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-hard.

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 with non-zero leading principal minors is factored into the product of a lower triangular matrix L and an upper triangular matrix U by a sequence of rank-one transformations, resulting in the sequence of matrices $A^{(0)} = A$ and $A^{(k)} = \left(a_{ij}^{(k)}\right)_{i,j=1}^{n-k}$ for k equals 1 to $n-1$, satisfying

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

The sequence $A^{(0)}, \ldots, A^{(n)}$ 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, under the normalization that L is 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

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

where $\|\cdot\|_{\max}$ is the entrywise maximum norm (if no LU factorization exists, set $\text{growth}(A) = +\infty$). 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 $\log(\text{growth}(A))$, 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 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]. 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, which "became one of the most famous open problems in numerical analysis" [Hig02, pg. 181]. This conjecture was later 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 set of matrices with growth 2^{n-1} [HH89]. An exponentially large growth factor can be a prohibitive barrier to accurate computation in floating point arithmetic. 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 concerning partial pivoting regards whether large growth can occur in practice [TS90, DM07, Tre12, HT24], and to what extent large growth is unstable (i.e., its possible resolution by the addition of a small random perturbation) [San04]. We refer the reader to [TB22, Urs25] for more details.

Due to the exponential 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, and 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. Also like complete pivoting, there is a substantial gap between the upper and lower bounds, with the largest known growth factor for rook pivoting being only of order $n^{1.669}$, proven by Edelman and Urschel [EU24]. 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 that 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

$$\left|a_{11}^{(k)}\right| \geq \max_{i,j\in[n-k]} \left|a_{ij}^{(k)}\right| \qquad \text{for} \quad k = 0, \ldots, n-1,$$

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

$$\left|a_{11}^{(k)}\right| \geq \max_{j\in[n-k]} \max\left\{\left|a_{1j}^{(k)}\right|, \left|a_{j1}^{(k)}\right|\right\} \qquad \text{for} \quad k = 0, \ldots, n-1,$$

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

$$\left|a_{11}^{(k)}\right| \geq \max_{j\in[n-k]} \left|a_{j1}^{(k)}\right| \qquad \text{for} \quad k = 0, \ldots, n-1,$$

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, for $p > 0$, which is fully defined

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

It is not obvious from the definition of Vol that a distribution with those prescribed marginals exists; 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, which is tight. Matrices that are locally sparse, i.e., with at most k non-zero entries per row and column, can also have exponentially large growth under partial pivoting. For example, there exist matrices with at most three non-zero entries 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-zero entries 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 possible value of $\mathrm{nnz}(A)$ for any matrix in $\mathrm{PP}_n(\mathbb{R})$ with growth 2^{n-1}. In addition, for all integers $2 \leq k < n$, 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 $n \in \mathbb{N}$ 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 $(A, g) \in \mathrm{Mat}_n(\mathbb{Z}) \times \frac{1}{2}\mathbb{Z}_+$ 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 a 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 $A \in \mathrm{RP}_n(\mathbb{R})$ with $\mathrm{growth}(A) \geq n^{(\frac{1}{4} - o(1)) \log_2 n}$.*

Theorem 1.7 (Theorem 7.12 / Quasi-polynomial growth of complete pivoting)**.** *For every n, there is a $A \in \mathrm{CP}_n(\mathbb{R})$ with $\mathrm{growth}(A) \geq n^{(\frac{1}{16} - o(1)) \log_2 n}$.*

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. 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 [Wil65]. 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-zero entries 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. 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 $T_{11} \neq 0$, 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. 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 $n = 6$, the matrices associated with $k = 2$ and $k = 3$, respectively, in Theorem 2.3 are

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

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 in magnitude 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 let 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 $A \in \text{PP}_n(\mathbb{R})$ with at most two non-zero entries per row and column. We prove the following claim by induction:

> The matrix $A^{(k)}$ 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, $A^{(k+1)}$ 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, $\left|a_{1j}^{(k)}\right| \le \|A\|_{\max}$, and, by $A \in \text{PP}_n(\mathbb{R})$, $\left|a_{i1}^{(k)}\right| \le \left|a_{11}^{(k)}\right|$. If $a_{ij}^{(k)} = 0$, then $\left|a_{i-1,j-1}^{(k+1)}\right| \le \|A\|_{\max}$, and there are still 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

$$\left|a_{i-1,j-1}^{(k+1)}\right| \le \left|a_{ij}^{(k)}\right| + \left|\frac{a_{i1}^{(k)} a_{1j}^{(k)}}{a_{11}^{(k)}}\right| \le 2\|A\|_{\max}.$$

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 $A \in \text{PP}_n(\mathbb{R})$ 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 $\|L\|_{\max} = 1$. 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, complete pivoting requires row and column exchanges, and still does not produce a polynomial growth factor. Here, we consider a fundamental existence question: does there always exist a row permutation that produces polynomial growth? By analyzing the randomized pivoting strategy $P \sim \text{Vol}(A)$, we answer this question affirmatively.

This question is closely related to rank-revealing factorizations and maximum volume pivoting in numerical linear algebra (see [GE96, Pan00, SG20, DGTY24]). A $k \times k$ submatrix 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:

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

But 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, but these choices need not be nested as k varies; 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 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 or a fixed size determinantal point process (DPP) [KT12]. Volume sampling has previously appeared in a variety of applied linear algebraic contexts, see, for instance, [DRVW06, DR10, DW18, DM21, BBC20].

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

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

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

Theorem 3.1. *For every $A \in \text{GL}_n(\mathbb{R})$, there is a probability measure on permutations $\text{Vol}(A)$ with marginals (1.2), where if $P \sim \text{Vol}(A)$ then $PA = LU$ satisfies*

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

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

Proof. The existence of the distribution 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 bounded by Lemma 3.3, so we simply sum these expressions,

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

□

Lemma 3.2. *Given $A \in \text{GL}_n(\mathbb{R})$, 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 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

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

Proof. The equivalence of the formulas 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 element in $S_i \backslash S_{i-1}$ to obtain the desired distribution $\textsc{Vol}(A)$ on permutations. □

Lemma 3.3. *Let $A \in \mathrm{GL}_n(\mathbb{R})$ and 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 $(\pi_1, \dots, \pi_n)$ be the permutation from Lemma 3.2 and set $S = \{\pi_1, \dots, \pi_k\}$, $S' = \{\pi_1, \dots, \pi_{k+1}\}$. By Cramer's rule,

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

We first handle L. We have

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

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

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

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

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

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

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

where the outer sum is over $S \in \binom{[n]}{k}$. Note that $S + \{\pi'\}$ 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 $B = A^\top A$,

$$\mathbb{E}\,\|U[k+1,:]\|^2 \le (k+1) \sum_{j=k+1}^{n} \frac{\det B[C+\{j\}, C+\{j\}]}{\det B[C,C]} \le (k+1) \sum_{j=k+1}^{n} B_{j,j},$$

where the final step used the bound $\det B[C+\{j\}, C+\{j\}] \le \det(B[C,C]) \cdot B_{j,j}$, 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. Whenever an $n \times n$ Hadamard matrix exists, it gives a lower bound of n for the growth factor under any row or column 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 exactly the 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 $AA^\top = 2^k I$, so $UU^\top = 2^k L^{-1} L^{-T}$ 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 consist of the n pairs $(\boldsymbol{e}_i^\top H_{k-1}, \boldsymbol{e}_i^\top H_{k-1})$ and $(\boldsymbol{e}_i^\top H_{k-1}, -\boldsymbol{e}_i^\top H_{k-1})$ for $i \in [n]$. The first n rows of A consist of precisely one row from each of these pairs and possibly re-signs them, and the last n rows consist 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 $\widetilde{H}$ and $\widehat{H}$ 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 $Q\widetilde{D}Q^\top = -\widehat{D}$, 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

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

□

When both rows and columns are reordered, 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 a 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, $\nabla f(\boldsymbol{x}^*)^\top(\boldsymbol{y}-\boldsymbol{x}^*) \geq 0$ and $\nabla g(\boldsymbol{x}^*)^\top(\boldsymbol{y}-\boldsymbol{x}^*) \geq 0$ 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 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 unique 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 $(M_{\boldsymbol{b}})_{ij} = [[(x_i = \boldsymbol{b}_i) \implies c_j]]$ and define

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

Notice that A_f has dimension $2n+m+2$ and m and can be constructed from f in polynomial time.

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

Proof. Suppose that f is satisfiable. We construct a permutation P such that 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 swaps rows $1+i$ and $1+i+n$ if and only if $\boldsymbol{b}_i = 1$. Then

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

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

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

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

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

We have $\|L\|_{\max} = 1$ and $\|U\|_{\max} = \max\{1, \left\|\frac{3}{2}\mathbf{1} - \frac{1}{2}M_{\boldsymbol{b}}^\top \mathbf{1}\right\|_\infty\}$. 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 $\operatorname{growth}(\widetilde{P}_{\boldsymbol{b}} A_f) = 1$, as desired.

Conversely, suppose that there exists a permutation P with $\operatorname{growth}(PA_f) < \frac{3}{2}$. 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 $\boldsymbol{b}_i = 0$ or $\boldsymbol{b}_i = 1$. 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 $\operatorname{growth}(PA_f) < \frac{3}{2}$, we must have $\left\|\frac{3}{2}\mathbf{1} - \frac{1}{2}M_{\boldsymbol{b}}^\top \mathbf{1}\right\|_\infty < \frac{3}{2}$. In particular, every entry of $M_{\boldsymbol{b}}^\top \mathbf{1}$ must be positive, which means that $\boldsymbol{b}$ is a satisfying assignment of f. □

Theorem 4.2 immediately implies our desired Theorem 1.4. Scaling A_f by a factor of two places all entries in $\mathbb{Z}$ and preserves the growth factor of it and all its permutations. Additionally, 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 is non-singular, 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 performed in polynomial time, and its bit length is bounded. In particular, if the entries of $A \in \operatorname{Mat}_n(\mathbb{R})$ have bit length at most ℓ, then, by Hadamard's inequality, every minor of A has bit length $O(n(\ell + \log n))$. 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, 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 and $p \in (0, \infty)$. 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,

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

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 analogue to $\textsc{RandPP}_p$ is to select both a row and column index, in proportion to a power of the entry, $\propto |a^{(k)}_{ij}|^p$; the case $p = 2$ was introduced in [GW26]. In addition, there is a more specialized analogue 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, $\propto |a^{(k)}_{ii}|^p$ [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 under 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

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

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 contained in $\{\pi_1, \ldots, \pi_k\} \subset [k + 1]$. The set $\{\pi_1, \ldots, \pi_k\}$ has exactly one element missing from $[k+1]$. For 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 top left $(m_{k+1}-1)\times(m_{k+1}-1)$ block is lower triangular with diagonal entries equal to z, and the lower right $(k-m_{k+1}+1)\times(k-m_{k+1}+1)$ 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 are zero. Note 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:

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

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

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

Stated differently, each excursion δ_k makes from a given 0 value has length at most n^α with probability at least $1-e^{-n^\alpha/(e+1)}$. 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 $1 - ne^{-n^\alpha/(e+1)}$, 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 $1 - ne^{-n^\alpha/(e+1)}$. □

Because PQ is orthogonal, $\|PQ\|_{\max} \leq 1$, and so $\text{growth}(PQ) \geq \|U\|_{\max}$.

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: for each matrix A exhibiting large growth under rook-pivoting, we construct a larger rook-pivoted matrix B with smaller entries which reveals A in the course of performing eliminations. The key strategy for constructing such a gadget is focusing on the block-nonzero pattern

$$B = \begin{bmatrix} W & X \\ Y & 0 \end{bmatrix}.$$

This way, A corresponds to the factorization $A = YW^{-1}X$. We take W to be a multiple of the identity matrix, so that the behavior under successive elimination steps is easy to determine explicitly. We take X to be a random matrix, so that its entries (and, when $W = \alpha I$, the entries of $Y = \alpha AX^{-1}$ as well) are small via a concentration argument, thereby ensuring B is rook pivoted.

Our construction in the next section uses a similar idea, and so we present the concentration results needed for this section and the next here. Both concern properties of Haar-distributed matrices, though this section only makes use of the first bound (6.1) for $Y = I$.

Lemma 6.1. *Fix any $X \in \mathbb{R}^{n\times p}$, $Y \in \mathbb{R}^{n\times q}$, and $M \in \mathbb{R}^{n\times n}$. Let $\delta \in (0,1)$ and $L = \log(2pq/\delta)$. If $Q \sim \text{Haar}(\text{O}(n))$, then*

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

and

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

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 vector uniformly distributed on the unit sphere $\mathcal{S}^{n-1}$, which, by 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 $\gamma = \sqrt{2L}$ 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 $\boldsymbol{x}$ and $\boldsymbol{y}$ in an orthonormal basis for their span, e.g., $\boldsymbol{x} = E\boldsymbol{a}$ and $\boldsymbol{y} = E\boldsymbol{b}$ for $E^\top E = I$ and $\text{Col}(E) = \text{span}\{\boldsymbol{x}, \boldsymbol{y}\}$. Then $Z = QE$ is a Haar-distributed sample from the Stiefel manifold of order at most two, and

$$\boldsymbol{x}^\top Q^\top MQ\boldsymbol{y} = \boldsymbol{a}^\top Z^\top MZ\boldsymbol{b} = \text{vec}(Z)^\top \frac{\boldsymbol{a}\boldsymbol{b}^\top \otimes M + \boldsymbol{b}\boldsymbol{a}^\top \otimes M^\top}{2}\text{vec}(Z).$$

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

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

for $c = 7/10000$. Observe that $\operatorname{tr}(\widetilde{M}) = \boldsymbol{x}^\top \boldsymbol{y} \operatorname{tr}(M)$, $\|\widetilde{M}\|_F \leq \|\|x\|\|y\|M\|_F$, 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 a 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 $\operatorname{growth}(A) \geq n^{\frac{1}{4}\log_2(n) - C\log\log n}$.*

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 $Q_k \sim \operatorname{Haar}(\mathrm{O}(2^k))$ 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 $X_k \in \operatorname{Mat}_{2^k}(\mathbb{R})$ 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

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

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 applying 2^k steps of Gaussian elimination 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

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

which establishes our growth claim. We now turn our attention to pivoting. Let $\mathcal{E}_k$ be the event that $(\|X_k Q_k\|_{\max} \leq s_k$ and $\|Q_k\|_{\max} \leq s_k)$. 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 uses the trivial $\max(s_j : j < k) \geq 1$. By a union bound, the probability of either event occurring 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 $X_{K+1} \in \mathrm{RP}_{2^{K+1}}(\mathbb{R}) = \mathrm{RP}_n(\mathbb{R})$ 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. Our lower bound has exponent $\frac{1}{16} \log_2 n$, which is nearly within a factor of two of the upper bound $0.207 \log n \approx 0.141 \log_2 n$. Our construction is similar in spirit to that of Theorem 6.2, with a few keys differences. For complete pivoting, we must bound the magnitude of all entries at each step, which requires a top-left block with sufficiently large pivot sizes (Lemma 7.5). In addition, to build a large growth factor, we require control on the norm of the input matrix, and, therefore, control on the norm of the output matrix as well. This is achieved by a block 4×4 structure (Definition 7.6) whose norm is well-controlled (Lemma 7.7).

First, we argue the existence of an orthogonal matrix, determined in Lemma 7.5, with pivots of sufficiently large magnitude under complete pivoting via a probabilistic argument. Lemmas 7.1, 7.2 and 7.4 are preparation for this. Then, given this matrix, we define our recursive structure, Definition 7.6.

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

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

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

Lemma 7.2 ([CD05, Lemma 4.1])**.** *Let $n \geq m \geq 2$. If G is an $n \times m$ Gaussian matrix, then*

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

for all $t > 0$.

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

$$\big\|X^{-1}\big\|_F^2 \geq cm\sqrt{n}$$

for an absolute constant c.

Proof. We first 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),

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

Applying Lemma 7.1 in dimension $m \geq \sqrt{n}$, with $t = C\sqrt{n}$ for a large absolute constant C, gives

$$\Pr\left(\big\|G_1^{-1}\big\|_F^2 \leq \frac{\sqrt{m}}{Cn^{1/4}}\right) \leq 2\exp\big(-c_0 C^2 n\big), \tag{7.2}$$

where c_0 is the constant from Lemma 7.1. On the other hand, Lemma 7.2 applied to $\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}$ gives

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

where the last inequality used $m \leq \frac{n}{2}$. Set $t = \frac{1}{100}$. It follows from (7.1) to (7.3) that

$$\Pr\left(\big\|X^{-1}\big\|_F^2 < \frac{\sqrt{n}m}{100^2 C^2}\right) \leq 2\exp\big(-c_0 C^2 n\big) + \left(\frac{2e}{100}\right)^{n/2}.$$

Now, choose C such that $c_0 C^2 > \log 4 + 1$. The total number of submatrices of an $n \times n$ matrix is at most 4^n, so a union bound provides our desired result. □

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

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

Moreover,

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

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 $(n-m)^2 \|(W/X)\|_{\max}^2 \geq \|(W/X)\|_F^2 \geq n-m$, establishing (7.5).

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

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 Gaussian elimination, then

$$\left\|W^{(j)}\right\|_{\max} \geq c\frac{j}{n^{5/4}} \qquad \text{for} \quad j = 0, \ldots, n-1. \tag{7.7}$$

Proof. Let $\widehat{W} \sim \mathrm{Haar}(\mathrm{O}(n))$. By Lemma 6.1, (7.6) holds for $\widehat{W}$ with high probability. By Lemma 7.3, with probability at least $1 - e^{-cn}$, every square submatrix Y with order in the interval $[\sqrt{n}, n/2]$ satisfies

$$\|Y^{-1}\|_F \geq cn^{1/4}\sqrt{\dim Y}. \tag{7.8}$$

Now, consider a realization $\widehat{W}$ that satisfies both these properties, and let P and Q be permutation matrices such that $W = P\widehat{W}Q \in \mathrm{CP}_n(\mathbb{R})$. Note that $W \in \mathrm{O}(n)$ and the above bounds still hold for $\|W\|_{\max}$ and square submatrices of W with order in $[\sqrt{n}, n/2]$.

We break our analysis of $W^{(j)}$ into three regimes, depending on the index j. If $\sqrt{n} \leq j \leq n - \sqrt{n}$, then, by (7.8) and (7.4), there exists a square submatrix Y of order $\min\{j, n-j\}$ with

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{\left\|W^{(j)}\right\|_F}{n-j} \geq \frac{\|Y\|_F}{n-j} \geq \frac{cn^{1/4}\sqrt{\min(j, n-j)}}{n-j} \geq \frac{cj}{n^{5/4}}.$$

If $j < \sqrt{n}$, then, by (7.5),

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{1}{\sqrt{n-j}} \geq \frac{1}{\sqrt{n}} \geq \frac{j}{n^{5/4}}.$$

Finally, if $j > n - \sqrt{n}$, then, again, by (7.5),

$$\left\|W^{(j)}\right\|_{\max} \geq \frac{1}{\sqrt{n-j}} > \frac{1}{n^{1/4}} \geq \frac{j}{n^{5/4}}.$$

□

Definition 7.6. For $A \in \mathrm{GL}_n(\mathbb{R})$, let

$$\widehat{A} = \frac{A}{\|A\|}, \qquad \alpha = \frac{n^{1/4}}{\log n}, \qquad \beta = \frac{n^{1/4}}{\log^{3/2} n}, \qquad \gamma = \frac{n^{1/4}}{\log^2 n},$$

and note that $\alpha\gamma = \beta^2$. Let $B(A)$ denote the $4n \times 4n$ random matrix

$$B(A) = \left[\begin{array}{cc|cc} I & 0 & \alpha Q^\top & \gamma\widehat{A} \\ 0 & I & \beta W Q^\top & 0 \\ \hline \alpha W Q & 0 & 0 & 0 \\ 0 & \beta W^\top & 0 & 0 \end{array}\right],$$

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

It is useful to first record the residuals of $B(A)$ after a certain number of steps of Gaussian elimination. Let $E_k \in \mathbb{R}^{n\times(n-k)}$ be the trailing $n-k$ columns of the identity matrix and $P_k = I - E_k E_k^\top$ be the projection matrix onto the first k coordinates. By (1.1),

$$B(A)^{(k)} = \left[\begin{array}{cc|cc} I_{n-k} & 0 & \alpha E_k^\top Q^\top & \gamma E_k^\top \widehat{A} \\ 0 & I_n & \beta W Q^\top & 0 \\ \hline \alpha W Q E_k & 0 & -\alpha^2 W Q P_k Q^\top & -\beta^2 W Q P_k \widehat{A} \\ 0 & \beta W^\top & 0 & 0 \end{array}\right] \qquad \text{for} \quad 0 \le k < n. \tag{7.9}$$

The $(4,3)$ block becomes non-zero only after the first identity block is completely eliminated,

$$B(A)^{(n+k)} = \left[\begin{array}{c|cc} I_{n-k} & \beta E_k^\top W Q^\top & 0 \\ \hline 0 & -\alpha^2 W & -\beta^2 W Q \widehat{A} \\ \beta W^\top E_k & -\beta^2 W^\top P_k W Q^\top & 0 \end{array}\right] \qquad \text{for} \quad 0 \le k < n. \tag{7.10}$$

After the second identity block is eliminated,

$$B(A)^{(2n)} = -\begin{bmatrix} \alpha^2 W & \beta^2 W Q \widehat{A} \\ \beta^2 Q^\top & 0 \end{bmatrix}. \tag{7.11}$$

In both cases, we benefit from the cancellations $QQ^\top = I$ and $W^\top W = I$ once the projection P_k becomes identity. After the remaining n steps of Gaussian elimination, we are left with a single block

$$B(A)^{(3n)} = \gamma^2 \widehat{A}. \tag{7.12}$$

In order for $B(A) \in \mathrm{CP}_{4n}(\mathbb{R})$, we must have $B(A)^{(2n)} \in \mathrm{CP}_{2n}(\mathbb{R})$, $\left\|B(A)^{(k)}\right\|_{\max} = 1$ for $k = 0, \dots, n-1$, and $\left\|B(A)^{(n+k)}\right\|_{\max} = 1$ for $k = 0, \dots, n-1$. Conditional on some properties of A, we show that, with high probability, this is the case.

Lemma 7.7. *For n sufficiently large, $\alpha \le \|B(A)\| \le 3\alpha$ surely.*

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

$$\alpha = \|\alpha W Q\| \le \|B(A)\| \le \left\| \begin{pmatrix} 1 & 0 & \alpha & \gamma \\ 0 & 1 & \beta & 0 \\ \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \end{pmatrix} \right\|_F = \left(2 + 2\alpha^2 + 2\beta^2 + \gamma^2\right)^{1/2} \le 3\alpha.$$

□

In the following three lemmas, we implicitly invoke Lemma 6.1, (6.1) and (6.2), order n 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 $\|A\| \geq \frac{\alpha}{4}$ and $\|A\|_{\max} \leq 1$, then $\left\|B(A)^{(k)}\right\|_{\max} = 1$ for $0 \leq k < n$ with probability at least $1 - o(1)$.*

Proof. We start with the non-zero blocks in the lower-left quadrant,

$$\begin{aligned} \left\|\alpha W Q E_k\right\|_{\max} &\leq \alpha \left\|WQ\right\|_{\max} \lesssim \alpha \sqrt{\frac{\log n}{n}} = o(1), \\ \left\|\beta W^\top\right\|_{\max} &\leq \beta \left\|W\right\|_{\max} \lesssim \beta \sqrt{\frac{\log n}{n}} = o(1). \end{aligned} \tag{7.13}$$

Next, consider the non-zero blocks in the top-right quadrant,

$$\begin{aligned} \left\|\alpha E_k^\top Q^\top\right\|_{\max} &\leq \alpha \left\|Q\right\|_{\max} \lesssim \alpha \sqrt{\frac{\log n}{n}} = o(1), \\ \left\|\gamma E_k^\top \widehat{A}\right\|_{\max} &\leq \frac{\gamma \left\|A\right\|_{\max}}{\|A\|} \leq \frac{4\gamma}{\alpha} = o(1). \end{aligned} \tag{7.14}$$

Finally, consider the non-zero blocks in the lower-right quadrant,

$$\begin{aligned} \left\|\alpha^2 W Q P_k Q^\top\right\|_{\max} &\lesssim \alpha^2 \left\|W \,\mathbb{E}\, Q P_k Q^\top\right\|_{\max} + \alpha^2 \left\|WQP_kQ^\top - W\,\mathbb{E}\,QP_kQ^\top\right\|_{\max} \\ &\lesssim \frac{\alpha^2 k}{n} \sqrt{\frac{\log n}{n}} + \alpha^2 \left(\frac{\sqrt{k}\sqrt{\log n}}{n} + \frac{\log n}{n} \right) = o(1), \\ \left\|\beta^2 W Q P_k \widehat{A}\right\|_{\max} &\leq \beta^2 \sqrt{\frac{\log n}{n}} = o(1). \end{aligned} \tag{7.15}$$

Note that all entries are smaller in magnitude than the pivot, which is 1. □

Lemma 7.9. *If $\|A\| \geq \frac{\alpha}{4}$ and $\|A\|_{\max} \leq 1$, then $\left\|B(A)^{(n+k)}\right\|_{\max} = 1$ for $0 \leq k < n$ with probability at least $1 - o(1)$.*

Proof. We start with the non-zero blocks in the off-diagonal part,

$$\begin{aligned} \left\|\beta E_k^\top W Q^\top\right\|_{\max} &\leq \beta \left\|WQ^\top\right\|_{\max} \lesssim \beta \sqrt{\frac{\log n}{n}} = o(1), \\ \left\|\beta W^\top E_k\right\|_{\max} &\leq \beta \left\|W\right\|_{\max} \lesssim \beta \sqrt{\frac{\log n}{n}} = o(1). \end{aligned} \tag{7.16}$$

Next, consider the top two blocks in the lower-right quadrant,

$$\begin{aligned} \left\|\alpha^2 W\right\|_{\max} &\lesssim \alpha^2 \sqrt{\frac{\log n}{n}} = o(1), \\ \left\|\beta^2 W Q \widehat{A}\right\|_{\max} &\leq \beta^2 \left\|WQ\widehat{A}\right\|_{\max} \lesssim \beta^2 \sqrt{\frac{\log n}{n}} = o(1). \end{aligned} \tag{7.17}$$

Finally,

$$\left\|\beta^2 W^\top P_k W Q^\top\right\|_{\max} \lesssim \beta^2 \sqrt{\frac{\log n}{n}} = o(1).$$

Note that all entries are smaller in magnitude than the pivot, which is 1. □

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

Proof. In order to express the matrix $B(A)^{(2n+k)}$, we decompose W and Q into additional blocks. For a fixed $0 \leq k < n$, consider the decomposition

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

where $W_{11} \in \mathbb{R}^{k\times k}$. Note that $W_{2:} = E_k^\top W$ and $Q_{2:} = E_k^\top Q$. Next, we split the first block row and column of $B(A)^{(2n)}$ according to this further partitioning:

$$B(A)^{(2n)} = -\left[\begin{array}{cc|c} \alpha^2 W_{11} & \alpha^2 W_{12} & \beta^2 W_{1:} Q\widehat{A} \\ \alpha^2 W_{21} & \alpha^2 W_{22} & \beta^2 W_{2:} Q\widehat{A} \\ \hline \beta^2 Q_{1:}^\top & \beta^2 Q_{2:}^\top & 0 \end{array}\right]. \tag{7.18}$$

Let

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

Applying the Schur complement formula directly to (7.18), we obtain

$$\begin{aligned} B(A)^{(2n+k)} &= \begin{bmatrix} -\alpha^2 W_{22} & -\beta^2 W_{2:} Q\widehat{A} \\ -\beta^2 Q_{2:}^\top & 0 \end{bmatrix} - \begin{bmatrix} -\alpha^2 W_{21} \\ -\beta^2 Q_{1:}^\top \end{bmatrix} \left(-\alpha^2 W_{11}\right)^{-1} \begin{bmatrix} -\alpha^2 W_{12} & -\beta^2 W_{1:} Q\widehat{A} \end{bmatrix} \\ &= \begin{bmatrix} -\alpha^2 S & -\beta^2 S Q_{2:}\widehat{A} \\ -\beta^2 (Q_{2:}^\top - Q_{1:}^\top W_{11}^{-1} W_{12}) & \gamma^2 Q_{1:}^\top W_{11}^{-1} W_{1:} Q\widehat{A} \end{bmatrix}. \end{aligned} \tag{7.19}$$

Next, we simplify the bottom blocks of (7.19). By the block inverse formula applied to W, and the fact that $W^{-1} = W^\top$, we have

$$S^{-1} = W_{22}^\top \qquad \text{and} \qquad W_{11}^{-1} W_{12} = -W_{21}^\top S. \tag{7.20}$$

Using (7.20),

$$\begin{aligned} Q_{2:}^\top - Q_{1:}^\top W_{11}^{-1} W_{12} &= Q_{2:}^\top + Q_{1:}^\top W_{21}^\top S \\ &= Q_{2:}^\top W_{22}^\top S + Q_{1:}^\top W_{21}^\top S \\ &= Q^\top W_{2:}^\top S. \end{aligned}$$

Let

$$M = I - W^\top E_k S E_k^\top, \tag{7.21}$$

or, in block coordinates,

$$M = \begin{bmatrix} I & -W_{21}^\top S \\ 0 & I - W_{22}^\top S \end{bmatrix} = \begin{bmatrix} I & W_{11}^{-1} W_{12} \\ 0 & 0 \end{bmatrix}.$$

Then $\operatorname{tr} M = k$ and $Q_{1:}^\top W_{11}^{-1} W_{1:} = Q^\top M$. We may rewrite (7.19) as

$$B(A)^{(2n+k)} = \begin{bmatrix} -\alpha^2 S & -\beta^2 S Q_{2:}\widehat{A} \\ -\beta^2 (W_{2:}Q)^\top S & \gamma^2 Q^\top M Q\widehat{A} \end{bmatrix}. \tag{7.22}$$

Note that this formula holds when $k = 0$, as, in that case, $S = W$ and $M = I - W^\top W = 0$.

Because $W \in \mathrm{CP}_n(\mathbb{R})$, $|S_{11}| = \|S\|_{\max}$. What remains is to show that $\alpha^2\|S\|_{\max}$ is the largest magnitude entry of (7.22). We begin with the top-right block, which we rewrite as

$$SQ_{2:}\widehat{A} = (E_k S^\top)^\top Q\widehat{A}.$$

Every column of $E_k S^\top$ is a row of S, and so $\left\|E_k S^\top\right\|_{1\to 2} \le \sqrt{n-k}\,\|S\|_{\max}$. In addition, $\left\|\widehat{A}\right\|_{1\to 2} \le \left\|\widehat{A}\right\| = 1$. Therefore, by (6.1),

$$\left\|SQ_{2:}\widehat{A}\right\|_{\max} = \left\|(E_k S^\top)^\top Q\widehat{A}\right\|_{\max} \lesssim \sqrt{n-k}\,\|S\|_{\max}\sqrt{\frac{\log n}{n}} = o\left(\frac{\alpha^2}{\beta^2}\right)\|S\|_{\max}.$$

Next, consider the lower-left block, which we rewrite as

$$(W_{2:}Q)^\top S = Q^\top W^\top E_k S.$$

Because $W^\top E_k$ has orthonormal columns,

$$\left\|W^\top E_k S\right\|_{1\to 2} = \max_{j\in[n-k]} \|S\boldsymbol{e}_j\| \le \sqrt{n-k}\,\|S\|_{\max},$$

and so, by (6.1),

$$\left\|(W_{2\cdot}Q)^\top S\right\|_{\max} = \left\|Q^\top W^\top E_k S\right\|_{\max} \lesssim \sqrt{\log n}\,\|S\|_{\max} = o\left(\frac{\alpha^2}{\beta^2}\right)\|S\|_{\max}.$$

Finally, we consider the lower-right block. By (7.21) and the orthogonality of W,

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

By (6.2) with $X = I$, $Y = \widehat{A}$, and M as itself,

$$\begin{aligned}\left\|Q^\top MQ\widehat{A} - \frac{k}{n}\widehat{A}\right\|_{\max} &\le C\left(\frac{(\sqrt{n} + (n-k)\,\|S\|_{\max})\sqrt{\log n}}{n} + \frac{(1 + (n-k)\,\|S\|_{\max})\log n}{n}\right)\\ &\lesssim \sqrt{\frac{\log n}{n}} + \frac{(n-k)\,\|S\|_{\max}\log n}{n}\\ &\lesssim \|S\|_{\max}\log n,\end{aligned}$$

as, by (7.5), $\|S\|_{\max} \ge (n-k)^{-1/2}$. Because $\|A\| \ge \frac{\alpha}{4}$ and, by (7.7), $\|S\|_{\max} \ge c\frac{k}{n^{5/4}}$, we have

$$\begin{aligned}\left\|Q^\top MQ\widehat{A}\right\|_{\max} &\le \left\|Q^\top MQ\widehat{A} - \frac{k}{n}\widehat{A}\right\|_{\max} + \left\|\frac{k}{n}\widehat{A}\right\|_{\max}\\ &\lesssim \|S\|_{\max}\log n + \frac{4k}{n\alpha}\\ &= o\left(\frac{\alpha^2}{\gamma^2}\right)\|S\|_{\max}.\end{aligned}$$

This concludes the analysis of $B(A)^{(2n+k)}$ for $k = 0, \dots, n-1$. Since $B(A)^{(3n)} = \gamma^2\widehat{A}$ and $\widehat{A} \in \mathrm{CP}_n(\mathbb{R})$, we have $B(A)^{(2n)} \in \mathrm{CP}_{2n}(\mathbb{R})$. □

The above three lemmas hold simultaneously with positive probability, yielding:

Proposition 7.11. *For sufficiently large n, if there exists $A \in \mathrm{CP}_n(\mathbb{R})$, $\|A\|_{\max} = 1$, with $\|A\| \ge \frac{\alpha}{4}$, then there is a realization of $Q \in \mathrm{O}(n)$ for which $B(A) \in \mathrm{CP}_{4n}(\mathbb{R})$, $\|B(A)\|_{\max} = 1$, $\alpha \le \|B(A)\| \le 3\alpha$, and*

$$\mathrm{growth}(B(A)) \ge \frac{\gamma^2}{\|A\|}\,\mathrm{growth}(A). \tag{7.23}$$

Theorem 7.12. *There is a universal constant C such that for each $n \in \mathbb{N}$, there is a completely pivoted matrix $A \in \mathrm{CP}_n(\mathbb{R})$ with $\mathrm{growth}(A) \ge n^{\frac{1}{16}\log_2 n - C\log\log n}$.*

Proof. It suffices to prove our desired result for geometrically spaced n, for then it would then hold for all n simply by padding the resulting matrices with the identity matrix, $I \oplus A$.

Let n_0 be sufficiently large so that Proposition 7.11 holds. Let

$$\alpha(n) = \frac{n^{1/4}}{\log n}, \qquad \gamma(n) = \frac{n^{1/4}}{\log^2 n}, \qquad s = \frac{\alpha(n_0) - 1}{n_0 - \alpha(n_0)}, \qquad \text{and} \qquad A_0 = \frac{I + s\mathbf{1}\mathbf{1}^\top}{1+s}.$$

We have $\|A_0\|_{\max} = 1$, $\|A_0\| = \alpha(n_0)$, and

$$A_0^{(j)} = \frac{1}{1+s}\left(I + \frac{s}{1+js}\mathbf{1}\mathbf{1}^\top\right) \qquad \text{for} \quad j = 0, \dots, n-1.$$

Therefore, $A_0 \in \mathrm{CP}_{n_0}(\mathbb{R})$ and $\mathrm{growth}(A_0) = 1$. Now, let $n_k = 4^k n_0$ and note that, if there exists an $A_k \in \mathrm{CP}_{n_k}(\mathbb{R})$, $\|A_k\|_{\max} = 1$, with $\frac{1}{4}\alpha(n_k) \le \|A_k\| \le 3\alpha(n_k)$, then we may choose a realization $A_{k+1} = B(A_k)$ provided by Proposition 7.11 such that $A_{k+1} \in \mathrm{CP}_{n_{k+1}}(\mathbb{R})$, $\|A_{k+1}\|_{\max} = 1$, and

$$\frac{1}{4}\alpha(n_{k+1}) \le \alpha(n_k) \le \|A_{k+1}\| \le 3\alpha(n_k) \le 3\alpha(n_{k+1}).$$

The growth factor of A_{k+1} is at least

$$\mathrm{growth}(A_{k+1}) \ge \frac{\gamma(n_k)^2}{3\alpha(n_k)}\,\mathrm{growth}(A_k).$$

Applying this repeatedly, and noting that $k = \frac{1}{2}(\log_2 n_k - \log_2 n_0)$, we obtain

$$\text{growth}(A_k) \geq \prod_{j=0}^{k-1} \frac{\gamma(n_j)^2}{3\alpha(n_j)} = \prod_{j=0}^{k-1} \frac{\left(4^j n_0\right)^{\frac{1}{4}}}{3\log^3(4^j n_0)} \geq \frac{4^{\frac{k(k-1)}{8}} n_0^{\frac{k}{4}}}{3^k \log^{3k}(n_k)} = \frac{n_k^{\frac{k}{4}}}{3^k 4^{\frac{k(k+1)}{8}} \log^{3k}(n_k)} = n_k^{\frac{1}{16}\log_2(n_k) - C\log\log n_k}$$

for some constant C and all k sufficiently large. □

Remark 7.13. One can construct A achieving the bound of Theorem 7.12 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 of conversation and collaboration on the growth factor problem in Gaussian elimination, as well as illuminating conversations with Alex Townsend, Nick Higham, and Nick Trefethen 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 academic 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 blends purely human-based mathematics with some mathematics that involved the use of ChatGPT Sol 5.6 and Claude Fable 5. This entire manuscript, including the results and proofs, was written by the authors. We aim to give a clear and detailed picture of AI involvement, section by section:

- Section 1 (human): No AI involvement.
- Section 2 (human/AI): Authors proved results regarding sparse matrices and partial pivoting in 2023. Interaction with AI improved the human-only results.
- Section 3 (human/AI): AI provided explanations of DPPs, examples of stochastic domination, and answered questions regarding the strong Rayleigh property. Authors had the idea of using a projection DPP, a Sylvester Hadamard matrix, and Hadamard matrices. Interaction with AI led to the proof that the choice of row permutation does not affect the Frobenius norms of L and U.
- Section 4 (human/AI): Authors had a proof of NP-hardness for a slightly different definition of growth factor in 2023. Interaction with AI led to a hardness proof for growth$(\cdot)$, which does not appear in the paper. Inspired by this proof, which shares many similarities with the original human-only proof, authors proved, without AI, the construction that appears in the manuscript.
- Section 5 (human/AI): AI was given the measure and prompted to compute expectations using determinant formulas. These were unsatisfactory. AI suggested a Jordan block after a prompt for a proof of boundedness. Authors performed simulations, which suggested that the expectation was tail dominated (AI confirmed this). Authors instead considered high probability bounds, still with a Jordan block. Authors mistakenly thought taking fixed z near 1 worked. Interaction with AI led to the suggestion of $z = 1 - 1/\sqrt{n}$. The above was for $p = 2$; authors generalized this without AI.
- Section 6 (human): No AI involvement. The earliest proof that the growth factor under rook pivoting is quasi-polynomial was completed in 2022.
- Section 7 (human/AI): Interaction with AI led to a quasi-polynomial lower bound construction. This interaction strongly followed the authors' research program/architecture for proving that complete pivoting has quasi-polynomial growth, which was a human-only endeavor from 2022 until a week prior to the first draft of this manuscript. Indeed, the resulting lower bound follows the human-generated road map that was provided. The resulting construction was quite complicated and hard to follow. Authors produced, without AI, the much simpler and intuitive construction that appears in the manuscript. This research program for the growth factor under complete pivoting could not have been completed 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 AI interactions, as no universal norms seem to currently exist yet.

References

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

[BBC20] Ayoub Belhadji, Rémi Bardenet, and Pierre Chainais. A determinantal point process for column subset selection. *Journal of Machine Learning Research*, 21(197):1–62, 2020.

[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.

[DM07] Tobin A Driscoll and Kara L Maki. Searching for rare growth factors using multicanonical Monte Carlo methods. *SIAM Review*, 49(4):673–692, 2007.

[DM21] Michał Dereziński and Michael W. Mahoney. Determinantal point processes in randomized numerical linear algebra. *Notices of the American Mathematical Society*, 68(1):34–45, 2021.

[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.

[DW18] Michał Dereziński and Manfred K. Warmuth. Reverse iterative volume sampling for linear regression. *J. Mach. Learn. Res.*, 19(1):853–891, January 2018.

[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.

[San04] Arvind Sankar. *Smoothed analysis of Gaussian elimination*. PhD thesis, Massachusetts Institute of Technology, 2004.

[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 $\Omega(2^n/n)$. *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.

[Tre12] Nick Trefethen. The smart money's on numerical analysts. *SIAM News*, 45(9), 2012.

[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.