparent utility of randomization.

Here, we disprove any such conjecture that sampling $P \sim \textsc{RandPP}_p(A)$ results in low growth with high probability. In particular, we provide an example for which the growth is nearly exponential. Our construction is the Q factor from the QR decomposition of the following ill-conditioned transposed Jordan block

$$N = \begin{pmatrix} z & & & & \\ 1 & z & & & \\ & 1 & z & & \\ & & \ddots & \ddots & \\ & & & 1 & z \end{pmatrix} = QR \tag{5.2}$$

for $z = \exp\left(-\frac{1}{pn^\alpha}\right) \approx 1 - \frac{1}{pn^\alpha}$, for a tunable $\alpha \in (0,1)$ which is allowed to depend on n. Our strategy is to show that the final pivot chosen by $\textsc{RandPP}_p(Q)$ concentrates toward the end of the matrix. That is, if $\pi_1, \ldots, \pi_n$ is the random sequence of pivot indices chosen by $\textsc{RandPP}_p(Q)$, then π_n will concentrate on the set $[n - o(n), n]$. Under this event, the growth of Q under $\textsc{RandPP}_p$ pivoting will resemble the growth of Q with no pivoting up to $n - o(n)$ iterations, which is exponential in z.

Lemma 5.1. *If $(\pi_1, \ldots, \pi_n) \sim \textsc{RandPP}_p(Q)$, then $\Pr(\pi_n \geq n - n^\alpha) \geq 1 - ne^{-n^\alpha/(e+1)}$.*

Proof. The central insight is that the expression for the distribution of the next pivot index from (5.1) is invariant to the multiplication of A on the right by non-singular upper triangular matrices. In particular,

$$\det(Q[S,:]R[:,[k]]) = \det(Q[S,[k]])\det(R[[k],[k]]),$$

and so

$$\textsc{RandPP}_p(N) = \textsc{RandPP}_p(Q).$$

This is quite convenient, as $\textsc{RandPP}_p(N)$ is easier to control.

Note that the first k columns of N only have non-zero entries in the first $k+1$ rows, and so the first k pivot indices are tightly contained in $\{\pi_1, \ldots, \pi_k\} \subset [k+1]$. The set $\{\pi_1, \ldots, \pi_k\}$ has exactly one element missing from $[k+1]$. Four our purposes, it is convenient to think not of sets of pivot indices, but of the sequence of row indices that are *not* pivots: let m_{k+1} be the unique element of $[k+1] - \{\pi_1, \ldots, \pi_k\}$. Note that $\pi_n = m_n$.

The relevant submatrices have a particularly nice block diagonal structure

$$N\left[[k+1]-\{m_{k+1}\},[k]\right]=\left(\begin{array}{ccc|ccc} z & & & & & \\ 1 & \ddots & & & & \\ & \ddots & z & & & \\ \hline & & & 1 & z & \\ & & & & \ddots & \ddots \\ & & & & & 1 \end{array}\right),$$

where the upper left $(m_k-1)\times(m_k-1)$ block is lower triangular with diagonal entries equal to z, and the lower right $(k-m_k+1)\times(k-m_k+1)$ block is upper triangular with diagonal entries equal to 1.

Consequently,

$$\det N[[k+1]-\{m_{k+1}\},[k]]=z^{m_{k+1}-1},$$

which translates to

$$\begin{aligned} &\Pr\left(\pi_{k+1}=m_{k+1}\,|\,\{\pi_1,\dots,\pi_k\}=[k+1]-\{m_{k+1}\}\right)\propto z^{(k+2-m_{k+1})p}, \\ &\Pr\left(\pi_{k+1}=k+2\,|\,\{\pi_1,\dots,\pi_k\}=[k+1]-\{m_{k+1}\}\right)\propto 1, \end{aligned} \tag{5.3}$$

and all other probabilities equal to 0. Notice that this uniquely determines m_{k+2}: if $\pi_{k+1}=m_{k+1}$, then $m_{k+2}=k+2$, and if $\pi_{k+1}=k+2$, then $m_{k+2}=m_{k+1}$. By changing variables to $\delta_k=k-m_k$, this becomes a Markov chain

$$\delta_{k+1}=\begin{cases}\delta_k+1 & \text{with probability } \dfrac{1}{1+z^{(1+\delta_k)p}} \\ 0 & \text{otherwise}\end{cases}.$$

From this, one can bound the height of the excursions of δ_k from 0

$$\Pr(\delta_{k+t}=t\,|\,\delta_k=0)=\prod_{j=1}^{t}\left(1+z^{jp}\right)^{-1}.$$

Notice that $e^{-1}\le z^{jp}\le 1$ is satisfied for $j\le 1/(p\log(1/z))=n^\alpha$, so for $t\ge n^\alpha$ we have

$$\Pr(\delta_{k+t}=t\,|\,\delta_k=0)\le\prod_{j=1}^{n^\alpha}\left(1+z^{jp}\right)^{-1}\le\left(\frac{e}{e+1}\right)^{n^\alpha}\le\exp\left(-\frac{n^\alpha}{e+1}\right).$$

Stated differently, each excursion δ_k makes from a given 0 value has length at most n^α with probability $1-e^{-n^\alpha/(e+1)}$. By applying a union bound over starting points of the excursion, we have that the maximum of the δ_k (and therefore δ_n) is bounded by n^α with probability $ne^{-n^\alpha/(e+1)}$. Finally, $\delta_n\le n^\alpha$ is equivalent to $m_n=\pi_n\ge n-n^\alpha$. □

Theorem 5.2. *Let Q be as in* (5.2) *for $\alpha\in(0,1)$ and $P\sim\textsc{RandPP}_p(Q)$. Then for $PQ=LU$ we have*

$$\Pr\left(\|U\|_{\max}\ge\exp\left(\frac{n^{1-\alpha}-1-n^{-\alpha}}{p}\right)\right)\ge 1-ne^{-n^\alpha/(e+1)}.$$

In particular, with $\alpha=(\log 8+\log\log n)/\log n$, we have

$$\Pr\left(\|U\|_{\max}\ge\exp\left(\frac{n-16\log n}{8p\log n}\right)\right)\ge 1-n^{1-8/(e+1)}.$$

Proof. Because PQ is orthogonal, we have for $PQ=LU$ that

$$U_{nn}^{-1}=(U^{-1})_{nn}=((PQ)^{-1})_{nn}=((PQ)^\top)_{nn}=(PQ)_{nn},$$

which is the π_n^{th} entry of the last column of Q. By construction $Q\boldsymbol{e}_n$ is the unique unit vector (up to sign) orthogonal to the leading $n-1$ columns of N, and so must be proportional to

$$Q\boldsymbol{e}_n\propto\begin{pmatrix}1 & -z & (-z)^2 & \cdots & (-z)^{n-1}\end{pmatrix}^\top.$$

Note that this vector has norm at least one. When $P \sim \textsc{RandPP}_p(Q)$, Lemma 5.1 says that π_n is at least $n - n^\alpha$ with probability $1 - ne^{-n^\alpha}$, and so the π_n^{th} entry of $Q\boldsymbol{e}_n$ is at most $z^{n-n^\alpha-1}$ in absolute value. Therefore,

$$\|U\|_{\max} \geq |U_{nn}| \geq z^{-(n-n^\alpha-1)} = \exp\left(\frac{n^{1-\alpha} - 1 - n^{-\alpha}}{p}\right)$$

with probability $1 - ne^{-n^\alpha/(e+1)}$. □

6. Quasi-Polynomial Growth of Rook Pivoting

Here, we construct a sequence of matrices for which Gaussian elimination with rook pivoting exhibits growth $n^{(\frac{1}{4}-o(1))\log_2 n}$. This is nearly within a factor of two of the exponent in Foster's upper bound $n^{\frac{3\log 2}{4}\log_2 n} \approx n^{0.52\log_2 n}$ [Fos97]. Our construction is recursive. At each step, random orthogonal matrices are used to spread the mass sufficiently to ensure the rook pivoting condition is satisfied, with the orthogonal factors arranged so that they cancel as elimination proceeds, allowing for a recursive structure.

First, we present two concentration bounds for the entries of products involving Haar-distributed matrices, which we use in this section and the next. Here, we only need the first bound for the special case $Y = I$.

Lemma 6.1. *Fix any $X \in \mathbb{R}^{n\times p}$, $Y \in \mathbb{R}^{n\times q}$, and $M \in \mathbb{R}^{n\times n}$. Set $L = \log(2pq/\delta)$. If $Q \sim \mathrm{Haar}(\mathrm{O}(n))$, then*

$$\Pr\left(\left\|X^\top Q Y\right\|_{\max} \leq \|X\|_{1\to 2}\|Y\|_{1\to 2}\sqrt{\frac{2L}{n}}\right) \geq 1 - \delta \tag{6.1}$$

and

$$\Pr\left(\left\|X^\top Q^\top M Q Y - \frac{\operatorname{tr} M}{n} X^\top Y\right\|_{\max} \leq C\|X\|_{1\to 2}\|Y\|_{1\to 2}\left(\frac{\|M\|_F\sqrt{L}}{n} + \frac{\|M\|L}{n}\right)\right) \geq 1 - \delta \tag{6.2}$$

for an absolute constant C.

Proof. Consider an arbitrary pair of columns $\boldsymbol{x} = X\boldsymbol{e}_i$ and $\boldsymbol{y} = Y\boldsymbol{e}_j$. By rotational invariance, $\boldsymbol{x}^\top Q\boldsymbol{y}/(\|\boldsymbol{x}\|\|\boldsymbol{y}\|)$ is distributed as the first coordinate of a uniform random vector sampled from the unit sphere $\mathcal{S}^{n-1}$, which, as a standard concentration result, satisfies

$$\Pr\left(\frac{|\boldsymbol{x}^\top Q\boldsymbol{y}|}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} > \frac{\gamma}{\sqrt{n}}\right) \leq 2e^{-\gamma^2/2}$$

for every $\gamma > 0$, see, e.g., [B+, Lemma 2.2]. Taking $x = \sqrt{2L}$ and applying a union bound over the pq pairs of columns of X and Y gives the first result (6.1).

For the second estimate, express x and y in an orthonormal basis for their span, e.g. $x = Ea$ and $y = Eb$ for $E^\top E = I$ and $\mathrm{Col}(E) = \mathrm{span}\{x, y\}$. Then $Z = QE$ is a Haar-distributed sample from the Stiefel manifold of order at most two, and

$$x^\top Q^\top M Q y = a^\top Z^\top M Z b = \mathrm{vec}(Z)^\top \frac{ab^\top \otimes M + ba^\top \otimes M^\top}{2} \mathrm{vec}(Z).$$

Set $\widetilde{M} = \frac{ab^\top \otimes M + ba^\top \otimes M^\top}{2}$. In particular, we may apply the Stiefel Hanson-Wright inequality [GS23, Theorem 7.1(1)], which claims

$$\Pr\left(\left|\mathrm{vec}(Z)^\top \widetilde{M}\,\mathrm{vec}(Z) - \frac{1}{n}\operatorname{tr}(\widetilde{M})\right| \geq t\right) \leq 2\exp\left(-c\min\left(\frac{(n-2)^2 t^2}{\|\widetilde{M}\|_F^2}, \frac{(n-2)t}{\|\widetilde{M}\|}\right)\right)$$

for $c = 7/10000$. Observe that $\operatorname{tr}(\widetilde{M}) = x^\top y \operatorname{tr}(M)$, $\|\widetilde{M}\|_F \leq \|x\|\|y\|\|M\|_F$, and $\|\widetilde{M}\| \leq \|x\|\|y\|\|M\|$. Setting t to be a sufficiently large multiple of $\|x\|\|y\|\left(\frac{\|M\|_F\sqrt{L}}{n} + \frac{\|M\|L}{n}\right)$ and applying union bound over pq pairs of columns of X and Y gives (6.2). □

Theorem 6.2. *There is a universal constant C such that for each $n \in \mathbb{N}$, there is a rook pivoted matrix $A \in \mathrm{RP}_n(\mathbb{R})$ with $\mathrm{growth}(A) \geq n^{\frac{1}{4}\log_2(n) - C\log\log n}$.*

Proof. It suffices to prove the theorem statement when n is a power of two by considering the matrix $\|A\|_{\max} I_{n-2^{\lfloor \log_2 n \rfloor}} \oplus A$ where $A \in \mathrm{RP}_{2^{\lfloor \log_2(n) \rfloor}}(\mathbb{R})$. Let $Q_0 = \begin{bmatrix} 1 \end{bmatrix}$, $Q_1 = I_2$, and $Q_k \sim \mathrm{Haar}(\mathrm{O}(2^k))$ be independent random Haar orthogonal matrices for $k \geq 2$. Define $s_k \in \mathbb{R}_+$ recursively by $s_0 = 1$ and

$$s_k = 4\sqrt{\frac{k}{2^k}} \max(s_j : j < k).$$

Define $X_k \in \mathrm{Mat}_{2^k}(\mathbb{R})$ recursively by $X_0 = \begin{bmatrix} 1 \end{bmatrix}$ and

$$X_{k+1} = \begin{bmatrix} s_k I & Q_k^\top \\ -X_k Q_k & 0 \end{bmatrix}.$$

Our construction is simply the matrix X_{K+1} for $K = \log_2(n) - 1$, conditioned on a particular positive probability event. Our analysis is in two parts: showing that X_{K+1} has the desired growth and that X_{K+1} is rook-pivoted with positive probability. Both parts require control on the row norms of X_k, i.e. the matrix norm $\|X_k\|_{2\to\infty}$. To this end, note by definition that

$$\|X_{k+1}\|_{2\to\infty}^2 \leq \max\Big(\|X_k Q_k\|_{2\to\infty}^2, s_k^2 + 1\Big) \leq \max\Big(\|X_k\|_{2\to\infty}^2, 2\max(s_j : j < k)^2\Big)$$

so by induction,

$$\|X_k\|_{2\to\infty} \leq \sqrt{2}\max(s_j : j < k).$$

Furthermore, s_k is eventually decreasing so $\sup_{k\in\mathbb{N}} s_k =: C$ is finite. We therefore have

$$\|X_{K+1}\|_{\max} \leq \|X_{K+1}\|_{2\to\infty} \leq \sqrt{2}C.$$

On the other hand, observe that 2^k steps of Gaussian elimination applied to X_{k+1} results in $\frac{1}{s_k}X_k$, and thus Gaussian elimination on X_{K+1} eventually results in the scalar $(s_K s_{K-1} \cdots s_1 s_0)^{-1}$. Additionally, $s_k \leq 4C\sqrt{\frac{k}{2^k}}$ so

$$\mathrm{growth}(X_{K+1}) \geq \frac{(s_K s_{K-1} \cdots s_1 s_0)^{-1}}{\|X_{K+1}\|_{\max}} \geq \frac{2^{\frac{1}{4}(K^2+K)}}{(4C)^{K+2}\sqrt{K!}} \geq 2^{\frac{1}{4}K^2 - (K+2)\log_2(4CK)},$$

which establishes our growth claim. We now turn our attention to pivoting. Let $\mathcal{E}_k$ be the event that $\left\| \begin{bmatrix} X_k \\ I \end{bmatrix} Q_k \right\|_{\max} \leq s_k$. Observe that the intersection $\mathcal{E}_0 \cap \cdots \cap \mathcal{E}_K$ implies that X_{K+1} is rook-pivoted. $\mathcal{E}_0$ and $\mathcal{E}_1$ occur by selection of $Q_0 = \begin{bmatrix} 1 \end{bmatrix}$, $Q_1 = I_2$ and $s_0 = 1$, $s_1 = 4/\sqrt{2}$. For each $k \geq 2$, conditioned on the values of $Q_1, \ldots, Q_{k-1}$, the matrix X_k is a fixed matrix with $\|X_k\|_{2\to\infty} = \left\|X_k^\top\right\|_{1\to 2} \leq \sqrt{2}\max(s_j : j < k)$. Lemma 6.1 therefore gives

$$\begin{aligned} \Pr(\|X_k Q_k\|_{\max} \geq s_k) &= \Pr\left(\|X_k Q_k\|_{\max} \geq \sqrt{\frac{8k}{2^k}} \cdot \sqrt{2}\max(s_j : j < k)\right) < \frac{1}{2}, \\ \Pr(\|Q_k\|_{\max} \geq s_k) &\leq \Pr\left(\|Q_k\|_{\max} \geq 4\sqrt{\frac{k}{2^k}}\right) < \frac{1}{2}, \end{aligned} \tag{6.3}$$

where the second inequality used the trivial $\max(s_j : j < k) \geq 1$. By union bound, the probability of either event occuring is strictly less than 1. Finally, $\Pr(\mathcal{E}_K \cap \cdots \cap \mathcal{E}_0) = \prod_{k=1}^K \Pr(\mathcal{E}_k \,|\, \mathcal{E}_{k-1} \cap \cdots \cap \mathcal{E}_0) > 0$, which establishes $X_k \in \mathrm{RP}_{2^k}(\mathbb{R})$ with positive probability. □

7. Quasi-Polynomial Growth of Complete Pivoting

Here, we produce a sequence of completely pivoted matrices with quasi-polynomial growth factor. Combined with the quasi-polynomial upper bound of approximately $n^{0.207\log n}$ [BEU25], our result settles the asymptotic behavior of the growth factor under complete pivoting.

Lemma 7.1 ([LTTJ20, Theorem 1.3])**.** *If G is an $n \times n$ Gaussian matrix then*

$$\Pr\Big(\left\|G^{-1}\right\|_F \leq \min\left\{n/t, \sqrt{n/t}\right\}\Big) \leq 2e^{-ct^2}.$$

for all $t > 0$ and some absolute constant c.

Lemma 7.2 ([CD05, Lemma 4.1])**.** *If G is an $n \times m$ Gaussian matrix then*

$$\Pr\big(\sigma_m(G) \le t\sqrt{n}\big) \le \frac{(tn)^{n-m+1}}{\Gamma(n-m+2)}$$

for all $t > 0$.

Lemma 7.3. *Let $W \sim \mathrm{Haar}(\mathrm{O}(n))$. Then with probability $1 - e^{-cn}$, every $m \times m$ submatrix X of W with $n^{6/7} \le m \le \frac{n}{2}$ simultaneously satisfies*

$$\left\|X^{-1}\right\|_F^2 \ge cn^{8/7}$$

for an absolute constant c.

Proof. We prove the claim for a fixed $m \times m$ submatrix X and then take a union bound over all such submatrices. Let $G_1 \in \mathbb{R}^{m \times m}$ and $G_2 \in \mathbb{R}^{(n-m)\times m}$ be independent standard Gaussian matrices. Then

$$X \overset{d}{=} G_1\big(G_1^\top G_1 + G_2^\top G_2\big)^{-1/2},$$

and therefore for the appropriate coupling of X and (G_1, G_2),

$$\left\|X^{-1}\right\|_F \ge \sigma_m\left(\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}\right)\left\|G_1^{-1}\right\|_F. \tag{7.1}$$

Applying Lemma 7.1 in dimension m, with $t = n^{5/7} \le m$, gives

$$\Pr\left(\left\|G_1^{-1}\right\|_F^2 \le \frac{m}{n^{5/7}}\right) \le 2\exp\Big(-cn^{10/7}\Big). \tag{7.2}$$

On the other hand Lemma 7.2 applied to $\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$, gives

$$\Pr\left(\sigma_m\left(\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}\right) \le t\sqrt{n}\right) \le \frac{(tn)^{n-m+1}}{\Gamma(n-m+2)} \le \left(\frac{ent}{n-m+1}\right)^{n-m+1} \le (2et)^{n/2}, \tag{7.3}$$

where the last inequality used $m \le \frac{n}{2}$. We will set $t = \frac{1}{100}$ here. It follows from (7.1) to (7.3) and $m \ge n^{6/7}$ that

$$\Pr\left(\left\|X^{-1}\right\|_F^2 < \frac{n^{8/7}}{100^2}\right) \le 2\exp\Big(-cn^{10/7}\Big) + \left(\frac{2e}{100}\right)^{n/2}.$$

Notice $\sqrt{2e/100} < \frac{1}{4}$ and that the total number of submatrices of an $n \times n$ matrix is at most 4^n, so union bound gives the result. $\square$

Lemma 7.4. *Let $W \in \mathrm{O}(n)$ be any orthogonal matrix. For every $m \times m$ non-singular submatrix X of W, there exists a square non-singular submatrix Y of W of order* $\min(m, n-m)$ *with*

$$\|(W/X)\|_F \ge \left\|Y^{-1}\right\|_F. \tag{7.4}$$

Moreover,

$$\|(W/X)\|_{\max} \ge \frac{1}{\sqrt{n-m}}. \tag{7.5}$$

Proof. Decompose

$$W = \begin{bmatrix} X & W_{12} \\ W_{21} & W_{22} \end{bmatrix}.$$

By the Schur formula for the inverse, $((W^{-1})_{22})^{-1} = (W/X)$. But since W is orthogonal, $W^{-1} = W^\top$ so

$$(W/X) = (W_{22}^\top)^{-1}.$$

Since W_{22} is the submatrix of an orthogonal matrix, all of its singular values are at most 1. Thus, all singular values of (W/X) are at least 1, so $(n-m)^2\,\|(W/X)\|_{\max}^2 \ge \|(W/X)\|_F^2 \ge n-m$, establishing (7.5).

Now consider $m \ge n/2$. Since taking the transpose does not change the Frobenius norm, we may take $Y = W_{22}$ to obtain (7.4) with equality. Now consider $m \le n/2$. The $n-m$ singular values of W_{22} are exactly the m singular values of X with $n-2m$ copies of 1. In particular, $\|W_{22}^{-1}\|_F \ge \|X^{-1}\|_F$ so we may take $Y = X$. $\square$

Lemma 7.5. *For all sufficiently large n, there is $W \in \mathrm{O}(n) \cap \mathrm{CP}_n(\mathbb{R})$ such that*

$$\|W\|_{\max} \leq C\sqrt{\frac{\log n}{n}}. \tag{7.6}$$

If $W^{(j)}$ is the residual matrix after j steps of complete pivoting on W, then

$$\left\|W^{(j)}\right\|_{\max} \geq c\frac{j}{n^{10/7}}. \tag{7.7}$$

Proof. Sample $W \sim \mathrm{Haar}(\mathrm{O}(n))$. We will show it has the desired properties with positive probability. Recall that $W^{(j)}$ is a Schur complement of W. First we control $W^{(j)}$ for $j \in [n^{6/7}, n - n^{6/7}]$. By Lemma 7.4, (7.4), for each j, there is a $\min(j, n-j) \times \min(j, n-j)$ submatrix Y_j of W with

$$\left\|W^{(j)}\right\|_F \geq \left\|Y_j^{-1}\right\|_F.$$

By Lemma 7.3, $\left\|Y_j^{-1}\right\|_F \geq cn^{4/7}$ for all j simultaneously with probability at least $1 - e^{cn}$. Using $(n-j)^2 \left\|W^{(j)}\right\|_{\max}^2 \geq \left\|W^{(j)}\right\|_F^2$, this becomes

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{cn^{4/7}}{n-j} \geq 4c\frac{j}{n^{10/7}} \tag{7.8}$$

where the second inequality uses $j(n-j) \leq \frac{1}{4}n^2$. Now, consider $j < n^{6/7}$. By Lemma 7.4, (7.5), we directly have

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{1}{\sqrt{n-j}} \geq \frac{1}{\sqrt{n}} \geq \frac{j}{n^{10/7}}. \tag{7.9}$$

Finally, consider $j > n - n^{6/7}$. Again by Lemma 7.4, (7.5), we directly have

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{1}{\sqrt{n-j}} \geq \frac{1}{n^{3/7}} = \frac{n}{n^{10/7}} \geq \frac{j}{n^{10/7}}. \tag{7.10}$$

Together, (7.8) to (7.10) give (7.7) with high probability. Next, (7.6) holds with high probability by Lemma 6.1. Finally, permuting rows and columns preserves $\|\cdot\|_{\max}$, so we may freely permute the rows and columns of W so that $W \in \mathrm{CP}_n$. Taking the union bound over these events gives the desired construction. ☐

Definition 7.6. For $A \neq 0$, let $\widehat{A} = A / \|A\|$. Denote by $B(A)$ the $4n \times 4n$ random matrix

$$B(A) = \left[\begin{array}{cc|cc} I & 0 & n^{7/30}Q^\top & n^{2/15}\widehat{A} \\ 0 & I & n^{11/60}WQ^\top & 0 \\ \hline n^{7/30}WQ & 0 & 0 & 0 \\ 0 & n^{11/60}W^\top & 0 & 0 \end{array}\right]$$

where W is the particular orthogonal matrix of Lemma 7.5 and $Q \sim \mathrm{Haar}(\mathrm{O}(n))$.

It will be useful to have written down what B looks like after a certain number of Gaussian elimination steps (without pivoting) have been performed. For the first $2n$ iterations, the upper-left identity blocks cause a leading row and column to be combined as an outer produced and placed on the lower right, as the Schur complement. Let E_k be the trailing $n-k$ columns of the identity and $P_k = I - E_k E_k^\top$ be the projection matrix onto the first k coordinates. Then

$$B^{(k)} = \left[\begin{array}{cc|cc} I_{n-k} & 0 & n^{7/30}E_k^\top Q^\top & n^{2/15}E_k^\top \widehat{A} \\ 0 & I_n & n^{11/60}WQ^\top & 0 \\ \hline n^{7/30}WQE_k & 0 & -n^{14/30}WQP_kQ^\top & -n^{11/30}WQP_k\widehat{A} \\ 0 & n^{11/60}W^\top & 0 & 0 \end{array}\right], \qquad 0 \leq k \leq n \tag{7.11}$$

We start getting population into the (4,3) block after the first identity is completely eliminated,

$$B^{(n+k)} = \left[\begin{array}{c|cc} I_{n-k} & n^{11/60}E_k^\top WQ^\top & 0 \\ \hline 0 & -n^{14/30}W & -n^{11/30}WQ\widehat{A} \\ n^{11/60}W^\top E_k & -n^{11/30}W^\top P_k WQ^\top & 0 \end{array}\right], \qquad 0 \le k \le n \tag{7.12}$$

and after the second identity is eliminated,

$$B^{(2n)} = -\begin{bmatrix} n^{14/30}W & n^{11/30}WQ\widehat{A} \\ n^{11/30}Q^\top & 0 \end{bmatrix}. \tag{7.13}$$

In both cases, we benefit from the cancellations $QQ^\top = I$ and $W^\top W = I$ once the projection $P_k = I$ has filled out. After the remaining eliminations,

$$B^{(3n)} = n^{4/15}\widehat{A}. \tag{7.14}$$

In order to show that this is the complete-pivoting path, we need to argue that the upper left entry of each $B^{(k)}$, $B^{(n+k)}$ for $0 \le k < n$ is equal to the max-norm.

Lemma 7.7. *$n^{7/30} \le \|B(A)\| \le 3n^{7/30}$ surely.*

Proof. For the lower bound, simply inspect the $(3,1)$ block. For the upper bound, one can replace each block of B with its spectral norm and take the Frobenious norm of the resulting 4×4 matrix,

$$n^{7/30} = \left\| n^{7/30}WQ \right\| \le \|B(A)\| \le \left\| \begin{bmatrix} 1 & 0 & n^{7/30} & n^{2/15} \\ 0 & 1 & n^{11/60} & 0 \\ n^{7/30} & 0 & 0 & 0 \\ 0 & n^{11/60} & 0 & 0 \end{bmatrix} \right\|_F \le 3n^{7/30}.$$

$\square$

In the following lemmas, we implicitly invoke Lemma 6.1 several times, where δ is some sufficiently small polynomial in n, say $\delta = n^{-10}$, so that the stated overall probability claims hold by a simple union bound.

Lemma 7.8. *If $\|A\| \ge \frac{1}{4}n^{7/30}$ and $\|A\|_{\max} \le 1$ then $\left\|B^{(k)}\right\|_{\max} \le 1$ for $0 \le k < n$ with probability at least $1 - 1/n$.*

Proof. Start with the nonzero blocks within the lower right quadrant,

$$\begin{aligned} \left\| n^{7/30}WQE_k \right\|_{\max} &\le n^{7/30}\|WQ\|_{\max} \lesssim n^{-8/30}\sqrt{\log n} \\ \left\| n^{11/60}W^\top \right\|_{\max} &\le n^{11/60}\|W\|_{\max} \lesssim n^{-19/60}\sqrt{\log n}. \end{aligned} \tag{7.15}$$

Now for the nonzero blocks in the upper right quadrant,

$$\begin{aligned} \left\| n^{7/30}E_k^\top Q^\top \right\|_{\max} \le \left\| n^{7/30}E_k^\top Q^\top \right\|_{\max} &\le n^{7/30}\|Q\|_{\max} \lesssim n^{-8/30}\sqrt{\log n} \\ \left\| n^{2/15}E_k^\top \widehat{A} \right\|_{\max} &\le n^{2/15}\|A\|_{\max}/\|A\| \le 4n^{-3/30}. \end{aligned} \tag{7.16}$$

Finally, the nonzero blocks in the lower right quadrant,

$$\begin{aligned} \left\| n^{14/30}WQP_kQ^\top \right\|_{\max} &\lesssim n^{14/30}\left\|W\,\mathbb{E}\,QP_kQ^\top\right\|_{\max} + n^{14/30}\left\|WQP_kQ^\top - W\,\mathbb{E}\,QP_kQ^\top\right\|_{\max} \\ &\lesssim kn^{-31/30}\sqrt{\log n} + n^{14/30}\left(\frac{\sqrt{k}\sqrt{\log n}}{n} + \frac{\log n}{n}\right) \\ &\lesssim n^{-1/30}\sqrt{\log n} \\ \left\| n^{11/30}WQP_k\widehat{A} \right\|_{\max} &\le n^{-2/15}\sqrt{\log n} \end{aligned} \tag{7.17}$$

Notice that all entries are strictly less than the pivot, which is 1. $\square$

Lemma 7.9. $\left\|B^{(n+k)}\right\|_{\max} \leq 1$ *for* $0 \leq k < n$ *with probability at least* $1 - 1/n$.

Proof. Start with the nonzero blocks in the off-diagonal part,

$$\begin{aligned} \left\| n^{11/60} E_k^\top W Q^\top \right\|_{\max} &\leq n^{11/60} \left\| W Q^\top \right\|_{\max} \lesssim n^{-19/60} \sqrt{\log n} \\ \left\| n^{11/60} W^\top E_k \right\|_{\max} &\leq n^{11/60} \left\| W \right\|_{\max} \lesssim n^{-19/60} \sqrt{\log n}. \end{aligned} \tag{7.18}$$

Now for the top two blocks in the lower right quadrant,

$$\begin{aligned} \left\| n^{14/30} W \right\|_{\max} &\lesssim n^{-1/30} \sqrt{\log n} \\ \left\| n^{11/30} W Q \widehat{A} \right\|_{\max} &\leq n^{11/30} \left\| W Q \widehat{A} \right\|_{\max} \lesssim n^{-4/30} \sqrt{\log n} \end{aligned} \tag{7.19}$$

Lastly,

$$\left\| n^{11/30} W^\top P_k W Q^\top \right\|_{\max} \leq n^{-2/15} \sqrt{\log n}$$

Notice that all entries are strictly less than the pivot, which is 1. ☐

Lemma 7.10. *If* $A \in \mathrm{CP}_n(\mathbb{R})$, $\|A\|_{\max} \leq 1$, $\|A\| \geq \frac{1}{4} n^{7/30}$, *then* $B^{(2n)} \in \mathrm{CP}_{2n}$ *with probability at least* $1 - 1/n$.

Proof. In order to express the matrix $B^{(2n+k)}$, we first break W and Q into additional blocks. Fix $0 \leq k < n$, and split W and Q after the first k rows and columns,

$$W = \begin{bmatrix} W_{1,:} \\ W_{2,:} \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, \qquad Q = \begin{bmatrix} Q_{1:} \\ Q_{2:} \end{bmatrix},$$

where $W_{11} \in \mathbb{R}^{k \times k}$. We can also write $W_{2:} = E_k^\top W, \qquad Q_{2:} = E_k^\top Q$. Splitting the first block row and column of $B^{(2n)}$ according to this further partitioning gives

$$B^{(2n)} = -\left[\begin{array}{cc|c} n^{14/30} W_{11} & n^{14/30} W_{12} & n^{11/30} W_{1,:} Q \widehat{A} \\ n^{14/30} W_{21} & n^{14/30} W_{22} & n^{11/30} W_{2,:} Q \widehat{A} \\ \hline n^{11/30} Q_{1,:}^\top & n^{11/30} Q_{2,:}^\top & 0 \end{array}\right]. \tag{7.20}$$

Denote the Schur complement

$$S = (W/W_{11}) = W_{22} - W_{21} W_{11}^{-1} W_{12}.$$

Applying the Schur-complement formula directly to (7.20) gives

$$\begin{aligned} B^{(2n+k)} &= \begin{bmatrix} -n^{14/30} W_{22} & -n^{11/30} W_{2,:} Q \widehat{A} \\ -n^{11/30} Q_{2:}^\top & 0 \end{bmatrix} \\ &\quad - \begin{bmatrix} -n^{14/30} W_{21} \\ -n^{11/30} Q_{1:}^\top \end{bmatrix} \left(-n^{14/30} W_{11}\right)^{-1} \begin{bmatrix} -n^{14/30} W_{12} & -n^{11/30} W_{1,:} Q \widehat{A} \end{bmatrix} \\ &= \begin{bmatrix} -n^{14/30} S & -n^{11/30} S Q_{2:} \widehat{A} \\ -n^{11/30} (W_{2,:} Q)^\top S & n^{8/30} Q_{1:}^\top W_{11}^{-1} W_{1,:} Q \widehat{A} \end{bmatrix}. \end{aligned} \tag{7.21}$$

In the last equality, the lower-left block is simplified using $Q_{2:}^\top - Q_{1:}^\top W_{11}^{-1} W_{12} = (W_{2,:} Q)^\top S$, which follows from the block inverse formula and $W^{-1} = W^\top$. Because W is completely pivoted, the upper left entry $|S_{11}|$ is $\|S\|_{\max}$, which is controlled by Lemma 7.5, $\|S\|_{\max} \geq k n^{-10/7}$. We must show the other three blocks have max-norms smaller than

$$n^{14/30} \|S\|_{\max}$$

so that complete pivoting on $B^{(2n+k)}$ uses S_{11}. To this end, note that in addition to $k n^{-10/7}$, we have another lower bound on $\|S\|_{\max}$ from the Frobenius norm: Schur formula for the inverse gives $S^{-1} = W_{22}^\top$ so every singular value of S is at least 1; thus, $\|S\|_{\max} \geq \frac{1}{n-k} \|S\|_F \geq \frac{1}{\sqrt{n-k}}$.

We first control the upper right block. Observe that it is equal to $n^{11/30}(E_kS^\top)^\top Q\widehat{A}$, and $\left\|E_kS^\top\right\|_{1\to2} \le \sqrt{n-k}\ \|S\|_{\max}$, $\|\widehat{A}\|_{1\to2} \le \|\widehat{A}\| \le 1$ so

$$\begin{aligned}\left\|n^{11/30}SQ_{2:}\widehat{A}\right\|_{\max} &\lesssim n^{11/30}\sqrt{n-k}\ \|S\|_{\max}\sqrt{\frac{\log n}{n}} \\ &\le n^{11/30}\sqrt{\log n}\,\|S\|_{\max}\end{aligned} \tag{7.22}$$

Next we control the lower-left block. Observe it is equal to $n^{11/30}Q^\top W^\top E_kS$, and $\|W^\top E_kS\|_{1\to2} = \|S\|_{1\to2} \le \sqrt{n-k}\,\|S\|_{\max}$, so

$$\begin{aligned}\left\|n^{11/30}(W_{2:}Q)^\top S\right\|_{\max} &\lesssim n^{11/30}\sqrt{n-k}\ \|S\|_{\max}\sqrt{\frac{\log n}{n}} \\ &\le n^{11/30}\sqrt{\log n}\,\|S\|_{\max}\end{aligned} \tag{7.23}$$

Finally, the main component of the proof is controlling the lower right block. Set $M_k = I - W^\top E_kSE_k^\top$ and notice $\mathrm{tr}(M_k) = k$. The Schur formula for the inverse means this block is equal to $n^{8/30}Q^\top M_kQ\widehat{A}$. Since $W^\top E_k$ and E_k are isometries, the nonzero singular values of $W^\top E_kSE_k^\top$ and S coincide. Consequently,

$$\|M_k\|_F \le \sqrt{n} + (n-k)\,\|S\|_{\max}\,, \qquad \|M_k\| \le 1 + (n-k)\,\|S\|_{\max}\,.$$

Therefore,

$$\begin{aligned}\left\|n^{8/30}Q^\top M_kQ\widehat{A}\right\|_{\max} &\le n^{8/30}\left\|Q^\top M_kQ\widehat{A} - \mathbb{E}\,Q^\top M_kQ\widehat{A}\right\|_{\max} + n^{8/30}\frac{k}{n}\|\widehat{A}\|_{\max} \\ &\lesssim n^{8/30}\left(\frac{(\sqrt{n}+(n-k)\,\|S\|_{\max})\sqrt{\log n}}{n} + \frac{(1+(n-k)\,\|S\|_{\max})\log n}{n}\right) \\ &\quad + n^{-22/30}k/\|A\|.\end{aligned} \tag{7.24}$$

The lower bounds on $\|S\|_{\max}$ mean that the first term is dominated by $O(n^{8/30}\,\|S\|_{\max}\log n)$. By $\|A\| = n^{7/30}$ and $\|S\|_{\max} \ge kn^{-10/7}$, the second term is bounded by $n^{7/15}\,\|S\|_{\max}\cdot n^{-1/210}$. Both are less than the target $n^{7/15}\,\|S\|_{\max}$ as required.

This means that the pivot for n iterations on $B^{(2n)}$ are all the upper-left entry. After these iterations are performed, the resulting matrix is

$$B^{(3n)} = n^{8/30}Q^\top W^{-1}WQ\widehat{A} = n^{8/30}\widehat{A}.$$

Since $n^{8/30}\widehat{A} \in \mathrm{CP}_n(\mathbb{R})$, we conclude $B^{(2n)} \in \mathrm{CP}_{2n}(\mathbb{R})$. ☐

Theorem 7.11. *For each sufficiently large n, there is a matrix $A \in \mathrm{CP}_n(\mathbb{R})$ with* $\mathrm{growth}(A) \ge \exp\big(\Omega(\log^2(n))\big)$.

Proof. If we obtain the result for geometrically spaced n, then it follows for all n by padding with the identity. Lemmas 7.8 to 7.10 imply that $B(A) \in \mathrm{CP}_{4n}(\mathbb{R})$ with high probability for n sufficiently large if $\|A\|_{\max} = 1$ and $\|A\| \ge \frac{1}{4}n^{7/30}$. Let n_0 be sufficiently large. Set A_0 to be the complete pivoting of $\frac{I+sJ}{1+s}$ for $s = \frac{1}{2}n_0^{-23/30}$, so that $A \in \mathrm{CP}_{n_0}(\mathbb{R})$, $\|A_0\|_{\max} = 1$, and $\frac{1}{4}n_0^{7/30} \le \|A_0\| = \frac{1+n_0s}{1+s} \le 3n_0^{7/30}$. Then set $n_k = 4^kn_0$ and $A_{k+1} = B(A_k)$ conditioned on the positive probability event that $B(A_k) \in \mathrm{CP}_{4n}(\mathbb{R})$. Notice by Lemma 7.7 that one maintains $\frac{1}{4}n_k^{7/30} \le \|A_k\| \le 3n_k^{7/30}$ and by Lemma 7.8 $\|A_k\|_{\max} \le 1$. Then

$$\mathrm{growth}(A_{k+1}) \ge \frac{n_k^{4/15}}{\|A\|}\,\mathrm{growth}(A_k) \ge \frac{1}{3}n_k^{1/30}\,\mathrm{growth}(A_k).$$

Repeatedly apply this, we obtain

$$\mathrm{growth}(A_{k+1}) \ge 3^{-(k+1)}(4^kn_0)^{1/30}(4^{k-1}n_0)^{1/30}\cdots(n_0)^{1/30} \ge 3^{-(k+1)}4^{\frac{k(k+1)}{60}}n_0^{\frac{k+1}{30}} = e^{\Omega((k+1)^2)}.$$

Since $k = \frac{\log(n_k/n_0)}{\log 4} = \Theta(\log(n_k))$, we obtain the desired result. ☐

Remark 7.12. One can construct A achieving the bound of Theorem 7.11 so that it is the unique element in $\{PAQ\} \cap \mathrm{CP}_n(\mathbb{R})$, i.e., there is no freedom in implementing complete pivoting. To do this, just multiply on both sides by $\mathrm{diag}(1, e^{-\varepsilon}, e^{-2\varepsilon}, \ldots)$ for sufficiently small $\varepsilon > 0$.

Acknowledgements

The authors would like to thank Louisa Thomas for improving the style of presentation. The second author would like to thank Alan Edelman for years' worth of conversation and collaboration on the growth factor problem in Gaussian elimination, as well as illuminating conversations with Alex Townsend, Nick Trefethen, and Nick Higham on the subject. This material is based upon work supported by the Institute for Advanced Study and the National Science Foundation under Grant No. DMS-1926686 – the second author is grateful to the IAS for their membership for the 2021-2022 year, where a number of preliminary results and ideas which inspired this final manuscript were developed. This material is based upon work supported by the National Science Foundation under grant no. DMS-2513687. Some of the research presented here is the result of human-AI interaction, see the AI usage statement below.

AI Usage Statement

This manuscript was created with the assistance of ChatGPT Sol 5.6 and Claude Fable 5. The abstract and introduction are written without AI input, and every mathematical result and proof is written by the authors. This manuscript blends purely human based mathematics and ideas with some mathematics that is the result of human-AI interaction. We aim to give a clear and detailed picture of this, section by section:

- The results of Section 2 began with human only results regarding sparse matrices and partial pivoting proven before the advent of either ChatGPT or Claude. Interaction with AI significantly improved the human only results.
- The results of Section 3, were proven with the use of AI. In particular, AI provided explanations of determinantal point processes, examples of stochastic domination, and questions regarding the strong Rayleigh property. The idea of using a projection DPP was human-generated. The use of a Sylvester Hadamard matrix and general Hadamard matrices was a human-generated construction, the proof that the choice of row permutation does not affect the Frobenius norm of L and U was the result of human-AI interaction.
- The results of Section 4, were the result of human-AI collaboration. The authors had a proof of NP-hardness for a slightly different definition of growth factor which was proven before the advent of either ChatGPT or Claude. Interaction with AI lead to a hardness proof for $\mathrm{growth}(\cdot)$, which does not appear. Inspired by this proof, which shared many similarities with our original human-only proof, the authors proved, without AI, save for the aforementioned use, the current construction that appears in the manuscript.
- The results of Section 5 were the result of human-AI collaboration. The AI was given the measure and prompeted to compute expectations using the determinant formulas for L and U. These were unsatisfactory, and, after prompting for a proof of boundedness, a Jordan block was suggested. Human-only numerical simulations were performed, which suggested that the expectation was tail dominated. This was confirmed by AI, and the authors switched to considering high probability bounds instead, still with a Jordan block. The authors did human-only calculations and mistakenly thought taking fixed z near 1 worked. Human-AI interaction led to the suggestion of $z = 1 - 1/\sqrt{n}$, a key stepping stone to the construction appearing here. All of the above was specifically for $p = 2$, which was generalized for without further use of human-AI interaction.
- The results of Section 6 were proven without the use of AI, and AI was not involved in any way. The earliest proof that the growth factor under rook pivoting was quasi-polynomial was completed in 2022. The current lower bound is an improvement over that original construction, though all improvements were made exclusively without the use of AI.
- The results of Section 7 are the result of human-AI interaction, but strongly followed a research program/architecture for proving that complete pivoting has quasi-polynomial growth that has been an active area of research since 2022 as a human-only endeavor, up to only a week before the first draft of this manuscript. Based on our experience, the current AI systems, without additional expert advice and a clear roadmap for the research program, was unable to prove such a result on its own. The resulting lower bound does follow the human roadmap which the AI systems were provided, and which was in development prior to the existence of ChatGPT or Claude. However, the authors have to remark that these tools are remarkably effective, and this research programme for the growth

factor under complete pivoting would not have been completable in such a short time period without the involvement of AI.

The authors welcome questions regarding the role of AI in this work and comments regarding the best methodology for disclosing human-AI interactions, as no universal norms seem to currently exist yet.

References

[B$^+$] Keith Ball et al. An elementary introduction to modern convex geometry.

[BEU25] Ankit Bisain, Alan Edelman, and John Urschel. A new upper bound for the growth factor in gaussian elimination with complete pivoting. *Bulletin of the London Mathematical Society*, 57(5):1369–1387, February 2025.

[BKS13] Matthijs J. Bomhoff, Walter Kern, and Georg J. Still. A note on perfect partial elimination. *Discrete Mathematics*, 313(14):1558–1563, 2013.

[CD05] Zizhong Chen and Jack J. Dongarra. Condition numbers of gaussian random matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):603–620, January 2005.

[CETW25] Yifan Chen, Ethan N. Epperly, Joel A. Tropp, and Robert J. Webber. Randomly pivoted cholesky: Practical approximation of a kernel matrix with few entry evaluations. *Communications on Pure and Applied Mathematics*, 78(5):995–1041, 2025.

[ÇMI09] Ali Çivril and Malik Magdon-Ismail. On selecting a maximum volume sub-matrix of a matrix and related problems. *Theoretical Computer Science*, 410(47–49):4801–4811, 2009.

[CP07] Viviana Cortés and Juan Manuel Peña. Growth factor and expected growth factor of some pivoting strategies. *Journal of Computational and Applied Mathematics*, 202(2):292–303, 2007.

[DGTY24] Anil Damle, Silke Glas, Alex Townsend, and Annan Yu. How to reveal the rank of a matrix?, 2024.

[DR10] Amit Deshpande and Luis Rademacher. Efficient volume sampling for row/column subset selection. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science*, pages 329–338. IEEE, 2010.

[DRVW06] Amit Deshpande, Luis Rademacher, Santosh Vempala, and Grant Wang. Matrix approximation and projective clustering via volume sampling. *Theory of Computing*, 2(12):225–247, 2006.

[Ede92] Alan Edelman. The complete pivoting conjecture for Gaussian elimination is false. *The Mathematica Journal*, 2(2):58–61, 1992.

[EU24] Alan Edelman and John Urschel. Some new results on the maximum growth factor in gaussian elimination. *SIAM Journal on Matrix Analysis and Applications*, 45(2):967–991, 2024.

[EUZ24] Alan Edelman, John Urschel, and Bowen Zhu. On a perturbation analysis of higham squared maximum gaussian elimination growth matrices. *arXiv preprint arXiv:2406.00737*, 2024.

[Fed25] Semyon Sergeevich Fedorovskii. Nonlinear growth of the l_∞-norm of matrices for maximal cross approximation. *Sbornik: Mathematics*, 216(10):1483–1492, 2025.

[Fos94] Leslie V Foster. Gaussian elimination with partial pivoting can fail in practice. *SIAM Journal on Matrix Analysis and Applications*, 15(4):1354–1362, 1994.

[Fos97] Leslie V. Foster. The growth factor and efficiency of gaussian elimination with rook pivoting. *Journal of Computational and Applied Mathematics*, 86(1):177–194, 1997.

[GE96] Ming Gu and Stanley C. Eisenstat. Efficient algorithms for computing a strong rank-revealing qr factorization. *SIAM Journal on Scientific Computing*, 17(4):848–869, 1996.

[Gol93] Herman H Goldstine. *The computer from Pascal to von Neumann*. Princeton University Press, 1993.

[Gou91] Nick Gould. On growth in Gaussian elimination with complete pivoting. *SIAM Journal on Matrix Analysis and Applications*, 12(2):354–361, 1991.

[GS23] Friedrich Götze and Holger Sambale. Higher order concentration on Stiefel and Grassmann manifolds. *Electronic Journal of Probability*, 28:1–30, 2023. Paper No. 79.

[GVL13] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU press, 2013.

[GW26] Marc Aurèle Gilles and Heather Wilber. Low-rank approximation by randomly pivoted LU, 2026.

[HH89] Nicholas J. Higham and Desmond J. Higham. Large growth factors in gaussian elimination with pivoting. *SIAM Journal on Matrix Analysis and Applications*, 10(2):155–164, 1989.

[Hig02] Nicholas J Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.

[HT24] Han Huang and Konstantin Tikhomirov. Average-case analysis of the gaussian elimination with partial pivoting. *Probability Theory and Related Fields*, 189:501–567, 2024.

[KT12] Alex Kulesza and Ben Taskar. *Determinantal Point Processes for Machine Learning*, volume 5 of *Foundations and Trends in Machine Learning*. Now Publishers, 2012.

[LTTJ20] Alexander E Litvak, Konstantin Tikhomirov, and Nicole Tomczak-Jaegermann. Small ball probability for the condition number of random matrices. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2017-2019 Volume II*, pages 125–137. Springer, 2020.

[Lyo03] Russell Lyons. Determinantal probability measures. *Publications Mathématiques de l'IHÉS*, 98:167–212, December 2003.

[MG15] Christopher Melgaard and Ming Gu. Gaussian elimination with randomized complete pivoting, 2015.

[Pan00] Chan-Tong Pan. On the existence and computation of rank-revealing lu factorizations. *Linear Algebra and its Applications*, 316(1–3):199–222, 2000.

[PN00] George Poole and Larry Neal. The rook's pivoting strategy. *Journal of Computational and Applied Mathematics*, 123(1-2):353–369, 2000.

[SG20] Lukas Schork and Jacek Gondzio. Rank revealing gaussian elimination by the maximum volume concept. *Linear Algebra and its Applications*, 592:1–19, 2020.

[Sha22] Rikhav Shah. Determinants of binary matrices achieve every integral value up to $\Omega(2/\mathrm{n})$. *Linear Algebra and its Applications*, 645:229–236, 2022.

[SST06] Arvind Sankar, Daniel A Spielman, and Shang-Hua Teng. Smoothed analysis of the condition numbers and growth factors of matrices. *SIAM Journal on Matrix Analysis and Applications*, 28(2):446–476, 2006.

[Str65] V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, April 1965.

[TB22] Lloyd N Trefethen and David Bau. *Numerical linear algebra*. SIAM, 2022.

[TS90] Lloyd N. Trefethen and Robert S. Schreiber. Average-case stability of gaussian elimination. *SIAM Journal on Matrix Analysis and Applications*, 11(3):335–360, 1990.

[Urs25] John Urschel. Numerical stability in gaussian elimination. *Notices of the American Mathematical Society*, 72(6):605–613, 2025.

[Vav89] Stephen A. Vavasis. Gaussian elimination with pivoting is P-complete. *SIAM Journal on Discrete Mathematics*, 2(3):413–423, 1989.

[vNG47] John von Neumann and Herman H Goldstine. Numerical inverting of matrices of high order. 1947.

[Wil61] James H. Wilkinson. Error analysis of direct methods of matrix inversion. *Journal of the ACM*, 8(3):281–330, 1961.

[Wil65] James H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.

[Wri93] Stephen J Wright. A collection of problems for which gaussian elimination with partial pivoting is unstable. *SIAM Journal on Scientific Computing*, 14(1):231–238, 1993.

[Yan81] Mihalis Yannakakis. Computing the minimum fill-in is NP-complete. *SIAM Journal on Algebraic and Discrete Methods*, 2(1):77–79, 1981.